A new project makes it possible to access information anywhere

The world in a Web

Joe Levy

IMAGINE a system that links all the text, data, digital sounds, graphics and video on all the world's computers into a single interlinked hypermedia "web". This is the potential of the Internet-based World Wide Web (WWW or W3) project which stems from CERN, the Centre Européene pour la Recherche Nucleaire.

Of course the Internet doesn't connect all the world's computers, being primarily a network of university and research machines. But this is changing fast, and it is becoming easier for home users to access the network through service providers like Demon.

The WWW project is based on the principle of "universal readership": if information is available, then any (authorised) person should be able to access it from anywhere in the world, using one simple program. The system allows for different information formats so that existing and future information need not be especially tailored to be accessible.

Suppose you need to know if there are any recent articles on high definition television available on the Internet. Your friend in the US is an expert on the subject, so you open the "home page" provided by her university's W3 server — easy since you happen to know its address. The page provides information on the university and has links to several departments; you mouse-click on the link to your friend's department.

A diagram of the department's floorplan appears on the screen and you click on the region representing her office. Up pops a page with her picture, a message recorded in digital audio, details of her work, and a set of links to the information on the net that she uses regularly, including topics in the subject that you are interested in.

A few more clicks take you to an abstract of a paper that looks interesting. This might be a paper written in hypertext with links to other documents or data such as sounds, graphics or digital video. But more likely at the moment, it will be the .name of a file on a "remote computer". You click on the link and the file transfer protocol is initiated and the file is transferred to your local machine. You have found the information by a combination of vague knowledge and the interlinked nature of the Web. On the way you

have had the chance to explore other useful information.

The Web is based on the idea of hypertext: that is, computerised text with links to other bits of text. Thus a heading might be a link to a particular section of text, a word might have a link to its definition in an on-line dictionary, a citation might have a link to an on-line version of the publication referred to, and so on. A hypermedia system extends the idea to pictures, sounds and video, with the running of separate applications such as databases and communications programs.

The Web runs on the usual client-server model. Most information providers run a server which provides information to a client somewhere else on the Internet. The two communicate using a simple protocol, HTTP (Hypertext Transmission Protocol). One of the features of the protocol is that the client can send information on which data formats it understands so that the server can respond in the correct way.

Client programs are called "browsers". The most popular is probably Mosaic, which is provided free by America's NCSA (National Center for Supercomputing Applications). At the moment this is mostly used on X-based Unix workstations, but Apple Macintosh and Microsoft Windows versions are being developed.

Mosaic provides a windowed interface and access to several "home pages" to get you started in accessing the Web. Like other browsers, Mosaic also provides telnet, ftp, archie, gopher and other services. When used to browse a source of information, Mosaic responds appropriately: a hypertext document will be displayed in the correct format. Plain text, graphics, digital audio and MPEG-standard digital video are passed to the appropriate program on the machine where they are displayed or output.

Information can be provided on the Internet by simply making a file available for anonymous file trans-

fer (ftp). However, the most flexible method of embedding information into the web is to provide it in a hypertext format using HTML, the hypertext mark-up language. This can be interpreted by a browser so that text appears in a reasonable format (section headings, paragraphs, bulleted lists, in-line graphics etc).

The other important feature of HTML is, of course, that it allows you to insert links to other files on the Internet. The file is located using a format that allows the address of the machine on the Internet, the file location and file format to be specified. The link is highlighted on the screen and can be accessed with a simple click of the mouse. Mosaic allows each file to be personally annotated with either voice or text.

The Web was initially developed at CERN to serve its constituency of high energy physicists but is now used to make all sorts information accessible, from poetry and history of art to the documents provided by the White House and US satellite weather maps.

The traffic on the CERN Web server is doubling every four months, and there are now at least 100 servers on the Internet with a potential user base of 15 million users in more than 50 countries.

The World Wide Web's simple formats and protocols have allowed a very fast expansion. The future will bring more sophisticated but compatible systems that allow more elaborate hypertext, security features and, inevitably, a system whereby information providers can charge users for information.

To access the World Wide Web, telnet to info.cern.ch (128.141.201.74). This will bring up a simple text-based query system with pointers to information on the Web. Mosaic is available by anonymous ftp from ftp.ncsa.uiuc.edu in /Web/xmosaic. Demon Internet Services: 081-349 0063. Dr Levy works at the Human Communication Research Centre, Edinburgh University. Email: joe@cogsci.ed.ac.uk.

