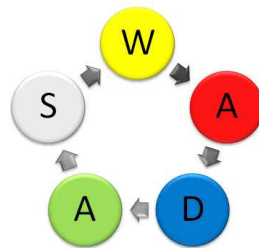


Web Analytics Dashboard and Analysis System

Raiza Daniela Fernandez Palacios

Student id 41371518

Final Project



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. Several vendors have focused on dashboards as a mechanism to provide the most relevant information to the users to make it easier for businesses to analyse the metrics exposed and improve the effectiveness of their websites. Unfortunately, not all dashboards have a positive impact because of failures in their designs. Based on our study and user testing, we have developed a Web Analytics Dashboard and Analysis System (WADAS), which logically presents a dashboard and automated analysis of configuration patterns for a third party Web Analytics tool. Hence, users will be able to implement improvements for their websites. This dissertation studies the main metrics used by Web Analytics and the best practices to design dashboards; allowing the reader to identify what should be done and what should be avoided in Web Analytics dashboard designs. Additionally, this document presents the software development process including: requirements analysis, design, implementation and discussion of observations noted after the tests. We found that WADAS provides an actionable dashboard and facilitates the familiarization of advanced configuration options in Google Analytics.

Acknowledgments

I would like to express my gratitude to Michael Johnson and Robert Dale for their outstanding supervision, guidance and feedback provided in the development of this project. I owe much to Benjamin Mangold and Michael Mangold for giving me advice and allowing me to access resources such as books, forums, papers and tutorials that made the completion of this work easier.

I am also indebted to many people that participated in the dashboard tests and to those with whom I have discussed my project. Their enthusiasm, ideas and qualitative evaluations have been invaluable.

Table of Contents

Abstract.....	2
Acknowledgments.....	3
1 Introduction	8
1.1 Motivation.....	8
1.2 Character and Organization	9
2 Foundational Concepts	10
2.1 Background and Related Work	10
2.3 Measurements Used in Web Analytics	11
2.3.1 Metrics	12
2.3.2 Dimensions.....	14
2.4 Web Analytics Dashboards	15
2.4.1 Best Practices to Create Dashboards	15
2.4.1 Common Mistakes	17
3 User Requirements and Specification.....	19
3.1 Software Requirements Specification.....	19
3.1.1 Functional Requirements.....	19
3.1.2 Design Requirements.....	21
3.1.2 Non – Functional Requirements	21
3.2 Requirements Analysis.....	22
3.2.1 Actors	22
3.2.2 Functional Areas	23
4 Design Phase	25
4.1 WADAS System	25
4.1.1 Model Component	25

4.1.2	Process Control Component	26
4.1.3	System Interface Component	26
4.2	Web Analytics Dashboard	28
4.2.1	Sales Dashboard	30
4.2.2	IT – Design Dashboard	33
4.2.3	Marketing Dashboard	35
4.2.4	Management Dashboard	37
4.3	Suggestion Module	39
5	Implementation Phase	41
5.1	General Implementation of WADAS System	41
5.1.1	Security Issues.....	42
5.2	Dashboard and Suggestion Module Implementation.....	44
5.2.1	Get data from GA API.....	45
5.2.2	Dashboard.....	46
5.2.3	Suggestion Module	48
6	Test Phase and Results.....	50
6.1	Design Tests	50
6.1.1	Limitations.....	50
6.1.2	Results	50
6.2	Implementation Tests	52
6.2.1	Security Issues.....	52
6.2.2	Requesting and Retrieving Data.....	53
6.2.3	User Interaction and Interface.....	53
7	Conclusions	54
	Summary	54
	Future Directions	55
	References	56

Appendix	58
Appendix A – Main Use Cases Definition.....	58
Appendix B – Google Authentication Process (AuthSub)	61
Appendix C – Questionnaires.....	62

Table of Figures

Figure 1 Finding patterns – Dots vs Numbers.....	16
Figure 2 Pie chart comparison	18
Figure 3 User Management – Use Case Diagram	23
Figure 4 Data Reporting Services – Use Case Diagram	24
Figure 5 Model Packages	25
Figure 6 Controller Packages	26
Figure 7 View Packages - Interface	26
Figure 8 User Flow Diagram.....	27
Figure 9 Header Section.....	29
Figure 10 Content Section.....	29
Figure 11 Footer Section.....	29
Figure 12 Sales Dashboard.....	33
Figure 13 IT – Design Dashboard	35
Figure 14 Marketing Dashboard	37
Figure 15 Management Dashboard	39
Figure 16 Suggestion Module example.....	40
Figure 17 Implementation Process Scheme	44

1 Introduction

E-marketing trends change continuously, and to maintain successful businesses on the Internet, website owners need to monitor and analyse their visitors in order to create effective marketing strategies. Businesses make use of Web Analytics tools to know more about their website such as, what is failing, which pages are the most visited, what are the trends, and of course what do their clients (visitors) want.

Without Web Analytics tools businesses cannot have real knowledge of their web traffic. In contrast, identifying what visitors are looking for when they navigate through the website and cross matching that information with the business goals makes it easy to design successful marketing strategies.

1.1 Motivation

Currently, every business owning a site on the Internet and desiring to obtain benefits 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 that do not offer the Web Analytics users meaningful metrics to help them 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 motivators of this project. Thus, with this document we aim, first of all, to expose the factors involved in the success and failure of Web Analytics dashboards. We also propose as solution a Web Analytics Dashboard and Analysis System (WADAS) 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 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 this project, different dashboard designs have been created, but only one is implemented currently in WADAS.

1.2 Character and Organization

Each dashboard makes use of several metrics that are calculated after collecting data from the website's visitors. In order to help the reader identify the factors involved in the success and failure of Web Analytics dashboards, it is necessary to have knowledge of the metrics and descriptive source of data employed by the Web Analytics tools in their dashboards.

Chapter 2. This chapter illustrates the foundational concepts of Web Analytics Dashboards. It is divided into three sections. In the first part, we provide a brief background, outline of the role of Web Analytics and related work. Through the exploration of different alternatives available in the market, we present a discussion about the reasons to choose Google Analytics (GA) as the third party tool to collect the data for the system. Then, we explain the evolution and the importance of the metrics as part of the foundational concepts used in Web Analytics for strategic decisions. In addition, this section contains the synthesis of the most relevant definitions about Web Analytics metrics found in several sources.

In the third section of Chapter 2, we consider best practices and common mistakes in creating Web Analytics dashboards based on different resources and comparisons between analytics tools. This will provide the opportunity to explore significant improvements in graphic design techniques, identified by experts, to create dashboards.

Chapter 3. Because of the outcome of this project was to produce a computer tool, software engineering procedures were applied to our research to develop a functional Web Analytics dashboard. Chapter 3 enumerates the functional and specific requirements of WADAS.

Chapter 4. The design of the application, which involves the Dashboard and the Suggestion Module, are shown in this chapter. Although the outcome is a single application, the two elements that compose it are very different in terms of design. For this reason, they are explained independently. Section 4.1 describes briefly the elements involved in the WADAS system design phase. Section 4.2 illustrates the four Web Analytics dashboards designed for WADAS, and Section 4.3 gives details of the design for the Suggestion Module and the main purpose of this.

Chapter 5. In this chapter, the techniques and process for the implementation of WADAS are studied. WADAS is implemented using different programming languages and application programming interfaces. This combination of programming paradigms and languages involve security and performance issues that we discuss in this chapter.

Chapter 6. Tests were performed to assess the effectiveness of the dashboard designs and the application. Therefore, the tests and results are explained in Chapter 6.

Chapter 7. In this, the final chapter, we present some concluding remarks and future opportunities for related projects.

2 Foundational Concepts

2.1 Background and Related Work

In the last 15 years the use of the Internet for marketing has increased significantly because of the small amount of money needed to invest and the short time to achieve a return on that investment. This fact has caused the growth of business competition and as a consequence people owning websites have become more concerned about improving their site to at least match the level of their competitors. Web Analytics tools were created to analyse and report data collected from websites visitors.

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 grown quickly because of the impact in businesses with a presence on the Internet. According to Kaushik [2007a, pp. 2 - 5] the term 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 for the users. Some of those companies are: Coremetrics, Omniture, WebTrends, WebSideStory, Google, Microsoft and Yahoo.

The impact of dashboards in Web Analytics has led different specialists to spend time identifying 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 Brath & Peters affirm [2004], each dashboard has a unique design, and it is very difficult to fulfill all user requirements. All of these tools (including the one that we designed for different kind of businesses) need to establish parameters and be predefined to certain levels. Yahoo for example 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 more actionable one and has a better approach to the space on 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 colours and highlight 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 Yahoo has over Google Analytics (GA); however, the latter 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. Among the new features are the motion charts, which combining observation, dynamic objects and creativity, allow the users to perform an advance data analysis [Gillis, 2008][Fernandez, 2009]. In addition, it has been easier to find documentation about the metrics, dimension, reports

and other concepts implemented in Google Analytics than in the other tools. Finally, 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 Web Analytics data.

At the beginning of this project we did not have a clear definition to collect the data from GA because there was not an interface to get the data in line. Fortunately, 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 launched YouCalc¹ and Polaris Desktop Widget² can be considered similar to WADAS in that they intend to make the most important Web Analytics data available. Polaris is a desktop application that makes use of the GA API providing standard reports with the aim of maintaining the information available from GA at any moment. Because it is a desktop application, it needs the Polaris software to be downloaded, and some technical specifications associated with the Platform and Operating System are required.

YouCalc, on the other hand, is a web application and therefore, does not require any particular installation as Polaris does. One of the big advantages on this tool is the flexibility to create our own dashboard in iGoogle³ making use of their modules (called apps). However, it fails for users who are not experts because it does not define which metrics should be displayed to have an actionable dashboard.

As WADAS, YouCalc does not store data related to the authentication and authorization credentials of the user GA account. However, unlike WADAS, every time that the user requests for a report (apps) needs to log in. As many times that the report is closed and opened the user needs to sign in with the Google Account although the session in YouCalc is still active. Another feature that makes WADAS different from other tools is the fact that it offers a Suggestion Module and dashboards for specific areas.

2.3 Measurements Used in Web Analytics

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. There are currently 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, 2007a, p4][Peterson, 2004, pp 106-114]. These metrics are: *number of visits, visitors, time on site, page views and hits or impressions* [Sostre & LeClaire, 2007, p19]. However, the way how they are measured can vary from tool to tool.

Generally, simple count metrics not related to the Key Performance Indicators (KPIs) of the organization are considered *content metrics*. Levene [2007, p177] states that these kind of metrics are useful to measure the quality of the site because they provide information about the top entry or exit pages, number of repeat visitors and time of sessions. On the other hand, count and ratio

¹ <http://www.youcalc.com>

² <http://www.desktop-reporting.com/polaris.html>

³ See for more information <http://www.google.com/support/websearch/bin/answer.py?hl=en&answer=20324>

metrics related to the business objectives are considered *commerce metrics* because they represent the conversion of visitors to customers, such as the proportion of visitors that make a purchase (conversion rate).

In a more detailed classification, the Web Analytics Association (WAA) [Burby et al., 2007] divides the metrics into three types. These types are Count, Ratio and KPI where the last two are the result of several calculations more than only addition, contrary to the Count metrics. All these metrics are measured using dimensions. This term makes reference to a descriptive source of data that permit to filter, label and group numeric measures. Using this classification and purposing as three main metrics - *Unique Visitors, Visits/Sessions and Page Views* - the WAA defines other metrics derived from them, which will be illustrated later.

Another point to underline from the last categorization is the fact of defining KPI as a type of metric. According to the classification, a KPI is a metric, but it does not mean that a metric is always a KPI. That is why, even though number of *Visits* and *Page Views* are one of the main metrics, they are not part of the business metrics; they only provide information about the traffic in the website [Burby and Shane, 2007]. The major difference between KPI and the other metrics is that the first one reflects business goals, provides context and meaning, is easy to understand, is actionable, and the management department has knowledge about its existence. Distinguishing these concepts is crucial because they are involved in the steps to create successful dashboards, which are explained in the Section 2.4.

Perhaps this classification was necessary due to the number of metrics and dimensions that vendors started to employ in the reports. Currently, there are more than 20 metrics defined [Kaushik, 2007b] that are used in the most common Web Analytics tools, and the level of importance of them is strongly associated with the interests and goals of the particular business.

2.3.1 Metrics

The next is a selection of the metrics and dimensions included in Web Analytics tools and defined in different sources [Burby et al., 2007][Peterson, 2004, pp. 45 – 57][Sortre & Le Claire, p 18 – 28]. We use an example similar to the real world with the aim of making the reader to understand easier the concepts for some metrics.

Scenario: *The website www.example.com was launched three months ago. This site is composed of 3 pages, the main one or landing page, the contact us page which displays an online form and the thank you page. The landing page contains 3 pictures and a pdf file. In the last three months the site has been visited 250 times, from this 250 times, 100 are from different sessions. The Web Analytics user has found that not many people fill the form in, even when they reach the contact us page. According to the statistics, 60% of the visits are new, but they do not spend in average more than 2 minutes on the site.*

- a. *Page Views (Count):* It is defined as the number of times a page is loaded [Burby et al., 2007 p7]; thus, it can be used to determine the popularity of the website. *For the scenario, the*

number of page views is 250 because it is the number of times that different pages in the site were visited.

