Yaşar Tonta, Internet: an innovative environment for information dissemination, access and retrieval in distance education. *European Journal of Distance Education* (2): 6-32, 1995. *http://yunus.hun.edu.tr/~tonta/yayinlar/disteduc.htm*

INTERNET: AN INNOVATIVE ENVIRONMENT FOR INFORMATION DISSEMINATION, ACCESS, AND RETRIEVAL IN DISTANCE EDUCATION

by

Yaşar Tonta
Hacettepe University
Department of Library Science
06532 Beytepe, Ankara, Turkey
email: tonta@eti.cc.hun.edu.tr

Paper presented at the 1995 Annual Conference of the
Association of European Correspondence Schools
23-26 May 1995, Istanbul, Turkey

Introduction

Information sources available through the networks such as the Internet and Bitnet have proliferated and diversified in recent years. Nowadays, in addition to textual information, multimedia databases containing `documents' with graphics, images, sound and animation can often be found on the network. Government agencies, various institutions, commercial companies, and, more recently, individuals themselves, can disseminate information and open their own databases to the use of others who have access to the Internet services. End-users can easily get access to such information through the Internet, download it, and use it for their own purposes. Needless to say, availability of information through the Internet in various forms is having a profound impact upon how we work, educate and entertain.

This paper provides a brief overview of the Internet. We summarize the types of information that can be found on the network, describe basic Internet services (e.g., electronic mail, remote login, file transfer protocol), and introduce some of the information discovery and retrieval tools such as the Internet Gopher, WAIS, and the World-Wide Web. We then go on to discuss the impact of the Internet on resource sharing, document delivery, electronic publishing, telecommunications, and education, More specifically, we discuss the impact of

the Internet on distance education and investigate the ways by which distance education can benefit from this new information dissemination, access, and retrieval environment. We conclude by drawing attention to some of the issues with regards to using the Internet in distance education projects and make some recommendations.

What is the Internet?

Polly & Cisler (1994) suggest a number of metaphors to describe the Internet:

> There's the "Internet as highway" model: a blisteringly fast, multilane roadway where the vehicles are traveling in at least three dimensions at once, the directional signage changes all the time, and there are no rest stops.
>
> There's the "Internet as house" model suggested by Mimi King and developed by Peter Graham on the Coalition for Networked Information Big-ideas listserv discussion group:
>
> ...it's a household with shared keys but where the lights are off and we have to grope around. Some of us are looking for the fuse box, some don't care (and there are a few of us over in one of the darkrooms doing things others don't want to know about). There's a madwoman in the attic and some ranters on the stairs. Some of our publisher neighbors are worried about their property values...and others have asked the press to look into it. Some of us are having a good deal of success in getting the lights on and settling the zoning issues, which is moving others to light out for the territory. The guys building the extension out in back aren't talking to the committee in front, though they have agreed to use the same plumbing pipes.
>
> Our favorite analogy, though, is John Perry Barlow's. He is a lyricist for the Grateful Dead, cattle rancher, and cofounder of the Electronic Frontier Foundation. At NationalNet '93 he said, "it's a biological phenomenon. Internet is not a vertebrate. It acts a lot like slime mold, growing without anyone in charge. Every time I try to describe Internet to anyone, everyone assumes I'm having a hippie mystic vision! (Polly & Cisler, 1994: 38)

But, what **is** the Internet?

The Internet is known as the "network of networks." It comprises about 20.000 separate networks (federal, national, regional, campus) and interconnects almost five million computers all over the world. Every 30 seconds, another network of computers joins the network. Some 160 nations are interconnected to each other through the Internet. The number of people that can be reached via the network is said to be around 30 million. (But who is counting?)

The Internet was born about 25 years ago, as a U.S. Department of Defense network called the ARPAnet (Advanced Research Projects Association Network). "The ARPAnet was an experimental network designed to support military research--in particular, research about how to build networks that could withstand partial outages (like bomb attacks) and still function" Krol, 1993: 11). The Internet Protocol (IP), developed in 1973, makes sure that the message put in an envelope (called "packet") that is intended to be received by another

computer reaches its destination ("address") successfully. The Transmission Control Protocol (TCP), developed in 1978, on the other hand, deals with the data inside the message and breaks it into pieces (packets) and puts them into separate envelopes so that data can efficiently be sent over the network. On the receiving side, a TCP software package collects the envelopes, extracts the data, and puts it into proper order. "The communicating computers--not the network itself--were also given the responsibility to ensure that the communication was accomplished. The philosophy was that every computer on the network could talk, as a peer, with any other computer" (Krol, 1993: 11).

