

# Leveraging Web Analytics To Understand The Customer Behavior

NVS Chaitanya<sup>1</sup>, MCA,CBIT, Hyderabad  
G.kalpana<sup>2</sup>, Asst.Professor, Dept of MCA,CBIT

In order to meet the increasing customer shifting preferences and to study the customer behavior, web analytics has come to light with the new source of sustainable advantage. Web analytics allows companies to uncover customer insights and leverage them for strategic decisions. Yet, despite its recognized value, many companies are not prepared to accept the change through web analytics. Therefore, a trans-disciplinary research that crosses the boundaries of the field of change management and web analytics is necessary to provide a deeper understanding of the key elements of change. First, to highlight the importance of analytics, the field business intelligence and analytics will be discussed. Second, a brief introduction to web analytics will be presented and the main challenges of web analytics will be explored. Recognizing the challenges of web analytics will facilitate our understanding of the problems behind the transition towards a data-driven culture. Further, Davenport and Harris' (2007) framework of being an analytical competitor will be combined with Peterson's (2009) discussion on the same topic to provide an account of the characteristics that companies need to possess in order to successfully implement a web analytics strategy.

**Keywords:** Analytics, Classification, data mining, metrics, web analytic

## I. Introduction

Analyzing the data available to your business to increase profitability is a practice as old as trade itself (Peterson, 2009a). The rise of the internet has changed the rules of analytics by enhancing the information collected offline with tons of new data online (Peterson, 2009a). Our "digital ID" allow companies to tie our behaviors together, thus allowing them to create an integrated view customers crossing over communication channels (Peterson, 2009a:5). Thus, the availability of data combined with an organization's increasing willingness to gather the information for insights, is leading to an analytical revolution (Peterson, 2009a).

### 1.1 Web Analytics- History, Current Scenario and Challenges

Web analytics dates back to the 1990s. However, web analytics was established as a discipline in the year 2000 (Kaushik, 2007). Basic web metrics, such as visits and web page views are considered to be the start of the web analytics concept. With the increased involvement of customers, companies found out that due to its simplistic and misleading nature, web metrics is not enough for measuring web site success (Phippen et al, 2004). Phippen et al (2004) argues that the internet empowers customers and for this reason companies today are adopting a more customer-centric approach. As customers are the key to both online and offline success, business need to understand how a customer engages with the Web site, and thus develop a marketing strategy based on customers rather than the organization (Phippen et al, 2004). Employing web analytics leads to understanding of the interactions between the visitor actions and the website and thus allows the company to use that insight to optimize the business and the website for improved customer loyalty.

Web analytics has grown tremendously since its rise but there is a lot of change and growth in front of it. During the recent years, companies have invested enormous amount of money on web analytics tools, support and personnel yet only a small percentage of those investment have translated into positive results. Avinash Kaushik, one of the most cited web analytics experts, contends that web analytics today is "just a toddler" (Kaushik, 2007:6). The lack of actual practical knowledge and real world practitioners driving strategy has had a detrimental effect on the field. Companies have access to the tools, but there is a lack of people and processes that would enable companies to garner actionable insights (Burby & Atchison, 2007). The problem is not that companies are not gathering web data, just the opposite. Somewhere between data collection and business improvement companies do something wrong in their analytical efforts (Burby & Atchison, 2007). Often the problem is that analysis is not integrated in the decision making of the organization. Many businesses look at analytics in isolation, as part of one department not of the whole organization (Burby & Atchison, 2007). To completely harvest the benefits of web analytics, the whole organization needs to become data-driven.

**1.2 Web Analytics Definition:** It is the methodological study of online/offline patterns and trends. A technique that can be employed to collect, measure, report, and analyze website's data. It is normally carried out to analyze the performance of a website and optimize its web usage. Web Analytics is used to track key metrics and analyze visitors' activity and traffic flow[1]. It is a tactical approach to collect data and generate reports. Web Analytics is an ongoing process that helps in attracting more traffic to a site and thereby, increasing the Return on Investment[4]. Web Analytics is used to assess the success rate of a website and its associated business[5]. **Web Analytics**, can be used to Assess problems in web content that can be rectified, Have a clear cut idea of trends that exist in website, Monitor web traffic and user flow, Demonstrate goals acquisition, Figure out potential keywords, Identify segments for improvement, Find out referring sources[2].

**1.3 Google Analytics:** The tracking and measuring number of visitors, traffic sources, conversion, goals, and other metrics are done by Google analytics tool (as shown in the above image). It provides a detailed statistics of the web traffic is used by more than 60% of website owners. It basically generates reports on: 1) Audience Analysis 2) Acquisition Analysis, 3) Behavior Analysis 4) Conversion Analysis

**1.3.1 Audience Analysis:** Audience analysis is used to know about the audience who visit the site along with their session history, bounce rate, page-views, etc. It can be used to trace the new user and returning users along with their geographical locations.

- 1) The age and gender of the audience under Demographics.
- 2) Language and location under Geo.
- 3) The affinity reach and market segmentation under Interests.
- 4) New and returning visitors, their frequency, and engagement under Behavior.
- 5) Custom variable report under Custom.
- 6) Browsers, Operating systems, and network of the audience in Technology.
- 7) Mobile device information under Mobile.

