The birth of a new industry: entry by start-ups and drivers for firm growth the case of web analytics software

Keij, M.P.A.

Award date:
2008

The birth of a new industry: entry by start-ups and drivers for firm growth

The case of web analytics software

MSc Thesis Technology and Policy TU/e

Matthijs Keij
ID: 478985
August 2008

Supervisors

dr. A. Nuvolari	Faculty of Technology Management
dr. ir. R. Bekkers	Faculty of Technology Management
B. Nieme	Managing director of Adversitement

Section of Technology and Innovation Policies
Faculty of Technology Management
Eindhoven University of Technology
The Netherlands

PREFACE

Visitor engagement is one of the buzz words the web analytics market is famous for. It indicates the extent to which a website is able to attract and hold an active visitor base. Key performance indicators to define engagement are time spent on site, visitor interaction metrics and lifetime value. These key performance indicators can also be applied to the process of writing this thesis, with some slight adaptations. Let's start with the first metric, time spent on thesis. The combination of a full-time position at Adversitement and writing this thesis has been a guarantee for a roller-coaster ride with regard to the energy put in work and put in study. I have to thank all of my supervisors, Alessandro Nuvolari, Rudi Bekkers and Bob Nieme, for their very flexible cooperation and the helpful and quick replies to my inquiries.

Visitor interaction indicates the vast list of resources I have used, both online and offline, to perform research on this subject. Many valuable resources are available, but information is often scattered or too vendor-specific. Interaction with colleagues from Adversitement, students from the TU/e and many business relationships have helped tremendously in achieving the research goals. Without them I would not have had access to many industry reports, customer surveys and best practices in web analytics.

Ultimately, this thesis has definitely contributed to my personal lifetime value. One of my trademarks is the fact that I love to search for answers. Fortunately this thesis provided enough challenges to do so. In addition I hope to share my knowledge even more by continuing to present web analytics workshops at universities and writing about web analytics related topics. The last point of this preface brings us back to the rollercoaster ride. I have to thank my girlfriend Annemiek for her support, advice and patience during the writing of my thesis. Even when the last corrections have been made sitting at a camping table in the Pyrenees, she has been very motivational and inspiring.

EXECUTIVE SUMMARY

This research analyzes the web analytics sector, an emerging market characterized by a multitude of innovations. Web analytics is defined as the measurement, collection, analysis and reporting of Internet data for the purposes of understanding and optimizing Web usage. For the analysis of this sector the history, the underlying technologies, policies and regulations and the largest web analytics vendors have been studied. The goal of this research is to assess which drivers are responsible for firm entry and firm growth in the web analytics sector. Drivers for firm entry have been determined by looking at the web analytics history and the characteristics of firms present in the early stages of sector development. Drivers for firm growth have been determined by analyzing the growth path of leading start-ups which have managed to sustain or develop their market position. The role of technological innovations on market development is interwoven in this analysis.

The collection of web analytics data can be done by parsing web server log files, analyzing all network packets sent to and from a server (packet sniffing) and by executing JavaScript code in the browser of a web visitor. The first web analytics tools used a logfile analysis to collect data, later followed by commercial offerings using JavaScript as the prevalent method for data collection. In general the JavaScript data collection method is better suited to collect data from visitor interaction on a website. The commercial offerings start to gain market share from around 1997. The year 2002 can be seen as the starting point for rapid developments in the web analytics sector. Market demand grows and web analytics vendors develop many innovative new features. The market development since 2004 is characterized by acquisitions of competitors and web analytics related companies. From a technological point of view web analytics is becoming the hub of all information flow with a strong focus on integration of web analytics with other data sources. Policies and regulations regarding web analytics have mainly been derived from existing privacy laws and do have a limited influence on firm entry and growth in this sector.

The sector analysis has been split in two different periods, first the period from 1997 to 2003 and secondly the period from 2004 to 2007. The first period can be demarcated by the "feature race" among web analytics vendors. Entry barriers are still low, customer demand starts to increase and vendors very actively release new versions of their products and try to outpace competitors with new features. Patents and market concentration have been analyzed and do not seriously hamper firm entry or firm growth. Main drivers for firm entry and growth are the increasing market opportunities and the competitive advantage by offering unique features. The second period shows that new entrants face some strong entry barriers, such as substantial capital requirements in building a robust and scalable network for data collection and analysis. Drivers for growth are the development of web analytics platforms which provide additional products and services related to web analytics. Those services include web analytics training, implementation support and business consultancy. Additional drivers for firm growth are large investments in research and development, skilled personnel and an extensive market strategy. Pricing strategies include offering a free tool, but only Google Analytics succeeds to some extent in competing with the top vendors.

This research identifies the most important drivers for firm entry and for firm growth in the web analytics sector. Main drivers for firm entry are market demand, budget allocation by website decision makers and the ability to offer software as a service (SaaS). Main drivers for growth include large investments in technology, turn data into actionable analyses in the user interface and access to funding to provide a scalable and robust network for data collection and analysis. New technologies, such as cloud computing and online storage can help to foster firm entry and growth by providing a scalable platform for web analytics applications.

TABLE OF CONTENTS

1 INTRODUCTION

The online channel (Internet, email, digital TV, etc.) has become an essential part of many people's lives. Internet is used to search for information, to read the news, to manage bank accounts, to compare and purchase products, to share ideas and opinions and the list goes on. To properly respond to the demand of website visitors it is important to understand their behavior. One of the methods to accomplish this is by measuring the usage of a website. Many solutions are available to perform this task, depending on the requirements and the budget of a customer. The market for these solutions is very dynamic and emerged only ten years ago. This makes it an ideal candidate for a sector analysis of an emerging market. Often data from early stages of market development are not available or hard to retrieve, but given its short history this market has a definitive advantage in terms of information and data availability.

Another contributing factor to this research is Adversitement, a Dutch firm specialized in web analytics and related services. The company has about twenty-five employees and is selling products and services to customers with a strong online presence, such as Vodafone, Funda and Nuon. Adversitement uses tools from different suppliers as well as software developed in-house to meet customer requirements and collect all relevant data about website visitors. Thanks to the close relation to different software suppliers Adversitement has access to a lot of market-specific information and company data. On the other hand a better understanding of market trends, entries and exits is very valuable to Adversitement.

1.1 Reasons for measuring a website

Suppose that an online travel site sells 1000 holiday packages a year. The diagram below shows a hypothetical simplified version of marketing expenses and sales revenues in one year. For the sake of clarity additional costs are not incorporated. Sales are divided in low, mid and high value packages representing different prices and margins on holiday packages.

To get an idea of costs and revenues in one year, have a look at the fictional example of a balance sheet below. This balance sheet is based on realistic figures (from Adversitement customers).

Costs (1 year)		**Revenues (1 year)**			
Marketing		**Packages**	**Items**	**Unit Price**	**Total**
Google Adwords	€ 100000	Low value packages	600	€ 100	€ 60000
Online banner campaigns	€ 80000	Mid value packages	300	€ 400	€ 120000
Email marketing	€ 30000	High value packages	100	€ 1.000	€ 100000
Totals	€ 210000				€ 280000

Under the assumption that costs and revenues mentioned in the balance sheet are the only costs and revenues it is easy to calculate the return on investment (ROI) by dividing revenues by costs. That leads to a ROI of 1.3. Now web analytics come into play to help identify the most valuable campaigns. Web analytics allows marketers to track from which marketing campaign a website visitor originated before ordering a holiday package. The following table (example data) shows the distribution of orders over the various marketing campaigns.

	Low value		Mid value		High value		Total revenues	Costs	ROI
	Items	Revenues	Items	Revenues	Items	Revenues			
All campaigns	600	€ 60000	300	€ 120000	100	€ 100000	€ 280000	€ 210000	1,3
Google Adwords	400	€ 40000	150	€ 60000	20	€ 20000	€ 120000	€ 100000	**1,2**
Online banners	100	€ 10000	100	€ 40000	20	€ 20000	€ 70000	€ 80000	0,9
Email marketing	100	€ 10000	50	€ 20000	60	€ 60000	€ 90000	€ 30000	**3,0**

This table clearly shows the difference in ROI measured on a campaign level. Without web analytics it is very hard to tell which marketing campaign is most effective in terms of ROI. The table shows it might be worthwhile investing more money in email marketing campaigns and reduce money spent on online banner campaigns. In general it will be a combination of total revenue and ROI which will determine the optimal distribution of the marketing budget over various campaigns. Other reasons to measure the behavior of website visitors might be to enhance usability of a website, detect and analyze performance problems and find opportunities to increase sales or customer service.

1.2 Research goals

The goal of this research is to get better insights in the dynamics of an emerging market driven by innovation. Therefore it is important to have data from the early stages of market development. Since the web analytics sector is quite young, more data about the emergence of this sector might be available. Additionally the author of this thesis has considerable working experience in this specific market and access to non-public data sources and customer information. The objectives to reach this goal are the collection of market data from the most relevant companies, relate this data to technological advancements and assess the response from web analytics companies to (changing) market demands.

Although the web analytics sector has many players, this research will only focus on companies selling web analytics tools, leaving consultancies and end users mostly out of the picture. Attention will be paid to those stakeholders, but an analysis of data from all market players would be out of scope for this research.

1.3 Research question

The main research question for this thesis is:

What are the drivers for firm entry and firm growth in the emerging market of web analytics software?

In order to assess which drivers exist for firm entry the characteristics of the firms present in the early stages of this market can be a valuable resource. These characteristics might indicate the importance of specific drivers. Therefore the following sub question has been defined:

What are the characteristics of firms present in the initial stage of the emerging market?

Not all companies have been successful in this market, but some have managed to grow from a start-up to a leading player. Especially this group of companies can contribute to the analysis of drivers for growth. Apparently their actions and reactions in this dynamic market have led to successful company development. To analyze this behavior the next sub question is formulated:

What is the growth path of leading start-ups?

Aside from firm characteristics this market is heavily influenced by technological developments and innovations. Web technologies often emerge faster than existing market leaders have anticipated for, resulting in possibilities for new entrants. To assess the influence of technological innovations the following sub questions has been defined:

What is the role of technological innovations in the development of this market?

1.4 Research demarcation

This research will focus on companies offering software to analyze websites. Most companies have an international scope, since software is often not bound to geographical locations. The majority of the companies included in this research are European or United States based. Although many companies are offering products and services related to the web analytics market, only the software suppliers (referred to as vendors) are included in this research. Consultancy firms and resellers will not be included in the detailed sector analysis, but their role will be described as well as their effect on market developments.

1.5 Research design and methodology

Compared to a related market, for example internet search engines, the market for web analytics is not very well documented and hardly appears in scientific articles. To give a general indication of the popularity of the topic "web analytics" compared to "internet search engines" a search query has been performed on Google Scholar. For "web analytics" just 639 results will appear, by contrast "internet search engines" (queried as: internet "search engine") which renders almost one hundred thousand results. Many industry reports indicate the importance of this sector, yet it hardly appears in scientific literature.

One of the challenges of this research is to find relevant data. Thanks to Adversitement many company data and industry reports are available, but scientific research is limited. One of the possible reasons might be the rapid developments in this market, which can make it difficult to perform a long-term analysis. Various data sources have been used, including books and articles about sector analysis and innovation systems, industry reports about web analytics, SEC filings from various companies and multiple websites and email groups to get a coherent view of the market and to use as many relevant resources as possible. First the technologies related to web analytics will be described. Secondly the history of this market is outlined, followed by a general chapter about policies and regulations. In the actual sector analysis these backgrounds are integrated with the structural framework for the analysis of innovation systems as proposed by Malerba (2004). The analysis is backed up with business cases from Adversitement to close the gap between theory and empirical evidence. Finally the conclusions of the research will be presented.

1.6 Report outline

The report outline will contain a short description of report, illustrated the diagram below which shows the structure of this thesis and the correlation between different chapters.

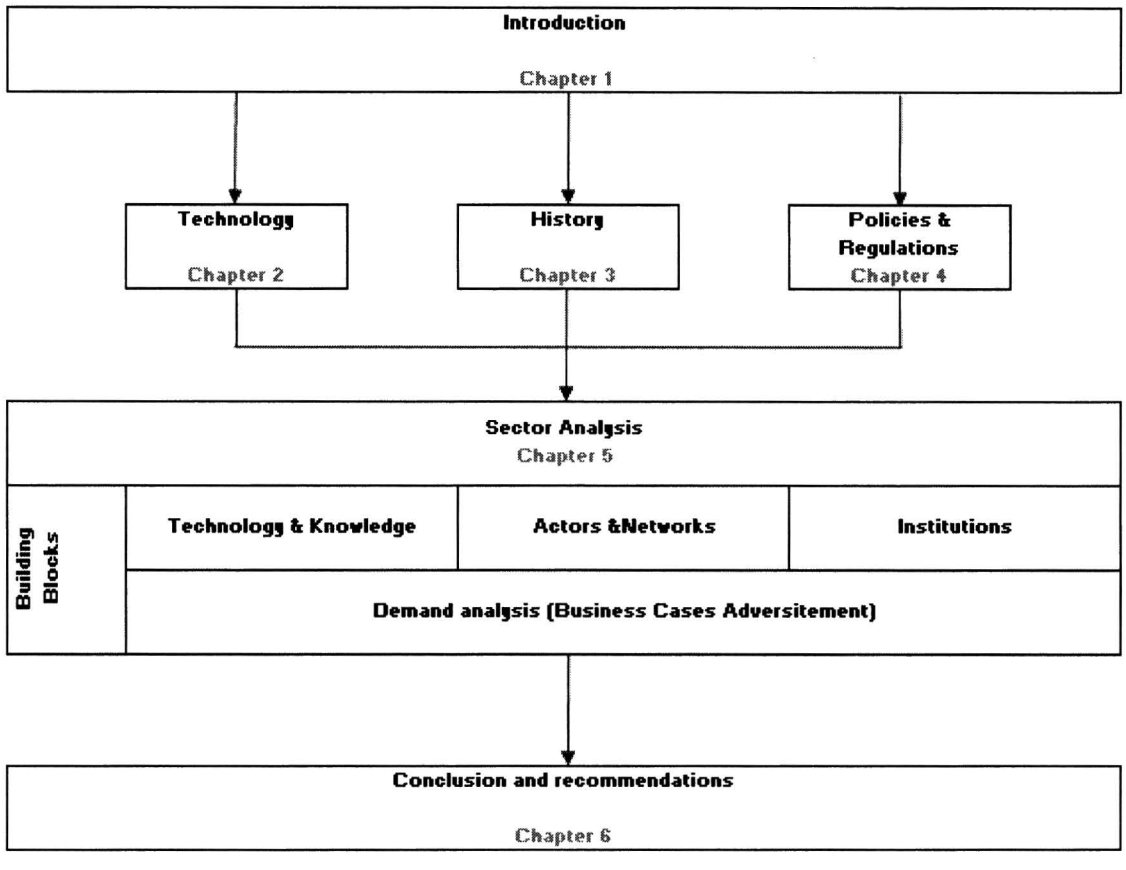

Figure 1: Structure of thesis

2 WEB ANALYTICS TECHNOLOGIES

2.1 Introduction

Web analytics is a very broad concept. To clearly demarcate the web analytics sector it is important to use a solid definition of this concept. This research will build upon the following definition for web analytics, formulated by the Web Analytics Association (WAA):

Web analytics is the measurement, collection, analysis and reporting of Internet data for the purposes of understanding and optimizing Web usage.

Many more definitions exist and some are worth mentioning. A very common, but narrow definition can be found on Wikipedia (Web analytics, 2007) and says: "Web analytics is the study of the behavior of website visitors". A good definition to explain web analytics to non-technical people is: "It is the tool web site owners use to measure the success (or not) of their web site - a web site thermometer" (Whitingjo, 2007). All definitions have at least one thing in common. They indicate the importance of knowing what visitors are doing on a website or online application. There are several reasons why this information is important, think of web site usability, knowing why your online visitors are buying (or not buying) and optimizing customer service. All these reasons are closely related to the nature of the website which will be analyzed. Therefore the next paragraph will present a breakdown into four different types of websites. The chapter ends with a comparison between the three main technologies which are used for in web analytics.

2.2 Website types

The following diagram shows an industry-standard (Eisenberg, 2002) breakdown into four different types of websites. The diagram lists the characteristics and goals of these types.

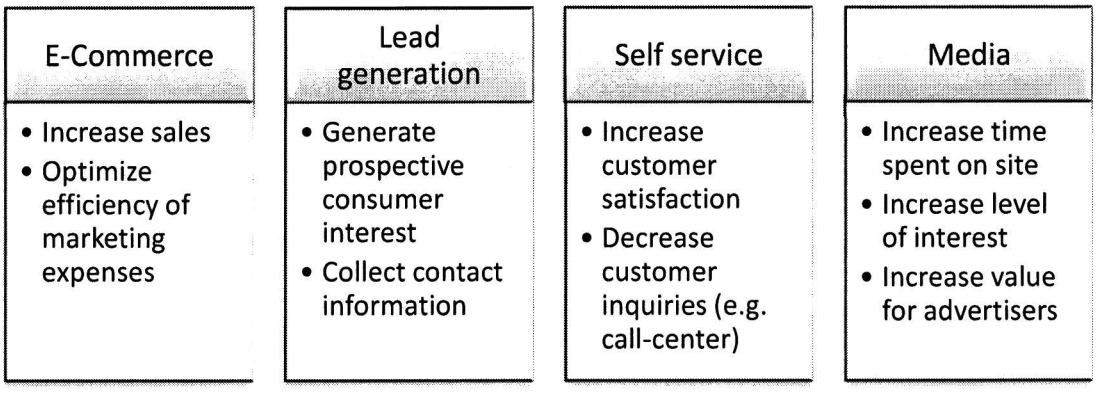

Figure 2: Website types (Eisenberg)

Of course many websites are hybrid, which means that their functionality extends to more than one type of website. For instance a telecom provider website will likely sell mobile phones and subscription plans, but also offer customer services, like an online bill and frequently asked questions. Every type has specific key performance indicators (KPI's). Before the KPI's will be discussed some general definitions of common terms in web analytics will be presented.

2.3 Core definitions

Definitions can vary based on the method which is used to collect visitor data. Most prevalent definitions are included in the definition's chapter at the end of this document (Definitions

). Please refer to this chapter when a specific term is not explained in detail.

2.4 Data collection methods

As indicated in the start of this paragraph, there are multiple ways to measure and collect internet data. The three most common methods will be explained in detail, these are logfile analysis, packet analysis and JavaScript analysis. All three methods measure and collect Internet data in a different way. The illustration in Appendix 1 shows the differences in a simplified drawing.

