Let your fingers do the walking

The Internet is, at its very simplest, a massive information storage depot. Data of all kinds, from all over the world, is stored here and finding it can present quite a problem. Everyone has trouble locating files on their hard disks sometimes, and most of them are probably only 500Mb in size. Imagine trying to locate that same small file on a hard disk that is thousands of times bigger, remembering that the file may be stored in any one of many disparate locations on a worldwide basis, and you can see just what a problem finding information on the Internet could be. Luckily it doesn't have to be quite as horrendous as you think. This month we'll show you exactly how to find what you are looking for, quickly and without fuss.

The Internet is no longer dominated by the text document, and while there is still a tremendous amount of textual information to be found spread over the Net, you will also find binary files in the form of complete software (public domain and shareware predominantly, but many companies also make software patches available online); graphics files, including video and animation as well as still images; and audio files. Returning to textual information, this can take the form of complete books, abstracts from published articles, biographical data, company information and stock market prices. In fact, just about every type of text-based material you can think of is accessible somewhere online.

STARTING THE SEARCH

The Internet allows its users to download files using the File Transfer Protocol (FTP). The computers that store files on the Internet are called FTP sites, but you cannot simply connect to any of these and download files at will. Many sites are open only to those who have the necessary access privileges, and these might be commercial concerns or scientific institutes. There is, however, a massive number of Anonymous FTP sites, which allow anyone who is online to look around and download files. Some of these sites are quite small and only store a modest number of files, some may specialise in holding files of a specific nature and yet others are vast shareware depositories that hold hundreds of gigabytes of files.

While you can connect to an Anonymous FTP site and browse the public directories for interesting programs, this is not only rather a hit-and-miss affair but also extremely time-consuming. Much better would be a method of querying a database for a specific keyword and getting a list of site addresses and directory paths holding the file. This is where Archie comes in. Developed by the McGill School of Computer Science in Canada, Archie gets its name from the word 'archive'. The Archie database holds information on files stored at more than 1,000 Anonymous FTP sites worldwide, and will return a list that gives details of file location, including the site, directory path and filename.

Archie can be accessed in many dif-

Archie commands

Archie has a number of command options, some of these are vital if you are connecting to a server using Telnet and are still useful to know about even if you are using one of the new breed of point-and-click graphical clients.

Help: probably the most useful command of them all, 'help' will return a list of all commands that are valid and available.

List: this reports on all the Internet sites that are stored in the Archie database, and includes the date that site information was last updated (which can be useful to know).

Prog: the most important command of the lot, 'prog' is the command used to initiate the Archie search in the first place.

Set: the 'set' command is used to set the various defaults for use during your Archie session. Use the unset command to toggle any of the options.

These are some of the parameters that can be used with the 'set' command:

Maxhits: this restricts the number of matches that Archie will return in any one search. The default is 1,000, which we would advise you to reduce as a matter of course. A default of 100 we find to be perfectly adequate.

Pager: enables screen paging, thus preventing all that useful information from scrolling off your monitor before you can read it.

Search: this sets the search method that you will use, for example 'set search sub' would make sub searches the default.

Sortby: sets the order by which the results of an Archie search are listed; you can choose from none, filename, hostname, size or time.

Whatis: whatis allows you to search the 'software description database' for keywords. This is a database of short file descriptions submitted by the uploaders of files to Anonymous FTP sites. The descriptions are not checked, not updated and searches can be somewhat haphazard in their success. However, a whatis search can be useful if you don't want to make a thorough Archie search.

The Internet contains more information than a million hard disks, so it comes as no surprise that finding a small text file, a modest shareware application or one of the 30 million cyberspace cadets out there can seem like a somewhat mountainous task. Davey Winder reveals the secrets of successful Net-searching

