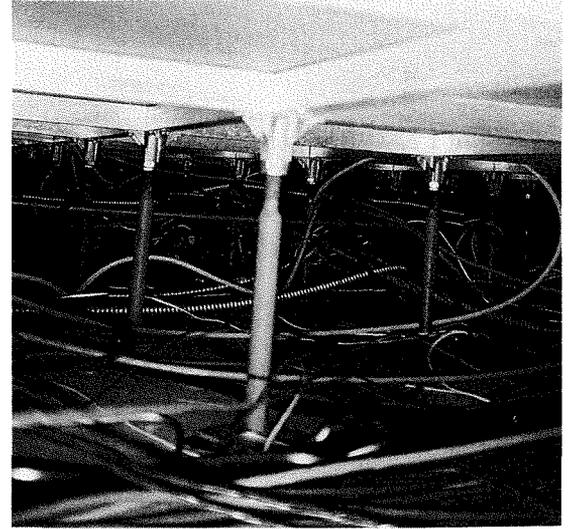


“Infinite” Information on the Internet

by Roy D. Williams



One might find the Internet worthwhile because it connects millions of computers, so that any one of them can exchange data with any other: but the really important thing is that it connects millions of *people*, enabling them to communicate with one another in ways they never could before. Like the telephone and jet aircraft, it will change our lives. The Net is a tool you can use to enhance your professional, social, and perhaps even spiritual lives. Don't expect the miracles proclaimed by the media, but don't be afraid of it either.

Like many new technologies, the Internet started out as something for the elite, but now it's becoming ordinary, and soon it will be necessary. In 1878, when the telephone was a new technology, it was very elite, very special, to have a telephone; it was a beautiful, expensive object of wood and brass—but now we take cheap, plastic telephones to the beach. The Net is still special, but it's getting less so—I'll bet that some people reading this haven't just taken a laptop computer to bed, but have taken to bed a laptop connected to the Internet! It's getting integrated into the fabric of everyday life—ubiquitous computing, ubiquitous Internet. (But some things never change: at left are phone cables along Broadway in Manhattan in 1889; above is a look beneath the machine-room floor in the Booth Computing Center.)

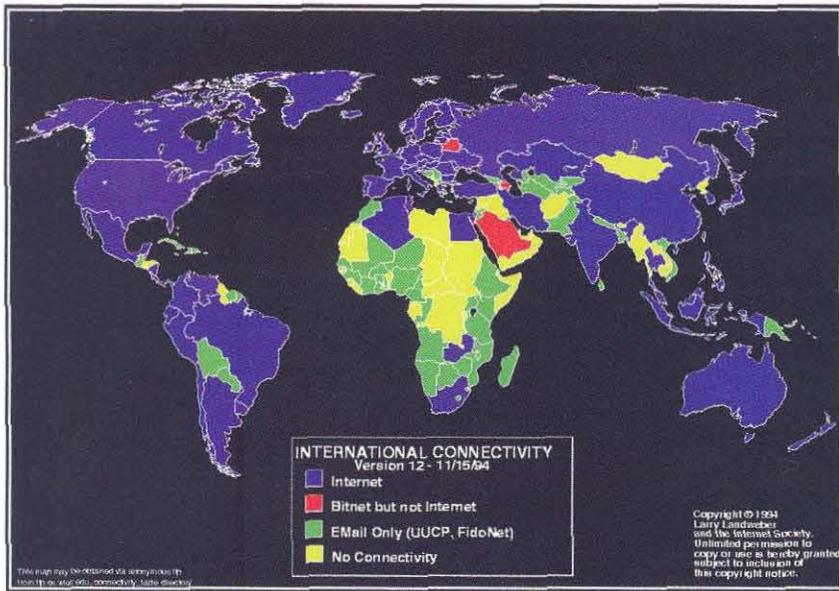
In this article, I'll talk briefly about the Internet's infrastructure—the hardware and how it works. Then I'll describe some of the services that are available on the Net, and give a feeling for cybersurfing—the art and recreation of

exploring the Net. And finally, I'll mention some implications for the future—issues of publishing, commerce, advertising, and privacy.

The number of machines on the Internet has exploded in the last year—it's doubling every 18 months—and if it were possible for this rate of increase to continue unabated, there will be a computer on the Net for every human being in the world by the year 2010! There is no center to the Internet, except in certain minor ways, and there is no hierarchy of control. It is a radical decentralization that works, and I think this is something unusual in human creations. Indeed, this lack of a center is vital to the robustness of the Internet; it was built during the Cold War to resist nuclear attack, and yet it's ironic to contrast the military parent with the free-spirited child to which it has, perhaps inadvertently, given birth. From its conception 25 years ago at UCLA until fairly recently, the Internet was run by ARPA (the Advanced Research Projects Agency, an arm of the Defense Department). Commercial activity was first allowed just a few years ago, and now that the technology is mature and demand assured, the private sector has taken over completely. Although the Internet is worldwide, the United States has a large and very clear lead in the computer-networking industry, which will bring in many billions of dollars to this country over the next few years. In my opinion, this lead exists because the government has had sufficient vision to encourage the relevant research since the early days, together with the foresight to give up control when necessary.

The Internet is actually a network of computer

Opposite: If the Internet is the information superhighway, then the proliferation of telephone service in the 1880s must have been the information macadam street.



Speed (bits per second)	Device	Time to transmit Complete Works of Shakespeare (5 Megabytes)
1,200	old modem	9 hours
14,400	new modem	45 minutes
128,000	double ISDN	5 minutes
1,544,000	T1	26 seconds
45,000,000	T3 (NSF backbone)	1 second
155,000,000	ATM	0.25 second
800,000,000	HIPPI	0.05 second

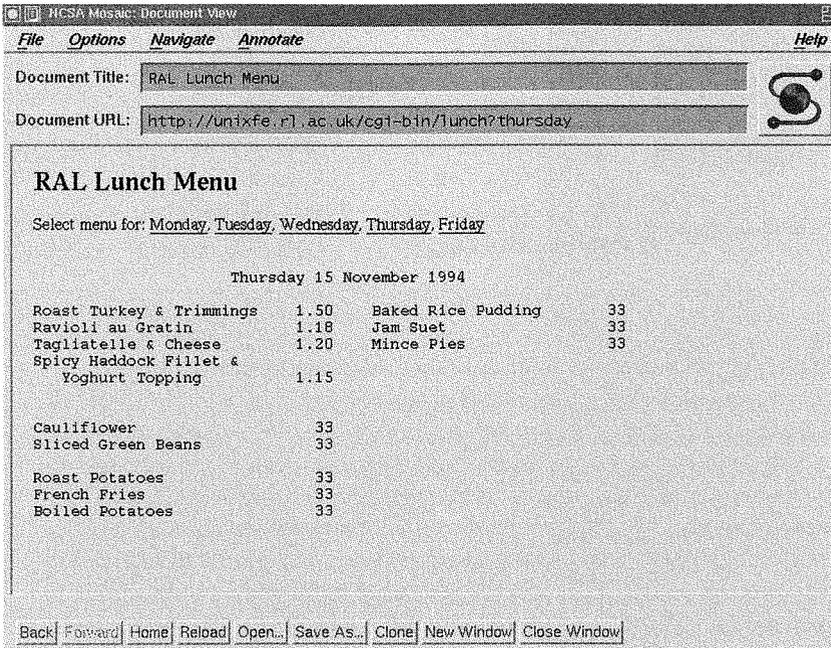
Top: The world according to Internet. (Bitnet is a more cumbersome cousin of the Internet.)

