

CHAPTER 4

File Transfer over the Internet: File Transfer Protocol

File Transfer Protocol, or FTP, is the method by which files are moved between computers connected to the Internet. FTP is a popular activity on the Internet because it is interactive and relatively easy to use, and there are literally thousands of programs and data, sound and image files available for you to copy and use. For example:

- The Computer Oriented Geological Society or COGS ([URL - ftp://ftp.csn.org/COGS/](ftp://ftp.csn.org/COGS/)) archive contains a variety of MS DOS and Macintosh programs and source codes. You will find programs related to geochemistry, geophysics, hydrology, LANDSAT imagery, mapping software, and mineralogy, among others.
- The World Paleomagnetic Database archive ([URL - ftp://earth.eps.pitt.edu](ftp://earth.eps.pitt.edu)) permits you to download databases and programs related to paleomagnetism.

Once you have connected to a few FTP sites you will be amazed at the number of programs and databases that you have access to. A whole new resource for enriching your education and conducting research is at your fingertips.

File Transfer Protocol Servers and Their Resources

One of the original purposes for building the Internet was to provide a means for people to share resources between computers. The Internet was created to permit researchers to archive and send files between each other's computers as a distributed computer system. Within a distributed computing environment, a program that several people use need not be stored on each individual computer. Instead, the program is located on one computer and accessed by several users individually or simultaneously. Users of a distributed system have the flexibility to run a program or examine data on the remote computer or to copy the files to their local workstation when access is needed. Having programs distributed on remote computers negates the need for each person to have large hard drives to store and use the same information.

Like most other Internet services and systems, FTP is built around the client-server paradigm. To send (*upload*) and receive (*download*) files requires an FTP client program running on your local computer. The client program takes the commands issued by your computer, converts them into a "universal" set of commands that the remote FTP server can understand, and requests the server to perform the actions. The server receives the

commands, processes the request and returns the items requested.

File transfer is somewhat of a misnomer because you don't transfer the actual file. When you transfer a file you issue a command to the FTP server to read the contents of the file and send a copy of it over the Internet. The local desktop computer listens to the server's communication and puts the file into the local computer's storage banks. Internet network software ensures that no errors or noise interrupt the transfer and destroy the contents of the file. But just as with any kind of technology, problems can occur. Network connections get broken or delayed to the point that your system "times out" before an entire file can be received. This is more likely to occur with large files (one megabyte and larger) or during peak hours of network use.

In many ways, FTP is the complement to Telnet, the topic of Chapter 5. Where Telnet can offer powerful command structures for examining information online, FTP has fewer commands that perform similar though more limited functions. Moreover, Telnet is designed to let you see the information on your computer screen while the information stays on the remote computer. In most cases, FTP requires that you send files to your computer before viewing. Some FTP client programs let you view ASCII text files before transferring to your desktop computer. In contrast to many Telnet-based services, which are organized around a menu structure, FTP systems are organized by directories and subdirectories. This makes finding and navigating the information a little harder. FTP server system administrators often provide an index file listing the files that are contained in a particular directory.

FTP is so popular because the client software is widely available and many File Transfer Protocol servers do not require special accounts and passwords. These servers make their resources available through an *anonymous* FTP accounting system. An anonymous FTP service lets you log on to the server under an anonymous user name and generally without a special password. Most anonymous FTP servers will tell you what to enter as a password, and more often than not, the FTP server will request your electronic mail address as the password. In some cases, restrictions are placed on what directories and files you can access on the anonymous FTP server.

File Transfer Protocol software is efficient and easy to use. There are just a few simple commands that are required to move files between computers. Better yet, the interface has been designed so another computer program can use FTP to automatically send or retrieve files. For example, a program can be written to scan files on a computer once a night and then use FTP to transfer the updated versions to another location. This process would be particularly useful in updating files of information like weather data. A program can be written to scan the data that has been collected from a remote weather observation site and its contents sent by FTP to a main data bank for archiving.