Gopher commands

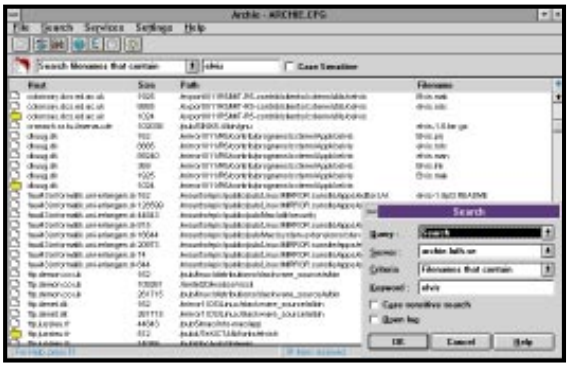
There is a number of commands that will be useful to know when using this type of client and here are some of them. The cursor keys will let you navigate around the Gopher menu:

UpArrow	Moves to the previous line
DownArrow	Moves to the next line
RightArrow	Selects the current item
LeftArrow	Exits the current item
>	View the next page of information
<	View the previous page of information
m	Return to the main menu
D	Download a file
n	Find the next item during a search
q	Quit
s	Save the current item to a file
=	Display information about the current item
/	Search for a menu item

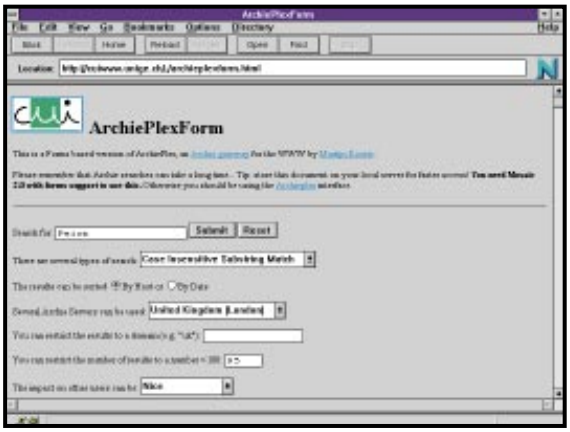
ferent ways, including by Telnet, by E-mail, by a special program or via a Gopher. Using a graphical Archie client, such as the one provided with the Internet Chameleon suite from Netmanage, makes the whole process very simple indeed. To connect via Telnet to an Archie server such as the one at Imperial College in London, first type in the destination, which in this case is archie.doc.ic.ac.uk. Then, when you get a reply, use the login of 'archie'. E-mail requests should be directed to archie@doc.ic.ac.uk. A message consisting of the word 'help' will return a useful file that tells you all you need to know about using this service.

To use a Gopher to perform an Archie search, just look out for a Gopher menu item that mentions Archie. To use Archie from the World Wide Web (WWW), you'll have to first access an Archie site like the Archieplex site. This can be accessed with a forms-compliant Web browser such as Netscape, and it's situated at http://cui→www.unige.ch/meta-index.html.

Whichever method you use to access Archie, it will still work in the same basic way. There are four types of searches that can be performed, the most useful (and certainly the easiest to use) being the 'sub' search. This gives a non-case-sensitive search for filenames that either match or contain a specified keyword. If you know the full name of the file you're looking for, then you could use the 'exact' search method. 'Subcase' searches are the same as 'sub' searches except they are case-sensitive, and the final method is the 'regex' search which uses regular expressions from Unix as wildcards. It is to be avoided unless you are familiar with Unix commands, and it should be remembered that Archie doesn't use Dos wildcards, so searching for a file called quote.txt with quo*.txt would not work (although searching for 'quote' or 'quote.txt' would).



Using the Archie client from Internet Chameleon is as straightforward as using any directory utility



An alternative view of Archie, as seen from the WWW

FINDING FRIENDS

You may find locating people on the Internet a difficult task. This is one area that has not been successfully addressed as of yet, mainly due to the enormous number of people who are actually connected to the Net. An electronic version of the telephone directory, containing E-mail addresses of more than 30 million people would be quite a feat of organisation. However, there are some methods that are open to you if you want to find a person or organisation on the Internet.

If you know the site where the person concerned is based, such as a university for example, then send E-mail addressed to the 'postmaster' at that address. Ask if they know the correct E-mail address for the person, and you may be in luck. You can also try to 'finger' the person at a site. A finger is a simple way of checking for details of users at specific sites, but it relies on you knowing the site and the name of the ▶

person. A number of service providers actually disable the ability to finger them, so as to prevent unwanted invasions of privacy. Assuming that you have a Finger client, the command syntax is simple: finger <name>@<site>.