2.5 Logfile analysis

2.5.1 Methodology

A logfile is a text file which stores information about all requests or transactions to the web server. This information is stored in a specific format (depending on the web server) and can be read and analyzed by a separate program. The amount of information which can be extracted from the logfile depends on how much information is stored by the web server.

2.5.2 Advantages

Since log files are already available, implementation of a logfile solution is relatively easy. Changes in website coding are not necessary. Additionally the web server records every transaction on the website, irrespective of the browser used by the visitor. Visits from robots and spiders will also be counted. This information is very useful for search engine optimization (optimizing a website to get a higher ranking in search engine results).

2.5.3 Drawbacks

The main drawback of logfile analyses is the limited possibility of recognizing unique visitors. If visitors are recognized by IP address, accuracy will be very low (Abraham, 2007). This is due to the fact that many visitors have a dynamic IP address which changes during the month. If this address is used for identification, the results will not be a valid indication of unique visitors. Known web robots and spiders need to be filtered, because they do not represent human visitors. This step requires a lot of testing and analyzing, because many non-human visitors try to identify themselves as human. A good example is a real estate website which is "scraped" for new content by automated agents. It is interesting to analyze this behavior, but it needs to be filtered for a meaningful analysis of visitor behavior.

Due to the nature of logfile analyses, site activity will always be measured server-side. This inevitably leads to impossibilities in the analysis. For example, viewing a Flash movie is only a single request for the web server, but can involve many interactions between the visitor and the Flash movie. This information is not stored in the logfile and thus not available for further analysis. Caching can also pose problems when static pages are served from a caching server. These pages will not be reflected in the logfile of the main server. Technically it is possible to prevent caching of pages, but that can have a dramatic impact on web site performance and visitor perception. Many vendors which have been using logfile software have switched to more hybrid systems (e.g. logfiles and page tagging) which allow for more detailed analyses not possible with just interpreting logfiles.

2.5.4 Economic factors

There are many free logfile solutions (AWstats, Webalizer), which only require a server to run on. However they often have a limited set of features which makes it hard to fulfill all information needs. Commercial solutions involve buying a license and installing the software. In most cases this is a single investment, but some companies charge a unit price based on the number of pages which need to be processed. Another important factor is hardware costs. Storing logfiles only takes disk space and not a lot of processing power. The next step, analyzing and reporting on these logfiles, can be very (database) intensive, depending on the volume of the data. An in-house solution means that a customer needs to have appropriate hardware and perform regular maintenance, which can lead to a steep increase in costs when information needs expand. To combat this problem it is possible to use a hosted solution to take care of hardware and maintenance costs.

For a website which has about 10000 visits a day, a regular database server will be sufficient to collect and analyze log files. This server needs to communicate with the web server, which can lead to data traffic costs when both servers are not at the same location. Appendix 2 shows a cost estimate for a logfile analysis of a website with 10000 visits. The database server which runs the analysis is expected to run for three years. The cost estimate also includes server hosting (the costs to maintain a server and the costs for generating data traffic). Not included are the costs to perform analyses. A common saying in the web analytics market is that tools are only accountable for 10% of the budget. The other 90% is for people who actually perform the analysis and work with the tool. Therefore one can argue about the need for a price comparison of different technologies, but in practice it proves to be an important factor, regardless of other (staff) expenses.

2.6 Packet sniffing analysis

2.6.1 Methodology

Packet sniffing means that all packets which are sent over the network are analyzed. It is also referred to as network or protocol analyzing. The software makes sure that only relevant packets will be recognized and processed. This can be realized by checking the header of each packet to decide if further analysis is relevant. The captured data is stored into a database for additional processing. If website data is served from different platforms all platforms need their own black box for packet sniffing (Pols, 2006). When a load balancer is used to divide traffic among several web servers, every web servers needs its own dedicated packet sniffer.

2.6.2 Advantages

Packet sniffing can collect all data which is sent to and from the network. Instead of reading a logfile with information about web server requests, a packet analysis reveals which information is sent to and from this server. A lot of data captured with packet sniffing is IT related, like request sizes, time to serve a page and login procedures. This information is not (entirely) available with logfile or JavaScript methods. Additionally this technique is not obtrusive to website visitors. They will not notice any impact on their browsing session. Packet sniffing enables the collection of all clickstream (Clickstream, 2007) data which can be used for additional analysis.

2.6.3 Drawbacks

Packet sniffing gives a wealth of information, but a lot of this information is strongly IT-related. Analyzing visitor behavior may prove difficult, because packet sniffing can not collect all client-side data. Another drawback is the fact that the software can capture sensitive information, like passwords, as well. A well structured configuration needs to be in place to get relevant data out of

large data streams. In relation to that scalability can also be an issue. As indicated before, every server that sends information to clients (visitors) needs to be analyzed. Packet sniffers are often dedicated machines and they need to be robust enough to handle and process all traffic.

Finally, it can be difficult for an analyst or investigator to get the right data out of the system, because packet sniffing is strongly bound to the IT department. This can have a negative impact on the speed of analysis and the potential of the system.

2.6.4 Economic factors

Packet sniffing needs specific hardware which runs in front of every operational web server. The more advanced the web infrastructure, the higher requirements for packet sniffing devices are. They need to reliably collect data, but also analyze packets real-time to get only relevant information in the database. Also implementation and consultancy costs tend to be relatively high compared to other methods, due to the extensive configuration and on-going tuning of the system. If an analyst needs an IT specialist to assist in getting information out of collected data, operational costs can rise quickly and diminish profitability of the analysis. A detailed table can be found in Appendix 2. In general packet sniffing tends to be used by larger companies which have resources to install additional servers and require a more advanced analysis of their website behavior, including website performance measuring (which is only to some extent possible with JavaScript). From personal experience there seems to be a trend at (sometimes partly) switching from packet sniffing to JavaScript technology.

2.7 JavaScript analysis

2.7.1 Methodology

JavaScript is a scripting language and widely used on websites. JavaScript runs client-side, which means it is active on the computer of the client (visitor). JavaScript has access to certain browser properties, like screen resolution and operating system. It also enables interaction between a client and a server. By responding to specific events JavaScript can execute functions or notify the server. Google Maps is a well-known example of a website which relies heavily on JavaScript for presentation and interaction.

JavaScript can also be used to collect data from the client and send it to a server. This can be the web server which is serving the content, but also a separate server which primary goal is to collect data. With a small snippet of JavaScript code all relevant information is collected and sent to the server. The predominant method of sending this information to a server is by constructing a 1x1 pixel image which is requested from the server. A query string with all collected data will be appended to this request. The server responds with a HTTP status code and collects the request. Subsequently this information needs to be processed to combine different requests and store information to a database or proprietary data storage.

2.7.2 Advantages

JavaScript analysis generally involves less configuration of the web server than logfile analysis and packet sniffing do. Since the collection of data is client-side, much more information which is available on the website (by interaction with the client) can be sent to the data collection server. For example, a logfile will not show whether a visitor abandoned a form in the first or last field on the same page. With JavaScript this information is available. Caching pages does not have an impact, because even when content is served from a caching server the JavaScript code will still be executed.

Interactive content, such as Flash movies and streaming media, can be analyzed with JavaScript as well. Depending on the implementation specific actions of visitors can be tracked to get a better view of their behavior. Cookies are used to recognize unique visitors. The JavaScript handles the procedures to set, update and delete cookies (which needs to be configured server-side when using other methods). Given the fact that web development is moving more and more to dynamic websites with lots of interactions between the website and its visitors, JavaScript offers the best options to analyze this data.

2.7.3 Drawbacks

JavaScript runs client-side, which means that clients (visitors) can decide to turn off JavaScript and therefore effectively block analysis of their behavior. The acceptance of cookies is also important, because recognizing unique visitors relies on cookies (small text files which reside on the computer of a visitor and contain information about his visit). In general not a lot of people will disable JavaScript, since a lot of websites depend on JavaScript. Cookies however tend to have a bad reputation, while they can be very functional. They allow for storing login information on the computer of a visitor, permitting him not to login again after every visit to a specific website. They are also used to store preferences for presentation or information. On the other hands cookies have been used to track visitors over multiple domains to analyze their website history and display targeted advertisements. Adding a privacy policy to the cookie information and serving a cookie from the same domain as the visitor is browsing will help in increasing cookie acceptance, but numbers of unique visitors are often inflated, due to not accepting cookies.

The implementation can be straightforward by installing a standard script, but when information requirements are complex, the implementation will be more time consuming. Generally many variables in the script are configurable. Filling these variables with the proper values can mostly be done automatically when a content management system is used.

2.7.4 Economic factors

As with the other methods, many solutions exist. A well-known example is Google Analytics, which is free for personal and business use. Other solutions generally charge a fixed price per page view or per event. Many JavaScript solutions collect data on a different server than the web server. This server is specifically configured for handling large amounts of requests. This reduces maintenance costs, since that is the responsibility of the web analytics provider. Exporting this data for further analysis or integration with other systems can be costly, depending on the data storage model and the contract. Updates usually involve placing a new script, instead of installing new software, which can be very economically for larger web sites. Telling the difference between the three described methods with regard to economic factors is difficult, because costs depend on a lot of variables. However the general market trend is that most major customers use page tagging (JavaScript) as their preferred method of data collection.

In theory it is possible to run a proprietary server to perform this analysis, but in practice almost any commercial JavaScript analysis offering is using an ASP model (Applicaton Service Provider). This means that the server is outsourced and will process requests sent over the Internet. The table in Appendix 2 shows a cost analysis for a JavaScript analysis.

Setup costs include page tagging (placing the JavaScript code on the web pages) and software setup to correctly identify all variables passed to the server.

2.8 Hybrid methods

Nowadays more and more tools are available, ranging from free tools like Google Analytics, to commercial offerings, like Clicktracks. Since many tools for performing web analytics are complimentary, a trend can be seen that companies use multiple tools. This observation is based on the current list of customers of Adversitement (January 2008) and market research. In many cases Google Analytics is used in conjunction with a commercial tool. Main reasons to do this are an additional check for data integrity and the tight integration between Google Analytics and Google Adwords (Google's advertising model). Additionally, some tools provide functionality not available in many popular tools, like link click analysis (analyzing where a visitor has clicked on a web page).

2.9 Linking different methods to user characteristics

The choice for a specific data collection method depends on the available budget, the analysis requirements and the option to customize website code. The rise of Google Analytics has enabled many companies to try a full-functional JavaScript web analytics tool. Google Analytics is free to use, although implementing the code might induce costs for updating the website code. Google Analytics

Small-and medium sized businesses will often use JavaScript or logfile methods to collect data about website usage. Especially since the growth of Google Analytics the barrier to use web analytics software has been minimized. Larger companies might use Google Analytics as a backup tool to do trends comparisons with data from their main web analytics tools. In addition to or as a substiture for Google Analytics many small companies still use logfiles, although the entry level logfile solutions are limited in their functionality compared to JavaScript based tools. Many hosting providers include a free logfile based web analytics tool, but it is often not configurable and will not collect all relevant data.

Packet sniffing is mainly used by larger companies, since it requires administrative access to a web server or a separate server to analyze the network packet flow. It is more prevalent in the financial sector, because out-sourcing data collection and analysis is considered to be a security risk by some. There are hardly any open source packet sniffing tools, which means that there is a higher barrier to step into this specific technique. Most companies use packet sniffing web analytics tools also for performance testing, since the packet flow can be analyzed into detail, in contrast to the logfile and JavaScript solution.

Finally, there is a lot of variation in tools which use the same technique. The best example is the JavaScript technique. As said before, Google Analytics is considered to be an entry-level tool and quickly adding features which were previously only available in high-end JavaScript web analytics tools. The current version has its limitations, which means that some companies will prefer alternative (commercial) solutions, like Omniture SiteCatalyst or Clicktracks. Reasons to prefer a high-end solution can be integration of web analytics data with other data sources (e.g. a customer database, warehouse inventory, etc.) and more versatility in exporting and presenting the collected data. In general the adoption of a web analytics tool takes some time and requirements may change over time, leading to a different choice of web analytics tool.

Most of the recent documentation (mainly specialized books and industry papers) will point out that JavaScript is the prevalent technology in web analytics. Although in some cases, like performance monitoring, packet sniffing or a logfile analysis might be the better choice (Sterne, 2002). In most cases, interaction with a visitor using a browser will be analyzed with JavaScript. Sterne argues that

the technology should serve the needed business metrics and not the other way around. Rather than advocating an industry approach he shows many examples of specific situations and the applied technology for web analytics.

3 HISTORY

3.1 Emergence of web analytics sector

Documentation about the early stages of the web analytics sector is scarce. Websites are limited in their information and references, while scientific articles about this topic hardly exist. However one article mentions the allegedly very first tool to measure a website, the Digits WebCounter (D'Alessi, 1997). According to Ballardvale Research (2004) the web analytics market originated a couple of years after the launch of the Mosaic browser in 1993. The Ballardvale research does not include web counters, as mentioned in the article of D'Alessi. It can be argued that web counters were the first tools to measure at least something on your website.

A web counter is a very simple tool. A website owner just has to place one line of code on the website. An example code can be found below:

```
<IMG SRC="http://counter.digits.com/wc/-d/4/counterguide" ALIGN="middle"
WIDTH="75" HEIGHT="20" BORDER="0" HSPACE="4" VSPACE="2">
```

This is regular HTML (Hypertext Markup Language) code and will retrieve an image from the counter.digits.com server. By retrieving this image the website visitor is automatically passing some information to the counter website. Since this code is very basic, the only unit of analysis is a hit. A hit indicates that a visitor has seen (hit) a page. It is not a very sophisticated method, because one visitor can generate many hits by just refreshing the page or writing an application to do this automatically. This will skew the accurateness of the total number of hits.

Around 1995 and 1996 the first web counters appear. The Digits WebCounter is the first massively used counter (counterguide.com, 2007). From a technological perspective it can be seen as the predecessor of the JavaScript analysis, since no software has to be installed on the user's server. It proved to be a very user-friendly tool to get insight in website popularity. Many website owners liked the counter, leading to a rapid increase in sign-ups for this counter. Consequently the demand for bandwidth was overwhelming and incurred high operating costs to offer this free service. At the same time many other counters appeared, offering more functionality or custom counter images. Given the fact that profitability was not easy with this business model market entry and exit was high. Accurate numbers do not exist, but given the large numbers of available counters and the fact that only a couple of them are still in use, market movements have been quite turbulent.

Another widely-used approach appeared around 1997 with the launch of Webalizer (mrunix.net, 2007). Opposite to the web counters, this tool requires installation on the server which runs the website (the website which needs to be analyzed). Many hosting companies started to include Webalizer as a part of their service to enable customers to quickly check the website statistics. Webalizer is an example of a logfile tool. It reads the logfiles which are collected by the web server. After that the software will analyze and visualize the results in tables and charts. Compared to the first web counters this provides much more information and is visually appealing.

3.1.1 First commercial offerings

Two companies, WebSideStory and WebTrends, have been pioneers in this new market and managed to survive the Dotcom crisis in 2001. Both companies started using a different technology. WebTrends analyzed logfiles with their first software version WebTrends 1.0, while WebSideStory's product, Hitbox Professional used JavaScript as its measurement technology.

WebTrends was incorporated on August 31, 1993 (SEC filling WebTrends Corp, 1999). Their initial logfile analysis offering, WebTrends Logfile Analyzer, was introduced in February 1996. Before 1996 the majority of their revenue was generated by the product AuditTrack to manage Novell networks. The chart below clearly shows the increase in licensing revenues after the launch of WebTrends Logfile Analyzer.

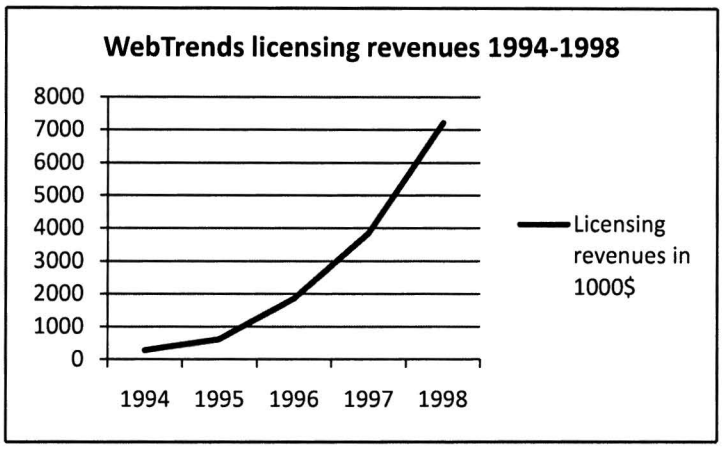

Figure 3: Based on WebTrends SEC filling 1999

This product was mainly used by professionals as opposed to the (free) web counters used by many home users. Nevertheless configuration of the WebTrends product is fairly easy. Users just have to select a logfile (information gathered by the web server) and after processing the logfile several predefined reports are available. Since this methodology requires installation on a web server, this was not suitable to users hosting their website at an external company (not allowing installation of custom programs). Appendix 2 shows some screenshots from WebTrends Log Analyzer 3.0, not the first version, but very similar to it.

Another company with a strong focus on professional users was Netgenesis. Founded in 1994 by MIT graduates (entrepreneurship.mit.edu, 2008) the product range of Netgenesis was closely related to data mining technologies. They launched their first website analysis package in 1995. According to Forrester research (Root, 2006) the name Netgenesis was almost synonymous with web analytics. In 2001 SPSS has acquired Netgenesis and has integrated it into a set of products for "predictive web analytics". Many market analysts have questioned this strategy given the steady loss of market share and the inability to expand the Netgenesis technology to more main-stream customers (Root, 2006). The 2006 Forrester report describes the situation as "SPSS NetGenesis fades out in the web analytics market". It is interesting to note that in the web analytics Yahoogroup (Schiffers, 2005), a forum used by many web analytics practitioners, Netgenesis was considered to be one of the most powerful, but complex, web analytics solutions. Failure to meet market demands, especially regarding out-of-the-box reporting and visualization of data, has turned the former market leader into a company with a market niche solution.

A completely different product was the Hitbox counter from WebSideStory. This counter was offered as a free service to any user looking for some basic statistics. WebSideStory generated revenues from displaying ads on pages which included their free counter and in the online reporting interface. The history of WebSideStory is quite turbulent. Founded in 1996 by Blaise and Agnes Barrelet their revenues were skyrocketing by generating ad revenues. This eccentric couple lured employees by

promises of champagne parties and free BMWs for those exceeding expectations (Weintraub, 2001). Many websites included a Hitbox counter to display website statistics to visitors. The vast majority of ad revenues came from adult-oriented sites, which led to some raised eyebrows, especially with investor relations. In 1999 venture capitalists Summit Partners and TA Associates funded WebSideStory and demanded a new CEO and cutting of all leads to adult-oriented sites (Weintraub, 2001). From that year WebSideStory began using a subscription model aiming at the professional market. After moving away from the online ads to the subscription model, profitability came at risk and since the second half of 1999 until 2004 they haven't been profitable at all (Hellweg, 2004). Apparently losing the ad revenues put a heavy burden on the company's financials. In 2000 WebSideStory was planning to go public at the Nasdaq, but their timing was not perfect. On the eve of the Dotcom crash, investors were not as willing as before to invest in a company which only had experienced losses so far. As a result WebSideStory decided to withdraw the registration on October 2001 (money.cnn.com, 2007).

