

Guide on using Xenu's Link Sleuth

Table of contents

- [Xenu's Brief Description](#)
- [Using Xenu to Check a Web Site](#)
- [Results](#)
- [Saving the Results](#)
- [Checking Multiple Sites](#)
- [Checking URL List](#)
- [Options](#)
- [Frequently Asked Questions](#)
- [Appendix 1: Effects on Hit Counters \(and other things to keep in mind\)](#)
- [Appendix 2: Future feature List](#)
- [Appendix 3: Bug List](#)
- [Appendix 4: Story of Xenu's Link Sleuth](#)
- [Appendix 5: Keyboard shortcuts](#)
- [Appendix 6: Credits](#)

Trademarks

Xenu, *Xenu's Link Sleuth* and *Link Sleuth* are trademarks used by Tilman Hausherr for software products and services. These products are not associated in any way with services licensed by RTC, CoST, BPI, CSI, etc.

Xenu's Brief Description

Xenu's Link Sleuth™ is a freeware program that is available for download at <http://home.snafu.de/tilman/xenulink.html>.

Xenu's spidering software checks Web sites for broken links. Link verification is done on "normal" links, images, frames, plug-ins, backgrounds, local image maps, style sheets, scripts and java applets. It displays a continuously updated list of URLs which you can sort by different criteria. A report can be produced at any time.

Additional features:

- Simple, no-frills user-interface
- Can re-check broken links (useful for temporary network errors)
- Simple report format, can also be e-mailed
- Executable file less than 500K
- Supports SSL websites ("https://")
- Partial testing of ftp and gopher sites
- Detects and reports redirected URLs
- Site Map

It is available for Microsoft Windows 95/98/ME/NT/2000/XP/Vista, but other operating systems are not supported.

Using Xenu to Check a Web Site

Doubleclick the Program icon to open Xenu's Link Sleuth, or type the path and the xenu.exe filename in the "Run" command found under the Windows "Start" button.

A Xenu window will appear on your desktop. You can now use Xenu's Link Sleuth to check for broken links.

1. Under the **File** menu, select **Check URL**. Or, click on the white 'page' icon located on the left end of the toolbar.
2. A dialogue box called "Xenu's starting point" will appear.
3. Enter the URL of the page you would like to check links on in the first box. You should enter the URL in the format **http://www.server.edu/~username/** (enter the correct URL and server information). If you'd like Xenu to check an entire website, you should enter the main or index page of the site. Xenu can check any website, so it will work on www.microsoft.com and so on.
4. The **Browse** button will allow you to check a local HTML file by browsing to its location on your computer. You can check links before uploading your pages to the web server by using this function.
5. The **Check external links** checkbox is selected by default; if you don't want Xenu to check external links (links to sites outside of your site) then deselect this box.
6. You can set certain URLs as 'internal' in the next box.
7. You can exclude URLs that you don't want checked in the next box.
8. Click **OK** button to begin checking.

Once you've checked a website, a small **DOC** icon will appear to the left of the **File** menu, and the main window will show up the [report](#), containing the list and color-coded status of all the links checked by Xenu.

Results

After you enter the URL for Xenu to check and click the **OK** button in the "Xenu's starting point" dialogue box, a rapid listing of the links being checked will appear within the Xenu main window.

Xenu shows good links in **green**, broken links in **red**, excluded types of links in **dark blue**, and links yet to be checked in **grey**. When the program has finished checking the links, it will inform you that "Link sleuth finished" and ask you if you want a report.

HTML Report

Click **Yes** if you'd like the HTML page report, otherwise click **No**. You can also generate the HTML report later by using **Report** command in **File** menu.

Xenu's report format is an HTML page that will open in a default browser window to show you the root URL that was checked and to provide the following summary for that URL (depending on enabled [options](#)):

- Broken links, ordered by link
- Broken links, ordered by page
- Broken local links
- List of redirected URLs (URLs that have moved)
- List of ftp and gopher URLs
- List of valid URLs you can submit to a search engine
- Site Map of valid HTML pages with a title
- Orphan files (requires FTP access for remote files checking)
- Statistics for managers

Non-HTML Report in the Xenu Main Window

You can view the detailed results in the main Xenu window. The report shows the following details for each link:

- **Address** - Full URL for current link.
- **Status** - shows whether the link is **OK**, **broken**, **skipped** etc.
- **Type** - MIME Type for current link.
- **Size**
- **Title** - shows Title for retrieved web pages, link text for skipped URLs.
- **Date** - Date-Modified stamp of the file/URL.
- **Level** - 0 for start page, 1 for URLs linked from start pages, 2 for URLs linked from that pages etc., depending on the number of levels allowed in [Options](#) dialog.
- **Links out** - number of outgoing links on checked page.
- **Links in** - number of inbound links on checked page.
- **Server** - server software for internet URLs (empty for local URLs)
- **Error** - error description.
- **Duration** - the duration of retrieving process.
- **Charset** - character encoding.

You can sort the results by clicking on these headers, which is useful when you want to get a quick impression of the number and type of errors, if certain types of links have been checked, or skipped, etc.