An Internet tool called 'Netfind' can be queried with the name of the person, together with a number of 'keys', such as the site name and the country. Although not perfect, and certainly not comprehensive, Netfind is one of the better methods of locating people. You can Telnet the Netfind site at Imperial College in the UK, based at monolith.cc.ic.ac.uk, and use the login 'netfind'. Alternatively, you can access Netfind from the World Wide Web if you have a forms-compliant browser such as Netscape. Point it at:
http://cuiwww.unige.ch/meta-index.html

A reliable method of finding people, providing that they are users of the Usenet system (even very occasional ones), is the database of Usenet posters known as the Usenet Address Server. You just send an E-mail message to mail-server@rtfm.mit.edu which contains text in the format of send usenet-addresses/<name>. If you want more comprehensive details of this service send a message saying 'send' to:
usenet-addresses/help.

GET GOING WITH A GOPHER

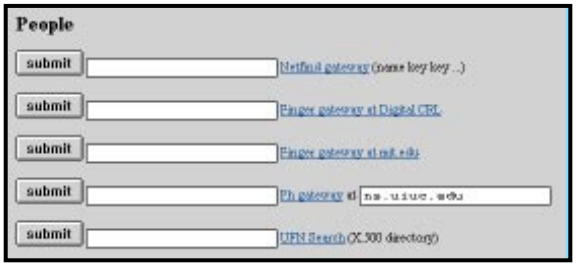
A Gopher, named after a burrowing rodent that is the mascot of the University of Minnesota where the application was developed, provides the user with a menu-driven interface with which to explore the Internet's information stores. There are hundreds of Gopher servers around the world, and there are some attractive graphical Gopher clients that make the application a joy to use.

If you can use the Windows File Manager, you will have no difficulty in using a Gopher client, as they look very similar. A Gopher presents the user with a menu of items, from which you can work your way through further menus to arrive at the information you require. Gophers can be used to search for information of any variety, but can also perform Archie searches, read Usenet Newsgroups and Telnet to other computers. It is important to remember, though, that Gopher can only search for filenames and directory names, and not for keywords contained within documents stored under those filenames. Providing you remember this limitation, you should have no problems.

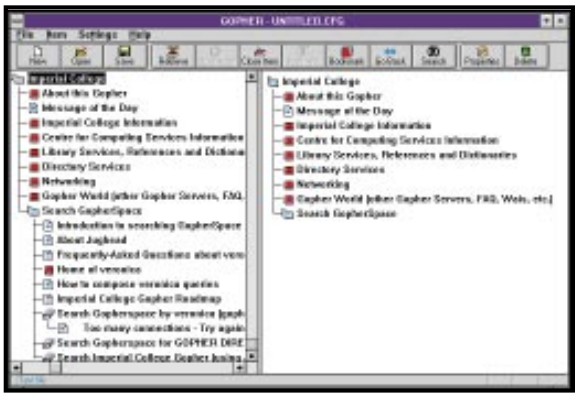
An entry in a Gopher menu can be any of the previously mentioned application types (selecting an entry may initiate a Telnet connection for example), as well as being a binary file, an image or even a MIME encapsulated E-mail file. To get the best out of this tool, you really need to make use of a graphical client, as 'text-and-cursor' type browsers will be restricted to plain text, obviously. All in all, it's an extremely useful tool and, because a Gopher only maintains connections to other sites while it is making requests or receiving any information (in the meantime it closes the connection), it saves on valuable network resources as well.

