Know-It WebTrak™

Information in this document is subject to change without notice. Companies, names, and data used in examples are fictitious unless otherwise noted. No part of this document may be reproduced or transmitted in any form by any means, electronic or mechanical, for any purpose, without the express written permission of Know-It.

Product and company names mentioned herein my be the trademarks of their respective owners.

Know-It™ WebTrak™ v1.5 Guide

Introduction

WebTrak provides basic HTTPD log analysis and reporting. WebTrak reads any given CERN formatted log (Common Log File Format (NCSA)) from a text file and produces a report in HTML format suitable for viewing with your Web browser. WebTrak produces statistical report and bar graphs in HTML.

FEATURES:
1) Reads CERN formatted HTTPD log files.
2) Produces HTML formatted reports and graphs for:
. Top Users
. Activity by Date
. Activity by Day
. Top Pages accessed
. Top Directories accessed
. Activity by nation
. Hourly Activity Report
. Analysis by Server Response Code
. Top Not Found Report

3) Reports include numeric and graphical (bar chart) representation.
4) Easy to install and use, WebTrak is made to be run in batch mode with parameters defined in a easy to edit text file.
5) Administrators my specify up 100 address exclusions, including 'wildcard' specifications, Input file name, output file name, file type exclusions, and may specify whether to perform IP address/name conversion or not.   Administrators may also exclude file types (.gif, .jpg, etc..) from the Top Page report.
6) IP address/name conversions are cached thereby decreasing reporting time required. This cache is saved in a file which is preloaded for the next reporting run, so subsequent report production times are dramatically decreased.
7) If extra data is found after the standard format, WebTrak will automatically produce table reports as needed. For example some Web Servers supply additional information about the referring site or the browser used. WebTrak will detect this information and automatically and produce tables EXTRA1 and EXTRA2 if needed. The titles to these table can be changed as needed (see below).
8) Any title or heading in the WebTrak output can be customized as needed including the use of foreign languages.

Section 1 - Installation and Setup

There are three steps to successfully setting up a reporting system using the WebTrak software. These steps are:

1) Install WebTrak Kit.
2) Edit the WebTrak parameters file.
3) Setup the WebTrak executable to run manually or automatically.

Step 1 - Install WebTrak kit:

Initially download the latest version of the WebTrak installation kit from the Know-It Internet Web site. The kit should be called 'KIWT_KIT.EXE' and is a self-installing archive. Execute the kit file start the installation process. You can execute the kit file by typing it's name in the command line of a DOS window, or double-clicking it from the file-manager, or using the RUN option from the file-manager or Explorer.

Before running KIWT_KIT.EXE, exit any programs that may interfere with the installation of this product. Upon executing KIWT_KIT.EXE, you will need to provide the following information to the screens when prompted:

A target directory for placing the executable files. A separate directory just for the WebTrak executables may be created if desired, but is not necessary, as WebTrak contains a complete uninstall utility should the need ever arise to remove the WebTrak system. There is one 32 bit executable program and a parameter file (WTPARAM.DAT) that are placed in this directory. (NOTE: If you are performing an upgrade, you should save your WTPARAM.DAT file.) After the initial run of the WebTrak program, a third file will be created which is called WTIPNAME.DAT which serves as the IP address-to-name cache file between runs if you set the LOOKUP flag to 'Y'.

A target directory for placing the provided HTML file.   The WebTrak kit provides a single file called HSTB.GIF which is required to be accessible to the browser that is used to view the HTML output report file. The report file assumes that the .GIF file is available at the top level of the server, otherwise you need to specify the HSTB_LOC parameter (explained below).

After entering the appropriate information, the files are copied from the kit to the target directories.

STEP 2 - Editing the parameter file:

The parameter file (WTPARAM.DAT) may be easily edited to adjust the report requirements as desired. Parameters are set using commands placed on individual lines in the file. The file can be edited with any text editor, but must be saved as a text file (i.e. not in a word processor format). When editing the file, lines beginning with ';' (or blank lines), are comments. You can add as many comment lines as you need for your own reference.