The information about WebTrends and WebSideStory might lead to the idea that these companies were the frontrunners in this market, but that is hardly true. Especially at the end of the nineties more and more companies emerged or entered the web analytics market with complementary products. However WebTrends and WebSideStory succeeded in building a sustainable company, at least in this period of rapid market developments. The period of 1993 until 2002 can be demarcated as the emergence of the web analytics sector.

The last company to mention regarding the start of the web analytics sector is Omniture, nowadays one of the largest web analytics companies. The Omniture website does not provide too many details about the corporate history, but the Ecommerce guide has an excellent article with detailed information (Cox, 2003). The former name of Omniture was MyComputer, Inc., founded by Josh James en John Pestana. In 1996 they released their flagship product SuperStats as an ASP (Application Service Provider) solution. At first SuperStats was aimed at small businesses, but interest from larger companies quickly increased. In 2002 the small business division was sold and the company changed its name to Omniture. The last years Omniture has taken over a number of companies related to web analytics and even its biggest competitor, Visual Sciences.

3.2 Growth of web analytics sector

Around 2002 web analytics started to become a hot topic among marketing professionals. For a sector which started with free web counters and IT-related analysis tools this proved to be a turnaround. After the Dotcom crisis investors were not very eager anymore to fund "money burning" start-ups without having seen a solid business plan. Part of a solid business plan had to be accountability and web analytics is one of the ways to measure cost effectiveness of a website or online application. An interesting view about the topics covered at industry meetings from 2002 to 2007 comes from Jim Sterne. Jim Sterne has organized summits, called eMetrics, since 2002 and his website lists a history which demarcates many market developments (Sterne, July 2007). More information about market developments is available from my working experience at Adversitement and personal communication with Bob Nieme (Adversitement) and will be described below.

The year 2002 can be seen as a starting point for the rapid growth in the web analytics market. Many industry professionals were aware that web analytics could have great potential, but there were not many best practices yet. Jim Sterne describes the first eMetrics summit as: "Everybody was thrilled to discover that they were not the only ones who understood the language and the possibilities and

the difficulties." (www.emetrics.org, 2007). One of the difficulties was the difference between log file and JavaScript analyses. The eMetrics summit in 2003 had some heated debates about this topic without a clear outcome. Since the Internet became more and more integrated in people's lives, internet technologies were emerging quickly and consequently the need for new and different analysis methodologies. Many vendors responded in increasing the rate at which updates and new versions of their tools were released to stay ahead of the competition in terms of product features. An example is the use of Flash technology on websites, for example to place interactive elements or streaming video on a website. Flash has originally been developed by Macromedia (acquired by Adobe in 2005). Flash technology allows websites to interact with visitors in different ways, leading to new analysis requirements. Around 2003 not many solutions offered native Flash tracking (the analysis of Flash elements on a web site), but today it would be considered as a must-have for any web analytics solutions.

An important moment in the history of web analytics is the year 2004. That year the Web Analytics Association (WAA) has been founded by Bryan Eisenberg, web analytics consultant Andrew Edwards and web strategy consultant Jim Sterne (Sterne, 2007). The association is sponsored by many major web analytics companies and web analytics professionals. The founding members include two executives from WebTrends and WebSideStory. Nowadays all major web analytics companies are a regular or premium member of the association. This gives the WAA a broad platform for web analytics research and industry supported standards for specific metrics. An example is the definition for a visit which is set to expire after 30 minutes of inactivity. The majority of web analytics solutions adhere to this definition of a visit. The WAA also plays an important role by representing the web analytics sector in matters of policy or regulation. In addition the WAA also develops training and certification programs to strengthen the professionalism of the web analytics sector (Sterne, 2007.).

The sponsoring of the WAA by many big players in the web analytics sector (logos are prominently displayed on the homepage) makes it hard to call the association completely independent. After a thorough research of the WAA website and member section not one article showed up that could put one of the founding members in a bad light. As such, the WAA seems to be more of an industry board, defining standards, setting up educational programs and collecting relevant resources. Competition seems not be affected as almost all web analytics companies are already a member or can still join the association. The WAA might have a positive effect on market entry by providing a platform with a lot of specific information. However the most recent articles about some topics (e.g. spyware and web analytics) are from 2005, which is quite outdated given the rapid pace of development.

According to the 2004 Forrester Research report (Chatham, 2004) about user feedback and vendor offerings, more and more vendors appear, which leads to a price decrease. Reported problems among web analytics users are mainly reporting collected data and organizational challenges (Chatham, 2004). Those challenges include turning insights into practice, gaining IT cooperation and determining what data and metrics to collect. The Forrester report identifies two types of web sites in 2004: marketing mega sites (like BestBuy.com) and service oriented sites (like Intel.com) (Chatham, 2004). Marketing mega sites require every feature available in web analytics solutions and will push for more, while service oriented sites might use a limited set of features, for example to do a scenario analysis for an online support portal.

The description of the eMetrics summit in 2005 reflects the interest in organizational requirements for a successful use of web analytics (www.emetrics.org, 2007). One of the main questions was how to make business decisions based on web analytics. In 2005 Adversitement also experienced an increase in market demand, indicated by many requests for proposals and new customers. Some customers already had advanced business intelligence systems, but lacked the technology to include web site measurements in their decision making. Some other customers already had access to web analytics, but were not able to use it effectively in their organization. Management often needed proof that web analytics could be a profitable investment. An example is analyzing the frequently asked questions on the website and matching this data with call center information. When certain questions are frequently asked by phone, call center costs can be reduced by providing a better answer on the web site.

In 2006 web analytics was not considered to be a topic on its own anymore. More and more interest was paid to integration with other data sources. Many companies have different channels to interact with their customers. In addition to the web site this can be a call center, an (offline) store or traditional marketing, like newspaper advertisements. If technically viable the combination of data sources gives a much broader insight in customer behavior than just the analysis of the website. Adversitement has a number of customers inclusing Vodafone, Nuon and Essent which have a very large self service portal. Access to this portal might be limited by requiring login information. The login procedure enables a customer to create many relationships to other (internal) data sources, for example from their call center.

The level of integration and possibility to track visitors on an individual level also poses privacy concerns. Although most analyses are performed on an aggregated level, theoretically it is possible to identify a specific web site visitor. To protect the privacy of individuals several acts have been defined in the US as well as Europe. Detailed information about privacy regulation can be found in chapter 7. One might expect that privacy regulations should have an impact on the various web analytics tools available, but that is hardly true. In fact, most solutions offer the possibility to perform analyses down to the level of individual visitors, provided that information on that level is available. That means the protection of privacy is ultimately in the hands of the individual or company using web analytics. In addition to regulation it becomes a matter of business ethics as well, since the option to track visitors on an individual level is technically available. It is the responsibility of companies to use that information in compliance with regulation and privacy statements.

In 2007, the focus has shifted to an integral approach. Web analytics is still the primary requirement to perform analyses, but there are more options to measure online success (www.emetrics.org, 2007). Adversitement has also seen a growing interest in new and promising technologies, like website search optimization and behavioral targeting. Two examples will explain these technologies. Website search optimization means optimization of the usability of the search functionality of a site and additionally optimizing the search results. Depending on the search context, for example a fashion e-commerce website, the items which generate most revenue can be placed on top of the search result list. Tests have shown that website visitors are reluctant to click on a result when presented with just a long list of options. Providing visitors with five search results, illustrated with a product image, can render much better click-through ratios. To know which elements to optimize web analytics will certainly prove its value, but other data, such as usability studies or fashion trends can be a very valuable resource to the optimization process. Behavioral targeting is another relatively new technology. It enables building anonymous profiles of individual visitors, based on the pages

they have viewed and the links they have clicked. An example can be the website of a car manufacturer. Based on already collected data a visitor interested in sports car may be less likely to visit the section of family cars and vice versa. Consequently visitors which exhibit a certain behavior on the website can be placed in a profile and be presented with targeted content or advertisements, hence the name behavioral targeting. This example is simplified, in reality data is processed with complex algorithms to continuously create, modify and delete visitor profiles. Behavioral targeting is not undisputed, because it has a definitely individual approach not all visitors will appreciate. On the other hand it enables a marketing team to offer more relevant content to a website visitor which can enhance the online experience.

3.3 Overview of history in timeline

The following timeline has been adapted from Omega Digital Media (Clifton, 2007). Appendix 1 contains a visualization of this timeline. The table below contains only breakthrough technological developments as well as important market developments.

Year	Technological developments	Market developments
1993	Release of Mosaic browser with support for image tags	Foundation of WebTrends (no release of software until 1996)
1994	Foundation of Netscape and progress on Netscape browser	Foundation of NetGenesis
1994	First commercial log file analyzer appears, I/PRO	
1995	First use of web counter to track page hits Release of Netgenesis Product Mix tool	
1996	Release of Web-Counter, first widely-used commercial offering Release of WebTrends Logfile Analyzer Release of WebSideStory Hitbox	Start of measuring web site popularity with free and commercial tools Foundation of WebSideStory
1997	Release of Omniture SiteCatalyst	
1999	NetTracker uses open database to enable integration of web analytics data with other systems	Many vendors announce major contracts with (ecommerce) websites (Chatham, 2003)

2001	Analysis of streaming media and Flash by NedStat and WebSideStory Tracking of mobile visitors by WebSideStory	First shakeout in web analytics market with mergers (NetGenesis acquired by SPSS) and new leaders (Forrester, 2003)
2003	Moniforce integrates performance measurement and web analytics (packet sniffing)	
2004	Focus on marketing and campaign analysis	Foundation of Web Analytics Association (WAA) Decreasing price because of intense competition and more vendors (Forrester, 2004)
2005	Integration of Adwords and web analytics by Google	Launch of Google Analytics brings web analytics to a mass public Market development to integrate web analytics in business processes
2006	Integration with other data sources (common practice for 49% of US web analytics users, Forrester 2006)	Web analytics sector is concentrating with a number of major players (Forrester 2006) Visual Sciences acquired by WebSideStory
2007	Measuring of "social media", like weblogs, RSS feeds and user forums. Definition of new metrics, other than page views.	Omniture acquires multiple companies, including Visual Sciences (former WebSideStory) and becomes the largest player in the market
2008	New technologies emerge to measure mobile and video analytics, understandable visualizations and reports become increasingly important	Yahoo acquires IndexTools and rebrands tool to Yahoo Analytics. This tool will be a free service and will mainly compete with Google Analytics

3.4 Current developments

A current development which is unmistakable is market consolidation. A typical example is the company WebSideStory, one of the first movers in the web analytics market. In 2005 it acquired the company Atomz for 39 million dollar. Atomz has been a top player in site search solutions. The Ataomz software enables a company to index their website and optimize the search functionality by using a search engine which runs on the Atomz servers. One year later WebSideStory acquired Visual Sciences for 57 million dollar. Visual Sciences offered multiple analytics solution of which Visual Site has been best known. That year Brian Eisenberg, chairman of the Web Analytics Association had described Visual Sciences as: "Visual Sciences has been acting in stealth mode for the past three years, and they by far have the most phenomenal product on the market." (Newcomb, 2006). Both acquisitions caused quite some stir in the web analytics market. The acquisition of Atomz could be explained by a need to diversify, but the solutions of Visual Sciences stretched far beyond web analytics. A common term used to describe their line of products is "customer analytics". In addition to site analysis tools, Visual Sciences also had software to analyze call-centers, email systems and retail stores. This allows for a much broader and deeper analysis of customer data, collected in different channels. The experience of the author with Visual Site, the flagship solution of Visual Sciences, lead to the preliminary conclusion that it is the most powerful tool for data visualization and segmenting.

Immediately after those acquisitions many market analysts expected a tremendous growth for Visual Sciences (WebSideStory rebranded itself to Visual Sciences). On the other hand competitors like Omniture did not hold back and also acquired a substantial amount of companies, like reporting supplier InStadia, site optimization vendor Touch Clarity and multi-variate testing company Offermatica (www.e-consultancy.com, 2007). Omniture was not as profitable as Visual Sciences, but expanded much more rapidly. In October 2007 Omniture surprised friend and foe with the announcement of the acquisition of Visual Sciences. In what has been the largest transaction in the web analytics market to date, Omniture paid approximately 394 million dollar to acquire Visual Sciences (Omniture, 2007). This acquisition has also caused considerable stir among Visual Sciences customers wondering what will happen to product development, because Omniture and Visual Sciences have some overlapping product ranges. According to Omniture the Visual Sciences products, like HBX Analytics for web analytics will be blended into Omniture's web analytics tool, called SiteCatalyst. The high-end tools of Visual Sciences, like Visual Site, will remain largely unchanged and will be sold under the Omniture brand name. There are still many competitors in this market, but Omniture has built a very strong position with its acquisition strategy.

Another remarkable development is the emergence of Google Analytics. Unlike the offerings of Omniture, Google Analytics is free. It is developed on the basis of Urchin, a former web analytics company bought by Google in 2005. At the end of 2005 Google released the first version of Google Analytics. After the introduction Google had to put a stop on new users due to excessive demand. Today Google Analytics is available without restrictions. Installation is very straight-forward and the reports generated by Google Analytics are relatively easy to understand. The success of Google Analytics can be contributed to Google's strong brand name, but also to the close relation with Google's advertising program called Google Adwords. To enable advertisers to analyze their Adwords performance (clicks on paid keywords in Google) in relation to website behavior, Google Analytics can easily be set up. It lacks some of the functionality of the top-tier tools, but provides enough features for many average users.

Google has been criticized for being a media seller and at the same time tracking the media buy's effectiveness (Newcomb, 2005). For an advertiser there is a trade-off between the position of the Google Adword advertisements and the cost for a specific keyword. If Google can perform an analysis on marketing costs and associated revenues, the cost per advertisement could be increased if profitability is very high. Not a likely scenario, but still a reason for some companies to keep advertising and analytics separated by using different tools.

3.5 Future developments

Future development will focus on combining data from different channels, e.g. Internet, television and call-center, to perform customer-centered analyses. Davenport (2007) argues that every organization has the opportunity to gather large amounts of data, but competitive advantage will only be possible with the right tools and the right analysis to work with this data. Davenport includes web analytics data as a valuable source for analyses and states that strategies should be built around data-driven insights.

Web analytics vendors acknowledge this need for actionable analyses and are expanding their solution portfolio in a horizontal and vertical way. Horizontal expansion means that additional tools, which are directly related to web analytics, are available to generate more data or better predictive modeling. An example is multi-variate testing. This technique enables a company to show different versions of the same page to a visitor. For example, 40% of the visitors will see a green "Buy this" button, another 40% will be presented with a red "Buy this" button and 20% of the visitor will see the default button (to act as a control group). Google is offering this solution as Google Website Optimizer, Omniture has a similar product called Test and Target. Both products extend the value of their respective web analytics offerings.

Vertical expansion can be reached by including not only the Internet in analyses, but also different channels, like television and call-centers (Sen, 2006). Customers of Adversitement, like energy company Nuon, already combine online customer data with call-center data to leverage the call-center volume and enhance the customer service by analyzing which information on the website leads to additional phone calls. Google is experimenting with radio and television advertisements, which can be as easily created as search engine advertisements. The effect of a television commercial on website traffic is immediately visible in the Google Analytics interface. Both examples show that more emphasis is placed on a customer analysis instead of just a website analysis.

Another development is further consolidation in the sector. Rumours that Google or Microsoft will buy one of the established web analytics firms, like Omniture, Webtrends or Coremetrics, might be farfetched. But given the need for constant product innovation and maintaining scalability it could become a reality in the next two years (Batra, 2008). The sector analysis will describe the current situation and potential outcomes in more detail.

4 POLICIES AND REGULATIONS

In the previous chapter the technological and historical backgrounds of the web analytics sector have been described. Policies and regulations affect the web analytics sector and will be studied to assess the impact on drivers for entry and growth. Web analytics practices have been subject of privacy concerns since the first tools have become available. The main reason for these concerns is the ability of web analytics tools to collect (potentially) personally identifiable information. An example which speaks for itself is Google's take on the question if IP-addresses are personal. Alma Whitten, software engineer at Google, wrote a blog posting in which he argued that IP-addresses are non-personal information (Whitten, A., 2008). Whitten argues that IP addresses recorded by every website without additional information should not be regarded as personal information, because it is not possible for these websites to identify the visitor who is behind this address (it is for internet service providers). Whitten received a multitude of replies, most of which did not agree with him. Many website visitors have a static IP address (an address which does not change over multiple browsing sessions) which means that this address can be considered "personal". In contrary to Google's view, the European Union regards IP-addresses as personal information, which means they will fall under the provisions of EU privacy directives.

By looking at random website's privacy policies it is not hard to notice that either not many updates have been applied to the privacy policy or that it is just an obligatory website item not considered to be very relevant. Many websites use advanced website analysis tools (to be checked by looking at the source code) which allow a very detailed analysis of visitor behavior. It is hard to check whether this information is actually used, but even the notice that this data will be collected is lacking from many privacy policies. Even when the privacy policy is legally covering all privacy aspects a visitor still has little control about how this information will be processed and related to other data sources. To protect the privacy of website visitors policies and regulations exist, in the United States as well as in the European Union which attempt to safeguard the online privacy. Additionally it is also a matter of business ethics for companies who are responsible for collecting, analyzing and commercializing online data as indicated by Claburn (2007) in his article about behavioral targeting.

From a technical perspective there are many ways to breach privacy regulations with readily available tools. Despite privacy and security settings in browsers there are multiple ways to compromise the privacy of a website visitor. A relatively unknown method is the use of cache cookies (Felten, 2000). It is using the browser cache as a place to store persistent information about the website behavior of a visitor. When deleting cookies, a visitor can still be recognized as a returning visitor (under the assumption that the cache has not been cleared). A detailed analysis about visit frequency and time spent on site can be performed without the consent of a visitor. Another method which is well-documented is a browser history hack (Grossman, 2006). With this technique a website owner can not only see the previous website (often called referrer), but also other websites a person has visited. The only requirement is to check this against a list of websites. When you make this list large enough, it is possible to get a much broader profile of website visitors. While it is obvious that these methods are not legal, it is often hard to get evidence and take action when privacy is jeopardized. The next part of the chapter will show which privacy regulations exist and how they protect the privacy, paying specific attention to website visitors.

