Site Analytics in IBM WebSphere Portal

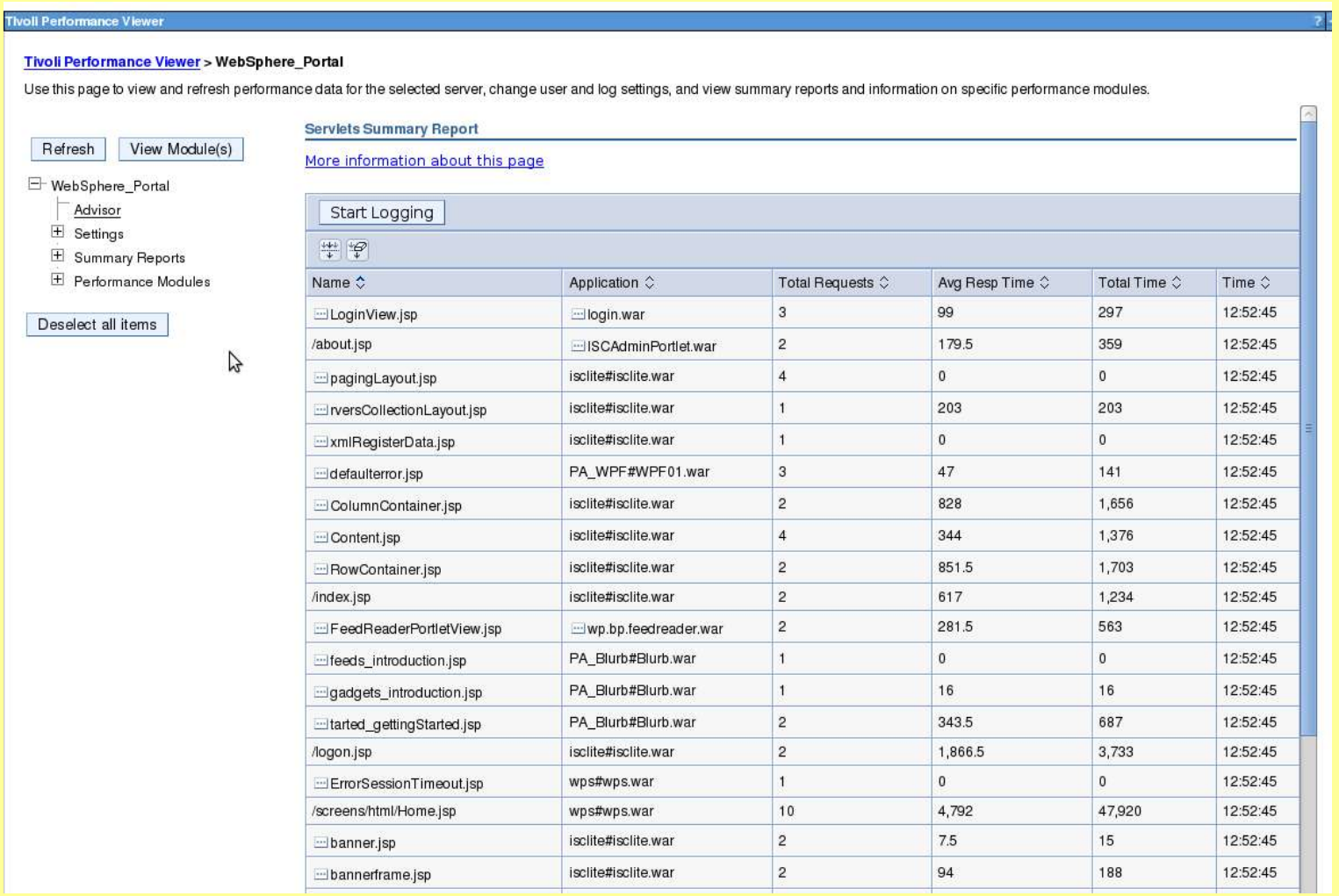
Dave Hay
Lotus Portal and Collaboration Architect
david_hay@uk.ibm.com
+44 7802 918423

This is a work-in-progress. It represents my own ever-increasing understanding and experience in this area.

