

Aleksi Ruuskanen

IBM Coremetrics – Web Analytics and Digital Marketing Optimization

Helsinki Metropolia University of Applied Sciences

Bachelor of Engineering

Information Technology

Bachelor's thesis

6.3.2013

Author(s) Title	Aleksi Ruuskanen IBM Coremetrics – Web analytics and digital marketing optimization
Number of Pages Date	40 pages 6 March 2013
Degree	Bachelor of Engineering
Degree Programme	Information Technology
Specialisation option	Software Engineering
Instructor(s)	Vladislav Jumppanen, Project Manager Auvo Häkkinen, Principal Lecturer
<p>The objective of this thesis is to gain deeper understanding of web analytics and how to use it to improve conversion rate of web sites and optimize digital marketing. The main objective is to identify the pros and cons of four analytics tools with a special focus on the one called IBM Coremetrics Web Analytics.</p> <p>This study is based on exploring and evaluating the IBM Coremetrics Web Analytics tool and comparing it to other major web analytics tools. It also draws on some third-party reviews and reports discussing the analytics tools covered in this study. Also included are two Coremetrics case studies, which give insight about how Coremetrics products perform in real life situations. Besides Coremetrics, the tools covered in this study are Omniture SiteCatalyst, Webtrends Analytics and Google Analytics.</p> <p>All four analytics tools covered in this thesis are hosted, web based applications that depend on JavaScript tagging. All of them provide sophisticated data that can be used to improve websites as well as optimize digital marketing. Choosing the right web analytics tool is highly situational due to certain critical differences. For instance, IBM Coremetrics requires a store built on IBM WebSphere Commerce, but it is easy to implement, whereas Omniture SiteCatalyst implementation requires a lot of planning and analysis. With Webtrends, implementation is quite easy through their implementation service but the tool has one major flaw: it's missing the common file formats that its competitors support, namely HTML and Word documents. Google Analytics is the only tool available for free. It offers some basic selection of metrics and conversion events, but more advanced advertisers might find themselves limited with Google's selection.</p> <p>Based on the findings, Coremetrics would be the obvious choice for a complex e-commerce site using WebSphere Commerce whereas a smaller site, which is not so complex would do well choosing Google Analytics. This study should be helpful for any web business wishing to collect specific information on their customers in order to improve their performance through digital marketing.</p>	
Keywords	Web analytics, IBM Coremetrics, Omniture, Webtrends, Google Analytics, Digital Marketing

Tekijä(t) Otsikko Sivumäärä Aika	Aleksi Ruuskanen IBM Coremetrics – Web analytics and digital marketing optimization 40 sivua 6.3.2013
Tutkinto	insinööri (AMK)
Koulutusohjelma	Tietotekniikka
Suuntautumisvaihtoehto	Ohjelmistotekniikka
Ohjaaja(t)	Projektipäällikkö Vladislav Jumppanen Yliopettaja Auvo Häkkinen
<p>Opinnäytetyön tarkoituksena on saada perusteellinen näkemys kävijäseurannasta, ja siitä kuinka sitä voi käyttää hyödyksi verkkosivujen vierailijoiden käännyttämiseksi vakituisiksi vierailijoiksi ja digitaalisen markkinoinnin parantamiseen. Päätaivoitteena on tunnistaa neljän kävijäseurantatyökalun hyvät ja huonot puolet, keskittyen erityisesti työkaluun nimeltä IBM Coremetrics Web Analytics.</p> <p>Opinnäytetyön perustana on IBM Coremetrics Web Analytics työkaluun tutustuminen, sen arvioiminen, sekä sen vertaaminen muihin suosittuihin kävijäseurannan työkaluihin. Työ sisältää myös muiden osapuolten arvosteluja ja raportteja käsitellyistä työkaluista. Lisäksi työ sisältää kaksi tapaus tutkimusta IBM Coremetricsista, jotka antavat kuvaa siitä, kuinka Coremetricsin tuotteet toimivat oikeissa tilanteissa. Coremetricsin lisäksi työssä käsitellään Omniture Sitecatalystia, Webtrends Analyticia, sekä Google Analyticia.</p> <p>Kaikki tässä työssä käsitellyt työkalut ovat verkkopalveluja, jotka keräävät tietoa verkkosivun kävijöistä JavaScript tagien avulla. Kaikki työkalut soveltuvat erinomaisesti erilaisten tietojen keräämiseen vierailijoista ja kerättyä tietoa voi käyttää niin verkkosivun kuin digitaalisen markkinoinnin parantamiseen. Oikean työkalun valitseminen on hyvin tapauskohtaista. IBM Coremetrics vaatii IBM WebSphere Commercella rakennetun verkkokaupan, mutta se on helppo ottaa käyttöön, kun taas Omniture SiteCatalystin käyttöönotto vaatii tarkkaa suunnittelua. Webtrends tarjoaa palvelua jolla käyttöönotto on helppoa, mutta työkalusta puuttuu tuki yleisille tiedostoformaateille, joita muut työkalut tukevat. Google Analytics on ilmainen työkalu, mutta sen ominaisuudet eivät välttämättä ole riittävät kokeneemmalle käyttäjälle.</p> <p>Tulosten perusteella IBM Coremetrics on selvä valinta IBM WebSphere Commerceen perustuvalla laajalla verkkokaupalla, kun taas Google Analytics soveltuu hyvin pienemmille verkkosivuille. Tämä työ antaa tietoa verkossa toimiville yrityksille jotka haluavat parantaa digitaalista markkinointiaan keräämällä tietoa sivuilla vierailevista kävijöistä.</p>	
Avainsanat	Kävijäseuranta, IBM, Coremetrics, Omniture, Webtrends, Google Analytics, Digitaalinen markkinointi