Filter the results

Under the **View** menu you can switch to **Show broken links only**, which will help you to see any broken links more easily. (Note: if you have selected this option but do not have any broken links, nothing will be displayed in the window. Turn it back to off to see all the links that were checked.)

Locate the referring page

Right-click the broken link and select **Properties** to see what pages in your site contain the link. The URLs for the referring pages will appear in the bottom window under **Pages Linking to this one**.

Export to other formats

You can easily export Xenu report to a couple of other file formats by using the appropriate **Export to...** commands in **File** menu. The following formats are currently supported:

- Tab separated file (best to view in Excel and other spreadsheet programs)
- GraphViz file (open format for representing structural information as diagrams)
- Google Sitemap file (can be submitted to Google and other search engines to better crawl your site)

Saving the Results

You may want to save the results of Xenu's link check, particularly if there are many links that need to be updated on your site.

To save the results of the link check to a file:

1. Check a URL using Xenu.
2. Click **No** when Xenu asks if you'd like a report.
3. Under the **File** menu, select **Save As**. Or, select the small disk icon from the toolbar. Enter a filename and enter the location where you'd like to save the file on your computer's hard drive.
4. Later, you can **Open** the file to view it again (either click on the small folder icon on the toolbar, or under the **File** menu select **Open**).
5. When you're viewing the saved file at a later time, use Xenu to retry broken links. This is especially useful to find links that may have been only temporarily broken, or if you want to verify that links you have fixed are indeed working properly. Under the **File** menu select **retry broken links** to have Xenu revisit links that were broken in the initial check.

Checking Multiple Sites

It is easy to check several different sites using Xenu Link Sleuth. Once Xenu has completed checking a URL you can enter another URL in the top box of the "Xenu's starting point" dialog box.

Even if you have an [existing report](#) and have not saved it yet, Xenu will create a new report for additional sites checked. You'll still be able to view the previous results too. By default, Xenu names these various reports XENU1, XENU2, etc.

To access the different sites' (non-HTML) reports within the Xenu main window, you can either click on the **DOC** icon and select **Minimize** (the separate [results reports](#) will appear in small windows), or type Control+F6 to view the **Next** report in the full-size Xenu window.

If you have checked several sites and have several different results reports, you may wish to [save](#) them with different file names to avoid confusion.

Checking URL List

Xenu also provides you with the ability to scan a limited set of links (e.g. a subsection of your site, or a couple of related pages etc.):

1. Create a text file with the URLs you want to check, one URL per row.
2. To load this file in Xenu, use **Check URL list** command in **File** menu.
3. Xenu will check all links on submitted URLs and generate the [report](#).

When you scan links using this method, all the URLs in your start text file get level 0, the links on these pages get level 1, and so on.

Note: When running a scan for URL list, it is not possible to exclude "external" links from the scan like when checking a single URL!

Options

There are some options you can set in Xenu to get better results. These can be accessed from the "Xenu's starting point" dialogue box by clicking the **More Options** button, or from the **Options** menu on the main toolbar.

1. Click on the **Options** menu on the main toolbar.
2. Click on **Preferences**.
3. You'll see a **Parallel Threads** slider. The default setting is 30. The author of the software suggests setting this slider at about 70 to get the best results, but depending on how much memory your computer has you may need to experiment to see what setting works the best with your particular machine.
4. The default setting for Maximum Level is 999 - in other words, Xenu will go down 999 levels to check links in a site. You can change this value to a lower number if you don't want Xenu to check items that are located deep within your directory structure.
5. The Options panel also contains a series of checkboxes, all of which are selected by default. You can deselect any of the options that you don't want to use.

Frequently Asked Questions (FAQ)

1. [Who is Xenu?](#)
2. [Is Xenu's Link Sleuth™ better than WebAnalyzer?](#)
3. [Is Xenu's Link Sleuth™ better than Net Mechanic?](#)
4. [Can I support the author financially?](#)
5. [Why does Xenu's Link Sleuth™ report `http://www.site.com/./page/index.html` as broken?](#)
6. [How can I configure a proxy?](#)
7. [Why does Xenu's Link Sleuth™ report an URL with a space in it?](#)
8. [I use Mozilla 3.0 Gold and can't get rid of file: URLs for images. What can I do?](#)
9. [What is the maximum number of websites that can be checked?](#)
10. [Can the software check my site locally?](#)
11. [Does it work on Windows NT 3.51?](#)
12. [How is it so damn fast?](#)
13. [Can I have the source code?](#)
14. [Can I buy the source code?](#)
15. [Just for fun, I checked Tilman's web site, and found many broken links. Why?](#)
16. [How do I correct broken links?](#)
17. [What about ftp and gopher sites?](#)
18. [Why can't I launch URLs?](#)
 1. [Why does the browser not open a new window?](#)
 2. [Why does Link Sleuth freeze when launching the report?](#)
 3. [Why does Link Sleuth freeze when launching the report or a URL?](#)
19. [What about cookies?](#)
20. [Why are some links reported as "broken" by Xenu, that can be displayed within my browser?](#)
21. [Why can't I connect to "secure" \(https\) sites?](#)
22. [Any known problems with Windows 95?](#)
23. [Any known problems with Windows 2000?](#)
24. [Why can't I configure the timeout?](#)