WebTrak Commands are:

EXCLUDE_ADDR   - (any part of an IP address(es) to skip
                example: 123.123.123 will skip all log records that have
                an IP address <u>beginning</u> with 123.123.123
                (such as 123.123.123.7). If your server pre-translates
                addresses, so that you have names instead of numbers
                in your log file, you can enter names to exclude also.
                (such as FRED.ABC.COM). Enter one exclusion per line.
LOOKUP   - (Y or N) turns on or off IP address to name lookup
                (if you don't care about the names, reporting is faster off)
                You can also set the flag to 'X' which means only use the
                existing IP name table (no new lookups). This is useful if
                you want to rerun a report quickly. The default is 'Y'.
INFILE   - designate the log file to be read. If not present program
                will ask at runtime.
OUTFILE   - designate the HTML file to be created. If not present
                program will ask at runtime.
DEBUG   - 'N'(default), 'Y' show IP address lookup activity, 'F'ull to show
                too much.
TITLE   - The title you want placed at the top of the report
LICENSE   - The license for your site. If not valid, will only
                process 1000 log records - good for testing. A license
                key can be obtained from Know-It.
TIMEOUT   - Maximum number of seconds to perform a DNS lookup
                of an IP address (if LOOKUP is on).
EXCLUDE_TYPE - Exclude certain file types from the top page report. Typical
                files to exclude are .gif, .jpg, etc..   Be sure to include the
                dot ('.').   One exclusion per line.
HOFFSET   - Allows designation of an offset amount from the time in the log
                file to the actual local time. Value may be positive or
                negative.
NOSHOW   - No parameters required. This turns off the report at the end of
                input parameters if desired.
EXCLUDE_EXTRA1 - (any text string or set of characters to skip or leave

out of the EXTRA1 report.

example: www.mysite.com will skip all log records that have extra 1 data containing 'www.mysite.com'.

Enter up to 100 exclusions - one per line.

EXCLUDE_EXTRA2 - (any text string or set of characters to skip or leave out of the EXTRA2 report.

example: www.mysite.com will skip all log records that have extra 2 data containing 'www.mysite.com'.

Enter up to 100 exclusions - one per line.

MAXUSERS - Designate the size of the Top Users report. (10 - 9999)

MAXPAGES - Designate the size of the Top Pages report. (10 - 9999)

MAXDIRS - Designate the size of the Top Directoriess report. (10 - 500)

MAXNATIONS - Designate the size of the Top Nations report. (10 - 500)

MAXNOTFOUND - Designate the size of the Top Not Found report (10 - 1000)

MAXEXTRA1 - Designate the number of rows in the EXTRA 1 report if present.

MAXEXTRA2 - Designate the number of rows in the EXTRA 2 report if present.

ADJWIDTH   - Adjust width of bar graphs relative size. Width Factor starts at 400. You can use this parameter to adjust the size by setting it from -350 to 350. Using the line ADJWIDTH=-100 for example would decrease the width of

the bar graphs by about 25 %. (that's a minus 100).

ROBOTS=RPT - causes the Top Users report to indicate if a user is a Web Spider.

PLOW   - Sets the priority of WebTrak to LOW to allow other jobs to run first. WebTrak runs when the system is idle.

PAGES=FULL - If set, will include parameters passed to forms in the top page report. You can choose to separate forms with different data submitted or not.

HSTB_LOC - Allows specification of an alternate location for the hstb.gif file. If not specified, WebTrak assumes the file is in the top-level location of the server (i.e. /hstb.gif). With this command you can place the hstb.gif file anywhere. The text is used as a prefix. Example: HSTB_LOC=/images would cause WebTrak to use /images/hstb.gif . Note: do not use a trailing slash.

Additionally, you can modify the text sections of the report if you wish to change the words to something clearer for your purposes, or to a foreign language. You do not need to include any of these in your WTPARAM.DAT file, unless you want to change them. They will default to the default English text. The text fields are all shown in appendix A. For Example:

TEXT1=Report produced

To change a text, simply include the line in your WTPARAM.DAT file. To change TEXT1 to the Spanish equivalent for example, you would add a line to your WTPARAM.DAT file that looks like this:

TEXT1=El informe produjo sobre

NOTE: A limit of 50 characters applies to each text field. That is, you can only use up to 50 characters of replacement text.

Here is an example file:
EXCLUDE_ADDR=123.456
EXCLUDE_ADDR=678.010.123
LICENSE=0000000000
INFILE=C:\HTTPD\LOGS\ACCESS.LOG
OUTFILE=C:\HTTPD\HTML\WEB_RPT.HTM
TITLE=My Company's WebTrak Report
TIMEOUT=15
EXCLUDE_TYPE=.gif
EXCLUDE_TYPE=.jpg
HOFFSET=-5
NOSHOW
LOOKUP=Y
MAXUSERS=30
ROBOTS=RPT
PAGES=FULL
PLOW
HSTB_LOC=/images

STEP 3 - Setup to run manually or automatically:

Produce a report by running the program, the KIWT.EXE . This can be accomplished by executing the KIWT.EXE file, that is typing it's name in the command line of a DOS window, or double-clicking it from the file-manager, or using the RUN option from the file-manager or Explorer. It is likely that this process will be automated by having the program invoked using a scheduling tool such as AT under Windows NT, or the System Agent under Windows95. No parameters need be given to the program at runtime since all can be defined in the parameters file. However, if the input file or output file are not designated in the parameters file, they will be asked for at runtime.