4.1 United States

United States privacy legislation has mostly concentrated on the government, rather than the private sector (Gormley, 1992). It is characterized by an uneven development, led by specific market developments and interest groups. Instead of one coherent privacy law, current legislation is a collection of targeted acts, for example for the protection of personal health or financial information. The basis is the 1974 Privacy Act, which originated from studies by consumer groups and the Department of Health, Education and Welfare. According to Reidenberg (2003) the aggregation of these specific acts leaves many significant gaps and few clear remedies of privacy violations. Additionally the increase in Internet access and personal information stored online, has led to more and more privacy concerns among the public.

Many stakeholders want to have their say regarding privacy regulation, ranging from consumer groups to industry advocates, often with opposing interests. Some claim strong privacy regulation is too expensive or will hinder economic growth, while others argue that the current privacy protection is too limited and not focused on consumers. Internationalization leads some American companies to adhere to much stricter EU privacy directives to service European countries and comply with their privacy protection laws.

Two important United States Acts will be briefly discussed: the Health Insurance Portability & Accountability Act of 1996 (HIPAA) and the Financial Modernization Act of 1999 (Gramm-Leach-Bliley Act). The Privacy Act and the Electronic Communications Act have a strong focus on governmental agencies and do not directly contribute to the protection of privacy-related online data.

The HIPAA is aimed primarily at addressing the use and disclosure of protected health information. Therefore it contains some guidelines about how to safeguard protected information on the internet. It is by no means a guideline which has to be followed by every website; in practice it is mostly relevant to websites which use personal health information and at the same time use web analytics to analyze visitor behavior. Common privacy concerns are that companies like Google get access to sensitive health information and combine it with other data gathered from different (online) sources. This is illustrated by the reactions to Google's announcement of a pilot with the Cleveland Clinic to store patient's health information in their Google accounts (Paczkowsi, 2008). Google is not governed by the HIPAA, even if it stores health information, opening up opportunities to share or use this information with third parties.

The Gramm-Leach-Bliley Act can be compared to the HIPAA, but with a focus on protected financial information. This means that a mortgage website is not allowed to store the annual income of a visitor and tie it to personally identifiable data, for example when a visitor is using an online calculator. It also includes the obligation that companies send an annual overview to consumers, which shows all stored data and how it is used. The last years internet banking has become very popular, which explains the heightened awareness of potential issues and how this act can protect consumers.

Finally the US Internet Caucus Advisory Committee organized a panel discussion in 2003 which concluded (Reidenberg, 2003) that there should be no distinction in online and offline data, privacy legislation should be clear and cover all privacy aspects, privacy statements on websites should be an obligation and that consumers should have the option to remove their data from information systems which use this data to track behavior (e.g. a web analytics tool). Regarding online privacy

statements the World Web Consortium (W3C) has developed the P3P (Platform for Privacy Preference Project). Many web analytics tools use this framework to explain their privacy policy in such a way that visitors can automatically accept or decline that their information is being stored and analyzed by a web analytics tool.

4.2 Europe

The main privacy directive in Europe is directive 95/46/EC (1995). It was developed to harmonize national privacy regulations of the EU member states. The Directive applies to "any operation or set of operations which is performed upon personal data," called "processing" of data. This includes the collection, storage and disclosure of personal data. In contrary to the United States regulation, this directive is more generalized and is very clear about what is and what is not allowed regarding the use of personal data. An advantage of this approach is that the regulation can be applied onto new technologies, without requiring a completely new revision.

In addition to this directive there is a specific directive (Directive 97/66/EC, 1997) for the protection of privacy in telecommunications. This directive states that any unauthorized tapping, listening or other unauthorized interception is illegal. This includes the Internet, but it might be overruled by other directives which allow government agencies to tap information for the protection of national security. In the United Kingdom tests are taking place where internet service providers collect browsing information (anonymously) and serve targeted advertisements on websites. The moment they use any personal information, this policy is considered to be illegal by Directive 97/66.

The introduction of this chapter shows that lawmakers (European Union) and companies (Google) might have a radically different view about what to define as personal information. On many websites a lot of information is collected and stored with a web analytics tool. Technically it is perfectly feasible to store a (anonymous) customer ID when a visitor is logging in, allowing to track a full navigational path on an individual level. Even without storing other personal information, it is possible to link this information to a specific person by combining the web analytics data with the customer database of the company. Although privacy statements should clearly indicate what information is stored and how it is used, the technical limits to breach privacy regulations are almost non-existant.

Recent market developments have shown that the European Union takes privacy very seriously. The acquisition of advertising company DoubleClick by Google has raised concerns among many members of the European parliament about the possible abuse of privacy-related information. The European Commission took a much longer time to study the effects of the acquisition than their US counterpart in this case, the Federal Trade Commission. In the US as well in Europe many consumer groups have opposed this acquisition (Anderson, 2008) with the main argument that there is too much privacy-related information in one company.

Similarities in US and European privacy regulations exist as well. Both have a strong focus on the option to remove your information from specific data sources, for example email lists, and to prevent information systems to store any information. Specifically for the latter, many web analytics tools have (often well-hidden) opt-out pages where a website visitor can indicate not to be included in any analysis by rejecting cookies from a specific domain or erasing stored information. Sometimes privacy-related issues which are not protected by US or European regulation are remedied with

state-level (US) or country-level regulation. For this case the privacy regulations in the Netherlands will be briefly discussed.

4.3 The Netherlands

The Netherlands do not have a general privacy law. The most import law is the law for the protection of personal information (Wet op bescherming persoonsgegevens). An excellent introduction is given by Engelfriet (2008). This law is the result of European directive 95/46/EC. Engelfriet shows a distinction of three different rules related to privacy protection, formulated in this law (abbr. WBP). The first rule is that personal information may only be used after prior consent, which means that a company has to be clear about the fact that personal information is collected, stored or analyzed. Secondly all organizations which use personal information have the obligation to give information about how this information is used whenever this is possible. And finally every person has the right to edit or remove information which is stored when it is not correct, not complete or irrelevant.

Regarding the use of web analytics tools it has the same implications as the European directive, which means that unauthorized use of personal information is not allowed. Quite often this restriction is circumvented by a specific "Terms and Conditions" which visitors have to agree upon, e.g. before they can enjoy the rights of registered users. Those terms and conditions often include the statement that anonymous information will be used to enhance customer service, optimize the website design and offer relevant advertisements. Especially the last point reflects the thin line between personal and non-personal information. Anonymous data may be used, but it can still be linked to the same visitor, which makes it quite personal again.

4.4 Impact on firm entry and growth

Privacy regulation seems to have little effect on firm entry and firm growth. Some new entrants are pushing the limits of privacy regulation by using additional data sources, sometimes provided by cable companies or telecom providers who know exactly who is behind a specific IP address. Given the discussion whether an IP address is personal information or not, the IP address definitely becomes personal with this method. In this case many privacy organizations have responded against these practices, which can turn into major entry barriers if the technology can not be easily adapted. For established firms further restrictions in the use of privacy-related data might pose a serious business risk. This is highlighted in many risk analyses by analysts and in annual reports of public companies (e.g. Omniture annual report 2006). Up until now, no cases exist where an established vendor had to exit the market or lost its market position due to privacy regulation.

5 SECTOR ANALYSIS

5.1 Introduction

As becomes clear from the historical and technical background the web analytics sector has changed rapidly. To perform a sector analysis it is necessary to focus on a specific time frame for which all influential factors are more or less the same. In other words, the current situation is hardly comparable to the early stages of sector development. This research will briefly analyze the very beginning of the sector and then move on to the major lift-off around 2000 and 2001. Finally the current situation and implications for market development will be described. The sector analysis will be a synthesis of the information from the previous chapters, complemented with economic figures and business cases.

Malerba's framework (2004) will be used to analyze the web analytics sector in different stages of development and to identify the building blocks of the innovation systems. These building blocks are identified by Malerba as knowledge and technologies, actors and networks and institutions. A short description of the building blocks will explain their relevance in this sector analysis. The first building block includes technology and knowledge, which are at the base of innovation in a sector. Multiple technologies may be prevalent in a sector (Steinmueller, in Malerba) and the web analytics sector is no exemption (as described in the technology background). These technologies affect the nature and organization of a sector. In addition to technology, knowledge plays an important role in innovation, dependant on accessibility, cumulativeness and opportunity (Malerba, 2004).

The second building block consists of the actors and networks in a sectoral system. Firms are the most important actors and include also users and suppliers. Targeted at the web analytics sector examples of firms are Omniture and Google. Users are companies which use their software to analyze the behavior of website visitors. Examples of suppliers in this sector are server and network suppliers which enable the infrastructure for Omniture and Google to collect, store and analyze web data. Firm heterogeneity also plays an important role as it indicates the degree in which firms are different or more alike and how this will influence innovation in a sector. Finally other types of agents are non-firm organizations, like the Web Analytics Association (WAA), privacy agencies and universities. Actors can be connected in different networks, often split in market and non-market networks (Malerba). Another view at networks is at the level of interaction, indicated by formal cooperation or informal interaction. This research will show that different networks have a great effect on innovation in the web analytics sector.

The third and last building block of Malerba's framework is focused at institutions. Institutions can be norms, routines, rules, laws, established practices which all have an effect on the interaction among agents and their behavior in the sector. Institutions in the web analytics sector include (but are not limited to) privacy regulations, software and interface standards and patent systems. Institutions exist on sectoral, national and global levels and can interact in different directions. In the case of web analytics national privacy regulations supersede privacy guidelines from the Web Analytics Association. On the other hand established practices and standards can lead to changes in the strategy of web analytics software suppliers. An example is the use of page tagging in favor of logfile analyses. Even logfile-only analysis tools had to start offering page tagging as an alternative data collection method (Chatham, 2003).

The building blocks can be identified by performing a structure-conduct-performance (SCP) analysis. Additionally the role of demand in the web analytics sector will be assessed by using specific Adversitement business cases and personal experience with customer's web analytics strategies and demands. Although Adversitement's customers can not be seen as a valid sample of the global web analytics market, it is the most detailed resource in terms of specific demand, allocated budgets and long-term strategies. The global market is described in different industry and market reports, since web analytics has started to move in the spotlight of software analysts. For the early stages of sector development not many independent resources are available, limiting the options for the structure-conduct-performance. Therefore several resources will be used, including weblog postings, interviews with company founders and SEC fillings to analyze revenues of private firms (turning from private to public in a later stage).

5.2 The early stages of the web analytics sector

As described in the history of the web analytics market the start of the web analytics sector can be dated around the mid nineties. At the time there was a very clear division between commercial and free offerings and their respective customer base. The first commercial tools were mainly used to assess web server performance (bandwidth management, traffic spikes, traffic demand per day, etcetera) and not really aimed at the analysis of website visitor behavior. The free tools on the other hand provided just some basic metrics, like number of times a page had been viewed in exchange for displaying an advertisement on the web site. The technology used in the first web analytics tools is not exactly rocket-science. Logfiles are already available when running a web server and a logfile tool is just performing an analysis of these files. Web counters used standard HTML code to display an image (optionally an invisible pixel image) which sent a standard GET request to the data collection server.

Summarizing, the innovative activities in this period can be characterized as a Schumpeter Mark I model. Entry barriers are very low, innovation is driven by entrepreneurs and new firms, knowledge is changing rapidly and uncertainty is high (Malerba, 2004).Many firms which have been active around the start of the web analytics market have been acquired or exited the market. The most important firms in terms of revenue were the suppliers of free web counters, which generated revenue from displaying banner advertisements. Hardly any institutions can be named and are not mentioned in interviews of that time. If any institution should be mentioned, it has to be the business model of offering a free tool in return for displaying advertisements. In that light the tool was just a vehicle to generate revenues from different online activities (advertising). Up until 1996/1997 there is not a distinct web analytics sector to speak of, because of many fragmented actors, a lack of standards and an unclear customer base. Except for WebTrends there are no firms from that stage which play a major role today. Therefore the sector analysis will focus on the periods 1997 to 2003 and 2004 until 2008.

5.3 Development of sector from 1997 to 2003

From 1997 to 2003 the web analytics sector has moved from its infancy to a very promising sector, illustrated by the funding web analytics companies have been able to gather (Chatham, 2003). Some startups from this stage have been able to become current market leaders. Forrester research (Chatham, 2003) shows a list of the major players in the web analytics market, see Appendix 3. This will be a guideline for market development and the firms involved as some of them are still present in the list for 2007. According to this report a number of firms are not included, because they employ

activities that are only partially related to web analytics (e.g. IBM and SAS which offer business intelligence tools). To perform a structure-conduct-performance analysis the three building blocks of the sectoral system will be analyzed, starting with technology and knowledge.

5.3.1 Technology and knowledge in the 1997-2003 web analytics sector

Market reports as well as the emergence of new tools show that technology starts to play a bigger role in the web analytics sector. Data collection has not changed much, but the processing, automated analysis and visualization of data sets new technological standards. Many ecommerce sites first used page views as an indication of the popularity of a specific page, not keeping in mind the relation with other pages, time spent on a page or the bounce ratio (proportion of visitors which leave the page immediately after landing on it). Automated funnel analyses show the click path of a visitor through a fixed scenario, e.g. from landing page, to category page, to product page, to shopping basket, to checkout page and order confirmation page. Using existing knowledge about data processing and analysis, new technologies are used to visualize the data in an attractive way. Clickable images help to quickly navigate through the data and drill-down to detailed reports. There are a number of commonly accepted definitions for basic metrics, like a page view, but no exhaustive set of definitions has been formulated. Since entry barriers are still very low, new entrants appear in the market offering features which are not available in existing tools. An example is WebSideStory with its Hitbox Professional offering, including a detailed ecommerce analysis without the need for very complex integrations. The rising popularity of the Internet is also an incubator for web analytics tools, since most tools report data in an online accessible interface. Knowledge about interface design is available from literature, but also shared online by practitioners and shown by the design of successful websites (IDC, 2005).

A basic patent analysis has been performed to analyze if the increased interest in new features is also reflected in a higher number of patent applications (Appendix 3, patent analysis). The USPTO database is used to perform a search for patents assigned to the Forrester 2003 top 10 companies (Appendix 3). This analysis shows that none of the market leaders in the period 1997 until 2003 applied for a substantial amount of patents. The maximum number of patents which were assigned directly to a single company was three. The analysis also adds some remarks regarding licensing technology from competitors and buying relevant patents from other companies. Other fast-growing markets with a rapid pace of innovation, e.g. semi-conductors, show that patents play a much larger role than they do in the web analytics market. The next chapter will analyze the web analytics market from 2004 until 2007 and indicate if things have changed regarding the importance of owning or licensing patents.

Summarizing, new technologies stimulate the development of web analytics tools, but no remarkable advancement in knowledge building can be reported. Patents do play a role, but not in such a way that it poses inevitable entry barriers to newcomers.

5.3.2 Demand analysis of the 1997-2003 web analytics sector

Opposed to the mid nineties the demand for feature-rich web analytics tools is increasing. Sterne (2002) argues that a company should focus on measuring the success of a website, rather than just collecting data. Analyzing and reporting in a meaningful way was not a common practice among many website owners. New web analytics tools proved to have many useful features which could speed up reporting and provide more valuable insights, directly from the web interface (Chatham, 2003). This period demarcates the start of what can be described as the feature race among web

analytics software suppliers. Innovative features in competing products forced suppliers to include similar functionality in their tools. An example is the measurement of Flash movies on a website. Not many vendors included this functionality; supposedly WebSideStory and Nedstat were the first to offer this feature (Clifton, 2007). Not long after, other vendors included similar technologies or changed their measurement code to enable data collection from Flash movies. In many cases this feature race had more to do with marketing efforts than actual customer demand. Based on several business cases around 2003 most end users effectively used about 20 to 30% of all available functionality. First, because most tools were more or less standardized to include many options, not all of which were relevant to the end user. Secondly not many companies using web analytics tools had in-depth knowledge about the (successful) use of web analytics tool, nor used an external agency to out-source this task.

Although Europe was a bit behind the developments in the United States, the seemingly lack of interest in web analytics was very apparent in the first years of the new millennium, based on the history of Adversitement, since 2001. Many large companies did not use a web analytics tool at all or worked with tools which provided very little functionality. In this period Adversitement had to actively spread the word about the benefits of web analytics to a company and the reasons to invest in a tool. Compared to the situation in 2008, where requests for proposals by prospective customers of Adversitement are flying in at an increasing rate, this period showed that market demand (at least in The Netherlands) was not very high. On the other hand, the first signs of change were apparent when e-commerce oriented websites began to develop an interest in the use of web analytics tools to optimize their online order process (Chatham, 2003). Adversitement's first customers could all be classified as e-commerce websites, with a focus on telecom and travel. A recurring pattern can be seen where a wave of increasing popularity of e-commerce websites is followed by a wave of increasing interest in web analytics tools to collect, analyze and optimize website behavior in order to generate more revenues.

An important question which should be considered is the extent to which awareness of web analytics tools might contribute to a low(er) market demand. Until 2003 many websites used basic tools which only indicated totals of page views and visits. Not many firms were familiar with higher-end tools with more functionality, which is evidenced in the Dutch market by Adversitement. Another argument for this hypothesis is the fact that the retention rate for customers starting with more advanced web analytics tools has been almost 100%. Apparently the majority of new customers acknowledged the value of these tools to get or hold a competitive advantage. The influence of this competitive advantage can also be explained by rapid increases in market demand from specific sectors, e.g. travel. When the word about the value of web analytics started to spread in the online travel market, more and more customers showed a genuine interest in the possibilities of web analytics for their companies (indicated by the number of new Adversitement customers in this specific sector).

5.3.3 Actors in the 1997-2003 web analytics sector

The table of the Forrester report (Appendix 3) lists the most important players in this period, limited to web analytics vendors. To assess the market concentration, a concentration ratio and Herfindahl index have been calculated (Appendix 3, concentration ratio and Herfindahl index). The concentration ratio (C4) shows that the top four players have about 55% of the total market. According to US competition law this would be considered an oligopoly. To get a more precise measurement of market concentration the Herfindahl index has been calculated as well. This shows

that the Herfindahl index is 0.11. According to the US department of Justice this indicates a moderately concentrated market, where a value above 0.18 is considered to be a concentrated market.

