# Web Analytics Dashboard and Analysis System

**Raiza Daniela Fernandez Palacios**
Department of Computing
Macquarie University
Sydney, Australia
daniela.fernandez@gmail.com

## Abstract

Modern web analytics tools collect vast amounts of information about website visitors; these reporting systems make it difficult to identify trends in data due to the number of reports available. By developing a system that logically presents automated analysis based on configuration patterns, web analytics users will be able to implement improvements to their websites. This paper presents a brief explanation of common metrics generated in Web Analytics tools and a summary of related work. Additionally, it describes the process followed to develop the tool across the design, implementation and test stages. Hence, it notes in conclusion that having an actionable dashboard linked to a suggestion module provides to the businesses key benefits to improve their sites.

## 1 Introduction

Currently, every business owning a site on the Internet and desiring to obtain a benefit from it is interested in measuring and analysing the website traffic to improve its performance. Due to the variety of metrics available, experts have found an excellent technique to condense and display the most important information in a single screen, commonly known as a dashboard. Paradoxically, despite the efforts of vendors of Web Analytics tools to present relevant and comprehensible data in a single screen, users continue having difficulties understanding the information quickly and making it practical to improve their sites.

Unsuccessful dashboards are the result of poor design techniques, which do not offer to the Web Analytics users meaningful metrics to help them to reach the objectives defined by the business in terms of site's performance. This fact and the need to have a successful solution for the users have been the motivation of the project.

This document exposes as solution a Web Analytics Dashboard and Analysis System (WADAS), software to automate the processing of web analytics data based on recognisable patterns. This tool combines two main components, a web analytics dashboard and a suggestion module.

WADAS will help users to measure their website's performance and take decisions based on the information displayed in the tool. During the research and analysis stage to develop WADAS we found that each dashboard is different from each other because they vary by business goals and department functions. On the other hand, we noticed how they can be classified in groups to display actionable data for diverse type of users (e.g, Marketers, Developers, CEOs…) [Few 2006, pp 175 – 178]. Hence, as was stated in the scope of the project, different dashboard designs have been created, but only one is implemented currently in WADAS.

In order to help the reader to identify the factors involved in the success and failure of Web Analytics dashboards and comprehend why the dashboard designs for WADAS were created like that, this paper will illustrate in Section 2 a brief background of Web Analytics and the main metrics employed in Web Analytics tools.

Section 3 presents a discussion about some related work to this project and the reasons to choose Google Analytics (GA) as the third party to collect the data for the system.

Section 4 is divided in two parts; the first one will show a Web Analytics dashboard designed for WADAS. The second part will explain the design for the suggestion module and the main purpose of this.

The techniques and process for the implementation of WADAS will be shown in Section 5.

Subsequently, in Section 6 the tests for the designs and results are discussed. Finally, conclusions and opportunities to work in future related to projects are presented.

## 2 Background

Web Analytics is a term used frequently in the e-marketing world. The history of Web Analytics is relatively short but the importance of this area has growth quickly because of the impact in businesses with presence on the Internet. According to Kushik [2007, pp. 2 - 5] the term of Web Analytics appeared when the server logs captured other information different to the number of visits in the websites; and then, it was necessary to define metrics. Later, vendors created tools to unify that information and display it in reports that could mean something for the users. Some of those companies are Coremetrics, Omniture, WebTrends, WebSideStory, Google, Microsoft and Yahoo.

Nowadays there are numerous metrics, but most of them are derived from others. There are some representative metrics that are always present in Web Analytics dashboards and can be considered as Foundational metrics [Kaushik, 2007, p4][Peterson, 2004, pp 106-114]. These metrics are: *number of visits, visitors, time on site, page views and hits or impressions*. However, the way they are measured can vary from tool to tool.

Understanding metrics and dimensions used in Web Analytics is essential to comprehend the role they play in the dashboard in order to help the business to evaluate its site's effectiveness. For example, most of the Web Analytics tools agree that *Visits* are calculated bearing in mind the period of time or activity executed when the visitor interact with the site [Burby et al., 2007, pp 9 -12]. Alternatively, *page views* are considered the number of times a page is loaded [Kaushik, 2007, pp133-140]. In a *visit* or *session* the pages can be loaded several times. Therefore, it is common to find the number of *page views* much bigger than *visits*. However, the measure of content in pages can be deeper because each page has a number of elements such as pdf files, images, videos and text that are requested to the web server every time an individual page is loaded. The number of requests made is counted as *hits*.