NOTE: The first time the program is run, if the input log file is large and contains many diverse IP addresses, it will take quite awhile to complete due to the time required to lookup each new IP address. These 'name server' lookups will be saved for subsequent

runs, and only incrementally new addresses requested from the name server. Typically, new addresses represent only a small percentage of subsequent runs. To produce quicker results, the LOOKUP option may be turned off (see STEP 2 above) if the name of the sites is not required.

Section II - Maintenance

No maintenance is required by the WebTrak software. The WTIPNAME.DAT file may be edited if desired to delete entries for addresses that have changed, or if the editor wishes to change the name reported for a particular address. It might be desired for example to change an machine name (such as machine12.company.com) to a particular user's name (such as John Smith). The file can be edited with any text editor, but must be saved as a text file (i.e. not in a word processor format), and the format of the file (comma delimited) should be maintained.

Additional features are planned for version 2.0 of WebTrak. The ability to switch off and on certain reports will be available. If you desire these options or others please send an Email to KI_TEAM@KNOW-IT.COM with the details of your request.

Ordering Information:

To receive a registration key for WebTrak, please send the system node name of the system on which you will be running WebTrak, plus the registration fee ($49 commercial, $39 non-profit groups (in US$ via check or money order) *), to

Know-It / WebTrak Registration
14651 Dallas Parkway
Suite 404
Addison, TX 75240

Upon receipt, we will return your registration key which you can enter into the WTPARAMS.DAT file to allow processing of unlimited records. The system node name of your system, can be observed in the first few lines of the screen output of WebTrak when it is invoked. The line says "running on:" and the system node name of your system. For a WindowsNT system, it is usually one word (e.g. 'myserver'), but for a Windows95 system, it is the complete network name (e.g. 'myserver.xyz.com'). The system node name can also be seen at the end of the WebTrak report.

* Always check the Know-It Web Site for the latest information, prices and releases at:

APPENDIX A - Text Fields - default values

TEXT1=Report produced
TEXT2=Log From
TEXT3=Thru
TEXT4=Total Records (Hits) Processed
TEXT5=Top Users
TEXT6=Unique Users Found
TEXT7=Total Pages Requested
TEXT8=Activity by Date
TEXT9=Rank
TEXT10=User
TEXT11=Hits
TEXT12=Graph
TEXT13=Top of Report
TEXT14=Date
TEXT15=Total Days in Log
TEXT16=Average Activity by Day of Week
TEXT17=Day of Week
TEXT18=Average Hits
TEXT19=Percent
TEXT20=Sunday
TEXT21=Monday
TEXT22=Tuesday
TEXT23=Wednesday
TEXT24=Thursday
TEXT25=Friday
TEXT26=Saturday
TEXT27=Hourly Activity Load
TEXT28=Hour
TEXT29=Parameters
TEXT30=Top Pages
TEXT31=Page
TEXT32=Total Unique Pages
TEXT33=Top Directories
TEXT34=Directory
TEXT35=Total Unique Directories
TEXT36=Top Nations
TEXT37=Nation
TEXT38=Total Unique Nations
TEXT39=Analysis by Server Response Code
TEXT40=Top Documents not found
TEXT41=Code
TEXT42=Description

TEXT43=K bytes

(the following fields are used for the optional extra reports which are only
 generated when extra fields are found in the log file.
TEXT44=Extra 1
TEXT45=Occurance
TEXT46=Total Extra 1
TEXT47=Extra 2
TEXT48=Total Extra 2

(the following fields are all the server reported codes)
TEXT49=Unknown Code
TEXT50=OK
TEXT51=Created
TEXT52=Accepted
TEXT53=Moved Permanently
TEXT54=Moved Temporarily
TEXT55=Not Modified - OK
TEXT56=Bad Request
TEXT57=Unauthorized
TEXT58=Forbidden
TEXT59=Not Found
TEXT60=Internal Server Error
TEXT61=Not Implemented
TEXT62=Bad Gateway
TEXT63=Service Unavailable
TEXT64=Average K bytes