More market characteristics should be taken into account to assess whether this moderate concentration could lead to problems. One of these characteristics is the existence of entry barriers. Around 2003 entry barriers to the web analytics market were very low. This is indicated by the Forrester report and the rise of numerous startups in the market. That means that while the market might be a bit concentrated there is more than enough room for new entrants. Investments mainly consist of software development, hardware and web hosting. Hardware and web hosting can be out-sourced and are (up to a certain degree) quite easily scalable. New entrants do not face specific requirements or certifications (such as an ABC Electronics audit), but this might be necessary in a later stage. Given the open character of the Internet, distribution of products is fairly easy, at least for hosted software solutions. As Steinmueller describes (Malerba, 2004) distribution and sales of many software products can benefit from economies of scale, since subsequent copies are only a marginal cost given the electronic distribution.

Economies of scale also influence price elasticity of web analytics software. In general price decreases in this sector lead to an increase in demand, which is very clear in the case of Google Analytics. This tool will be described in detail in the 2004-2007 sector analysis and is a good example of the tremendous effect of a free tool on market demand. On the other hand many of the top 10 companies in 2003 upgraded their free or basic services to enterprise-level tools with a major increase in price. This did not lead to a drop in market demand, indicating that customers had high expectations of the added functionality over time. Another reason could be vendor lock-in (Chatham, 2004), which can prove to be an obstacle for customers to switch between different tools.

Several market characteristics have been taken into account which apply to all vendors. Other market characteristics are vendor-specific, such as the investments in research and development and the number of employees. Both characteristics can indicate drivers for firm entry or growth. Based on the WebSideStory business statement (SEC filing, 2005) the number of employees in 2003 was 135, of which 47 were active in technology development. Not exactly the size of a startup anymore, but reflecting the establishment of larger firms in the sector. It would not be easy for a new entrant to be an immediate competitor to a firm with 47 technology developers and large research and development investments. In theory bright new ideas could be developed into a sustainable business model with a limited number of employees, attracting venture capitalists to expand the business. In that scenario there should be successful companies which entered the market around 2003. As the analysis of the sector from 2004 until 2007 will show there are no entrants which can challenge the current market leaders.

The demand analysis shows that the market for web analytics was still young around 2003. The Forrester 2003 report mentions several trends in market demands boiling down to real-time reporting, data visualization and strong statistical possibilities. Companies with data-intensive business processes or strong online presence exhibit a high demand for tools. An additional factor which influences market demand is vendor lock-in. Many of Adversitement's first customers had relatively small teams, responsible for websites with traffic exceeding that of much larger firms. These customers were not bound to specific software applications which were purchased years ago or developed in-house and thus more flexible in choosing new and better tools. An example of

vendor lock-in is the use and export of historical data. Many larger companies already did some sort of data collection (mainly server-side) and had their reporting based on these numbers. Switching to a web analytics tool (in-house or outsourced) would lead to trend breaks in reporting and analyses, a considerable reason not to switch to another tool at that time. Finally one of the most important obstacles for the adoption of web analytics tools was the budget allocation. Server performance analyses and to some extent logfile analyses were on the budget of the IT department. Outsourced web analytics tools which used client-side JavaScript often had a strong focus on marketing and had to be purchased from the marketing budget. From 2003 to today not much has changed in the proportion of IT to marketing budgets, which means that in most cases the IT budget is much larger than the online marketing budget. Around 2003 many companies attributed ownership of web analytics to the IT department, where the marketing department would have been a more logical choice, given the analysis requirements and optimization requirements of most companies.

Suppliers as well as customers have been named as actors. Both operate in networks with formal and informal relationships. The most important network in this case is the Internet (what's in a name). Suppliers and customers are active on weblogs, in discussion groups and online seminars (webinars) to attract customers (suppliers) and to share experiences or perform market orientation (buyers). Many industry reports are sponsored by the market leaders and employees are often actively engaged in online discussions regarding web analytics tool. It is in the nature of this market that suppliers and buyers are analyzed, discussed, recommended and dismissed. Press releases show that many companies understand the value of a strong online presence and engagement with prospects and customers. In this stage of market development not many formal relationships existed, except for contracts and acquisitions.

5.3.4 Institutions in the 1997-2003 web analytics sector

Up until the 2004 foundation of the Web Analytics Association there have not been similar industry associations in the web analytics sector. Many web analytics vendors organized summits for customers where customer cases were presented and web analytics users could share their experiences. Additionally many vendors had user forums which were accessible from the user interface of the tool. User experiences at Adversitement show that this functionality was hardly used up until 2003. Adversitement has also organized client meetings around 2003, one of them with industry guru Eric T. Peterson. The knowledge level was quite high, but since the (Dutch) market was in its early stages, only the frontrunners which acknowledged the value of web analytics were present.

Several other institutions played an important role in the web analytics network by addressing privacy implications of some web analytics tools. Especially in the early days of the web analytics market many solutions used cross-domain tracking which means that they could track the visitor behavior across every website where there code had been installed. Some other tools used shady techniques to capture the browser history without any user intervention. The lack of an industry organization was a drawback for reliable web analytics companies which more than often had to defend themselves against false accusations. One of the SEC filings from WebSideStory even notes being listed on a spyware list as a potential business risk. Those spyware lists were often compiled with the general idea that a false inclusion would be better than leaving some potential harmful tools out of this list. After 2003 the Web Analytics Associations (WAA) has defined business practices and privacy guidelines to prevent web analytics companies from being included on anti-spyware lists.

5.4 Development of sector from 2004 to 2007

Similarly to the 1997-2003 analysis the Forrester web analytics report (Burns, 2007) will be a guideline for market development and the most important players. For this period much more information is available about technology, market characteristics, including market demand, and networks and institutions. Appendix 4 contains data about the vendors which are analyzed for this period.

5.4.1 Technology and knowledge in the 2004-2007 web analytics sector

The basic underlying technology for many tools and the way data is collected, has not changed much in the period from 2004 to 2007. Logfile analysis, packet sniffing analysis and JavaScript analysis are still prevalent in the market, although JavaScript is increasing its dominant position as the preferred technology for data collection (Rizzotto, 2007). Much more technological advancement has been made in data reporting and data analysis. Reporting includes processing the data and outputting it in a meaningful way. Just as in the previous period reporting heavily relies on existing and new internet technologies to visualize and display information.

Web analytics reporting is no longer limited to an online interface or Excel. An example of a new reporting option is a widget. A widget is a small program which can run as a desktop application. Yahoo (Yahoo Widgets), Adobe (Adobe Air and Flex) and Microsoft (Vista gadgets and Silverlight) are companies which provide widget platforms which can be freely used by widget developers. A widget needs to pull data from the web analytics tool to be able to display interactive information. By supplying a data feed (e.g. a daily file which contains the most relevant web analytics data) or a reporting API (see Definitions) a web analytics tool can be used as a data source for other applications. Both a data feed and a reporting API are no technological breakthroughs, but rather an application of existing (internet) technology (Hao, 2006). Therefore any new entrant would be able to offer similar services in a new product.

The main challenge for a new entrant would be to provide a scalable solution. Developing a sustainable web analytics solution starts with having some key features, but heavily relies on the scalability, especially regarding the needs of large corporate customers. One of the latest trends which can help new entrants to provide a scalable solution, are services like Amazon Web Services (overview by Srinivasan, 2005). Amazon is using part of its enormous network, storage and computing power to offer computing (EC2) and storage (S3) services. The Amazon platform can be used to enable auto-scaling of an application (www.amazon.com/aws, 2008). The application will run on a virtual server which resides in the Amazon environment. If the application hits the limits of the virtual server, it will automatically spawn a new instance of the virtual server to extend computational or storage requirements. The down-side to this solution is the fact that one has to trust the reliability of the Amazon platform to guarantee a specific up-time. Many internet applications use the Amazon platform and downtime of this platform has a serious impact on these applications. An alternative would be to perform data collection with another server cluster and perform additional reporting and analysis on the Amazon platform, thus reducing the risk that customer data will be lost.

As indicated before, the underlying technology of many web analytics tools has not changed radically since the 1997-2003 period. Most noticeable technological differences are in storage and computing by using large distributed networks and server clusters (in-house or outsourced), as shown by the various descriptions of web analytics supplier's uptime and data protection in their respective service

level agreements (SLA). Vendors which appear both in the 1997-2003 and 2004-2007 list all have some technological characteristics in common. They have all adopted new technologies, for example the web technology AJAX to optimize the user interface and enable drag and drop functionality. More importantly they have succeeded in using the available technology to build a web analytics platform which can integrate different data sources, provides a usable interface and can answer the questions their customers have. The five vendors (Google Analytics and Clicktracks are classified as new entrants) have done substantial investments in their technology to become a central hub in web-related data. In other words, the analysis is no longer limited to what people do on the website, but extends to other areas, such as automated keyword bidding, visitor segmentation and behavioral targeting. Three of these areas will be highlighted. Automated keyword bidding means that the budget for advertising systems, such as Google Adwords, can be automatically adjusted based on different rules, for example the net profit a keyword generates. Therefore successful keywords will receive more budget than keyword which do not lead to a success, where success can be defined as a new order, a submitted lead form or a question answered by the online customer service. Secondly, visitor segmentation is a technology to filter the data based on specific visitor characteristics, such as age group, gender, time spent on site or a combination of characteristics. This functionality is extremely powerful in determining the exact behavior of a small set of visitors. An example is the analysis of all visitors that have made a purchase online. From their behavior an analyst can determine which factors contribute to a purchase and apply optimizations to the website or the online marketing strategy. Finally, another technology related to web analytics is behavioral targeting. This technology uses an automated system to deliver the right content to a specific visitor by finding patterns in web analytics data (Claburn, 2007). An online bank can choose to offer a specific banner to visitors which indicate by their browsing behavior that they are interested in an investment plan. This technology needs web analytics data to function correctly. The combination of these technologies enable a vendor to offer a complete package from a single user interface, which might be an advantage over using individual tools for every task.

Many technologies have been described and additionally the influence of patents on knowledge building will be assessed. The 2004-2007 patent analysis (Appendix 4) shows that none of the vendors is very active with patent applications. From all the vendors in Appendix 4, only Omniture has a new patent. Google has multiple patents, but none of them specifically mentions web analytics. The patent analysis shows some additional queries which lead to the conclusion that patents do not play a very important role in the web analytics sector. One of the reasons could be that most of the technology which is used is publicly available and technological innovations by a specific company do not have to be disclosed to be used effectively. A vendor might have a radically new technology in analyzing the web analytics data, which could leads to savings in hardware and reduce computing time. By using this technology only in its own data center, this innovation will never be publically available, reducing the need for patents. This indicates that new entrants will probably not experience high entry barriers, such as expensive patent licenses. On the other hand the patent analysis shows that big investments in patents apparently are not necessary to achieve sustainable growth and develop new functionality. An important remark to this analysis is the fact that a company might have multiple patents, which have been issued to a different company which later has been acquired. Additionally many companies have patents pending, which are not issued yet. The Omniture annual report of 2006 shows that the company has 19 issued patents and 85 patents pending (50 US and 35 international). According to the annual report the patents are related to "data modeling and classification; online messaging optimization; Web site traffic analytics and predictive

modeling; online behavior prediction and analysis; real-time monitoring and aggregation of Web activity; Web content management and optimization; and compilation of data records relating to Web site visitation sessions". This indicates that some of the patents of Omniture are indeed a result of their acquisitions (such as Touch Clarity for predictive modeling and online behavior prediction and analysis) In addition to patents intellectual property is also secured by copyrights, trademarks, service marks, trade secret laws and contractual restrictions.

Knowing how to use all available technology is also an important factor, for vendors as well as for end-users of the software. Knowledge is available in product documentation, in standards related to internet (IEEE) and on public or closed user forums. The importance of knowledge is reflected in the fact that end users of web analytics software not only rely on the vendors, but also use services of specialized agencies (such as Adversitement) for implementation, user training and advanced analyses. Adversitement has done several complex web analytics implementations, where implementation teams of different vendors were not able to use the full potential of their own products to meet customer requirements. This indicates that in-depth knowledge about the technology and its opportunities can give an agency, but also a vendor a competitive advantage.

Tacit knowledge includes best practices, experience with web analytics projects in specific industries and personal skills to work with a web analytics tool. Codified knowledge includes implementation documents, written procedures for specific analyses and books or blog postings about web analytics. A lack of knowledge can seriously hamper firm growth or rise potential entry barriers. Given the rapid developments in this market a vendor, established vendors as well as new entrants, should be up to date with new technologies, such as video or mobile analytics. Proof for this is the fact that many logfile based applications have exited the web analytics market, by realizing too late that a switch to JavaScript data collection is necessary to collect data from interactive applications, such as Flash. Webtrends started with a pure logfile solution, but offered a mixed model just in time, supporting also JavaScript analysis. Part of this move is explainable by market demand, but knowledge about technological advancements in this sector is crucial to develop or hold the market position.

5.4.2 Demand analysis of the 2004-2007 web analytics sector

Afther the analysis of technology and knowledge this part of the sector analysis has its focus on the role of customer demand. Where the 1997 to 2003 period was demarcated by the "feature race" in which web analytics vendors challenged the market with tons of additional features with every new release, the period from 2004 to 2007 can be characterized as the "demand for answers". This characterization is derived from personal experience at Adversitement and working with various customers, as well as industry reports such as another Forrester web analytics report (Menzies, 2007) and the Aberdeen paper (Lovett, 2007). Both reports indicate that despite a wealth of features in available web analytics tool, customers find it particularly hard to use the tools in the right way. The Aberdeen study (Appendix 5) shows that many customers struggle to get the full potential from their web analytics tool and that many customers find it hard to interpret the data. In other words, the web analytics tool is not giving them answers to their questions. The Forrester report (Menzies, 2007) shows that many companies complain that the web analytics tool is giving them interesting, but not valuable reports. It indicates that almost all relevant data is available in web analytics reporting, but that it is not easy to get value out of these reports by using it to optimize the website. Finally many customers of Adversitement have indicated that they need consultancy services to speed up the integration of web analytics in their organization and turn their investments into profit.

From these examples it is clear that customers have many questions, but to embed the demand analysis in a well-organized model a distinction will be made based on firm size (based on revenue) and firm industry. The source for firm size will be a Jupiter report which described the web analytics constellation in 2007 (Daniels, 2007). This research shows that in 2007 only 12% of the companies (which took part in Jupiter's Executive Survey) are seeking to upgrade or replace their current web analytics tool. Another 22% indicate that they will employ a new solution in the next 12 months. That means there are good opportunities for web analytics vendors. A further analysis of the data shows the influence of company size on market opportunities and demand. The largest companies (revenue over 3 billion USD) have deployed web analytics applications more than in any other segment with a share of 66% (Daniels, 2007). This market segment also has the highest concentration of web site decision makers which are planning to replace or upgrade their web analytics tool (16%). Another good opportunity might be the small and medium-sized business segment (revenue between 50 and 500$ USD), which has the greatest concentration of web site decision makers planning to start with a web analytics solution.

The Jupiter research is based on US companies, but the situation in Europe is quite comparable as is shown by the Forrester report of Daniels (2007). Opposed to the previous period Europe is no longer far behind the United States in adoption and use of web analytics tools. Given the fact that Europe had to catch up there might be even more opportunities for web analytics vendors for new deployments of their tools. Based on the customers of Adversitement another classification of demand can be made by identifying different industries. Note that most of these customers are based in The Netherlands, but given the wide range of customer types this list is quite a good indication of the European market. Sectors that are using web analytics more than average are especially travel and telecom industries. Since the start of Adversitement these sectors have proven to be early adopters of web analytics tools and to foster the importance of data in decision making. One of the reasons are the high marketing expenses in this sector and therefore the need to identify successful and failed marketing campaigns. Another reason is that margins can be very small, which drives the need for continuous optimization. Firm sizes vary from small to corporate companies, but in general smaller companies show more rapid developments than corporate, among others to the lack of very time-consuming business procedures. The table and charts in Appendix 6 show the combination of firm size and firm industry and their influence on the adoption of web analytics tools and procedures (web analytics involvement, see explanation in Appendix 6). The charts show that the telecom and travel sector are most active regarding web analytics, reasons for this have been described. When looking at company size, the total web analytics involvement can be misleading as for some combinations of sector and company size data is not available. Estimations are used to achieve a full data set, which still results in a very similar chart (Appendix 6, Figure 8). Small and medium sized companies are almost equal, while corporate companies prove to have the highest web analytics involvement. An important remark is that this data is skewed towards larger companies, since the tools Adversitement is using for her customers are not free (like Google Analytics). Many smaller companies will prefer to use Google Analytics which requires no investments in licenses and can provide a good starting point for web analytics. Practice and industry reports have shown that corporate companies have started to invest more and more in web analytics over the last years. This provides new opportunities for existing vendors and new entrants. On the other hand it can also be an entry barrier for new entrants, since corporate companies tend to choose recognized brand names with a proven track record and a strict service level agreement. That

might be potential show stoppers for new entrants aiming at medium and corporate sized companies.

In addition to licensing a specific tool the market is asking for further support in the form of (online) training, seminars and business consultancy (Lovett, 2007). All five vendors which maintained their position from the start of the web analytics sector until 2007 have a strong focus on delivering a full suite of tools, product training and professional services. Despite their efforts market demand shows that there is a considerable gap between demand and supply. As stated by the Aberdeen survey (Lovett, 2007) the vendor training is often limited to using their own tools without a broader context of defining business objectives and asking the right business questions. The gap is filled by agencies like Adversitement (and many others) which provide dedicated web analytics services and can use multiple tools to achieve the objectives of a customer. Given the fact that more and more sophisticated tools are offered as a free service, proper training, education and tailored product implementations can be crucial in holding or increasing the market position. Many customers ask for a tailor-made solution, which often includes much more than just the installation of software. Think about custom-made applications which use the web analytics data, localized training and integrations with company-specific applications. This might be a good opportunity for new entrants to pick the fruit which is left hanging by the established players.

Finally the demand analysis includes one of the main reasons market demand is generated, namely marketing by web analytics vendors. The period from 2004 to 2007 has shown that marketing is extremely important to obtain sustainable growth or enter the market successfully. For example, Omniture has spent more than 40% of its revenues on sales and marketing. The other vendors show comparable figures. One of the reasons for this budget allocation is the fact that many web analytics solutions are available. Potential customers will often start their search for a new or replacement tool on the Internet. The volume of related questions in (among others) the web analytics Yahoo Group mailing list is illustrative for this customer orientation process. Other popular sources for information are seminars and industry reviews by research companies. It takes a lot of time and money to have a solid presence in all these areas. As described in the history of the web analytics sectors, some tools with superior features have not made it, partly because they did not have enough market presence. WebTrends is currently battling the same problem by dropping a few places in some review rankings and losing marketing momentum (Menzies, 2007). That translates to less media coverage (online and offline) and a higher risk that prospective customers will not assess their solution. The marketing strategy can be developed up to a certain degree, but the last years have shown that the power of indirect marketing is increasing. Indirect marketing is described as product advocates who write favorably about a specific tool on a weblog or in a magazine without having a direct relationship with the vendor (Westervelt, 2007). The same goes for user forums where experiences with various tools can be exchanged and buyers can obtain more objective information. Some companies actively engage in such discussions (e.g. on the web analytics Yahoogroup) to provide more detailed information or exert damage control in response to criticism. New entrants often manage to get a lot of attention and free publicity by promoting unique features of their tool, for example having an alliance with cable companies to track visitor behavior more accurately. Established companies trust on press releases, user seminars and tradeshows to strengthen their brand name and increase its awareness.