Data transmission speeds (above) depend on your equipment. Caltech's campus network runs at T1 speed over fiber-optic lines—the orange cables in the steam tunnels (right). The Internet's "backbone," or main transmission lines, run at T3. (By comparison, a CD-ROM delivers one million bits per second.) Another frequently used unit of information is the "Library of Congress," which is 10^{15} bits, or one million billion bits. It would take almost 35 hours to transmit a Library of Congress at HiPPI speeds.



networks connected by data links, the speed of which determines the user's frustration level. A telephone line equipped with a fast modem can transmit data at perhaps 20,000 bits per second; if we take the complete works of Shakespeare (about 40 million bits) as a unit of information, then it would take less than an hour to be transferred at that speed, which is clearly less time than it would take to recite! But this transmission rate is not satisfactory to today's cybersurfer. It may be that we only want a small quote from Shakespeare, and if we need to get the complete works before we can see the quote, then an hour is definitely too long. Furthermore, images, sounds, and moving pictures take an inordinate number of bits to represent: one full-color image can easily require millions of bits, and compressed TV-quality video needs five million bits per second. (The dedicated Net user might be tempted by ISDN, a service offered by the phone company that is about ten times faster than a normal telephone.) Paul Messina, Caltech's assistant vice president for scientific computing, and I are working on a superfast channel called HiPPI (High-Performance Parallel Interface), which runs at 800 million bits per second. This type of channel allows supercomputers to exchange huge volumes of data at the fastest possible rates, and can also be used as a backbone service, simultaneously transmitting many customers' much smaller volumes of data over a shared trunk line.

Most of the strands connecting the Internet are fiber-optic cables, such as these at left in Caltech's fabled steam tunnels. (Fiber-optic cables are faster, and carry more signals at once, than copper wires.) These cables connect computers to one another and to switching machines called routers. In order to get sent across the Net, your data is broken up into "packets," each of which carries your computer's return address and the address of the recipient computer. A router reads the address and sends the packet to another router to which it is physically connected, and which is (hopefully) closer to the destination. To make this decision about which machine might be closer, the router needs some knowledge of its local environment; this knowledge is updated without human intervention, so that when part of the Internet is damaged, data automatically flows around the crippled link. Thus two messages from one machine—or even two packets from the same message—may reach the same destination by very different routes. The Net's protocol (known as TCP/IP) also assumes that packets may get lost, so each time a destination machine receives a packet, it sends an acknowledgment



Eating a virtual lunch at the Rutherford-Appleton Labs in Oxfordshire, England (above) required a 17-stop journey (right). The top line of each entry is the Internet address of a computer, followed in parentheses by its IP (for Internet Protocol) address, which is what the router actually reads. The second line of each entry shows the machine's type and location. The three columns of numbers have to do with the delay times, in milliseconds, at each step. But use of technology doesn't necessarily imply literacy—somebody at Sprint can't spell "Anaheim."

```
% traceroute unixfe.rl.ac.uk
traceroute to unixfe.rl.ac.uk (130.246.12.16)
sampson-151 (131.215.151.254)          2 ms  2 ms  1 ms
  Caltech Concurrent Supercomputing Facilities
cisco (131.215.145.254)                 5 ms  4 ms  5 ms
  CALTECH CCGS/CCO cisco box
cerfnet-cit-gw.ln.net (130.152.104.1)  5 ms  3 ms  4 ms
  CERFnet Caltech Backbone Cisco Router
134.24.120.100 (134.24.120.100)        21 ms 25 ms 23 ms
  Dr. Zog CERFnet SDSC CISCO Router
mobydick.cerf.net (198.17.46.153)      18 ms 10 ms 14 ms
  (San Diego)
ucop-sdsc.cerf.net (134.245.2.112)     45 ms 46 ms 43 ms
  UCOP-SDSC CERFnet Cisco Router
sl-ana-3-S26-T1.sprintlink.net (144.228.73.81) 59 ms 61 ms 57 ms
  (Anaheim)
sl-ana-2-F00.sprintlink.net (144.228.70.2) 58 ms 74 ms 64 ms
  SprintLink DS-3 Backbone Router in Anaheim, CA.
sl-stk-6-H20-T3.sprintlink.net (144.228.10.25) 127 ms 181 ms 61 ms
  SprintLink DS-3 Backbone Router in Stockton, CA.
sl-stk-5-F00.sprintlink.net (144.228.8.233) 177 ms 147 ms 68 ms
  Sprint SL-STK-5 router in Stockton, CA
sl-dc-6-H10-T3.sprintlink.net (144.228.10.1) 149 ms 122 ms
  (Washington DC)
icm-dc-1-F00.lcp.net (144.228.20.101)   129 ms 132 ms 121 ms
  Sprint ICM-DC-1 router in Washington, DC
icm-london-1-S1-1984.lcp.net (192.157.85.114) 219 ms 224 ms 204 ms
  Sprint ICM-LON-1 router in London, UK
smds-gw.uiccja.net (128.86.1.12)       216 ms 207 ms 213 ms
  (University of London Computing Center)
smds-gw.rlja.net (193.63.203.65)       218 ms * 265 ms
  (Rutherford Lab)
gw.ccl.ac.uk (130.246.4.1)              214 ms 242 ms 227 ms
  (Rutherford Lab)
unixfe.rl.ac.uk (130.246.12.16)        205 ms 216 ms 215 ms
  (Rutherford Lab)
```

back to the sender. If the sender doesn't get the receipt back, it waits a while and sends the same packet again, and again; however, the time between these repeated packets gets longer and longer in order to avoid saturating the system with packets sent to a machine that may be broken. At the destination, the recipient machine collects all the packets and throws out any duplicates. It also puts the packets back together in the right order, if necessary, since they can arrive out of sequence after their journey.