In terms of popularity evaluation, the last metrics are useful by themselves. Nevertheless, combined with other metrics such as *time on site, click through rate, bounce rate, exit rate, revenue per click* and of course *ROI (return of investment)* it is possible to identify which pages are causing visitors to leave the site, assess the loyalty, measure how effective is an ad [Sostre & LeClaire, 2007 p 28], among others.

Using descriptive source of data or dimensions that permit to filter, label and group numeric measures, it is possible to contextualize and give meaning to the metrics.

*Entry* and *Exit Page*, *Language*, *Medium*, *Keyword* and *Source* are some of the most common dimensions that allow us to distinguish the first and last pages viewed by a visitor, the primary language used in the visitor's browser, the type of referral to access the website (Search Engine, Twitter, Facebook, Email...), and the word or group of words used to find the site in a search engine.

There are other dimensions defined by the Web Analytics tools that permit the identification of patterns among the visits, such as: *Browser, Browser Version, City, Connection Speed, Country, Continent, Day of last Visit, Domain, Flash Version, and Platform or Operating System.*

As was mentioned before, the metrics isolated do not provide value for the performance improving process, this value is added when they are analysed together and in a context. However, the number of combinations using the metrics and dimensions is giant and could produce numerous reports for the user. It is at this point where the main problem makes its appearance. Which methods to use to show the information collected in a simple way and also help the web analytics user to take action over their websites to improve them? Dashboards are an answer for this problem.

## 3 Related Work

Because of the impact of Dashboards in Web Analytics, different specialists have spent time identifying a number of steps that people should follow with the aim of avoiding the frequent mistakes in designing a Web Analytics Dashboard.
A dashboard ought to be totally customizable; however, as is affirmed by [Brath, 2004], each dashboard has a unique design, and it is very difficult to fulfill all the user requirements. Because of all these tools (including the one that we designed) are created to be used by different kind of businesses they need to establish parameters

and be predefined in certain level. Between some of the common vendors, Yahoo provides a blank dashboard, which is an advantage, while Google and Microsoft have preset display layouts.

However, according to the Director of Data Insights in Yahoo [Mortensen, 2008], comparing the dashboard provided by the three vendors, Google is the one more actionable, and that has a better approaching of the space in the screen. Better, but not perfect, because none of them satisfies the first rule of a successful dashboard, which is to display the important data in a single screen.

Although the texture and graphics offered by Microsoft have a simple but nice use of colors and highlighting elements; the flexibility in contrast to Yahoo and Google is inferior. In addition, this tool ceased to be an option for the project because on the 19th of March, the adCenter AnalyticsBeta [Carson, 2009] announced the closure of the program.

Flexibility is an advantage of Yahoo over Google Analytics (GA); however, the last one offers a better way to understand and communicate the meaning of the metrics. Furthermore, GA has found new ways to display the data in a dynamic way which lets users have a new perspective in interpreting graphics. In addition, GA allows users to link information from other products in the same application. For all these reasons GA was chosen as the third party tool to obtain the data.

At the beginning of this project there was not a clear definition to collect the data from GA; because there was not an interface to get the data in line. Fortunately, the last 21st of April the GA API (Application Programming Interface) was launched as well as featured examples from customers that use the API for different aims. From these applications youcalc and Polaris Desktop Widget can be considered similar to WADAS; however, they do not offer a suggestion module nor a dashboard for specific areas. A good feature from youcalc in terms of functionality is the possibility to create your own dashboard adding gadgets to a website and accessibility in line to the GA account.

## 4 Design

The design of WADAS was divided into two components. First of all, following the best practices to create Web Analytics dashboards, bearing in mind the most common mistakes [Few, 2006, pp 48 -75] and based on related tools, we designed four dashboards. The second component of the design stage was the Suggestion Module which is only one for all the dashboards. Below both design processes are explained.