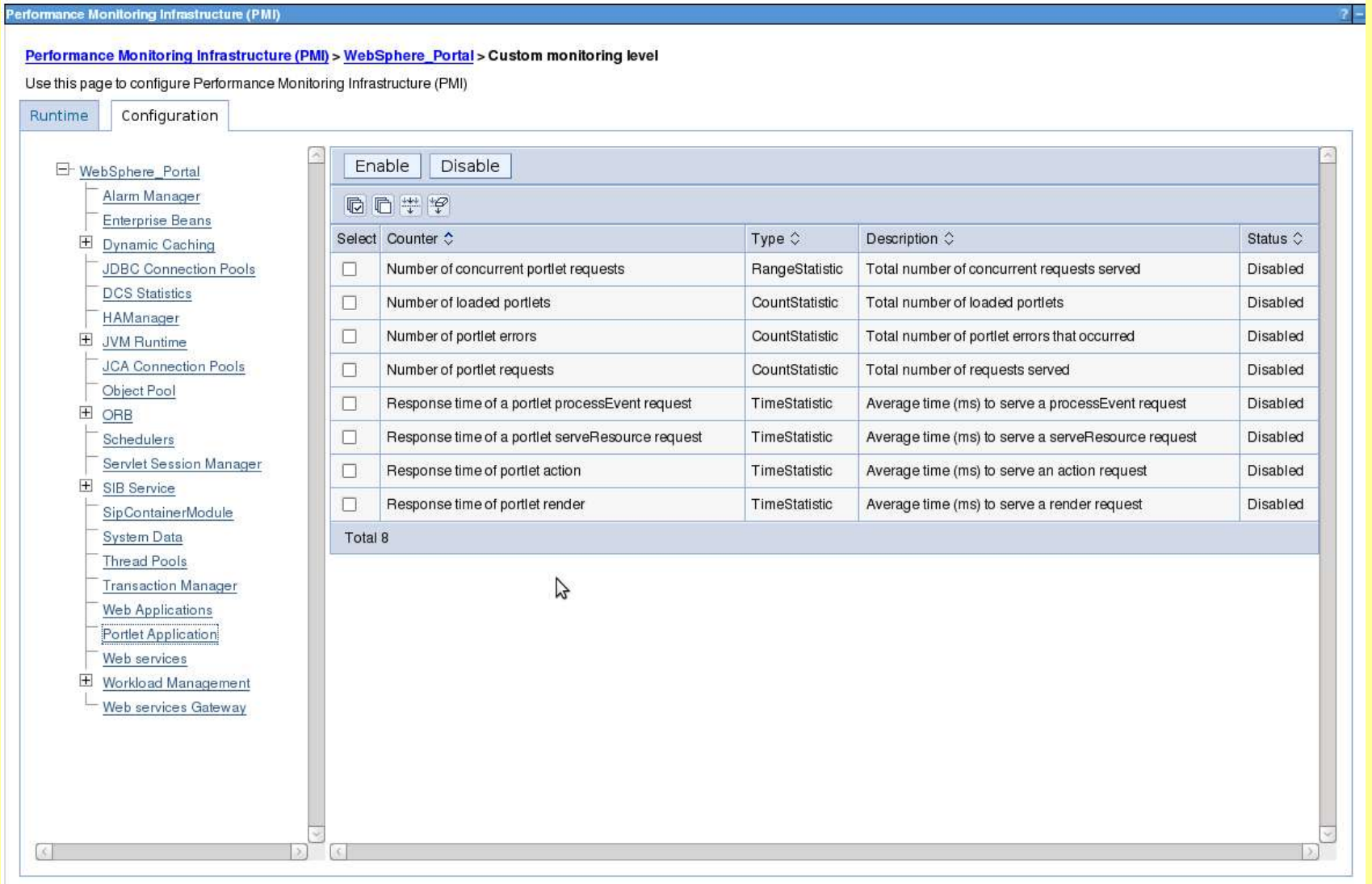
In brief, this shows how Site Analytics may be set up using IBM WebSphere Portal in order that usage and error logs can be analysed using tools such as AWStats, Nihuo and WebTrends (other tools are available)

The term "portal analytics" describes a process that can help you understand how your portal is used. IBM WebSphere Portal writes usage records to a dedicated log file. Because the format of the log follows industry standards ("NCSA Combined"), you can integrate portal usage data with your preferred reporting and analytics tools.

