

Configuration Options

[List Files You Already Have](#)
[Prompt on Overwrite](#)
[Leave Disk-space Free](#)
[Delete Partial Transfers](#)
[Retry Connections](#)
[Time-out in Seconds](#)
[Show File Type Icons](#)
[Sort Files Found List](#)
[Sounds](#)
[Netscape Bookmarks File](#)
[Log File Options](#)
[Make Links Local](#)
[Scan Directories Not Directly Linked](#)
[Auto Number "index.html" Files](#)
[Include HTML Bookmark Links](#)

List Files You Already Have

This check box determines whether or not files that already exist are displayed in the "Files Found" list. This takes into account building directories and remote to local filename conversions.

Prompt on Overwrite

Determines whether or not you are prompted to overwrite files on "Save Files". "Ignore Errors" will also ignore "overwrite" prompts.

Leave Disk-space Free

This determines if and how much disk space is saved if you are doing a large transfer.

Delete Partial Transfers

This determines whether or not files that are not 100% transferred are saved on the disk. Unsuccessfully transferred files are usually useless unless they are large text files.

Retry Connections

Determines if and how many times inSITE should try to re-initiate connections after failures and time-outs. You might want to set this value very high for some busy FTP sites.

Time-out in Seconds

How soon should inSITE give up on a transfer. Note that since FTP sites take a long time to log-in you should probably keep this value around 60 seconds.

Sort Files Found List

This determines whether or not inSITE keeps the “Files Found” list sorted all the time.

Show File Type Icons

This determines whether or not inSITE should assign icons based on the file extensions. This may slow some things down slightly.

Sounds

If you get tired of the sounds delete the ones you don't like or simply replace them with wav files you do like. To enable/disable sounds, see the configuration menu. The ".wav" files are located in the "\Wavs" sub-directory.

Netscape Bookmarks File

This should be the location of your Netscape bookmarks file. This is important for the "Starting URLs" pop-up menu options.

Log File Options

If you wish to have a log of transfers inSITE will create a log file. The server name for this URL:

`http://www.somewhere.com/asdfsdf/asdf.txt`

would generate:

`www_somewhere_com.log`

The "Starting URL" is where the server comes from.

The specified file is in the text box below the "Log File Options".

Make Links Local

This attempts to convert all links in any HTML files saved to local references. For example:

`http://www.somewhere.com/frednet/something/`

might become:

`file:///c:/junk/frednet/something/`

Scan Directories Not Directly Linked

**If this is checked sub-directories not necessarily directly linked will be scanned.
For example if this link was found:**

`ftp://ftp.fred.com/pub/junk/misc/something.txt`

Then not only that page but also:

`ftp://ftp.fred.com/pub/junk/misc/`

`ftp://ftp.fred.com/pub/junk/`

`ftp://ftp.fred.com/pub/`

`ftp://ftp.fred.com/`

would be scanned for files.

Auto Number "index.html" Files

When you save a file such as "ftp://ftp.fred.com/pub/junk/" you are saving a directory listing in HTML format. It does not have a specific name so "index.html" is assigned to it. If an "index.html" file already exists then a number is appended to the file name for example:

"index.html" might become "index1.html"

This is useful when downloading pages from multiple servers into the same directory.

Include HTML Bookmark Links

This refers to links which have a “#” sign in them. You usually don't want to download multiple copies of a file so the bookmark information is chopped off. They can be useful for local browsing sometimes as the inSITE file list lets you launch them directly into Netscape.

The Files Found Tab

[File Selection](#)

[Filtering Files](#)

[The File Mask](#)

[Show Full Paths](#)

[Ignore Errors](#)

[Build Directories](#)

[Get Sizes](#)

[Selected Information](#)

[Save Files](#)

[The Files Found List](#)

File Selection

The "Select" button goes through the "Files List" and selects all files which match the "File Mask" below it.

Filtering Files

The "Filter" button goes through the "Files List" and temporarily deletes any files which do NOT match the "File Mask" below.

The File Mask

The "File Mask" is used by the "Select" and "Filter" buttons. The mask can use "?" or "*" marks.

- The "?" signifies any one character.
- The "*" represents any number of any characters, including possibly no characters.

