The Internet and the World Wide Web...19 questions

http://www.cnet.com/Content/Features/Techno/Networks/index.html

1. WHAT THE INTERNET IS
2. HOW DOES THE NET WORK
3. WHAT IS THE WEB
4. HOW DOES THE WEB WORK
5. WHO STARTED THE NET
6. WHO CONTROLS THE INTERNET
7. HOW ARE ONLINE SERVICES DIFFERENT FROM THE WEB
8. WHY IS THE WEB SO SLOW
9. IS THE NET SAFE?
10. WHAT IS A SEARCH ENGINE
11. WHAT ARE JAVA AND ACTIVE X
12. WHAT IS AN INTRANET
13. HOW DOES EMAIL WORK
14. WHAT ARE NEWSGROUPS
15. WHAT'S ALL THE FUSS ABOUT PUSH?
16. HOW DO I MAKE A WEB PAGE
17. HOW DO I TALK TO PEOPLE ON THE NET
18. HOW ELSE CAN I USE THE NET
19. WILL THE WEB BECOME MORE LIKE TV?
WHAT THE INTERNET IS

The Internet is a loose association of thousands of networks and millions of computers across the world that all work together to share information. Like many complex systems, the Internet is easiest to explain through the use of metaphors, and the Net has inspired its fair share. The one that has stuck is the "information superhighway," and while it has become a cliché, the transportation analogy really does hold up pretty well. Think of the Internet as a mondo version of a mass transit system like Boston's T, with a few main subway lines that intersect at certain points. Connecting to the subway lines are commuter rails, bus lines, and ferry boats that spread out and crisscross the metropolitan area.

On the Net, the main lines carry the bulk of the traffic and are collectively known as the Internet backbone. The backbone is formed by the biggest networks in the system, owned by major Internet service providers (ISPs) such as GTE, MCI, Sprint, UUNet, and America Online's ANS.

By connecting to each other, these networks create a superfast pipeline that crisscrosses the United States and extends to Europe, Japan, mainland Asia, and the rest of the world. But that doesn't mean that the network is equally well developed at every point along the route. The U.S. backbone has so many intersecting points that if one part fails or slows down, data can be quickly rerouted over another part, a feature called redundancy. Overseas, the network may have less redundancy and so be more vulnerable to slowdowns or breakdowns.

In the United States, there are five points--located in San Francisco, San Jose (California), Chicago, New York (actually, Pennsauken, New Jersey), and Washington, D.C.--where the main lines intersect, kind of like how the major U.S. airlines have hub cities. Confusingly enough, three of these are called network access points (NAPs), while the other two are called metropolitan area exchanges (MAEs), but they basically do the same thing: use high-speed networking equipment to connect the backbone to other networks (see Figure 1). These networks are owned by smaller regional and local ISPs, which in turn lease access to companies and individuals in the areas they serve.

Government agencies and universities are also actively involved in running the parts of the Internet that link supercomputer centers devoted to the research and education communities.
While this used to be the main purpose of the Net, the explosion of private and corporate use has caused a huge traffic jam on the backbone. Academics now complain that they can't get their work done because the network is too packed with everybody else.

With help from these communities, as well as financial support from the private sector, Congress has been actively planning the Next Generation Internet. It aims to deliver on President Clinton's pledge to build and promote a new and faster network that will form a second backbone over the next five years. This Internet II, as it's sometimes called, will not replace the existing Net, but will provide alternate routes for academics and government agencies to share information without getting caught in commercial traffic.

**HOW DOES THE NET WORK**

The secret of the Net is a network protocol called TCP/IP--that is, a kind of coding system that lets computers electronically describe data, like the contents of this story, to each other over the network.

The term actually refers to two separate parts: the transmission control protocol (TCP) and the Internet protocol (IP). Together they form the Esperanto of the Internet. Every computer that hooks to the Internet understands these two protocols and uses them to send and receive data from the next computer along the network.

TCP/IP creates what is called a packet-switched network, a kind of network intended to minimize the chance of losing any data that is sent over the wires.
First, TCP breaks down every piece of data--such as an email message or instructions from a Java applet--into small chunks called packets, each of which is wrapped in an electronic envelope with Web addresses for both the sender and the recipient. The IP protocol then figures out how the data is supposed to get from point A to point B by passing through a series of routers--sort of like regular mail passes through several post offices on its way to a remote location.

Each router examines the destination addresses of the packets it receives and then passes the packets on to another router as they make their way to their final destination. If your email was broken into ten packets, then each of those may have traveled a completely separate route. But you'll never know it, because as the packets arrive, TCP takes over again, identifying each packet and checking to see if it's intact. Once it has received all the packets, TCP reassembles them into the original. (See Figure 2.)

TCP/IP is the most important of a long list of Internet protocols. It is sometimes used as a global term to describe additional protocols, including simple mail transfer protocol (SMTP), file transfer protocol (FTP), and Telnet protocol.
Although the terms Web and Internet are often used synonymously, they're actually two different things.

The Internet is the global association of computers that carries data and makes the exchange of information possible. The World Wide Web is a subset of the Net—a collection of interlinked documents that work together using a specific Internet protocol called HTTP (see "How does the Web work?"). In other words, the Net exists independently of the Web, but the Web can't exist without the Net.