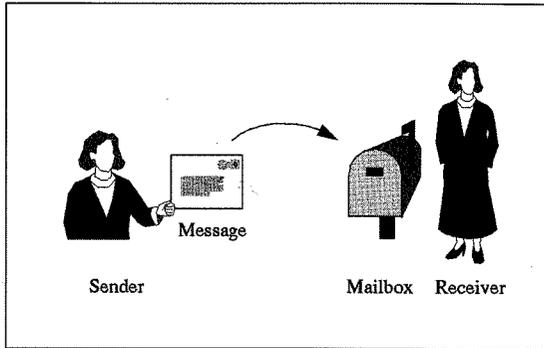
I'll give you an example of routing. At left is a lunch menu for the Rutherford-Appleton Laboratories in Oxfordshire, England, where I used to work. (You can see that the famous jam suet is available. It's a quintessential British pudding—a great solid steaming lump of cholesteric calories and hot custard that stays in your belly for hours.) I connected to the Rutherford Lab to download that menu, and I used a utility called `trace-route` to show where the packets went, as shown in the lower figure. So from Caltech, a router owned by a company called CERFnet sent packets to the San Diego Supercomputing Center. Another router in San Diego took us to Anaheim, where we got on a line owned by Sprint, a long-distance phone service, that took us to London via Stockton, California, and Washington, DC; thence to the University of London Computing Centre and finally through a couple of machines at the Rutherford Lab.

Once your computer is on the Internet, you can use it in four basic ways:

- **Electronic mail**, or e-mail, is like a letter written on paper. You can write and respond to letters, forward them, send copies to others ("cc" the letter), receive form letters from mass mailing lists, and do all the other things you can do with paper letters.
- **Usenet and the "chat" groups**, where one person informally talks to many people, is like addressing a meeting, sitting at a dinner table with others, or standing on a soap-box.
- **Telnet and FTP services** have the feeling of a person talking to a computer: to access a database such as a library catalogue, for example, or to connect to a supercomputer on the other side of the world.
- **The World Wide Web** is like going to the library, except that it's a single, incredibly heterogeneous global library where millions of people have the opportunity to publish whatever they want. You can browse, and get caught up by other topics, and find something that you didn't even know you were looking for. More difficult is to set out on a search and actually find what you were looking for!

Right: Sending e-mail is identical, in concept, to sending physical mail, also known as "snail mail." Only the addresses are different.

Below: When you're typing an e-mail message, it's very easy to make a wisecrack that the recipient may not recognize as a joke because your intonation and facial expression don't transmit over the Net. Thus a whole vocabulary of "Smileys" have evolved to convey your state of mind. Other Smileys are just for fun. Smileys are read by tilting your head to the left.



Some Domain Names

.edu	U.S. Educational Institution
.mit.edu	Caltech of the East
.com	U.S. Commercial
.mil	U.S. Military
aol.com	America Online
.ac.uk	Academic, in the U.K.
.jpl.nasa.gov	JPL (part of NASA)
.edu.au	Educational Institution, Australia

```

:-) smile
:-D grin
;-) wink
:-X kiss
:-X big, wet kiss
:-( frown
;-( going to cry
:'-( crying
%-( confused
:-S nonplussed
:-O surprise
8-O astonishment
:-] forced smile
   (sarcasm)
:-P aside
:-| apathy
:-|| tight-lipped
   anger
:-b sticking out
   tongue
:-(O) shouting
:-@ cursing
!-( black eye
%+{ lost a fight
@-S cybersurfed
   too long

/:-) beret
d:-) baseball cap
<:-) dunce cap
8-) wears glasses
=:-) punk rocker
&:-) curly hair
:~> goatee
':-) toupee
y:-( bad toupee
*~:-) party hat
*~%*} partying too
   much
@%-\ hung over
   (ice bag)

=):-) Uncle Sam
5:-o Elvis
) Cheshire Cat

```

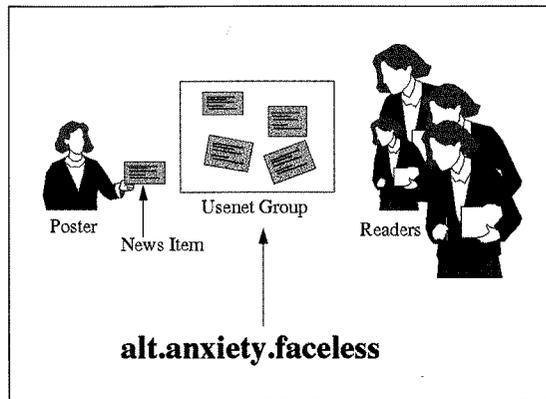
With electronic mail, the sender writes a message and mails it via the Internet to a mailbox on another computer, where it waits for the addressee to read it. The address to which a piece of mail is sent consists of a person's name followed by an "at" sign (@) and a number of words separated by dots. (They are not called *periods*, but *dots*—if you learn anything from this article, you'll learn how to pronounce an e-mail address.) As you read the address from left to right, the words refer to increasingly larger domains, just as when you read down the lines of a snail-mail address, the lines refer to larger and larger domains. So, in the address "jill@rabbit.uea.ac.uk", "rabbit" is the name of the computer containing Jill's mailbox, ".uea" is the University of East Anglia, ".ac" is the academic part of the United Kingdom Internet, and ".uk" is the United Kingdom. In the United States, ".edu" refers to educational institutions, ".com" is commercial organizations, ".mil" is the military, and ".gov" is government. Traditionally, e-mail addresses in the United States have no national domain. You may have noticed that British postage stamps do not have "Great Britain" printed on them, which is because it was the British who first made postage stamps. In exactly the same way, at Caltech we are ".caltech.edu", not ".caltech.edu.us", because the United States invented the Internet.

E-mail is trendy now, but soon it will become necessary. "You don't have an e-mail address yet?" is a disdainful question increasingly heard by the have-nots. For the haves, it is increasingly diffi-

cult to bother to communicate with the one member of the collaboration who does not have e-mail. In science, at least, e-mail makes long-distance collaborations easy. I collaborate with a colleague in Atlanta, for example, and we exchange e-mail two or three times a day. It's a nice interpolation between the formality and solidity of a paper letter and the undocumented ad-libbing of the telephone. You have time to compose the message carefully, but it is then delivered very quickly. And you don't have to talk to answering machines!