This helps in the study of the activity through custom modules that is created to capture the selections.

**1.3.2 Acquisition Analysis:** To find the sources from where your web traffic originates can be identified by using Acquisition analysis

Traffic sources – In the acquisition, all the sources of the traffic needs to be checked.

Major sources of the traffic are:

- i) Organic traffic is the traffic coming through all search engines (Google, Yahoo, Bing....)
- ii) Social traffic is the traffic coming through all social media platforms (like- Facebook, Twitter, Google+, ...)
- iii) Referral traffic is the traffic coming through where your website is linked.
- iv) Direct traffic is the traffic coming directly to your website. For example, typing the url of your website, clicking on the link of your website given in emails, etc.

Acquisition analysis is used to:

- 1) Capture traffic from all channels, particular source/medium, and from referrals.
- 2) Trace traffic through Google AdWords (paid search).
- 3) Determine traffic from search engines.
- 4) Queries, triggered landing pages, and geographical summary can be seen.
- 5) Track social media traffic.
- 6) It helps in identifying networks where the users are engaged.
- 7) Referrals from where the traffic originates can also be seen.
- 8) A view of the hub activity, bookmarking sites follow-up, etc.
- 9) In the same tab, A look at the endorsements in details.
- 10) It helps to measure the impact of social media on a particular website.
- 11) See which plug-ins gave a traffic to a particular traffic.

**1.3.4 Behavior Analysis:** It monitors users' activities on a website. Behavioral data can be categorized under the following four aspects:

**i) Site Content:** It shows how many pages were viewed. A detailed interaction of data across all pages or in segments like content drill-down, landing pages, and exit pages can be seen. *Content drill-down* is breaking

up of data into sub-folders[7]. *Landing page* is the page where the user lands, and exit page is where the user exits your site. Behavioral flow can be measured in terms of content.

**ii)Site Speed:** Page load time, execution speed, and performance data can be captured here[8]. The speed with which the browser can parse through the page can be seen. Further, the page timings, user timings, and get speed suggestion can be measured.

**iii)Site Search:** It gives a full picture of how the users search across the site, what they normally look for, and how they arrive at a particular landing page[9]. And their search patterns before landing onto the required page on the website can be analyzed.

**iv)Events:** Events are visitors' actions with content, which can be traced independently[10]. Example: downloads, sign up, log-in, etc.

**1.3.5 Conversion Analysis:** Conversion is a goal completion or a transaction by a user on your website. For example, download, checkout, buy, etc[11]. To track conversions in analytics, a goal needs to be defined and a URL that is traceable needs to be set.

**i)Goals** – Metrics that measure a profitable activity that the company want the user to complete. It can be set to track the actions[14]. Each time a goal is achieved, a conversion is added to the data. Goal completion, value, reverse path, and goal flow can be observed. □

**ii)Ecommerce** – ecommerce tracking can be set to know what the users buy from your website. It helps to find product performance, sale performance, transactions, and purchase time. Based on these data, decisions can be taken. □

**iii)Multi-channel funnels** – Multi-channel funnels or MCF reports the source of conversion; what roles the website plays, referrals' role in that conversion; and what all slabs did when users pass through landing page to conversion[15]. For example, a user searched for a query on Google search page, he visited the website, but did not convert. Later on, he directly typed your website name and made a purchase. All these activities can be traced on MCF.

**iv)Attribution** – Attribution modeling credits sales and conversions to touch points in conversion tracking. It lets one decide what platforms or strategy or module is the best for the business[17]. A person may visit a particular website so many times but might not make any purchase. In the last stage he may visit and directly makes purchase. Here, the last interaction model will credit direct for the conversion, whereas first interaction model will assign credit to paid medium[18]. This way, one can analyze what module should be credited for a conversion.

## II Web Analytics WA Tools

The market for Web Analytic tools is quite large. There are many products available which can be differentiated by the level of sophistication, the capabilities, the qualities and the reliability (Nakatani & Chuang 2011). Server logs which provided the raw data to be worked with were only the beginning of analysis.

**2.1 Optimizely :** An optimization platform to test and validate changes and the present look of a particular webpage[19]. It also determines the final layout. Multipage Testing , A/B Testing and Multivariate Testing can be performed by this tool to improve and analyze the website. It provides an administrative and management functionality that lets One create account, organize projects, and experiment[20]. This facility helps in tracking clicks, conversions, signups, etc. users are allowed to run tests and use custom integrations with Optimizely interface.

**2.2 KISSmetrics:** A powerful web analytics tool that delivers key insights and user interaction on a particular website. A clear picture of users' activities on the website and collects acquisition data of every visitor is obtained[21]. KISSmetrics helps in improving sales by knowing cart-abandoned products. It is used to find exactly when to follow up your customers by tracking the repeated buyers activity. KISSmetrics also helps in identifying the following:

- 1) Cart size
- 2) Landing page conversion rate
- 3) Customer activity on your portal
- 4) Customer bounce points

- 5) Cart abandoned products
- 6) Customer occurrence before making a purchase
- 7) Customer lifetime value, etc.