File Transfer Protocol also is flexible, enabling you to transfer any kind of file between different types of computer platforms. The originators of the Internet recognized that people would be using different styles and brands of computers and thus designed FTP, as well as other Internet software protocols, to be “transparent” to computer system design. The client program is written for a specific type of computer like a Macintosh, but it doesn’t matter what kind of computer the server you communicate with is. You can transfer files from your Macintosh to a Sun Sparc Station or Pentium server. You can transfer:

- programs either in source or executable form
- ASCII text files, electronic mail messages, documents list, and so on
- graphic images (e.g., BMP, TIF, GIF)
- spreadsheets and databases
- audio, video and multimedia files

Many World Wide Web browsers have the ability to link to an FTP server for downloading files. So why should you use FTP software instead of some other Internet software? FTP software is exclusively designed for transferring files across the Internet. You could attach data files and documents to an electronic mail messages and send them to a remote location. However, if there is incompatibility between the sender’s and recipient’s email systems, then you might not be able to detach and read the file. Faxing information has become a popular way of distributing information. Faxing involves sending an image or picture of the message between individuals. The main drawback is that the information is difficult to reuse and equally hard to edit. Information that is already in digital form can be easily imported into a text editor or word processing package, read, edited, printed off or exported and sent back to the point of origin in a fast and efficient way. Above all, transferring information over the network using FTP is fast and inexpensive.

Earth Online Tip: If you download shareware programs, be courteous and send the author(s) the registration fee should you decide to continue to use the program.

Logging on to FTP

You transfer files via FTP with a client program. The client program tells the remote FTP archive what actions to take when you’re connected to it. Client programs are either text-based or graphically based interfaces. Web browsers enable you to save files to a local computer but uploading is not available. If you log on to an Internet provider under a shell account you can initiate an FTP session by simply typing “ftp” at the system prompt. Graphical user interface client programs running on your desktop computer are used over SLIP/PPP accounts and direct connections. Next, you’ll need the address of the system you want to establish communication with. There are a variety of ways of getting FTP site addresses. You can start by looking in Chapter 10, “Internet Resources for the Earth Sciences.” There are a number of other books that publish extensive lists of FTP sites.

People often publish lists of FTP sites in frequently-asked-question files (FAQs). Try sending a message to an electronic mail discussion list or a Usenet group asking for useful FTP sites. Use one of the several Internet search services like Archie, as explained in Chapter 6. Some FTP clients let you save the addresses along with descriptive information into a “hot list” file to make connecting the next time easier.

Once you have the address for an FTP server, you’ll need an account, a user ID and an account password for the target FTP system you’re connecting to. Most of the time you will be logging on to an anonymous FTP server so a user ID and password is not important. When prompted for a user name you simply enter the word “anonymous.” The server will ask you to enter either your full electronic mail address or a tab. It is proper netiquette to enter a correct email address if asked. Many systems like to keep track of the number of users and where they are connecting from. FTP has created more traffic on the Internet than any other application but soon will be replaced by the ever expanding World Wide Web. Most FTP sites limit the number of anonymous logins, so you might encounter difficulty when trying to connect. Have patience and try again later, maybe during off-peak hours.

After you have successfully logged on to the server, you will need to know the names of the files you want and in what directories they reside. Be aware that system administrators delete old or unused files and change subdirectory structures. Be forewarned that the address and directory location you obtain may be out of date or that the file you are looking for has been deleted or moved. You can get a listing of files and directories by entering the “dir” or “ls” command. Some system administrators will leave an index file in a subdirectory explaining what files reside there. You will also need to know the file format and if the file is compressed or uncompressed. Most files are either ASCII or binary. An accurate transfer of the file depends on knowing the file type. FTP does not understand the format or contents of a file. If a user requests FTP to transfer a file using an incorrect format, the resulting copy may be corrupted. The major file types are:

- ASCII: a file that contains text which is only printable characters and assorted punctuation and other characters
- binary: also referred to as an image format; the format used by most programs, spreadsheets, databases, and word processing documents

The File Transfer Protocol default transfer type is usually set to text. Binary transfer works for both ASCII and binary formats. Set the FTP transfer to binary to ensure the transfer of any non-ASCII characters if you’re not sure what the file format is. If the file you’re transferring is compressed, you will need a decompression program compatible with the format of the compression program used on the file.

A lot of time can be saved if you do a little planning before you start your session. Know the location to store the transferred file to, whether on your hard drive, floppy diskette, or a shared disk space on a local server. It’s advisable to place downloaded files into a temporary directory and then move them at some other time. File transfers can take

awhile if you have large files or a number of smaller files to transfer. Remember, someone else may be waiting for you to log off so they can get into the system.