The Internet's second aspect, the Usenet, consists of newsgroups, also known as bulletin boards, each of which is devoted to a specific subject. You send a message to the newsgroup, where it gets posted, and anybody who subscribes to the newsgroup can read it. Your posting disappears after days or weeks; otherwise the system would fill up. Newsgroup names look somewhat like e-mail addresses, except that the words get more specific as you go from left to right—for example, "alt.clothing.sneakers". This is a real newsgroup—just one of some 5,000 accessible from Caltech. There's a newsgroup for everything, it seems. There's one for *The Simpsons*—the cartoon—and one for O. J., too. There's a Newt Gingrich newsgroup. There's even one called "alt.tv.dinosaurs.barney.die.die.die". Others are more serious—"comp.sys.intel" has been bursting with the Pentium brouhaha. Readers of a newsgroup will frequently follow a "thread," or topic of discussion, that continues

alt.adoption
 alt.angst
 alt.books.reviews
 alt.clothing.sneakers
 alt.education.home-
 school.christian
 alt.fan.letterman
 alt.final.suicide
 alt.movies.silent
 alt.music.naked-
 barbies
 alt.pantyhose
 alt.politics.usa
 .newt-gingrich
 alt.privacy.clipper
 alt.scottish.clans
 alt.sports.baseball
 .sf-giants
 alt.support.tourette
 alt.tv.dinosaurs
 .barney.die.die.die
 alt.tv.simpsons
 ca.earthquakes
 caltech.announce
 .alumni
 caltech.class.math2a
 clari.biz.economy
 clari.biz.industry
 .construction
 clari.feature
 .miss_manners
 clari.news.flash
 clari.world.mideast
 .turkey
 comp.infosystems.www
 .misc
 comp.parallel
 comp.society.privacy
 comp.sys.intel
 de.alt.fan
 .haraldschmidt
 de.market.jobs
 fr.petites-annonces
 .vehicules
 k12.chat.elementary
 la.eats
 misc.kids.vacation
 rec.arts.startrek
 .current
 rec.bicycles.racing
 rec.crafts.textiles
 .quilting
 rec.gardens.orchids
 rec.woodworking
 sci.math.num-analysis
 sci.physics.cfd



One person can reach many people by posting on an electronic bulletin board, or newsgroup, like "alt.anxiety.faceless" at left. This bulletin board, a support group for featureless cartoon characters, is fictitious, but isn't much more specialized than some of the real newsgroups (below, left) available through Caltech's computers.

through several messages; angle brackets along the left margin of a message are used to mark material that has been included from somewhere else, generally from a previous posting on the same subject, and you can get many layers of angle brackets. Also, a lot of scientific conferences are planned and advertised through the Internet: agendas are set and speakers recruited electronically—it's only when you arrive that you get deluged with paper. And the Net isn't just for grown-ups—a 10-year-old girl in New Zealand can become key pals with a nine-year-old in Springfield, Virginia, through a group called "k12.chat.elementary".

Because they put so many people in touch at the same time, newsgroups can produce adult friendships as well in a way that paper letters and the telephone cannot. For example, two years ago, before my wife and I went to Moscow, I sent a posting to the Usenet group on Russia, "mlist.russia", asking, "How do I get from the Moscow airport into the center of the city?" The replies ranged from "You just get on the bus" to "Don't do this—you will be shot." But one reply was from a biochemist at MIT who has a sister in Moscow, and who asked me if I would deliver some medicine because the Russian mails are so unreliable. She sent the medicine to us in California, and we met her sister in Moscow, enjoyed her company, and made friends. This relationship was formed because of the Net.

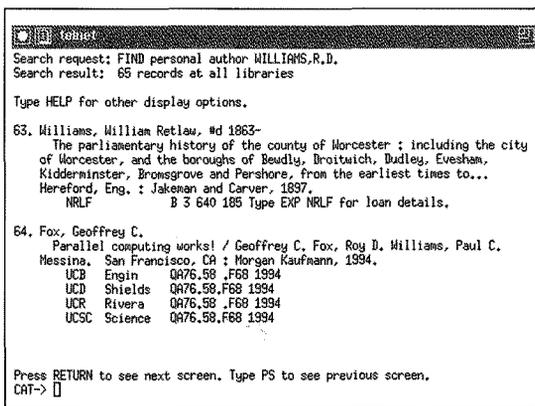
Similar to the Usenet, but real-time, is the idea of "chat." Chat means that whatever you type immediately appears on the screens of all the other people in the chat session. And you

don't have to type anything—you can be a "lurker," which is somebody who listens but does not contribute. Chat is more like a pub than a town meeting—sometimes it's a party, sometimes it's a fight. We have gotten used to having a relationship with someone by telephone without ever seeing them—travel agents, for example. Now we can meet people by Internet without hearing them either, which has had the beneficial side effect of making racism, ageism, sexism, and other types of discrimination more difficult!

The third use of the Internet is Telnet and a related program called FTP, which are methods for exchanging files between computers. For example, some of the world's fastest supercomputers are here at Caltech, but people all over the country use them via Telnet. A Telnet connection allows you to load your software and data files into someone else's computer and run them, even if it's thousands of miles away. People can even fix your software from their office, without an expensive housecall. Telnet can be used for searching databases, such as MELVYL, which is the on-line catalogue for all of the University of California's libraries.

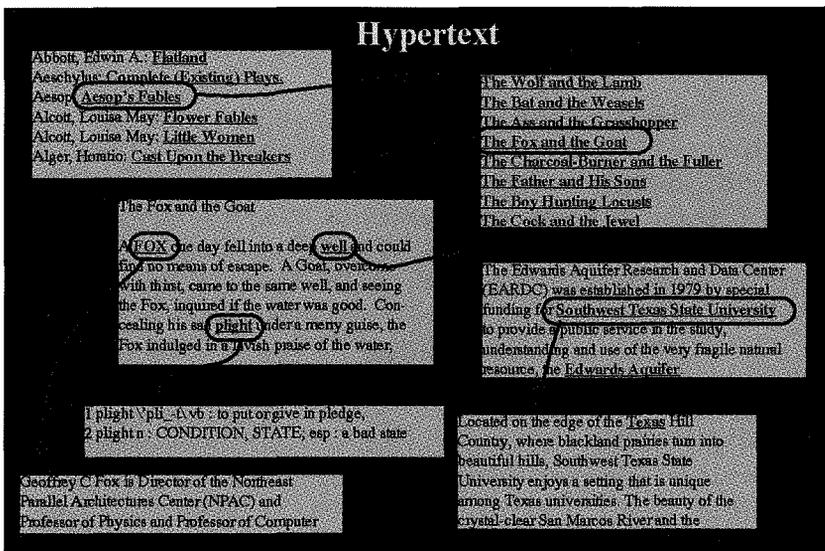
The last—and currently the most talked-about—feature of the Internet is the World Wide Web and the similar services such as Gopher. You can look up almost anything on the Web—today's Senate calendar, how to make an origami frog, weather forecasts for Siberia, a history of the vacuum cleaner. Even the *Encyclopedia Britannica* is on-line, but that you have to pay for. The Web is based on the idea of hypertext, in which multimedia documents—images,

Right: A Telnet search for authors named Williams in the MELVYL database yielded, among other things, a parliamentary history of Worcester, as well as this excellent book on parallel computing.



sound, and video clips, as well as text—are linked to other documents. Links can appear on your screen as underlined words or phrases, as push buttons, as icons, as images, or even as different parts of the same image. When you point your cursor and click on a link, the document you're viewing is replaced by the one at the other end of the link. A link can lead to a document on a machine anywhere in the world, which is how the World Wide Web got its name. The Web is growing even faster than the Net itself—doubling every three months, versus every 18 months for the Internet. Every human on the planet will be on the Web within five years at this rate, and every cockroach by 2002!