Contents

Abbreviations

1	Introduction	1
2	Web analytics	1
2.1	History and evolution of web analytics	1
2.2	Data-gathering methods	4
2.2.1	Web server logging	4
2.2.2	Web bug	5
2.2.3	Page tagging	6
2.2.4	Packet sniffing	8
3	Digital marketing optimization	9
3.1	Defining digital marketing	9
3.2	Ways to optimize digital marketing	10
3.2.1	Search Engine Optimization	10
3.2.2	Social Media Marketing and Optimization	14
3.2.3	Paid Search Marketing	15
4	IBM Coremetrics	17
4.1	Introducing Coremetrics	17
4.2	Implementing Coremetrics	20
4.2.1	Configuring communication between the store and Coremetrics	20
4.2.2	Category Definition File	21
4.2.3	Tagging store pages for Coremetrics	22
4.3	Using Coremetrics Web Analytics	24
4.3.1	Application overview	24
4.3.2	Dashboards	25
4.3.3	Reports	25
4.3.4	Workbooks and Manage	27
5	Comparing Coremetrics to other web analytics tools	28
5.1	Coremetrics and Omniture	28

5.2	Coremetrics and Webtrends	29
5.3	Coremetrics and Google Analytics	30
5.4	Comparison conclusion	31
5.5	What others thing	31
6	Coremetrics success stories	33
6.1	Boden increases trouser sales by 160 percent with IBM Coremetrics	34
6.2	Ace Mart reduces online marketing spend by 25 percent to improve ROI	35
7	Conclusion	35
	References	37

Abbreviations

AJAX	Asynchronous JavaScript and XML
API	Application Programming Interface
CDF	Category Definition File
CSS	Cascading Style Sheet
CSV	Comma Separated Values
HTML	Hypertext Markup Language
IP	Internet Protocol
ISP	Internet Service Provider
JSON	JavaScript Object Notation
JSP	JavaServer Pages
KPI	Key Performance Indicator
PDF	Portable Document Format
PPC	Pay Per Click
ROI	Return On Investment
RSS	Rich Site Summary (originally Really Simple Syndication)
SaaS	Software as a Service.
SEO	Search Engine Optimization
SMM	Social Media Marketing

SMO Social Media Optimization

XML Extensible Markup Language

1 Introduction

A good website has objectives that define the purpose of the site. Objective is an event that the site wishes for users to complete. Objectives differ from site to site, but some common objectives are for example registering to the site, subscribing to marketing newsletter or making an order in web store. Websites seldom are perfect and they have room for upgrades. Web analytics is used to measure, collect, analyze and report website usage data in order to improve the satisfaction of website objectives. Website usage data can include for example the number of unique visitors to the site, total number of page loads, information about how the visitors reached the site, and information about the visitors' browser and operating system.