After connecting to an FTP server you can obtain a listing of the contents of the directory you have logged into by issuing either of two commands having the same format:

```
ftp>dir <directory-name local-file-name>
```

or

```
ftp>ls <directory-name local-file name>
```

Both commands list the files in the given directory on the remote machine. Directory and file name commands are optional. Invoking the directory name command will cause the FTP server to display the contents of the requested directory. If the command is left blank then the contents of the current directory will be displayed. Using the second command (local file name) instructs the server to put the listing into the given file name on your local computer. Omit the command if you want the listing to be displayed on your computer screen. Wild-card characters are permitted with these commands. On most computers, an asterisk (*) is used to display all files with the same name or file extension. For instance, issuing the command

```
ftp>dir *.txt
```

displays all files with the extension “.txt.”

Common FTP Commands

There are several dozen FTP commands for working with and moving files between computers, some of which accomplish the same task. Many are similar to common UNIX commands. Thankfully you only need to know a few of them to get your job done. The more common commands, along with an example, are provided below:

ASCII or text	Set file transfer type to ASCII ftp> ascii
binary or image	Set file transfer type to binary ftp> binary
bye, close or exit	End FTP session ftp> close
cd directory-name	Change working directory on remote system ftp> cd hydrology
dir, ls	List directory on remote system ftp> dir
get file-name	Transfer copy of file from remote FTP system

help or ?	ftp> get mydata.dat Show FTP online help ftp> help
lcd	Change working directory on local (your) system ftp> lcd mydir
mget	Transfer copies of multiple files from remote system ftp> mget mydata.dat yourdata.dat
mput	Transfer copies of multiple files to remote system ftp> mput mydata.dat yourdata.dat
open	Open connection to specified remote system ftp> open ftp.remote.edu
put	Transfer copy of file to remote system ftp> put mydata.dat
pwd	Display working directory on remote system ftp> pwd
user	Specify username to remote system ftp> user

Now that you have a notion of what File Transfer Protocol is and how it works, let's look at an example session. The example shows the steps and keyboard input through a shell account on a UNIX machine. We will log on to a University of Illinois FTP server with a local client and download a file. There are two ways this can be accomplished. You can start the FTP client first and then use the open command to establish the connection to the FTP server:

```
%ftp
ftp>open vmd.cso.uiuc.edu
```

The alternative way is start the FTP client software and make the connection at the same time:

```
%ftp vmd.cso.uiuc.edu
```

In either case your client attempts to make the connection:

```
Connect to vmd.cso.uiuc.edu
220 FTP server at vmd.cso.uiuc.edu, 03:54:38 CDT Wednesday 10/18/95
222 Connection will close if idle for more than 5 minutes
Name (vmd.cso.uiuc.edu:mritter):anonymous
331 Send your local userid instead of password for identification, please
Password:
```

Your password will not appear on screen, even with an anonymous FTP session, when

typed. The remote FTP server responds with:

```
230 Welcome to the University of Illinois anonymous FTP
230
230
ftp>
```

After reaching the FTP prompt I'll type in the command "dir" to get a directory listing.

```
ftp>dir

200 Port request ok
125 List started ok
read me 80 67 6 2/08/94 10:41:28 anon
wx Notes 77 260 12 2/08/94 10:39:13 anon
```

The file I'm after is called "sources.zip." It is a FAQ file for meteorology resources that's in compressed format and located in the wx subdirectory. I'll change directories before attempting to download:

```
ftp>cd wx
250 Work directory is wx191 (Read Only)
```

I'll send another "dir" command to check and see if the sources.zip file is in the subdirectory.

```
ftp>dir
```

A long list of files is displayed (only the first couple are reprinted here to illustrate what is displayed on screen).

```
ACUS1 DOC v 75 82 2 1/02/93 16:24:59 wx
AURORA DOC v 77 226 4 4/29/95 9:44:37 wx
```

Finally, near the bottom of the list we find:

```
sources zip v 8192 9 17 10/11/95 9:02:45 wx
```

Now I'll set the file transfer to binary mode and download the file to my hard disk.

```
ftp>bin
200 Representation type is IMAGE
```

Check the spelling of the file before issuing any transfer commands. Notice that there

is a gap between the words “sources” and “zip” in the file listing above. Assume there is a period between the file name (sources) and the extension (zip). Now issue the “get” command with the full file name to transfer the file to your computer.

```
ftp>get sources.zip
```

```
200 Port request ok  
150 Sending file 'sources.zip'
```

It will take a short time for the file to be transferred. When the transfer is completed the remote FTP server responds with:

```
250 Transfer completed successfully.
```

```
local:sources.zip remote: sources.zip
```

```
66189 bytes received in 5.8e+02 seconds (0.11Kbytes/s)
```