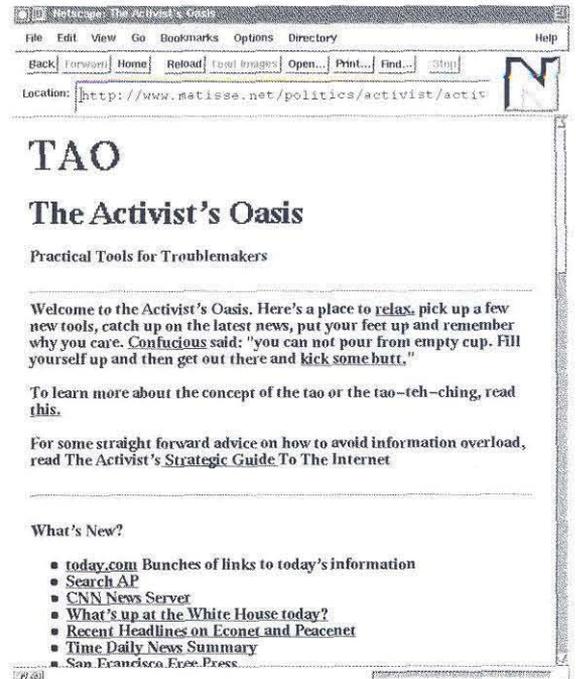
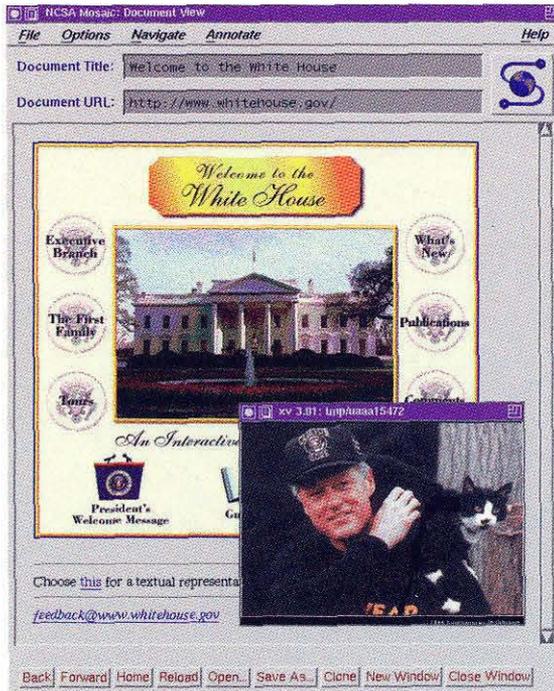
Hypertext is not really different, in principle, from a footnote or an endnote in text. With a footnote, you go to the bottom of the page for more information. With an endnote, you go to the end of the document. With a link, you jump to another document. An example of this is shown at left. Suppose we start with a list of books, and we want to look up Aesop's fables. We click on Aesop, and our screen is replaced with a list of Aesop's fables—fairly quickly, we hope. Now we click on the fable we want to see, for example, "The Fox and the Goat," and the text comes up. We can make significant jumps very easily. For example, "Fox" might point to Geoffrey Fox, the director of the Northeast Parallel Architecture Center at Syracuse University, which is a long way from the fable of the fox and the goat. Links are an associative way of moving through information, similar to the way human memory works by jumping from concept to associated concept. The old way of using the Internet involved remembering long, complicated addresses in order to find things. Now you can just follow your nose—even if you only vaguely remember how you found something before, you can probably find it again by following the same path.



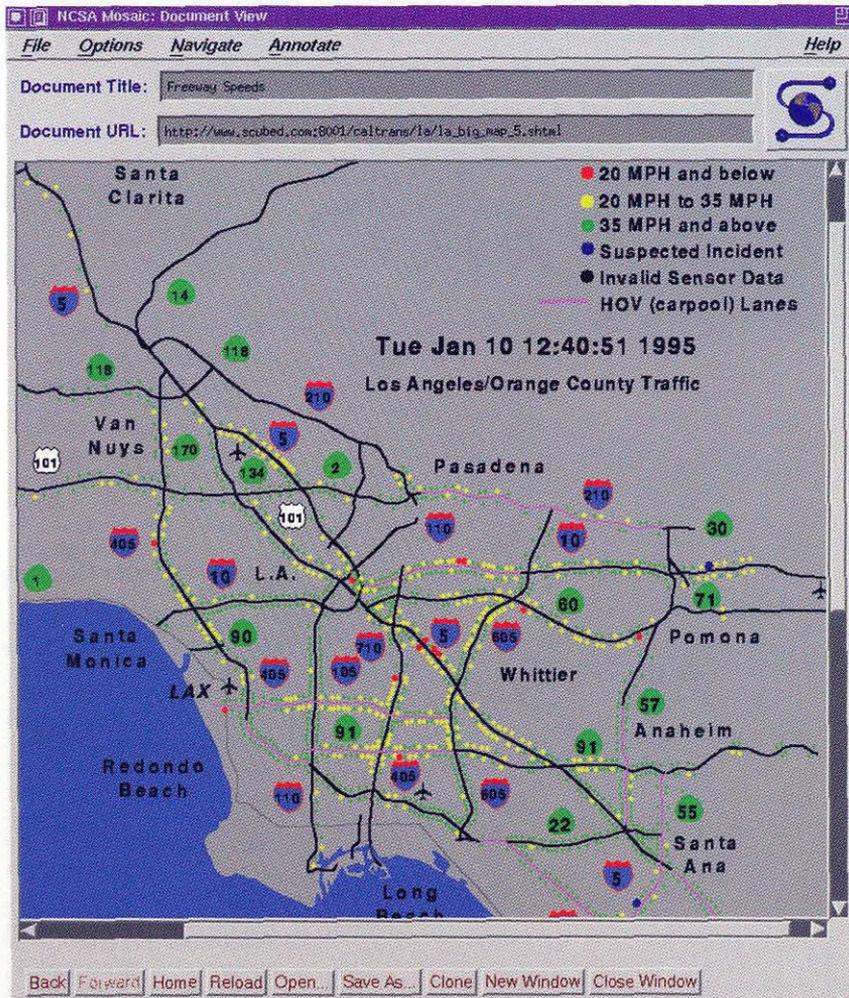
Hypertext (above) provides a way of navigating an ocean of information. The underlined words mark links (red arrows) to other documents; selecting a link and clicking on it brings the document on the arrow's other end up on your screen. Some links are short hops—looking up the word "plight," for example—while others are strides in seven-league boots.

