DRTC Annual Seminar on Electronic Sources of Information
1-3 March 2000

Paper: AF

SCHOLARLY COMMUNICATION AND THE INTERNET

Meenakshi Joshi, Documentation Research and Training Centre, Indian Statistical Institute, 8th Mile, Mysore Road, Bangalore 560 059

Information technologies of the 20th century have affected the process of scholarly communication. They have not only changed the way in which two scholars interact with each other but also the nature of contributions of others like publishers, booksellers and research libraries. Internet offers a variety of services, which can help achieve communication in an efficient and effective manner. This paper tries to trace the evolution of Internet from the perspective of scholarly communication.

1. INTRODUCTION

Communication has taken place in many ways and in many forms throughout the history of mankind. Beginning with the sign languages or expressive body languages it has evolved through spoken words, written words and printed words all the way to electronic form. Electronic communication began with the telegraph. Telephone, radio, television were the other media of electronic communication to follow. Some important media used to store and disseminate intellectual information were the microforms like microfiche and microfilm. The introduction of computers ushered in the age of punched cards, magnetic tapes, floppy disks, CD-ROMs and now DVDs. Computer networks was a step forward in speeding up the process of communication across the globe.

During the last twenty years, the techniques for information handling - Information Technology or IT - have developed very considerably. Information Technology involves a combination of computer techniques, telecommunication techniques and electronics that is used in the production, distribution, storing and retrieval of information. The main features of the developments within the field of information technology are:

Increased computer power - resulting in quicker and cheaper computer handling,

Cheaper storage media, for example, optical media such as CD-ROM (Compact Disc-Read Only Memory),

The digitalization of information - text, pictures, speech, sound and video - facilitating the integration of different types of information,

Increased use of networks - local, national and international - for data transmission and greater network capacity, and

Smaller apparatus.

The Internet is the greatest innovation in human communication since the invention of printing using movable metallic types. It has made the world a global village. People can communicate with each other at global distances. The Net is the gateway to world's information resources like databases, e-journals and catalogues and archives of thousands of libraries. Full text, catalogues, graphics, images, audio and video clips and countless other types of information can be accessed, exchanged, e-mailed, downloaded etc. It can be used for video-conferencing and as a telephone medium. On the millions of host computers, millions of individuals use Internet for accessing/exchanging some or all of the above types of information sources spread out globally. [5]

2. INTERNET - A MEDIUM FOR SCHOLARLY COMMUNICATION

"Scholarly communication" can be broadly defined, as the various means by which information exchange takes place in academia--including the formal publication of research; informal discourse among colleagues; class discussions and lectures; data retrieval through local and global networks; and continuing access to the scholarly record in print and digital libraries. [6] Until very recently, scholarly information needs have been served almost exclusively by the technology of printing. The Internet now offers many different resources, which can be effectively used for scholarly communication.

2.1 Important Internet Resources

The important resources available on the Internet include

1. E-mail - It can be achieved independent of global time and without the receiver having to be in place besides some other benefits. Many people can be reached through the "mailing lists". [8]

2. Usenet - A system of discussion groups in which individual articles are distributed throughout the network. Each subject collection of posted articles is known as a newsgroup.

3. Ftp service - allows one to copy files from one computer to another in either direction.

4. Gopher

5. Telnet - a means to log in and use a remote computer allowing to use specific applications and data residing in the remote computer.

6. World Wide Web - A hypertext based system to facilitate worldwide information sharing. What makes the web page special is they contain links to other web pages.

3. SCHOLARLY COMMUNICATION AND INTERNET

Scholarly activity - research - creates a need to spread and share information about the results, methods, new processes and products. The findings are shared and evaluated by colleagues and students. There is a need for both informal and formal communication, both locally and on a worldwide scale. [9] Scholarly activity ordinarily involves communication to colleagues and students of results, observations, and interpretations emerging from one's

research. It usually culminates in the publication of the results. It also involves a lot of communication during the time when research is going on. The Internet has made this much easier now. Today informal communication takes place to a great extent through the electronic invisible colleges, based on the use of e-mail, electronic conferences and discussion groups.