Many websites also implement digital marketing. Digital marketing consists of web, email and mobile marketing. The most visible form of digital marketing is the ads on websites, other forms are for example Search Engine Optimization (SEO), Paid Search Marketing, and Guerilla marketing taking place in social media. Data gathered with web analytics can also be used to optimize and personalize the digital marketing.

The purpose of this thesis is to focus on one of the web analytics tools called IBM Coremetrics Web Analytics and describe it to the reader along with general idea of web analytics and digital marketing optimization. We intend not only to learn how to use and implement it, but also compare it to other popular web analytics tools and see if there is room for improvement. We will also take a look into how other parties perceive Coremetrics in the form of reviews, reports and case studies.

2 Web analytics

2.1 History and evolution of web analytics

Web analytics was born in the 1990s as the Internet became commercialized and more and more people had access to it. The early websites were static, created by a single file that could contain text and links. As simple as these sites were even they sometimes encountered errors and it didn't take long before someone came up with the idea to take advantage of server logs to obtain information about the number of client re-

quests (also known as hits) made to the web server. Server logs also had additional information about the requests, such as time of request, web browser and operating system used, and a referrer (website sending the request) if a visitor followed a link to the site [1].

As a growing number of nontechnical people became interested in the data contained in the server logs, it became clear that the data needed to be presented in more understandable form. Soon the first scripts were created that automatically parsed the server log files and created metrics from the gathered data, and thus web analytics was officially born. One of the first server log analysis programs was Analog. Analog is a free software created by Stephen Turner in 1995 and is still widely used. Figure 1 shows a report from an early version of Analog, and how basic the first metrics were.

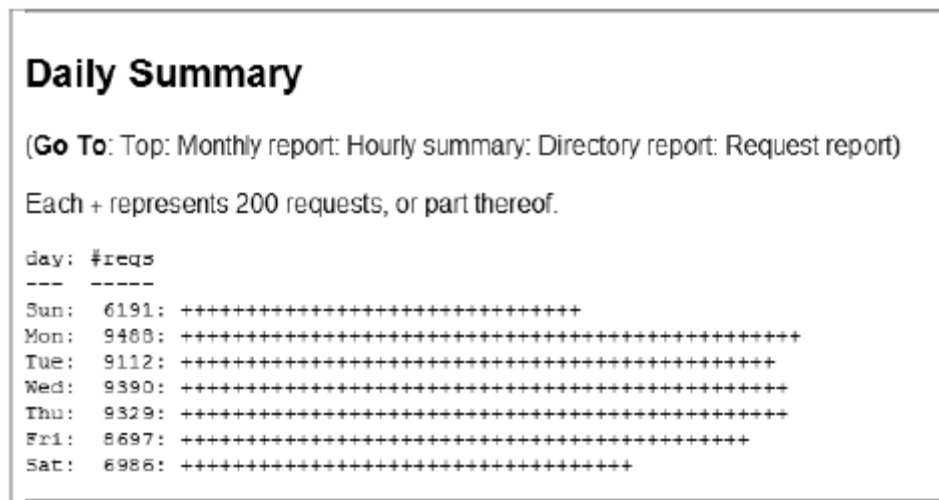


Figure 1. A sample report from Analog, version 0.9 beta [1].

Now that the marketing people could understand what was happening on the website and web analytics popularity quickly grew, reports also improved appearance. One of the pioneers of commercial web analytics, Webtrends, not only made improvements for the standard log file parsers, but also added tables and graphs to the reports, making server log data even more accessible [1].

Soon website visitors too could see web analytics in action, as web counters (also known as hit counters) became very popular around 1995-96. A web counter tracks the number of hits made to the page and shows them visually on the site, usually with digital images or plain text. Figure 2 shows web counter examples of both types.