To use the Web, you use a program called a browser, such as Mosaic or Netscape. When the browser is started, it brings up your "Home Page." The Home Page is your point of entry to the Web, and you can always jump back to it by pushing the Home button on your browser, so you can't get too badly lost. You can use an institutional Home Page, such as the Caltech one, or you can create your own personal Home Page. You can put all sorts of personalized stuff on the Web, including links to whatever you think is interesting. Somebody once said, "I didn't know what to do with it, so I put it on the Web." Having said that, let's look at a few of the pages we can find on the Web...

This is the White House Home Page. Among other things, you can get a picture of the President with Socks, the First Cat. The Clinton/Gore administration has been responsible for making a great deal of Federal information accessible by Internet.

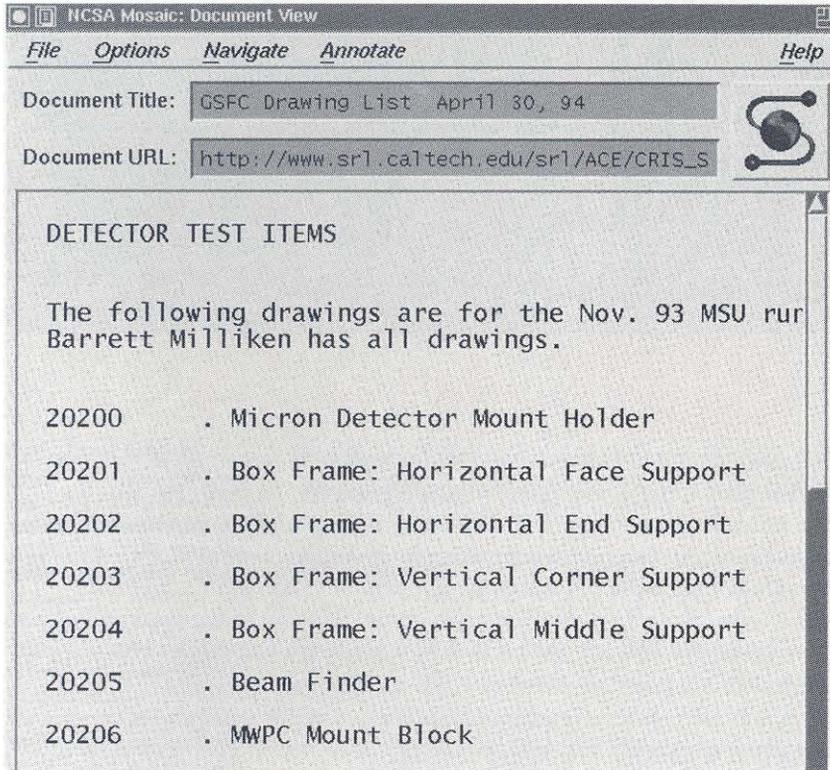


Here is a page devoted to political activism. The free flow of information makes it much easier and cheaper for groups to lobby Congress.

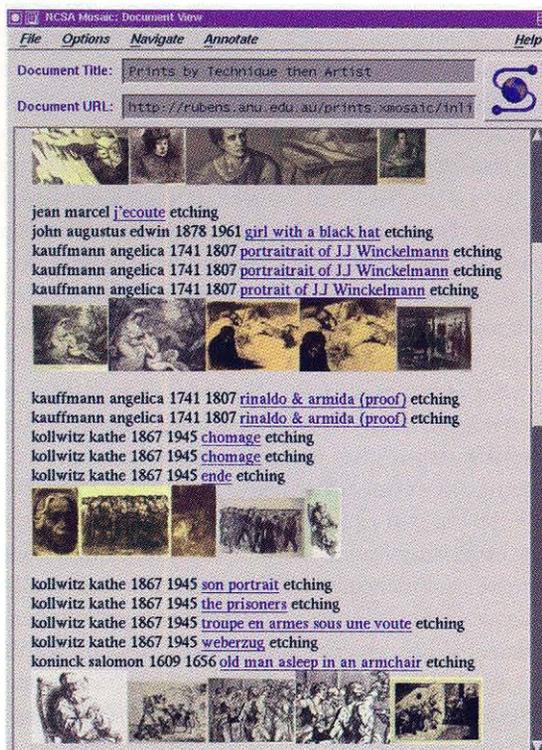


This page, created by Caltrans, shows freeway speeds in the Los Angeles area and is updated every five minutes. This particular image came from the noon hour on a day when we had three inches of rain. Each dot represents a set of sensors embedded in the roadway; the dot's color indicates the traffic flow there. This page was really handy when it came into being a few months ago, but now so many people use it that the connection gets easily overloaded.

The Cosmic Ray Isotope Spectrometer is a collaboration by Caltech, Johns Hopkins, NASA Goddard, and others. They use the Web to manage the project and keep track of how various tasks are progressing.



The Australian National University maintains a large selection of prints and drawings on-line, so you can visit the museum virtually. How does the existence of a virtual museum affect attendance at the real museum?

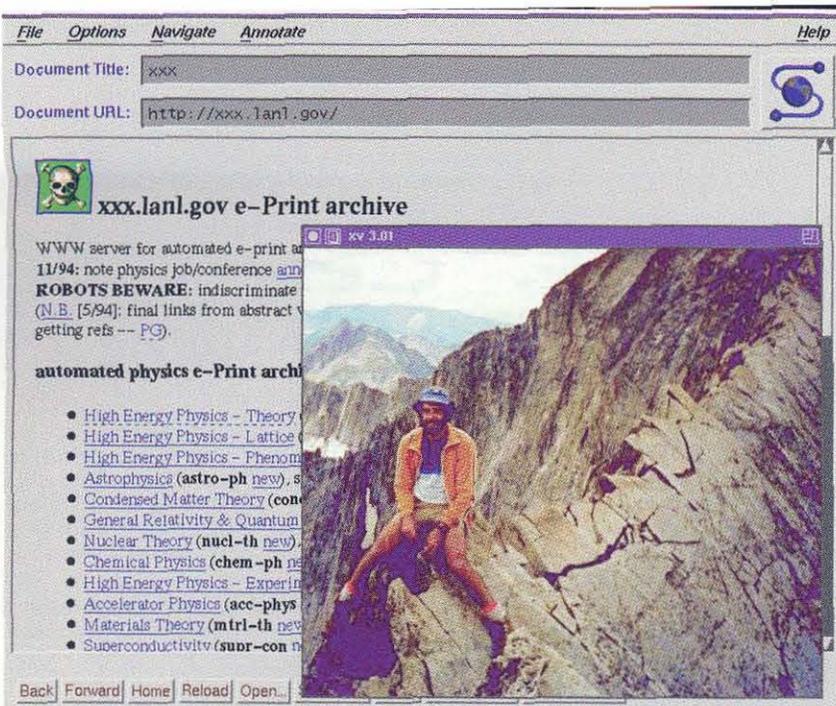


- | | |
|------------------------------------|---------------------------------------|
| American Demographics | Revue de Presse |
| Daily Texan | Sand River Journal |
| Boston Review | Technology Review |
| Buzz | The Chronicle of Higher Education |
| Electronic Antiquity | The Drama Review |
| Federal Communications Law Journal | The Nation |
| Internet World | The New Yorker |
| Mother Jones | The Reader |
| New Perspectives Quarterly | The Times Higher Education Supplement |
| New Republic | The Times Literary Supplement |
| Policy Review | USA Today |
| Postmodern Culture | Verbiage Magazine |
| Quanta | Whole Earth Review |
| Review of Early English Drama | Wired |