### 4.1 Web Analytics Dashboard

The business' mission and vision can be aligned, but the ways that each department in the organization work to achieve the objectives set are different. Bearing in mind this fact and the information available from the Web Analytics tool chosen (GA), the dashboards design was classified into four groups. The dashboards designed serve to fill the needs of the Marketing, Sales, IT – Design and Management department. Some of the information is exclusive for a type of dashboard, but the majority of that data is shared among the four dashboards. The main difference is the mode how it is displayed according to the purpose of each dashboard. For example, data associated with the Operating System and browser used to access the site probably is relevant for the IT – Design department, but not for the Sales department.

Below it is shown in the figure 1 the Web Analytics dashboard design for the sales department.

The main objective of this dashboard is that the sales user can identify the significant data expressed in the easiest way to allow the user taking decisions to improve the revenue. To achieve this goal, it presents data of the current revenue compared to the one in last periods, the top 5 products performance, the campaigns that are bringing more ROI to the business, percentage of goal completion per visit, average cost paid for each click in case of using AdWords and top 5 sources based on the revenue rather than visits as occurs in the IT – Design's dashboard.

### 4.2 Suggestion Module

Perhaps, one of the problems faced by the users is that some of the Web Analytics tools are designed for experts in the area but are offered to both types of users, with basic and advanced acknowledge. Hence, basic users start using software with numerous options, tools and functions that are not familiar to them, causing frustration. Such frustration and dissatisfaction is more visible when the results shown in the reports depend largely in the tool configuration and the adequate use of the options available.

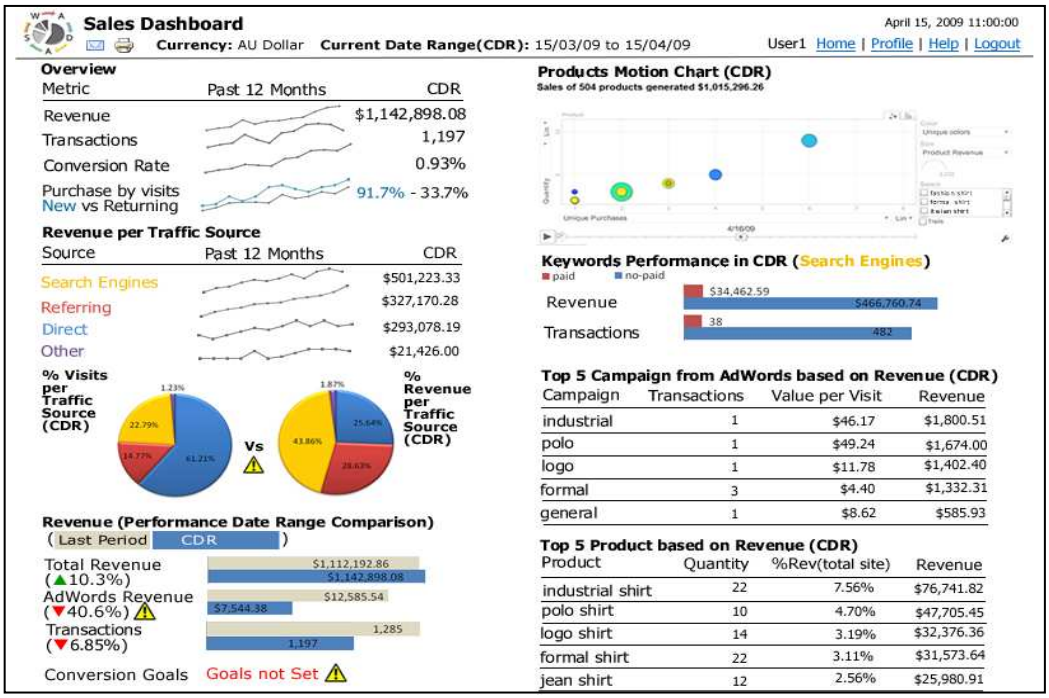


*Figure 1. Sales Dashboard*

Assuming that the users have a basic acknowledge to interpret the metrics shown in the dashboard or at least know the purpose of their website, the suggestion module does not desire to teach the users to be better sellers, developers, designers, or managers.