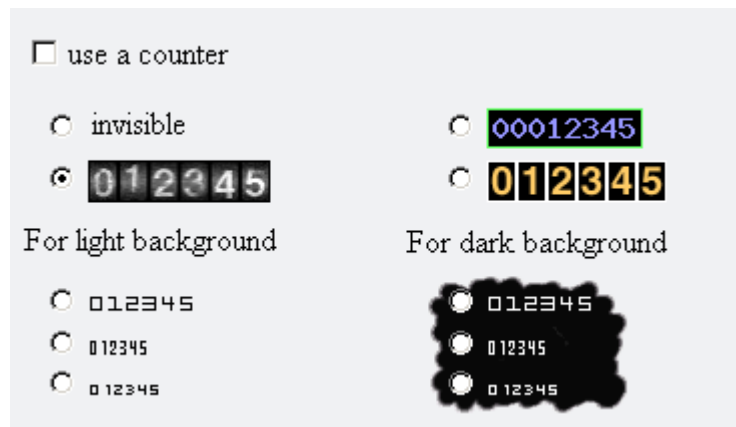


Figure 2. A webmaster's counter options [2].

Web counters were considered a very appealing feature that could show how popular your website was and were included on nearly every site.

In the late 1990s server log analyzers started to experience challenges in terms of providing accurate data. As the popularity of search engines continued growing, so did the usage of web crawlers. Web crawler, also known as search bot, search engine spider and many other names, is a computer program that browses the Internet automatically or orderly. It is often used by search engines for providing up-to-date data and fast searches, website maintenance for validating Hypertext Markup Language (HTML) code and to ensure that the links work, and by spammers for gathering email addresses [3]. Problem was that the crawler visits to the site were counted by the log analyzers while they were not real visitors.

One other challenge was page caching made by Internet Service Providers (ISP). When ISP had copy of the page, all succeeding pages would be served from the ISP and log analyzers would not register the entries for those pages. ISPs also assigned dynamic Internet Protocol (IP) addresses for users. This in addition to the increased usage of web proxies made identifying unique visitors very difficult [1]. While these challenges could be countered to some degree by tracking visits by cookies and ignoring requests made from known web crawlers, concerns about log data accuracy and desire to outsource the web analytics services led to another way for collecting data from the site called page tagging [4].

Page tagging uses JavaScript for gathering data and sending it to a data collection server. Page tagging is very simple to use as it is usually just a few lines of JavaScript

code embedded to each page, making web analytics implementation a lot easier. Using JavaScript also made it easier to gather additional information (such as the visitors' screen size or the price of the purchased goods) [4]. Companies could also outsource the collecting and processing of data to web analytics vendors.

The next big step in web analytics was the introduction of click analytics. As the name implies, click analytics is a special type of web analytics that focuses on clicks made on the site [4]. Rather than going through piles of charts and tables, decision makers could now see exactly what the users were doing on the site. This also led to increased usage of analytics, since now anyone, not just web analysts, could understand what was happening on the page. Click analytics also made it easier to improve the site based on the visitors' behavior [1].

The evolution of web analytics continues on, as new innovations are made for presenting the complex data gathered from site in newer and easier ways. One of the more recent innovations is a web heat map, which shows the areas where users have clicked most frequently by using colors. The more clicks, the brighter the color. Another recent innovation is a session replay, which can replay what the visitor did on site, including mouse movements, clicks and form entries.

2.2 Data-gathering methods

This section introduces several methods for collecting data, how they work and their pros and cons.

2.2.1 Web server logging

The oldest data collection method, i.e. server logs, were first used as a means to capture errors encountered on the site and later enhanced to capture additional information. It works in all simplicity as follows:

- A visitor types the address of the site.
- A request for the page is sent to web server.

- The request is accepted by the web server and an entry about the request is created in the logs. The entry can contain different kinds of information, such as visitor's IP address, browser used et cetera.
- The web server sends the requested web page to the visitor.

Because every web server comes with the means to capture data and create logs, web server logging is probably the easiest way to access data. There are also many log analyzers available for free, so it is easy to start producing basic metrics. Web server logs are usually taken from the server on a set interval, so log analyzers can be configured to produce automatic standard reports. Web server logs are also the only way to gather data about web crawlers, as they do not execute JavaScript tags. It is good to ensure that the web crawlers crawl and index the pages correctly. It is also good to note that data produced with server logs are the property of the website, while most other data-gathering methods sends data to web analytics vendors who capture, analyze and store it. In case of switching web analytics vendors, it is easier to re-analyze old data with new tools if the data is owned by the website [1].