You can separate masks using a semi-colon:

fred*.gif;ted*.jpg;cool??.mpg

You can use (optional) parentheses to write some comment text:

Something interesting (file*.txt;cool*.txt)

So in sum:

fred?*.gif

Would match:

fred1.gif

fred1asdfasdfasdfasdf.gif

fred5asdfasdfasdfasdf.gif

fredasdfasdfasdfasdf.gif

But not:

fred.gif

That is because the "?" demands a character between "fred" and ".gif".

Edit the "\\windows\inSITE.ini" file to add (or change) standard unchanging file masks.

Show Full Paths

This determines whether or not the full URL is displayed for files which can be shortened to simple file names. For example:

`http://www.somewhere.com/something/theFile.txt`

becomes:

`theFile.txt`

If the URL ends in a "/" then obviously there is no short file name to use and thus the whole URL is still used. The short file name is also the one the file is downloaded to. If no file name is present "index.html" is used for the download name.

Ignore Errors

This determines whether or not inSITE will pop-up a dialog box asking for your response to an error message. Some errors will stop the transfer and wait for your response. On long transfers it is wise to check this box.

Build Directories

This determines whether or not sub-directories are built as demanded by the URL. The server name is not used and the starting directory is the one you have listed in "Save To".

For example:

`http://www.somewhere.com/something/is/here/really/farOut.txt`

would build the directories:

`\something\is\here\really`

on top of whatever directory you listed in "Save To". If you listed "c:\junk" then it would become:

`c:\junk\something\is\here\really\farOut.txt`

when you pressed "Save Files".

Get Sizes

This attempts to connect to each site and get the size of the select URLs. The size is in kilobytes (KB). Sometimes you may want to do this more than once to verify the correctness. Also some sites will not return a size.

Note: "RightClick" on the list of "Files Found" to change to "Report Style". This is the only view that displays the sizes. Otherwise you can see the sizes on the Hint/Satus Bar.

Selected Information

This simply lists the "number of files selected" versus "total number of files".

Save Files

This will initiate the sequential downloading of all the selected files. A "Red Light" indicates failure to download the file. An "Orange Light" indicates that part of the file was transferred. A "Green Light" means the files was successfully saved. The files are saved to the "Save To" directory or sub-directories of "Save To" based on whether or not "Build Directories" is checked.

The Files Found List

This list displays all the files found on the links scanned. There are a number of options not immediately visible. "RightClick" on the list to start the pop-up menu. From there you can:

- Choose the style of display
- Clear the list of "Files Found"
- Sort the list of files alphabetically
- Regenerate the list, this will rebuild a list after the effects of "Filter"ing
- Open the URL in Netscape
- Toggle the selected files to the opposite selection state
- Or remove files based on various factors
- Press the "Delete" key to delete all select files

To see the sizes found you must be in the "Report Style" of display. "RightClick" to select the style.

Note: for options that manipulate the selected state it is important to right click on a selected item to maintain the selected state before the operation.



Contents

[Copyright and Business Information](#)
[Ordering inSITE](#)

[General Usage](#)
[Special Keys](#)

[The Links Tab](#)
[The Files Found Tab](#)
[The Save To Tab](#)
[Menu Options](#)
[Configuration Options](#)

Copyright and Business Information

"inSITE: Into the Internet" and all contents etc. are ©Copyright Alexander Schonfeld, 1996. All rights reserved.

Contact Information:

Alexander Schonfeld
5615 SW 42nd Ave.
Portland, OR 97221

schonfal@engr.orst.edu

<http://www.engr.orst.edu/~schonfal/>

Ordering inSITE

See the "Order Form" option from the "File" menu, or the "Order Form" section of the manual.

Special Keys

Selection keys:

- Shift + Arrow or Mouse Click: This will select multiple items at once.
- Ctrl + Mouse Click: Will select one additional item.
- Left Click + Drag: Will draw a box to select more than one item

Other keys:

- Ctrl+Tab: Will cycle through the tabs
- Alt+F4: Exit

Note: many windows support more keys or options than shown, so read the hint bar.

General Usage

Overall tip: inSITE is a complement to Netscape, not meant to replace it. Use the Netscape bookmark system. It is much nicer.

Overall tip: Read the hint bar!