5.4.3 Actors and networks in the 2004-2007 web analytics sector

Similar to the previous period an analysis of market concentration has been performed to assess the concentration ratio and the Herfindahl index. For the period from 2004 to 2007 the concentration ratio is 53% (55% in 1997-2003) and the Herfindahl index is 0.09 (0.11 in 1997-2003). Both the concentration ratio and the Herfindahl index have decreased a little bit. Remarkable is the fact that the four largest companies have about the same concentration ratio when the two periods are compared, but different vendors appear in the list. In the period from 2004 to 2007 Google has reached a second position with a strong market share. The Herfindahl index indicates that there have not been very strong fluctuations in the market concentration. While the United States uses fixed market concentration cut-offs the European Union watches the differences. In both cases market concentration can not be considered to be a serious entry barrier. Google has shown that a new tool, Google Analytics, can establish a strong market position, although it should be kept in mind that the sector is still quite turbulent. Of course Google Analytics has built on an existing tool (Urchin) and Google has a very strong brand name, but in a very concentrated market even Google would not be able to get to a second position.

As described for the previous period economies of scale still play a role in software distribution (Malerba, 2004). With the rise of SaaS applications this is even more prevalent than in the prevous period of analysis. Price elasticity is best illustrated by the Google Analytics case. After Google renamed their acquired software tool Urchin to Google Analytics and decided to offer it for free, they have welcomed an incredible amount of new customers. It has been a couple of years since the introduction of Google Analytics, so it is possible to assess the influence of this free tool on paid tools to some extent. Many analysts believed that the rise of Google Analytics could be the end of paid tools. The vendor list shows that Google Analytics is the only free tool among the big players. Google Analytics definitely has an effect on the market, but rather as a catalyst for web analytics adoption, than as a replacement tool for paid solutions. Adversitement currently works with many customers who have hit the limits of Google Analytics and are looking for a more robust solution. These companies might have chosen a more basic tool if they did not start with Google Analytics.

Recent developments include the acquisition of Indextools by Yahoo and the announcement that there will also be a free Yahoo Analytics (with more powerful features than Google Analytics). Given the history of Google Analytics this introduction of another major free tool will probably have the same effect. Yahoo Analytics will possibly attract some companies using paid tools, but also leading customers to more sophisticated web analytics tools or platforms with multiple applications. Indicative for this development is the fact that Omniture has announced a new contract with Yahoo itself, even though they have their own analytics tool. Finally there are more and more open-source web analytics tools, such as Piwik. They might be the perfect start for small and even some medium-sized companies due to their flexibility, support for open standards and active community engagement. However, just like the other vendors it needs hardware to perform data collection and data analysis. All major vendors indicate in their risk analysis that hardware and network conditions can be a bottleneck when more capacity is needed. Making an open source tool scalable requires a lot of effort on the hardware and networking side, which is one of the reasons more and more customers prefer a SaaS solution.

Networks in the web analytics sector are getting more and more connected, due to alliances and partnerships. Web analytics vendor Coremetrics works with e-commerce platform supplier ATG to offer a bundled solution. Omniture has its Genesis network by which many third parties can connect

their tool to Omniture its web analytics tool. Additionally many customer relationship management systems (CRM) have out of the box options to integrate with different web analytics tools to combine customer data with their web behavior.

Finally, the structure of the companies present in the vendor list for 2004 to 2007 has been analyzed. Similar to the strong growth in revenues and number of customers, the number of employees and R&D investments has also increased rapidly. The Omniture annual report for 2006 shows that they have 353 full-time employees, almost triple the number WebSideStory had around 2003. The research and development expenses of Omniture have continued to grow with the revenue, being between 10% and 15% of total revenue over the years 2004, 2005 and 2006. No exact information is available for other vendors, but considering the fact they are continuously releasing new products which compete with the offerings of Omniture, it is probably close to these investments. Related to the investments, another group of important actors are investment banks and groups. Many web analytics vendors need substantial amounts of cash to setup and maintain their servers and networks. These expenses are not immediately reflected in new revenue, because it often takes a while before a new customers has finished implementation. Some vendors are public companies and some are private, but in both forms they are continuously trying to get new funding to secure new developments and build competitive advantage.

5.4.4 Institutions in the 2004-2007 web analytics sector

Unlike the period before 2004 the current web analytics sector has many institutions. First, the vendors organize many seminars and industry meetings. The most important ones are named by Lovett (2007) and include the Coremetrics University, Google's Conversion University, Omniture University, Visual Science's former Digital Marketing University, now called Active!nsights and WebTrends' Professional Training Courses. The term "university" is used quite often, but is not an indication for cooperation with academic institutions. This cooperation is taking place, but on a different level. For example, Omniture is offering the Omniture Academic Program including course material, training for tutors and live access to the web analytics tool. Several other vendors transfer part of their knowledge to academic institutions by master classes and seminars, while simultaneously marketing their products and employ their recruitment strategy.

The cooperation with universities can help in achieving growth, as technology transfer can work both ways. Adversitement (although not a vendor) has worked with several institutions, including the Eindhoven University of Technology (TU/e) where the university benefits from a wealth of real life web analytics data from very large companies. At the same time a company like Adversitement can benefit from the knowledge available at the TU/e regarding data mining, modeling and visualization. New entrants could also benefit from a similar cooperation, since a lot of university knowledge can be applied in a web analytics tool (based on the origin of current web analytics technologies), under the assumption that this knowledge can be translated in a usable interface.

An important institution and a milestone for the web analytics sector is the birth of the Web Analytics Association (WAA) in 2004. The WAA is the first and up until now the only industry organization. It provides many useful resources to her members (corporate and individual), organizes web analytics courses and defines guidelines for privacy and information security. Membership of the WAA can be obtained by paying the membership fees and should not be regarded as an approval or quality label for a specific vendor. Certifications and audits are done by other institutions, such as ABC Electronic. An ABC Electronic audit guarantees that a vendor can meet the service level

agreement, that data is collected and processed in a structured way and that the data can be considered reliable (e.g. not missing 5% of the data received by the tool).

Finally, organizations which protect the privacy of web visitors are very important institutions in the web analytics sector. Web analytics vendors and privacy organizations often seem to have conflicting interests. Web analytics tools try to gather as much information from website visitors as possible (and allowed by privacy laws), while privacy organizations argue that every website visitors should be able to opt-out of web analytics tracking. The integration of web data with additional data sources, like CRM systems, can be a reason for additional privacy implications. To achieve sustainable growth web analytics vendor should cooperate with privacy organizations to maintain their business model while respecting the privacy of website visitors.

6 CONCLUSION, DISCUSSION AND RECOMMENDATIONS

This chapter contains the main conclusions of this research. Additionally the research methodology will be discussed and the chapter ends with recommendations for further research.

6.1 Conclusion

The main objective of this research is to identify drivers for firm entry and firm growth in the web analytics sector. To assess which drivers exist for firm entry the history of the web analytics sector has been described as well as the initial stage of sector development (1997-2003). In the very first stages of sector development (before 1997) new entrants are characterized by their business model of offering a free service in return for the opportunity to display advertisements. That business model has proven not to be sustainable, forcing these companies out of the market or pushing them to another business model. Firm entry after 1997 is driven by market demand and by having specific features, but is still relatively independent of the technology used for data collection. WebTrends is using logfiles and WebSideStory has JavaScript for data collection. New entrants which manage to get a strong market position show an impressive list of features and continue to develop more of them. Many opportunities exist in the 1997-2003 market and limited investments are enough to enter the market without any problems. A software as a service (SaaS) model decreases entry barriers even more by reducing distribution costs.

Budget allocation for web analytics tools is shifting from IT to marketing departments. One of the reasons for this change is the increase in marketing expenses and the need to analyze the effectiveness of marketing campaigns. New entrants with a strong focus on online marketing analyses will be more likely to sustain or develop their market position. Institutions do not play a major role in firm entry, which is indicated by the fact that no industry organization exists until 2004.

The development of the web analytics sector after 2004 shows a revival of the free pricing model. Google Analytics is one of the examples of a tool which is available free of charge. Google quickly accumulated a large market share, but still is the only free tool in the list of top vendors (based on revenue). Technology still plays an important role, since new entrants often show unique features not available in existing tools. On the other hand firm entry is seriously hindered by huge capital requirements to build a robust network and offer a scalable application. Many new entrants appear, but none of them have been able to compete with the current top vendors or only service the lower end of the market (e.g. Clicktracks).

Drivers for firm growth have been analyzed by the sector analysis of the period from 2004 to 2007, paying special attention to vendors who have been able to sustain their position since the start of the web analytics sector. From a technology perspective multiple drivers for growth can be identified. First, the data collection technology becomes more and more important, since more data (in general) means better analysis opportunities. Vendors which originated from logfile technologies have at least partially switched to JavaScript to provide analyses of customer interaction. Vendors who failed to do so have left the market or lost their position. Secondly, technological drivers for growth include an appealing user interface which provides out-of-the-box analyses. The growth path of all successful startups shows that they have invested much effort in their user interface and analytical capabilities. Secondly, in addition to the web analytics tool itself, an important driver for growth is the existence of web analytics platforms which add related products, such as predictive modeling tools, dynamic search applications and integrations with third party systems. These platforms are an essential part of the market strategy. The top vendors invest large shares of their revenues in sales and marketing.

Patents still do not play a major role, but are used with other protections such as copy rights and trademarks to protect intellectual property. Firm growth seems not be limited by having none or just a limited amount of patents. More important for firm growth is the offering of additional services, such as training, implementation support and consultancy services. The rise of many web analytics agencies proves that the market needs those services and the top vendors all actively engage in providing (tool-specific) training and business consultancy. To further stimulate growth many vendors have strategic alliances with other companies which provide CRM or advertising systems, enabling customers to easily integrate data and perform more detailed analyses. In addition to business relationships with other companies, drivers for growth can also be found in cooperation with academic institutions. Technology transfer enables vendors to gain competitive advantage by incorporating scientific knowledge into their products or services.

Finally, an important driver for growth is funding. Funding can be received from investment groups or by going public and is essential in building a solid business. Maintaining and upgrading a robust network demands a vast amount of capital. All top vendors (excluding new entrants Google Analytics and Clicktracks) aim at the high end of the market with large corporate customers, which require a scalable and solid platform. New technologies like cloud computing and online storage (for example, Amazon Web Services) can help to foster firm entry and growth by providing a platform which is easily scalable and reduces the need for huge capital investments.

6.2 Discussion

To indicate the impact of the hypotheses and the available data on the conclusions, some remarks will be placed regarding this research. To assess drivers for firm entry and firm growth, more company data would have been desirable. Specific information regarding balance sheets, R&D investments and total revenues is often not available (especially for private companies). This research has built upon data from several industry reports, but would benefit from the inclusion of more companies, specifically new entrants. From the start of the web analytics sector many new entrants have appeared. Some have left the market, but many of them are still present. However, they are not large enough to be included in the top vendor lists or serve a specific market segment.

Another point of discussion is the influence of different geographical markets on this research. United States based vendors have been analyzed in conjunction with European market demand. The US and European market have many similarities, but also exhibit their own characteristics. An example is the Google search engine, which is much more popular in Europe than it is in the United States where it competes with AOL and Yahoo. Another example is the corporate culture in the United States with a strong focus on targets and decision making on the basis of data. These differences might have an influence on specific drivers for firm entry and growth in both markets.

Finally, the role of web analytics agencies has become very important over the last years. They have speed up the adoption of web analytics tools and highlighted the importance and relevance of web analytics. It would be out of the boundaries of this research to also include agencies, but their behavior certainly influences drivers for firm entry and growth.

6.3 Recommendations

After the discussion some recommendations for further research will be described. First, the patent analysis should be performed with more attention to specific technologies. Since the patent analysis in this research did not yield much data, it would be a wise choice not to limit the patent analysis to

US patents only, but also include European and national patents. Additionally a patent citations analysis might be interesting to assess links between specific technologies or a specific set of companies.

Secondly, the role of web analytics agencies should be included in a follow-up research. These agencies have a very good view of the market demand and can be a valuable source of the budget allocation for web analytics tools among their customers. A stronger focus on web analytics customers and their allocated budgets would also be preferable, since revenues of web analytics vendors do not tell the whole story. Customers often use agencies to perform part of the implementation, have one or more employees working full-time with web analytics tools and spend money to integrate other data sources with web analytics data. A thorough analysis of large customers would provide valuable insights in the economic characteristics of the web analytics sector.

Finally, a very interesting topic for further research is the rise of free and open source web analytics tools. In the current market both act as catalysts for high-end tools, but in the future they might replace paid tools to some extent. New technologies might be able to provide a scalable platform without the need for building large networks. Other emerging markets, such as the software industry in general, have shown similar developments and might provide valuable data to build this research on. Considering the rapid developments in the web analytics sector over the last ten years, there will be a rich source of data to explore in future research.

7 REFERENCES

7.1 Literature

Abraham, M., Meierhoefer, C., Lipsman, A., 2007. *The impact of cookie deletion on the accuracy of site-server and ad-server metrics: an empirical Comscore study.* Comscore Press.

Burby, J., Brown, A. & WAA Standards Committee, 2006. *Web Analytics "Big Three" Definitions*. Web Analytics Association, Washington.

Burns, M., 2007. *The Forrester Wave™: Web Analytics, Q3 2007.* Forrester Research.

Chatham, B., 2003. *Why Website Analytics Matter.* Forrester Research.

Chatham, B., 2004. *What Matters To Web Site Analytics Users.* Forrester Research.

Claburn, T., 2007. Call of the wolves. In *InformationWeek,* Nov. 12, 2007, 1162, pp. 69.

Clifton, B., 2007. *Web Traffic Data Sources & Vendor Comparison.* Omega Digital Media.

D'Alessio, D, 1997. Use of the worldwide web in the 1996 US election. *Electoral studies*, Vol. 16, No. 4, pp 489-500. Elsevier Studies, 1997.

Daniels, D., 2007. *Web Analytics Constellation 2007. Assessing web analytics applications.* Jupiter Research, a division of JupiterKachan.

Davenport, T., 2007. *Competing on Analytics: The New Science of Winning.* Harvard Business School Press, 1st edition.

Directive 95/46/EC, 1995. European Union. *Concerning the protection of individuals with regard to the processing of personal data and on the free movement of such data.*

Directive 97/66/EC, 1997. European Union. *Concerning the processing of personal data and the protection of privacy in the telecommunications sector.*

Felten, E. and Schneider, M., 2000. Timing attacks on Web privacy. In *Proceedings of ACM Conference on Computer and Communications* Security '2000. pp.25-32

Gormley, K., 1992. One Hundred Years of Privacy. *Wisconsin Law Review, pp . 1335-1440.* University of Wisconsin.

Hao, W., Gao, T., Yen, I., Chen, Y., Paul, R., 20006. An Infrastructure for Web Services Migration for Real-Time Applications. *Second IEEE International Symposium on Service-Oriented System Engineering (SOSE'06)*, pp. 41-48.

IDC, 2005. *An integrated approach to digital marketing through web analytics*. IDC Go-to-market services, doc. No 32917.

Lovett, J., 2007. *Sector Insight – Web Analytics University*. Aberdeen Group, a Harte-Hanks company.

Malerba, F., 2004. *Sectoral Systems of Innovation: Concepts, issues and analyses of six major sectors in Europe.* Cambridge: Cambridge University Press.

Menzies, C., 2007. *Web analytics needs people.* Forrester Research.

Mowery, D.C. & Nelson, R.R. (Eds.), 1999. *Sources of Industrial Leadership: studies of seven industries.* Cambridge: Cambridge University Press.

Reidenberg, Joel R., "Privacy Wrongs in Search of Remedies". *Hastings Law Journal,* Vol. 54, 2003.

Rizzotto, F., 2007. *Adding intelligence to web analytics.* IDC White paper, IDC Italy.

Root, N.L., 2006. *The Forrester Wave™: Web Analytics, Q1 2006.* Forrester Research.

Sen, A., Dacin, P., and Pattichis, T., 2006. Current trends in web analytics. *Communications of the ACM.* November 2006, Vol. 49, No. 11.

Srinivasan, L., Treadwell, J., 2005. An overview of service-oriented architecture, web services and grid computing. *HP Software Global Business Unit Publications*, v02 11/2005.

Sterne, J., 2002. *Web Metrics: Proven methods for measuring website success.* Wiley, 1st edition.

Weintraub, A, 2001. WebSideStory: A happy ending? *Business Week*, July 2001. McGraw Hill Publishing.

7.2 SEC filings

Netratings, Form: 10-Q, Filing Date: 5/10/2006
Online copy at http://sec.edgar-online.com/2006/05/10/0001104659-06-033332/Section11.asp .

WebSideStory Inc., Form: 10-K, Filing Date: 3/29/2005
Online copy at http://sec.edgar-online.com/2005/03/29/0000936392-05-000073/Section2.asp .

WEBTRENDS CORP, Form:S-1/A, Filing Date:1/29/1999
Online copy at http://sec.edgar-online.com/1999/01/29/15/0000891020-99-000116/Section10.asp

7.3 Websites

Amazon Web Services, 2008. Elastic cloud computing (EC2) and storage (S3) services. Retrieved May 25, 2008 from
http://www.amazon.com/aws

Anderson, Nate, 2008. Consumer advocates try to derail Google/DoubleClick merger. Retrieved April 24, 2008 from
http://arstechnica.com/news.ars/post/20070420-consumer-advocates-try-to-derail-googledoubleclick-merger.html

Ballardvale Research, 2004. Market Trends - Web Analytics: History and Future. Retrieved July 25, 2007 from Ballardvale Research
http://www.ballardvale.com/free/WAHistory.htm

Barret, B., 2001. The Webalizer: Credits. Retrieved November 20, 2007 from
http://www.mrunix.net/webalizer/credits.html

Batra, A., 2008. Move over Google Analytics, here comes Yahoo Analytics. Retrieved from the Webanalysis weblog,
http://webanalysis.blogspot.com/2008/04/move-over-google-analytics-here-comes.html

Clickstream. (2007, July 10). *In Wikipedia, The free encyclopedia.* Retrieved July 10, 2007, from http://en.wikipedia.org/wiki/Clickstream

www.clickz.com, Retrieved February 11, 2008, from
http://www.clickz.com/showPage.html?page=3564521
Google acquisition of Urchin

Counterguide, 2004. A history of web counters and web analytics. Retrieved December 11, 2007 from Counterguide
http://www.counterguide.com/article/7_history_of_web_counters_and_web_analytics.html.