This module aims to guide the user through the Web Analytics tool (in this case GA) to obtain more clear and accurate results in the reports and consequently make right decisions to improve the website performance

An example of the results shown in the suggestion module is:

- *You have not set goals for this profile. Goals can help you to know how many visitors are completing your goal. (e.g. average of visitors that have purchased x product).*
- *You should set your default page in your profile settings to avoid distinct entries in the report referencing the same page (e.g www.mysite.com and www.mysite.com/index.php can be the same but are shown twice because the default page 'index.php' is not set).*
- *The campaign 'Camp 1' is not generating any revenue since it was created, you should consider make changes for that.*

Since the suggestions affect the results for the entire website, this is the same for each dashboard displayed in a specific profile.

## 5 Implementation

The implementation for the tool involves different elements since the registration of the user to the application. Although for the actual operation of WADAS has been implemented some functions related to the user management (e.g registration, deregistration, login, logout, profile updating and permission assignation), in this paper is explained the implementation related to the process of data to generate the Dashboard and Suggestion Module.

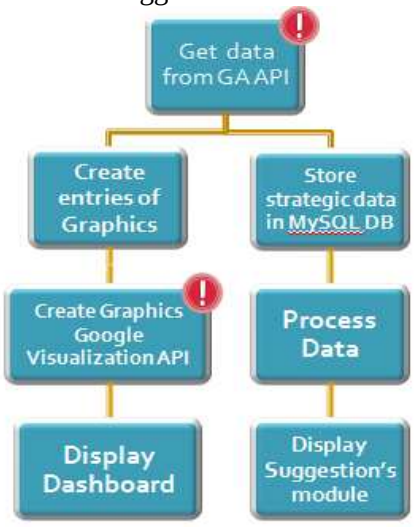


*Figure 2. General Process Scheme*

The figure 2 shows the main processes involved from data collection until the dashboard and analysis system are displayed to the user.

As you can see in the figure, there are two processes highlighted (!). The reasons to make emphasis in these two processes are that they depend in a third party and have a level of complexity bigger than the others. Hence, any change made in the Google Analytics API or the Google Visualization API can affect the operation of WADAS.

The following is an explanation in more detail about the process to generate the Dashboard and suggestion module:

**I.** ***Get data from GA API:*** The information is collected from the GA account subscribed. A GA account regularly has more than one profile, which can contain filters that affect the data displayed; therefore, the data collected needs to be classified by profiles and account.

The information received by the GA API has two types of feeds, one associated with the account information and the other with the actual data. PHP is used to read and decompose the information in the feeds. Here the process is divided to create the Dashboard and the Suggestion module.

***I.1. Dashboard***

I.1.1. Create Entries of Graphics: According to the type of user (Sales, IT, Marketing, CEO) the data is formatted to create the entries for the graphics used with the Google Visualization API. Until now this API only supports two ways to send a request to the data source, the first one is using Google Spreadsheet and the other with JavaScript code [Google Code, 2009]. In our case, the second option is used. Therefore, using the feeds with PHP and JavaScript the table object called *DataTable* is created.

I.1.2. Create Graphics Google Visualization API: This process consists in the use of the methods that define the type of graphic (e.g PieChart, Table, Barformat) and the *draw* function. To the second method are passed as parameters the *DataTables* defined in the step I.1.1, and a number of options that help to define the graphical appearance (e.g background, width, height, 3D option).

I.1.3. Display Dashboard: Finally the graphics, tables, labels and other components are placed in the dashboard and displayed to the user.

***I.2. Suggestion Module***

I.2.1. Store strategic data in MySQL DB: Some of the data collected from the GA API needs to be stored in a database because it requires complex and historical data.

I.2.2. Process Data: In this step variables such as the configuration of goals in the account, default page, ecommerce option, campaigns and keywords are evaluated.

I.2.3. Display Suggestion module: In a separate screen to the dashboard is displayed the suggestions that the user should follow to improve the analysis of the performance for the website.

## 6 Tests and Results

In order to test the designs were created four (one for each type of dashboard) mockups in paper. Each mockup was presented sequentially with an additional page. That page consisted of multiple choice questions to assess that the evaluators understood the data displayed in the dashboard, and a block for final comments to allow the evaluator give feedback about the dashboard and provide ideas. The group of 10 people that responded the test was composed of Web Analytics experts, designers, developers and others, which were chosen because they have a site on Internet. The time spent in each dashboard to answer the questionnaire was measured.