The Web began in March 1989, when Tim Berners-Lee of the European Particle Physics Laboratory (a collective of European researchers better known by its original name CERN, or Conseil Européen pour la Recherche Nucléaire) proposed the project as a means to better communicate research ideas among members of the far-flung organization.

The Web uses a metaphor of individual pages, usually combined to make up sites. Web pages are written in HTML, or Hypertext Markup Language, which tells the Web browser how to display the page and its elements. The defining feature of the Web is its ability to connect pages to one another—as well as to audio, video, and image files—with hyperlinks. Just click a link, and suddenly you're at a Web site on the other side of the world. (Before the Web, you had to type in exact Net addresses or wade through a series of menus to get where you wanted to go.)

Despite its cool hyperlinking ability, the early Web labored for a while in obscurity, a little-known alternative to the less technically advanced Gopher protocol. But in February 1993, Marc Andreessen, then developing for the National Center for Supercomputing Applications, introduced the first graphical Web browser, called Mosaic. (Andreessen went on to cofound Netscape Communications in April 1994.) And the rest, as they say, is history.

**HOW DOES THE WEB WORK**

The Web is based on a set of rules for exchanging text, images, sound, video, and other multimedia files, which is collectively known as HTTP, or hypertext transfer protocol. Web pages can be exchanged over the Net because browsers (which read the pages) and Web servers (which store the pages) both understand HTTP.

But everything would still be chaos if the Web didn't have an addressing scheme that every computer on the network understands. An IP address is a 4- to 12-digit number that identifies a
specific computer connected to the Internet. The digits are organized in four groups of numbers (which can range from 0 to 255) separated by periods. Depending on how your ISP assigns IP addresses, you may have one address all the time or a different address each time you connect. Web servers have the same kind of addresses: if you type http://204.162.80.183/ in your browser, you'll get the same result as if you had typed http://www.cnet.com/.

Internet domain names are the next level of Internet addressing, just as the street name is followed by the city and state. Domain names create a single identity for a series of computers used by a company or an institution. So while there may be 38 servers at a given company, each with its own IP address, they all share a common domain name, such as CNET.COM.

The domain name identifies all the computers in a group. But if you want to get to a specific page stored on any of those computers, you'll need an even more precise address. That's why every Web page on the Internet, and even the objects you see displayed on Web pages, has its own unique address, known as a Uniform Resource Locator (URL), which tells your browser exactly where to go on the server to find a page.

WHO STARTED THE NET

No one person or organization can claim the sole credit for the Internet. But the first germ of the Internet was a series of memos written in 1962 by MIT's J. C. R. Licklider about what he called the "Galactic Network" concept. He envisioned a global network through which everyone could share and access data and programs. Only a few months later, Licklider became the head of the computer research program at the United States Department of Defense's Advanced Research Projects Agency (ARPA), the institution that largely spearheaded and funded the Internet's development.

In 1961, a series of independent research teams began developing packet switching and the beginnings of what would eventually become TCP/IP, the basic protocol that defines how information is exchanged over the Net. (See "How does the Net work?" for an explanation of these technologies.)

In 1967, ARPA's Lawrence Roberts published his "Plan for the ARPANet" computer network, which built on these new technologies to propose an architectural design for a worldwide network.
By the end of 1968, the company that would become BBN Planet (a major backbone ISP recently bought by GTE) was well into the development of the first hardware that could route data over the ARPANet. In late 1969, the first tests were made at UCLA and then at Stanford.

Over the next several years, this test-tube Internet grew steadily but unremarkably as government agencies, universities, and corporations continued to develop and hammer out protocols and architectures. Email and the Internet made their first public appearances in 1972 at the Internet Computer Communication Conference. In 1973 and 1974, the protocol known as TCP/IP emerged in essentially its current form, although the same group of collaborators would continue to refine it through the early 1980s.

Once the protocols were in place, the various developers formulated much of the software and services that make up the Internet. The basic services for connecting to files remotely (via Telnet), transferring files over the Net (via FTP), and sending and receiving electronic mail appeared in the mid- and late 1970s. The Usenet news system first appeared in 1979 as an offshoot of the rise of Unix. The World Wide Web began in 1989.

In 1990, the U.S. government officially decommissioned ARPANet, and the National Science Foundation (NSF) took over the role of managing the Internet backbone, which was then called the NSFNet. In 1995, the NSF in turn withdrew, turning the backbone over to a consortium of commercial providers.

**WHO CONTROLS THE INTERNET**

No one person, company, institution, or government organization owns the Internet. No one source foots the bill for it. No one entity governs it, or even has a controlling interest. The Internet is truly a collaborative, collective enterprise.

Many institutions and companies donate their computer resources in the form of servers and computer technicians to hold up some part of the Internet—such as, critical links between different regions. Governments around the world are also starting to exert their influence through legislation. And every computer on the Net has to understand a basic set of technologies, which several organizations are involved in maintaining, updating, and disseminating.

There are a handful of organizations that are truly influential and that taken together form a sort of checks-and-balances system:
The World Wide Web Consortium (W3C) sets the standards for HTML and other specifics of the Web.

The Internet Engineering Task Force (IETF) focuses on the evolution of the Internet with a specific eye toward keeping the Internet running smoothly as a whole.

The Internet Engineering Steering Group (IESG) is a related organization responsible for managing IETF activities and the Internet standards process.

The Internet Architecture Board (IAB) is responsible for defining the overall architecture of the Internet (the backbone and all the networks attached to it), providing guidance and broad direction to the IETF.