3.1 E-publishing

The most important and established means of making the scholarly work known to the world is through publishing an article in a journal. The advent of World Wide Web made it possible to overcome some of the inherent disadvantages of a printed journal. Traditional publication costs have risen sharply within the last few years, whereas computing and communication costs have fallen. The vast majority of research workers produce material for publication with the help of word-processors. The text (and in some cases the graphics) is available in digital form. There are a number of alternatives for electronic publication. They include publishing on some form of transportable memory such as an optical disk, for example CD-ROM or CD-I, or making full-text, (with or without graphics and pictures) available online via communication networks. The Internet offers an obvious potential for electronic publishing and distribution of scientific and technical information. Scholars now have access to global networks from their workstations and personal computers. These networks are increasing in speed and capacity. We have moved from the kilobit networks to megabit networks and are now approaching the era of the Gigabit networks. The research networks offer a possibility for the electronic publishing of
information.

In some instances, research results are not published by conventional, printed means because the results can't be printed and still be meaningful. This is true, for example, when the results are three dimensional, graphic, moving simulations, or animations, or when the outputs are dynamic visual representations of variable processes or theoretical constructs. Traditional, printed publication is completely inadequate for disseminating research of this kind. Yet, this research should be included in what we refer to as "the literature".

3.2 Problems of the Internet

It is true that the Internet is presently the fastest mode of scholarly communication available. It is a data transport system, which has to physically transfer the data over some medium. It is coming under severe strain from the perspective of newer applications. Some of the issues, which are forcing the research community to have a second thought regarding Internet as a means of communication, are

1. Rapidly growing demand for bandwidth.

2. Rapidly increasing number of hosts and networks participating in Internet.

3. Major shifts in the business model and structure of the Internet Service Providers (ISPs).

4. Internet has evolved into an almost completely commercial operation.

5. Connectivity to the network by individuals and smaller institutions becoming a serious bottleneck.

6. The problems of having more than one ISPs - peering issue being the most important one.

7. Assignment and management of network numbers - almost anything can be now connected to the Internet.

8. Overall reliability and robustness of the Internet - physical disruptions, problems with software, configuration and operational errors etc.

9. Attacks on computers - the Net infrastructure becoming both a victim and a tool of attack on computers. [1].

3.3 Success of the Internet

Inspite of all its problems Internet has been hugely successful. From the scholarly communication perspective also it has gained importance. In the early 1990s the higher education community in the US accepted the vision of electronic access to share, network based information resources. Many professional and scholarly societies are offering large amount of content on the Net. This is in addition to the innumerable e-journals now available. Many newer, content oriented applications are being developed. Some of these are

1. **Web indexing services** - which are like a global knowledge database. These services try to index all the pages available on the WWW. Information is chosen by a gathering program called a "spider" or a "robot". This program is given a seed of URLs and sent to bring back data on those pages and pages linked to those pages often to a certain depth or number of times to follow additional links. Some services add pages manually or index pages whose URLs are submitted by the page's developer. Information stored may include the full text or just the important or "weighty" words in the web page as determined by a formula in the gathering program. Information may be stored with the accompanying HTML tag: title, headings, alternate image title, mete tags, or it may be stored as a text-file copy of the page. The search engine matches the search request against the page and determines the score for the page based on an algorithm or formula.

2. **Internet telephony** - which replaces the costly long-distance voice telephone calls with digitized, packetised voice carried over the Internet. It can be conducted end-to-end between two computers connected directly to the Internet. The gateways bridge an Internet telephony service and voice phone network. This allows one to send messages through Internet to a person who does not have a computer connected to the Internet but has only a telephone.

3. **Streaming audio and video technologies** - which basically use the same packetising approach, used for Internet telephony. These technologies try to send enough information before play starts at the recipient's machine, so that buffering covers network delays and occasional packet losses. Streaming a-v suffers from bandwidth constraints and packet loss and is an extremely inefficient provider of "broadcast" material to large number of network users at once.

4. **Push technology** - popularized by Marimba and Pointcast- releases one from navigating the web for information. Instead user establishes an interest profile and the pertinent information is sent by a network server on an ongoing basis. It is relatively inefficient suffering from the same information broadcasting limitations that affect streaming audio and video. [1].