**2.3 Crazy Egg:** An online analytics application that provides eye-tracking tools. It generates heatmaps based on where people clicked on the website therefore gives an idea on where to focus. It helps in filtering data on top 15 referrers, search terms, operating systems, etc. Crazy Egg requires a small piece of JavaScript code to be placed on the site pages. Once the code is on the required site, Crazy Egg will track user behavior. servers will create a report that shows the clicks on the pages is being tracked[34]. Reports can be reviewed in the dashboard within the member's area of the Crazy Egg site. Crazy Egg offers insights in four different ways:

□

**Heatmaps:** It gives a defined picture of where visitors who clicked on a particular page. Where there is a need to make changes so as to improve conversions.

**Scrollmaps:** Gives insights of to what length people scroll down on a page. Crazy Egg, ensures where people leave the page and where to hold them exactly and where to add more to hold them for longer.

**Overlay Tool:** It gives an overlay report of the number of clicks occurring on a website. **Confetti:** Confetti distinguishes clicks segmented by referral sources, search terms, etc[37]. Now, the origin of the clicks to a particular site can be known, so the traffic sources is uncovered.

**Installation:** Insert JavaScript code on source code of a website. Crazy Egg will by default track the user behavior[39]. The servers generate reports providing you the view. Set dashboard to review the reports.

#### 2.4 Key Metrics:

Every business needs to find key metrics for its proper functioning. Just by having a tracking code in a website is not sufficient[40]. Analyzing may help in retaining the customer and hold them. The types of metrics can be distinguished by the quantity, the ratio which looks at metrics and their relation and the value which e.g. represents an URL or search referrer (Hassler 2010, p.105). Most commonly the following web metrics are used:

**Page View:** The page view represents the number of times a single page was viewed (Web Analytics Association 2008). In the past **hits** which represented requests to the server were measured from log files[41]. They are probably the oldest metric in Web Analytics (Hassler 2010, p.90). But since web pages nowadays include more than just text, one request can easy lead to 5 or more hits as every picture for example is counted separately[42]. Thus, the metric hit is not that important any more. Instead the page as a whole is looked at. It is a good metric to estimate the general demand of a website. The question to ask is if it is good or bad to have a high number of page views[43]. If the structure of a website is not leading to answers the number of page views might be quite high. But people might not be happy with the site. Only from the number of page views it cannot be stated if a site is successful or not (Kaushik 2007, pp.9, 140). Kaushik points this out by saying “... a page view is a page view; it is just a viewed page[44]. Nothing more, nothing less” (2007, p.77). One aspect which may become more and more important in the future is the information value of page views when using Rich Internet Applications like AJAX. They can lead to an underestimate of page views as it can happen, that a user only request one page and parts of the site are only reloaded while surfing (Hassler 2010, p.93).

**Most viewed Page:** Also called most popular pages or most requested URL the most viewed page simply counts which page had the greatest number of visits (Kaushik 2007, p.150).

**Visitors and Visits/Sessions:** These metrics are some of the basic ones in Web Analytics and provided by every tool[47]. The name thereby often is different and ranges from visitors and visits, over total visitors or unique visitors to sessions and cookies. It is important to understand how a tool really measured the specific metric. Most commonly the differentiation is the following:

**Visits/Session:** represents the number of visits (sessions) on a website during a specified time period[48]. One visit thereby is an interaction by an individual with one or more pages of a website (Kaushik 2007, p.133; Web Analytics Association 2008).

**Visit Duration/Time on Site/Length of Visit/ Time Spent per Visitor:** It is counted as the difference between the timestamp of the first minus the last activity in a session and shows the length of time in a session[51]. This time-based metric often is hard to evaluate unless it is known what the user did in between. Did he really viewed the page or was he doing something else? (Web Analytics Association 2008; Kaushik 2007, p.136). Furthermore, when a user stays on a page for e.g. five minutes that can mean that he is interested, that he is distracted, or cannot find what he was looking for (Sullivan 1997).

**Visitors:** counts the number of users of a website (Hassler 2010, p.98). It is important to notice that for example the number of unique visitors can be quite different from the total number of visitors and the question needs to be asked how the visitors were identified (Kaushik 2007, p.132).

**Unique Visitors:** (unique browser) tries to count the number of distinct users within a given time-frame. Each individual is only counted once. To distinguish between individuals authentication, IP-addresses or cookies can be used. All have different limitations so that it is important to understand how the number was made up in the specific case and to which timeframe it is counting (Web Analytics Association 2008; Kaushik 2007, p.133).

**New Visitors:** shows the number of unique visitors who viewed the site for the first time ever (Web Analytics Association 2008).

**Return(ing) Visitor:** shows the number of unique visitors who visit the site, but also have visited it before (Web Analytics Association 2008).

**Visits per Visitor:** The number of visits divided by the number of unique visitors in a given time period (Web Analytics Association 2008).

**Click-through:** This number indicates how often a link was clicked by a user (Web Analytics Association 2008).

**Referrer:** Referrer can show where users came from and what they searched for. The referring website shows the page where the user came from before viewing the current page (Web Analytics Association 2008). Search key phrases show what users searched for in search engines before they got to the current site. But it is not always possible to get referrer information. Around 40 to 60% of referrer data will be empty. Reasons are, among others, that the page was bookmarked or the entry was through direct type in of an URL (Kaushik 2007, p.145).