The Internet Society (ISOC) is a supervisory organization made up of individuals, corporations, nonprofit organizations, and government agencies from the Internet community. The group comments on Internet policies and practices and oversees a number of other boards and task forces--including the IAB and IESG--dealing with Internet policy issues.

The Internet Assigned Numbers Authority (IANA) and the Internet Network Information Center (InterNIC) lead the organizations responsible for assigning IP addresses and domain names, respectively.

Regional and long-distance phone companies, backbone ISPs, cable and satellite companies, and the U.S. government all contribute in significant ways to the telecommunications infrastructure that supports the Internet. Some of these companies, such as UUNet, BBN, Sprint, and MCI, have found ways to make lots of money by leasing access to the Internet to other companies. As the commercial potential of the Net matures even more, these companies might begin to throw their weight around. Several major ISPs have already banded together to hash out industry-wide technical issues.

But no one can wrest away total control of the Net, which is why it's not just a communications medium, but a metaphor for the new global economy.

HOW ARE ONLINE SERVICES DIFFERENT FROM THE WEB
To the naked eye, there's not that much difference between online services and the Web anymore. But the differences used to be much more marked.

Customers used to pay the four big online services--America Online (AOL), CompuServe, Prodigy, and the Microsoft Network (MSN)--to get access to worldwide networks with libraries of digital information, to send email around the world, and to join online communities where people with similar interests could communicate for business or pleasure.

The online services offered all this by building and maintaining long stretches of interconnected wires and servers that no one but their customers could use. Each system had a different interface and customers nearly always paid by the hour or the minute. The companies competed over which one was easiest to use and which one had the best content.

But then the public noticed the Web, which offered a vast network with an interface--the browser--that was the same for everybody. Where the online services relied on vast private networks, the Web wasn't owned by anyone in particular. To get on America Online's network, you had to pay them. To get on the Web, you had to pay any regional or national ISP for a local dial-up number. Suddenly, the online services were competing with hundreds of ISPs, instead of only each other.

So the online services adapted: they started letting users roam the Web as well as their own networks. And they got cheaper, substituting monthly flat fees for per-hour surcharges. To most users, online services are now simply big ISPs. Prodigy has even taken to calling itself "the SuperISP."

Of the big four, only one--AOL--is still determined to maintain its own private network. MSN, Prodigy, and CompuServe have all instead constructed huge Web sites to serve as new homes to all of their exclusive content. They still want you to pay, but for access to the Web and to content you can't see anywhere else instead of as an entrance fee for their private networks.

CompuServe and Prodigy are using the wires they still control to sell "premium" Internet services--that is, guaranteed Net service to businesses that demand total reliability. MSN, on the other hand, doesn't care which ISP you use: it is more interested that you sign up to get at its content.

There still seems to be a market for private networks among users who have never used the Net before; this is where most of AOL's new customers come from. Disney is also launching a new online service aimed at children and families. But the distinction between subscribing to a private network like Disney's Daily Blast and subscribing to a private Web site like MSN is becoming more and more blurred.

WHY IS THE WEB SO SLOW
One minute you're flying along the Web, happily swinging from link to link, and the next you've slammed into a tree, "waiting for reply" or ponderously "transmitting data." Why?

Part of it is the Internet's fault: its ability to handle an enormous amount of data every day trades flexibility for speed. Everyone who uses the Net shares bandwidth--the data-carrying capacity of a network. Every time you send an email or download a file, you're contributing to the load. Web pages are particularly bad bandwidth hogs because they are loaded down with graphics and multimedia.

Any number of things can go wrong on the Web, from a squirrel chewing on your phone line to a breakdown of a Web site's server to a traffic jam on the Internet's backbone that affects everybody. Everything could be peachy again minutes later, and usually is.

Still, knowing what's wrong can relieve your frustration a bit. The following are explanations of the most common server error codes and browser error messages you'll see:

common server error codes 404: a host server responded to your browser, but it cannot find the requested URL, which usually means the document was moved or even deleted. Or it could mean that you clicked a bad link; you may be able to fix this by simply starting over and typing the URL into the address field.

403: the requested resource is forbidden, which generally means you don't have the privileges needed to access that page. Recheck the URL and try again, or look around elsewhere on the site for another way to get to the page you need.

503: the server is probably too busy to handle an additional request for service, or it could be some other problem along the Internet. Try back in a few minutes.

common browser error messages Unable to locate server: your browser wasn't able to verify that the domain name exists. The server name in the URL you sent is probably incorrect.

Host unavailable: the actual wording of this message depends on the browser you're using. Anything similar to this means that the domain name you're attempting to access does exist, but it's currently not responding to your request. This usually occurs when a server is so busy that it is effectively offline, when it is down for maintenance, or when it is temporarily unavailable because of some other problem. Wait and try again later.
IS THE NET SAFE?

The Net is about as safe as a dark alley. Maybe there are bad guys lurking in the shadows; maybe there aren't. But you aren't defenseless.

There are two types of trouble on the Net: threats to security and threats to privacy. Potential security bogies include viruses contained in file downloads, rogue ActiveX controls that can crash your computer, malicious email attachments, holes in Java that let hackers read data on your hard disk, and a host of other weaknesses in your Internet software.

Despite all the press reports, the odds are against your becoming the random victim of a hacker. You're much more likely to run into a virus, but installing and using antivirus software should take care of this. Avoid opening email attachments from people you don't know, and use good judgment about paying with credit cards on the Web: if you don't know that the company is reputable or don't know where to find it in the real world, then keep your card in your wallet.