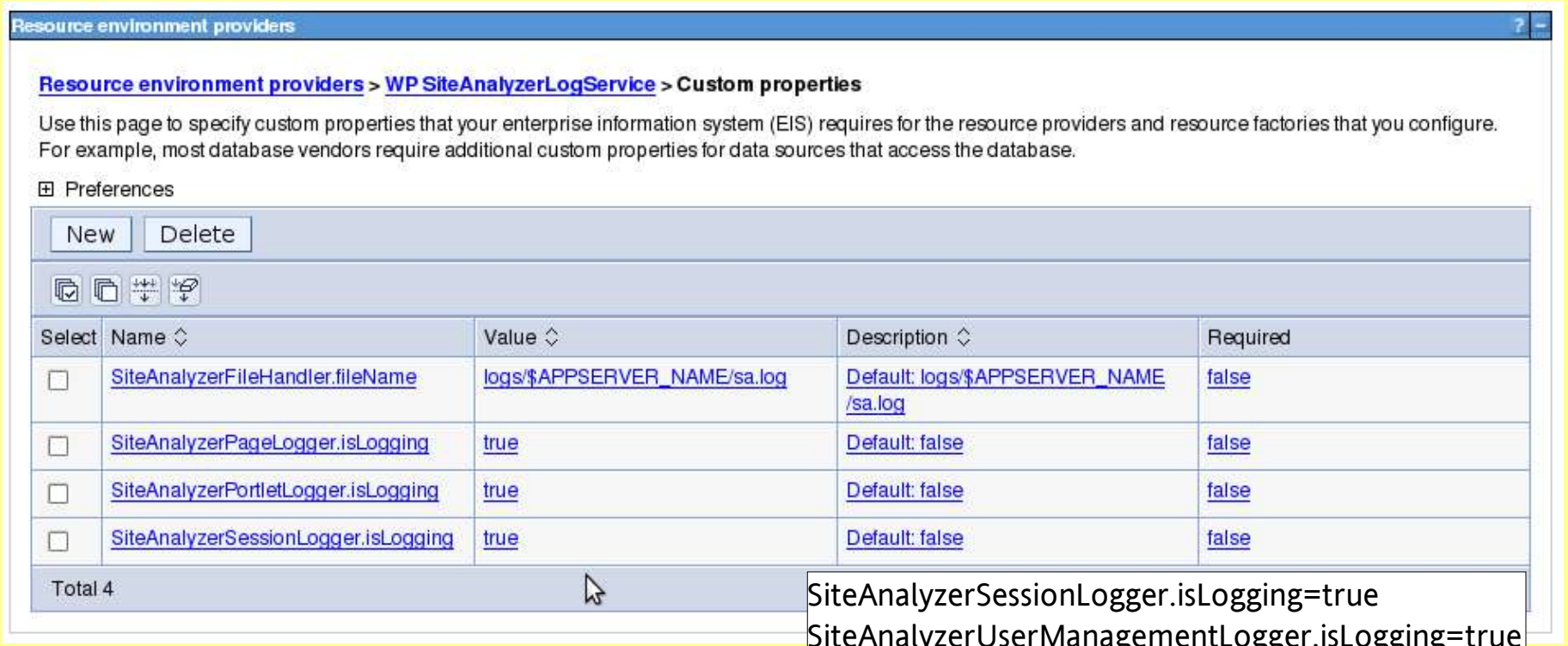
There are a number of options in terms of logging activity, including the performance of the Java processes within the underlying WebSphere Application Server



Or more specific portal/portlet-focused metrics: -



Or via a built-in service within WebSphere Portal, which can generate the NCSA log files for further external analysis




```
SiteAnalyzerSessionLogger.isLogging=true
SiteAnalyzerUserManagementLogger.isLogging=true
SiteAnalyzerPageLogger.isLogging=true
SiteAnalyzerPortletLogger.isLogging=true
SiteAnalyzerPortletActionLogger.isLogging=true
SiteAnalyzerApplicationActionLogger.isLogging=true
SiteAnalyzerErrorLogger.isLogging=true
```

Some sample reports from AWStats

Statistics for: dmht60p
Last Update: 17 Dec 2008 - 13:22
Reported period: Month Dec 2008

When: Monthly history Days of month Days of week Hours
Who: Countries Full list Hosts Full list Last visit Unresolved IP Address Robots/Spiders visitors Full list Last visit
Navigation: Visits duration File type Viewed Full list Entry Exit Operating Systems Versions Unknown Browsers Versions Unknown
Referrers: Origin Referring search engines Referring sites Search Search Keyphrases Search Keywords
Others: Miscellaneous HTTP Status codes Pages not found

Summary		
Reported period	Month Dec 2008	
First visit	17 Dec 2008 - 11:33	
Last visit	17 Dec 2008 - 12:33	
	Unique visitors	Number of
Viewed traffic *	2	2 (1 visits/vis
Not viewed traffic *		