Data gathered with server logs is mostly technical, however, and not very good for capturing business or marketing information. Cookies are needed for identifying visitors accurately. As server logs capture all the requests, some requests need to be filtered out, such as image or Cascading Style Sheet (CSS) file requests, page errors or web crawler activity in order to get accurate reports about website traffic. ISP caching also causes some traffic data loss because the ISP serves the web page instead of the web server and no data is recorded in the server logs [1].

Kaushik (1, pg. 66) recommends that web server logging would be mainly used for gathering data about web crawlers in order to measure the success of SEO, and that other methods would be used for all other types of web analysis.

2.2.2 Web bug

Web bug, also known as web beacon, is an object included in the website or email and is usually invisible to the user. Web bugs are often 1x1 pixel transparent images, but newer web bugs can use HTML IFrame, style, script, embed, object, and other tags for tracking activity [5]. When the page or email is opened, a request is made for the image, which is usually hosted on a third-party server. This server then sends the image back to the web browser, along with some code that can read cookies and gather ano-

ymous visitor data, like the fact that page was viewed, time of viewing, and previously set cookies. Web crawlers do not make image requests, so they are not included in the reports [1].

Web bugs are often used by third parties, who have no access to the server log files, in order to monitor the effects of ads and banners on the website. They can also be used by publishers, who have content on many websites, to collect and store data on one server that is hosting the image. Web bugs are also used a lot in spamming and phishing and by email marketers to verify that the email containing the web bug has been sent to a valid email address, it has passed the spam filters and the email has been actually viewed [5].

Web bugs are not very reliable form of collecting data on websites. They often rely on third-party cookies, which in turn are affected by increasingly strict privacy constraints. Web browsers might not even accept the third-party cookie or an Antispyware program might delete the cookie altogether [5]. Email programs typically have images disabled when opening the mail or the email could be viewed as plain text, which means that no request is made. The collected data is also less rich than what could be collected by using page tagging.

2.2.3 Page tagging

Page tagging, also known as JavaScript tagging, is a piece of JavaScript code on the page. Most web analytics vendors and programs are using it to gather data [1]. When a page is loaded, JavaScript code is executed and information about visitors' session along with cookies is sent to a data collection server. The data collection server can in some cases send back additional code in order to set new cookies and gather more information. While data collecting and processing is usually done in third-party data collection server, usually web analytics vendors' servers, some vendors offer the chance for JavaScript driven solutions that keep the data in-house, giving greater control regarding privacy and setting cookies, and maintaining the ownership of the data [1].

Example code 1 shows how one of the more popular web analytics tools, Google Analytics, is added to the site.

```

<script type="text/javascript">
  var _gaq = _gaq || [];
  _gaq.push(['_setAccount', 'UA-XXXXXXXX']);
  _gaq.push(['_trackPageview']);
  (function() {
    var ga = document.createElement('script'); ga.type =
'text/javascript'; ga.async = true;
    ga.src = ('https:' == document.location.protocol ? 'https://ssl' :
'http://www') + '.google-analytics.com/ga.js';
    var s = document.getElementsByTagName('script')[0];
s.parentNode.insertBefore(ga, s);
  })();
</script>

```

Example code 1. Including Google Analytics on the site [6].

As seen in example code 1, including page tagging is a very simple task. Only the tracking code (UA-XXXXXXXX) needs to be changed for a code provided by Google [6]. While the code needs to be on every page in order for it to work, it can be added to a global site element that is included on every page (like footer), making it easy to maintain. In cases where websites are hosted on third-party web servers, therefore restricting or downright denying access to web server logs, page tagging is the only way to gather web usage data [1].

Some other pros of page tagging are that it completely ignores ISP caching. Since JavaScript is executed every time a page is loaded, it does not matter from where the website is delivered, data will be collected anyway. Another big plus for page tagging is the excess control over collected data and the possibility to add custom tags on certain pages, such as shopping cart or checkout, in order to capture additional information like the name, price and quantity of purchased product. JavaScript tagging also has lots of potential for new innovations and ways of gathering data. Considering all this, it is easy to see why most web analytics vendors prefer JavaScript tagging as their method for collecting information.