25. [What about JavaScript?](#)
26. [What about passwords entered in a FORM?](#)
27. [How about a WAP version?](#)
28. [What about these error codes?](#)
29. [Why do I get broken links with filelist.xml, editdata.mso and oledata.mso?](#)
30. [Why do I get "file not found" on remote checks?](#)
31. [Can I make a foreign language version?](#)
32. [Why isn't Xenu detecting missing URLs?](#)
33. [Running Xenu with Norton Internet Security](#)
34. [Why timeouts?](#)
35. [Any Spyware, Adware, Malware?](#)

1. Who is Xenu?

See [here](#).

Do you want to be a **Knight of Xenu**? Then [join that team](#) in the worldwide [RC5-64 key decryption effort](#), which uses "idle time" on thousands of computers all around the world. Here is how to do it:

1. Download the [client](#).
2. Configure the client:
 1. Chose "1" ("General Client Options"), then choose "1" ("Your e-mail address") and enter your e-mail address (without "<" or "("), then go back to the main menu with "0".
 2. Chose "2" ("Buffer and Buffer Update Options"), then chose "9" ("Load-work precedence"), then press the backspace key to delete everything, and enter "RC5,DES=0,CSC=0,OGR=0".
 3. If you are behind a firewall, or use a dial-up line, you will have to chose "6" ("Keyserver <=> Client connectivity options") to configure the use of a proxy, or to watch for an dial-up active connection. Go back to the main menu by entering "0" twice.
 4. Enter "0" to save the settings and exit.
3. Start the client and watch him work.

After one day or two, you can go the [main statistics page](#) and enter your e-mail address. You can then access your very own statistics page. Scroll down and click "Please e-mail me my password". You will receive an e-mail with your ID and your password. Now [click here to join team #3504](#) and enter ID and password when requested.

Please do not participate in this project if you are on a corporate machine without first getting permission from your supervisor and the security people.

[E-mail me](#) if you have any problems with RC5-64. But read the description above first.

2. Is Xenu's Link Sleuth™ better than WebAnalyzer?

Yes and No. Xenu's Link Sleuth™ does not have the graphic capabilities of WebAnalyzer 2.0 ("Wavefront view"). But here are some of the advantages of Xenu's Link Sleuth™:

- It is free
- Simple user-interface
- Better error reports (not just "network error")
- "Save" works also while the software is busy
- The "broken links view" shows only broken links; In WebAnalyzer you'd have to press the button again and again as the window fills with crap.
- While Xenu does not offer an "update" facility (which doesn't work anyway), it has a "recheck broken links" function that works fine.
- It is small, written by one person with 5 years experience of Windows development and 15 years of professional experience as software developer. This means that bugs will be corrected quickly. This is a matter of honour.
- The report can be viewed easily, even when you have long URLs.
- Uses much less disk space for intermediate files, executable file much smaller
- Loading of saved files much faster (WebAnalyzer loses time by displaying the extra graphics)
- Supports SSL websites ("https:// ")
- Partial testing of ftp and gopher sites
- Search for local orphan files
- Special handling of redirected URLs
- Site Map
- Randomization of checking order, means less concurrent requests on a single server

Xenu sez: check your website both with this product and with another product ([Linkbot](#), [InfoLink](#), [LinkScan](#), [LinkAlarm](#) and [Web Link Validator](#) offer trial versions - WebAnalyzer is no longer available since February

2002 and hasn't been updated for years), and decide what you need and what you are willing to pay.

3. Is Xenu's Link Sleuth™ better than [Net Mechanic](#)?

Years ago, Net Mechanic was a free WWW based service, and was useful to check very small web sites. It is no longer free. The free trial is too small, and reports about all links, instead just the broken ones.

4. Can I support the author financially?

No need to. If you feel the software is useful, you may donate money to causes I support.

- [International Cultic Studies Association \(ICSA\)](#) is a nonprofit, [tax-exempt](#) research center and educational organization founded in 1979. ICSA's mission is to study psychological manipulation and cultic groups, to educate the public and professionals, and to assist those who have been adversely affected by a cult-related experience. I suggest a donation of \$20 for individuals and \$200 for corporations. In the US, your donation can be deducted from your income. (ICSA does not endorse this site in any way, did not develop this software, does not sell this software, and the use of this software does not depend whether or not you make a donation.)

Germans can make a tax deductible donation to the [Dialog Zentrum Berlin e.V.](#), Konto-Nr. 1551390051, Bank für Kirche und Diakonie BLZ 35060190.

Or visit the [Xenu bookstore](#).

Or send me a T-Shirt of your city, university, employer in XL size. Please don't send anything that is more expensive than \$40 (including shipping). Take into consideration that I'll be wearing your T-Shirt at work. USPS "airmail letter post" is fast, reliable and unexpensive (and large sizes are allowed!), so please don't use FedEx or UPS, because this could result in me having to pay money for customs.