For a complete rundown of security dangers on the Net--and an interactive security check--check out CNET's feature "Net crime: don't be a victim."

Threats to your privacy are more subtle, but here again you can define some limits. For instance, you can make your email safe from prying eyes by using an encryption program such as Pretty Good Privacy (PGP). Encryption software translates your message into a secret code so that it can be read only by the person who has the correct decryption key--that is, the person you're sending it to.

But more people are probably going to want to know your buying preferences than the contents of your email. Information circulated on the Web can help a marketer construct a consumer profile of you. Once they have your email address, they can then besiege you with sales pitches.

Many Web sites are actually programmed to harvest information about any visitor who comes through. They send small files called cookies to your hard drive. Each cookie can be read only by the Web server that created it, for up to days or months later when you revisit the site. Among other things, cookies permit Web sites to track your name, your email address, your ISP's name, the last site you visited, your operating system, and your browser's specific make and version number. They can also help you out by storing passwords so that you can get into subscription-only sites without having to type the password every time. But the savvy consumer will want to
control what information is being collected. You can have your browser refuse to accept any cookies or use a program such as Kookaburra Software's CookiePal to track and manage your cookies.

Online telemarketers are annoying, but information collected online can also be used in more sinister ways, such as sending you obscene emails. For example, every time you post to a Usenet newsgroup, your email address becomes available to everyone who reads the group--and some newsgroup postings stick around for years. One way around this problem is to use an anonymous remailer service such as Replay, which forwards your email without your address.

Whenever you enter your name, address, and phone number in a form on the Web, that information could be going to people you don't know, so think twice before revealing personal info, especially your home address or your phone number. The Federal Trade Commission and Congress have been at work formulating privacy policies, but none of this is regulated yet.

WHAT IS A SEARCH ENGINE

When you absolutely, positively have to know something about Cretaceous Mongolia (and, after all, who doesn't?), there's nothing quite like the instant gratification offered by the Web. But there's searching, and then there's finding. Understanding how various types of search sites and programs work can make for more efficient info quests.

There are three primary types of search sites on the Web: search engines, Web directories, and parallel and metasearch sites.

Search engines such as Excite and HotBot use automated software called Web crawlers or spiders. These programs move from Web site to Web site, logging each site title, URL, and at least some of its text content. The object is to hit millions of Web sites and to stay as current with them as possible. The result is a long list of Web sites placed in a database, which users search by typing in a keyword or phrase. (For more about how search engines work, see CNET's feature "can you trust your search engine?")

Web directories such as Yahoo and Magellan offer an editorially selected, topically organized list of Web sites. To accomplish that goal, these sites employ editors to find new Web sites and work with programmers to categorize them and build their links into the site's index.

Since both approaches make sense, all the major search engine sites now have built-in topical search indexes, and most Web directories have added a keyword search.
Parallel and metasearch sites ride piggyback on the Web crawler sites. Parallel search programs, such as Vironix Software's WebFerret, launch simultaneous searches on all the popular search engine sites, returning all the results in a single window.

Metasearch sites go a step further. One of the problems with searching on the Web is that the searching vocabulary varies from search site to search site. For example, when you search for Cretaceous Mongolia on Yahoo, the search term should look just like that. But the same search performed at Infoseek would be more effective if you entered +Cretaceous +Mongolia; at Galaxy, it should be Cretaceous AND Mongolia. Metasearch sites, such as Metasearch.com, take care of this for you. They let you enter a term in a single field and then automatically account for all the particulars for half a dozen or more popular search sites.
WHAT ARE JAVA AND ACTIVE X

Both Java and ActiveX are technologies that let programmers create animated and interactive Web pages--the kinds that move, flash, and play games. HTML is the language that describes all the basic elements of a page (such as text and graphics), but its current incarnation can't do much to make a page interactive; Java and ActiveX fill that void.

Sun's Java is a programming language similar to the popular C language used to make applications like word processors or spreadsheets. Java applications--known on the Web as applets--have the unique ability to run on any operating system, from Windows to Mac to Unix.

While lots of applets are available as shareware and can be plugged ready-made into Web pages, programming Java from scratch is not a Saturday afternoon activity.

(Don't confuse JavaScript with Java. JavaScript is a scripting language, a special kind of programming language used to tie other components together or to accept user input. While the names are confusingly similar and both technologies are designed to make Web pages more interactive, Java and JavaScript are two very different things.

ActiveX is a little harder to define than Java. That's because Microsoft has chosen to make the term active a major part of its Internet marketing campaign, but there are so many active things from Microsoft that the word begins to lose meaning.

However, ActiveX controls are roughly equivalent to Java applets, in that they run in Web browsers and are designed to enhance Web pages. But where Java is a full-fledged programming language, ActiveX controls are created in Microsoft's Visual Basic development environment. That means that ActiveX is closely linked to other Microsoft technologies; for example, it lets you look at Word and Excel documents from a browser window. But where Java is designed to be cross-platform, ActiveX was created primarily for Windows and works only with Internet Explorer, not Netscape's Navigator.

Java and ActiveX are not mutually exclusive--they can work together. But Sun and Microsoft are both trying hard to make Web developers loyal to one over the other.
WHAT IS AN INTRANET

First, we had the Internet. But then, as if there weren't enough other buzzwords to learn, everyone began talking about intranets.