All computers connected to the Internet switched to the TCP/IP protocols in 1983. The ARPAnet was discontinued in 1990 after a new network (NSFNET) was set up in 1980s by the National Science Foundation, an agency of the U.S. Government. As Shaw (1994) said, it has grown into a "chaotic and amorphous network of networks."

Here are some recent statistics to help us understand how big the Internet is and how fast its use is growing in every field:

Growth in world usage of the Internet last year (1994) = 95%

Nations joining the network in 1994 = 22

Number of nations interconnected = 159

Number of computers connected to the Internet:

U.S. = 3.2 million; Britain = 241,191; Germany = 207,717; Canada = 186,722; Australia = 161,166; Japan = 96,632; France = 93,041

Number of organizations worldwide interconnected = 56,000;

businesses = 32,000; business computers online: 1.3 million; school and university computers = 1.1 million; government computers = 209,345;

These are approximate figures. As Krol (1993) indicates, it is difficult to know exactly what comprises the Internet since the numbers would include various federal networks, a set of regional networks, campus networks, and some foreign networks. More recently, having seen that the Internet was good, some non-IP based (Bitnet, DECnets, etc.) networks joined the network, too.

As can be seen from the above figures, the Internet is no longer a network to support military research only. Today the Internet is an international network used for research, commerce, education, entertainment, and so on. The U.S. Vice President Al Gore envisioned the Internet as an "information superhighway" and saw the network as an environment to get access to rich information sources (archives and library catalogs, medical records, police records, etc.). In fact, he thought of a poor little girl in grade school in Tennessee who comes home from school and needs information to complete her homework. Through the network Gore envisioned to open all the resources of the Library of Congress to that little girl so that she could complete her homework and get a first-class education. He saw the endless opportunities the network may offer long before anyone else and continues to promote a stronger global information superhighway. Most recently, in February 1995, he was in Brussels for the meeting of the Group of Seven (G-7), the nations of Germany,

Japan, Britain, France, Italy, Canada, and the United States where they discussed, among others, how they could cooperate to design a global information superhighway.

Who Governs the Internet?

This brings the question of who governs the Internet. There is no single authority, entity or organization for the Internet as a whole. In many ways Krol (1993) likens the Internet to a church with its council of elders.

> Every member has an opinion about how things should work and you can either take part or not. The Internet Society (ISOC), a voluntary member organization, has the ultimate authority for where the Internet is going. Its purpose is to promote global information exchange through Internet technology. It appoints a council of elders, which has responsibility for the technical management and direction of the Internet.
>
> The council of elders is a group of invited volunteers called the *Internet Architecture Board*, or the IAB. The IAB meets regularly to "bless" standards and allocate resources, like addresses. The Internet works because there are standard ways for computers and software applications to talk to each other. This allows computers from different vendors to communicate without problems.
>
> As in a church, everyone has an opinion how things ought to run. Internet users express their opinions through meetings of the Internet Engineering Task Force (IETF). The IETF is another volunteer organization; it meets regularly to discuss operational and the near-term technical problems of the Internet. When it considers a important enough to merit concern, the IETF sets up a "working group" for further investigation (Krol, 1993: 13-14)

Who Pays for the Internet?

> • • The old rule for when things are confusing us "follow the money." Well, this won't help you to understand the Internet. No one pays for "it"; there is no Internet, Inc. that collects fees from all Internet networks or users. Instead, everyone pays for their part. NSF pays for NSFNET. NASA pays for the NASA Science Internet. Networks get together and decide how to connect themselves together and fund these interconnections. A college or corporation pays for their connection to some regional network, which in turn pays a national provider for its access (Krol, 1993: 15).

What Does This Mean for Me?