Cox, B., 2003. Case Study: Half.com. Retrieved from the Ecommerce Guide on February 8, 2008.
http://www.ecommerce-guide.com/news/trends/article.php/10417_1565681

E-consultancy, December 2007. E-Business briefing December 2007: Omniture's Neil Weston on the threat of Google. Retrieved on January 8, 2008 from
http://www.e-consultancy.com/news-blog/newsletter/3571/omniture-s-neil-weston-on-the-threat-of-google.html#36751

Eisenberg, B., September 27, 2002. What should you measure. Retrieved from the Clickz Network on July 23, 2007.
http://www.clickz.com/showPage.html?page=1471031.

Engelfriet, A., 2008. Persoonsgegevens op Internet. Retrieved from Ius Mentis on April 10, 2008 from
http://www.iusmentis.com/maatschappij/privacy/persoonsgegevens/

Grossman, J., 2006. I know where you've been. Retrieved on March 11, 2008 from
http://jeremiahgrossman.blogspot.com/2006/08/i-know-where-youve-been.html

Hamel, S., 2008. Web analytics vendors market shares. Retrieved July 5, 2008 from Immeria
http://blog.immeria.net/2008/01/web-analytics-vendors-market-shares.html

Hellweg, E., June 9, 2004. WebSideStory's Second Chapter. Retrieved December 11, 2007 from CNN Money
http://money.cnn.com/2004/06/09/technology/techinvestor/hellweg/

MIT, 2007. MIT spin-offs. Retrieved November 20, 2007 from MIT Entrepreneurship Center
http://entrepreneurship.mit.edu/mit_spinoffs.php

Newcomb, K., November 16, 2005. Google Analytics: rising tide or tsunami? Retrieved from the Clickz Network on February 11, 2008 from
http://www.clickz.com/showPage.html?page=3564521

Newcomb, K., February 2, 2006. WebSideStory buys Visual Sciences. Retrieved from the Clickz Network on December 10, 2007.
http://www.clickz.com/showPage.html?page=3582476

Omniture, October 25, 2007. Omniture to acquire Visual Sciences. Press release retrieved from http://www.omniture.com/press/417

Paczkowsi, J., 2008, February 21. New from Google: "Google Privacy Disaster Waiting to Happen". Retrieved February 28, 2008 from http://digitaldaily.allthingsd.com/20080221/google-health/

Page, R., 2007, December 12. Nostradamus and Web Analytics - 2008 Predictions! Retrieved February 11, 2008, from http://rich-page.com/web-analytics/nostradamus-and-web-analytics-2008-predictions/

Pols, A., 2006, March 31. Web analytics: What about packet sniffing? Retrieved July 24, 2007 from http://webanalytics.wordpress.com/2006/05/31/web-analytics-what-about-packet-sniffing/

Schiffers, O., December 1, 2005. [webanalytics] Netgenesis. Message posted on the Web Analytics Yahoogroup, retrieved October 10, 2007 http://tech.groups.yahoo.com/group/webanalytics/message/4355

Sterne, J., 2007. The WAA story. Retrieved December 10, 2007, from the Web Analytics Association http://www.webanalyticsassociation.org/en/cms/?306

Sterne, J., July 2007. The history of the eMetrics marketing optimization summit. Retrieved November 20, 2007 from http://www.emetrics.org/history.php

Sramanamitra, 2008, July 16. Deal radar 2008 – Coremetrics. Retrieved July 2, 2008 from http://sramanamitra.com/2008/07/16/deal-radar-2008-coremetrics/

Telecompaper, 2007, December 12. Vodafone verstevigt nummer 2 positie in Nederland. Retrieved February 9, 2008, from Telecompaper http://www.telecompaper.com/nl/article.aspx?id=16031

Web analytics. (2007, July 22). *In Wikipedia, The free encyclopedia.* Retrieved July 23, 2007, from http://en.wikipedia.org/wiki/Web_analytics

Westervelt, R., 2007. What is indirect marketing? Retrieved December 11, 2007, from http://ubrander.wordpress.com/2007/02/01/what-is-indirect-marketing/

Whitingjo, January 15, 2007. How would you explain web analytics?. Message posted on the Web Analytics Yahoogroup, retrieved October 10, 2007 http://tech.groups.yahoo.com/group/webanalytics/message/9024

Whitten, A., 2008. Are IP addresses personal? Retrieved March 15, 2008, from the Google Public Policy Blog http://googlepublicpolicy.blogspot.com/2008/02/are-ip-addresses-personal.html

World Wide Web Consortium (W3C). Retrieved March 20, 2008 from http://www.w3c.org

DEFINITIONS

API
See Reporting API

Bounce rate
The bounce rate for a specific page is the number of visitors which only see a single page and exit the website after visiting this page.

Conversion
A conversion can be defined as reaching a goal on your website. This goal can be a purchase, but also acquiring contact information or answering a question.

All definitions above are inextricably linked to the basics of every analysis. Some metrics are expressed in numbers of page views, some in numbers of visits. Without an understanding of these core definitions many other metrics are meaningless.

Clickstream analysis
Analysis of the path a visitor has taken on the website. Requirement is that all steps have been recorded, with one or more of the available measurement technologies.

Cookie
A small text-file residing in the browser cache of a visitor. Cookies are used to store visitor preferences, login information and can be used to recognize returning visitors. Cookies can be deleted by a visitor, regardless of the browser which is used. Regular cookie deletion can lead to skewed data regarding returning and unique visitors. Serving a cookie from the same domain the visitor is browsing is called a first party cookie. If the cookie is served by another domain (often the case with outsourced web analytics solutions), the cookie is called a third party cookie. Cookie acceptance for third party cookies is much lower than for first party cookies, primarily caused by default web browser security settings.

Counter
A counter (or web counter) is used to indicate the number of visits on a particular website. The counter presentation can be textual or graphical, the latter being dominant. Counters have been very popular in the nineties, but their information is limited and counters are often related to advertising business models.

First party cookie
See Cookie

JavaScript
JavaScript is a web programming language used to add interaction to a web page or web application. The JavaScript code is executed client-side and is used in web analytics software to retrieve visitor specific information, such as browser type, display resolution. The code can be complemented with server-side information, such as order value and items.

Hit
A hit has been the first unit of measurement for websites. It indicates that a specific page has been "hit", in other words retrieved by a visitor. In early analytics programs a hit was a hit, regardless of

the visitor being a human or search robot. The term hit has largely been replaced by the page view as default unit of measurement.

Logfile
A text file on a web server which stores all website related information, such as page requests, image requests and form postings. Log files have a specific format which can be changed to gather more or very specific information. Content served from a different server (e.g. a caching server) will usually not be reflected in the log files.

Packet sniffing
Packet sniffing is a method to gather data about website usage by "sniffing" all traffic flowing between a visitor and a web server. The technique can be used for many applications, but related to web analytics the goals is to analyze which information is actually served to a customer.

Page view
A page view has been the default unit of measurement for many years. It indicates that a single page has been viewed. Since a single page can be viewed multiple times by the same visitor, the visit metric is often used to indicate in how many visits a page appeared, regardless of the number of page views. When a visitor sees the search page of Google, one page view is counted. After pressing "Search" and viewing the results page, a second page view will be counted.

Reporting API
An API is an Application Program Interface. A reporting API is an interface which can be used by external programs to import or export information from a data source (e.g. a web analytics tool). There are different API technologies, but web services seem to be prevalent in the web analytics sector. A web service provides a framework to query a database and retrieve a response, often in XML format from an online application.

SaaS
Software as a Service. SaaS is a term used to describe a type of contract where the supplier is offering a service and software being a part of that service. More and more software companies are using the SaaS model because of the scalability and convenience to customers.

Segmentation
Segmentation can be used to segment website visitors based on specific characteristics. Segmentation is primarily done during analysis of the data, but solutions exist to dynamically allocate visitors to a certain segment, for example based on their product preferences.

Server
A server is a computer which can serve data. In relation to web analytics servers are used to serve websites and to collect data. Multiple applications can run on a single server, but often servers have a dedicated task, such as a web server.

Service Level Agreement (SLA)
The SLA of a web analytics vendor shows the guaranteed uptime for data collection and data reporting, the way data is stored and secured and often includes a paragraph about data ownership. It provides interesting details about the way data collection and processing is organized.

Third party cookie
See Cookie

Uniques
Uniques is a term used to describe unique visitors. Based on the software used, unique visitors can be aggregated on a daily, weekly, monthly or quarterly basis. Often a cookie is used to recognize a unique visitor. Therefore a quarter is a generally accepted maximum to accurately identify a unique visitor, because cookie deletion will cause the same visitor to be counted as a new customer. By analyzing different authentication systems, such as login information, data for unique visitors can be enhanced with great margins.

Visit
A visit is counted when a visitor visits a web page. If the visitor leaves the website and does not return to the website in 30 minutes or if the website visitor is inactive for 30 minutes, the visit will automatically expire. This is a standard definition set by the Web Analytics Associations. However many web analytics tools will allow users to manually set the visit timeout, for example to one hour.

Visitor (unique visitor)
A visitor is counted by using some form of authentication (for example a cookie). The same visitor can visit one website multiple times a day. In that case multiple visits will be counted, but only one visitor will be counted. The accurateness of the data greatly depends on the authentication used. If only a cookie is used, the same visitor will not be recognized when browsing the website from a different computer or after deleting the browser cache.

APPENDIX 1 – WEB ANALYTICS HISTORY

Web analytics technology and timeline

The picture below shows a simplified version of the differences between logfile, packet sniffing and JavaScript technology to collect data about the use of a website.

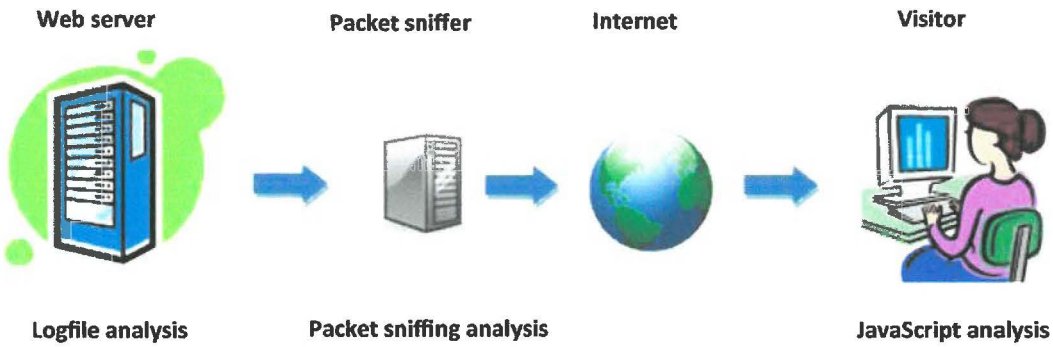

Figure 4: Different methodologies for analysis (company documentation Adversitement)

Web analytics timeline

The timeline below is an indication of the development of the web analytics sector.

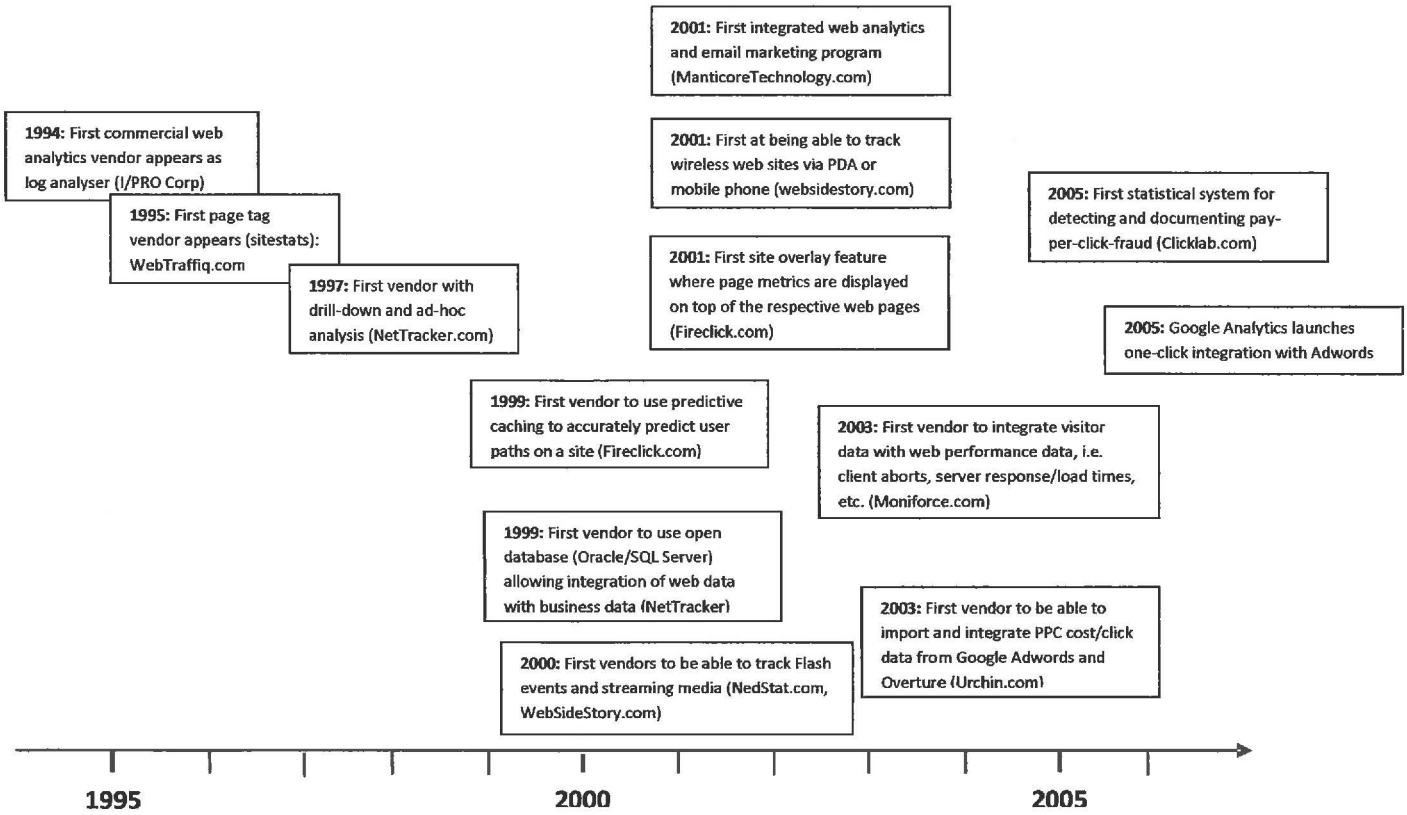

Figure 5: Web analytics technology timeline, adapted from Omega Digital Media (Clifton, 2007)

APPENDIX 2 – COST ESTIMATES

Details of cost estimates

The costs estimates below are based on a small-sized website with 10 million events (e.g. page views) a year. Estimates are based on personal communication with various customers and experience with multiple web analytics tools.

Logfile analysis is the easiest technique to set up, under the condition that web server logfiles are already available and provide enough details. Setup includes loading logfiles, naming of pages and optionally building a navigation hierarchy.

Logfile analysis (costs per year)	Cost (in euro)
Hardware (database server)	2000
Server hosting	1000
Software licence	2000
Setup	2000
Total	7000

Packet sniffing needs more hardware because it analyzes web data at a network packet level. Therefore it requires a so-called sniffing server, often in addition to a database server for the analysis. In many cases sniffing takes place directly on the web server, but that might have a performance impact. In most cases setup involves installing a plug-in for the web server which can capture this information and send it to a different server for analysis.

Packet sniffing analysis (costs per year)	Cost (in euro)
Hardware (sniffing and database server)	4000
Server hosting	1000
Software licence	2000
Setup	2000
Total	9000

JavaScript analysis requires installing JavaScript code on every page which needs to be analyzed. Setup includes placing a generic code on every page and adding optional code to track custom events (like cart adds or purchases). When this code is in place, almost all of the further configuration can be done in a web interface.

Javascript analysis (costs per year)	Cost (in euro)
Software licence	2000
Setup	3000
Total	5000

APPENDIX 3 – VENDOR AND PATENT ANALYSIS 1997-2003

Top vendors in 2003

The table below shows the top web analytics vendor according to 2003 Forrester Research (Chatham, 2003).

Forrester Rank	Vendor	Web analytics revenue (million USD$)
1	NetIQ/WebTrends	$50
2	SPSS/NetGenesis	$38
3	WebSideStory	$15
4	Omniture	$10
5	Coremetrics	$8
6	digiMine	$8
7	Sane Solutions	$8
8	Accrue Software	$5
9	Buystream	$3
10	Speed-trap	$2
Total revenue		**$147**

Source: Forrester Research, 2003

Concentration ratio and Herfindahl index (2003)

Total market revenues (million dollars): 205

Note: this number is higher than the total revenues of the top 10 companies. It also includes smaller companies of which revenues do not exceed two million (Forrester 2003). Revenue has been used as a measure of output, since this is the most consistent data across the companies in the list.

	2003 revenues (million dollars)	Proportion of total
NetIQ/Webtrends	50	24%
SPSS/NetGenesis	38	19%
WebSideStory	15	7%
Omniture	10	5%
Coremetrics	8	4%
digiMine	8	4%
Sane Solutions	8	4%
Accrue Software	5	2%
Buystream	3	1%
Speed-trap	2	1%

Concentration ratio (C4): 55%

Herfindahl index: 0,11

Patent analysis for 2003 companies

The USPTO database has been used, since most companies in the list are US-based. The query has been performed by filtering on assignee name. Another option could be using a specific class, but web analytics is a broad term and related patents cover many classes. Important classes (US Classification) are 705 (Data processing: financial, business practice, management, or cost/price determination), 707 (Data processing: database and file management or data structures) and to a lesser extent 709 and 710 (both related to digital processing systems).

Queries used at USPTO database:

(AN/(((((((((omniture OR websidestory) OR coremetrics) OR webtrends) OR netgenesis) OR digimine) OR (sane AND solutions)) OR accrue) OR buystream) OR speed-trap) AND APD/19920101->19993112)

This query only renders three results, all from WebTrends. It shows that from 1992 to the end of 1999 only WebTrends applied for patents. The following query alters the date range to show results from 2000 to the end of 2003.