The idea was to come up with a new word to describe corporate offshoots of the Net--digital cul-de-sacs that connect to the Net but are closed off to the general Net-using public.

Intranets work like the Web, with browsers, Web servers, and Web sites, but they're used internally by companies or organizations. Companies use them because they let employees share corporate data, but they're cheaper and easier to manage than most private networks--no one needs any software more complicated or more expensive than a Web browser, for instance. They also have the added benefit of giving employees access to the Web. Intranets are closed off from the rest of the Net by firewall software, which lets employees surf the Web but keeps all the data on internal Web servers private.

The very latest buzzword is extranet, a term coined by Netscape cofounder Marc Andreessen. Extranets are several intranets linked together so that businesses can share information with their customers and suppliers.

This is still an evolving concept, but suppose, for example, that two companies form an alliance to jointly develop a product. They might connect their intranets or parts of the intranets to each other, using a private, leased telephone line or even the public Internet. A company's extranet could include shared content in the form of private newsgroups that let representatives from two or more companies hash out ideas and coordinate projects.

HOW DOES EMAIL WORK

Email is not that different from regular mail, actually: you have a message, an address, and a carrier that figures out to get it from here to there.

The difference is that email messages--and any attachments--are broken down into small chunks of data called packets, which travel independently, weaving their way along with innumerable other packets traveling to different destinations. It's as if each page of a letter was mailed separately. On the way, the packets are passed from one server to the next until they reach their final destination. Any given message's packets and attached file may travel by several different
routes, so the components often arrive out of order and at different times. Once all the packets have arrived, they are recombined into their original form.

This makes sending the message faster, because it doesn't require transmitting one big, bandwidth-hogging piece of data. But it also means that an entire message can be held up if one little piece is missing. Usually, however, this entire process, traveling 3,000 miles or more, takes less than a minute to complete.

But what if you don't know a person's email address? It's actually pretty easy to track someone down on the Net. Web sites such as WhoWhere, Four11, and Bigfoot list individual and business email addresses; all you have to do is type in the name. Most search engine sites offer similar features, as do the Netscape Messenger, Microsoft Outlook Express, and Eudora email programs.

If you want to send or receive a mass mailing, you need to subscribe to a mailing list, also known as a listserv. Usenet newsgroups, on the other hand, are publicly stored messages that anyone can look at; you don't have to be a member of a list to read these messages.

**WHAT ARE NEWSGROUPS**

Say you have an overwhelming passion for rose gardening. (Just imagine here that you do.) How do you find other rose gardeners to swap tips with and brag to? Easy: join a newsgroup.

Newsgroups are publicly posted discussion forums--kind of an electronic clubhouse for people with shared interests. The messages are presented in a list, known as a thread, that shows the original message, the responses to the message, and the responses to the responses, so that you can follow an entire conversation or just the parts you're interested in.

Your browser alone won't let you get to newsgroups. You can read and post messages using either standalone newsreader software, such as Forté's Free Agent, or a newsreader that's a separate part of a Web browser package, such as Netscape's Collabra or Microsoft's Outlook Express.

Your newsreader lets you check newsgroups the way your browser lets you surf Web sites. The Usenet is the world's largest collection of public newsgroups. The newsgroups go by a complex set of abbreviated names, with the first set of letters of a newsgroup's name indicating its primary subject, such as rec (recreation), soc (society), or comp (computers). Additional abbreviations are
separated by periods and are tacked on to indicate subtopics. It's not uncommon for an individual newsgroup to have five, six, or more elements in its name.

The messages in newsgroups are stored on news servers owned by ISPs, universities, companies, and other large entities all over the world. Most news servers keep only the more recent posts; they'd soon run out of storage space otherwise.

What happens if you can't find a newsgroup that covers your favorite topic? (Not rose gardening--there are plenty for that.) Well, you could create a new newsgroup--but not without a little effort. If you want your group to be a standard Usenet newsgroup (those whose names begin with comp, misc, news, rec, soc, sci, and talk), you must submit a highly bureaucratic document, called a Request For Discussion (RFD), to the news.groups newsgroup. The group then organizes a straw vote where anybody who wants to can vote on your proposal. (To find out more about the RFD process and how to write an RFD proposal, see Jon Bell's Creating New Newsgroups page.)

The alt newsgroup hierarchy was created because many people felt it was too difficult to create an ordinary newsgroup. (Contrary to popular belief, alt does not mean "alternative topics"; it means "alternative newsgroup management structure.") If you want to create a newsgroup without all the hassle, you post a suggestion in the alt.config newsgroup and leave it up to the news administrators--the ones who make the ultimate decision about carrying new alt newsgroups. For a guide to creating an alt newsgroup, read David Barr's So You Want to Create an Alt Newsgroup.

There are also such things as local and private newsgroups. A discussion group created on a corporate intranet is an example of a private newsgroup. Most ISPs offer a handful of local newsgroups where they make tech support announcements that no one but their customers would want to see.

WHAT'S ALL THE FUSS ABOUT PUSH?

Some people say the Internet is just too much work: you have to go out and look everything up for yourself, and lots of stuff is hard to find. The industry came up with push technology to solve this problem by delivering--pushing--the information you want directly to your computer so that all you have to do is read it.
To understand push, you first need to understand pull, the basic delivery model of the Web. Pulled content is what you get when your browser looks at a Web page—it goes out to the Web server where the page is stored and pulls it to your desktop so you can look at it. In other words, you go out and look for Web sites on your own.

