

EICSTES DELIVERABLE D4.1

PROTOTYPE OF THE WEBSITE ANALYSIS TOOL

Peter Morris

P.Morris@Surrey.ac.uk

University of Surrey.

January 2002

Web analysis prototype

These are the instructions for how to install the Web Log Analysis prototype system

Protecting your system

I checked all of these files using four virus scanners: McAfee, F-Secure, VFind and a specialist heuristic scanner called Skeptic from MessageLabs.Com. However, you should check the files once again prior to loading them onto your system. It goes without saying that I cannot accept any liability for problems with viruses that are introduced into your machine from this or any other source.

Likewise, this software is a prototype and does not have all the full system checking and testing that I would do with a full production. Although I am sure that there are no intrinsic problems with what I am delivering, you may wish to try the code on a PC that is not mission critical.

Background

The executable (log.exe) is the output from Microsoft Visual Basic and needs some supporting runtime files. These are:

tabctl32.ocx - support for the tabbed interface on which the analysis program is based
comdlg32.ocx - support for browsing for files
scrrun.dll - support for file I/O routines

These are added to your system automatically by the installation routine. However, if you believe your machine is set up in an unusual manner, please email me to discuss how the installation might be done safely on your machine.

Installation

This version only runs on Windows platforms and has been tested on Windows 98 and Windows 2000. There is no reason why it should not work on other versions and I will help you get it working

on those should you have any problems.

Using "My Computer" or Windows Explorer, navigate to the directory WP4Proto on the CDROM and double click on Install.bat

This will create a directory on your C drive called c:\eicstes and one underneath it called c:\eicstes\wp4proto

In the second of these two directories, the executable (log.exe) and a sample log file (sample.log), will be placed.

If you wish to move these files about, you can do so safely. No programs are registered nor are there any .ini files or Windows Registry settings.

The supporting run-time files are copied into Windows/System, Windows/System32 or WINNT/System32 directories (the latter is only for NT4 and Win2K). These are the normal places for such support files. However, your system may have some or all of these files in other directories. If the code does not work first time, you may wish to investigate whether any of the support files live in directories other than these three.

Using the Prototype

Not all of the features on all screens have been implemented yet. The nature of the prototype is to see sample results, not to provide full bells and whistles. Feel free to try anything. However, the machine on which you run the program will need Internet access if you want to resolve IP/Host names, since the program makes a call to your Domain Name Server.

If you have a log file of your own, either copy it into the c:\eicstes\wp4proto directory or point at it from the "Files" tab using the "Browse" button.

The results are quite simple at the moment. Don't expect to see much in the way of familiar paths in your output. I have found that the percentage of repeated paths is very small (typically less than .01%, although this will depend very much on the nature of the web site).

As for how to use it, use the "Browse" button to point at the log file from the "Files" tab and select "Start Analysis". As long as the log file format is standard Apache, the program should be able to read it. The sample file included with the prototype is in "Combined" log format.

Problems

There are no known problems with the prototype. However, National Language Support may get in the way of the program producing results. The prototype only supports UK English and, if you get no results from the analysis you may need to make the following change. How do you know if there are no results? There are two indicators: (i) the "Lines Analysed" field on the "Status" tab will be 0 and (ii) the "View Results" tab will have two empty white boxes.

How to fix this: Go to the control panel and change the language in "Regional Options" to be "English (United Kingdom)". The system may prompt you to restart your computer. You may ignore this. Run the program and all should be well. After you have finished running the program, you should change the language in Regional Settings back to your home language. Again, the computer may prompt you to reboot but you may ignore this once again.

Enhancements

I will be improving the program over the next month or two, prior to starting the port across to the chosen programming language. If you would like to see an update before then, please let me know

and I will include you in a future distribution.

Likewise, if you have any features that you would like to see in the package, let me know and I'll see what I can do to incorporate them.

Contacting me

The easiest way to contact me is by email but feel free to call me as well.

email: P.Morris@Surrey.ac.uk

phone: +44 (0)1483 683972

Peter Morris