**Entry Page:** It shows the first page of a visit (the URL). As usually many entry pages exist it is often displayed by a list of URLs with the number of visits (Web Analytics Association 2008).

**Exit Page:** The exit page represents the end of a visit. It is the last page accessed on a site before leaving it (Web Analytics Association 2008).

**Exit Link:** Within a page often external links are included. When a user clicks an external link and exits the own website the link is called an exit link.

**Top Exit Page:** The top exit page is the page where most visitors leave the site from. The exit rate is the percentage number of how many people left from a specific page. However, how can these numbers be interpreted? Why do people leave at a specific site?

The answer can be either because they found what they were looking for (easy to discover with a thank you page of an e-commerce site) or they left because they couldn't find it (Kaushik 2007, p.9).

**Page Exit Ratio:** It is the number of exits from a specific page divided by total number of page views of that same page (Web Analytics Association 2008).

**Event:** An event is every recorded action with a specific date and time directed to the browser or the server. It can be either seen as a count (number of events) or as a dimension (how many visits can be associated with a specific event) (Web Analytics Association 2008) and often is associated with a specific page or session[52].

**Completion/Order Conversion Rate:** The percentage of visitors who purchase on the site (Ogle 2010; Hasan et al. 2009, p.703).

**Bounce Rate:** Depending on the author the bounce rate is defined differently. It is either the percentage of single page visits (Hasan et al. 2009, p.701) or the percentage of visits that only stayed on the site for a very short time (Kaushik 2007).

**Assessment of Content Popularity:** It is a ranking of the content of a website e.g. the most frequently showed texts (Ogle 2010).

**Path Analysis:** The path analysis shows how users click through a site (Ogle 2010). It is one kind of pattern mining as used within pattern discovery. This list could be endless, as each tool is able to generate hundreds of different metrics[53]. For example it could go on with the most popular pages, internal referrers and internal search phrases. The list above should only give an idea of what is possible[55]. Each Web Analytic tool measures the metrics a little bit differently and two different tools will never have the same numbers for one website. When trying to compare metrics the same analytic tool should be used. Keep in mind the measurement inaccuracies when comparing between websites[57]. Benchmarking on own measurements from today and the past are relatively resistant against measurement inaccuracies. The different metrics and the way they were measured are very important in the Web Analytic process.

**Demographic info** – Demographic data shows Age and Gender. With the help of Demographic Info, the percentage of Male/Female visitors coming to a website can be found out[58]. Analyzing the ratio of this data, can be used to make a strategy according to genders. Age group data helps in finding what percentage of age group visit a particular website. So, that a strategy can be made for highest percentage of age group visitors.

**Devices** – This data shows the devices information. Devices information, can be used to find how many percentage of visitors come from mobile, how many come from desktop, how many come from tablets, etc[59]. If mobile traffic is high, then the website needs to be more responsive.

The market for Web Analytic tools is quite large. There are many products available which can be differentiated by the level of sophistication, the capabilities, the qualities and the reliability (Nakatani & Chuang 2011). Gassman 2011) and browsing through the web, many websites can be found which describe “The best tools”, or the most used ones<sup>2</sup>. Reducing these to a common group it can be stated that the most used Web Analytic tools on the market right now are the following:

Adobe SiteCatalyst (Ominiture)*	Coremetrics	Sitestats
	Analytics (IBM)*	StatCounter
Analyzer NX (AT Internet)*	Etracker	SiteCensus
Click Tale	Google Analytics*	Webtrekk Q3
ComScore Digital Analytix*	NetInsight (IBM)	Webtrends Analytics*
Piwik	Netmind GO	Yahoo! Web Analytics*

### III Emerging Analytics Trends:

**3.1 Social Media Analytics:** Social Media Analytics comprise of gathering data from social media platforms and analyzing it to derive information to make business decisions[60]. It provides powerful customer insight to uncover sentiments across online sources. You tend to take control of Social Media Analytics in order to predict customers' behavior, discover patterns and trends, and make quick decisions to improve online reputation. Social Media Analytics also let you identify primary influencers within specific network channels. Some of the popular social media analytics tools are discussed below.

**3.2 Google Social Analytics:** It is a free tool that lets you add social media results to your analysis report. You get to know what is in air about your business[61]. How many people interacted with your website through social media and how many liked and shared your content.

**3.3 SumAll:** It combines Twitter, Facebook, and Google Plus into one dashboard to give you an overall view of what people are talking about you on social media. Facebook Insights Facebook plays a major role in your marketing campaign. You need to familiarize yourself with Facebook data to mark a flag[62]. You need to set up a page for your business to get the insights. It gives you information about who visited your page, saw your post, liked your page, and shared it.

**3.4 Facebook Insights:** Facebook plays a major role in your marketing campaign. You need to familiarize yourself with Facebook data to mark a flag[60]. You need to set up a page for your business to get the insights. It gives you information about who visited your page, saw your post, liked your page, and shared it.

**3.5 Twitter Analytics:** Twitter Analytics show how many impressions each tweet received, what is the engagement status, and when a particular user is at his peak.

**3.6 E-commerce Analytics:** Business owners need to survive and thrive among tough competition. They have to become big decision makers in order to survive in the market. This is where Web Analytics play a critical role[61].