4 INTERNET NEEDS OF THE SCHOLARLY COMMUNITY

Since much teaching, learning, and collaborative research may require real-time multimedia and high-bandwidth interconnection, sufficient network infrastructure to support such applications will be required. Some of the possibilities envisioned by the higher education and research community, which are based on the use of Internet, are

Distributed learning modules: Conceptually, teachers and students can share materials in cyberspace with students learning in a self-directed manner under the supervision of an educational system or teacher. Development of some kind of Instructional Management System (IMS), a standard process for using the Internet in developing and delivering learning packages and tracking outcomes should be possible.

New ways to envision and retrieve information: In the future, today's text-oriented models of information structure could be replaced by interactive pictures of information structure. (illustrations of interlinked and explodable animal forms, for example). This will require very high bandwidth availability.

Virtual environment sharing: Sometimes called tele-immersion, participants in teleconferences could share the perception that everyone was in the same physical place, possibly with virtual (but somewhat real) models of shared work objects such as architectural models or multimedia storyboards.

Virtual laboratory: A virtual laboratory would allow scientists in a number of different physical locations, each with unique expertise, computing resources, and/or data to collaborate efficiently in an ongoing way. For example, data from a telescope might be transmitted in real time to a group of astronomers working at different sites around the world. They might collaborate in real time to analyze the data and to decide how to adjust the aim of the telescope to optimize the value of the data being collected during the session. The same data might be multicast to amateur astronomers who have "subscribed" to the appropriate "channel". Virtual laboratories are also envisioned for the design and manufacturing of complex systems such as airplanes and for studying and forecasting weather patterns. [4].

Current Internet cannot reliably support the existing massive volume of primarily character based, interactive traffic or web access. Yet there is a whole selection of new highly interactive, and multimedia rich Internet applications, of great interest to the research and educational community. These are important to those who are focussed on ways in which networks and networked information can potentially transform teaching, learning and scholarly communication. The newer visions of Internet use demand a quality of service that cannot be extracted from the current Internet. That quality of service sometimes may depend solely on the high-probability availability of raw bandwidth, but it often depends on the availability of differentiated network services designed, for example, to ensure against problems of latency in delivering 30-frames-per-second, synchronized video from distant server to personal computer. Many people involved in higher education envision a future where, on-line tools for communication and collaboration and on-line learning resources will be used, not only to strengthen the traditional classroom, library, and laboratory experience, but also to extend the higher education's reach with convenient and flexible anytime, anyplace modes of instruction, research, and public service.

5 NEW INITIATIVES

The need for differentiated network services and related quality-of-service guarantees loomed large for the leaders from higher education information technology services and information resources communities. Some of them gathered in 1995 at a conference in Monterey, California, with experts from network industry and the federal and regional communities. The differentiated network service needs and their implications for higher education inspired a group of conference attendees to initiate a movement that evolved into the Internet2 project. The resulting series of meetings and workshops motivated over 40 university CIOs to announce the Internet2 project in the fall of 1996. The Networking and Telecommunications Task Force of EDUCOM, an association of higher educational institutions concerned with information technology applications, developed Internet2 (I2 for short) during 1997. The University Corporation for Advanced Internet Development (UCAID), which looks after the I2, was established in late 1997. [3]

Internet2 is a new effort to create the next generation of Internet computing infrastructure and the applications that use that infrastructure. It could be popular for two basic reasons: enhanced communication between research institutions and a new distribution system for academic instruction.

Some new technologies used in this and other new initiatives are:

1. ultra fast all-optical fibre networks.

2. high speed routers and switches.

3. new version of Internet protocol to prevent shortage of Internet addresses.

4. multicast technology to conserve bandwidth by disseminating data to multiple users at the same time.

5. software for replicating information throughout the network to reduce bottlenecks.

6. software for measuring network performance.

7. software to ensure reliability and security of information transmitted over the Internet. [5]

5.1 Goals of Internet2 initiative

The intent of I2 is not to develop a separate replacement for the present Internet and then disconnect the participating institutions from the existing worldwide Internet. The goal is to create a testbed for advanced applications. The I2 sites will continue to be connected to the present Internet. The main goals of this project are:

1. enable a new generation of applications
2. recreate leading edge research and education network capability
3. transfer new capabilities to the global production Internet.