Or send me a "thank you" letter on company paper, if you work for a well-known company. Make sure that you are authorized to send such a letter. This is my street address:

Tilman Hausherr

Hauptstrasse 15
10827 Berlin
Germany

5. Why does Xenu's Link Sleuth™ report <http://www.site.com/./page/index.html> as broken?

The key is the `./` part. It means you have e.g. a top level page that links to a page in a directory above, which doesn't exist. It is true that Mozilla will not have any problems with such a page; but I am less tolerant.

6.How can I configure a proxy?

You can configure a proxy in the control application of Windows. Double-Click on the "internet" symbol, then click on the "card" of the dialog box that is named "Connection". You may need a proxy if you are sitting "behind a firewall". This is usually so in big corporate networks.

One user with Windows 2000 always had a timeout, he solved it by checking "Use HTTP 1.1" and also "Use HTTP 1.1 through proxy connections" in the "Advanced" tab of the Internet Options in the control panel. However, this may not work for everyone, because some web servers do not support HTTP 1.1.

7. Why does Xenu's Link Sleuth™ report an URL with a space in it?

Either because you do have a space in the URL, or because you have a carriage return / newline in it. Although Mozilla tolerates this, I do not.

8. I use Mozilla 3.0 Gold and can't get rid of `file:` URLs for images. What can I do?

Re-edit the page, double-click on the picture, remove `file:` from the picture location and take care to uncheck "copy image to document's location" in the "properties" dialog box (at the bottom left) before you save and exit the dialog box.

9. What is the maximum number of websites that can be checked?

There is no maximum. It is limited by the memory on your computer.

10. Can the software check my site locally?

Since september 1998 (1.0n), you can do so without a local web server (your address would then be <http://127.0.0.1>). Use the "Browse" button in the "New" dialog box.

The results will not always be the same as a "remote" check:

- Sometimes you'll get "error 3". It happens because the WININET.DLL is unable to handle directories, i.e. links that end with "/". You can avoid this by linking to the actual "main file", usually `index.html` or `default.html`. That your browser can handle local directories and display them nicely, is because he does *additional* work, which I do not.
- Mixups of higher/lower case characters in links won't be found, since Windows does not make a difference. But UNIX does!
- The main reason that you still need to make occasional "remote" checks is because you might have forgotten to upload your files to your WWW server.

A user of IE 4.0 reported that when not online, the software checks every "remote" URL like a local file. This is a problem of the newer version of the WININET.DLL; the version with IE 3.0 reports "no connection" or "no such host" instead, which is more logical.

11. Does it work on Windows NT 3.51?

One user said it worked fine after he copied a version of WININET.DLL from a Windows 95 system standing nearby, and put it into the directory where Xenu's Link Sleuth™ was installed.

12. How is it so damn fast?

Because it uses a ([possibly patented](#), see patents [here](#) and [here](#)) technique known as **preemptive multithreading**. It means that the link checking software retrieves several web pages at the same time; the competition uses the same technique. The maximum count of threads is initially set to 30, but you can configure it to any number between 1 and 100. A number that is too high might result in failed connections or in timeouts, which means you will have to recheck the broken links. At the time I had a dial-up connection, I got good results with 70. Now I have a DSL connection, and I have to set the number to 1-5. I suspect that my DSL provider has installed a brake somewhere to prevent "commercial" customers from using the unexpensive "private" service.

13. Can I have the source code?

Hahahahahaha!

14. Can I buy the source code?

Sure, make me "an offer I can't refuse".

15. Just for fun, I checked Tilman's web site, and found many broken links. Why?

I check my own web site every week on friday. Nevertheless there are always broken links:

- Links that I know to be broken: I keep them like that to remind me to find these people some day. The web page itself has a notice that the link is broken.
- Temporary unreachable hosts: these are temporary routing errors.
- Really broken links: I will usually correct the link or remove it within the next few days.

16. How do I correct broken links?

Repairing broken links (i.e. getting the correct ones) is a difficult task that takes time, but with experience, you'll get it done faster and faster.

- if you have the e-mail address of the site owner (because you know him), try an e-mail. Sometimes the address still works, even if the web site is gone.
- find the home page of the site you link to, to see if the site has a "sorry we moved" message. If you linked to `http://www.host.com/~user/page888.html` and this is broken, look at `http://www.host.com/~user/` to see if there is a message, or to see if the site has been reorganized. Some sites reorganize their user pages differently, e.g. `http://www.host.com/homepages/users/page888.html`. Sometimes the web switches changes between the two methods. Other sites are owned by the user himself, e.g. `www.user.com`, so the home page is the root page. If the site exists but you cannot find your page, send an e-mail to the owner.
- use search engines to find the site or the name of the site owner (if you know). To find where the site is, use web search engines (like [Google](#) or the [Internet Archive](#)) and usenet search engines (like [Google Groups](#)).
 - You find the site you searched for
 - You find a site that links to the site you searched for
 - You find the site in the Google Cache or the Internet Archive (simply enter the URL in the search box!), and can use the contents to search for the name of the owner
 - You find a site that links to the site you searched for, but is also broken. E-mail the site owner, and tell him that the link is broken. Bookmark the site and revisit it in a week, to see if the other person has found it. If not, you have nevertheless succeeded in making the other person feel as bad as you, which brings some relief :-)
 - You find the new e-mail address of the user. Either e-mail him,

or try to construct the URL yourself (user@host.com leads to
http://www.host.com/~user/)