**3.7 E-commerce Analytics:** Helps in figuring out customers' acquisition, users' behavior, and conversion. Google Analytics, can be used to get relevant information about the volume of sales, product with revenue, and sources of conversion occurred. All this information needs to be kept to find out where the business stands and to boost e-commerce sales, generate leads, and enhance brand awareness[62].

**3.8 Mobile Analytics:** Mobiles have emerged as one of the most significant tools in the past two decades. It changes the way people communicate and innovate. This has led to marketing driven by mobile apps. Mobile apps have proved easy to access and engaging[63]. Webmasters and online business makers need to take support of mobile apps to make their way perfect. Once making of a mobile app is done, it needs new users to be acquired, engage with them, and earn revenue. For this, mobile analytics needs to be used. It helps marketers to measure their apps better[76].

For example: How many people are using a particular app, How to optimize user experience, How to prioritize, What operating system to focus on, How to visualize navigation path, etc. Here are 10 case studies of companies that used insights from web analytics and drove ROI.

### IV Case Studies Comparison

**4.1 Bt Financial Group:** Is a leading provider of superannuation, investment, and insurance products. The BT website focuses on service and usability with an online application form as one of the key conversion points[68]. Landing pages with different combinations of the design elements for testing are created to optimize the user experience and maximize conversions. With conversion testing, BT increases form completions by more than 60%.

**4.2 BuildDirect:** does business in more than 100 countries with an expanding portfolio of building materials[69]. Though the company is growing rapidly, management is eager to improve the efficiency of its online spending. Through web analytics, BuildDirect finds home buyers who purchase a sample have a 60% likelihood of returning to the site within the next 30 days and placing a full order, BuildDirect uses GA's A/B testing capabilities to perfect its marketing approach. With insights from web analytics, BuildDirect increases sales by 50%.

**4.3 Harvard University:** To expand the digital reach of two established schools, Harvard Summer School and Harvard Extension School, Harvard ran a 12-month SEO and PPC campaign. They use web analytics as an audit to identify technical setbacks, content positioning, to create new landing pages for search traffic,

and top-of-funnel paid search awareness campaigns. The result are: 1) 89% increase in visits from organic search, 2) 75% increase in registrations from organic search, 3) 30% increase in CTR with AdWords, and 4) 124% increase in ROAS with AdWords[75].

**4.8 Puma:** Has rich, dynamic web site; but, just as PUMA constantly improves its products, it also believes in making site changes that help visitors easily achieve their goals. While testing its web site header, it finds a variation that increases online orders by 7.1%. Puma more than doubles the amount of time visitors spend interacting with PUMA brand content, such as news, videos, and photos. It results in 47% more traffic.

#### 4.9 Netflix

The core job of analytics is to help companies gain insight into their customers. Then, the companies can optimize their marketing and deliver a better product[70]. (Without analytics, companies are in the dark about their customers.) Analytics gives businesses the quantitative data they need to make better, more informed decisions and improve their services[71]. At current count, Netflix has 109.25 million worldwide streaming customers. Having this large user base allows Netflix to gather a tremendous amount of data. With this data, Netflix can make better decisions and ultimately make users happier with their service.

**Example,**“ In early 2000’s Netflix could not figure out what the core demographic was of their users and so invested a lot of time analyzing their customers trying to figure out under which of the following categories do they fall into 18-24 year old group and 24-35, etc. And found out that their customer was higher income, older, and more educated. What analytics and data can do is give you insight so you can run a better business and offer a superior product.

#### 4.11 HP

Hewlett Packard (HP) is one of the largest technology companies operating in more than 170 countries around the world. HP markets its products and services to both businesses and consumers via direct distribution, online distribution, consumer-electronics and office-supply retailers, software partners and major technology vendors (Hewlett Packard, 2010).

HP implemented web analytics in their online strategy in 2006. In the beginning, it was only one HP EMEA representative who received external training in internet marketing and web analytics. Further, a special EMEA team was established that had the main responsibility to collect information about the online data needs of the different business units[72]. Based on the needs of the company a web analytics tool was chosen. The EMEA team produced a report template with understandable metrics and Key Performance Indicators for the employees and soon after began to distribute a monthly web analytics report to all the countries. This was a strategic move to introduce the employees to the data-driven way of thinking[73]. As HP has both corporate and private customers, it is important that the employees understand correctly the enormous amount of online data. *“Top management is focused mainly on sales figures and not how results are delivered. They don’t pay much attention to which online campaigns lead to sales.”*

#### V Conclusion

Web Analytics is not just a project that can be undertaken. It is part of the continuous improvement process of an organisation (Gassman 2008) and should thus be seen as part of the whole. However the market is still without open standards for tools and metrics and each vendor has a unique approach (Herschel 2010). This is noting this thesis can or aimed to change, but what could be done was to establish a framework for developing a sustainable WA programme that widens the focus of the Web Analytic process to fully account for business requirements and to focus attention on achieving actionable outcomes.