The push model, however, is much more like television; the information you're interested in arrives automatically on your desktop. Even the vocabulary is adopted from TV; the content provider "broadcasts" its information via "channels" that your browser "tunes in to." Each channel has different kinds of content—news, entertainment, corporate press releases, stock information, and so on—and users pick the channels that they're interested in. There's no guarantee that you'll want to read every single story, but you won't have to go looking for any of it. The latest information is always there, no matter which channel you choose, and all you have to do is be connected to the Net.

The PointCast Network is the pioneer of push. It uses its own Web client to tune into its channels, but is also organizing "networks" of channels to be delivered via browsers.

There's also what is called smart pull or push-pull hybrids. A good example is Lanacom's Headliner. It pushes up-to-the-minute headlines to your desktop, but to get the stories behind the headlines, you must click them, which takes you to the Web site where the full story can be found.

Push can also be used to distribute software. Marimba, for example, is promoting push technology that can update software automatically over corporate networks; before you even know that your word processor is out of date, the network delivers a new one.

The problem with push is that it clogs up the wires. Unlike Web surfing, push delivers information to you whether or not you want it at that particular moment; it can impose an especially heavy burden on corporate networks if everyone is receiving data broadcasts several times a day. Push technology vendors are devising various techniques to alleviate this problem, but it's still an issue.

HOW DO I MAKE A WEB PAGE
ou can make this easy or hard on yourself, depending on how complex you want the final product to be. At the very least, you need some content formatted in HTML, a place on the Internet to store your pages, and a way to transfer new or updated content.

If you want to put up a page with a list of your ten favorite movies of all time, a picture of your dog, and a link to your best friend's Web page, then creating your page should be easy and quick. If you want to do something fancier, like include interactive forms or animated graphics, things get more complicated--but there are lots of books and how-to guides to help you out.

To create a simple Web page without learning HTML, use a Wysiwyg HTML page creator, such as Netscape's Composer or Microsoft's FrontPage 98. If you know or are willing to learn HTML, use a tool where you can type in the HTML code, such as HomeSite or BBEdit. For the pros and cons of various HTML editors.

Once you've got your pages created, you need a place to store them. Most ISPs offer their customers from 1MB to 10MB of server space as part of their basic Net access package.

Although using your ISP's server will work just fine, your site's URL, or Web address, will most likely be long, ugly, and hard to remember--for instance: http://www.tiac.net/users/sfinnie/index.html. If that bothers you, check out your ISP's domain name service. For an additional monthly fee plus an up-front charge in the $50 to $100 range, you'll get a friendly URL, such as www.mycoolsite.com.

After you've launched your site, you'll want to get people to look at it. The easiest way is to register it with the various search engines and Web directories. Submission services handle the job of submitting to many search engines at once.

HOW DO I TALK TO PEOPLE ON THE NET

Who says that computers isolate people? The Net is a great place for a good, old-fashioned chat: real-time (albeit written) communication with a friend--or a complete stranger.

Online services have public chat groups, or rooms, devoted to a variety of subjects; if you find someone you want to talk to one-on-one, you can also go into private chat rooms. But while AOL and the other online services make chatting easy, there are lots more ways to talk on the Net.
Internet Relay Chat (IRC) lets groups of people chat on any of hundreds of public chat servers. First off, you need an IRC chat client, such as mIRC. You can start a chat group (called a channel) or join an existing one.

Using IRC does require a fairly involved setup process, but it offers the most established--and largest--chat forums on the Net. If the hassle isn't worth it to you, there are hundreds of newer chat sites right on the World Wide Web, and all you need is your browser to join in. Just pick a site, become a member (this is usually free), choose a channel, and start chatting.

For something completely different, Telnet to a Multiple-User Dimension (MUD) or two. These text-based virtual worlds, many of which feature a dungeons-and-dragons-type theme, usually stress the environment and the action over the actual chatting.

But no matter what kind of chat you choose, keep in mind that the world of chat can be a wild and woolly one. You can avoid most trouble--such as flames or sexual harassment--by obeying the rules of chatiquette and taking some simple safety precautions. And if you don't like the tone of a chat session, just leave and find one that's more to your taste. Or try a chat with a moderator--such as a public chat with a celebrity.
HOW ELSE CAN I USE THE NET

The average Net user surfs the Web, sends email, subscribes to a newsgroup, and maybe gabs in online chat rooms--and that seems enough to keep the ISPs in business. But today's Net applications also have some colorful ancestors, such as Archie, Veronica, and Jughead, that are still in action, although most are not as popular as once they were.

It all started with file transfer protocol (FTP), a method of moving files around the Net. Internet files, mostly software intended for downloading, are stored on FTP servers; you use an FTP client program, such as Ipswitch's WS_FTP, to get at them--that is, download, upload, move, and/or rename the files.

Archie is a search tool that hunts down specific files by name in FTP archives. (Archie equals archive; get it?)

Telnet lets you log on over the Net to the Unix-based servers often found at public and private libraries, as well as university resource centers. Telnet lets you access and search directories; it also lets you read, copy, and sometimes even add files. Basically, it opens up rich databases of information that aren't available anywhere else on the Net.

