A Resource for Teaching Internet Access

Natarajan Narasimhamurthi

*Department of Electrical and Computer Engineering,
University of Michigan-Dearborn, USA
e-mail: nnarasim@umich.edu*

S Amba

*Information Area, Central Leather Research Institute
Madras 600 020
e-mail: samba@niclai.ernet.in*

Abstract

This article presents the basics of Internet and its benefits to the users who want to avail the opportunities offered by the net in simplified questionnaire form. The text lists out various login sites of interest to information professionals, librarians and documentalists.

1. INTRODUCTION

However hackneyed it might sound, we are in the midst of the information revolution and the Internet is the force that is driving this revolution. The Internet is a network of literally millions of computers in over 75 nations and accessed by millions of users. It is estimated that every month two million new users try the Internet [1]. While a majority of the machines use the Internet as a communication medium, a good percentage of them are information service providers.

At present, information on the Internet is stored in a haphazard, relatively unstructured manner thus making the access to information tedious, and a hit or miss proposition. Jacobson [2] compares the Internet to a large rapidly growing library where the books have been placed randomly on the shelves and have no call numbers or other classification schemes and with new collections being added every day. Hence, it is important that librarians understand and exploit this information resource which is literally swamping the world with electronic data and information. Of particular interest to librarians is the problem of citing electronic documents. Standards are being developed for this. Another area of interest is of cataloguing Internet resources. The OCLC has brought out a manual for cataloguing Internet resources. This manual was developed to aid those participating in

the OCLC/US Department of Education funded project 'Building a Catalogue of Internet Resources'. A searchable catalogue of Internet resources called the InterCat will be available at URL http://www.oclc.org:6990 till March 1996. It is essential, therefore, for library and information workers to actively participate in this revolution and not just be carried along with the tide. For this, it is necessary to know about the Internet and how to effectively utilise this vast 'comparatively unorganised' library.

In this paper, we provide information that will be useful to the instructors. We have selected only what we consider as the core information that every user of Internet should know. At a minimum, the students should be able to use the Internet to send and receive electronic mail, *telnet* to remote sites, search for files using a search program such as *Archie*, and retrieve files using *ftp*. Although the World Wide Web with its hypertext technology will make searching and retrieval easier, not all documents are available on the Web and the primitive methods are still necessary. However, we feel that the students should be introduced to the Web and strongly encouraged to browse through the Web. We have indicated some interesting sites to visit on the network. The instructor may add additional sites and also make it a requirement for the students to atleast visit a few and retrieve information from them.

One final note to the instructors: 'Teaching' Internet is more like teaching swimming or bicycling. There is not much of theory to teach and the students learn only by doing. It is therefore important to ensure that the students are given adequate computer time and the course be taught in a laboratory setting (with an Internet connected computer).

2. INSTRUCTIONAL MODULES

2.1 Preliminaries and Pre-requisites

2.1.1 What is FAQ?

FAQ stands for Frequently Asked Questions. Most tutorial and instructional materials on the Internet are often in the form of questions and answers and the documents are often called FAQs. The name of the documents often have the extension *.faq* or *.FAQ*.

2.1.2 What is Internet?

Internet is a backbone connecting networks of computers. Typically, one machine in each network, usually called the gateway machine, is connected to the Internet.

2.1.3 Who regulates the Internet?

Nobody! Internet is an *unregulated* network. Any machine that follows the Internet Protocol (IP) can be on the network. Each machine on the network is given a *unique* Internet address. Most users access the Internet through an Internet Service Provider (ISP) who has a machine on the Internet. Currently, in India, the three providers are the Educational and Research Network of the Department of Electronics (ERNET), Vidhesh Sanchar Nigam Limited (VSNL) and the Software Technology Parks of India (STPI). Of the three, only VSNL is available to the general public.

2.1.4 What is an Internet Service Provider?

Internet Service Provider (ISP) gives you an account on a machine that is connected to the Internet. You log on to the machine and access the Internet from there. Most of the users will connect to the machine using modems and telephone lines. Dedicated users may have round the clock connections using leased telephone line

connections, or radio links or even satellite links to the provider.