There are some additional specific objectives, which relate to the new applications and the network services like Quality of service (QoS) guarantee. [2]

5.2 Internet2 initiatives

The major I2 initiatives are

1) **QBone (Quality of service Backbone testbed)** - the aim is to find technologies to allow computer network administrators to assign priorities to certain Internet traffic. Ultimately, QBone will give users a guarantee that their certain piece of Internet traffic, like video-on-demand will get good service without random delays. With QBone, user essentially asks the network for a contract or a guarantee that this traffic will not be delayed or disrupted. This is not a contract in the legal sense, but more of an assurance of delivery of certain level of service. QBone effort aims to bring quality of service across networks rather than in a single network where it is relatively easy.

2) **I2-DSI (Internet2 Distributed Storage Infrastructure)** - which is a replicated hosting service for Internet content and applications. The channels are replicated across a distributed infrastructure consisting of servers with substantial processor and storage resources. Each user request is directed to the server closest to the requesting client in networking terms. The result is that network traffic is kept local and load is balanced among the distributed servers.

3) **I2-DVN** (Internet2 Digital Video Network) - to provide high quality of digital video services to the I2 community through advanced technologies. The term video includes not only traditional video but also simulations, animations, virtual reality movies, images with audio sound tracks and other types of digital media objects.

4) **I2-MI** (Internet2 Middleware Initiative) - The term middleware can be defined as those sets of tools and data that help applications use networked resources and services. The need for middleware comes from increasing number of applications, in the customization within those applications and the number of locations in our environment. the middleware components include some core functionalities like identifiers, authentication and authorization. It also includes other functionalities like secure multicast, bandwidth brokering, services for omnipresent computing,

services for research computing and administrative computing. [2]

5.3 Participants in the Internet2 project

Begun in 1996 by 34 US research universities, I2 today has over 160 member universities which are working with corporate and affiliate members. There are many non-US members also, mainly from Europe and others from Japan, Middle east etc. The major universities, which are participating in I2 include the University of California, Massachusetts Institute of Technology, Purdue University etc. The big corporates like Microsoft and IBM are also part of this initiative. [2]

5.4 Beginning of Internet2

Abilene is a project of the University Corporation for Advanced Internet Development (UCAID). This most advanced research and education network in the US was launched on February 24, 1999 delivering high performance network services to 37 universities. It serves as a testbed for newer applications such as distance learning, tele-medicine and digital libraries. A demonstration in February 1999 of Abilene showed the potential of a quality of service Internet,

when a doctor in Ohio performed gall-bladder surgery with computer network video assistance from a surgeon in Washington, D.C.

As recently as the fall of 1999, professors in the US and Japan together taught a course to 60 graduate students from three universities. University of Wisconsin, US, KEIO University, Japan and NARA Institute of Science and Technology (NAIST), Japan jointly developed a graduate school course using the next generation Internet technology and the class archiving technology. Two professors from the two universities served as guest lecturers for each other's graduate students. The advanced networking link enabled much higher quality video and audio than are typically possible today on the World Wide Web. Due to the full-screen TV-quality pictures and high-fidelity audio, a professor lecturing at one site could call on and answer questions from students in either country. Students could also review lectures throughout the semester because videos of lectures were stored digitally online. The network link between the two classrooms used IPv6, which is the next generation protocol, rather than the IPv4 used in most of today's commercial Internet. They used multicast and DVTS technology to transmit the Digital Video stream over the very high speed Internet test bed.

6 CONCLUSION

Internet, which started as an experiment for defence purposes, offered immense capabilities for scholarly communication to become faster and effective. But as it evolved to become an almost commercial operation, the research and higher education community found it rather inefficient to use for scholarly communication. Conceived by a coalition of universities, government institutions and high-tech businesses, the goals of Internet2 are to take today's Internet to next level and give its members a new, powerful way to communicate and share information. Although I2 is going to be the semi-private network of its members, once the technology is developed, it will be used to upgrade the existing Internet.

7 REFERENCES

1. Lynch, Clifford: The evolving Internet: applications and network service infrastructure. JASIS, 49(11), 1998, p. 961-972.

2. http://www.internet2.edu/

3. http://www.dlib.org/dlib/april98/graves/04graves.html

4. http:/www.whatis.com/

5. A. Neelameghan - Notes given by him and discussions with him.

6. http://libweb-01.ucsc.edu/scomm/

7. http://educate.lib.chalmers.se/cth/pathciv/course/channel/email.html

8. http://educate.lib.chalmers.se/cth/pathciv/texts/chap3.html