- post a message in a newsgroup that deals with the topic. Hopefully the site owner or one of his friends reads the messages there.
- if you are still unsuccessful, either delete your link to the site or repeat your attempts after a month (some sites might reappear in a search engine after some time). Sometimes it happens that a host is reorganizing its hard disk, and all user pages get back within a few days.

17. What about ftp and gopher sites?

Starting with version 1.0k I have implemented a new ftp checking method that is 100% reliable. Sadly, this method [does not work with proxies](#). The previous method I used (and still use for gopher) was unreliable, as it did not detect certain errors.

The method for checking gopher sites is still unreliable. When an ftp or gopher site is accessed through a proxy, this proxy builds up a web page. Sadly, it doesn't always bring up the information whether the URL exists or not. When you access a gopher site without a proxy, it brings an error message, but not an error code. This seems to be a bug of the `OpenURL()` function of `WININET.DLL`.

The output lists ftp and gopher sites as links, which allows you to make a manual check of these sites.

18. Why can't I launch URLs?

Starting with version 1.0g (Christmas 1997), URLs are launched with DDE ("dynamic data exchange", a windows method of communication between applications), to open many browser windows but to prevent the opening of several Netscape applications. This is done with the help of the Registry, by searching for HKEY_CLASSES_ROOT\http\shell\open. This has the path for the browser, the DDE application name (e.g. "Netscape"), the DDE topic (usually "www_OpenURL"), and a template for the DDE item (usually "%1"). If you cannot launch an URL, do not panic - export and e-mail me the segment of your registry (start REGEDIT.EXE, and search for "http").

The cause is usually that you have not installed Netscape properly (maybe you just transferred the files from another computer). Solution: reinstall Netscape over your current installation.

Starting with version 1.1b, I have stopped displaying an error message when the registry is incomplete, because there were too many complaints. Instead, the browser will simply be launched with the page. This has the disadvantage that the page won't be displayed in an extra window of the current active browser application.

One user with Microsoft Vista 64 (UAC disabled) was unable to launch URLs (message box: "Unable to open browser for 'URL': error 5: Access is denied"). The cause was COMODO Firewall Pro 3.0.25.378. Without the firewall, it worked fine. Please remember that ["Personal Firewalls" are mostly snake-oil](#). Set up an external firewall box instead - this is usually included in your router.

18a. Why does the browser not open a new window?

This is a problem with Microsoft Internet Explorer. Open your registry and search for HKEY_CLASSES_ROOT\http\shell\open\ddeexec. If the key

value is "%1",,,-1,0,,,,, then change it to "%1",,0,0,,,,, (i.e. you change the -1 to 0).

18b. Why does Link Sleuth freeze when launching the report?

If Link Sleuth freezes when launching the report, but not when double-clicking on a URL, the reason might be the site map. A site map can be HUGE if the site goes very "deep" (high level, see the "level" column in the Link Sleuth window). A very "deep" site can happen if you have a forum.

Solution: disable the site map in the options dialog, or exclude the "deep" parts of your website (e.g. a forum) in the initial dialog box.

18c. Why does Link Sleuth freeze when launching the report or a URL?

I do not know why this happens, but I have experienced this myself with Windows ME (but not with Windows XP), and have received similar reports from users. The problem goes away by rebooting Windows, but comes back later. You can also get rid of the problem by making a change in the XENU.INI file below the line with [options], enter this:
UseDDE=0

The only disadvantage is that it will not open a new window in the browser.

19. What about cookies

By default, cookies are disabled, and Xenu rejects all cookies.

If you need cookies because

- you have used Internet Explorer to authenticate yourself before starting a run
- to prevent the server to deliver URLs with new session id's

then you can enable the cookies in the advanced options dialog.
(This has been available since Version 1.2g)

Warning:

You should not use this option if you have links that delete data, e.g. a database or a shop - you are risking data loss!!!

20. Why are some links reported as "broken" by Xenu, that can be displayed within my browser?

Some servers read the "User Agent", i.e. the name of the software that tries to access a website. Some websites are programmed only for Netscape and Internet Explorer, and refuse everything else. Some may even specifically refuse Xenu because of past misuse. Andi has a [list of websites](#) that deny access to Xenu. A user-configurable "User Agent" would be the solution, but this would make abuse possible.

21. Why can't I connect to "secure" (https) sites ?

If you have [set your proxy](#) correctly, try to connect with IE. If this doesn't work, [read this usenet post for help](#). If this still doesn't work and you use Windows NT 4.0, [install the latest NT service packs](#) (up to SP5).