Trying to follow a heuristic evaluation [Kaushik, 2007, p 57], part of the group that took the test was able to participate in a feedback meeting about the dashboards. Thus, we could evaluate different point of views and discover new observations.

There were some limitations in this process. Because of the short time to implement the project, the dashboards were not implemented; therefore, some of the functionalities could not be tested. For example, each dashboard includes a motion chart and this is a dynamic chart for multi-dimensional analysis. Other interactive options with the chart, change of parameters and drill-down of data could not be tested either in this stage.

On the other hand, this fact helped to test simplicity, synthesis, use of space and identification of key metrics.

As results from these tests, most of the answers for the multiple choice questions were right. In contrast, the most difficult part to eva-

luate and that resulted in design changes were the comments made by the users. Some evaluators provided comments about the colors, type of graphs and signs used in the dashboard, suggesting new alternatives.

Additionally, a pattern found in the respondents was that after the first dashboard assessed, they spent less time answering the multiple choice questions for the others. It could be due to the design consistency across the dashboards. Hence, it seems like they spend some time in the first dashboard learning how to read it, time that was saved it in the next ones.

For the implementation have been done internal tests in order to ensure the data displayed is coherent with GA. For the user tests were chosen two people with an account in GA. Some changes have been implemented in the dashboards after the user tests with real data. The changes involved restrictions to display information when the GA accounts are not configured to display e-commerce data (e.g revenue, transactions and products) and goals.

On the other hand, the tests for the suggestion module require more time because are prepared for people with key skills. The main difference with the dashboard design tests is that to assess the results from the suggestion module it is required acknowledge of web analytics professionals in GA.

## 7 Conclusion and Future Work

This document described the process followed to create an application aimed at helping Web Analytics users to make strategic decisions in order to improve their website's performance. Although the tool probably will not resolve the site performance troubles, it will provide a summary status and will point users to do some actions to solve the problems.

The scope of the project was bound to the implementation of one of the dashboards. It allowed us to evaluate in real scenarios the design and how actionable was the dashboard. The primary future work will be the implementation of the other three dashboards in WADAS.

The suggestion module does not consider information from other sources that can be linked to GA, such as Google Website Optimizer, Google ad Sense and Google AdWords. A future project could include suggestions based on specific information that is not retrieved using only the GA API and that could be achieved adding the Google AdWords API for example.


### Acknowledgments

Thanks to Michael Johnson and Robert Date for their guidance and feedback provided in this project. I would also like to thank the participants in the dashboard tests for the qualitative evaluation.


## References


Brath R. & Peters M., 2004, *Dashboard Design: Why Design is Important*, InfoManagement Direct, 15 October 2004, viewed 18 March 2009, http://www.information-management.com/infodirect/20041015/1011285-1.html

Burby J., Brown A. & WAA Standards Committee, 2007, *Web Analytics Definitions Version 4.0*, 16 August 2007, viewed 10 March 2009, *http://www.webanalyticsassociation.org/attachments/committees/5/WAA-Standards-Analytics-Definitions-Volume-I-20070816.pdf*

Carson M., 2009, *adCenter Analytics Beta to Close*, Microsoft adCenter Analytics Beta, 12 March 2009, viewed 13 March 2009, http://www.adcentercommunity.com/blogs/analytics/archive/2009/03/12/adcenter-analytics-beta-to-close.aspx

Few S., 2006, *Information Dashboard Design: The Effective Visual Communication of Data*, O'Reilly, Sebastopol, CA.

Google Code 2009, *Google Visualization API*, 21st April 2009, viewed 23 April 2009 http://code.google.com/apis/visualization/documentation/queries.html

Kaushik A., 2007, *Web Analytics: An Hour a Day*, Wiley Publishing Inc, Indianapolis, USA.

Mortensen D., 2008, *A Reflection on Google, Microsoft and Yahoo Web Analytics Dashboards*, 11 September 2008, viewed 09 March 2009. http://visualrevenue.com/blog/2008/09/google-microsoft-yahoo-web-analytics-dashboards.html

Peterson E., 2004, *Web Analytics Demystified: A Marketer's Guide to Understanding how Your Web Site Affects Your Business.* Celilo Group Media.

Sostre P. & LeClaire J., 2007, *Web Analytics for Dummies.* For Dummies, USA.