> • • The concept that the Internet is not a network, but a collection of networks, means little to the end-user. You want to do something useful: run a program, or access some unique data. You shouldn't have to worry about how it's all stuck together. Consider the telephone system--it's an internet, too. Pacific Bell, AT&T, MCI, British Telecom, Telefonos de Mexico, and so on, are all separate corporations that run pieces of the telephone system. They worry about how to make it all work together; all you have to do is dial. If you ignore cost and commercials, you shouldn't care if you are

> dealing with MCI, AT&T, or Sprint. Dial the number and it works...You only care who carries your calls when a problem occurs (Krol, 1993: 15)

Information Sources Available on the Internet

The amount of information on the Internet currently occupies about a few terabytes (10^{12} bytes) worth of space. Yet the amount is increasing tremendously as more multimedia information sources are added to the network. As there are many different types of information resources available through the Internet, it is difficult to list them all here. Here are some of the basic categories of information sources:

Library catalogs and bibliographic databases

Catalogs of large libraries and information centers are open to public via the Internet. Bibliographic databases held in government and research institutions are also accessible through the network. They may also include indexes and abstracts of published articles in various subjects.

Archives of electronic texts, electronic journals and newsletters

The electronic copies of many printed books, dictionaries, encyclopedias and other such reference materials are also available through the Internet. The Project Gutenberg aims to convert major non-copyrighted works into electronic form so that people can get access to them free of charge. Parallel publication of both the electronic and printed versions of the same journals has started in early 1980s and a number of publishers such as the American Chemical Society and the American Mathematical Society offer the full-text of their journals through the Internet. More recently, the Internet is used as a means of publishing scholarly articles in electronic form. The use of networks as a medium of publication has proliferated especially after the "cold fusion" controversy of 1989. Hundreds of electronic journals are currently being "published" through, and archived on, the network. The fifth edition of the *Directory of Electronic Journals, Newsletters and Academic Discussion Lists* contain entries for "nearly 2500 scholarly lists, 675 electronic journals, newsletters, and related titles such as newsletter-digests--an increase in size of over 40% since the 4th edition of April 1994 and 4.5 times since the first edition of July 1991" (Okerson, 1995). The full-text of articles published in electronic journals can be downloaded from the archival sites.

Logs of discussion lists and newsgroups

Many researchers and scholars belong to specialized electronic mailing lists where tools of the trade, philosophical questions, and research queries are discussed. Similarly, less formal newsgroups are set up to discuss more mundane issues such as social and cultural concerns. The logs of these discussion lists and newsgroups can be searched retrospectively to find the relevant postings and they can be obtained via the network.

Software archives

The copies of many useful public domain computer programs are held in some sites and made available to other network users. Such software programs are usually free of charge. Some shareware software programs are also offered through the network.

Information about the users

Electronic phone books are put on the network so that users can obtain the (electronic) addresses of other network users by consulting such phone books. Assorted campus information can also be found in such databases.

Images and Sounds

As the Internet develops and gets more sophisticated, new electronic information sources such as documents with images or compound multimedia objects, weather maps, satellite pictures, electronic sensor feeds are also disseminated through the network. Television and film archives are likely to be available on the network soon once the economics of providing such services are determined. The Internet is also used to "broadcast" the radio speeches of important personalities (such as the interview with the "geek of the week").

Major Internet Services and Resource Discovery and Retrieval Tools

As described above, the Internet offers a wide range of information resources. Yet it's "messy and poorly coordinated. There are incredible resources, but there is no central coordination to help you find what you want" (Krol, 1993). "Resources can appear and disappear without notice, so considerable investment must be made in maintaining knowledge of what's there and what's where." (Shaw, 1994: 208).

Traditionally, *electronic mail*, remote login (*telnet*) and file transfer protocol (*ftp*) have been the first services offered by the Internet.

Electronic mail (e-mail) lets you send and receive messages over the network. E-mail comes directly into people's mailboxes. It doesn't require that people take the time to seek out information actively or use any special software other than their familiar electronic mail programs.

Telnet is used for (remote) logging into other computers on the Internet. It's used to access lots of public services, including library card catalogs and other kinds of databases.

File Transfer Protocol (ftp) moves files back and forth. It's most useful for retrieving files from public archives that are scattered around the Internet. This is called "anonymous FTP," because you don't need an account on the computer you are accessing.