2.1.5 I have a PC and a Modem. Can I place my machine on the Internet?

Generally, you use your PC as a terminal to the machine provided by the ISP. However, if your ISP provides the right access to you, you connect to the ISP using either PPP or SLIP connections (you use specialised software to do this). Once you are connected and the connection is authenticated, the ISP will provide your machine with a temporary Internet address. At this stage, your computer is physically a part of the Internet. Your machine will be on the network till you disconnect, at which time the temporary address is returned to the pool that is maintained by the ISP. This address will be assigned by the ISP to the next machine that connects using the SLIP/PPP.

Not all ISPs provide SLIP/PPP connections and those that do may not extend this privilege to all their clients. Service providers call the clients who are permitted SLIP/PPP connection as TCP/IP accounts.

2.1.6 What is TCP/IP?

TCP (Transmission Control Protocol) refers to the manner in which the control is transferred from one machine to the other (in this case the service provider and your machine). IP stands for Internet Protocol, the protocol that should be followed by any machine that wants to communicate on the Internet. Programs running on personal computers that allow the PC to use TCP/IP are often referred to as TCP/IP stacks.

2.1.7 What is the history of Internet?

The origin of the Internet can be traced to the ARPANET. ARPANET was a network created by the Advanced Research Projects Agency, the US Department of Defence Organisation. This network linked researchers in various organisations and enabled them to share resources and information. This network split into two — one for military use and the other for civilian purposes. The interconnection between them came to be known as the Internet. Other networks such as the BITNET, CSNET, NSFNET, JANET to name a few were developed and connected with the Internet, thus making it a huge network of computers.

2.1.8 What tools and equipment do I need to access Internet?

A PC, modem, communication software and a computer account on a machine connected to the Internet. If the access to the Internet is text based, any simple PC with modem speed of 2400 baud rate is adequate. However, if you are interested in graphics and sound, you need a PC with processor comparable to or better than Intel 386 chip and modems with speed of 9600 baud. If you are interested in a SLIP/PPP connection, you would need the appropriate software. Intensive users of the Internet should have a (Unix) workstation directly connected to the network on high speed link.

2.1.9 What pre-requisite (background) knowledge do I need to use Internet?

You need to know: How to use a PC and the communication programme and how to log on to the machine on which you have the account. If you plan to transfer files between your PC and the machine on the Internet you should become familiar with the file transfer programmes that comes with the communication software. Kermit, Xmodem, Ymodem and Zmodem are some of the widely used file transfer programmes.

2.1.10 What is available on the Internet?

1. Free software developed at various universities and research organisations.

2. Free software developed by individuals and placed on the Internet for general use.

3. Shareware/crippleware programmes.

4. Research articles and reports.

5. Electronic journals and books.

6. Bulletin boards.

7. Discussion groups.

8. Pictures.

9. Library catalogues.

10. Product information and specifications.

11. Upgrades and bug fixes for commercial software.

12. Newspapers.

13. Electronic shopping malls.

2.1.11 What is Shareware/Crippleware?

Shareware is a form of distributing commercial software. Typically you are given 30 to 40 days to try it out and at the end of the period you are expected to pay for the software if interested or delete it from your machine. Crippleware is a form of commercial software where certain key features are inactivated. This allows you to try the software before buying it.

2.1.12 What facilities are available for searching the Internet?

If you know the name of the file, you can use Archie to locate the file. In most cases it is enough if you can specify the name of the file using wild cards. Gopher is a menu driven search tool that also locates files. Even if the computer you use does not provide these services, you can telnet to the sites that do provide them (telnet is described in a later module). In addition to Archie and gopher, there are several other search tools such as veronica, wais, etc.

2.2 E-Mail, Electronic Journals, Bulletin Boards and Discussion Lists

2.2.1 What is e-mail?

Electronic mail or e-mail as it is popularly known is the electronic equivalent of the normal postal system. E-mail permits the transmission of text in machine-readable form from one computer system to another where it is stored in the recipient's mailbox. Every e-mail consists of atleast four elements :

a) the address of the recipient,

b) the address of the sender,

c) a subject line and

d) the letter/message.

