Information on the Internet: How selective should we be?

ROBERT W. ALLAN and PHILIP S. KOSTENBADER
Lafayette College, Easton, Pennsylvania

Vast amounts of information are available on the Internet, and programs such as Mosaic provide users with easy access to text and multimedia files. Once a user has set up a host system, consideration must be given to what information will be made available. In addition to legal restrictions, the interests of potential users (those accessing the host system) should be considered when determining the kinds of information to be made available. If those maintaining these host systems are somewhat selective when making information availability decisions, the potential information "glut" may be attenuated.

The Internet is a global network connecting a collection of more than 30,000 smaller networks worldwide. Most of these smaller networks are operated by universities, colleges, government agencies, and a newly burgeoning group of private corporations. These networks have unique addresses that presently identify more than 1 million hosts and probably more than 3.5 million users worldwide (the unique address assigned to each institutional host is called an Internet Protocol, or IP, address). If you have an IP address, you are part of the Internet and have access to all of the information offered by any other host on the Internet. With so many networks and worldwide users, a lot of information is available. Indeed, after logging on to the Internet, it is almost as if the user has picked up a magazine with an infinite number of pages and no table of contents.

In a recent news article in *Science*, Pool (1994) suggests that we are already deeply ensconced in what he calls an Internet "info-glut." Before 1992, during the decade of Internet development, roughly 750,000 hosts were identified on the Internet. From January 1991 to January 1992, more than 350,000 new hosts were given addresses on the Internet, with each of these hosts presumably offering information to all Internet users (Marine, 1992). Up through 1992, this 46% single-year growth was the largest ever experienced by the Internet, and through 1994 this growth pattern has continued. Part of this growth may be due to the easy availability of network interface programs such as Mosaic. Users no longer need to know an esoteric set of commands to access information on the Internet; now they simply access Mosaic (a graphical user interface), click on a starting option, and access is provided to the files on remote host systems. In addition, most hosts provide lists of other interesting Internet sites, which, again, may be accessed by simply pointing and clicking. As a result, the Internet is

healthy and growing. However, is it growing too quickly, and what are the implications of this growth for the local user?

Effects of Internet Growth

Internet users can access text files, pictures (both black and white or color), video clips (with and without sound), and sound files. Perhaps the biggest limitation of the Internet is that these files must all be transferred digitally via network equipment possessing limited operating characteristics. With access to the Internet becoming easier (via programs like Mosaic), and with an exponential growth in the number of hosts and users, the network equipment has to service more requests for data transfer than it was designed to handle effectively. Of course, faster data transfer equipment is available and could be installed, but the new equipment is expensive and would greatly increase the cost of data transfer, which, in turn, would increase the institutional network operating costs.

Although the major "backbone" of the Internet (in the United States) is largely funded by the National Science Foundation (NSF), all the regional networks into which institutions tap directly are funded by those participating institutions. For example, a small college with a local-area network may pay as much as \$20,000 each year to maintain access to the Internet through a state-wide internetwork. These costs will increase as the demand for increased speed and transfer volume increases, since the regional networks must pay for the equipment upgrades. Some of the cost increase may be offset by the registration of additional hosts to use the Internet. More hosts means that the costs can be spread over a larger user base; however, the cost of transferring exponentially increasing numbers of files may exceed the ability of host institutions to fund equipment that can handle the increasing transfer load.

This problem is compounded by the fact that individual users, with programs such as Mosaic, are increasingly tempted to transfer multimedia files that are substantially larger than the text files that used to be the

Reprint requests should be sent to R. W. Allan, Department of Psychology, Lafayette College, Easton, PA 18042.

bulk of what traveled over the Internet. For example, it is not unusual for users to transfer 2- to 4-Mb sound files, and, at 56 kbits per second, this is a time-consuming transfer that eats up valuable Internet resources. Frequently, Mosaic users will "browse" the Internet, transferring large files, thereby creating literal traffic jams on the Internet, only to discover that the files they have transferred are of little use.