Quick Start:

When you find a site you are interested in through a news article or web page etc. you then want to explore it. Netscape is the best way to do this. When you find that the site has many files you are interested in, inSITE becomes useful. Run inSITE and "RightClick" on the "Starting URL" box. Select "Copy from Netscape" and the URL will be pasted into the box. Then configure the other options on the "Links" page. Once configured you press the "Scan For Files" button and inSITE will begin scanning the pages recursively. This may take some time on a dial-up connection. Setting the recursion depth about 3 or 4 will no doubt turn up hundreds or thousands of pages if not narrowed by some of the limiting factors. Using the "Scan URL Mask" to stay on one server is one way to make sure your search isn't too broad. Once you have scanned the pages the files found will appear in the "Files Found" list. From there you can select which ones you want to download. There are a number of techniques you can use to remove the ones you don't want. Before saving the file,s it is probably wise to check where you are putting them. See the "Save To" tab for this purpose.

Once you have downloaded the files you can "DoubleClick" on them in the "Your Files" section of the "Save To" tab. This will request that Netscape view them. By configuring Netscape you can view almost any type of file. Sound files, Zipped archives, movies and many other types can be viewed with Netscape.

The concept of a "URL":

URL stands for Universal Resource Locator. A URL is the principal means of connecting data on the World Wide Web. A web page is made up of many URLs. An example URL is:

`http://www.somewhere.com/adirectory/afile.txt`

This can be broken down into:

`protocol://[username:[password]][@]host[:port]/[directorypath]/[filename]`

In the above example:

protocol = http

there is no username, password, or port specified

host = www.somewhere.com

directorypath = adirectory/

filename = afile.txt

The protocols generally available on the web are:

ftp = File Transfer Protocol, used to access so-called FTP sites, for example:

`ftp://ftp.cdrom.com/pub/demos/`

file = used to access local files

`file:///c:/windows/`

gopher = a simpler version of the WWW with mostly text and single graphics

`gopher://something.com/somewhere/`

inSITE will allow you to use all the above protocols in your searches.

Understanding how it's all connected:

Although you don't necessarily see Netscape involved in inSITE's operation, Netscape provides inSITE's interface to the Internet. All the above protocols are returned as a web page. This web page has links imbedded in it. I'll repeat, all the above protocols return web pages that Netscape generates. These pages are in HTML (Hypertext Markup Language). HTML documents have links imbedded in them at various points. The FTP site directory listing is actually an HTML document with links to all the files on the FTP site. InSITE breaks these links down into two groups: HTML documents and other files.

The HTML files and the other files:

When you select "Scan For Files" inSITE goes scans the "Starting URL" for links (which are URLs pointing to HTML documents) and files (which are not HTML documents). On the "Links" tab you will see the links to HTML documents found in the most recent "Scan For Files". If you have the recursion depth set to 1 you will only see the links found in the "Starting URL".

The "Scan URL Mask" is what determines which URLs are actually scanned. This mask may use the standard "*" and "?" characters to specify broad ranges of URLs. If you want all links scanned you could simply enter "*". Usually you want to narrow the scope to a single server and thus something like "*www.something.com*" would be appropriate.

Note, links are added to the Links Tree that do not match the mask. These links are simply there for reference. NOTHING is scanned without matching the "Scan URL Mask".

When links to files that are not HTML files are found, they are normally added to the "Files Found" list. Note that the "Scan URL Mask" has nothing to do with what files are added unless you check the "Apply Mask to Files" box. This is useful because quite commonly Web pages will have links to files not on their server imbedded in pages that are on their server. To narrow your scope to exactly one server and nothing else you would need to check the "Apply Mask to Files" box as well. Note, this still doesn't change the fact non-scanned links are added to the Links Tree. You can tell if a link has been scanned by the stop light icon next to it. If it is green it has been successfully scanned and will not be scanned again. The Links Tree is primarily a navigation tool, and as such it is cleared every time you scan a new page by DoubleClicking on the tree or pressing the "Scan for Files" button.

The Links Tab

[The Starting URL](#)

[The Scan URL Mask](#)

[Recurse Depth](#)

[Include HTML Files](#)

[Include Inline Images](#)

[Apply Mask to Files](#)

[Scan For Files](#)

[The Links Tree](#)

The Scan URL Mask

This mask lets you filter out certain URLs. For example if you only want to scan pages on one server you could specify:

`*www.someserver.com*`

or two servers and only ftp sites:

`ftp://*;*www.one.com*;*www.two.com*`

The "?" matches any one character, but always one character. The "*" can match any number of any characters. This is not case sensitive, so be careful.