The file has been successfully downloaded to my hard drive and I’m ready to decompress and examine it. I could have saved the file under a different name by simply typing a new name after the old. FTP will copy, transfer and save the file under the new file name on my desktop computer. For example:

```
ftp>get sources.zip metsource.zip
```

Transferring files via FTP is much simpler when using a graphical user interface client program like WS_FTP, a PC-Windows software application (Figure 4.1). When users start the program, a session profile window pops up for them to enter information about the FTP server they wish to connect to. Login configuration information can be saved for retrieval during future sessions. After they fill in the required information and click the connect button, the main application window appears. Directories and files on the local computer are presented in windows on the left hand side while the same information for the FTP server is displayed in the right hand side of the application. Users navigate their way through directories by either clicking on the directory title or clicking on the “ChgDir” button and filling in the correct path. Text files can be viewed by highlighting the file and clicking the view button, and soon the files are transferred to the local computer and displayed in a text editor. Buttons to handle making new subdirectories, removing old ones, and deleting and renaming files on either computer are provided. The user defines the type of file transfer by clicking on the appropriate button (ASCII, binary, L8). To transfer a file, the user simply highlights the file and clicks one of the arrow buttons. Clicking on the left arrow button will download a file to the desktop computer, while the right button is used to upload a file to the remote FTP server.

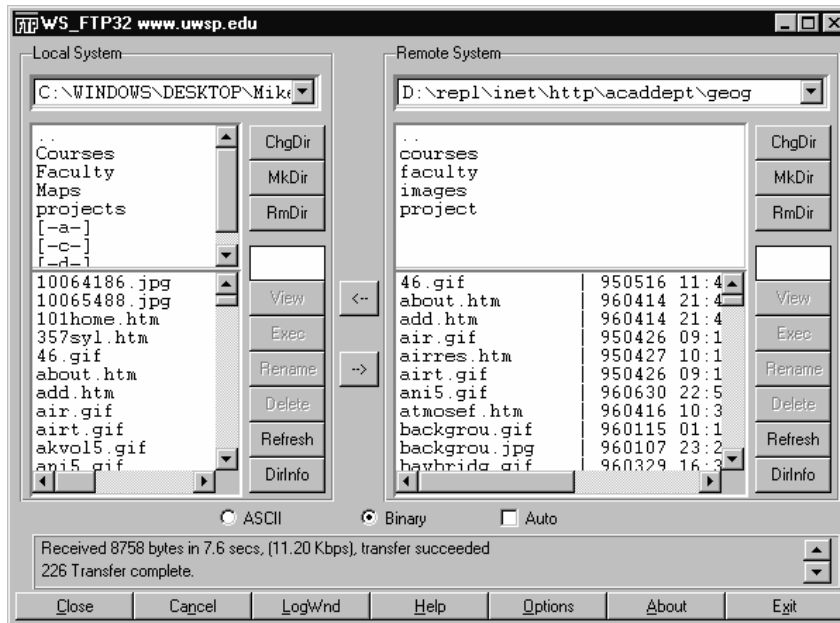


Figure 4.1 WS_FTP PC-Windows client

Files can be transferred via Gopher and World Wide Web browser applications too. For instance, over the World Wide Web you gain access to an FTP server by either clicking on a hyperlink (see Chapter 2) or typing in the FTP server address in the location field of the browser, preceding it with "ftp://" (without the quotes). This tells the browser that you are connecting to an FTP server rather than a World Wide Web server. Once connected the browser displays a list of subdirectories and files (Figure 4.1). Small icons indicate the entries as directories (folders), text files (document), and images (graphic). Additional icons for movies, sounds, and so on are used. Refer to the Apply It! section later in this chapter.

Compressed Files

Most files, like the one I downloaded in the example session above, are compressed to save on storage space. File compression programs scan files to reduce wasted space, duplicate characters, and so on in a file. There are a variety of compression techniques and thus a number of different compression programs available. Files run through one of these programs can be reduced from 30% to 70% in size.

Being smaller than their normal size, compressed files transfer more quickly across the Internet. Compressed files should always be transferred as binary files; to do otherwise is to risk corrupting the file. Compressed files can be identified by their extension. The more

common compression utilities are shown in Table 4.1.