2.2.2 What is in an e-mail address?

The e-mail address consists of your *id* and the name of the machine on which you have your account. For example, *nvijay@giasmd01.vsnl.net.in* would be the e-mail address of the user *nvijay* whose account is on the machine called *giasmd01.vsnl.net.in*. The format is *userid@host.subdomain.firstleveldomain*. The address is in several parts, each of which is separated by a period. The first part is the name of the machine, the subdomain indicates the location of the machine and the network and the first level domain indicates the country, and if in the US, the type of organisation. The system used by the Internet to assign addresses to computers on the network is known as the *Domain Name System* (DNS).

2.2.3 How are machines identified on the Internet?

Computers on the Internet can be identified either by names as described above or by numbers. The number consists

of four parts separated by decimal points. The number is the official address of the machine. On the Internet there are machines called domain name servers which have a database which map the names to the numbers.

2.2.4 How do I send e-mail?

You need a mail server to send e-mail. Unix machines usually have two programmes called *mail* and *Mail* that use SMTP (Simple Mail Transfer Protocol). Of the two the latter is more user friendly. To send mail using the traditional mail programme, type *Mail* followed by receiver's address. Type the message and use the commands specific to the system for sending it to receiver. Menu driven mail servers such as Pine, Elm, Pegasus which support MIME are available on most systems. These programs are generally easier to use.

2.2.5 What is MIME?

MIME stands for Multipurpose Internet Mail Extensions. The MIME standard permits you to incorporate different file formats such as text, graphics, sound and video in your e-mail.

2.2.6 What are e-Journals, bulletin boards, discussion lists?

Electronic Bulletin Boards also known as electronic conferences or lists, are just an extension of the electronic mail service. Some are restricted but generally these bulletin boards are open for all. Some bulletin boards are moderated whereas others are free-for-all discussions groups. Subscribers who participate in these electronic conferences come to know about others who face similar problems and the various solutions arrived at. Bulletin boards or computer conferences span a wide variety of subjects. For the librarian there are a number of 'lists' in which librarians discuss different aspects of library

work. The **lis-forum** managed by National Centre for Science Information at the Indian Institute of Science, Bangalore serves as a forum for wide ranging discussions on topics of interest to librarians. The **Pacs-L** discusses applications of computers in libraries. Closely related to computer conferences are electronic journals and newsletters. Some of these journals are subjected to peer review just like the conventional scholarly journal. Electronic journals can have print counterparts or may be in electronic format only. The Office of Scientific and Academic Publishing of the Association of Research Libraries has brought out a *'Directory of Electronic Journals, Newsletters and Academic Discussion Lists,* (Gopher Edition) in May 1995. Another electronic source for information about electronic journals is the *NewJour* which is an electronic announcement list which updates the ARL Directory between its formal printed and networked editions. There are electronic journals for the librarian such as the *Olive Garden* and the *TER (Telecommunications Electronic Reviews).*

2.2.7 How do I subscribe to them?

Bulletin Boards are managed by an account known as Listserv. The Listserv also deals with subscription requests. For example, to subscribe to Pacs-L you do the following

```
mail listserv@uhupvm1.uh.edu
subscribe pacsl [followed by] your full name.
```

2.3 Telnet and Archie

2.3.1 What is Telnet?

Telnet is an application that allows you to login to any machine on the Internet once you are on the net. Thus, once you have logged on to a machine on the Internet, it is possible to log on to any machine in the world.

2.3.2 Why would I want to use Telnet?

Often, you may have account on more than one machine which may not be geographically close. Telnet permits you to work on more than one machine. This is especially useful for collaborative work. Also, certain facilities may not be available in your local machine but only on some remote machine. As long as you have login privileges on the remote machine, you can make use of these facilities. Some machines might only provide restricted access to them for anyone. For example, you will telnet to a site that has an Archie server and login with the user id *Archie* to make use of the Archie service.

2.3.3 What is Archie?

Archie started out as an archive of files available to general public. Over the years, Archie has become a powerful search tool to locate publicly available files. You use Archie by connecting to an Archie site and running the Archie program.