- b. *Visits (Count):* This is also known as sessions as well because it keeps in mind the period of time or activity executed when the visitor interact with the site [Kaushik, 2007b]. Hence, if a visitor does not perform any activity in a period of time (defined by the Web Analytics tool), the visit session ends. Correspondingly, if the visitor executes an activity after the period of time it is counted as a new visit. *Because of 100 are the different sessions in the example, the number of visits is 100 instead of 250.*
- c. *Unique Visits (Count):* Each individual is counted as a unique visitor. Due to authentication not being a requirement for all websites, the most common technique used by the Web Analytics tools to count *unique visits* is via cookies that are stored in the user's computer [Burby et al., 2007, pp.9 - 12]. The only problem with this method is that users can delete cookies and the number of *unique visits* will be over counted. New, *repeat* and *return visitors* can be derived from this metric [Tyler, 2006, pp 173 - 180].
 - *New visitor:* All of unique visitors that are accessing the site by first time. *In the scenario exposed, 60 visits are from new users.*
 - *Repeat visitor:* All of unique visitors that are accessing the site for more than once in a period of time specified. *Assuming that the period of time specified for the report are the three months, the number of repeat visitors is 40.*
 - *Return visitor:* All of unique visitors that are repeat visitors and that also have visited the site before the period of time specified in the report.
- d. *Hits (Count):* It is considered as all requests from a web server or file (PDFs, Excel, images, text and graphics). Thus, as a visit to a page from the site can generate 8 hits, 1000 visits could count more than 10000 *hits*. Although that number is big and can be useful to know the amount of content in individual pages, it is not very representative for taking strategic decisions. Sostre & Le Claire [2007, p23] consider this metric as the most deceptive because many people tend to confuse it with the number of *visits* or *page views*. *In our example, the number of hits for the landing page is 1250 that corresponds to 5 times 250 page views (3 pictures, 1 pdf and 1 called to the html every time the page is loaded).*
- e. *Time on Site:* It is basically the duration of the visit. Normally, it is calculated by subtracting the timestamp of the last activity and the first activity of the session [Burby et al., 2007 p 17]
- f. *Click Through (Count):* Represents how many times a link was clicked by a visitor.
- g. *Click Through Rate (CTR) (Ratio):* It is used to know the rate at which an ad is clicked compared to the times that is viewed [Sostre & LeClaire, 2007 p 28].
- h. *Bounce Rate (Ratio):* This is the percentage of entrances to the site that has left resulting in the exit from the website. Although the GA API does not provide this metric, we calculate it

dividing the single *page view* visits by the *entry pages*. This is one of the most representative metrics if it is analysed with the actual *exit pages* because it identifies which pages are causing that visitors to leave the website. *If there were 100 visits and 20 bounces, the bounce rate is 20%.*

- i. *Exit Rate (Ratio)*: Normally, people confuse bounce rate with exit rate. The main difference is that the exit rate divides the number of exits by total number of *page views* without bearing in mind the length of the visit [Burby et al., 2007 p 28].
- j. *Page View per Visit (Ratio)*: Calculated with the metric of page view divided by the total number of visits in a period of time specified.
- k. *Return of Investment (ROI - KPI)*: This is a metric that allows assessing the efficiency of an investment. It is calculated by dividing the difference between the investment and the cost of this, by the cost of investment, but it is expressed as a percentage [Clifton, 2008, p 228]. With this metric it is possible to measure how much time the business should spend in Web Analytics, keeping the business goals as priority.
- l. *Revenue Per Click (KPI)*: This metric identifies the income from the number of clicks caused by access to the site [Teixeira, 2008]. However, this metric only provides information when the click comes from one ad (e.g. advertisement payed to any search engine to show the name of the website and Uri in the section of sponsored links).

2.3.2 Dimensions

The metrics explained in 2.3.1 are very useful **only** if they are applied in context because the numbers by themselves do not mean anything. That meaning can be reached across the dimensions. Below we describe the most common dimensions used in Web Analytics reports for segmenting data.

- a. *Entry Page*: As the name indicates, it is the first page viewed during a unique visit.
- b. *Exit Page*: The last page viewed by a visitor before leaving the site.
- c. *Language*: The language provided by the HTTP request for the browser to determine the primary language used by the visitor.
- d. *Medium*: The type of referral used to access the website, for example if it was accessed directly, or found using a search engine (Google, Yahoo...).
- e. *Organization*: Name of the service provider for accessing the Internet.
- f. *Keyword*: Word or group of words used to find the site in a search engine. This dimension is very representative for people that define campaigns and pay for clicks in their ads.
- g. *Referrer*: It is the page URL that was visited and generated the visit to the website [Burby et al., 2007, p 18]. When the user accesses this field directly is displayed as "not set". Each referrer is classified in a *medium*.

There are other dimensions defined by the Web Analytics tools that enable patterns of the visits to be identified, such as: *Browser, Browser Version, City, Connection Speed, Country, Continent, Day of last Visit, Domain, Flash Version, Java Enabled* (if the browser allow Java elements), and *Platform or Operating System*.

As was previously mentioned, the metrics isolated do not provide value for the performance improving process, this value is added when they are analysed together and in one context. Gibbons [2008] conveys some ideas of the advantages that can be reached merging some simple metrics. For instance, the effectiveness of the landing (main) page, keywords relevance and traffic quality from different sources can be assessed combining the *bounce rate, type of sources* and *visit duration* metrics. Nevertheless, the number of combinations using the metrics and dimensions is gigantic and could produce numerous reports for the user.

It is at this point where the main problem makes its appearance. Which methods should be used to show all the information collected in a simple way and also help the Web Analytics user to take action over their websites in order to improve them? Dashboards are a solution to this problem.

2.4 Web Analytics Dashboards

Few [2005, p1] defines a Dashboard as “a visual display of the most important information needed to achieve one or more objectives consolidated and arranged on a single screen so the information can be monitored at a glance”. In addition, he states that the goal of a dashboard is communication. Good communication requires unambiguity, which is one challenge of a successful dashboard design. To extend this definition, Eckerson [2005, pp 13- 15] also identifies monitoring, analysis and management as the three main applications in dashboards. These applications allow the users to monitor operational processes, analyse the large amount of information and guide the business in the right direction to improve the performance of the sites.

In a interesting way, Eckerson [2005 p XIII, 2006] presents a more detailed explanation of the performance dashboard’s concept as a puzzle of three layers and types that are associated with the three components mentioned above. In essence, he defines the monitoring, analysis and management layers as a sequential process performed by the businesses when any of the three types of performance dashboards (operation, tactical and strategic) are used. However, not all dashboards are compliant with these features; therefore, they have a negative effect in the performance of the business. As we mentioned before, experts have spent time identifying the best practices to follow when designing dashboards and the common mistakes that should be avoided.

2.4.1 Best Practices to Create Dashboards

Web Analytics Dashboards are used to understand and improve the performance of the business through the site usage. According to Saiffee[2005, p3] a dashboard should be intuitive, customizable, interactive and actionable, and part of the business intelligence strategy. Following this statement Vermehren [2008] in one of his articles proposes six steps to follow when designing a dashboard, steps that are aligned to the features exposed by Saiffee. Below we synthetise the steps suggested by Vermehren [2008] combined with related observations from other authors.

The first thing we should do when creating a dashboard is to define its purpose and to recognize which kind of dashboard should be designed according to the needs of the group of users that will use it. The second step consists in the selection of relevant metrics and KPIs. Vermehren [2008] classifies the Web Analytics KPIs in five groups: Marketing, Engagement, Usability, Conversion and Loyalty. Marketing KPIs are associated with number of *visits* and *traffic source* (e.g search engines, emails, and newsletters). Engagement, Usability and Loyalty KPIs can be measured with the information collected about the *time on site*, number of *new* and *return visitors*, and the actions that visitors execute when they navigate through the entire site. In contrast, Conversion KPIs are focused on specific pages that guide the visitors to become customers.

Each one of the last groups can be used by numerous metrics, but not all of them are relevant to quantify the performance of the site. Therefore, the critical task in this step is to select a group of important metrics to measure the organizational goals, then make effective use of the space to display the data is the next step.

Studies completed by different experts such as Tidwell [2006] have demonstrated that the human being can understand visual forms faster than numbers and that there are some variables that affect this result: Position, Brightness, Orientation, Saturation, Size, Texture and Shape. To illustrate this assumption, *figure 1* is an example from Tidwell [2006, p 279] where it is possible to confirm how fast the human eye responds to images. For the both images, if you try to pick up the items that have a different pattern from the others, the time spent in the square composed by numbers is longer.

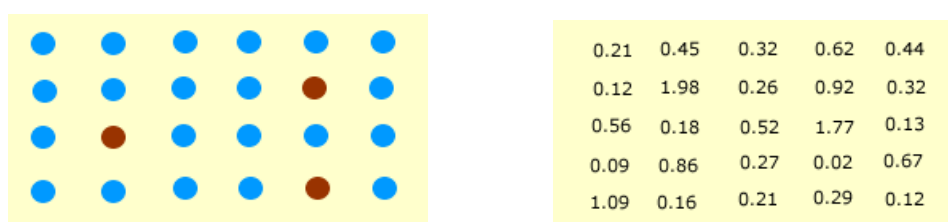


Figure 1 Finding patterns – Dots vs Numbers

After asking seven different people to find the different items in each one of the images, all of them spent less than 4 seconds to figure out that there are three red spots in the first image, while in the second image the average was 45 seconds to identify three numbers that start with 1 instead of 0.

Although graphics communicate a message faster, the main goal in this step is to use different colours, images, shapes, charts and graphs in the *healthiest* way allowing the user to have a clear and fast understanding in a single screen. Being more specific, Kaushik [2007c] affirms that all the data in a dashboard must fit in a single page (A4) and be understandable. Evidently, this step could be itself the most difficult because it requires creativity, and design skills without losing clarity about the objectives defined.

Done this, other resources should be used to expand the information. It can be achieved with special highlights, icons and underlines. An important contribution from Kaushik[2007c] is the idea to include a section for insights and recommended actions in the dashboards because it allows

identifying causes, effects and possible next steps. We could consider this idea as part of the motivation for the Suggestion Module in WADAS.

The last step in the process of designing a Web Analytics dashboard is the addition of interactivity. In other words, the dashboard should allow the user to customize some features related to order of data, view of some detailed information, filters, number of charts displayed and KPIs.

2.4.1 Common Mistakes

Despite the codification of steps required to create a dashboard, experts have emphasized the common mistakes in dashboard designs. Two authors that have worked in depth on this topic are Few [2006] and Kaushik [2008]. The first author affirms that one of the causes of cluttered dashboards is that vendors focus on displaying nice widgets full of bright colours and sophisticated shapes, missing the objective of showing key information to the user. In his book *Information Dashboard Design*, Few [2006, pp 48 -75] classifies thirteen mistakes in dashboard designs that combined with the characteristics of a well designed dashboard and the suggestions provided by Kaushik [2008] lets the reader have a better idea about the process to create dashboards, bearing in mind structure and visual features (what should be done and what should be avoided).

At this point of the document, we have made clear that a proper dashboard should not exceed the boundaries of a single screen. Ironically, this is the most common mistake made, and it becomes an error because navigating through more than one page is often the cause of missing significant data. Displaying data in an inadequate context and with excessive detail or precision are other two errors that people tend to make. In fact, in Chapter 6 we indicate this type of error in our Web Analytics dashboards and how it was fixed.

Similarly, using bad scales can result in perception illusions that lead to misunderstand the data. An example of this is when zero or negative values are not included in the scales when they should, to be compared objectively with others.

Showing a metric that does not clearly express what the user wants to measure is another mistake. Percentages are good notations to show levels of performance; however, they are not the answer for everything and sometimes the actual value of loss or gain has more meaning than a single percentage. This means, choosing efficient metrics needs to take into consideration the element to be measured and the units to express the metric. In the same way, choosing the wrong type of graphic to display the data can confuse the user.

Bar, Line, Pie charts, tables and maps were designed for different purposes. Thus, not all information is successfully shown in a specific chart. Pie charts are highly used because users are familiar with them. They are good when the information represented is part of a whole (100%), otherwise they can be confusing. Sometimes, displaying the information can be boring but using arbitrary display media is even worse. The key is the combination of few display media and the right data.

Even when the display media chosen is the right option, the way it is used determines its usefulness. A pie chart well designed can have a positive impact in the dashboard, for example. In contrast, pie charts designed using a colour palette from the same basic colour (e.g. different types of blues)

representing more than five elements can be quite confusing. In figure 2, for the first pie chart the human eye requires more time to identify what data match with the piece of the pie. Also, it is missing the percentages, which means that even identifying which colour represent the data the only possible assumption is to identify that one is bigger than others.

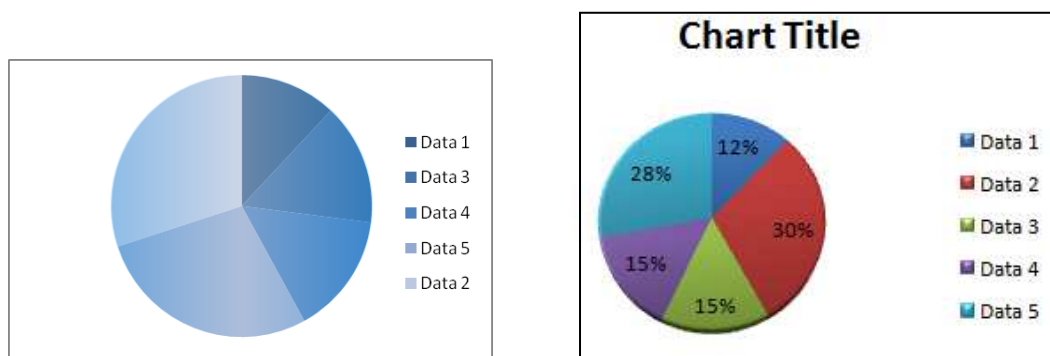


Figure 2 Pie chart comparison

The position of the charts, text, title, logo and other elements are crucial to ensure that the most important data is evident to detect. Generally, left, top and middle side are used to place important information because those are the first sides the human eye scans when looked at the screen. However, many Web Analytics vendors reserve a considerable part of the top or left side to place the logo of the company, space that ought to be used to display critical information.

In the same way, the decoration is good if it has meaning and it does not interfere with the purpose of the dashboard. The risk of using decorative elements in a dashboard is a distraction or worse, it can result in wrong interpretation of the data. Amongst the examples provided by Few [2006, p72], there is one really remarkable - where an electronic control panel is the graphic design base of the dashboard. Although the design looks ingenious, the actual information is overshadowed by the design.

Another mistake is to highlight important data ineffectively misusing or overusing the colours. Although it is necessary to highlight the critical information for decision making, this process must be considered strategically. Otherwise, other data that supports the critical one could lose meaning and become unusable. Colours play a vital role in the success of the dashboard. Hot (e.g. red, orange and dark yellow...) or cold (e.g. blue, green and purple..) colours can influence the actions of the observer in a positive or negative way. Normally, hot colours are used to highlight critical points; however, using them too much can make the important elements ineffective.

Finally, the creation of an unattractive visual display is an error that results in the prevention of using the dashboard. The design should be tested thoroughly before it is launched. These tests involve the levels of comfort and clarity with which the message is communicated.

3 User Requirements and Specification

We carried out a user requirement analysis to design and implement the tool. We asked different Web Analytics users what they would like to find in an application that is not provided by their Web Analytics tool. Particularly, we followed the requirements specified by the Business director of Mangold Sengers⁴, one of the Google Authorized Consultant companies in Australia.

3.1 Software Requirements Specification

We have classified the system requirements in three types. Functional requirements that are related to the application structure where we define the logic of high-level functions provided by WADAS. We enumerate the design requirements that involves rules for display the elements in WADAS and then, we present other technical specification as not – functional requirements of the system.

3.1.1 Functional Requirements

3.1.1.1 User Management

SRS-USR-001: The system shall allow the user to become a member of WADAS registering his/her GA account. The following information is required to become a member of WADAS

- First Name
- Last Name
- Username (email from GA account)
- Password (this is an independent password not related to the GA account)
- Type of User by Default (IT, Designer, CEO, Salesman, Marketing staff)
- Phone Number
- Store historical data from profiles of GA account for generation of report in Suggestion Module (Yes/No).

SRS-USR-002: The system shall allow the user to deregister from WADAS, inactivating the account.

SRS-USR-003: The system shall allow the user to change the personal information (profile). That information includes:

- First Name
- Last Name
- Password
- Type of User by Default (IT, Designer, CEO, Salesman, Marketing staff)

⁴ <http://www.mangoldsengers.com>

- Phone Number
- Store historical data from profiles of GA account for generation of report in Suggestion Module (Yes/No).

SRS-USR-004: The system shall allow the administrator (User with role of Administrator) to change the status of other users to:

- Active
- Inactive
- Blocked

SRS-USR-005: The system shall automatically assign a new password and send it by email to a user that requests for the 'forgot your password?' action.