Table 4.1 Common compression formats

Compression Program	Decompression Program	File Extension
compress	uncompress	.Z
pack	unpack	.z
Stuffit	unsit	.Sit
PKZIP	PKUNZIP	.Zip

The program you need to use to uncompress the file will depend on the computer you are using and the kind of compression program used. Uncompressing the file may be as simple as typing the program name and file name at the command line. For uncompressing a DOS file that has been compressed with PKZIP and is residing on your desktop computer's hard drive, type

c:pkunzip.exe source.zip

PKUNZIP will extract the file to your hard drive and your ready to use it. Consult your local computer systems administrator for the compression/decompression program you might need.

Earth Online Tip: Scan downloaded files for viruses before using them. Unfortunately, the Internet is a major source of computer virus infection.

Earth Science and FTP

Earth scientists are making good use of the opportunities afforded by File Transfer Protocol by making available valuable software and data files for research and educational use. For instance, Florida State University Supercomputer Computations Research Institute makes SciAn, a UNIX-based 3D-visualization package, available for downloading. Several important geoscience FTP servers have come online over the past several years including the Computer Oriented Geological Society or COGS ([URL - ftp://ftp.csn.org/COGS/](ftp://ftp.csn.org/COGS/)) server and the International Association for Mathematical Geology (IAMG) FTP site. The IAMG promotes international cooperation in the application of mathematics in geological research and technology. The IAMG also cooperates with the application of mathematics and statistics to biological and planetary sciences. IAMG's FTP site is an archive for algorithms and source codes of programs that have been published in Computers and Geosciences ([URL - ftp://www.iamg.org/pub/CG](ftp://www.iamg.org/pub/CG)) and Mathematical Geology ([URL - ftp://www.iamg.org/pub/MG](ftp://www.iamg.org/pub/MG)). Additional programs are available from Jim Carr's Numerical Analysis for

Geological Sciences (**URL - ftp://www.iamg.org/pub/Carr**). The FTP archive is intended to assist IAMG members in acquiring programs and data that have been previously difficult or time-consuming to obtain. Most programs and data have been compressed. Uploading files directly by the public is not possible as it is a “read only” site. However, it is accessible by anonymous FTP. New programs appear after the publication of each issue of Computers and Geosciences. Programs and data from older issues are being uploaded as time permits. Program submission guidelines are provided in the README file found in the /pub subdirectory. IAMG also has created a World Wide Web page interface to its FTP archives, making it much easier to obtain the programs (**URL - http://ftp://images.asc.nasa.gov**). From the FTP Web interface users can click on links to pages that contain the contents of each issue and the available programs and data. Links are provided from the issue pages to the programs available for downloading.

One of the greatest bonanzas for educators is the number of image archives that are linked to the Internet. Nearly 10,000 of NASA’s press release photographs from the manned space program are accessible via FTP (**URL - ftp://images.jsc.nasa.gov**), Gopher (**URL - gopher://images.jsc.nasa.gov/70**) and the World Wide Web (**URL - http://images.jsc.nasa.gov/**). The digital image collection includes pictures from the Mercury space program to the present. In 1995 some 30,000 earth observation photographs were loaded. For the most part these images are copyright free, like many other images loaded on the Internet by the government, and can be freely used by individuals. The National Geophysical Data Center at Boulder, Colorado, maintains an extensive set of photographs illustrating natural hazards around the world. Each set contains twenty color or black-and-white slides with accompanying background information. The NGDC makes the Earthquakes and the Volcano and Tsunami image sets and captions available for viewing online over the World Wide Web (see Chapter 2 for details about the World Wide Web), yet they are archived on an FTP server. Web pages display scaled-down versions of the photographs along with explanatory captions. Users download the image by clicking on the download link provided at the bottom of the page, which initiates an FTP.

What You Have Learned

- FTP is an interactive means to transfer files of any type between FTP clients (your computer) and remote FTP servers.
- FTP is a fast, efficient, time-saving and cost-effective way of sharing information.
- Anonymous FTP enables anyone to log on and interact with a remote server without the need for an individual account.
- FTP transfers can be done automatically by an external program.
- Only a few commands need be learned to transfer information.

Apply It!