**References:**

- [1] Abraham, M., Meierhoefer, C. & Lipsman, A., 2007. Cookie Deletion Whitepaper, comScore.
- Agrawal, R. & Srikant, R., 1994. Fast Algorithms for Mining Association Rules in Large Databases.
- [2] In VLDB'94, Proceedings of 20th International Conference on Very Large Data Bases. Conference on Very Large Data Bases. Santiago, Chile: Morgan Kaufmann, pp. 487–499.
- [3] Agrawal, R. & Srikant, R., 1995. Mining Sequential Patterns. In International Conference of Data Engineering. Taipei, Taiwan, pp. 3–14.
- [4] Aliaga, M. & Gunderson, B., 2002. Interactive Statistics 2nd ed., New Jersey: Prentice Hall.
- Araya, S., Silva, M. & Weber, R., 2004. A methodology for web usage mining and its application to target group identification. *Fuzzy Sets and Systems*, 148, pp.39–152.
- [5] Benbasat, I., Goldstein, D.K. & Mead, M., 1987. The Case Research Strategy in Studies of Information Systems. *MIS Quarterly*, 11(3), pp.369–386.
- [6] Berners-Lee, T., 1989. Information Management: A Proposal. The original proposal of the WWW. Available at: <http://www.w3.org/History/1989/proposal.html> [Accessed October 23, 2011].
- [7] Berners-Lee, T. & Fischetti, M., 2000. Weaving The Web: The Past, present and Future of the World Wide Web, London: Texere.
- Berry, M.J.A. & Linoff, G.S., 2011. Data Mining Techniques: For Marketing, Sales, and Customer [8] Relationship Management 3rd ed., John Wiley & Sons.
- [9] Berthon, P., Pitt, L.F. & Watson, R.T., 1996. The World Wide Web as an Advertising Medium. *Journal of Advertising Research*, 36(01), pp.43–54.
- [10] Blaikie, N.W.H., 2000. Designing Social Research, Wiley-Blackwell.
- Bruce, C.S., 1994. Research students' early experiences of the dissertation literature review. *Studies in Higher Education*, 19(2), pp.217–229.
- [11] Bundesministerium der Justiz, 2007. Telemediengesetz - Einzelnorm. Available at: [http://www.gesetze-im-internet.de/tmg/\\_15.html](http://www.gesetze-im-internet.de/tmg/_15.html) [Accessed April 1, 2012].
- [12] Burby, J. & Atchison, S., 2007. Actionable Web Analytics: Using Data to Make smart
- [13] Business Decisions, Indianapolis, Indiana: Wiley Publishing.
- Burks, A.W., 1946. Peirce's Theory of Abduction. *Philosophy of Science*, 13(4), pp.301–306.
- Cavaye, A.L.M., 1996. Case study research: a multi-faceted research approach for IS. *Information Systems Journal*, 6(3), pp.227–242.
- [14] CERN, 2008. CERN - Where the web was born. CERN - Where the web was born. Available at: <http://public.web.cern.ch/public/en/About/Web-en.html> [Accessed October 23, 2011].
- Checkland, P. & Holwell, S., 1998. Action Research: Its Nature and Validity. *Systemic Practice and Action Research*, 11(1), pp.9–21.
- [15] Chein, I., Cook, S.W. & Harding, J., 1948. The field of action research. *American Psychologist*, 3(2), pp.43–50.
- [16] Chen, W. & Hirschheim, R., 2004. A paradigmatic and methodological examination of information systems research from 1991 to 2001. *Information Systems Journal*, 14(3), pp.197–235.
- Chi, H., 2002. Improving Web Usability Through Visualization. *IEEE Internet Computing*, 6(3), pp.64–71.
- [17] Chishom, R.F. & Elden, M., 1993. Features of Emerging Action Research. *Human*, 46(2), pp.275–298.
- Connolly, D., 2000. A Little History of the World Wide Web. W3C: A Little History of the World Wide [18] Web. Available at: <http://www.w3.org/History.html> [Accessed April 11, 2011].
- Cooley, R., 2003. The Use of Web Structure and Content to Identify Subjectively Interesting Web Usage Patterns. *ACM Transaction on Internet Technology*, 3(2), pp.93–116.