3.1.1.2 Security

SRS-SEC-001: The system shall request for login and password to protect from unauthorised use of the system.

SRS-SEC-002: The system shall allow the user to login (creating a valid session) and logout (ending session) from the system.

SRS-SEC-003: The system must request the user to enter the Google Account password in order to retrieve information from the GA account.

NOTE: WADAS does not store passwords from the GA account, for this reason every time that the users want to initiate a session for retrieving and displaying data (Dashboard and Suggestion Module), they need to authenticate log in to the GA account. The GA session ends after one hour of inactivity or when the user log out from WADAS.

SRS-SEC-004: The system shall allow the administrator (User with role of Administrator) to reset the password of other users as required.

3.1.1.3 Retrieving and Displaying Data (Dashboard – Suggestion Module)

This subsection is concerned with the ways in which to retrieve and present information to the user.

SRS-REP-001: After the user is logged in the system presents the list of profiles from the user's account. The user can then select the profile to work with in order to display the dashboard.

SRS-REP-002: The system shall allow the user to enter the following selection criteria to retrieve the dashboard.

- Type of Dashboard (IT-Design, Marketing, Management, Sales): According to the type of User assigned to the user's profile, the type of dashboard by default is presented. However, the user has the option to select other kind of dashboard to display.
- Start Date: Initial date for the period that is retrieving (e.g 01/01/2009).
- End Date: Ending date for the period that is retrieving (e.g 31/01/2009)

SRS-REP-003: After selecting the profile from the account, if the user account has enabled the use of the Suggestion Module, the system shall allow the user to see the results from the Suggestion Module for that specific GA profile.

SRS-REP-004: The system shall allow the user to print the dashboard.

SRS-REP-005: The system shall allow the user to send the dashboard results by email as an attachment.

3.1.2 Design Requirements

SRS-DES-001: The system shall allow the user to submit a form that permits later store and retrieve persistent data related to the user account in WADAS, which includes:

- First Name
- Last Name
- Email (Username)
- Password
- Phone Number
- Type of User by Default (e.g Marketing, IT, Designer, Salesman, CEO)
- Store historical data from profiles of GA account for generation of report in Suggestion Module (Yes/No).
- Role in WADAS (e.g Administrator, Normal User)

SRS-DES-002: The system shall store and retrieve persistent historical data related to the GA account for the users that allow WADAS to do it.

SRS-DES-003: The system must be designed to work in MAC OS X, Windows and/or Linux for the next browsers:

- Internet Explorer 7.0 +
- Firefox
- Chrome

SRS-DES-004: The system shall be developed in any language that can interact with the GA API and Google Visualization API.

SRS-DES-005: The system shall be developed in a Web environment, not desktop. Therefore, the user should be able to use the application through the Internet without downloading any particular software related to the application.

SRS-DES-006: Each dashboard must fit in a page, no bigger than a page size A4.

3.1.2 Non – Functional Requirements

SRS-NFN-001: The application shall respond to any retrieval and generate a report in less than 1.5 minutes; otherwise the user will be redirected to a time exceeded error page.

SRS-NFN-002: The application shall allow the user to maintain the session for the application opened without any time limit. However, the session with the GA account will be closed after 1 hour if the system does not detect any activity.

SRS-NFN-003: The system will provide a Help and Frequent Asked Questions (FAQ) section to assist with the WADAS functionality.

3.2 Requirements Analysis

After the analysis of the above requirements we identified two major functional areas and three actors.

3.2.1 Actors

3.2.1.1 User

Aliases	Member, Client, Customer
Description	Everyone that has an account with user id and password, but does not have administrator privileges.
Type	Human

3.2.1.2 Administrator

Aliases	Admin, Super user
Description	Super user that can delete, block and update data from other users. Also this actor is able to reset the password of users when it is required.
Type	Human

3.2.1.3 WADAS Database

Aliases	WADAS System
Description	An external actor that is responsible of processing data, generating and sending passwords when is required to a specific email account.
Type	System/Device Actor

3.2.2 Functional Areas

3.2.2.1 User Management

This package involves tasks related to the Membership Management. Therefore, it contains procedures such as the membership registration/deregistration, login and logout. The figure 3 illustrates the Use Case Diagram with the actions that can be requested by the different actors in this package. The Appendix A presents the use case definitions for the user management and data reporting services modules.

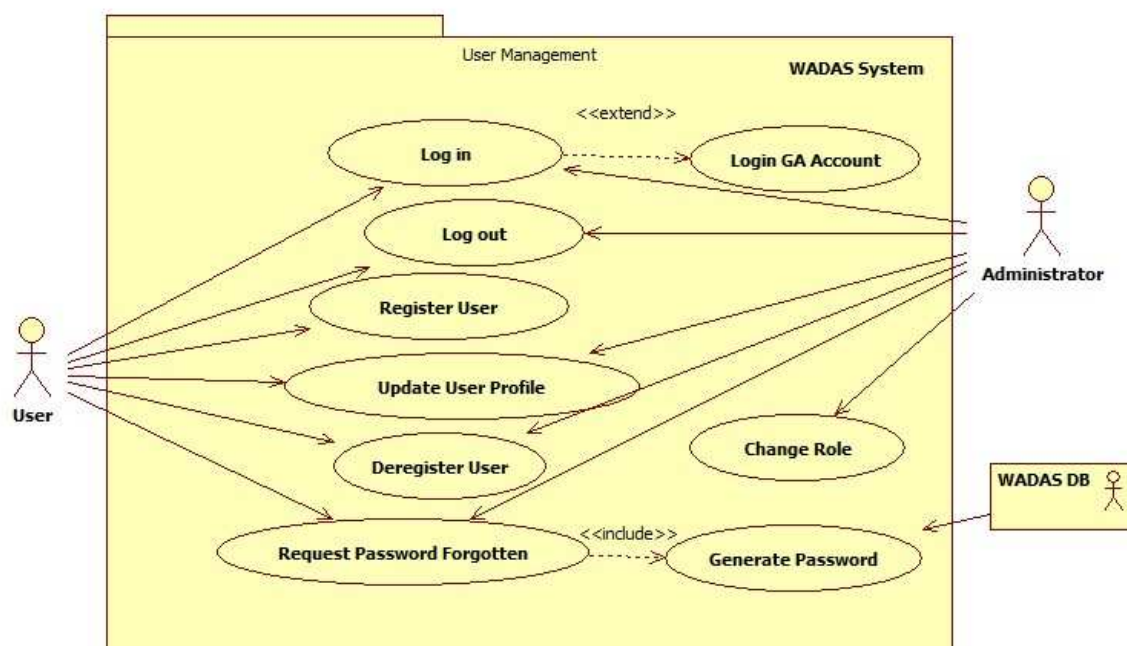


Figure 3 User Management – Use Case Diagram

3.2.2.2 Data Reporting Services

This module contains a range of services based on the information retrieved by GA and offered by WADAS to the users. These services include the dashboard presentation and the Suggestion Module report. The figure 4 displays the Use Case Diagram for the data reporting services.

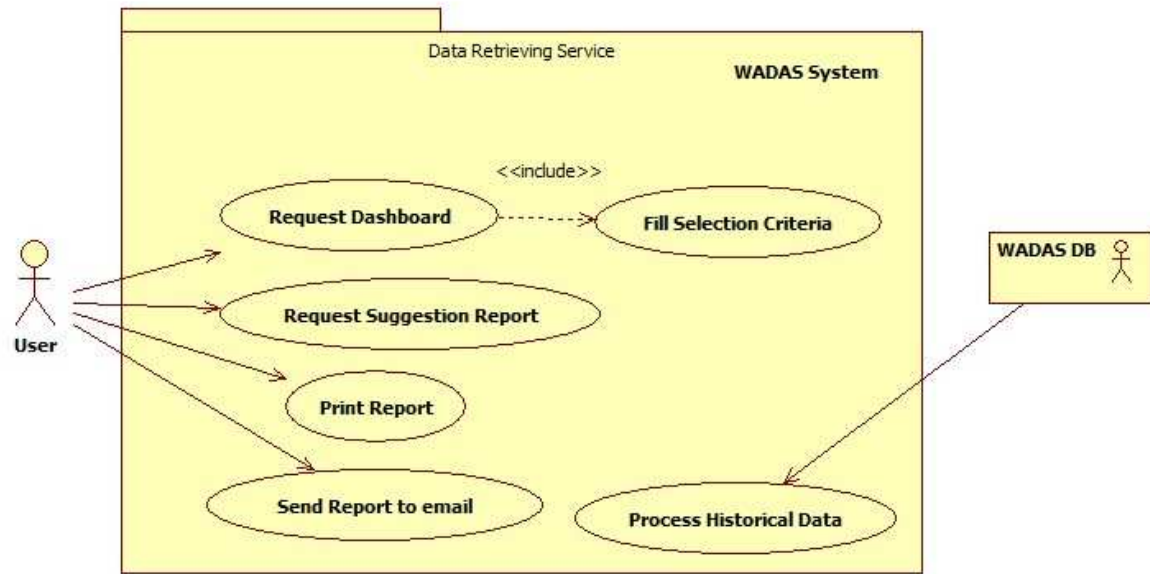


Figure 4 Data Reporting Services – Use Case Diagram

4 Design Phase

This chapter explains the design phase involved in the development of the application and the design of the two main components of WADAS. As part of the normal process to develop any application, this chapter describes the Design phase for the WADAS System following the UML (Unified Modelling Language) approach. However, according to the scope of the project, the main challenge established was the design of a successful Web Analytics Dashboard and the creation of the Suggestion Module. Therefore, in this chapter the design of these two components is explained.

4.1 WADAS System

The WADAS System design takes advantage of the model-view-control (MVC) object-oriented design pattern. Hence, the core processes that involve access to the Data Base and persistence of data are maintained in the Model packages. Such processes are called via the Control packages, which can create and modify the objects from the Model. On the other hand, the user only has interaction with the View packages, which consist of the User Interface components.

4.1.1 Model Component

The figure 5 shows the classes included in the model package. This classes are independent of the view, and because of WADAS is created under a web environment we make use of a passive model. It means that any operation executed in the model is notified to the view when the model requires the view component to be changed.

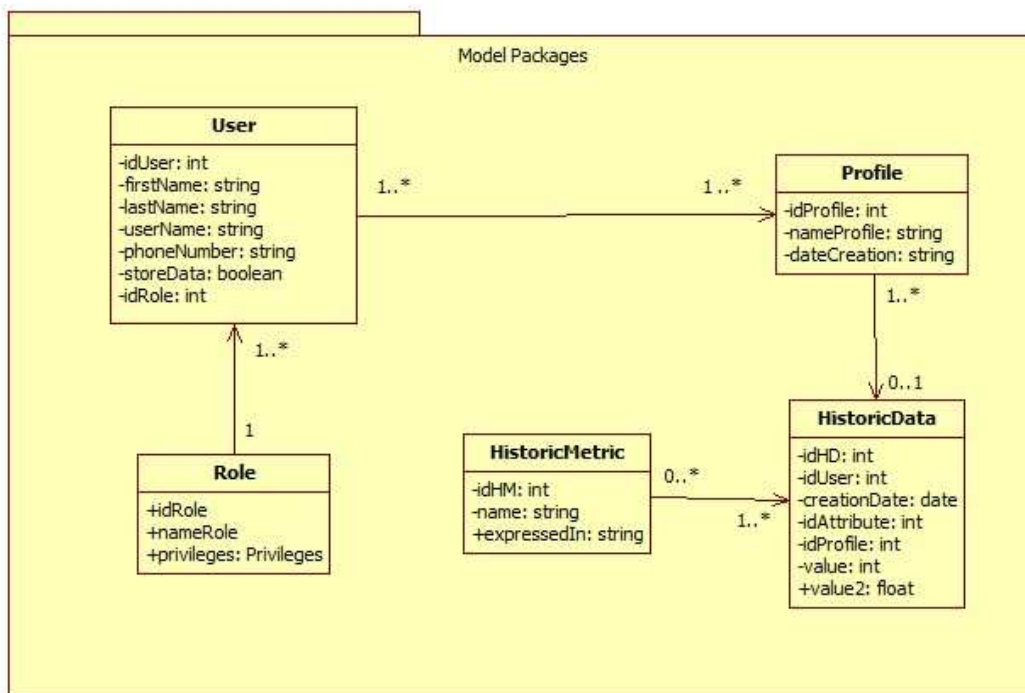


Figure 5 Model Packages

4.1.2 Process Control Component

The classes designed in this component allow the sending and reception of requests between the view and the model component. One of the most important controllers for our system is the DashboardController which executes the methods required to retrieve information from the GA API and format it to be displayed in the view component. The figure 6 illustrates the five controllers created to lead the communication between model and view packages.

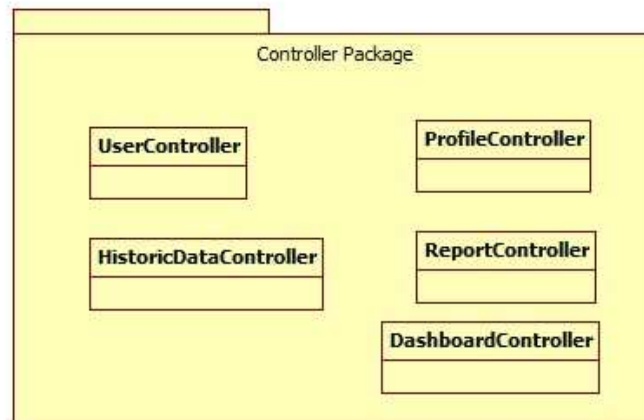


Figure 6 Controller Packages

4.1.3 System Interface Component

This component contains the scripts that output the application to the user. Each packet in the interface component contains several scripts that build the interface to the user. The figure 7 shows the 5 packages created to create the interface with the user.



Figure 7 View Packages - Interface

WADAS is a scalable web application where the user interacts with the tool as with any other website that requires an initial authentication. To make easier the understanding of a normal process followed by a user and the options available at the moment on the WADAS site, the figure 8 illustrates the user flow.