22. Any known problems with Windows 95?

Some people have reported crashes. These problems were usually solved by installing IE 3.0 (or higher) or the following service packs:

- [Windows 95 Kernel 32 Update](#) (29.7.1997)
- [Windows Socket 2 Update](#) (19.2.1998)
- [Microsoft DUN 1.3 and Winsock2 Year 2000 Update](#)
- [Patch for "File Access URL" Vulnerability](#) (12.11.1999)
- [Microsoft DUN 1.4](#)

One guy had problems with the WININET.DLL (v. 4.70.1300) installed with OEM Windows 95 (v. 95 4.00.950 C). Changing to version 4.70.1335 solved the problem.

A simpler solution is to go to <http://windowsupdate.microsoft.com> and install whatever they tell you (you need to have IE 4.0 or higher on your system)

23. Any known problems with Windows 2000?

Although I received many reports that it runs fine, one user reported a problem and a solution:

Windows 2000 automatically sets a configuration option to use HTTP 1.1 for connecting to web sites. Many, many web sites do not use that version but continue to use HTTP 1.0, so the automatic setting may prevent connections. This is the reason why Xenu would not run for me. When I disabled that setting, Xenu performed properly.

To disable that setting: Control Panel -> Internet Options -> Advanced (tab) -> HTTP 1.1 settings (list heading) -> Use HTTP 1.1 (checkbox: uncheck it)

24. Can I configure the timeout?

Enter the number of seconds in the [options] segment in XENU.INI, e.g. as `timeout=120`. The default value is 60. Note that this isn't "perfect". Microsoft Windows has a [bug](#) so that the timeout can't be set the way it should. I am using a [workaround advice from Microsoft](#). However I have observed that it doesn't work if the timeout "hits" while trying to find out if a host name exists.

Alternatively, try this:

- Start the Registry Editor (**REGEDIT.EXE**)
- Go to **HKEY_CURRENT_USER \ Software \ Microsoft \ Windows \ CurrentVersion \ InternetSettings**
- Select **New > DWORD** from the **Edit** menu
- Call it **ReceiveTimeout** with a value of <number of seconds >*1000

(The "hidden" default is 300000, i.e. five minutes, which is too long)

- Restart your system

Some users have complained that if one URL hits a timeout or a failed connection, all URLs from that host also do. Starting with version 1.2h, this behaviour can be disabled by unchecking "fail all URLs with same failed host" in the advanced options dialog. (The default behaviour is "checked")

25. What about JavaScript?

JavaScript is a programming language, not a formatting language. This makes web pages dynamic; they might depend on the mouse type, the screen size, etc... I have been begged to check simple JavaScript links that have the form

```
javascript:function('address', param1, param2, . . . , paramN)
```

My solution, which was first [announced in the user group](#), requires a change in the XENU.INI file. You need a basic understanding of regular expressions (regexp). You must put your regexp in the INI file, like this:

```
[Options]
Javascript=javascript:.*\(['"])(.*(/|s?html?|gif|jpe?
g|png|jsp|cfm|zip|exe|aspx?|pl|pdf|xml|ra|asx|ram|swf|php)(\?.*)?)
['"](.*)
```

In the example above, the substring within the first (...) must contain the valid URL.

Frank Visser [suggested](#) an improved regexp:

```
[Options]
Javascript=[j|J]avascript:[_a-zA-Z0-9]+ *\(['"])([^\"]+)[']
```

Frank Visser has also written a [better description on his site](#).

Note: there may be no blank between within "javascript:function", because the regexp wouldn't work.

26. What about passwords entered in a FORM?

The software is not able to enter passwords in a FORM. I just don't see a way to accomplish this easily. I assume it is possible if one combines a set of variable names, values, and a web page that would accept them with a POST command. But some alternatives might work:

- Log in with Internet Explorer, start Xenu, then enable cookies in the advanced options dialog (read the [details](#)), then start the check
- If the server accepts authentication with GET (should work with the Tomcat server), try sending such a URL. However, you might still have to activate cookies.

27. How about a WAP version?

Xenu does check .wml files since February 2001.

28. What about these error codes?

I identify only a subset of all possible error codes in the "Status" column. If you get an unknown error code in the Xenu application window, you can scroll to the right for an explanation text.

More information:

- [HTTP status codes](#) (100 through 505)
- [WinInet error codes](#) (12001 through 12156), also [here](#)

29. Why do I get broken links with `filelist.xml`, `editdata.mso` and `oledata.mso`?

Because Microsoft creates these broken links :-(Don't bother with them, or read Knowledge Base article Q219694: [Saving Office HTML File to a FrontPage Web Results in a Broken Hyperlink](#). Or try this tool: [Office 2000 HTML Filter 2.0](#)

You can also get rid of the problem by excluding them in the advanced options dialog.