Apply Mask to Files

This determines whether or not the URLs added to the "Files Found" list also have to match the "Scan URLs Mask". Sometimes pages link to files not on their server so if you only want files on a specific server you should check this box:

www.someserver.com

would match:

<http://www.someserver.com/~fred/something.gif>

The Starting URL

The "Starting URL" is the first URL scanned when you select "Scan For Files". The links to files and other HTML pages are the starting point of the search.

There are a number of options not immediately visible. "RightClick" on the text box to start the pop-up menu. From there you can:

- Copy URL from Netscape: This option simply copies the URL currently in Netscape to the Starting URL text box.
- Open URL in Netscape: This option loads the specified "Starting URL" in Netscape
- Add to Bookmarks: This option adds the current Starting URL to the Netscape bookmarks file as defined in the Configuration menu.
- Select from Netscape Bookmarks: This option pops-up a listing of the URLs in the Netscape bookmarks file.
- Select from Session List: This option pops-up a list of all pages found in the most recent scan for files.

Edit the "\\windows\\inSITE.ini" file to add (or change) standard unchanging URLs to the history.

Recurse Depth

Imagine a directory structure...

ftp://www.here.com/windows/system/junk

If you are viewing the

ftp://www.here.com/windows

directory and you select a recursion depth of "1", only the files in that directory will be listed. If you select "2",

ftp://www.here.com/windows

and

ftp://www.here.com/windows/system

will both be scanned for files. So will any other sub-directories of /windows. The files found in those directories would be added to the "Files Found" list.

Include HTML Files

This determines whether or not HTML files are added to your list of "Files Found".

Include Inline Images

This determines whether or not Inline Images are added to your list of "Files Found".

Scan For Files

This button will initiate the scan starting with the "Starting URL". The "Include" and "Apply Mask" options effect what files are added to the list of "Files Found". You can view all the files found on the next tab. For speed press "Ctrl+Tab" to move to the next tab.

The Links Tree

This tree attempts to represent in an intuitive fashion the hierarchical nature of the Internet. Links that have been successfully scanned will have a "Green Light" next to them. Partially scanned links will have an "Orange Light" and failed links will have a "Red Light". No light means that inSITE has not tried to scan that link. Thus no files will be added to the "Files Found" listing from those unscanned links. No successfully scanned link will be scanned twice.

There are a number of options not immediately visible. You can:

- "DoubleClick" on a link to start scanning that link
- "RightClick" on the list to sort it alphabetically
- Press the "Delete" key to delete a node and all sub-nodes

Menu Options

[Append List of URLs](#)

[Save List of URLs](#)

[Quick Launch Netscape](#)

[Show Error Window](#)

[Set Username and Password](#)

Append List of URLs

This menu option lets you load a saved list of URLs and append them to the current list of "Files Found". The format of the file is (*.lks):

*http://www.something.com/~fred/,16KB

http://www.something.com/,

Where the "*" specifies whether or not the URL is selected. The file is a plain text file editable with "edit". The string after the first comma is the size of the file if it is known ("16KB" in the first example, unknown in the second).

Save List of URLs

This menu option saves the current list of "Files Found" to a specified ".lks" file. This file is a text list of the URLs. See "Append List of URLs" for the format.

Quick Launch Netscape

This function quickly opens a Netscape window in compressed title-bar form.

Show Error Window

This checkable menu option determines whether or not the "Error Window" is open.

Set Username and Password

This dialog box attempts to set the Netscape username and password so that you can log into protected sites. Note: this only works on some sites.

Save To

This tab provides you with a directory selector. The currently opened directory is the one saved to. From the directory window you can:

- Press the "Delete" key to delete an empty unopened directory

The file listing at the right represents all the files in the current directory which match the file above file mask. You can:

- Press the "Delete" key to delete all selected files from the directory
- "DoubleClick" to open the local file in Netscape

The "New Folder" button will create a new sub-directory in the current directory. The name is specified in the text box at right of the button.