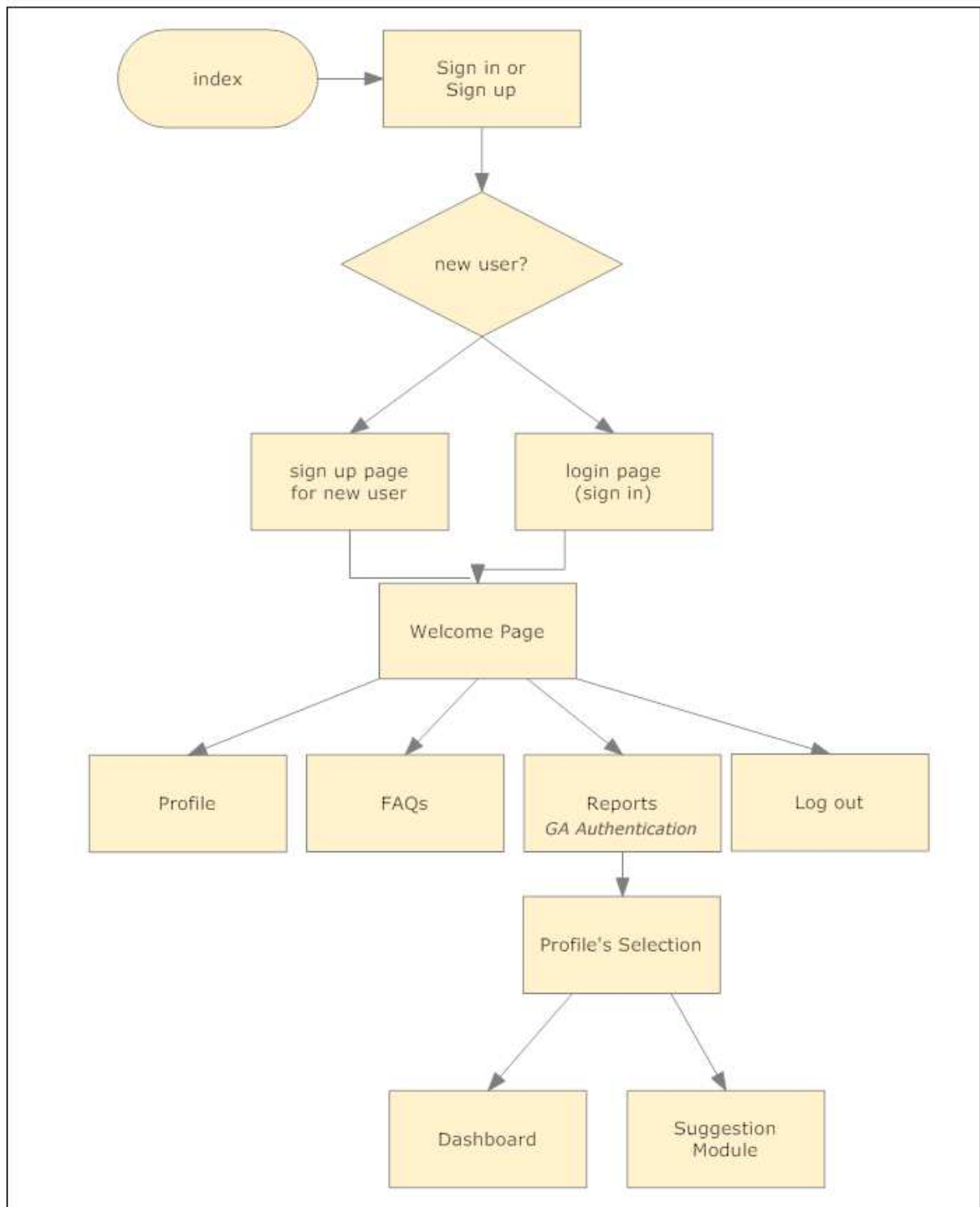


Figure 8 User Flow Diagram

When the users access the application, they have two options, log in using a username and password or submitting the registration form to become a member of WADAS and then log in. At the moment that users are authorized and authenticated in WADAS, the welcome page is displayed and from this point they have four options:

- View and Update Profile
- Visit the Frequent Asked Questions
- Log in to the GA account to generate the dashboard or the suggestion report.
- Log out

Other pages and functionalities can be added to WADAS in a future, allowing users to perform other tasks.

4.2 Web Analytics Dashboard

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 mission and vision of a business may be aligned, but the ways each department in the organization works to achieve the objectives set may be different. Bearing this in mind 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 departments. Some information is exclusive for a type of dashboard, but the majority of the data is shared among the four dashboards. The main difference is the mode in which 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 is probably relevant for the IT – Design department, but not for the Sales department. However, the four dashboards share a similar structure composed by

- **Header**

It is true that displaying the name of the Dashboard and the Logo of the product or business is important for the organization because it provides identification. However, we decided to allocate an 8% of the screen to display the following data:

- WADAS Logo
- Name of the Dashboard (e.g Sales Dashboard, Marketing Dashboard)
- Image menu with Hyperlinks to Print Dashboard and Send Dashboard to email
- Today's Date (System date including hour, min and seconds)
- Current Period Selected (e.g 01/01/2009 to 31/01/2009)
- Currency (e.g AU Dollar, US Dollar)
- Username
- Menu with Hyperlinks to Homepage, User Profile, Help Section, Log out

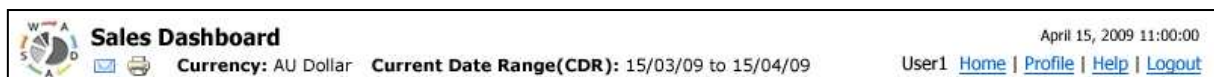


Figure 9 Header Section

- **Content**

The content area fits into 87% of the total space reserved on the screen. We divided the content area in six sections for each dashboard, and maintain patterns in them allowing the user to find the most important and less detailed metrics in the top – left of the screen. That section is called overview because it does not display detailed data but allows the behaviour of important metrics over the last 12 months to be viewed. In the same way, the section two and three show overall information, while the right side of the screen is used to display data in at a more detailed level.

Section 1 Overview	Section 4 Motion Chart
Section 2 Comparison Metrics	Section 5 Detailed or Geographic Information
Section 3 Performance & Comparison with last period	Section 6 Detailed Information

Figure 10 Content Section


- **Footer**

This part of the dashboard is not shown in the next designs because it was considered as optional and it is only included in the version for the presentation of the project in such way that the public can see this is part of a university’s project. It takes 5% of the space allocated on the screen.



Figure 11 Footer Section

Additionally, there are other aspects of the visual design that are shared in each dashboard.

-  This symbol appears as an alert for the user to highlight results that require special attention.
- Because of the lack of a grid to divide each section, we decided to highlight the titles (bold letters) and leave a small white space between the groups of data.
- CDR is an abbreviation that represents Current Date Range.
- All the sections except the first one have interactive functions allowing the user to:
 - Order the results for each column when the data is displayed in tables.
 - Check the actual value that produced the percentage shown in the pie charts. For instance, the percentage of revenue per direct traffic source is 25.64%. If the user wants to check the actual revenue generated from that traffic source only has to click the portion corresponding in the pie chart and the value (\$293,078.19) will be displayed.
 - Data in bar charts corresponding to the last period is shown in gray colour.
 - The motion charts allow the user to analyse data with multiple dimensions, finding opportunities and problems faster and checking the behaviour of the elements through the time.
- Unfortunately, some information cannot be displayed in the dashboards because the profiles in GA accounts are not configured properly or because the type of Website does not fit the goals of a specific dashboard. Therefore, in this kind of situation, the dashboard displays a red title informing to the user that such information was not found.

A clear example of this is a website that does not have activated the ecommerce option in the profile; therefore there is not any data about products or transactions.

NOTE: The designs displayed below are the ones considered before the tests. Therefore, some of the features shown in the next designs are not included in the actual implementation of WADAS because after the feedback received by the evaluators about the design some changes were required. Such changes will be explained in the Chapter 6.

4.2.1 Sales Dashboard

The main objective of this dashboard is that the sales user can identify the significant data expressed in the easiest way to improve the revenue. As expressed by Few S. [2006, p176], the sales dashboard is the second type of dashboard more commonly implemented after the management or executive one, because the core of most businesses is sales.

The information displayed in this dashboard was selected bearing in mind the data available in the GA API and studying the type of information that a sales manager looks for in a report to take strategic decisions. The next is a list of the data shown in the Sales Dashboard, which is displayed in the figure 12.

4.2.1.1 Section 1 - Overview

The Section 1 of the Sales Dashboard displays the next Metrics for the past twelve months and the current period:

- Revenue
- Transactions
- Conversion Rate
- Purchase by New Visits versus Returning Visits
- Revenue per Traffic Sources
 - Search Engines
 - Direct Traffic
 - Referring
 - Other

With this data Web Analytics users can review the behaviour of the website's clients over the last twelve months. They can also conclude if there has been an increase in the number of transactions and revenue, or not. Users are enabled to identify the traffic sources that are generating more revenue, whether or not it is a constant over the time or it is an unusual result.

With the spark lines that compare purchase by new and returning visits, users can identify the visitor's trend and measure loyalty.

4.2.1.2 Section 2 - Comparison Metrics

This Section shows two pie charts allowing the user to compare different metrics for the same dimension. The first pie chart displays the percentage of visits per traffic source in the current period. The second one displays the percentage of revenue obtained per traffic source.

Sometimes, users attribute the success of their sites only to the number of visits. Nevertheless, not always the traffic source with major number of visits is the one that generates best return of investment. The graphics displayed in this section help the user to identify the behaviour of the visits coming from different traffic sources and possible failures in the media utilized to promote the site.

4.2.1.3 Section 3 – Performance and Comparison with Last Period

The Section 3 displays information for the current period and last period of the next metrics:

- Total Revenue
- AdWords Revenue
- Transactions
- Conversion Goals

With the overview section we are able to review trends over the last twelve months. With this section we can observe in more detail the changes from the last period compared to the present. Also if we have configured an AdWords account, we will see how much revenue we are receiving

from AdWords and from the entire site. We will be able to determine how AdWords is impacting our return of investment.

Finally, for those who have followed the good practices of GA and have goals configured, this dashboard will report them the percentage of objectives reached.

4.2.1.4 Section 4 – Motion Chart of Products

This motion chart allows the user to monitor the behaviour of the products during the month. The advantage of the motion charts is the use of different dimensions (up to 5) in the same graphic. For the products motion chart the metrics available are:

- Product Revenue
- Quantity
- Unique Purchases
- Visits

Because this is an interactive feature, the user can select all products or some of them and visualize how their performance was during the current period (day by day).

4.2.1.5 Section 5 – Detailed Information of AdWords

Because of for ROI calculations it is important to check the revenue generated from paid (AdWords and other ads) and no-paid keywords (organic search engine). This section not only displays the performance for no-paid and paid keywords, but also for those who had configured an AdWords account it displays the top 5 campaigns based on the revenue obtained.

4.2.1.6 Section 6 – Detailed Information of Products

This Section displays the Top 5 products based on the revenue generated for the current period. With the information for these products the user can find the quantity of items sold for that product and the percentage of revenue compared to the entire site for each product. This information is important because the user can determine if the revenue obtained is really representative for the business and if the number of items sold for the product compared to others is on average the same.

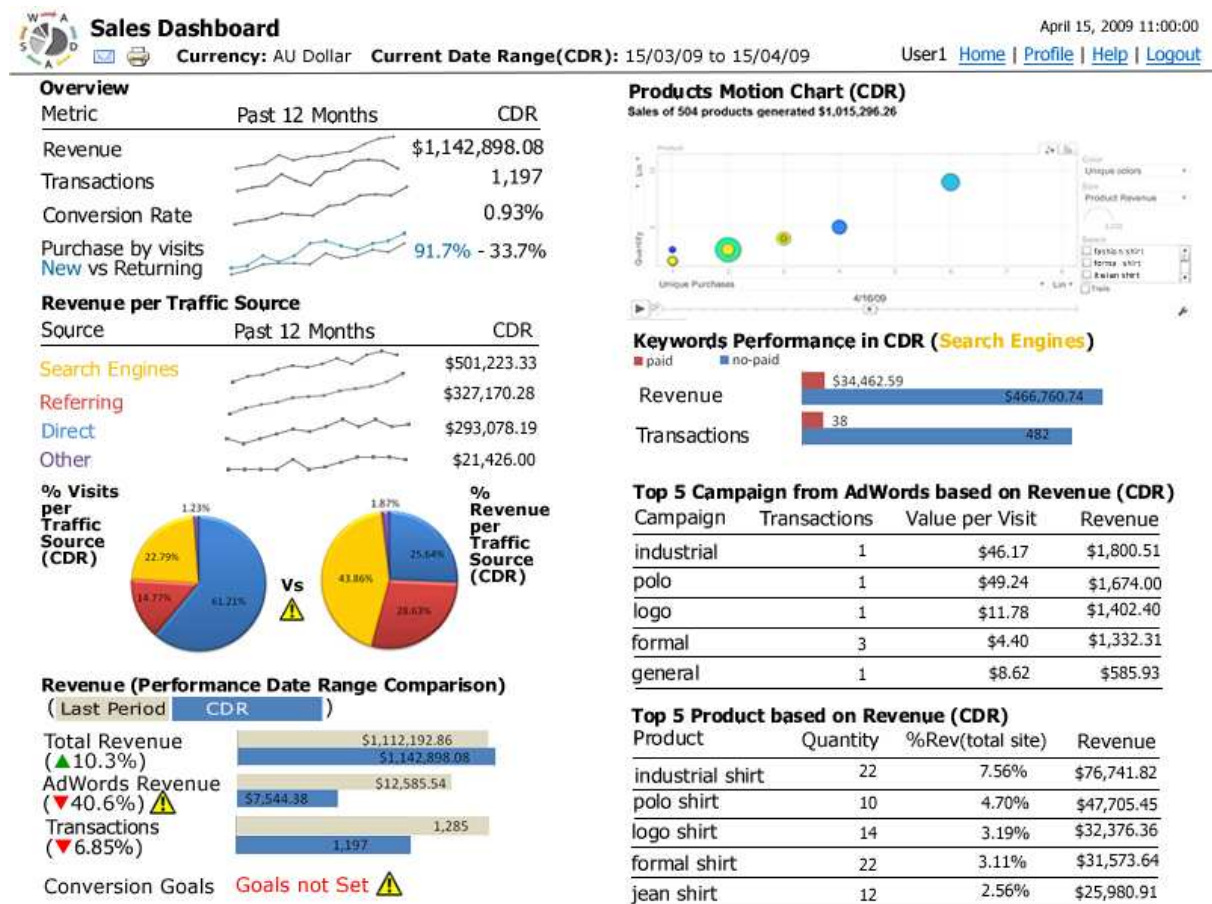


Figure 12 Sales Dashboard

4.2.2 IT - Design Dashboard

This dashboard is designed to assess and monitor technical aspects, and the mock up is displayed in the figure 13. Generally users interested in this dashboard are Designers and Chief Information Officers. The failure of many websites is caused by the appearance of the site in different browsers.

4.2.2.1 Section 1 - Overview

Similar to the Sales Dashboard this section displays an overview of the next metrics:

- Revenue
- Conversion Rate
- Visits
- Pages per Visits
- Purchase by New Visits versus Returning Visits
- Average Time on Site (ATOS)
- Bounce Rate

4.2.2.2 Section 2 - Comparison Metrics for Browsers and Operating System

Here, the user is able to identify the percentage of visits that use a specific browser and operating system. In the example, contrary to what many IT personnel thinks, Internet Explorer continue being

the browser most used by the public; therefore, the look and feel and functionality of the website should be correct in this browser.

4.2.2.3 Section 3 - Performance and Comparison with Last Period

The section 3 presents the same information than the Sales Dashboard in section 2 but using a bar chart instead of pie charts. The reason to design these dashboards with a different presentation was to test which one could be read quickly.