Also adding to this problem is the ease with which users can create their own host systems on which they may offer access to picture, video, and auditory files. With a program like Mosaic, host systems offer access to personal information, descriptions of departments and institutions, discussions of research and teaching, and almost anything else imaginable. However, are there any limitations on what a user may offer via the Internet?

Internet Limitations

There are some legal limitations on what users may make available on the Internet. First, since the major backbone of the Internet is funded by the NSF, purely commercial use is not allowed. If a company that markets a data analysis package, for example, wants to advertise a new product or release, the Internet would probably not be an appropriate outlet. However, if the same company wanted to inform users of a bug in their software, the Internet would probably be a good forum. If a company wants to use the Internet to advertise, there are commercial providers that give companies access to global networking services (Krol, 1994).

Second, export laws may apply to material you, as a user, might make available to others on your system. Before making programs available on the Internet, check with your regional network people to see whether export laws are being violated. After all, if you write program code which another researcher copies to a remote computer via the Internet, you are exporting a product.

Third, you must consider property rights and licenses. For example, if you make copyrighted photographs available to other users, you are probably violating copyright law. Once again, check with your regional network director before making copyrighted or licensed material available from your host machine (Krol, 1994).

Once information being offered meets all legal criteria, the ethics of the Internet (Krol, 1994) suggest that there should not be any further restrictions on what you may offer on the Internet; that is, the user must make the final decision. An important question logically follows: Without dictating what should and should not be made available, are there ways in which users can be more responsible about what they make available on the Internet, thereby reducing the traffic problem? For example, if every user cut the number or size of available files by 25%, this would immediately translate into a similar reduction in transfer time and traffic on the Internet.

Because the Internet respects the individual's right to be creative and to make almost anything he/she has pro-

duced available to other users, we would suggest that each user consider at least four additional issues: (1) Who might be interested in what you have to present on your host system? In other words, can you identify your target audience? If not, perhaps you shouldn't be offering much of anything. (2) Do these potential users have access to the Internet? If not, then you might consider some other method of transmitting your material (e.g., via journals). (3) More specifically, if you're offering pictures or sound files, do your potential users have access to Mosaic (or some comparable program)? If not, then it makes little sense to make this kind of information available on your host system. (4) Would your potential users think to look to the Internet for the information and files that you are making available? For example, if people accessing your host system are researchers interested in your latest accomplishments, would they think to look on the Internet for summaries of your experimental work? Of course, you must be careful about making complete transcripts of your publications available; this would probably represent a violation of your copyright agreement with the journal publisher.

If each host could answer these questions before making information available, transfer time and traffic on the Internet would be dramatically reduced. Users would be less likely to access large files only to discover that they were not useful; instead, each of us, as potential purveyors, should impose some limits before file transfers begin.

Conclusions

It is probably safe to predict that the Internet will continue to evolve and grow to meet the increased demands of users. Data transfer speeds will continue to increase, the cost of transferring data will probably increase, but additional users will probably pick up much of that increased cost. It is also safe to predict that more and more information will be available in more and more sophisticated, multimedia formats. With newly funded efforts to catalog and index what is available on the Internet (Pool, 1994), productive use of this global information "superhighway" will be realized (we may actually have a "table of contents" for all the information on the Internet). In addition, if each user thinks carefully about what to make available on the Internet, always keeping the target audience in mind, the probability of an "info-glut" should be minimized.

REFERENCES

- KROL, E. (1994). *The whole Internet user's guide and catalog*. Sebastopol, CA: O'Reilly & Associates.
- MARINE, A. (1992). *Internet: Getting started*. Menlo Park, CA: SRI International.
- POOL, R. (1994, October). Turning an info-glut into a library. *Science*, 266, 20-22.

(Manuscript received January 10, 1995;
accepted for publication January 31, 1995.)