Internet White Pages

If you're trying to find someone on the Internet, there are several Gopher and Telnet sites that collect the E-mail addresses and names of users on the Internet. Netfind, at:
http://alpha.acast.nova.edu/netfind.html
provides a simple way to search for someone's E-mail address. If you'd like to discover more about Netfind, read the Internet Tools Summary at:
http://www.rpi.edu/Internet/Guides/decemj/→
itools/nir-utilities-netfind.html
These all have Gopher menus that list many universities, some companies and other helpful indexes. Texas Tech is at:
gopher://cs4sun.cs.ttu.edu:70/11/Phone%20Books
Notre Dame University is at:
gopher://gopher.nd.edu/11/Non-Notre%→
20Dame%20Information%20Sources/Phone→
%20Books-Other%20Institutions"></A>
and MIT is at:
gopher://sipb.mit.edu:70/1
Home Page Directories at:
http://www.rpi.edu/Internet/Guides/decemj/→
icmc/culture-people-lists.html
is the page to try if you're looking for someone who has a personal home page available on the WWW.
Knowbot Information Service at:
telnet://info.cnri.reston.va.us:185
will enable you to use a command line to search for a variety of information. You can enter commands like "Query Clark" to look for someone with that surname. (This interface is not for the faint of heart.)
Four11 Directory Service at:
http://www.four11.com
is another useful White Page site. Services include searching, custom listing, group connections and search agents.



Netfind and other resources to help locate those hard-to-find people, all these can be found at the W3 Search Engines site on the WWW


A simple graphical Gopher client (in this case from Internet Chameleon)

choose from. Character-based clients can be found at most of the online systems, and these are perfectly adequate for performing information searches.

However, to get the best from a Gopher with the minimum of fuss, as is the case with much of the Internet, you need to use one of the newer graphical clients such as HGopher or the one provided with Internet Chameleon. Using this type of approach you can navigate around Gopherspace using just the mouse. You can easily initiate Archie searches, provided that there is an Archie option on the Gopher menu of course, and file associations can be set up in the same way as with Windows File Manager so that file types are automatically handled and displayed correctly. It provides the added benefits of sound and vision.

One menu option that is finding its way on to increasingly more Gopher menus is Veronica. This keeps an index of Gopher items which can be searched by keyword, the results are then displayed as a fully interactive Gopher menu. Veronica is a useful addition to the Gopher family, making it easy to get quickly to the information you want without having to navigate through endless Gophers first.

Another newcomer you may be fortunate enough to have discovered is called Jughead. This performs the same function as a standard Gopher, except that it limits its searches to the local server. Although this doesn't, obviously, give you access to all the information on the Internet, it is very useful if you know that there is a good

GOPHER GOLD

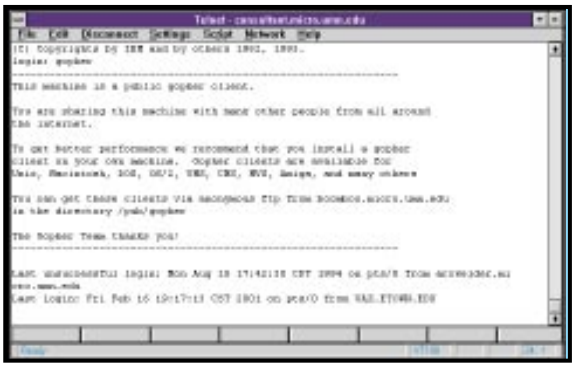
Gophers can be accessed by Telnet: for example, you could try consultant.micro.umn.edu and use 'gopher' as the login, but be warned that this method of access is not exactly speedy. If there is a lot of network traffic, the Gopher responses may slow to the point where they become unusable. Much better is to use a Gopher client of your own, and there are many of these to

chance that the information that you want will be stored in the near vicinity.

WIDE AREA SEARCHES

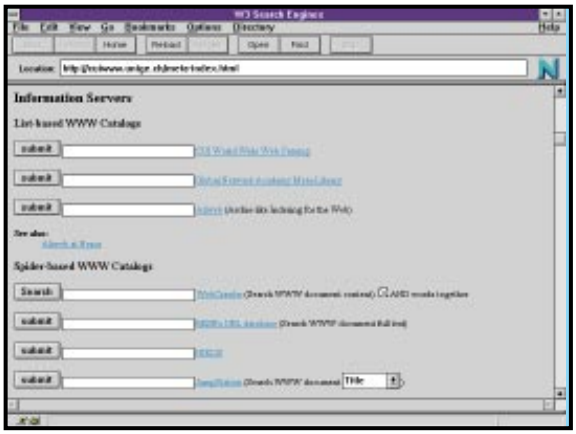
Wide Area Information Servers (WAIS) seem to be largely ignored by many Internet users. This is a shame, as they can be very useful tools if used properly. Rather than searching for a filename or directory name, as is the case with Gopher, WAIS allows searches for the text contained within those files and directories. This is accomplished by the searching of many disparate databases using the same standard interface. WAIS certainly isn't the quickest method of finding information, but when you consider that it is looking for your keywords among many millions of other words, perhaps you can be slightly more forgiving. Unfortunately, Boolean searches are not supported so you can't use 'and', 'or' and 'if' statements in your search criteria.