4.2.2.4 Section 4 – Motion Chart of Pages

The motion Chart for the IT-Design Dashboard displays detailed information about the page viewed. Thus, the user can analyse the behaviour of the pages contained in the site, combined with the next metrics:

- Number of Page views
- Unique Page views
- Bounce Rate
- Revenue

4.2.2.5 Section 5 – Detailed Information of Pages

This section lists the top 5 pages based on the number of times that a page has been loaded. For the IT-Design department is critical to know which are the pages more visited, the percentage of exits from those pages and the bounce rate, in order to improve them. However, it is also useful to know the average value per each one of those top 5 pages.

Note that in the table we display the bounce rate per page. We explained in Chapter 2 - Section 2.3.1 the concept of bounce rate. According to that, a high bounce rate is not a positive metric but it needs to be analysed with the average value per page in order to conclude if the page should be redesigned to another ad or link that is attracting users to it wrongly. Also, being able to see information about the actual name of the page, bounce rate, average of value and percentage of exits the user can determine if a big number of exits is a notion of failure or success. For example, in the figure 13, we can observe that for the 'thankyou.html' page the percentage of exits is 5.57%, for this particular case we would prefer a bigger percentage because it would mean that people is completing the purchase.

4.2.2.6 Section 6 – Detailed Information of Network Vendors

The vendor used to connect to the Internet, is a particular dimension that can be interesting to the organization to increase the performance of the site. This Section provides information about the top 5 Network Vendor based on the number of visits. Trends of user behaviour could be identified with this report.

In addition this section shows a bar chart comparing the number of visits and revenue generated for people that use the internal site search and those who not. Although this chart is not very detailed because the user cannot identify which words were used in the internal site, the user could determine how important the internal search is for the performance of the site. In the example for the figure 13, the number of visits that used the internal search is bigger than those that not;

however, the revenue is less. It can be caused because visitors did not find the correct information with the internal search tool.

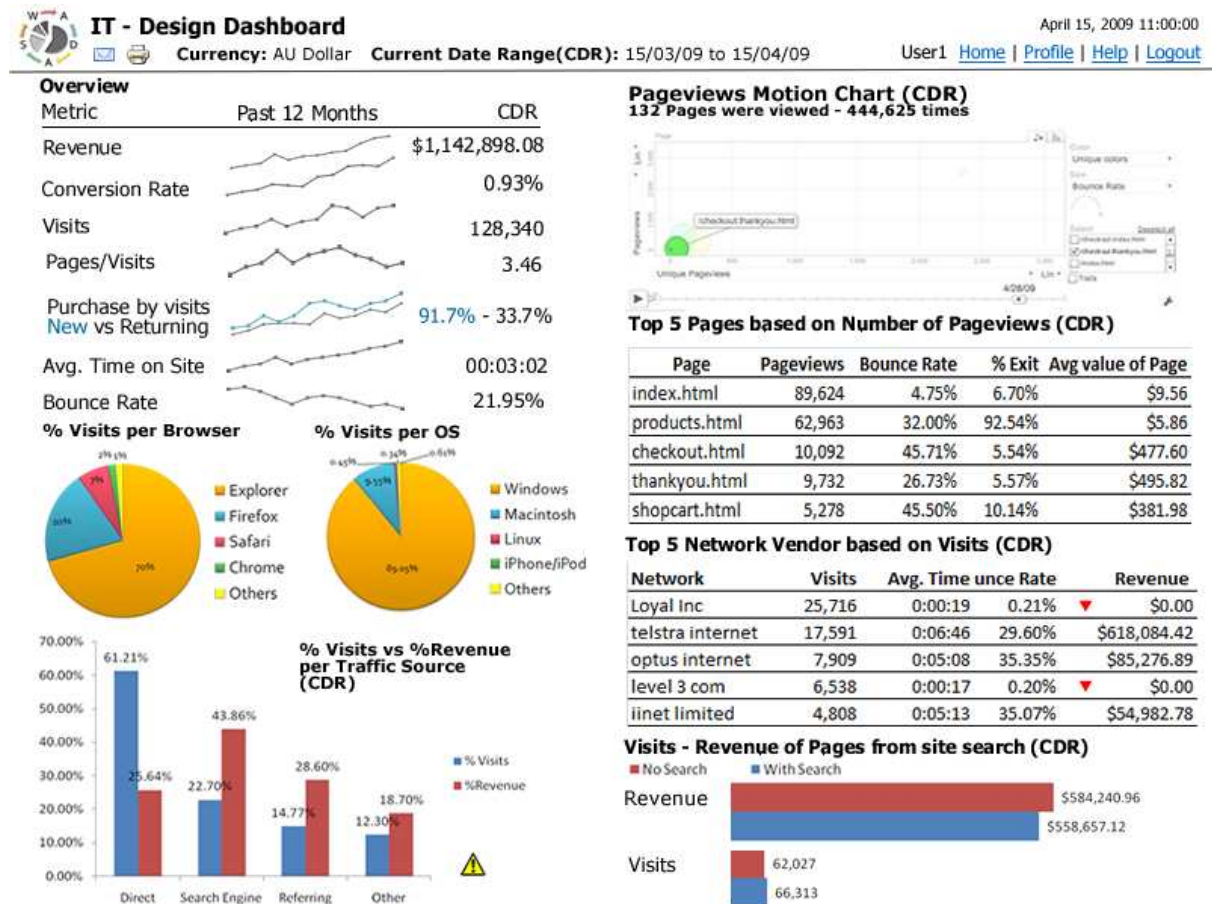


Figure 13 IT – Design Dashboard

4.2.3 Marketing Dashboard

The Marketing Dashboard is commonly used to analyse the clients' behaviour when they visit the site. Thus, the Marketing team is able to determine what factors are preventing customers to purchase, navigate and/or interact with certain elements the organization desires. The figure 14 displays the mock up of this dashboard, which is the one that we have implemented at the moment in WADAS.

We decided to implement this dashboard rather than the others for two main reasons. First of all, for a huge number of companies the reason to have a presence on the Internet is marketing and popularity rather than online sales for example. Additionally, this dashboard is composed by a set of metrics that can be calculated without requiring special configuration in the GA profile. Implementing the Sales dashboard for example, would not give us data in most of the tests performed because from our experience with clients that use GA, few of them have configured the profile and the Google tracking code to store ecommerce data.

4.2.3.1 Section 1 - Overview

The Section 1 of the Marketing Dashboard displays similar metrics than the other dashboards for the past twelve months and the current period:

- Revenue
- Conversion Rate
- Visits
- Pages per Visits
- Purchase by New Visits versus Returning Visits
- Average Time on Site (ATOS)
- Bounce Rate

4.2.3.2 Section 2 – Comparison Metrics

Once more, we found in this section a chart to compare the percentage of visits with the percentage of revenue obtained per each traffic source.

4.2.3.3 Section 3 – Performance of Internal Search

The number of people that make use of the internal search is also important for the Marketing department. If the number of visitors using the internal search was significant, the organization could try including marketing strategies in the results of these searches to address the user to the converting pages.

4.2.3.4 Section 4 – Motion Chart of Pages

The motion chart for this dashboard is the same that the one used in the IT-Design Dashboard. For the marketing department is important to identify the pages more valuable for the organization in order to create marketing strategies to improve the return of investment and popularity.

4.2.3.5 Section 5 – Detailed Information of Referring Sources

This Section displays the top 5 referring sources based on the number of visits. However, it also shows the revenue, number of transactions and percentage of new visits coming from that referring site. Sometimes, organizations spend a lot of money advertising their products in websites that are not generating any revenue. Hence, there are big surprises when they find out that free social media sites such as twitter, facebook and flickr generates more visits and revenue than others.

4.2.3.6 Section 6 – Detailed Information of Keywords and Products Performance

The Section 6 in the Marketing Dashboard presents two elements. The top 5 keywords based on number of visits is presented as a table, where the user can also review the revenue , number of transactions and bounce rate caused by each keyword utilized.

Subsequently, the user is able to compare between two periods the performance of the top 5 products sold based on the quantity of items purchased.

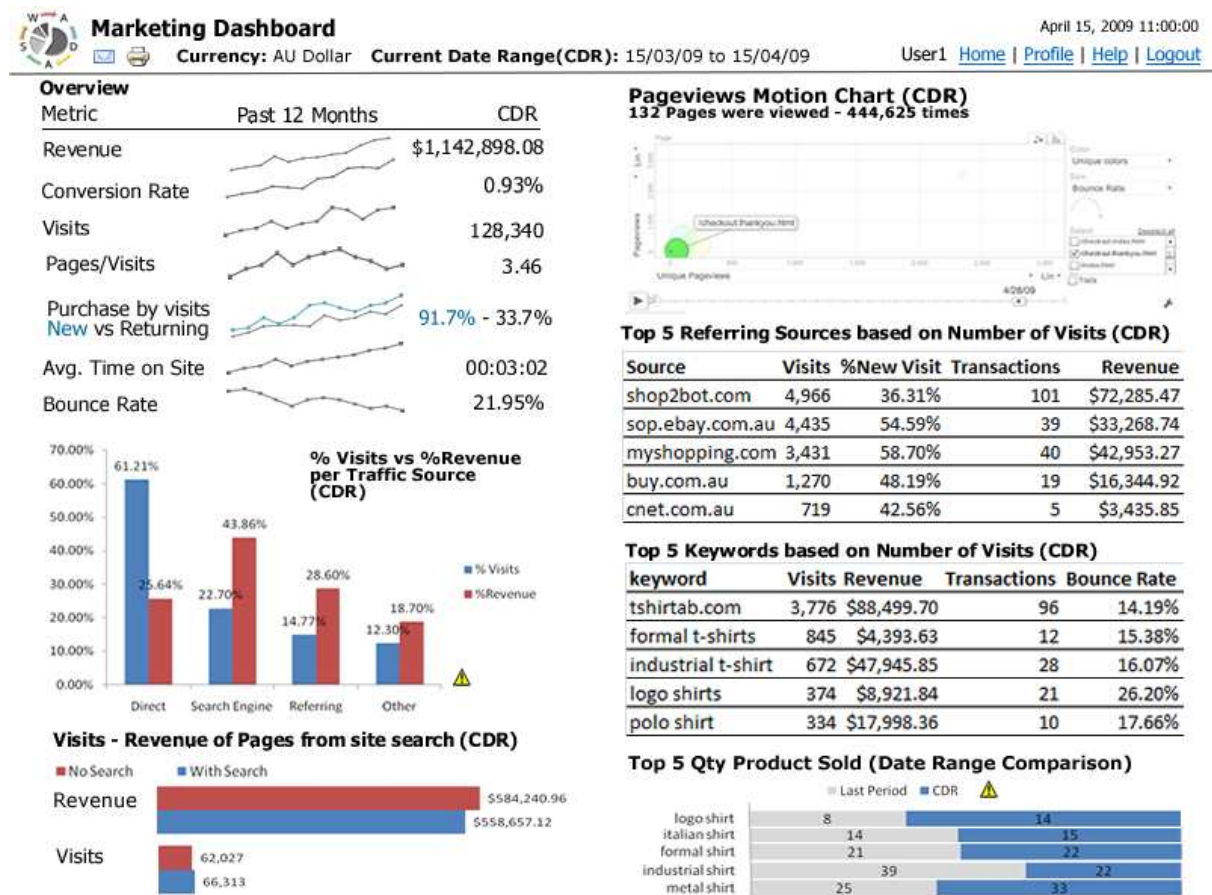


Figure 14 Marketing Dashboard

4.2.4 Management Dashboard

Probably this is one of the dashboard more used by any business because it has as objective allow the management executives to know at a glance the performance of the site, visitor trends, problems, opportunities and areas that might need attention.

4.2.4.1 Section 1 - Overview

The Section 1 of the Management Dashboard displays the next metrics for the past twelve months and the current period:

- Revenue
- Conversion Rate
- Transactions
- Purchased Products
- Visits
- Pages per Visits
- Purchase by New Visits versus Returning Visits
- Average Time on Site (ATOS)
- Bounce Rate

4.2.4.2 Section 2 – Comparison Metrics

The executive can determine in this section which traffic source is generating more revenue and which is the source most used to visit the site. In the example, the executive is able to conclude that although the percentage of direct visits is bigger than the ones that found the site using search engines, the revenue obtained from the search engine traffic source is the double than the direct traffic. It can be caused by many factors; one of them is if the profile in GA has not set the filter to exclude internal traffic. Frequently, people from the organization access the site to make tests and include new content writing the URL in the browser, or they have that bookmarked. If the profile does not have set the exclude internal traffic filter all these visits are counted as direct visits that do not generate any revenue.

4.2.4.3 Section 3 – AdWords Performance

Currently, people using AdWords to advertise their site can link the AdWords account with the GA account. Regular changes in the campaigns, keywords and ads are made month by month. Hence, for the organization is significant to compare the AdWords performance between the last and current periods.

In the example, although the number of clicks is similar between the current and last period, there is a huge difference in the return of investment obtained. This is a point that might require attention because probably any changes in the site or ads are causing that people click the ads but when they are in the site do not complete the cycle desired by the organization (e.g do not buy any product, do not make a donation, do not fill a registration form).

4.2.4.4 Section 4 – Motion Chart of Traffic Source

The traffic source motion chart allows the executive to analyse the behaviour of the different sources (source / type of source) using the next metrics:

- Number of Visits
- Number of pages per visits
- Revenue
- Bounces

4.2.4.5 Section 5 – Geographical Information

The Section 5 of the Management Dashboard is quite different to the other dashboards. In this opportunity, as you can see in the figure 15, a world's map is displayed. Hence, the executive can find the amount of visits to the site around the world. It can provide relevant information about problems and opportunities. For instance, a cake shop that only sales products to a local region and is using AdWords to advertise the site, securely is not interesting in receiving visits from other parts of the world because when those visitors arrive to the site find out that the cakes are only distributed locally, and then leave the site. On the other hand, it can be an opportunity for other kind of businesses.

Depending of the situation, the organization can take as action the restriction of the advertisements and impressions in organic searches to the local country and/or city.

4.2.4.6 Section 6 – Goals Performance and Detailed Information of Products

This Section, presents the performance between the last and current period for the Goals set in the profile and the top 5 quantity of products sold. Not all the sites implement an ecommerce options; which means that information about products is not displayed. In the same way, new GA users are not familiarized with the use of Goals; therefore, if they are not set in the profile, the dashboard shows a message letting the user to know that the Goal Performance chart cannot be displayed because the goals are not set.

Setting Goals are very important in Web Analytics, because they help the user to see the result of the site for specific metrics in relation to the target.