Jan 2008	Feb 2008	Mar 2008	Apr 2008	May 2008	Jun 2008	Jul 2008	Aug 2008	Sep 2008	Oct 2008	Nov 2008	Dec 2008

Month	Unique visitors	Number of visits	Pages	Hits	Bandwidth
Jan 2008	0	0	0	0	0
Feb 2008	0	0	0	0	0
Mar 2008	0	0	0	0	0
Apr 2008	0	0	0	0	0
May 2008	0	0	0	0	0
			0	0	0
			0	0	0
			0	0	0
			0	0	0
			0	0	0
			0	0	0
		1	23	23	163 Bytes
			23	23	163 Bytes

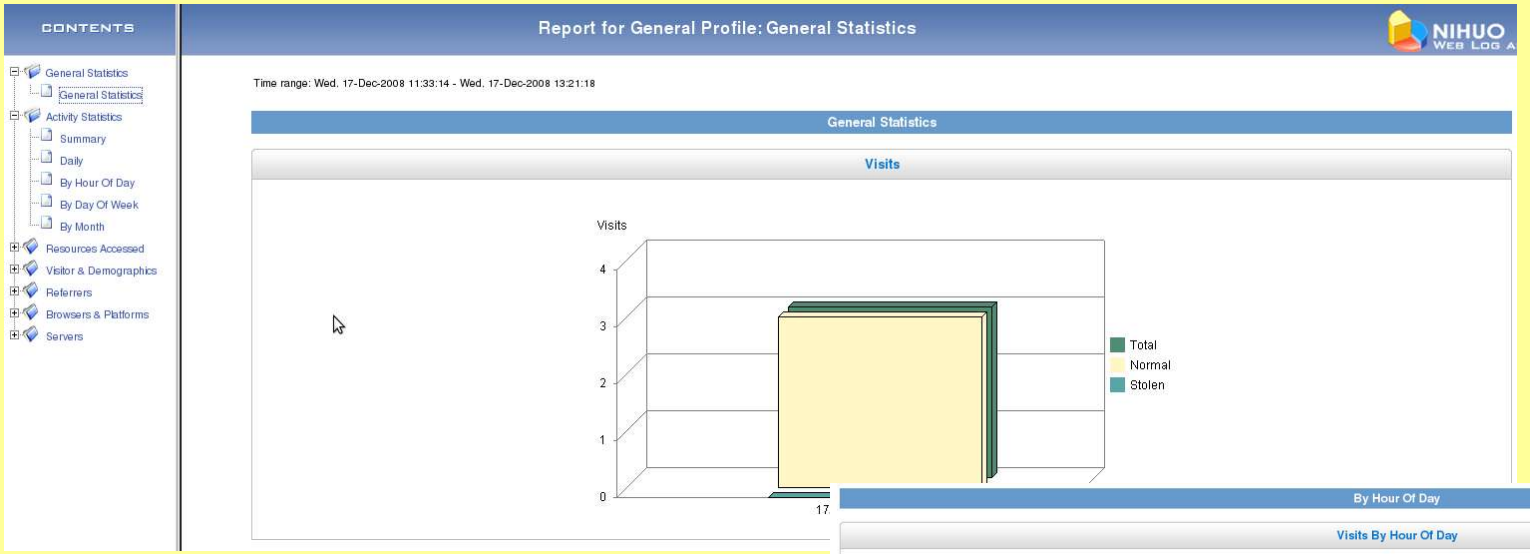
Pages-URL (Top 10) - Full list - Entry - Exit