Gopher is a tool for finding and transferring FTP and Telnet files on the Net. Developed at the University of Minnesota, Gopher is named both for the school's mascot and for its ability to "go fer" files. Once the most common means of navigating the Net, Gopher's popularity has been in steady decline since the rise of the flashier World Wide Web, but to get a sense of Net history, use your browser to take a look at the University of Minnesota Gopher.

Veronica and Jughead are searchable indexes for Gopher archives, named as a play on Archie. It appears the developers were fans of Archie comics--though we never did see tools named Betty, Reggie, or Big Moose.

WILL THE WEB BECOME MORE LIKE TV?

Some people used to say that the Web would eventually mean the end of TV. Why would you want to watch more reruns, after all, when you could be surfing the Net? Instead, the two technologies are becoming more like each other.
The idea of adding moving pictures to Web sites is irresistible. At the same time, TV broadcasters have imagined using the Net to grab and retain viewers' interest. The classic example is where someone watches a basketball game and then clicks the tattooed guy to find his rebound stats, his annual endorsement income, or pictures of him in drag.

To move toward this goal, companies such as WebTV (now owned by Microsoft) are working to bring the Internet to your living room via set-top boxes, small terminals that use phone lines to display Web pages on your television. Intel's Intercast technology, not yet in wide use, actually lets television programmers embed Web data in television signals, thus bypassing phone lines.

Meanwhile, streaming multimedia is making the Web more like TV. Instead of having to download a multimegabyte file before you can play it--a process that could take minutes or hours on a 33.6-kbps modem--a streaming file can start to play right away. CNET's own CNET Radio uses Progressive Networks' RealAudio, a popular streaming format for voice and music. Popular streaming formats for moving pictures include RealVideo (also by Progressive) and VXtreme's Web Theater (now owned by Microsoft).

But even streaming video is hampered by the limitations that modems place on data transfer. You just can't squeeze enough data through a standard phone line to play full-screen, full-motion video; most of it ends up looking, as the phrase goes, like dancing postage stamps.

The next big breakthrough in Web/TV integration will come when we can break this bandwidth barrier (see "Why is the Web so slow?"). ISDN helps, but to get television-quality video on a computer, we're going to need even more Internet throughput. This could eventually come from cable modems, which use cable television lines to transmit and receive Internet data, at speeds in the realm of 10 megabits per second.

In the meantime, do you really want to watch Seinfeld on your computer? Probably not. The pictures you get on your computer are small, choppy, and fuzzy. There are better uses for a $3,000 computer than imitating a broken-down $99 television with bad reception.
GLOSSARY

**ISDN**

Integrated Services Digital Network

The plain old telephone system doesn't handle large quantities of data, and the phone companies realized this a long time ago. So the ISDN spec was hammered out in 1984 to allow for wide-bandwidth digital transmission using the public switched telephone network. Under ISDN, a phone call can transfer 64 kilobits of digital data per second. But it's not always easy to adopt.

**POTS**

plain old telephone service

If you're logging on to the Net using a regular modem that employs your phone line, you have a POTS connection. POTS is the basic voice phone service you get from Ma Bell and her kids, and the term is used to differentiate this type of connection from ISDN or a leased line like T1.

**Unix**

Described by one of its developers as "a weak pun on Multics" (which was an experimental, time-sharing operating system at Bell Labs in the 1960s), Unix took off in the early 1970s as a general-purpose operating system. Since much of the Internet is hosted on Unix machines, the OS took on a new surge of popularity in the early 1990s.

Unix comes in many flavors--including Xenix, Ultrix, GNU, and Linux--and runs on a variety of platforms, which makes its development a subject of widespread discussion. But the truly great debate involves how to style the word itself: should it have an initial capital (Unix)? Or should it be in all caps (UNIX)? Since the operating system itself is case-sensitive, the debate rages. Bell Labs' implementation of Unix is trademarked in all caps; for the other implementations, it's optional.

**Telnet**
Telnet is an application that lets you log on to a Unix computer. Provided you have an account on that Telnet server, you can then use its resources. A drawback of Telnet is that it's character-based, so you need to speak Unix to the other computer.

**MUD**

Multiple-User Dimension

Originally known as a Multi-User Dungeon, a MUD is a text-based virtual environment in which users' "characters" interact in real time. Characters can navigate rooms described by text; type to other characters; create shared objects; and engage in games, puzzles, or combat. Most MUDs involve some level of role-playing, where characters adopt personas from books or medieval times, for example. A MUD can be accessed via Telnet or using a special MUD client.

**IRC**

Internet Relay Chat

IRC is a way of hooking up with other Net users to exchange written comments--live and in real time. To do this, you need an IRC client and an IRC server. Once connected to the server, you join a channel, or discussion group, which can include people from all over the world. IRC channels may hold discussions about anything under the sun (and some topics that shouldn't see the light of day). IRC can be accessed by a variety of downloadable software on both the PC and Mac.

**Usenet**

Usenet is a worldwide network of thousands of Unix systems with a decentralized administration. The Usenet systems exist to transmit postings to special-interest newsgroups covering just about any topic you can imagine (and many you wouldn't even want to imagine).

**firewall**

If you want to protect any networked server from damage (intentional or otherwise) by those who log in to it, you put up a firewall. This could be a dedicated computer equipped with security measures such as a dial-back feature, or it could be software-based protection called defensive coding.
**spider**

Also known as a Web spider, this class of robot software explores the World Wide Web by retrieving a document and following all the hyperlinks in it. Web sites tend to be so well linked that a spider can cover vast amounts of the Internet by starting from just a few sites. After following the links, spiders generate catalogs that can be accessed by search engines. Popular search sites like Alta Vista, Excite, and Lycos use this method.