Now that you've been introduced to software that is specifically designed to transfer files across the Internet, let's look at easier ways to do it. Netscape Navigator, like most Web browsers, lets you download files via anonymous FTP transfers. It's quite simple, and you don't need any other helper applications. First, you must tell Netscape to open a session

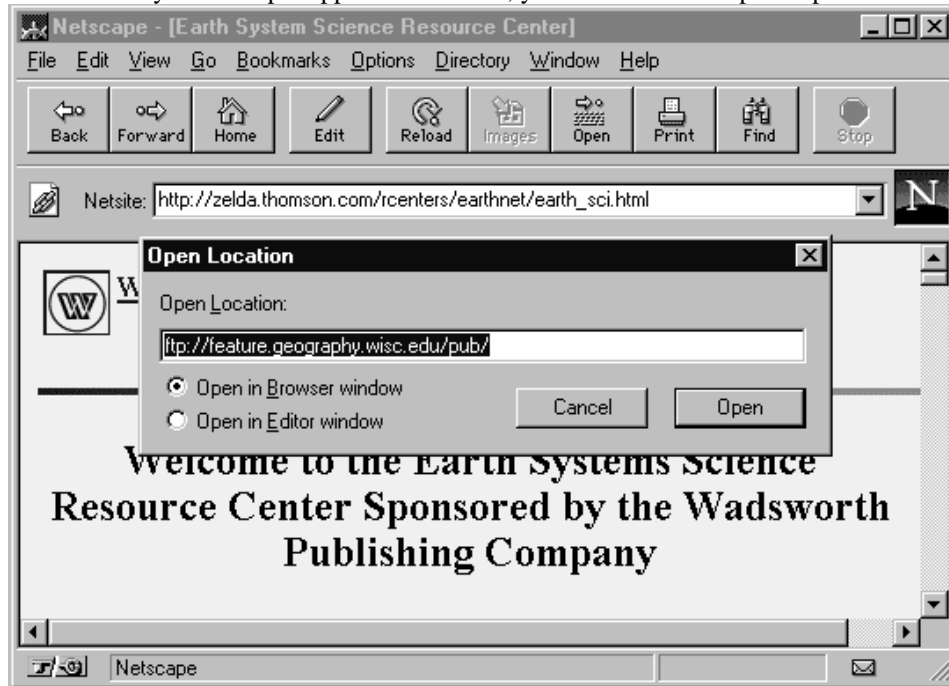


Figure 4.2 Connecting to an FTP site in Netscape

with an FTP site. You do this the same way you open a connection to a World Wide Web site. Go to the "File" menu and choose "Open Location." When the Open Location window appears, type the address of the site into the field and click the "Open" button. You can use the "Open" button on the toolbar to bring up the same window, or simply type the address in the URL Location field and hit the "Enter" key (Figure 4.2).

FTP archives are good places to look for data and graphics. I happened to find out that a graph of carbon dioxide concentrations is available at **feature.geography.wisc.edu in the pub/phys/** subdirectory. I'll open my FTP session by typing the URL into the Open Location window. Recall that the Internet address, **feature.geography.wisc.edu**, is preceded by **ftp://** to tell the Netscape browser that I'll be conducting an FTP session and not a Web session (e.g., **http://**). After clicking the "Open" button, communication is established and the FTP site appears on screen. Figure 4.3 shows the directory structure for

the FTP site. Netscape displays icons to represent the contents of a particular item on the list. Each entry on this page has a folder icon beside it indicating that they are all directories (Figure 4.3). I'll choose the "pub/" subdirectory (Figure 4.4); most pub subdirectories are used for general public access. The subdirectory list includes page icons indicating that an item is a text file. The file I'm interested in is the "phys" subdirectory, so I'll click on it to retrieve its directory listing.