Information Dissemination, Discovery and Retrieval Tools

We have pointed out earlier that the Internet can also be used to disseminate information. People often wish to share information with groups of colleagues, sometimes within their own work groups or departments, and sometimes within a larger sphere. Novogrodsky, et al. (1993) describes four information dissemination tools and compares them with one another: mailing lists, list server mailing lists, Usenet newsgroups, and the Internet Gopher. The following summary is based on their article. The Wide Area Information Server (WAIS)

and the World-Wide Web (WWW), relatively newer Internet services, are also introduced in this section (Krol, 1993; Roth, 1994). Software packages that allow users to search all the Gopher servers (menu items and titles) on the Internet using keywords (*veronica*, *jughead*) or searching for a given file name in public domain ftp archives (*archie*) are not discussed in this paper.

Mailing lists

Mailing lists allow you to `broadcast' an electronic mail message to a group of people easily. By sending your message to a single address--the address of the mailing list, you can have your message automatically sent to all members of the list. They work well for announcing timely material such as upcoming meetings, and for group discussions among the list.

Since more people are likely to have access to and know how to use electronic mail software, mailing lists are probably the best method to use to reach a group of people who use different types of computers and software.

List Server Mailing Lists

List servers, such as LISTSERV, listproc, and tulp, are programs that manage mailing lists. They provide all the capabilities of the simpler type of mailing lists described above, and also provide a number of additional features including automated list management and document distribution.

Usenet Newsgroups

A Usenet "newsgroup" is an electronic bulletin board on a specific topic. By using "news reader" programs, people can read messages ("articles") that have been posted by other users, and can participate in discussions by posting their own articles. This may sound obscure, but it's really what everyone else calls "bulletin boards" or discussion groups. Usenet is the world's largest bulletin board service.

Newsgroups are more efficient than mailing lists for large group discussions, since messages are stored on a central server. Since people can choose when to read articles on newsgroups, they will not become involuntarily inundated with large volumes of electronic mail.

Some newsreaders allow one to search for article by their subject lines or even within the full text of the articles. This makes it possible to locate specific articles in your area of interest easily.

The Internet Gopher

The Internet Gopher is a distributed document search and retrieval system. It provides a whole new way for beginners to get to know the Internet. Because Gopher servers are menu-based, there is no memorization of cryptic commands. When you find something you like, you can read or access it through the Gopher without having to worry about domain names, IP addresses, changing programs, etc. Because Gopher links you transparently to information services everywhere on the Internet, you can get information effortlessly from

all over the world regardless of where. If you can navigate through a simple menu, you hold the resources of the Internet in your hand.

Gopher information servers are most appropriate as places to "publish" information that does not change frequently. The big advantage of gopher isn't so much that you don't have to look up the address or name of the resources, or that you don't have to use several commands to get what you want. The real cleverness is that it lets you browse through the Internet's resources, regardless of their type, like you might browse through your local library with books, filmstrips, and phonograph records on the same subject grouped together. Gopher knows which application (**telnet**, **ftp**, etc.) to use to get a particular thing you are interested in and does it for you (Krol, 1993).

Sometimes a combination of these tools can be used to best reach your intended audience. For instance, you may choose to make documents available on a Gopher server and then send an announcement to a discussion list or a Usenet newsgroup.

Wide Area Information Server (WAIS)

WAIS, one of the newer Internet services, is a distributed text searching system. It is based on a standard (named Z39.50) that describes a way for one computer to ask another to do searches for it. WAIS is really a tool for working with collections of data, or databases. It's great for searching through indexed material and finding articles based on what they contain. That is: WAIS lets you search through Internet archives looking for articles containing groups of words.

Like Gopher, WAIS allows you to find and access resources on the network without regard for where they really reside. In Gopher, you find resources by looking through a sequence of menus until you find something appropriate. WAIS does the same thing, but it does the searching for you. You tell it what you want; it tries to find the material you need.

World-Wide Web (WWW)

Imagine sitting in front of your computer screen and browsing through a Smithsonian Institute exhibit, downloading up-to-the minute radar and weather maps, reading an online magazine about Internet news, watching the Olympic skaters and their scores, or becoming a virtual tourist.