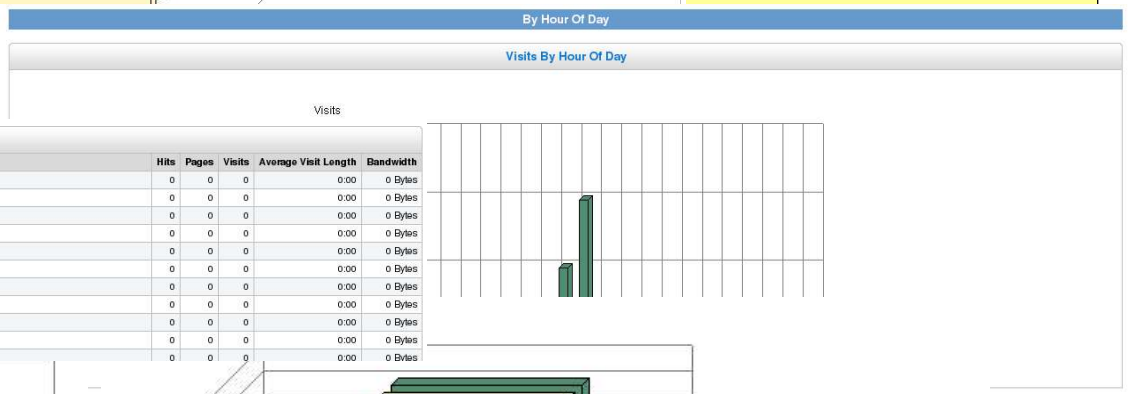
14 different pages-url	Viewed	Average size	Entr
/Page/6_CGAH47L00ODQC02F24OKIS0003/null	3	0	
/Portlet/5_CGAH47L00ODQC02F24OKIS00O3/Quickr_REST_Portlet	3	0	
/Page/6_CGAH47L0082M302B566VO730M6/null	2	0	
/Command/Login	2	92 Bytes	
/Portlet/5_CGAH47L0082M302B566VO730E4/Login_Portlet	2	0	
/Page/6_CGAH47L0009JD02B5MH2L10GQ6/null	2		
/Portlet/5_CGAH47L0009JD02B5MH2L10G60/Information_Portlet	2		
/Page/6_CGAH47L0009JD02B5MH2L10G67/null	1		
/Command/Logout	1		
/Portlet/5_CGAH47L0009JD02B5MH2L10GU4/Information_Portlet	1		
Others	4		

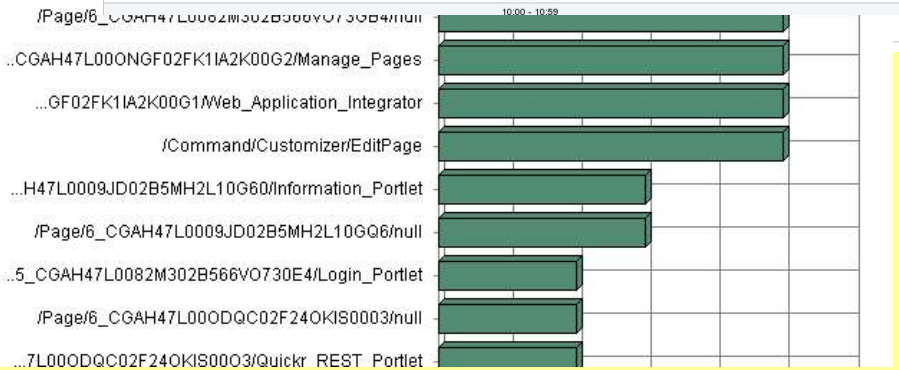
Direct address / Bookmark / Link in email...
Links from an Internet Search Engine - Full list
Links from an external page (other web sites except search engines) - Full list
- http://stargate.uk.ibm.com/Page/6_CGAH47L00ODQC02F24OKIS0003/nul... 3 3
- http://stargate.uk.ibm.com/Page/6_CGAH47L0009JD02B5MH2L10G67/nul... 3 3
- http://stargate.uk.ibm.com/Page/6_CGAH47L0082M302B566VO730M6/nul... 2 2
- http://stargate.uk.ibm.com/Page/6_CGAH47L0009JD02B5MH2L10GQ6/nul... 2 2
- http://stargate.uk.ibm.com/Page/6_CGAH47L0009JD02B5MH2L10GE1/nul... 1 1