JavaScript tagging does have some cons too. Some users prefer keeping JavaScript disabled, making the collecting of data with page tagging impossible. While page tagging is usually provided by web analytics vendors, some might consider doing them in-house, in which case creating the tags will take a lot of thinking and planning and they need to be maintained as the site evolves. There are also sites that store visitor session data on the server rather than in cookies or URL parameters, reducing greatly data collected with page tagging. And finally, in the worst case scenario, page tagging

might cause conflicts with other JavaScript components on the site, worsening the usability of the site or downright breaking it [1].

2.2.4 Packet sniffing

The final method for collecting data presented here is packet sniffing. Packet sniffer, also known as network analyzer, can collect data by intercepting the traffic between a web server and the outside world [7]. When a page request is sent to a web server, it first passes by a packet sniffer which collects the attributes that can help gather data about the visitor before passing the request to the web server. The request is then sent back to the packet sniffer, which can collect data about the page returned and possibly add a JavaScript tag that allows additional information to be gathered before sending it back to the visitor [1].

Packet sniffer can be software installed on web servers, but it can also be hardware connected to a data center, with all traffic to the web server being routed through it. Thanks to the nature of the solution, it does not need JavaScript or third-party cookies, thus eliminating the challenges that come with using either of them. Packet sniffing can also create lots of data instantly by combining technical data with page related business data [1].

While packet sniffing can work without JavaScript, it is still needed in order to truly collect optimal data for web analysis. This of course brings back the problems that come with page tagging. Another layer needs to be created between the visitor and a web server, and in cases where there are multiple web servers, packet sniffing can become quite an expensive solution. In addition, because packet sniffer collects the data from raw packets of Internet web server traffic, extensive configuration is needed to collect only relevant data. Raw data contains also Personally Identifiable Information (PII), such as passwords and credit card numbers, so the packet sniffer solution would need not only careful stress testing, but also legal review [1].

3 Digital marketing optimization

3.1 Defining digital marketing

Digital marketing is a very broad term, consisting of things like banners on websites, emails sent in marketing purposes, SEO, and many other things. In all simplicity it could be said that digital marketing is online advertising and it targets the users of computers, tablets, mobile phones, game consoles and all other devices that have access to the Internet [8]. There are two concepts of digital marketing: push and pull.

Push in digital marketing means delivering content to a consumer directly without permission. This is usually done by sending emails or text messages. Push has the advantage of targeting a specific segment of customers and therefore marketing tailored for that group can be more effective than mass marketing. Push marketing provides specific information about consumers' actions, such as whether the content has been viewed or deleted [9]. Push marketing is also used by spammers, with no intention of targeting certain consumers and instead just mass sending content.

Pull digital marketing on the other hand means that the consumer actively seeks the marketing content. Some examples of pull digital marketing are:

- Consumer navigates to a specific website.
- Consumer clicking on a link or banner ad in order to get more information about the advertized product.
- Consumer subscribing to marketing newsletters or Rich Site Summary (RSS) feeds and viewing received content.
- Using SEO to improve visibility of content and make it more appealing to consumers.

As the effectiveness of pull digital marketing is directly related to the advertisements' ability to arouse consumers' interest, some thought needs to be put into not only the advertisement itself, but also the location of advertisement, for example by placing a banner ad on a website visited mostly by the consumers targeted in the marketing content. Although the effectiveness of advertisement can be tracked with the number of clicks it has received, there is no way to trace individual consumers who view the mar-

keting content (unless they willingly give out their information), leaving no chance to contact consumer in the future and thus limiting the marketing scope.

3.2 Ways to optimize digital marketing

Audience of digital marketing is ever-growing; the latest update from Internet World Stats shows that on 30 June 2012 there were a total of 2.4 billion Internet users, which is over one third of the worlds' population [10]. Ryan and Jones write in the preface of their book as follows:

It took television 22 years to reach 50 million households; it took the internet just five years to achieve the same level of penetration. Things are progressing at an unbelievable rate, and we're approaching a pivotal point in marketing history – a time when digital marketing will overtake traditional mass media as the medium of choice for reaching the consumer of tomorrow. [11, pg. XIV]

From the quote and the statistics we can see just how large and important part of marketing digital marketing is and how vital it is for reaching the consumers. In order to maximize the effectiveness of digital marketing, some optimization is almost always needed.