2.3.4 How do I connect to an Archie site?

You telnet to an Archie site and login as *Archie*. Once you are logged in, the computer will automatically start the Archie program. Note that most Archie sites will not allow you to do anything beyond running the program.

2.3.5 Can you name some Archie sites?

Site	Location
archie.sura.net	Maryland, USA
archie.unl.edu	Nebraska, USA
archie.mcgill.ca	Canada
archie.funet.fi	Finland
archie.doc.ic.ac.uk	United Kingdom
archie.au	Australia
archle.ncu.edu.tw	Taiwan

2.3.6 How do I look up the location of a File using Archie?

Once you have Archie running, use the command *find* followed by the file name. You can use wild cards in the search. You can also use what is called a sub-string search in which the pattern you use for the file name will be matched with any file that has the pattern as a sub-string. Generally, only the first 99 files that match the pattern would be shown, but this number can be changed. Also, if you want a more permanent record of the search results, you can issue the command *mail to* followed by your e-mail address. In this case, the results of the search will also be mailed to you by e-mail.

2.3.7 Now that I know Where the File is, and How do I get it?

To get the file you need to use the program called *ftp*.

2.4 *ftp* and File Transfer

2.4.1 What is ftp?

File Transfer Protocol (*ftp*) is an application that permits the transfer of files between two computers. Typically, you run the program on one computer and open a connection to the other computer. When establishing the connection, you need to login to the other computer using a valid userid and the corresponding password. Once, the connection is established, you can transfer files from your local machine using the *put* command and retrieve files from the remote machine using the *get* command. You retrieve most of the public documents that are available on the Internet using 'anonymous' *ftp*.

2.4.2 How do I connect to an anonymous ftp site?

Anonymous ftp sites are computer sites which have public files. These files can be accessed by anyone. You do not have to be

a registered user to access these files. When you use anonymous ftp, you login as *anonymous* once you are connected to the ftp site. The password generally used is one's own e-mail address.

2.4.3 How do I put files for the use of others?

Some anonymous ftp sites will permit you to put files on their system. You can select one such site and place your documents there. Your file can now be used by others. It is also important to note that most anonymous ftp sites mirror other sites also; i.e., they will keep copies of files found in the site being mirrored. Thus, once you place the file on a site, it may be duplicated at several sites.

2.5 World Wide Web

2.5.1 What is WWW?

The World Wide Web, more commonly known as the Web or WWW, is a network of WWW servers that use the Internet as the backbone. The Web started as a small, hypertext-based network of documents. The documents themselves were distributed over several machines on the Internet, connected to each other through hypertext links. Although it started as a tool for researchers to exchange information, many organisations joined by adding documents to this network and providing links to these documents. At present, the Web is truly world wide with Web documents literally spanning the globe.

2.5.2 What are Hypertext and Hypermedia?

Hypertext is a document which permit 'non-linear' access to information. The documents contain programmed 'links' to other documents and selecting the link results in the display of the linked document. The 'path' that you follow in a hypertext document is determined entirely

by you. This is in contrast to the conventional documents which follow a 'linear' structure. Hypermedia refers to hypertext documents which in addition to text contain graphics, video and sound. The pages on the Web are hypertext documents as they contain links to other documents. What makes the Web truly world wide is that the documents could be anywhere in the world and are known to each other through their URL.

2.5.3 What is an URL?

Every document on the Internet has an unique identifier called the URL (Uniform Resource Locator). URLs contain three parts—the type of protocol that is used to fetch the document, followed by the computer site where the document is stored, followed by the directory path on the computer to the actual document. The last piece of the URL is likely to be the file name. For example, the URL *http://www.3w.com/3w/index.html* locates file called *index.html* in the directory *3w* on the machine *www.3w.com*. The initial section *http:* indicates that this document should be fetched using hypertext transfer protocol. Also, the suffix *html* in the file name indicates that it is a HTML document.

2.5.4 What is HTTP?

Hypertext Transfer Protocol (HTTP) defines how hypermedia files get transferred from the WWW servers to your computer. A computer that stores and transmits WWW documents has a software called HTTP server. Browsers on your computer have HTTP client software built in so that you can display these documents on your computer.