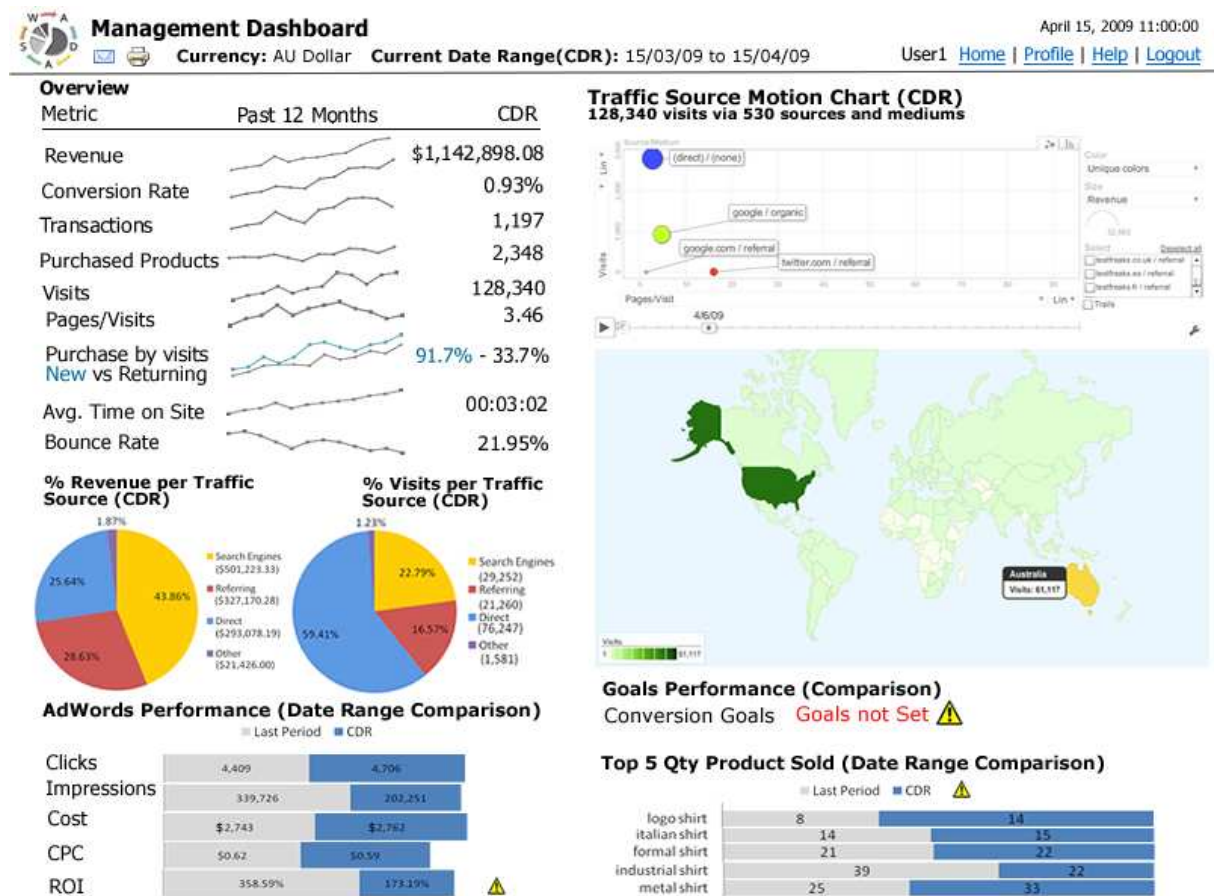


Figure 15 Management Dashboard

4.3 Suggestion Module

The second component of the design stage was the Suggestion Module, which is only one for all the dashboards.

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.

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.

With the dashboards users can visualize historical data, trends, patterns and areas that require special attention. In contrast, with the Suggestion Module, users will see a list of suggestions to improve the accuracy of the data, to understand advanced options enabled in GA and to take action about elements that do not seem to be obvious, but that probably are affecting the website's performance. An example of the results shown in the Suggestion Module is presented in the figure 16.

1. *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).*
2. *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).*
3. *The campaign 'Camp 1' is not generating any revenue since it was created, you should consider making changes for that.*

Figure 16 Suggestion Module example

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

5 Implementation Phase

The implementation for the tool involves different elements from the time that the user is registered to the application. As WADAS was designed using an object oriented methodology, we followed in the implementation this paradigm for the major operations.

The system makes use of the next programming languages:

4. PHP (Zend Framework)
5. JavaScript
6. HTML

Zend⁵ is a framework that implements the MVC pattern; therefore the packages and classes illustrated in the section 4 were implemented using this framework.

5.1 General Implementation of WADAS System

The classes shown in the subsection 4.1.1 (figure 5) were implemented using the Zend Framework. Although the methods to manage the data provided by the user were controlled by the framework, the actual storage of persistent data was achieved using a relational database (MySQL).

According to the user flow shown in the subsection 4.1.3 (figure 8) when a user visits the landing page of the WADAS application, the system asks the user to login in. When a guest user becomes member of WADAS the information is stored in a MySQL database. Part of such information is used to validate the login and password when a user attempts to login.

In order to access the dashboards and Suggestion Module reports, the user requires two levels of authentication. The first authentication allows the user to access the WADAS system. At this level the users can see and update their profile, access the FAQ – help page and log out from the application.

This first authentication was implemented with the Zend Framework (PHP) and MySQL, where the username and password entered by the user are verified with the username and password encrypted in the MySQL database.

The second authentication, which is against the Google account, was implemented using a special library (GDATA) included in the Zend Framework. Google provides four options to make use of the APIs authenticating an account:

7. Client Authentication
8. Authsub Authentication
9. OAuth Authentication

⁵ <http://framework.zend.com>

10. OpenID Authentication

We decided to implement the GA authentication function with the Authsub protocol, which allows our application to access the GA API without requiring that we keep password information of our Users related to their GA accounts. The Appendix B shows a diagram provided by Google that illustrates this process in more detail. For more information about this process check the Google AuthSub authentication page⁶.

After the user is authenticated with the GA account and in WADAS, the system enables the options to display the dashboards and the Suggestion Module.

5.1.1 Security Issues

5.1.1.1 Creation and Deletion of sessions for WADAS

A session in WADAS is created when the user log in using a username and password, and it is managed on the server side. In order to create a session we use abstract containers called namespaces through an element defined by Zend (Zend_Session_Namespace), which permits to store information associated to the user that has been authenticated in the application. To avoid problems of starting a session after the headers of the pages displayed to the user, we start the sessions in the bootstrap file.

At the moment, WADAS does not offer other functionality that involve critical confidential information rather than the dashboard and suggestion report, which require the previous authentication in the GA account.

5.1.1.2 Login and Logout process to WADAS

When the users visit the home page, they must enter a username and password to log in the application. That information is transferred from the interface to the controller that makes use of the user model to validate the credentials. Because of the MVC model, the source code that manage such validations is not placed inside the interface scripts; therefore, clicking view source code or saving the page in a local computer will not show the values entered by the user.

5.1.1.3 Requesting of the single-use token to Google

Our main objective for using Authsub as an authentication and authorization protocol to access the API information was to avoid handling GA user's account credentials. Instead of saving any information from the user, we request a single – use authentication through the function AuthSubRequest, which is called as an URL with four parameters explained below.

a. next -> URL of our application, which must be registered previously in Google as an authorized application.

b. scope -> URL of the GA API to access the feeds.

⁶ See <http://code.google.com/apis/accounts/docs/AuthSub.html>

c. secure -> For purposes of this project we are not using certificates. If WADAS is offered to the public it should use certificates (SSL) to provide a secure transactions. Therefore, at the moment this parameter is 0, which means that we are requesting a non-secure token.

d. session -> The value is 1 to indicate that the token may be exchanged for a session token.

<https://www.google.com/accounts/AuthSubRequest?next=<url of our application>&scope=https%3A%2F%2Fwww.google.com%2Fanalytics%2Ffeeds&session=1&secure=0>

5.1.1.4 Upgrading of the token for long-lived session using the AuthSubSessionToken method.

When the authentication token is received we exchange the single – use token for a session – token, because we require using the authentication more than once. However, to manage security issues we revoke the session – token calling the method AuthSubRevokeToken after one hour of creating the GA session.

5.1.1.5 Attempting to use different GA accounts in WADAS that do not correspond to the user registered in our system.

We decided to validate the GA account against the username created in WADAS. It means that when a client is logged with the username user@example.com, this user should sign in to this GA account in order to use the Dashboard or Suggestion Module. If the user attempts to use the dashboard using a different GA username to the one logged the interface will show an error page.

5.1.1.6 Verification of denial of access to other profiles when the user does not have a role of Administrator.

Each user has a role in WADAS. Normal users have a role with restricted privileges; therefore, some features associated with the administration of other WADAS accounts are not enabled for these users. Only users with role ‘Admin’ can access other accounts and modify their information.

Sometimes, attackers attempt to access special pages that privileged users have visited writing the URL in the browser. However, using the MVC design model, the name of the page displayed on the browser is not the actual interface script. It is the name of one of the methods from the controller class. We check the user’s role in this method and depending of the result we show the respective interface script. The following is an example to make clear this statement.

Scenario: A super user accessed WADAS and decided to update the profile of another user. When the administrator selects the user to update and presses clicks in the option ‘Update’.

*The hyperlink to make this action looks like this: <http://www.example.com/users/update>. Internally, in our UserController class, we have a method called update**Action**. Zend identifies the methods that are called from the interface as actions, for this reason we must add the word ‘Action’ after the name that we show in the link.*

Inside this method we validate if the user has an administrator role or not, and this validation is possible because we can access any information in the database from that user based on the session handled. When the user is an administrator, we call the page 'updateAdmin.phtml', which is the interface script that presents special attributes enabled for the administrator. However, the user will see in the browser the URL that was clicked (<http://www.example.com/users/update>). The same happen if a normal user attempts to update the own profile. The method updateAction verifies that the role is not from one administrator, and then call the page 'update.phtml', which does not display options such as 'change status' and 'change role'.

5.1.1.7 Use of cryptographic hash function to store user password in the WADAS database

The passwords are stored in the WADAS database using the hash function SHA1. Therefore, someone that has access to the physical database probably would be able to change the data but not to see the plain passwords from the users.

5.2 Dashboard and Suggestion Module Implementation

In this section is explained the implementation related to the data collection to generate the Dashboard and Suggestion Module. The figure 17 shows the main procedures 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.

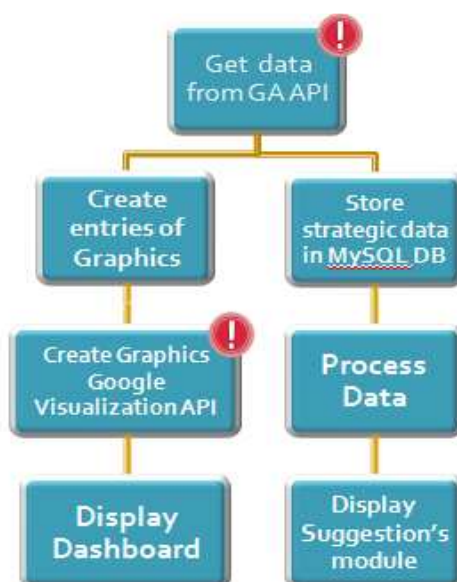


Figure 17 Implementation Process Scheme

The following is an explanation in more detail about the process to generate the Dashboard and Suggestions Module.

5.2.1 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 into the feeds.

The account feed is retrieved using the next structure of URL:

```
https://www.google.com/analytics/feeds/accounts/<username@gmail.com>
```

Where <username@gmail.com> is the email address or username of the client logged in WADAS, and that also corresponds to the account in GA.

The data feed has more parameters to retrieve the information correctly. The next is the structure of URL:

```
https://www.google.com/analytics/feeds/data?ids=<ga:profileid>&dimensions=<ga:namedimension>&metrics=<ga:namemetric>&start-date=<ga:startdate>&end-date=<ga:enddate>
```

As a normal URL address with parameters, the first parameter is after the question mark (?), from there, each parameter is separated by the symbol &.

ids=<ga:profileid> - We require to define the profile id from the one we want to get data. Each profile in the GA account has an id, those ids were retrieved previously when the account feed was requested.

dimensions=<ga:namedimension> - We can specify one or more dimensions to group the metrics. The limit of dimensions in a single query is seven. For example to get data divided by Operating System and Browser the dimensions parameter would look like this *dimensions=ga:browser,ga:operatingSystem*.

metrics=<ga:namemetric> - We can specify one or more metrics, up to 10 in a single query. For example, to get the number of visits, entrances and number of exits from the site the metrics parameter would look like this *metrics=ga:visits,ga:entrances,ga:exits*.

start-date=<ga:startdate> - We need to specify the range of date. The format is YYYY-MM-DD. For example, to set the initial date in the first day of the year 2009 it would look like this *start-date=2009-01-01*.

end-date=<ga:enddate> - Similar to the *start-date*, we require to specify the end of the period that wants to see.

There are other parameters (e.g order, filters, and maximum number of results) that can be added to the URL for retrieve the GA data feed. However, the ones described above are the most important.

At this point, WADAS only has made use of the account feed to display the available profiles that the user has assigned in GA. Later, the user must select one of them to retrieve data for the Dashboard or for the Suggestion Module. Consequently, here the process is divided to create the Dashboard and the Suggestion Module.

5.2.2 Dashboard

5.2.2.1 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 Spreadsheets and the other with JavaScript code [Google Code, 2009]. In our case, the second option is used.

The scope of the project specifies the implementation of one dashboard, the current dashboard implemented in WADAS is the one designed for the Marketing department. Here we explain the feeds used to retrieve the information required for the Marketing Dashboard.

In the Chapter 4, specifically the point 4.2.3, we showed and explained the Marketing Dashboard design and each one of its sections. To make the reader easier to understand the data feed requested to the GA API to collect the data related to the Marketing Dashboard we classified the feeds per section:

Section 1. Because of in this section we need to show data from the past twelve months and the current period, the dimensions are the Year and Month.

```
https://www.google.com/analytics/feeds/data?ids=<profileid>&dimensions=ga:year,ga:month&metrics=ga:visits,ga:pageviews,ga:timeOnSite,ga:bounces,ga:entrances,ga:transactionRevenue,ga:transactions,ga:newVisits&start-date=<startdate>&end-date=<enddate>
```

Section 2. In this case, we want to classify the number of visits and revenue for each traffic source. Because of the lack of a dimension that classifies the traffic source we require to get the data grouped by medium and then according to the type of referral to the website make the classification in the four different types of traffic source (Search Engines, Direct, Referring, Others).

```
https://www.google.com/analytics/feeds/data?ids=<profileid>&dimensions=ga:medium&metrics=ga:visits,ga:transactionRevenue&start-date=<startdate>&end-date=<enddate>
```

Section 3. Section three displays information grouped by the visits where the internal search activity occurred or not.

```
https://www.google.com/analytics/feeds/data?ids=<profileid>&dimensions=ga:searchUsed  
&metrics=ga:visits,ga:transactionRevenue&start-date=<startdate>&end-date=<enddate>
```

Section 4. This is one of the most complex data feed requested because return a big amount of information. This is due to the information in the motion chart displayed day by day for the period entered.

```
https://www.google.com/analytics/feeds/data?ids=<profileid>&dimensions=ga:pagePath,g  
a:year,ga:month,ga:day&metrics=ga:uniquePageviews,ga:bounces,ga:transactions,ga:trans  
actionRevenue&start-date=<startdate>&end-date=<enddate>
```

Section 5. In this section is displayed the top 5 referring sources based on number of visits. There is a limit of results to display, a condition that restrict the data only to referring traffic source and also a criterion to order the data. Therefore, we make use of additional parameters to request this data feed.

```
https://www.google.com/analytics/feeds/data?ids=<profileid>&dimensions=ga:source&metric  
s=ga:visits,ga:newVisits,ga:transactions,ga:transactionRevenue&start-date=<startdate>&end-  
date=<enddate>&max-results=5&filters=ga:medium%3D%3Dreferral&sort=-ga:visits
```

Section 6. Once more, in this section we make use of additional parameters to show the top 5 keywords based on the number of visits and the top 5 quantity of products sold. Below there are the three data feeds requested to get such information.

```
https://www.google.com/analytics/feeds/data?ids=<profileid>&dimensions=ga:keyword&met  
rics=ga:visits,ga:transactionRevenue,ga:transactions,ga:bounces,ga:entrances&start-  
date=<startdate>&end-date=<enddate>&max-results=5&sort=-ga:visits
```

The second part of the section 6 requires additional processing because we need to get the information of the top 5 products in the current period and later compare them with the one obtained in the last period. Therefore, the second third data feed of the section 6 must be called 5 times.

```
https://www.google.com/analytics/feeds/data?ids=<profileid>&dimensions=ga:productName  
&metrics=ga:itemQuantity&start-date=<startdate>&end-date=<enddate>&max-  
results=5&sort=-ga:itemQuantity
```

```
https://www.google.com/analytics/feeds/data?ids=<profileid>&dimensions=ga:productName  
&metrics=ga:itemQuantity&start-date=<startdate>&end-date=<enddate>&sort=-  
ga:itemQuantity&filters=ga:productName%3D%3D<productfrompreviousDataFeed>
```

The next step after retrieving all the data feeds is to format the information to create the charts and graphics. Using PHP and JavaScript the table object called DataTable is created for each chart.