- Cooley, R., Mobasher, B. & Srivastava, J., 1997.
- [19] Web Mining: Information and Pattern Discovery on the World Wide Web. In Proceedings of ICTAI' 97. Newport Beach, California.
- [20] Crabtree, B.F. & Miller, W.L., 1999. Doing Qualitative Research 2nd ed., Thousand Oaks, California: Sage Publications.
- [21] Creswell, J.W., 2003. Research Design: Qualitative, Quantitative, and Mixed Method Approaches 2nd ed., Thousand Oaks, California: Sage Publications.
- Crotty, M., 1998. The Foundations of Social Research: Meaning and Perspective in the Research Process, Sage Publications.
- [22] Crowder, P. & Crowder, D.A., 2008. Creating Web Sites Bible 3rd ed., Indianapolis, Indiana: Wiley Publishing.
- [23] Davidson, R.M., Martinsons, M.G. & Kock, N., 2004. Principles of canonical action research. *Information Systems Journal*, 14(1), pp.65–86.
- [24] Ding, W. & Lin, X., 2010. Information Architecture: The Design and Integration of Information Spaces, Morgan & Claypool.
- European Commission, 2006. Leonardo Da Vinci programme. Available at: [http://ec.europa.eu/education/programmes/leonardo/objectives\\_en.html](http://ec.europa.eu/education/programmes/leonardo/objectives_en.html) [Accessed December 26, 2011].
- [25] Facca, F.M. & Lanzi, P.L., 2005. Mining interesting knowledge from weblogs: a survey. *Data & Knowledge Engineering*, 53, pp.225–241.
- [26] Fallman, D., 2003. Design-oriented human-computer interaction. In Proceedings of the SIGCHI conference on Human factors in computing systems. CHI '03. New York, NY, USA: ACM, pp. 225–232. Available at: <http://doi.acm.org/10.1145/642611.642652>.
- [27] Fink, A., 2009. Conducting Research Literature Reviews: From the Internet to Paper 3rd ed., Thousand Oaks, California: Sage Publications.
- [28] Friedman, K., 2003. Theory construction in design research: criteria, approaches, and methods. *Common Ground*, 24(6), pp.507–522.
- [29] Gassman, B., 2009a. Five Best Practices for Web Analytics Initiatives, g.
- Gassman, B., 2008. Introduction to the Gartner Maturity Model for Web Analytics, Gartner Group.
- [30] Gassman, B., 2009b. Key Challenges in Web Analytics, 2009, Gartner Group.
- Gassman, B., 2011. Web Analytics Market Update, 2012, Gartner Group.
- [31] Geertz, C., 1973. The Interpretation of Cultures, New: Bas.
- Glaser, B.G. & Strauss, A.L., 1967. The Discovery of Grounded Theory. Strategies for [32] Qualitative Research, Transaction Publishers.
- [33] Grace, L.K.J., Maheswari, V. & Nagamalai, D., 2011. Analysis of Web Logs and Web User in Web Mining. *International Journal of Network Security & Its Application (IJNSA)*, 3(1), pp.99–109.
- [34] Greening, D.R., 2002. Data Mining on the Web: There's gold in that Mountain of Data.
- [35] Data Mining on the Web. Available at: <http://drdobbs.com/architecture-and-design/184413868?pgno=3> [Accessed October 26, 2011].
- [36] Gregor, S., 2002. A Theory of Theories in Information Systems. *Information Systems Foundations: Building the Theoretical Base*, Australian National University, Canberra, pp.1–20.
- [37] Kaplan, B. & Duchon, D., 1988. Combining qualitative and quantitative methods information systems research: a case study". *Manage. Inf. Syst. Q.*, 12(4), pp.571–586.
- [38] Kaplan, B. & Maxwell, J.A., 2005. Qualitative Research Methods for Evaluation Computer Information Systems. In J. G. Anderson & C. E. Aydin, eds. *Evaluating the Organisational Impact of Healthcare Information Systems*. New York: Springer, pp. 30–55.