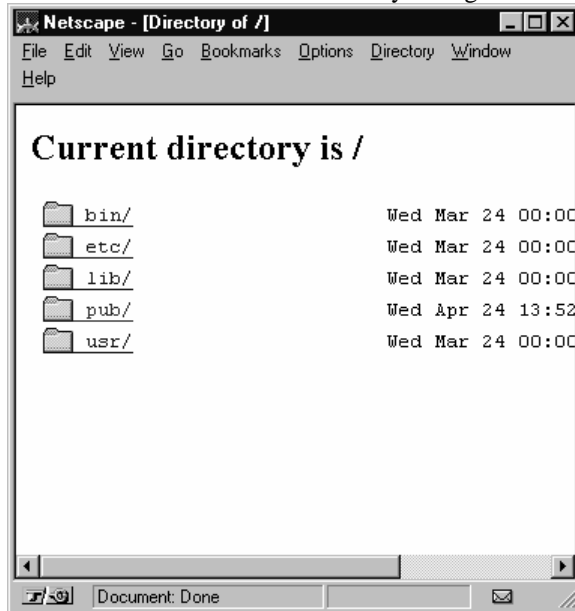


Figure 4.3 Directory structure of *feature.geography.wisc.edu*

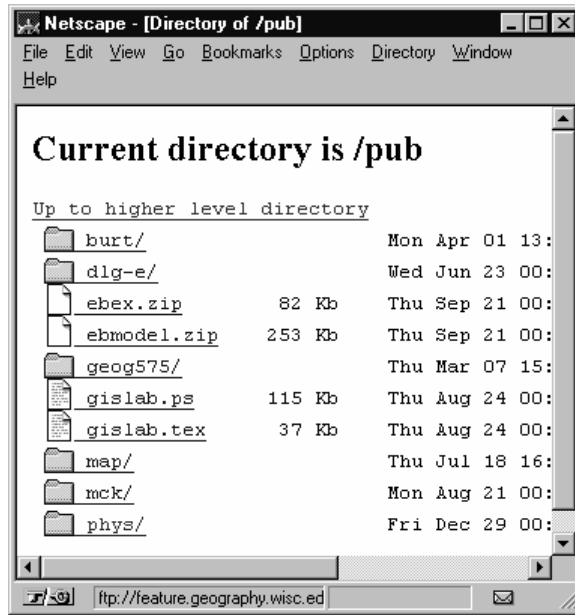


Figure 4.4 Directory structure of feature.geography.wisc.edu pub

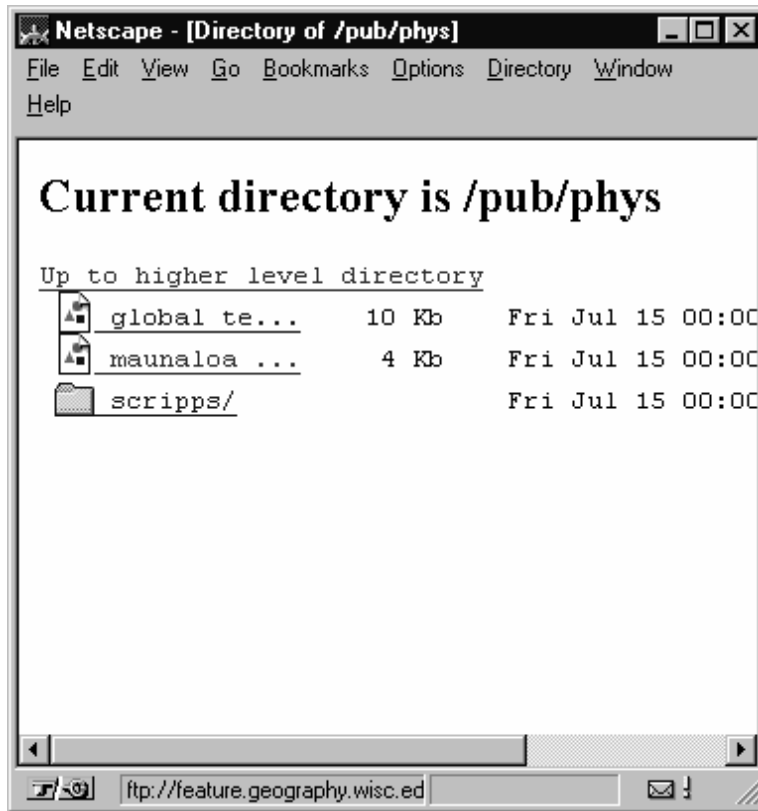


Figure 4.5 Directory structure of /pub/phys

Finally, I'm at the subdirectory I'm interested in (Figure 4.5). The graphic I've been looking for is the maunaloa image file. I'll click on the hyperlink to examine the image on screen first (Figure 4.6). It looks just like what I want so I'll download it by choosing "Save as" from the "File" menu. Netscape responds with a "Save as" window to confirm the file name and location for the file to be saved to. Clicking on the "OK" downloads and saves the file to my desktop computer. It's ready to load into an HTML or word processing document. Because it's in a digital format, I can load it into a graphics program to change the picture's attributes and add text or additional graphics. An alternative way to save this image is to position the cursor over the image, click the right mouse button and choose the "Save image as" option (Figure 4.6).

Try It Out!