In 1993, hundreds of universities and other organizations around the world have formed a worldwide information system, called the World-Wide Web (WWW) based on specifications originally developed at CERN, the European Laboratory for Particle Physics located at Geneva, Switzerland. This system allows computers on the Internet to access all this shared information. A computer with information to share runs the Web server software, while a machine accessing this information uses the Web client software (typically NCSA Mosaic). Both use a common language to communicate with the other called HTTP (hypertext transfer protocol).

The Web is a distributed information access service based on hypertext model of information representation. What you see on the computer screen are documents and links, which themselves contain other documents and links. Hypertext documents can also contain images, sounds, and animation. The hypertext structure allows you to navigate through webs of information. When you find a document, you can display the text contained within,

or you can print it or save a copy of it on your computer. Some Web servers allow you to perform a full-text search on a group of documents, displaying those documents that contain the matched word.

Web is an attempt to organize all the information on the Internet, plus whatever local information you want, as a set of hypertext documents. You traverse the network by moving from one document to another via "links." Many of the resources available through the Web are WAIS resources.

You really have to see it for yourself to appreciate what the Web has to offer.

The Impact of the Internet

The Internet is having a great impact upon almost every conceivable field one can think of. Exploring the detailed impact of the Internet on all aspects of life is beyond the reach of the present paper. Suffice to say the Internet has changed the way we communicate, share resources, deliver and access information, publish, do business and commerce, and educate. For instance, developed nations are trying to improve their telecommunications infrastructures so that they could move billions of bytes worth of information back and forth. For they see "information" as important and valuable a resource as energy for the development of the Gross National Product (GNP). They are aware of the relationship between economic development and the availability of information resources. They take the information dissemination, access and retrieval issues as a whole and plan such services accordingly.

The National Information Infrastructure (NII) of the U.S. is the result of such an initiative and it paves the way for sharing information among various components of the government and private sector (health, police, legal system, banking, insurance, education, media, etc.) Such sharing is not confined to printed information only. It can be in any form: X-ray films taken in hospitals, "digital cash," film and television archives, realtor information, architectural details of buildings, and so on.

The National Research and Education Network (1991) Act of the U.S. authorizes to set up and develop a network for the purposes of research and education. Students and faculty alike in all levels are encouraged to acquaint themselves with the Internet services as early as possible (e.g., in grade schools). The use of the Internet by all segments of the society as much as possible is the goal to be strived for. A similar approach can also be seen in other developed countries (e.g., United Kingdom and France).

The Internet and the Distance Education

Just as the remote diagnosis of diseases is becoming possible thanks to sharing health data between medical centers by means of the Internet, it should be possible to better educate and train people using the Internet facilities. In fact, this is being done already. A number of courses are offered through the Internet and some schools (online universities) offer degrees to those who complete the degree requirements using the Internet facilities. Residency is not part of the requirements! Anyone who has access to the Internet facilities, regardless of

where he or she leaves, can enroll in such degrees. All course material is offered through the network. Students complete and turn in their assignments in the same way. Professors could grade their work interactively when they have access to the same databases and information sources.

Over the years, the field of distance education seems to have accumulated a wealth of knowledge about the use of technology for educational purposes. Distance education and training literature is filled with articles on computer-mediated communication (CMC) and its use in distance education projects. Therefore I think that you are in the best position to appreciate what the Internet can offer to the distance education. I am not here to tell you what you should do to make use of the Internet facilities to improve the distance education curricula. I firmly believe that the quality of the correspondence school programs would increase and a new courses based on what the Internet has to offer could be designed in the near future. For instance, a European Electronic University that Professor Bates envisioned a few years ago could be supported best with the Internet's information access and retrieval tools. The central databases of course materials can be accessed by all the students via the Internet. Such repositories would include multimedia information as well as the electronic copies of course texts.

In the context of distance education, the Internet can be used in three main application areas: academic, management and research. Zorkoczy's framework for the use of CMC in distance education also applies to the Internet. Zorkoczy identified three factors that play important roles in the success of CMC in distance education:

-the features and characteristics of the technology itself

-the applications

-the characteristics of the users and their environment

We think his framework can also be used to characterize the Internet and its use in distance education. We have to deal with these three aspects when considering the use of the Internet facilities.