5.2.2.2 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 5.2.2.1, and a number of options that help to define the graphical appearance (e.g background, width, height, 3D option). The next is an example of the graphic creation for the percentage of visits per traffic source using the Google Visualization API for our Marketing dashboard.

First of all, we need to include the visualization API library and load the type of graphics that we are going to use. Since the Marketing Dashboard makes use of the table, bar chart, pie chart and a motion chart packages we need to load them. The next is a pseudo code for the algorithm implemented when creating a pie chart.

```
Call to Google Visualization API library  
Load graphic package (piechart)  
//creating an DataTable object  
dataPieVisits ← DataTable();  
number_rows ← size of feed  
add number_rows to dataPieVisits Object  
for each (element in feed)  
    set value of feed in dataPieVisits Object  
//Creating a Pie Chart  
chartPieVistis ← new PieChart(<div id area to show chart>)  
//Drawing a Pie Chart  
chartPievisits ← Draw(<feature parameters>)
```

5.2.2.3 Display Dashboard

Finally the graphics, tables, labels and other components are placed in the dashboard and displayed to the user. To display the chart we use divs in HTML. The pie chart created in the step 5.2.2.2 can be displayed in the body page with the simple called to the div id area specified when creating the pie chart.

5.2.3 Suggestion Module

5.2.3.1 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. Such data is stored in our database only if the user has specified in the profile that WADAS is allowed to store persistent information to display the Suggestion Module.

As was shown in the design section 4.1, the model package consists of the core items of the system. Part of these items are the profiles, and for the Suggestion Module the historic data.

Therefore, the WADAS database contains a table of profiles associated to the users, a table of historic data and an additional table that maintains the type of metrics stored in the historic data table.

5.2.3.2 Process Data

In this step variables such as the configuration of goals in the account, default page, ecommerce option, campaigns and keywords are evaluated making use of the information stored in the historic data table.

After processing the data and finding a number of suggestions for the user, they are passed from the report controller to the web page to be displayed.

NOTE: The evaluation of the data is made based on the comparison between the newest period stored against the others.

5.2.3.3 Display Suggestion Module

In a separate screen to the dashboard are displayed the suggestions that the user should follow to improve the analysis of the website's performance.

6 Test Phase and Results

This chapter describes the tests and results developed for the Design and Implementation of the WADAS System.

6.1 Design Tests

In order to test the dashboard designs we created four (one for each type of dashboard) mock-ups in paper, which were displayed in the chapter 4. Each mock-up 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 Appendix C illustrates the four questionnaires designed to evaluate the dashboards. The group of 10 people that responded to the test was composed of Web Analytics experts, designers, business directors, developers and others, which were chosen because they own a site on Internet. The time spent in each dashboard to answer the questionnaire was measured.

Trying to follow a heuristic evaluation [Kaushik, 2007a, 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.

6.1.1 Limitations

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 charts, change of parameters and drill-down of data could not be tested either in this stage. This kind of features requires the implementation of the dashboard to be tested. On the other hand, this fact helped to test simplicity, synthesis, use of space and identification of key metrics.

6.1.2 Results

As results from these tests, most of the answers for the multiple choice questions were right. In contrast, the most difficult part to evaluate and that resulted in design changes were the comments made by the users. Some evaluators provided comments about the colours, type of graphs and signs used in the dashboard, suggesting new alternatives.

The table 1 shows the total number of correct answers for each question classified by dashboard. Four of the evaluators had troubles answering the question 3 of the Sales Dashboard. When they were asked about the reasons for the selection they did, which was not the correct one, we found that the option b for this question was not clear.

Dashboard/Questions	Q1	Q2	Q3	Q4	Q5	Q6	Total
Sales	10	10	6	10	9	9	54
IT-Design	9	9	10	10	10	10	58
Marketing	8	10	9	10	10	10	57
Management	10	10	9	10	10	8	57

Table 1 Questionnaires Results – Correct Answers

The average of time spent for the total four questionnaires was 29 minutes. The overall time spent for each evaluator in the four questionnaires is displayed in the table 2.

Evaluator	Time Spent
E1	0:23
E2	0:35
E3	0:39
E4	0:25
E5	0:26
E6	0:40
E7	0:28
E8	0:25
E9	0:24
E10	0:31
Total Avg Time	0:29

Table 2 – Average time Spent Solving the Questionnaires

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.

These are some of the observations provided by the evaluators and that resulted in changes of the dashboard designs:

- a. What does it mean CDR in the dashboard?

CDR is an abbreviation for Current Date Range, which is explained at the beginning of the dashboard, in the header section. However, for the evaluators was less obvious this nomenclature. Also they suggested only using 'Current Period' was enough through the charts where it is not possible to determine in a glance that the data shown refers to the current period.

- b. In the sales dashboard the idea of the pie charts is good but the association of the labels coloured in the previous table with the pie is not really clear.

For this dashboard we wanted to save space using different colours that identify the type of traffic source without repeating the labels in the pie charts. However, most people are used to look for the small labels in the right hand or top of the graphics.

- c. In the Marketing Dashboard would be great to see the actual values corresponding to the percentage shown in the traffic source bar chart. Also it is easier with the pie charts.

Unfortunately one of the interactive functionalities that could not be assessed during these tests was the possibility to see the actual value corresponding to the percentages when the user clicks over the chart. For this observation we did not change anything in the implementation but we explained to the evaluator that it was considered in the interactivity features.

- d. The content section of the dashboard is smaller than the header. That tabulation is not necessary because it is clear with the separator line what the header is.

For this suggestion, the tabulations were deleted. In the current implementation the header, content and footer are aligned with the same width.

- e. If the revenue is always a value with more than four digits, probably it is not necessary to show the decimals

The revenue can be 0 or a big number. According to the common mistakes mentioned in The Chapter 2, section 2.1 we considered that two decimal digits are enough to compare information. However, analysing the suggestion provided by the Evaluator in the case of the revenue probably is not significant those decimal digits as it is for percentage values. Therefore, we decided to show the revenue without the two decimals.

6.2 Implementation Tests

For the implementation have been done internal tests in order to ensure the data displayed is coherent with GA. 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. The next are the elements and actions tested.

6.2.1 Security Issues

As was stated before, WADAS includes two authentication and authorization steps. First of all the user needs to login to the WADAS System with a user and password that is stored in our database. Subsequently, in order to access the GA API, the user needs to have and to be authenticated in the GA account. In order to test different security issues that are involved in the use of the application, we assessed.

- Creation and Deletion of sessions for WADAS
- Login and Logout process to WADAS
- Requesting of the single-use token to Google
- Upgrading of the token for long-lived session using the AuthSubSessionToken method.
- Attempting to use different GA accounts in WADAS that do not correspond to the user registered in our system.
- Verification of denial of access to other profiles when the user does not have a role of Administrator.

- Use of cryptographic hash function to store user password in the WADAS database

6.2.2 Requesting and Retrieving Data

6.2.2.1 Validation of charts that cannot be displayed because the feed did not return any data.

To test the display of GA accounts that do not provide enough data because of properties not configured in GA, we created a user with access to three GA accounts.

1. The first GA account was configured to use ecommerce and AdWords.

The dashboard was displayed correctly.

2. The second GA account was configured with three goals, but without ecommerce nor use of AdWords.

The dashboard was displayed correctly, but with some values in 0 because there were not products to display or transactions.

3. The third GA account did not include any advanced configuration.

The dashboard was displayed correctly, but without some information related to ecommerce or revenue.

6.2.2.3 Validation of feeds bad formed that can generate errors in the process

To validate possible feeds with a wrong structure, we created fictitious arrays with a similar structure to the feeds passed by Google but with some data missed. When the operation could not finish successfully, the controller returned an error message and the user was redirected to an error page.

6.2.2.4 Redirection to error page for any other error not identified

For any other errors that could not be identified, in the tests the users were redirected to a friendly page created in the in the bootstrap displaying the next message:

“An error has occurred, we apologise for the inconvenience. Please contact the WADAS administrator to report the problem (daniela.fernandez@gmail.com)”

6.2.3 User Interaction and Interface

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.

These tests were run in Firefox, Chrome, Opera and Internet Explorer 7 – 8. Hence could be assessed the look and feel and functionality for the WADAS website.

7 Conclusions

Summary

In this dissertation we have attempted to propose a solution for the problem that many users face when analysing the measurements and indicators to improve the performance of their websites. Hence, we have presented a Web Analytics Dashboard and Analysis System (WADAS) as a solution to this problem. We did this by studying and presenting a set of definitions for the main elements of Web Analytics. We also achieved our goal analysing the creation of successful dashboards as a technique to obtain the information required to take strategic decisions for the business. By identifying the best practices and common mistakes in designing Web Analytics dashboards we could start the creation of our own work.

There are a number of ways that the problem could be approached; therefore various methods for designing a solution. However, given our timeframe, we decided to base our solution solely on the requirements of one particular client. This fact required setting boundaries on the requirements specification that we could address. Hence, having the user requirements defined, we presented the software development process followed to implement our tool. The process included the analysis of the software requirements, which were classified in functional, not-functional and design requirements. Later, we described the design proposed for the system based on the Model – View – Controller (MVC) design pattern, which allows us to separately manage the behaviour of core data, the operations requested from the client and the display of the information. Furthermore, we explained the process in detail to design the Web Analytics dashboards for WADAS and the Suggestion Module.

Subsequently, we gave an understandable idea about the process followed to pass from the design to the implementation of WADAS. The description of this phase included the fundamentals of the main components involved in the implementation of the Web Analytics Dashboard and the Suggestion Module, bearing in mind that to get the final outcome we made use of third party applications.

From a research and analysis perspective, we intended to prove that successful dashboards linked to a suggestion module are an effective tool to minimize the time that organizations spend understanding complex reports. It could be verified with the tests performed to evaluate the Web Analytics dashboards designed. Such tests gave us the sufficient tools to know in a real environment what the users are looking for in a dashboard, what make them feel confused, and how long they take trying to understand a graphic report.

After the design and implementation tests we also found that even the simplest and explanatory Web Analytics tool cannot provide, on its own, the right actions to take in order to improve the website performance. The tools can possibly supply good options based on different configuration patterns and calculations. However, due to the different nature of each business and the external factors that affects the visitor's behaviour in the websites, web analysts are needed to provide insights and more accurate decisions. On the other hand, 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.

Future Directions

The scope of the project was bounded 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 would be the implementation of the other three dashboards in WADAS. The tool has been implemented having considerations of scalability; therefore, we expect to add new functionality to the application. In the same way, the analysis and solutions presented in this dissertation are not definitive and undoubtedly much remains to be done.

First of all, having more time to implement and measure the effectiveness of the four dashboards in a real environment could certainly result in adjustments to the designs and probably more ideas to add flexibility to their configuration without losing their purpose.

Secondly, improvements in the results obtained from the Suggestion Module could be implemented. At the moment, 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.

Finally, an approach to make the Suggestion Module more robust is to include the use of artificial neural networks or other similar models of processing to produce more accurate and actionable responses that not only analyse the GA configuration patterns but also external information provided by the organizations.

References

Brath, R and Peters, M. 2004. Dashboard Design: Why Design is Important. *InfoManagement Direct*. [Online] 15 October 2004. [Cited: 15 March 2009.] <http://www.information-management.com/infodirect/20041015/1011285-1.html>.

Burby, J and Shane, A. 2007. *Actionable Web Analytics: Using Data to Make Smart Business Decisions*. s.l. : John Wiley and Sons, 2007.

Burby, J, Brown, A and Committee, WAA Standards. 2007. Web Analytics Definitions Version 4.0. *Web Analytics Association*. [Online] 16 August 2007. [Cited: 17 March 2009.] <http://www.webanalyticsassociation.org/attachments/committees/5/WAA-Standards-Analytics-Definitions-Volume-I-20070816.pdf>.

Carson, M. 2009. 2009. *AdCenter Community*. [Online] Microsoft Advertising Community, 12 March 2009. [Cited: 13 March 2009.] <http://community.microsoftadvertising.com/blogs/analytics/archive/2009/03/12/adcenter-analytics-beta-to-close.aspx>.

Clifton, B. 2008. *Advanced Web Metrics with Google Analytics*. Indiana : John Wiley and Sons, 2008.

Eckerson, W. 2005. *Performance Dashboards: Measuring, Monitoring, and Managing Your Business*. s.l. : John Wiley & Sons, 2005.

Fernandez, D. 2009. New Features Added to Motion Chart, Google Analytics Results. *Mangold Sengers Official Blog - Google Analytics Results*. [Online] Mangold Sengers , 18 March 2009. [Cited: 18 March 2009.] <http://www.googleanalyticsresults.com/2009/03/new-features-added-to-motion-charts.html>.

Few, S. 2005. Intelligent Dashboard Design, Information Management Magazine. *Perceptual Age*. [Online] 1 September 2005. [Cited: 19 March 2009.] http://www.perceptualedge.com/articles/dmreview/intelligent_dashboard.pdf.

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

Gibbons, K. 2008. Google Analytics metrics to enhance your AdWords campaigns. *SEOptimize*. [Online] 21 February 2008. [Cited: 12 March 2009.] <http://www.seoptimise.com/blog/2008/02/5-google-analytics-metrics-to-enhance-your-adwords-campaigns.html>.