1. FTP to the University of Illinois's anonymous FTP server and download a file. This is a busy server so expect some delays. Download the "sources.zip" file if you're interested

in

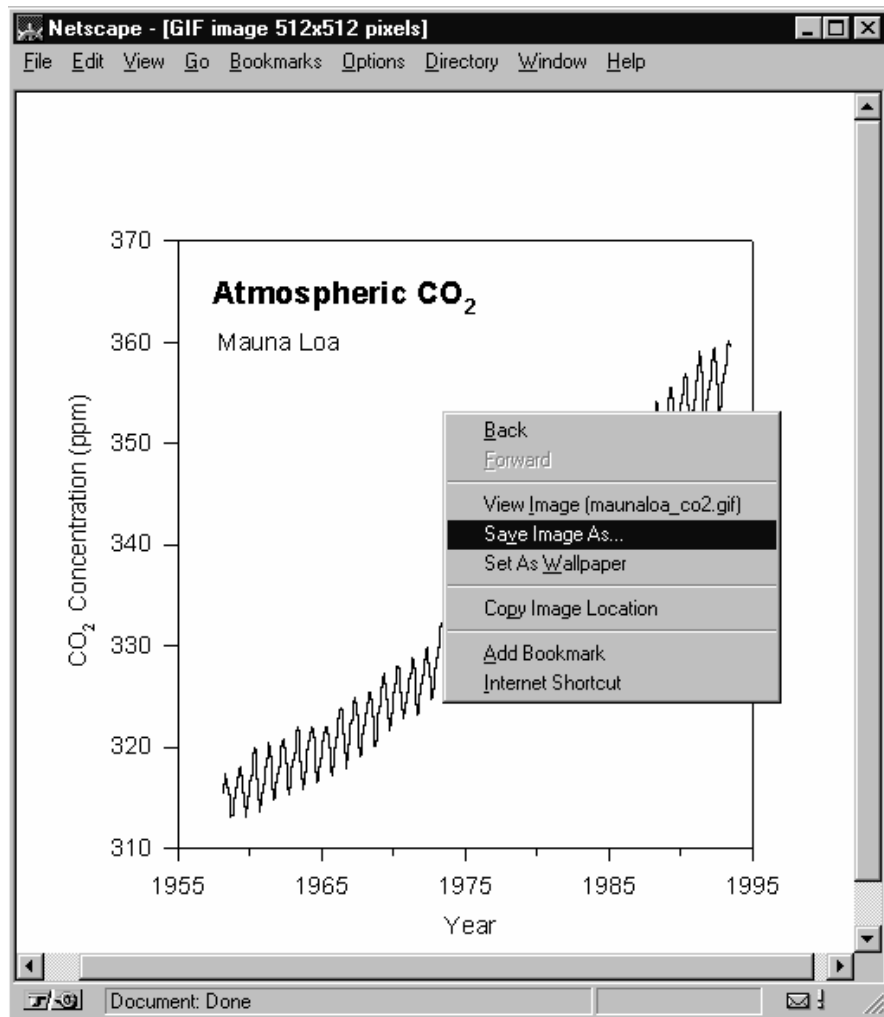


Figure 4.6 Saving the Mauna Loa CO₂ graphic

meteorology and/or climatology. You can follow the example FTP session and substitute your <filenames> for the ones used in it.

2. Try transferring the same file via electronic mail. Use the example for transferring files by electronic mail in the discussion above as a guide.
3. Massive shareware archives are connected to the Internet for people to download test programs. Connect to **URL - ftp://oak.oakland.edu**. Now follow the subdirectories by

changing directories (**cd**).

```
cd SimTel  
cd msdos  
cd astronomy
```

to such programs as

```
astrmt30.zip Astrometric data reduction for CCD images  
nalm1995.zip Nautical Almanac Data (hourly) for 1995  
sv115. SVGA vector graph. planetary, ephemer. & anim.
```

For geology software follow

```
cd SimTel  
cd msdos  
cd geology
```

to such programs as

```
acpl0021.zip Package for analysis & ctrl of pollution (7parts)  
np940107.zip NewPet: A geochemical data plotting system
```

And for mapping information follow:

```
cd SimTel  
cd msdos  
cd mapping
```

to programs like

```
topov300.zip 3-D topographic map plotting program
```

There are plenty of programs for other computer platforms too. Another good site to take a look at is **URL - <ftp://sunsite.unc.edu>**. Check out Chapter 10, "Internet Resources for the Earth Sciences," for other FTP sites.