■ **Mondo 2000: A User's Guide To The New Edge** *(Thames and Hudson, £12.95)*

WITH Mondo 2000 magazine not being available from most bookstalls, R. U. Sirius and Queen Mu have produced a paperback collection, edited by cyberpunk novelist, software designer and university professor Rudy Rucker. The 42 sections go from Aphrodisiacs via Brain Implants and Smart Drugs to Zines, which explains the disclaimer: "This book is not intended to replace medical or professional advice, and you should consult your doctor or other professional before embarking on any of the medical or other programs described herein." But the bulk of the content is about the new wave computing counterculture: artificial life, cyberspace, electronic freedom, hacking, multimedia, The Net, virtual reality and, of course, virtual sex. If all that isn't enough, the section on personal computing has a wonderful photograph of Bill Gates's head superimposed on a body-builder's bulging torso. Buy two copies and you'll have one to frame.

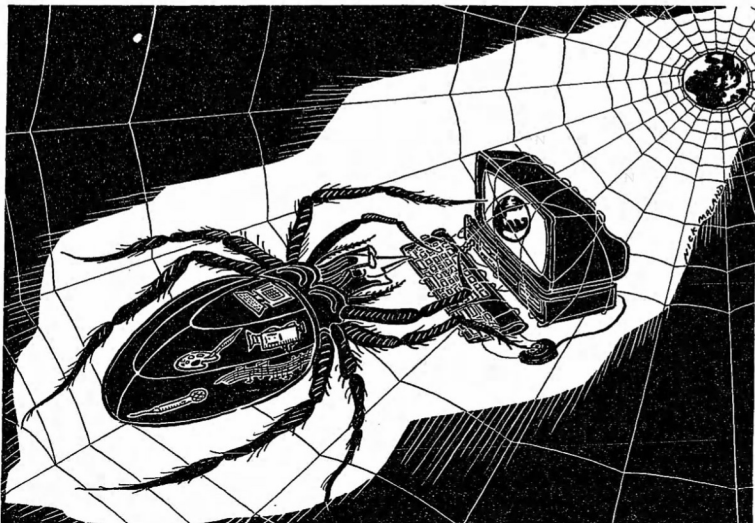■ **The Internet Navigator** *by Paul Gilster (John Wiley, £21.95)*

THE Internet is bursting with information about using the Internet, but there is a chicken and egg problem to obtaining it. The Internet Navigator is a good solution. It aims to be "the essential guide . . . for the individual dial-up user". It contains almost 500 pages of detailed information about using the net, complete with explanations of commands and example print-outs. As well as the usual stuff — telnet, ftp etc — coverage includes gopher, veronica and whois servers, and developments such as Project Gutenberg and World Wide Web. The book is up to date enough to include Mime, Internet Talk Radio and the President Clinton/White House connections. Unlike many American computer books, it is written in English and takes a global view. The copious contents pages and excellent index make it simple to find things.

■ **The PC Plus Modem and Communications Guidebook** *by Sue Schofield (Future Business Books, £19.95)*

FUTURE'S first book is a nicely presented beginners' guide. Sue Schofield (no relation) knows comms well, though her editor should have picked up a few minor errors in the associated chat. As well as covering the basics, the book is unusual in having chapters about topics such as viruses (etc), ISDN, teleworking, CompuServe, the Internet, and radio-based communications. It also covers Cix, though not Telecom Gold or AppleLink. While the book doesn't come with a free disc, you can send off for a shareware comms program (£2.50), or a free CompuServe introductory membership kit.

■ **The Way Computers and MS-DOS Work** *by Simon Collin (Microsoft Press, £14.95)*

BILLED as "The Ultimate All-Color Beginners' Guide", this is the first in a series produced in the UK by Dorling Kindersley, and typical of the company's work. The book is copiously illustrated with photographs and line drawings, beautifully laid out, and has small panels of simple text. It is up to date in covering MS-DOS 6, and future volumes will explain Microsoft Excel and Word for Windows. For people who really don't know anything about PCs, this series looks like becoming a good place to start.