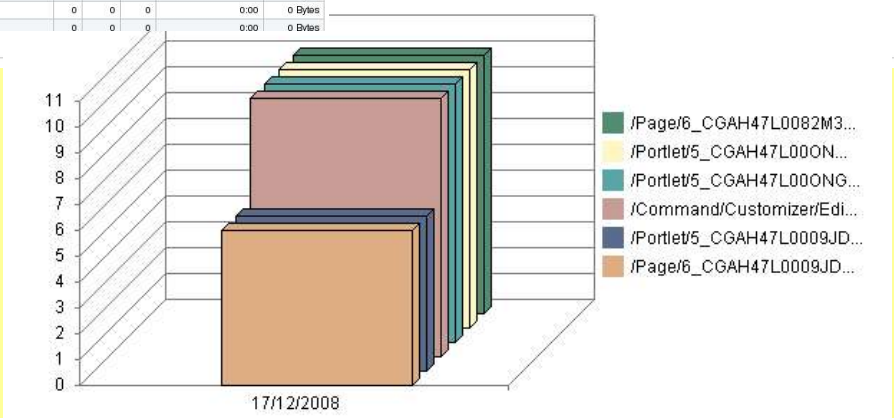
Unknown Origin

Some sample reports from Nihuo









More to consider ... Personalisation, Feedback and Likeminds

Feedback and analytics

Personalization provides a complete logging framework for collecting data on how visitors are using your Web site. If Feedback is enabled, data is automatically collected about each Personalization rule that is fired. In addition, development tools enable Web sites to collect a variety of data related to visitors' actions and behavior. By default this data is logged to a standard database schema for later analysis and reporting. The framework is also extensible, allowing Web sites to customize and supplement the way data is collected and stored to more fully meet their needs.

http://publib.boulder.ibm.com/infocenter/wpdoc/v6r1m0/topic/com.ibm.wp.ent.doc/pzn/pzn_feedbackanalytics.html

An introduction to LikeMinds

Personalization contains a dynamic recommendation system based on LikeMinds. LikeMinds is software that is used with your e-commerce applications. LikeMinds analyzes user interactions that occur on your Web site and generates real time predictions and recommendations to your Web site users.

Real time predictions are generated by three LikeMinds engines using recommendation rules within Personalization. These rules, called recommend content, base their predictions on transactions logged through Personalization's rating and action beans.

When a user visits your Web site, rating and action beans log captured transactional data. If your e-commerce Web site is set up so that users can rate content (or products), you use Rating beans to capture rating data. Similarly, if you use shopping cart technology, you use action logging beans to capture content affinity behavior to capture shopping activity. Both rating and action data is stored in your database.

http://publib.boulder.ibm.com/infocenter/wpdoc/v6r1m0/topic/com.ibm.wp.ent.doc/pzn/pzn_intro_likeminds.html

Further reading

Configuring WebSphere Portal for site analysis logging

http://publib.boulder.ibm.com/infocenter/wpdoc/v6r1m0/topic/com.ibm.wp.ent.doc_v6101/trouble/adsaconf.html

Understanding the site analysis log

http://publib.boulder.ibm.com/infocenter/wpdoc/v6r1m0/topic/com.ibm.wp.ent.doc_v6101/trouble/adsaundr.html

IBM WebSphere Developer Technical Journal: Using portal analytics with open-source reporting tools – Stefan Liesche and Steffen Uhlig

http://www.ibm.com/developerworks/websphere/techjournal/0609_liesche/0609_liesche.html

AWStats – Free real-time logfile analyzer to get advanced statistics (GNU GPL)

http://awstats.sourceforge.net

Nihuo Web Log Analyzer

http://www.loganalyzer.net/

WebTrends

http://www.webtrends.com/

NCSA Combined Log Format

http://httpd.apache.org/docs/2.2/logs.html#combined

Setting up AWStats on Ubuntu

Install Apache2 web server (using Synaptic Package Manager)

Apache automatically starts (runs as root) - can be stopped/started using **apache2ctl stop/start/restart** (again need to run as root so I used **sudo bash** to get a root shell)

Download AWStats (http://prdownloads.sourceforge.net/awstats/awstats-6.8.tar.gz)

Expand **awstats-6.8.tar.gz** into **/usr/local/awstats** (will need to rename expanded directory **awstats-6.8** or use **ln**)

Change directory to **/usr/local/awstats/tools**

Run the AWStats initial configuration process: -

perl /awstats_configure.pl

and enter the default Apache2 configuration filename e.g. /etc/apache2/httpd.conf when prompted

Choose to create a new configuration file and enter a suitable name e.g. the hostname of your machine and accept the other defaults.

This will generate a configuration file, **/etc/awstats/awstats.dmht60p.conf**, within which you'll see a reference to the log file that AWStats will be analysing e.g. **/var/log/httpd/mylog.log**.

Make a directory entitled **/var/log/httpd** and place the the **sa.log** file (generated by WebSphere Portal's SiteAnalytics service) into it as **mylog.log**

Run the following process to create the initial statistics database: -

perl /usr/local/awstats/wwwroot/cgi-bin/awstats.pl -config=dmht60p -update

Run the AWStats overview report generation process: -

perl /usr/local/awstats/wwwroot/cgi-bin/awstats.pl -config=dmht60p -output -staticlinks > awstats.dmht60p.html

Open the newly created/updated **awstats.dmht60p.html** in a web browser