In this section we introduce some common ways for digital marketing optimization.

3.2.1 Search Engine Optimization

SEO consists of many tweaks made to the website that together make the site more search engine friendly. However, as the main consumers of the site are its visitors, optimizations should be made so that the site remains user friendly. The optimal situation would be where the site ranks high on search results and converts most new visitors into regular visitors.

Every page on the site should contain descriptive and unique title. The <title> tag is placed within the <head> tag inside an HTML document. Not only does a title tell a visitor and a search engine what the page is about, it is also usually the first line in search results. If a title is too long, only a portion of it is shown, so it is good to keep titles relatively short, yet accurate. The title should also contain the most important keyword that relates to the content of the page. Cramping too many keywords in the

title or using keywords that do not have anything to do with the content gives the site unprofessional feel and reduces the visitor numbers.

Another important tag that goes into the header section of the page is the <meta> tag. It provides metadata for search engines and is not visible on the page itself. The <meta> tag consists of attribute value and content. For example if the attribute would be name and the value would be description, content would be summary for the page. In the Example code 2 the syntax for author <meta> tag can be seen.

```
<meta name="author" content="Aleksi Ruuskanen">
```

Example code 2. Author <meta> tag.

The most important attribute regarding SEO is description, giving search engines information what the page is about. In the 90's the keywords <meta> tag was also widely used. It was supposed to inform search engines about the site relevant keywords, but it was easy to misuse. It often provided misleading and unreliable information and in some cases even drew visitors to spam sites. Since 1998 search engines one by one dropped support for the keywords <meta> tag and in 2002 even the last major search engine to support them, Altavista, stopped relying on them. Nowadays keywords <meta> tag has little or none importance in ranking the site [12].

Good site structure, with main sections and subsections, is an important part of any site, and there are numerous ways to improve it. If possible, every page on the site should have a friendly URL that refers to it. A friendly URL is an easily readable web address that includes words relevant to the content. For example, a friendly URL for a web stores' jeans category could be something like www.webstore.com/clothes/pants/jeans. It is easy to remember and tells both the consumer and the search engine what the page is about. URLs are always displayed with search results and while more complex URL structures don't really matter to search engines, a URL that is long and contains different parameters or ID numbers may turn away potential consumers [13].

Good site structure also includes easy navigation. Navigation is meant to help consumers find the content they are looking for. Navigation also provides search engines information about content that a webmaster thinks as important. If the site has lots of content, dividing relevant content into subsections prevents the navigation menu be-

coming too overwhelming for a consumer. A good general rule is that it does not take more than three clicks for a consumer to find what he or she is looking for from the home page of the site. When a site has subsections or sections even below that, it should also have a breadcrumb and sitemaps. A breadcrumb is a row of internal links for the upper pages of subsection all the way to the home page, allowing the consumer to quickly navigate to previous sections. A sitemap has all the pages on the site in organized model, essentially showing the structure of the site. One sitemap should be created for consumers, and another one for search engines. For consumers there should be an HTML site map page, which helps consumers if they are having trouble finding the page they are looking for. While search engines crawl through HTML site map page too, an Extensible Markup Language (XML) Sitemap should be prepared for them. Example code 3 is a sample XML Sitemap containing one URL with additional information, such as when it was last updated.

```
<?xml version="1.0" encoding="utf-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>http://example.com/</loc>
    <lastmod>2006-11-18</lastmod>
    <changefreq>daily</changefreq>
    <priority>0.8</priority>
  </url>
</urlset>
```

Example code 3. Sample XML Sitemap [14].

An XML Sitemap can be as simple as to contain only URLs used on the site in order to ensure that search engines discover all the pages. However giving additional information with URLs can make search engines crawl the site in more optimal way.

When a site has many pages, it can happen that a consumer navigates to a non-existent page, either by typing a wrong URL or by following a broken link. In cases like these, instead of the general 404 Not Found error message, the site should have a custom error page. A good custom error page should have at least a way for consumer to navigate back to the home page of the site, but the best solution would be to have a full navigation menu. A custom error page should also be consistent with the style of the site and should not contain any cryptic error code messages. Instead a well written message, like an apology for the error encourages the consumer to continue using the site.