(AN/(((((((((omniture OR websidestory) OR coremetrics) OR webtrends) OR netgenesis) OR digimine) OR (sane AND solutions)) OR accrue) OR buystream) OR speed-trap) AND APD/20000101->20033112)

The result shows that in this period of three years the top vendors only applied for nine patents. However this data might be skewed, because many of these companies own patents which are originally assigned to another company. Additionally many companies license patented technology. For example Omniture is currently licensing 7 patents from AC Nielsen/Netratings (5,675,510, 5,796,952, 6,115,680, 6,108,637, 6,138,155, 6,643,696, and 6,763,386).

The description of some patents leaves much room for interpretation. Evidence for this case is given by a 2005 lawsuit from Netratings which accused WebSideStory, Visual Sciences, Omniture, Coremetrics, SageMetrics and Sane Solutions of patent infringements. All cases were settled in 2005 and 2006. Netratings had licensed certain of its patents in exchange for license or royalty fees (SEC filing Netratings, 2006).

APPENDIX 4 – VENDOR AND PATENT ANALYSIS 2004-2007

Top vendors in 2007

The table below shows the top web analytics vendors according to 2007 Forrester Research (Burns, 2007). This report does not include revenues per vendor, but only lists companies which have revenues of over 25 million US dollars. Revenues have been sourced from annual reports (public companies) and investment bank estimates (for non-public companies).

Forrester Rank	Vendor	Web analytics revenue (US$)
1	Omniture	$143
2	Google (Google Analytics)	$133
3	Unica	$102
4	Visual Sciences	$80
5	Webtrends	$80
6	Clicktracks (as part of Lyris)	$43
7	Coremetrics	$35
Total revenue		**$616**

Sources: Forrester Research, 2007 (Burns), Hamel, S., 2008, Sramanamitra, 2008

The Google Analytics revenue has been calculated by using data from the weblog Immeria.net (Hamel, 2008). The author of this weblog has developed a web analytics identification tool (WASP) which can be installed as a Firefox plugin. This plugin is widely used and collects data about the usage of specific web analytics tools and the number of pages on which a specific tool is present. From this data an estimated market share (measured by the number of pages on which a specific web analytics tool is present) can be derived. The ratio of the Google Analytics market share and the Omniture market share is then multiplied by the Omniture revenue to get a rough estimate of Google Analytic's market capitalization.

The Coremetrics revenue has been determined by using an investment bank analysis of this company in which the 2007 revenue (estimate) is given. Other revenues are sourced from annual reports. The 2007 vendor list contains five companies already present in 2003, which are Omniture, Unica (Unica acquired Sane Solutions), Visual Sciences (Visual Sciences acquired by WebSideStory, using Visual Sciences as new company name), Webtrends and Coremetrics. Newcomers are Clicktracks and Google Analytics.

Concentration ratio and Herfindahl index (2007)

Total revenue for all web analytics companies in 2007 is not available in analyst reports. Therefore the ratio between the revenue of the top vendors in 2003 and the total revenue in 2003 is used to calculate an estimate of total revenue in 2007. The combined revenues from the top vendors (616 million USD) is multiplied with a factor 1.4 which yields a total revenue of the 2007 web analytics market of 859 million USD.

	2007 revenues (million dollars)	Proportion of total
Omniture	143	17%
Google (Google Analytics)	130	15%
Unica	102	12%
Visual Sciences	80	9%
Webtrends	80	9%
Clicktracks	43	5%
Coremetrics	35	4%

Concentration ratio (C4): 53%
Herfindahl index: 0,09

Patent analysis for 2004-2007 companies

Similar to the 2003 patent analysis the USPTO database has been used and patents have been searched by filtering on assignee name. In the first query below Google has been left out on purpose, because its product range is much broader than that of the other firms.

AN/((omniture) or (unica) or (webtrends) or (visual and sciences) or (clicktracks) or (coremetrics)) and (APD/20040101->20073112)

This query only renders one result, a patent from Omniture. Surprisingly no new patents have been granted to the other companies listed in the query. One of the reasons could be that some core "web analytics" patents (e.g. the Netratings patents) are still used and there is no need to protect their intellectual property by patents. Technological innovations which relate to the way a web analytics vendor processes and analyzes data will normally not be disclosed by the vendor and this information is not publicly available, thus limiting the need for patents. Another reasons is that some patents might be still pending. In its 2006 annual report Omniture indicates that 50 US and international patents pending. Some of their patents might be registered under a different name, for example one of their acquired companies.

AN/(google) and (APD/20040101->20073112)

This query renders multiple results, but none of them is specifically related to web analytics. Google might use some technologies described in their patents for web analytics, but none of them mentions the term "web analytics" specifically. To assess whether other patents specifically mention web analytics the following query has been executed.

SPEC/(web and analytics) and (APD/20040101->20073112)

This query renders 28 results and after a more detailed analysis of all the individual patents, only 11 of them are relevant to the web analytics market. Vendors which specifically mention which patents they use (e.g. Omniture) do not mention any of these 11 patents.

For further research a more detailed analysis should break down all web analytics related technologies in separate categories and analyze the links between those patents. A patent citation analysis could be relevant to assess the influence of specific patents and specific technology categories.

APPENDIX 5 – DEMAND ANALYSIS

Aberdeen Group study which shows that many customers are not using web analytics tools in an optimal way.

	Coremetrics	Google Analytics	Omniture	Unica	Visual Sciences	WebTrends
% of customers who claim they are not getting the full potential from their current web analytics technology	50%	43%	50%	18%	33%	39%
% of customer that reported data is difficult to extract from analytics platform	25%	22%	17%	36%	13%	15%
% of customers that dedicate full time staff to managing web analytics	38%	21%	39%	36%	53%	12%
% of customers that state analytics data is difficult to interpret	25%	24%	17%	18%	13%	30%

Table 1: Source: Aberdeen Group (Lovett), 2007

Note: Metrics displayed in this table reflect a sampling of customers that responded to Aberdeen's survey on web analytics. This data is intended to provide directional insight into the pressures and actions of a small sampling of customers using specific vendor solutions. Customer may or may not be using the latest versions of vendor solutions.

APPENDIX 6 – WEB ANALYTICS INVOLVEMENT PER SECTOR AND COMPANY SIZE

Demand analysis based on company size and company sector of Adversitement customers. Company size "small" is defined as less than €1 000 000 revenue, "medium" is defined as between €1 000 000 and €20 000 000 revenue and "corporate" is defined as over €20 000 000 revenue

Web analytics involvement is rated on a scale from 1 to 10. A ranking of 10 means that all decision making of a company (related to the web) is based on web analytics. A ranking of 1 means that a company has a tool, but is not using it actively. A ranking of 0 means that not enough data is available. All data is derived from personal experience and from customer profiles which are maintained by Adversitement.

	Small	Medium	Corporate
Telecom	8	9	7
Travel	7	9	9
Banking			5
Insurance	8	6	4
Retail		6	9
Media	4	4	7
Real Estate		7	
Energy			6

Table 2: Web analytics involvement based on Adversitement customers

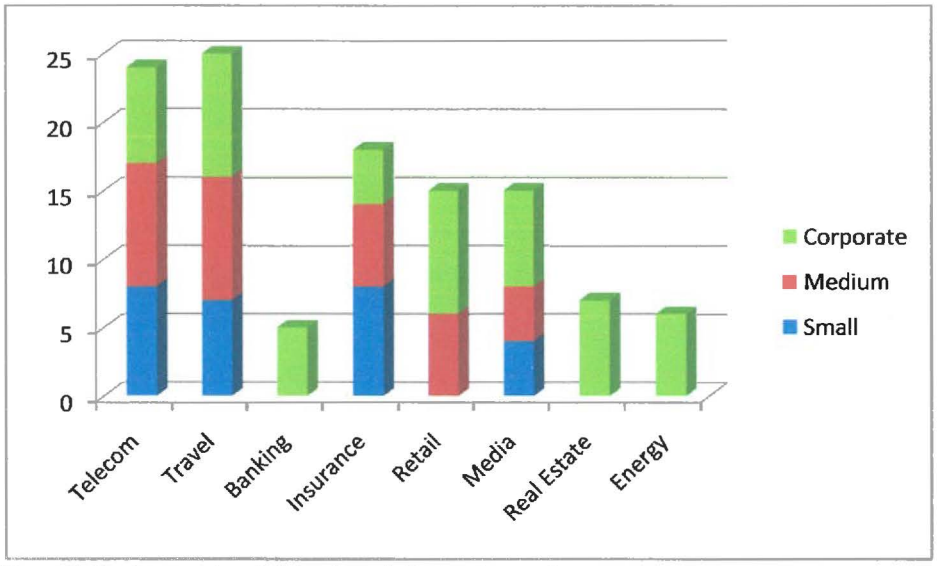

Figure 6: Chart per sector - total web analytics involvement

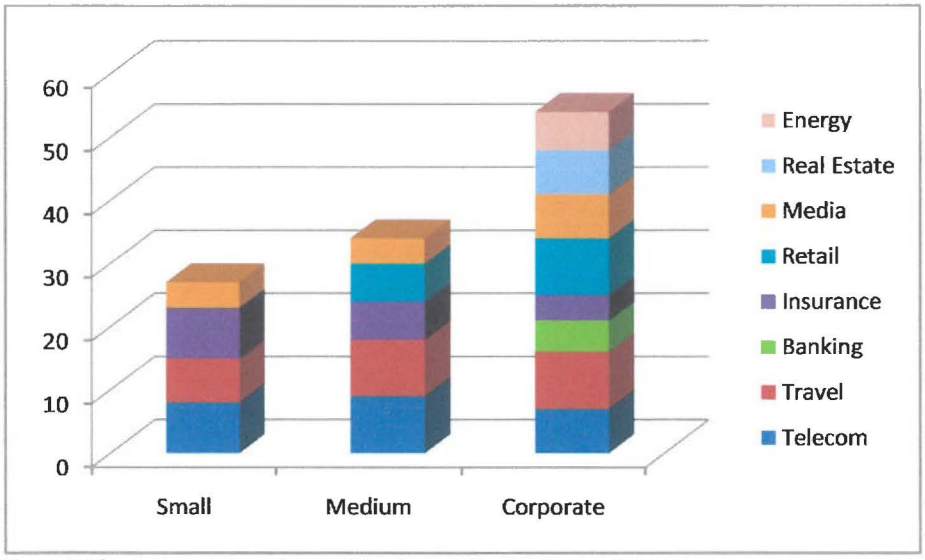

Figure 7: Chart per company size - total web analytics involvement

The chart below also includes estimates for missing data based on market knowledge and internal assessments of prospective customers by Adversitement. It looks quite similar to the chart which is displayed in figure 6.

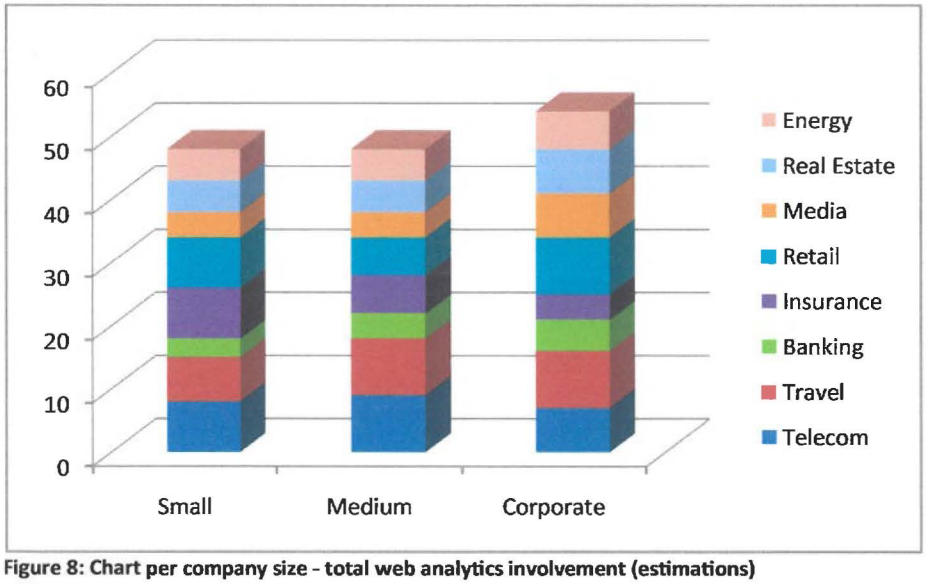

Figure 8: Chart per company size - total web analytics involvement (estimations)

APPENDIX 7 - COMPANY DATA

Data is derived from the Hoover's online company database. Full access to this database requires a subscription. All data is retrieved in December 2007.

Coremetrics

1840 Gateway Dr., Ste. 320
San Mateo, CA 94404
Phone: 650-762-1400
Fax: 650-762-1499
Toll Free: 877-721-2673

http://www.coremetrics.com/

Overview

Want to know who's clicking on your Web site, and where, and how? Coremetrics provides Web-based analytics and precision marketing software that companies use to analyze the behavior of their online customers and Web site visitors. The company's software captures and stores site visitor activity data, then converts the data into individual visitor profiles that give online marketers, information technology managers, and e-commerce company executives insight into how their Web sites are being used. Customers have included Bank of America, Columbia House, Motorola, and Williams-Sonoma.

Coremetrics has received venture capital funding from Accel Partners and Highland Capital Partners.

Financials

Company Type	Private
Sales (mil.)	$16.9 (est.)
Employees	215

Key People

President and CEO	Joe Davis
COO	Shawn Farshchi
CFO	Mark Resnick
VP, Worldwide Sales	Jeffrey (Jeff) Schmidt
Chief Marketing Officer	Kathi Kaplan

Top competitors

NetIQ
Netratings

Google

Overview

1600 Amphitheatre Pkwy.
Mountain View, CA 94043 Phone: 650-253-0000
Fax: 650-253-0001

http://www.google.com

If you've never Googled, you probably aren't finding what you want online. Google operates the leading Internet search engine, offering targeted search results from billions of Web pages. Results are based on a proprietary algorithm -- Google's technology for ranking Web pages is called PageRank. The company generates nearly all of its revenue through ad sales. Advertisers can deliver relevant ads targeted to search queries or Web content. Google also operates the Google Network, a network of third party customers that use Google's advertising programs to deliver relevant ads to their own Web sites. Founders Sergey Brin and Larry Page each have nearly 30% voting control of the company.

Financials

Company Type	Public (NASDAQ (GS): GOOG)
Fiscal Year-End	December
2007 Sales (mil.)	$16,594.0
1-Year Sales Growth	56.5%
2007 Net Income (mil.)	$4,203.7
1-Year Net Income Growth	36.6%
2007 Employees	16,805
1-Year Employee Growth	57.4%

Key people

Chairman and CEO	Eric E. Schmidt
President, Technology and Director	Sergey Brin
SVP and CFO	George Reyes
SVP Corporate Development, Chief Legal Officer,	David C. Drummond
SVP Global Sales and Business Development	Omid Kordestani

Competitors

IAC Search and Media
MSN
Yahoo

NetRatings

Overview

120 W. 45th St., 35th Fl.
New York, NY 10036
Phone: 212-703-5900
Fax: 212-703-5901
Toll Free: 888-634-1222

http://www.nielsen-netratings.com

NetRatings, which operates under the Nielsen/NetRatings brand as a unit of market research giant Nielsen, analyzes Internet traffic for clients in advertising, e-commerce, Internet publishing, and traditional media. Its measurement research covers not only Web browsing but also instant messaging, streaming media, and other online applications. The company measures online advertising effectiveness and provides its clients with demographic research about their online audiences. In addition, Nielsen/NetRatings offers Web site metrics and custom research. NetRatings is now part of the business unit Nielsen Online, together with BuzzMetrics.

Financials

Company Type	Subsidiary of The Nielsen Company
Fiscal Year-End	December
2006 Sales (mil.)	$81.8
1-Year Sales Growth	20.2%
2006 Net Income (mil.)	$2.8
2006 Employees	397
1-Year Employee Growth	0.3%

Key people

Chairman	John A. Dimling
EVP Corporate Development, CFO, and Secretary	Todd Sloan
EVP Global Operations and US Sales	Manish Bhatia
SVP Engineering and Operations and CTO	John A. Kleine
Director of Marketing and Communications, Asia	Elvira Lodewick

Competitors

Coremetrics
Omniture
comScore

Omniture

Overview

550 E. Timpanogos Cir.
Orem, UT 84097
Phone: 801-722-7000
Fax: 801-722-7001

http://www.omniture.com

Omniture aims its aperture at your Web site. The company provides Internet analytic software and services to corporate customers such as AOL, eBay, General Motors, and Microsoft. Omniture's primary product, SiteCatalyst, helps clients measure Web site traffic, visitor activity, advertising effectiveness, and e-commerce transactions. Other products include the Omniture Discover, Data, and SearchCenter line of products, designed to provide customers access to all of their data in real time. Director Mark Gorenberg owns about 23% of Omniture.

Financials

Company Type	Public (NASDAQ (GM): OMTR)
Fiscal Year-End	December
2007 Sales (mil.)	$143.1
1-Year Sales Growth	79.5%
2007 Net Income (mil.)	($9.4)
2006 Employees	353
Company Type	Public (NASDAQ (GM): OMTR)
Fiscal Year-End	December
2007 Sales (mil.)	$143.1
1-Year Sales Growth	79.5%
2007 Net Income (mil.)	($9.4)
2006 Employees	353

Key people

President, CEO, and Director	Joshua G. (Josh) James
EVP and CFO	Michael S. (Mike) Herring
EVP Products and CTO	Brett M. Error
EVP Business Development and Corporate	John F. Mellor
SVP Worldwide Marketing	Gail M. Ennis

Top competitors

Coremetrics
NetRatings
WebTrends

WebTrends

Overview

851 SW 6th Ave., Ste. 600
Portland, OR 97204
Phone: 503-294-7025
Fax: 503-294-7130
Toll Free: 877-932-8736

http://www.webtrends.com

WebTrends answers the who, what, when, and where of corporate Web sites. Customers use the company's software to analyze, manage, and report on who is using their site, what pages visitors view most, and when problems are about to occur. Its software also monitors security, determines how users access a company's Web site, and analyzes e-commerce functions. WebTrends' customers come from a wide range of industries and include T-Mobile, Royal Appliance, and Volvo . The company was originally acquired by NetIQ in 2001, which later sold the division to Francisco Partners in 2005.

Financials

Company Type	Private
Fiscal Year-End	December
Sales in 2007 (mil.)	$5.4 (est.)
Employees in 2007	45

Key people

VP and CFO	Bryan J. LeBlanc
VP and General Manager, EMEA	Nick Sharp
CEO	Daniel E. (Dan) Stickel
VP Engineering and Hosted Operations	Bruce Kenny
Chief Marketing Officer	Kathleen E. (Kathy) Brush

Competitors

Coremetrics
NetRatings
Omniture