Gillis, J. 2008. The new features are now available in your account! Let's take a deeper look, starting with Motion Charts. *The Official Google Analytics Blog*. [Online] Google, 10 November 2008. [Cited: 15 March 2009.] <http://analytics.blogspot.com/2008/11/new-features-are-now-available-in-your.html>.

Google. 2009. Google Visualization API. *Google Code*. [Online] 21 April 2009. [Cited: 23 April 2009.] <http://code.google.com/apis/visualization/documentation/queries.html>.

Kaushik, A. 2007a. *Web Analytics: An Hour a Day*. Indiana : John Wiley and Sons, 2007a.

Kaushik, A. 2007b. Web Analytics Standards: 26 New Metrics Definitions. *Kaushik Avinash Blog*. [Online] 23 August 2007b. [Cited: 14 March 2009.] <http://www.kaushik.net/avinash/2007/08/web-analytics-standards-26-new-metrics-definitions.html>.

Kaushik, A. 2007c. Five Rules for High Impact Web Analytics Dashboards. *Kaushik Avinash Blog*. [Online] 12 March 2007c. [Cited: 18 March 2009.] <http://www.kaushik.net/avinash/2007/03/five-rules-for-high-impact-web-analytics-dashboards.html>.

Kaushik, A. 2008. The “Action Dashboard” (An Alternative to Crappy Dashboards). *Kaushik Avinash Blog*. [Online] 30 April 2008. [Cited: 19 March 2009.] <http://www.kaushik.net/avinash/2008/04/the-action-dashboard-an-alternative-to-crappy-dashboards.html>.

Levene, M. 2006. *An Introduction to Search Engines and Web Navigation*. London : Addison-Wesley, 2006.

Mortensen, D. 2008. A Reflection on Google, Microsoft and Yahoo Web Analytics Dashboards. *Visual Revenue*. [Online] 11 September 2008. [Cited: 09 March 2009.] <http://visualrevenue.com/blog/2008/09/google-microsoft-yahoo-web-analytics-dashboards.html>.

Petterons, E. 2004. *Web Analytics Demystified: A Marketer's Guide to Understanding how Your Web*. s.l. : Celilo Group Media, 2004.

Saifee, Q. 2005. RoadMap - Dashboards. *WIT Inc*. [Online] 19 May 2005. [Cited: 19 March 2009.] <http://www.witinc.com/resource/attach/124/RoadShowMay18192005Quaid.pdf>.

Sostre, P and LeClaire, J. 2007. *Web Analytics for Dummies*. s.l. : For Dummies, 2007.

Teixera, J. 2008. Five Google Analytics metrics that you're probably not using. *Morevisibility Leaders in Search, Design & Interactive Marketing*. [Online] 11 September 2008. [Cited: 14 March 2009.] <http://www.morevisibility.com/analyticsblog/five-google-analytics-metrics-that-youre-probably-not-using.html>.

Tidwell, J. 2006. *Designing Interfaces: Patterns for Effective Interaction Design*. [ed.] O'really. 2006.

Tyler, M and Ledford, J. 2006. *Google Analytics 2.0*. Indiana : John Wiley and Sons, 2006.

Vermehren, C. 2008. Designing Web Analytics Dashboards. *The Art of Web Analytics*. [Online] 04 April 2008. [Cited: 09 March 2009.] <http://theartofwebanalytics.com/>.

Appendix

Appendix A – Main Use Cases Definition

Package	User Management
Id Use Case	UC-USR-01
Description	Validates and authorize username and password
Actors	User and Administrator
Preconditions	
Primary Scenario	Member requests to sign in. System displays form: Username (email) Password Member enters login and password Submit Authorise member to use WADAS.
Secondary Scenario	After step 4: Login not found. A.1. Extension point to UC-USR-03 B. After step 4: Invalid password B.1 Return to step 2. C. From step 1: User forgot password C.1 Call use case “Remember password” C.2 Return to step 1
Extends	Login to GA Account

Package	User Management
Id Use Case	UC-USR-02
Description	Member logout from WADAS
Actors	User and Administrator
Preconditions	Member has to be logged.
Primary Scenario	Member requests to logout. System delete Member’s temporal information Menu is redisplayed only with the public options.
Secondary Scenario	

Package	User Management
Id Use Case	UC-USR-03
Description	Guest can register in WADAS
Actors	Guest User
Primary Scenario	Guest user requests to register in WADAS. System displays form to fill

	Desired user name (email used for the GA account) First name Last name Password Phone Number Type of user by default Store Data (Yes/No) Guest user enters the information required Guest user saves the information by clicking on 'Submit' button.
Secondary Scenario	After step 2: The guest user enters a login that already exists. A.1 Return to step 2. After step 2: The guest user enters an invalid email address. B.1 Return to step 2. After step 2: The guest user didn't fill the information required (fields that cannot be left blank). C.1 Return to step 2, showing the fields where the information is required.
Notes	After submit the information the Guest become to be a Member of WADAS

Package	User Management
Id Use Case	UC-USR-04
Description	User updates profile
Actors	User
Preconditions	User has to be logged.
Primary Scenario	User requests to update the profile The System displays a form to update with current member's information. Password First Name Last Name Store Data (Yes/No) Phone Number Type of user by Default (Marketing,Sales...) Status (Active, Inactive) User updates the information available to change. Submit
Secondary Scenario	After step 4: Invalid information A.1. System displays invalid fields A.2. Return to step 2

Package	User Management
Id Use Case	UC-USR-05
Description	Administration Staff updates User Profile
Actors	Administrator
Preconditions	Administrator has to be logged.
Primary Scenario	1. Administrator requests the updating of a Member 2. System requests to choose a member to update. 3. Administrator chooses a member. 4. The System displays a form to update with current member's information. Password

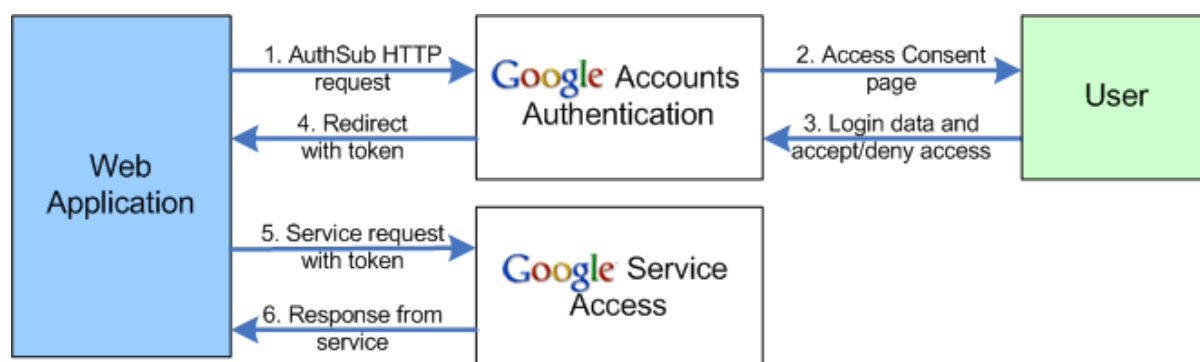
	First Name Last Name Address Phone Number Role Status (Active, Inactive, Blocked) 5. Administrator updates the information available to update. 6. Submit
Secondary Scenario	After step 6: Invalid information A.1. System displays invalid fields A.2. Return to step 4

Package	Data Reporting Services
Id Use Case	UC-SRV-01
Description	Member requests to see Dashboard
Actors	Member
Preconditions	Member is authenticated with the GA account.
Primary Scenario	1. Member issues display dashboard. 2. System requests selection of Profile. 3. Member selects a profile from the list. 4. System displays the Dashboard.
Includes	Fill Selection Criteria

Package	Data Reporting Services
Id Use Case	UC-SRV-02
Description	Member requests to see Suggestion Module
Actors	Member
Preconditions	Member is authenticated with the GA account.
Primary Scenario	1. Member issues display suggestion module. 2. System requests selection of Profile. 3. Member selects a profile from the list. 4. System validates permissions of persistent data storage in the user profile 5. System displays the Dashboard.
Secondary Scenario	After step 4: System found that does not have permissions to store historical data A.1 Return to step 1
Includes	Fill Selection Criteria

Appendix B – Google Authentication Process (AuthSub)

1. When the web application needs to access a user's Google service, it makes an AuthSub call to Google's Authentication Proxy service.
2. The Authentication service responds by serving up an Access Request page. This Google-managed page prompts the user to grant/deny access to their Google service. The user may first be asked to log into their account.
3. The user decides whether to grant or deny access to the web application. If the user denies access, they are directed to a Google page rather than back to the web application.
4. If the user grants access, the Authentication service redirects the user back to the web application. The redirect contains an authentication token good for one use; it can be exchanged for a long-lived token.
5. The web application contacts the Google service with a request, using the authentication token to act as an agent for the user.
6. If the Google service recognizes the token, it supplies the requested data.



Appendix C – Questionaries

Dashboard Design Assessment

(To meet the needs of Web Analytics users)

Dear Evaluator,

In advance, thank – you for taking the time to complete these questionnaires in order to assess the design of 4 Web Analytics Dashboards. Your response, together with the response from others will help to improve the effectiveness of them for different businesses.

Before you start you should know that:

- The time for answering each questionnaire will be taken.
- You are the assessor. Please remember that with this questionnaire we are testing the Web Analytics Dashboard design, we are not testing your knowledge in Web Analytics, Sales, Marketing or Programming. Therefore, if some of the answers are wrong, probably the dashboard has failures in its design.
- At the end of each questionnaire you are able to provide additional comments, observations and report experiences with other dashboards.

Scenario

A t-shirt company has used Web Analytics tools to measure its website performance. The business has different products and has implemented AdWords to display ads in Google and advertising network. However, they are using other marketing strategies such as social media (twitter, facebook and flickr) and email marketing to generate more revenue. The organization desires to take strategic decisions based on the data presented by WADAS. In order to do that, the business director distributed a dashboard in 5 departments (IT, Design, Marketing, Sales and Management).

Basic Information

First Name	
Last Name	
Occupation	
Initial Time	
End Time	

Instructions

The next is a multiple choice questionnaire; therefore, only one option can be selected. Please tick with 'x' the appropriate answer.

e.g. The total number of items sold for the product 'logo shirts' during this period was

- a. 25
- b. 12
- c. 14
- d. 12

IT-DESIGN DASHBOARD

1. A new flash component has been included in the Website and the team needs the test how well it works in different browsers. According to the graphs which are the operating system and browser most used by the users?

- a. Windows and Firefox
- b. Windows and Safari
- c. Linux and Explorer
- d. Windows and Explorer ✓

2. Which one of the next affirmation is true?

- a. The website receives more visits coming from search engines than referrals and direct traffic.
- b. The percentage visits from Search engines and Referring together is bigger than the direct traffic.
- c. Traffic from search engines and referring are generating more revenue than direct and other source traffics.
- d. a and c are correct. ✓

3. Visitors using Telstra and Optus have a better behaviour in terms of revenue produced than Loyal Inc.

- a. Yes ✓
- b. No

4. Between the top 5 pages from the site, which ones have more value?

- a. checkout.html, thankyou.html and products.html
- b. checkout.html, index.html and shopcart.html
- c. index.html, shopcart.html
- d. thankyou.html, checkout.html and shopcart.html ✓

5. The total number of visits where visitor made use of the internal search was:

- a. 62,027
- b. 54,000
- c. 66,313 ✓
- d. 68,000

6. Which network vendors have an average time on site greater than 4 minutes?

- a. Telstra, Optus and Loyal inc
- b. Optus, Loyal inc and inet limited
- c. Telstra, inet limited and Optus ✓
- d. Level 3 com, Telstra and Optus

Comments and Observations

MARKETING DASHBOARD

1. Most of the referring visits are coming from:

- a. sop.ebay.com.au
- b. shop2bot.com ✓
- c. buy.com.au
- d. cnet.com.au

2. From the top 5 products which one (s) have better performance in the current period.

- a. logo shirt and formal shirt
- b. metal shirt and Italian shirt
- c. industrial shirt and Italian shirt
- d. a and b. ✓

3. Which are the keywords more searched and that at the same time have generated more revenue?

- a. tshirtab.com and formal t-shirts
- b. industrial t-shirt, logo shirts and formal t-shirts ✓
- c. polo shirt
- d. tshirtab.com and industrial t-shirt

4. Industrial shirt is one of the products most sold but it had a better performance last period.

- a. Yes ✓
- b. No

5. In average how many pages are viewed in each visit.

- a. 3 to 4 ✓
- b. 6 to 9
- c. 1 to 2
- d. 5 to 6

6. Which one of the next affirmation is true?

- a. The conversion rate for this period is lower than 1% but it's better than the one in last period. ✓
- b. The average time on site (ATOS) is 5 seconds
- c. The number of items sold for the product 'formal shirt' in this period was 21
- d. The trend shows that the visits are decreasing during the last 12 months.

Comments and Observations

SALES DASHBOARD

1. Which source has generated more revenue this period (date range) and which one has had a significant increasing the last 12 months?

- a. Direct and Other
- b. Direct and Referring
- c. Search Engines and Referring ✓
- d. Other and Search Engines

2. In the last 12 months the number of visits has increased; however, most of the clients that make a purchase are new visitors.

- a. Yes ✓
- b. No

3. After comparing the percentage of visits and revenue per traffic source you have found that:

- a. Most of the direct visits don't generate revenue,
- b. People referred from other sites, blogs, social media, and emails generate more revenue although the number of visits is much lower than direct traffic.
- c. Search engines and Direct traffic are the key to increase the revenue
- d. a and b are right ✓

4. Industrial and formal shirts are the products more sold in this month.

- a. Yes, and logo shirt is the third product with more items sold. ✓
- b. No, industrial and formal shirts have generated good revenue but can be products with more items sold.
- c. Yes, but polo shirt has generated more revenue
- d. a and c are right

5. The percentage of visits that results in a purchase is:

- a. 0.39%
- b. 0.93% ✓
- c. 33.7%
- d. 10.3%

6. I can measure how well the site fulfils the objectives set in this dashboard

- a. No, because the Goals are not set ✓
- b. Yes, with the AdWords revenue percentage
- c. Yes, comparing the total revenue from the last month and this month.
- d. Perhaps, analysing the top 5 products.

Comments and Observations

MANAGEMENT DASHBOARD

1. The visits from which traffic source have generated more revenue?

- a. Referring
- b. Search Engines ✓
- c. Other
- d. Direct

2. Most of the visits come from

- a. Direct Traffic ✓
- b. Referring sites
- c. Other Traffic
- d. Search Engine Traffic

3. The ads were shown more in the last period but more people click on them in the current period.

- a. Yes ✓
- b. No

4. How many transactions were registered this period?

- a. More than 1,500 ✓
- b. More than 2,000
- c. Less than 1,500
- d. More than 2,500

5. The Return of investment for this period generated by AdWords was better than the last one.

- a. Yes
- b. No ✓

6. Bearing in mind that Bounce Rate is the percentage of visits in which the person left the site from the landing page, with the data displayed we can conclude that:

- a. The visit quality has improved the last 12 months
- b. The visit quality has decrease in a 30%
- c. The visit quality is less than 30%
- d. a and c are correct ✓

Comments and Observations