A small sample from the Net's newsstand.

In order to get connected to the Internet, you call a service provider—America Online, CERF-net, CompuServe, Netcom, Psi, or a host of others—who will charge you a fee and give you a phone number for your modem to call and software to let your computer talk to the Internet computers. (The Net is in some sense free, but you've got to pay both for the phone use and for your service provider to connect you to it.) Levels of service vary, but so far *Consumer Reports* hasn't done Internet Service Providers—I'm sure it's just a matter of time! Your service provider may offer on-line help, which the Net doesn't. Service providers may also offer extras like access to Sabre, the airline-ticketing database; legal databases such as Lexus; the world's magazines and newspapers through Nexus; specialized stock-market databases; and so on. You can get censorship from your service provider, if you want it—you can have a separate account with restricted access for the kids, which is like having the phone company prevent 976 calls being made from your phone.

Many magazines and newspapers are available on the Internet; there are even e-zines, as they are called, that exist only on the Net and aren't published on paper. A lot of publishers are transferring their paper offerings en bloc to the Net, even though they're not quite sure what they're doing or why. An on-line clone of a magazine is easy to read—there's no advertising, no perfume samples, no bits of paper dropping out on the floor when you turn the page (what does that tell us about how long this kind of service is likely to last?). An example of a more thoughtful



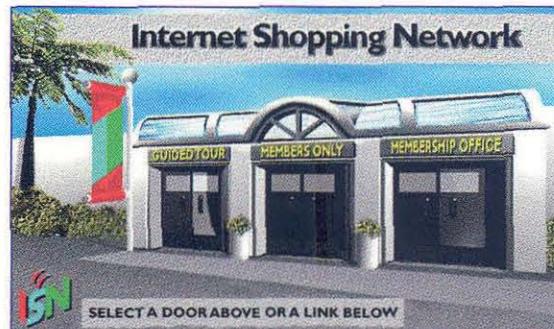
Above: Paul Ginsparg maintains a database of electronic preprints of high-energy physics papers.

Above, right: Click on the appropriate door of the Home Shopping Network's Home Page to get a demo, go shopping, or join the club.

approach is *Mother Jones Interactive*, which is a set of discussion groups and a database promoting grassroots activism. The magazine itself, *Mother Jones*, was one of the first to go on-line—you can download the cover picture if you're willing to sit and wait a few minutes, but besides that, it's all text and looks just like any other Web page. It doesn't matter whether you think the visual layout of a publication is important or not—it's simply not there in the Internet version, especially over a slow telephone link, with images switched off. It's difficult to retain brand identity under such circumstances. But new software, for example a product called Adobe Acrobat, is now available to recreate faithfully the original look of the printed page—even to the ads, which should make the publishers happy and the software popular.

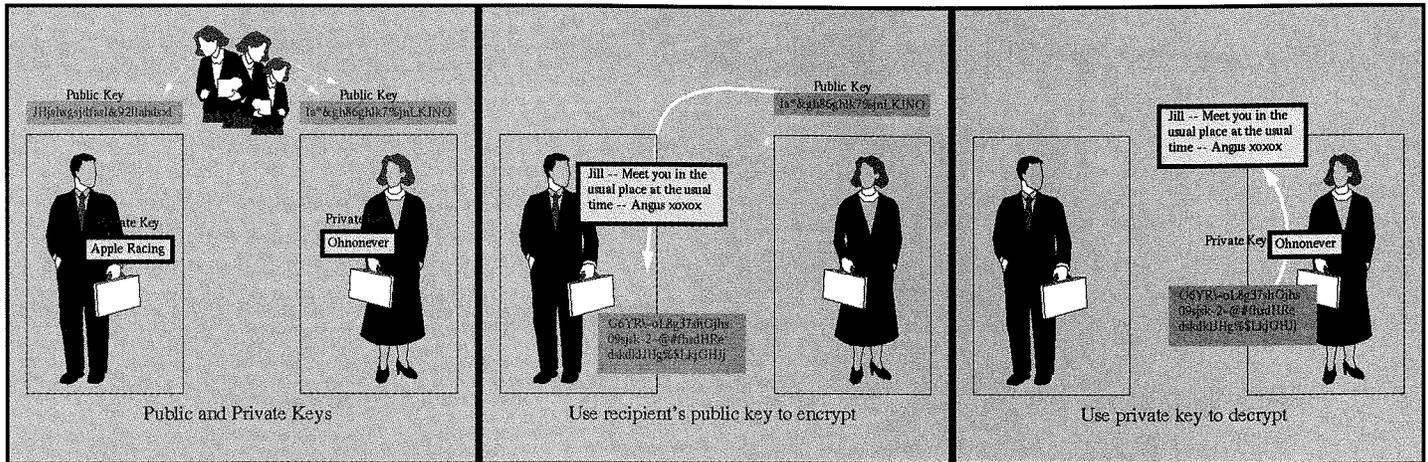
There are books on the Internet: the great classics certainly, but also modern, copyrighted works. The publisher hopes that people who see the book on the Net will go out and buy the real thing, but cheapskates can simply read it on-line or print it on a laser printer instead of buying it. Will we wind up reading more things directly from the machine, without printing them? I think the era of the bedside computer is not so far away.

Scholarly articles are not just appearing on-line, but their on-line "publication" is squeezing out the importance of paper journals. Paul Ginsparg, at the Los Alamos National Laboratory in New Mexico, runs a system called "xxx.lanl.gov", which contains a database of preprints of high-energy physics papers. People



download about 30,000 preprints per day from his system, and roughly 10 new preprints are added per day. This is really catching on because paper journals take a year or more to publish something, but when you send a paper to Ginsparg's system, it's available to the global scientific community immediately. Many high-energy physicists don't even look at the paper journals anymore, only the Internet sources. But on-line papers are not peer reviewed, and peer review is the quality assurance of the scientific enterprise; furthermore, peer-reviewed papers are what get you tenure! The future of on-line journals is a big question—how do you combine the rigorous prepublication scrutiny of peer review with the instant dissemination of your work? After-the-fact reviewing might be possible if it were true that the number of times a paper is downloaded is a useful measure of the quality of the paper.