Xenu will exclude URLs that end with `/filelist.xml`, `/editdata.mso` and `/oledata.mso`. (This feature has been available since Version 1.2g)

30. Why do I get "file not found" on remote checks?

There may be several causes for this:

- Your Internet Explorer isn't working properly, or is in offline mode, or is blocked by your firewall. Enter the URL you want to check into IE and see if it works.
- One user got it working by starting Internet Explorer first, and then starting Xenu. I believe that the cause is a broken setup of Windows, or of Internet Explorer.
- Your temporary directory is full: enter %TEMP% (not "c:\temp" !) into the Windows Explorer, check if there are many TGH*. * files, and delete them.

31. Can I make a foreign language version?

No, please don't. There's no guarantee that any of the message texts will be kept in the next version. The other problem is that I didn't write the software in a way to be language-independent. I could have done it - but I think most people on the web do understand english.

32. Why isn't Xenu detecting missing URLs?

A web server should return HTTP error 404 for non-existent URLs. Some servers are poorly configured: some redirect to an existing URL with an error message (bad!), others do show an error page, but the server doesn't return the 404 error (very bad!).

One user had the problem that his Microsoft IIS server didn't return the 404 error. He found help on [this page](#), and then sent me his solution, which only works in .ASP under IIS:

```
<%@LANGUAGE="VBScript"%>
<% Response.Status = "404 Not Found"
On Error Resume Next 'important in an error page to
strTarget = Request.ServerVariables("QUERY_STRING")
strReferer = Request.ServerVariables("HTTP_REFERER")
<HTML><BODY>The page doesn't exist, sorry dude!<BR><
```

The Apache web server has a different (and better) method of doing the same thing using native HTML code for the webpage. You simply set the correct config items in the http.conf file on your box.

33. Running Xenu with Norton Internet Security

A user had trouble to use Xenu with Norton Internet Security 7, and got error 12007 (no such host). After he added Xenu, it worked. This is what he did:

- added Xenu by opening Norton Internet Security by double clicking the Norton "Globe" Icon
- In right side panel, "Personal Firewall", Click CONFIGURE
- In the Personal Firewall pop-up, Click PROGRAMS
- In PROGRAM CONTROLS MENU
 - Settings for - "Home (Active)"
 - click "Turn on Automatic Program Control" box
 - Under "Manual Program Controls"
 - Scroll to "Xenu.exe"
 - Click Xenu once to Highlight it
 - Click MODIFY
- in the pop-up:
 - Click PERMIT
 - Click OK
 - Click Ok again
- Run Xenu

34. Why timeouts?

This is difficult to answer. The cause might be network overload; it might help to set a lower amount of threads, or to fine-tune the DoS detection of your firewall.

Check your firewall logs to see whether it detected a "SYN flood" DoS attack by you. SYN is the first data packet that is sent to a host when starting a connection. Theoretically, Xenu might send up to 100 SYN packets that are not immediately answered, so a firewall (that counts "unanswered" SYN packets) might think something "evil" is going on. My firewall box once claimed to have detected a SYN flood when I opened many newspaper articles in background browser windows.

35. Any Spyware, Adware, Malware?

This software exists since 1997 and never had any type of malware. It does not "phone home" or return any statistics to me. There are random "ads" in the HTML report for causes I support; however I don't get paid for this. Any passwords that you enter in the software (e.g. for orphan search) are not "remembered" after you close Xenu, nor are they passed to me.

Some debug output is stored in the file XENULOG.TXT which you will find in your %TEMP% directory. That file does not contain any passwords and it is used for support (I will sometimes ask you to attach it to an e-mail to me), primarily for problems with the launch of URLs in your browser (especially the report). The file is human-readable, so feel free to have a look. The file is not sent to me by Xenu, it just sits there and you can delete it if you wish.

Here's a green review by [McAfee Siteadvisor about Xenu's Link Sleuth](#). Note that until before July 11 2008, Yahoo Search (which uses input from McAfee SiteAdvisor) was redflagging every URL of the whole snafu.de domain, including my user site (this seems to have been corrected now). McAfee SiteAdvisor has [redflagged the snafu.de domain](#), but not the user pages. This was related to three downloads (CuteFTP, GoZilla and Nok2Phone) on the customer support ftp site of snafu.de, who has been my ISP for over a decade. These downloads have been removed since then and both Yahoo and McAfee have been notified. On July 30 2008, I noticed that the snafu.de domain has been greenflagged.

If you have any more questions about security, don't hesitate to [contact me](#).

Appendix 1: Effects on Hit Counters (and other things to keep in mind)

Hit counters will be affected by using Xenu. Every time you use Xenu to check the links on your pages your counter will record a hit for each page where Xenu checks links, so be aware that your hit counter totals will be increased.

Xenu's Link Sleuth also affects webservers since it sends a lot of requests very quickly. This adds to the load on the server and could ultimately affect performance for all users. If you have a large site with many links that need to be checked, please consider running your link checks at an off-peak time (early in the morning or on a weekend).