The Internet is still an evolving technology. Telecommunication networks of today are simply inadequate to transmit a large amount of non-textual information. For instance, it may take several minutes to transmit a single color picture when the network is busy. Yet multimedia information in the form of graphics, pictures, sound and video can be extremely important for research and teaching in some fields (e.g., scientific visualization, teaching foreign languages, cinema and television). It is highly unlikely that research and teaching without such visual information will be useful. Moreover, we should remember what Bates pointed out: "technological applications, no matter how powerful or sophisticated, must serve rather than determine the educational goals." Allow me to summarize his other thoughtful comments:

-`black box' approaches, based on the notion that all learning can be conveyed through a single piece of equipment, are naive and dangerous. A multi-media approach, including the direct human interaction is essential for effective distance education.

-the older, more familiar technologies such as print and audio still have much to offer in terms of learning and cost-effectiveness; newer technologies are not automatically superior, will still need to prove themselves.

-technology alone is not enough; there still needs to be a system of teachers, management and administration to ensure that courses are properly designed, delivered and supported in this field. Thus the technology needs a sophisticated educational sub-structure to maintain it (Bates).

As for the applications given in Zorkoczy's framework, I think we will see more and more hypertext and hypermedia applications designed for distance education students on the Internet. Yet it is not clear if the corresponding schools can easily incorporate such applications in their curricula. The economic and operational issues may hinder such efforts.

We are especially thinking of the characteristics of the users and their environment. Providing distance education students with needed skills and technology to utilize the Internet resources to complete their education is easier said than done. The computer (and information) illiteracy is a common problem everywhere. This is perhaps more so for distance education students as they may not have enough time or financial resources to polish their computer skills. Furthermore, from the management point of view, you may not just assume that everyone knows how to use the Internet services efficiently. Educating them until they get conversant with the Internet services requires additional resources (teachers, equipment, materials, etc.). As we pointed out earlier, the Internet is a messy environment and it is not always easy to find what you are looking for.

Not all courses offered by the distance education programs will be suited for the use of the Internet. Yet it is no exaggeration to suggest that distance education will become more and more Internet-assisted. As the computer literacy levels rise in our society, it will get easier to design distance education courses incorporating the Internet technology. Meanwhile, planners of contemporary distance education programs should strive for the maximum utilization of the Internet as an environment for information dissemination, access and retrieval.

Bibliography

Bates, A.W. "Towards a European Electronic University: Technology and Course Design for European-Wide Distance Education Courses,"data file available at the following address: telnet://sun.nsf.ac.uk (login: janet, password: icdl; hostname: uk.ac.open.acs.vax, username: icdl)

Krol, Ed. (1993). *The Whole Internet: User's Guide & Catalog*. Sebastopol, CA: O'Reilly.

Miller, Wayne. (1993 October). "The Academic Gopher: How to Tame It, Train It, and Do Tricks With It," *Berkeley Computing and Communications* 3(7): 14-16.

Novogrodsky, Seth, Aron Roberts, & Shuli Roth. (1993 October). "A Comparison of Four Information Dissemination Tools," *Berkeley Computing and Communications* 3(7): 10-13.

Okerson, Ann. (1995 May 18). "New Edition of Internet Journal Directory Available," e-mail message posted to the Public-Access Computer Systems Forum, PACS-L@UHUPVM1.UH.EDU.

Polly, Jean Armour & Steve Cisler. (1994 January). "Connecting to the Global Internet," *Library Journal* 119(1): 38-39.

Roth, Shuli. (1994 April). "Expanding Your World of Information with the World-Wide Web," *Berkeley Computing and Communications* 4(3): 11-12.

Shaw, Debora. (1994). "Libraries of the Future: Glimpses of a Networked, Distributed, Collaborative, Hyper, Virtual World," *Libri* 44(3): 206-223.

Zorkoczy, Peter. "CMC in Distance Education and Training: The Broader Context," data file available at the following address: telnet://sun.nsf.ac.uk (login: janet, password: icdl; hostname: uk.ac.open.acs.vax, username: icdl)