**Cookie**

According to Netscape, cookies are a "general mechanism which server side connections can use to both store and retrieve information on the client side of the connection." In English, that means cookies are small data files written to your hard drive by some Web sites when you view them in your browser. These data files contain information the site can use to track such things as passwords, lists of pages you've visited, and the date when you last looked at a certain page.

**ISP**

Internet service provider

Once upon a time, you could only connect to the Internet if you belonged to a major university or had a note from the Pentagon. Not anymore: ISPs have arrived to act as your (ideally) user-friendly front end to all that the Internet offers. Most ISPs have a network of servers (mail, news, Web, and the like), routers, and modems attached to a permanent, high-speed Internet "backbone" connection. Subscribers can then dial into the local network to gain Internet access--without having to maintain servers, file for domain names, or learn Unix.

**HTTP**

hypertext transfer protocol

The protocol used to transmit and receive all data over the World Wide Web. When you type a URL into your browser, you're actually sending an HTTP request to a Web server for a page of information (that's why URLs all begin with "http://"). HTTP1.1, the latest version, is currently undergoing revisions to make it work more efficiently with TCP/IP
protocol

Computers can't just throw data at each other any old way. Because so many different types of computers and operating systems connect via modems or other connections, they have to follow communications rules called protocols. The Internet is a very heterogenous collection of networked computers and is full of different protocols, including PPP, TCP/IP, SLIP, and ftp.

PPP

Point-to-point protocol

PPP is the Internet standard for serial communications. Newer and better than its predecessor, SLIP, PPP defines how your modem connection exchanges data packets with other systems on the Internet.

SLIP

Serial line Internet protocol

SLIP is a standard for connecting to the Internet with a modem over a phone line. It has serious trouble with noisy dial-up lines and other error-prone connections, so look to higher-level protocols like PPP for error correction.

TCP/IP

Transmission control protocol/Internet protocol

These two protocols were developed by the U.S. military to allow computers to talk to each other over long distance networks. IP is responsible for moving packets of data between nodes. TCP is responsible for verifying delivery from client to server. TCP/IP forms the basis of the Internet, and is built into every common modern operating system (including all flavors of Unix, the Mac OS, and the latest versions of Windows).
data packet

Although your computer and modem can send data one character at a time, when you're surfing the Internet, downloading files, or sending email, it's more efficient to send information in larger blocks called data packets. Modems generally send packets of around 64 characters along with some extras for error checking. When downloading files using a protocol like Xmodem, however, the packets are larger. And when using Internet protocols such as TCP/IP, the packets are larger still--around 1,500 characters.

asynchronous communication

This term describes how your computer uses a modem to connect with other computers. Back in the days of teletypes and dumb terminals, computers sent data synchronously--they operated using a shared timer that marked the transmission of each character. This didn't work very efficiently for large blocks of data over phone lines, however. So modern modems use asynchronous rules: instead of synching up to a time signal to mark a character, transmitting computers use a start bit, a stop bit, and an optional error-checking parity bit to indicate to receiving computers the boundary of each character. (The term is a bit of a misnomer, though, since all modems synch up with one another before they transmit data.)

SMTP

simple mail transfer protocol

When you're exchanging electronic mail on the Internet, SMTP is what keeps the process orderly. It's a protocol that regulates what goes on between the mail servers.

IP

Internet protocol

The Internet protocol defines how information gets passed between systems across the Internet.

NNTP

network news transfer protocol
Usenet news articles can't be just posted and accessed willy-nilly, so they conform to this protocol that runs interference between newsreaders and news servers. NNTP dictates the way in which news articles are distributed, queried, retrieved, and posted.

**Router**

This piece of hardware does what it says: it routes data from a local area network (LAN) to a phone line's long distance line. Routers also act as traffic cops, allowing only authorized machines to transmit data into the local network so that private information can remain secure. In addition to supporting these dial-in and leased connections, routers also handle errors, keep network usage statistics, and handle security issues.

**LAN**

local area network

A local area network is a short-distance network used to link a group of computers together within a building. 10BaseT Ethernet is the most commonly used form of LAN. A piece of hardware called a hub serves as the common wiring point, enabling data to be sent from one machine to another over the network. LANs are typically limited to distances of less than 500 meters and provide low-cost, high-bandwidth networking capabilities within a small geographical area.

**Ethernet**

Ethernet is a standard for connecting computers into a local area network (LAN). The most common form of Ethernet is called 10BaseT, which denotes a peak transmission speed of 10 mbps using copper twisted-pair cable.

**fast Ethernet**

Fast Ethernet is an upgraded standard for connecting computers into a local area network (LAN). It works just like regular Ethernet except that it can transfer data at a peak rate of 100 mbps. Also referred to as 100BaseT, fast Ethernet is more expensive and less common than its slower 10BaseT sibling.
Visual Basic

A high-level programming language from Microsoft that's graphically oriented and relatively easy to learn, Visual Basic can be used to create everything from simple database applications to commercial software packages.

hub

This chunk of hardware is used to network computers together (usually over an Ethernet connection). It serves as a common wiring point so that information can flow through one central location to any other computer on the network.