Appendix 2: Future feature List

Things I will do in the future (maybe when hell freezes over!):

- simple conversion from Unicode
- [ROBOTS.TXT](#) support
- Detect [remote loading](#) of images ([geocities](#) sabotages this)
- Custom views in Xenu Window
- Solution for leftover TGH*. * files in temp directory
- Command-line parameters (actually, this has already been done, for a client who agreed to pay my development time to two people I support. If you need something similar, e-mail me, the price is a \$300 donation to be split between two people I support)
- Names of last checked URLs in also file menu
- Server based Link Sleuthing, i.e. to be used as CGI application, so that ISPs could offer link sleuthing to their own users, i.e. users could check their own web sites.
- If you are an ISP who is willing to offer this, contact me to work out details.
- Automatic saving every minute
- A correctly working "Update" feature that rechecks changed sites (tricky, so I will never do it)
- Checking of internal anchors, checking for duplicate anchors
- Ideas from Chris:
 - What about identifying how many steps it takes to reach a particular page from the home page and how much kb had to be downloaded before one could reach there.
[TH: useful e.g. to which steps a user must take to reach the page of a particular product]
- [Read RFCs](#)
- [Your suggestions](#): e-mail me also if there is something of the above you'd like to have, and persuade me to do it.

Appendix 3: Bug List

The software works pretty well, but here the list of things that shouldn't be.

- the thread count is sometimes incorrect if the maximum is changed while active
- the thread count is sometimes incorrect at the end of the session
- The `` closing tag must not have spaces or newlines inside
- `<applet code="myclass.class" archive="jump.zip">` will produce a broken link if `myclass.class` exists, but only in the archive
- leftover TGH*. * files in the %TEMP% directory
- weird effects when INI file >64K

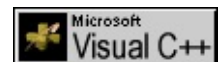
If you find another bug, [e-mail me a description](#), **please include the URL you are checking, and if possible try to save your work in a .XEN file and attach it.** Also check <http://windowsupdate.microsoft.com> to make sure that your system has all the updates.

Appendix 4: Story of Xenu's Link Sleuth™ (for fellow software developers)

In April and May 1997 my employer assigned me on an out-of-town job, because another department needed a guy with MFC experience. So from monday to friday I was away, and on the evenings I was bored to death. Every week-end I was back home, and I usually checked my web site for broken links with *WebAnalyzer*. Sadly the software had a lot of bugs, and their support was ignoring my e-mails, and I was mad as hell, as I had spent quite a lot of money on a product that wasn't worth it. My job was also the first contact with VC++ 4.2 (previously I had only worked with VC++ 1.5, because our customers have a lot of 16bit systems), which had some easy-to-use Internet access classes. I had already experience with WINSOCK programming, but these classes would spare me a lot of time evaluating HTTP result headers and other annoying stuff. On an evening after an excellent italian food with a good chianti I took some hotel letter paper and wrote down a concept for checking links. A month later I took some time to install the development software on my computer and started working, with the help of that hotel-room concept. The work was done on some evenings, but mostly on week-ends, when I had more time.

My philosophy on software development has always been "smaller, simpler, cheaper", long before the NASA realized this (in May 2002 I was told that the actual NASA philosophy was *Faster, Better, Cheaper* - oops!) Because of that, I need no fancy (but totally useless) graphics like in *WebAnalyzer*. Just results. And they'd better be 100% correct or I'd have to kill myself :-)

The application is written in Visual C++, and uses the MFC classes as much as possible: CDocument, CView, CListView, CObArray, CMapStringToOb, CArchive, CInternetSession, CHttpFile, etc, etc. That saved me a lot of time!



Appendix 5: Keyboard shortcuts

CTRL+N Check **N**ew URL

CTRL+O **O**pen Xenu report

CTRL+W Close Report **W**indow

CTRL+S **S**ave Report

CTRL+T Export to **T**ab separated file

CTRL+R **R**etry broken links

R Generate web based **R**eport

CTRL+Q Exit Xenu (**Q**="Quit")

CTRL+C **C**opy

CTRL+F **F**ind

F3 Find next

CTRL+V **P**aste

CTRL+B Show **B**roken links only

Appendix 6: Credits

- Icons in EXE file: Martin Hunt and Paul Campbell
- Icon on web page: Erik Plummer
- Idea to use banners in report: Marc Cross
- Xenu logo button: Fred C.
- second Xenu logo button: [Charles A. Upsdell](#)
- Volcano animated cursor: [Juan C. Pradas-Bergnes](#)
- Idea & help with SMTP integration: [Mark Findlay](#)
- SMTP class: P.J. Naughter
- Xenu artwork: [William C. Chenoweth](#)
- WinHelp version of documentation: [Andrew Schoenhofer](#)
- Regular Expressions: [Henry Spencer and Guy Gascoigne](#)
- Install and deinstall: [Inno Setup](#)
- help with Xenu banner: [Bruno Zacke](#)
- [wildcards](#): Jack Handy
- [sort icons for list columns](#) Thomas Holte
- new Xenu icon in May 2008: Dominic Rath of [Hitflip.de](#)
- Idea and code for GraphViz: [Kevin Niehage](#)
- Current HTML Help (CHM) version of documentation: [Andrey Aleksanyants](#)