The ability to list documents found in order of how closely they match your search requirements is a nice feature, as is the use of 'relevance feedback' which lets the user mark documents which are the most relevant. WAIS will then go away and look for documents that it feels are similar to those that you have marked. Although this can be a time-consuming process, the results are generally worth your time. You can access WAIS from many Gopher menus: just look for the WAIS menu item. Alternatively, you can use the World Wide Web to access WAIS by pointing your Web browser towards http://www.wais.com.



Telnetting to a text-based Gopher certainly isn't as pretty as the graphical clients, and it's much slower as well

Web WAIS: a simple yet powerful tool



Probably the most useful information-searching resource on the Internet: the W3 Search Engines site

In addition, you could also try http://home.mcom.com/home/→internet-search.html. If you're a Demon user, you'll find this as a button on your home page, which has access to a variety of Web search engines.

Each search engine functions in a different way and has access to different document sources. The Lycos search engine, served by Carnegie Mellon University, will allow you to search on document titles and content. The Lycos index is built by a WebCrawler that can bring in 5,000 documents per day. The index searches document title, headings, links and keywords it locates in these documents. The WebCrawler is part of the WebCrawler project, managed by Brian Pinkerton at the University of Washington. It allows searches by document title and content. The Centre Universitaire d'Informatique World Wide Web (CUI W3) catalogue, however, searches through directories (summaries) of other WWW documents it finds.

If you still haven't found what you're looking for and you'd like to try out other available search engines, check out these alternatives: http://cuiwww.unige.ch/→meta-index.html, which is the W3 Search Engine, published by the University of Geneva; http://web.nexor.co.uk/susi/→cusi.html, which is an extension to the NEXOR communications company engine that provides you with a search tool which can help you find software indexes, people indexes and dictionaries.

If you're using the WWW, you've already discovered the most pleasantly presented information on the Internet. The WWW is the fastest-growing application to use on the Internet, bringing together vast quantities of disparate information which can be accessed by a user-friendly graphical browser. It comes in a hypertext format that allows information within documents, and beyond documents as well, to be linked together. The user can then follow threads of information without worrying about where it is stored or how to get there. The only drawback is finding the information in the first place, because there is simply so much of it. True, you can join the WWW at any point and eventually arrive at the information pages you require, but this would not only take a lot of time and effort, it would also unnecessarily waste network resources.

USING WEB SEARCH ENGINES

The way to find the right starting points is by using a Web search engine. These come in many varieties, and all of the most popular and useful can be accessed from the same spectacular site. The W3 Search Engine site can be found at http://cuiwww.unige.ch/meta-index.html, and contains just about every search engine that is worth using. Providing you have a forms-compatible Web browser (Mosaic, Netscape, or one of the variants will suffice) you just have to enter your keyword in the dialog box next to the search engine you wish to use, and there are plenty to choose from.

As well as the CUI WWW Catalogue itself, you can also find Aliweb, WebCrawler, Lycos and the WWW Worm among many others. It's not just WWW site information that is accessible here either, you can search for people on the Internet using Netfind and Finger, or perform Veronica, WAIS, and Archie searches – all from this one amazing resource. If you are stuck with some acronym, there are even searchable dictionaries here to help you. It really is the most useful single site we have come across in thousands of hours of browsing.

USING USENET

Usenet, a collection of online conferences each covering a specific subject and used by millions worldwide, is often overlooked as a method of seeking out the whereabouts of online information. If you want to know where to find something, why not ask in a relevant Usenet newsgroup? The chances are that someone reading this will know exactly where you should go, and they will be only to glad to tell you. The Internet, and Usenet in particular, is built upon a sense of community. Make the most of it and gain from the experience of those who have gone before you.