You can shop on the Web, through many companies that have been set up in the last year or so for this purpose. Even the Home Shopping Network is available! You can buy all kinds of computer products, of course—you can even get free software demos. But you can also buy cookies, or even lingerie. To buy these products, a credit card is generally used—not because it's the most efficient payment method available, but because it's the only one. Credit cards do not provide sufficient security, they don't facilitate microbilling, nor do they provide anonymity. There's a problem with security because, unfortunately, the Internet is quite a leaky channel: the skills needed to tap into somebody else's Internet



Left: Jill's private key—"Ohnonever"—is converted by her computer into a string of gibberish that can be posted publicly for all to see.

Center: If Angus wants to send Jill a private message, his computer combines his text with her public key to produce an encrypted message that can be transmitted openly.

Right: Jill's computer then combines her private key with Angus's public message to retrieve the private message.

transactions and steal credit-card numbers aren't very rare. And you might want to buy a lot of very cheap things—if you look up something in the *Encyclopedia Britannica* on-line, for example, they might charge you a few cents, and credit cards don't work well with such small transactions. Anonymity will be increasingly important; the problem with electronic transactions is that people are going to figure out who you are, put you into a database, and sell you to marketers. Along with measures to make U.S. currency more difficult to counterfeit, the government is thinking of printing bar codes on our money. I, for one, don't like the idea of somebody scanning my bills and finding out everywhere I go and everything I buy. Several companies are trying to market the idea of electronic digital cash, known as cryptcash, that's secure, comes in small denominations, and is anonymous. Once there's trusted electronic cash, people will be able to start businesses on the Internet very easily, selling custom products to a global market with very little start-up cost.

Closely related to the question of security of information is the issue of encryption, which is a topic of heated discussion these days. The essential question is whether the government has the right, when sanctioned by a judge, to "wire-tap" a computer in the same way that the law allows telephone taps. We have to decide this soon, because technology is rapidly taking over. Software to produce military-grade, unbreakable encryption is already available on the Internet for free. The system works like this: you make up a "private key"—a phrase that you never tell

anyone, that's between you and your computer. Your computer then converts this private phrase into a public phrase, or "public key," which is a sequence of apparently random characters. You can't go backwards—you can't turn a public key back into a private key, even using all the computing resources in the world for the age of the universe. The public keys are available to everybody. Now, let's suppose that Angus wishes to send a message to Jill. Angus looks up Jill's public key and his computer combines that with his message to produce the encrypted text. The encrypted text is sent to Jill, who uses her private key to decrypt it and get the message back. The private keys never move across the Net, so nobody can intercept them. The only way for someone to get your private key is to look over your shoulder as you type, or to steal it if you're foolish enough to write it down. The government is trying to outlaw this kind of software—it's treated as munitions under some circumstances—precisely because they can't break it. But as the more anarchic citizens of the Net like to say, "If privacy is outlawed, then only outlaws will have privacy."

There are other legal issues as well. When Gutenberg invented the printing press in the 16th century, one of its first uses was to produce large quantities of pornographic woodcuts. The same vulgar objectives are fulfilled by any new medium, including, of course, the Internet. In October 1994, Carnegie Mellon University decided to censor the Usenet feed, some of which contains obscene material. There was an uproar in the campus community, and talk of free speech



"On the Internet, nobody knows you're a dog."

and so on. The question here is whether the Usenet feed is like a telephone company or a television station. Ma Bell is a common carrier and isn't expected to censor its traffic—you can say anything you like on the phone. Whereas a television station partakes of the limited resource of radio bandwidth, and therefore is held responsible for its content.

It's also possible to send e-mail and contribute Usenet postings completely anonymously, which can lead to very frank discussions—people can say things on the Internet that they don't say anywhere else. But it can also lead to antisocial behavior, and I think the Usenet is starting to suffer a bit from this anonymity. People don't take any responsibility for what they say anymore, and the few are spoiling it for the many. In so-called flame wars, people try to be as vicious to each other as they can with just words. In an extreme case, last Thanksgiving two journalists on Long Island not only had their e-mail "bombed"—that is, their e-mail mailboxes were filled up with rubbish—but the attackers also got into the phone company's computers and redirected the victims' incoming calls to an answering machine containing an obscene message. The Internet operates on the honor system, and if you flout that you can do people a lot of harm. But preventing this sort of thing is difficult when there's encryption and anonymity. If you can't see what's being moved across the Net, and you don't know who sent it, how can you possibly decide whether it should be there or not, and, if not, how to stop it?

In conclusion, in my opinion cybersurfing is

much more fun than watching television. You can go where you want to go, look at the information that you want to look at. You can be your own publisher: it is becoming easier and cheaper every month to set up a Web server.

The Internet will make disseminating individual artistic expression easy, and we will have access to information that can empower us. It is, perhaps, the gateway to a great new virtual culture.

We can expect journalists—and ordinary citizens!—to report on their findings from raw data, rather than predigested information. The bright light of media attention will become more penetrating, causing honesty in reporting the facts, but also more scope for fallacious statistical arguments. Retailers will adjust to the new medium, enticing us into their cyber-stores with giveaways of information, "frequent-visitor programs," and advertisements with ever-fresh "eye-candy" pictures.

The less-exciting alternative is that we'll be forced to spend time in some awful virtual space where we won't meet anybody, and every now and then our path will be blocked by an advertisement for laundry detergent or a feminine-hygiene product, and we'll have to wait for it to finish before we can continue.

Individuals who have the technical and creative abilities to do so should try to put something of themselves on the Internet, before it gets taken over by relentless corporate agendas. We must not simply cocoon behind security gates, with our computers and lots of software, having nothing to do with the nasty cold real world outside. We must use the Internet to build a virtual community and explore what it can do. But we must make sure that it enriches the physical community and real meetings between people, rather than replacing them. □

Roy Williams is a senior staff scientist at Caltech's Center for Advanced Computing Research. He earned his bachelor's in mathematics at Trinity College, Cambridge University, in 1979. His PhD in physics followed at Caltech in 1983, and he has been a research fellow here, at Oxford University, and at the Rutherford-Appleton Laboratory in Britain. He has been working with parallel computers at Caltech since 1986, researching fluid-flow algorithms, differential equations, parallel software, high-speed networking, and the presentation of information using the Internet. He may be found by e-mail at roy@caltech.edu or on the Web at

<http://www.ccsf.caltech.edu/~roy/>.