2.5.5 What is HTML?

If you want to put your information on the WWW you will need to use the Hypertext Mark-up Language or HTML. Every Web page is an HTML document. An

HTML document consists of text and HTML codes which define how the text will look. If you want to create your pages on the Web you will need to know HTML or procure a program that generates HTML files.

2.5.6 How to retrieve information from the Web?

To access information on the Web, you need a special program called the browser. The browser has the ability to render HTML documents and also follow links embedded in these documents. You start the session at a special document called the Homepage. Depending on the Internet Service Provider, you may or may not be able to choose the homepage you want to start with. Starting from your page, you follow the links, from one document to another until you get the information you need. Some URL's will have *ftp* in the place of *http* as the protocol to be used to access the document. In such cases, the browser will automatically invoke *ftp* and transfer the file to your computer.

2.5.7 What is a Homepage?

The Homepage is the starting page that is displayed when you access the World Wide Web. The home page is designed to have links to the various documents on the site. You can create your own home page with links to the documents you want to access.

2.5.8 What is a WWW Browser?

Browsers are programmes which help to navigate the Web. The browser can be on your ISP's computer or it can be a programme that you can install on your PC. Browsers are of different types, some of them do not permit graphics or other media while others are multimedia capable browsers.

2.5.9 What are the Browsers available?

Mosaic, Netscape, Lynx are the most commonly available browsers. To evaluate the browsers you should first be clear about the nature of the information you want to access (text or graphics or sound or video, etc.) and the computing power available on your own system. If you are interested in multimedia, then you would probably go in for a multimedia browser such as the Netscape or Mosaic. If you are only interested in text, or your resources do not permit multimedia, then you should probably choose Lynx. Also, it is important to note that the browser that is available is often dictated by the Internet Service Provider.

2.5.10 Can you name some Web documents of interest?

URL	Description
http://www.internic.net/infoguide/gopher/about-internet.html	The definitive source for information regarding the Internet. This should be the first site you should visit.
http://www.interramp.com/cool/ cool.html	The place to start if you want to explore the Web. To get the most out of your exploration, you would need a graphical browser.
http://www. bsdi.com/server/doc/web-info.html	The definitive source for information regarding the World Wide Web. Novices may be overwhelmed by the jargon.
http://www.hot wired.com	An online magazine for WWW users. A must for all but the casual user.
http://www.3w.com/3windex.htm	An online magazine on WWW. A good starting place for beginners.

URL	Description
http://chronicle.merit.edu	The Chronicle of Higher Education. Look here for an electronic copy of the journal.
http://www.cup.ac.uk	The Cambridge University Press. Look here for an online catalogue as well as links to other publishers.
http://www.sjmercury.com	Headline news stories provided by San Jose Mercury News. Somewhat regional but a good place to start for researchers of topical major international news.
http://www.childsoft.com/childsoft	Look here for high quality educational software for children.
http://www.cs.cmu.edu:8001/Web/books.html	Search for books by author or title.
http://info.isoc.org/home.html	Look here for policies on the use of Internet.

3. CONCLUSION

The modules presented above deal with the technical aspects of Internet. There is another important aspect to Internet usage which is of interest to library and information workers. This is concerned with the administrative and management aspect of use of the Internet by the organisation.

Some of the questions which would arise in this connection are those dealing with accounting and budgeting for the use of the Internet, the types of facilities to be provided for the members of the organisation, whether the library and information centre should charge for the services it provides etc. These issues are necessarily addressed on a case by case basis.

4. BIBLIOGRAPHY

1. Eager, Bill. Using the Internet. Indianapolis, Que Corporation, 1994.

2. Jacobson, Robert L. 'Taming the Internet,' The Chronicle of Higher Education, 1995, April 21, A29.

5. REFERENCES

1. New Riders' Official World Wide Web Yellow Pages. Indianapolis; New Riders Publishing, 1995.

2. PC Magazine (India) October 15, 1995 and October 30, 1995. Covers the Internet and the tools for using the Internet.

3. Weibel, Stuart L. The World Wide Web and Emerging Internet Resource Discovery Standards for Scholarly Literature. Library Trends, 1995, 43(4), 627.

INDEX TO QUESTIONS