- [39]Kaufman, L. & Rousseeuw, P.J., 1990. Finding Groups in Data: An Introduction to Cluster Analysis, Hoboken, New Jersey: John Wiley & Sons.
- [40]Kaushik, A., 2007. Web Analytics: An Hour a Day, Indianapolis, Indiana: Wiley Publishing.
- [41]Kohavi, R., Mason, L.I. & Zheng, Z., 2004. Lessons and Challenges from Mining Retail E-Commerce Data. *Machine Learning*, 57, pp.83–113.
- [42]Kosala, R. & Blockeel, H., 2000. Web Mining Research: A Survey. *SIGKDD Explorations*, 2(1), pp.1–15.
- [43]Kumar Jain, R., Kasana, R.S. & Jain, S., 2009. Efficient Web Log Mining using Doubly Linked Tree. *International Journal of Computer Science and Information Security (IJCSIS)*, 3(1).
- [44]Lawrence, D. & Tavakol, S., 2007. *Balanced Website Design: Optimising Aesthetics, Usability and Purpose*, Springer.
- [45]Lee, T.W., 1999. *Using Qualitative Methods in Organizational Research*, Thousand Oaks, California: Sage Publications.
- [46]Lewin, 1946. Action Research and Minority Problems. *Journal of Social*, 2, pp.34–46.
- [47]Lincoln, Y.S. & Guba, E.G., 1985. *Naturalistic Inquiry*, Beverly Hills: Sage Publications.
- [48]Lovett, J. et al., 2009. US Web Analytics Forecast, 2008 to 2014, Forster Research.
- [49]Madria, S.K. et al., 1999. Research Issues in Web Data Mining. In M. Mohania & A. Tjoa, eds. *Data Warehousing and Knowledge Discovery*.
- [50]Lecture Notes in Computer Science. Springer Berlin / Heidelberg, pp. 303–312. Available at: [http://dx.doi.org/10.1007/3-540-48298-9\\_32](http://dx.doi.org/10.1007/3-540-48298-9_32).
- [51]March, S.T. & Smith, G.F., 1995. Design and natural science research on information technology. *Decis. Support Syst.*, 15(4), pp.251–266.
- [52]Marshall, C. & Rossman, G.B., 2011. *Designing Qualitative Research 5th ed.*, Thousand Oaks, California: Sage Publications.
- [53]Masand, B. & Spiliopoulou, M., 2000. Web Usage Analysis and User Profiling B. Masand & M. Spiliopoulou, eds., Springer Berlin / Heidelberg. Available at: [http://dx.doi.org/10.1007/3-540-44934-5\\_1](http://dx.doi.org/10.1007/3-540-44934-5_1).
- [54]Merriam-Webster, 2001. Design - Definition and More. Design - Definition and More. Available at: <http://www.merriam-webster.com/dictionary/design> [Accessed September 12, 2011].
- [55]Miniwatts Marketing Group, 2011. Internet World Stats. World Internet Usage Statistics News and World Population Stats. Available at: <http://www.internetworldstats.com/stats.htm> [Accessed October 20, 2011].
- [56]Morville, P. & Rosenfeld, L., 2006. *Information Architecture for the World Wide Web 3rd ed.*, Sebastopol, CA: O'Reilly.
- [57]Muijs, D., 2004. *Doing quantitative research in education with SPSS*, London: Sage Publications.
- [58]Myers, 2011. Qualitative Research in Information Systems. *Qualitative Research in Information Systems*. Available at: <http://www.qual.auckland.ac.nz/> [Accessed May 12, 2011].
- [59]Nielsen, J. & Norman, D.A., 2000. Web-Site Usability: Usability on The Web Isn't ALuxury. Available at: <http://www.informationweek.com/773/web.htm> [Accessed October 20, 2011].
- [60]Norguet, J.-P., Zimányi, E. & Steinberger, R., 2006. Improving Web Sites with Web Usage Mining, Web Content Mining, and Semantic Analysis. *SOFSEM 2006, LNCS 3831*, Springer, pp.430–439.
- [61]Novak, T.P. & Hoffman, D.L., 1997. New metrics for new media: toward the development of Web measurement standards. *World Wide Web J.*, 2(1), pp.213–246.
- LDI NRW, 2009. Datenschutzkonforme Ausgestaltung von Analyseverfahren zur Reichweitenmessung bei Internet-Angeboten. Available at: [https://www.ldi.nrw.de/mainmenu\\_Service/submenu\\_Entschliessungsarchiv/Inhalt/Beschluesse\\_Duesseldorfer\\_Kreis/Inhalt/2009/Datenschutzkonforme\\_Ausgestaltung\\_von\\_Analyseverfahren\\_zur\\_Reichweitenmessung\\_bei\\_Internet-Angeboten/Analyse.pdf](https://www.ldi.nrw.de/mainmenu_Service/submenu_Entschliessungsarchiv/Inhalt/Beschluesse_Duesseldorfer_Kreis/Inhalt/2009/Datenschutzkonforme_Ausgestaltung_von_Analyseverfahren_zur_Reichweitenmessung_bei_Internet-Angeboten/Analyse.pdf) [Accessed March 1, 2012].
- [62]Nunamaker, J.F., Chen, M. & Purdin, T.D.M., 1990. Systems development in information systems research. *J. Manage. Inf. Syst.*, 7(3), pp.89–106.
- [63]Ogle, J.A., 2010. Improving Web Site Performance Using Commercially Available Analytical Tools. In *Symposium: AJS Carl . Brighton Workshop on Health Informatics*. Springer, p. 2604–261.
- [64]Orlikowski, W.J. & Baroudi, J., 1991. Studying information technology in organizations: Research approaches and assumptions. *Information Systems Research*, 2(1), pp.1–28.
- [65]Orlikowski, W.J. & Iacono, C.S., 2001. Research Commentary: Desperately Seeking the “IT” in IT Research - A Call to Theorizing the IT Artifact. *Information*, 12(22), pp.121–134.
- [66]Paliouras, G. et al., 2000. Clustering the Users of Large Web Sites into Communities. In *Proceedings of the Seventeenth International Conference on [67]Machine Learning. ICML '00*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., pp. 719–726. Available at: <http://dl.acm.org/citation.cfm?id=645529.657810>.
- [68]Pani, S.K. et al., 2011. Web Usage Mining: A Survey on Pattern Extraction from Web Logs. *International Journal of Instrumentation, Control & Automation (IJICA)*, 1(1), pp.15–23.
- [69]Power, N. & Moynihan, T., 2003. A theory of requirements documentation situated in practice. In *Proceedings of the 21st annual international conference on Documentation*.
- [70]SIGDOC '03. New York, NY, USA: ACM, pp. 86–92.
- [71]Rossi, M. & Sein, M.K., 2003. Design research workshop: a proactive research approach. Presentation delivered at IRIS, 26, pp.9–12.
- [72]Schubert, P., Kummer, M. & Leimstoll, U., 2004. Legal Issues of Personalized E-Commerce Applications. In *Proceedings of the European Conference on Information Systems (ECIS)*.
- [73]Sen, A., Dacin, P.A. & Pattichis, C., 2006. Current Trends in Web Data Analysis. *Communications*, 49(11), pp.85–91.
- [74]Souza, R. et al., 2000. *The Best Of Retail Site Design*, Cambridge, Massachusetts: Forster Research.
- [75]Spiliopoulou, M., 2000. Web Usage Mining for Web Site Evaluation. *Communications of the ACM*, 43(8), pp.127–134.
- [76]Spiliopoulou, M. & Pohle, C., 2001. Data Mining for Measuring and Improving the Success of Web Sites. *Data Mining and Knowledge Discovery*, 5(1-2), pp.85–114.