Besides making sure that the links are working properly they can also be optimized further. Anchor text is the clickable text that forms a link and it is placed inside the anchor tags in an HTML document. Links can be internal, pointing to the other pages on the same site, or external, allowing access to content provided by other sites. Much like a title, descriptive anchor text will help both the search engine and consumer understand what the linked page contains and most things that make a good title apply also to anchor text.

Unique and interesting content is probably the most important factor for making the brand known and increasing the consumer conversion rate. Good content will also affect positively on SEO by ranking the site higher, as well as social media marketing as the consumers who like the content are likely to pass on their experiences. Many consumers deem other consumers as trustworthy sources and testimonials from the latter will influence behavior like buying decisions of the former. Providing fresh content and easy-to-read text that is focused around the topic is the key for quality content, which will allow the site to grow recognition on its own.

Even the best content is useless if the consumers will not find it. Keywords and keyword phrases are the search terms which are used to find a site. Clever usage of keywords and keyword phrases in content ensures that a consumer finds the site he or she is looking for. It is also good practice to take into consideration that consumers who are familiar with the content might use different search terms than those who are new to the content. With good mixing of advanced and basic keywords and keyword phrases one can ensure that all kinds of consumers will find the site.

Most web analytics tools provide a way for keywords research and analysis, making it easier to find the optimal keywords that are not only the most used, but also unique enough to target the desired consumers and easy to optimize into content related sentences. Finding a few best keywords is important, as search engines tend to punish the sites that have too much keyword density (ratio of keyword usage and total number of words in content), treating them as spam. The right amount of keyword density is hard to determine since each search engine have their own recommendation, but common recommended density is between 2 and 4 %. Keyword placement and proximity is also an important part of content optimization. Having two or more keywords close to each other and placing them at the beginning of title, description, HTML headings and first and last sentences allow the site to achieve higher ratings.

3.2.2 Social Media Marketing and Optimization

Social media channels are great way for organizations to increase their trustworthiness, make their brand better known and engage consumers, but they can also damage an organization's reputation if used wrong. Social Media Marketing (SMM) is used to gain attention or traffic to the site via social media platforms like Facebook, Twitter or Youtube. Social Media Optimization (SMO) is used to optimize the site and its content for easier sharing in social media including things like page having Facebook Like button, and using RSS feeds. SMO works well with SEO, as search engines use sharing and "liking" the site as an indication of its quality, boosting the site's rating.

The first course of action is to determine which social media channels are suited for bringing in more consumers to the site. Facebook and Twitter are pretty much a norm nowadays and almost every organization utilizes them. If the site has content in video or audio form, platforms like Youtube are good places to share them. Blogging about new content and giving a chance to subscribe to RSS feeds about blog entries or marketing news are also good ways to increase consumer awareness. There are programs for social media monitoring which can give insight into what platforms are the most suited for the site and help identify topics of interest, making it easier to create and optimize content.

Once accounts for social media platforms are established, it is time to provide content for them or launch a social media campaign. All social media platforms should be kept up-to-date: updating some while neglecting others can make an organization look disorganized. Most social media accounts can be linked together and updating one account updates the others as well. To get the most out of social media, organization should also engage consumers. Conversation with consumers can increase the organization's reputation, not to mention getting direct feedback to help improve services and content. The organization's online communication should be transparent and honest. Sharing information in a truthful manner instead of keeping to oneself builds trust. Social media monitoring programs can also be used to monitor discussions about the organization and can help find negative comments, which can then be addressed properly [15].

Most social media channels also have tools that organizations can use to make the content on the site easier to share or follow. With a single button the consumer can

share the content and expand its reach. This is practically free advertisement and promotion, and should be used wherever possible.

3.2.3 Paid Search Marketing

Paid search marketing is a method for gaining more traffic to the site by displaying advertisements or sponsored links with search results. It is sometimes referred to as pay per click (PPC) marketing since most advertisements are sold based on the PPC advertising model. Advertisers specify a list of keywords and phrases, and when searched with them, the advertisement is shown. Advertisements are usually text based and are displayed on top of organic (nonpaid) search results or in a vertical column [16]. Figure 3 shows Google's PPC-program called Google AdWords in action.

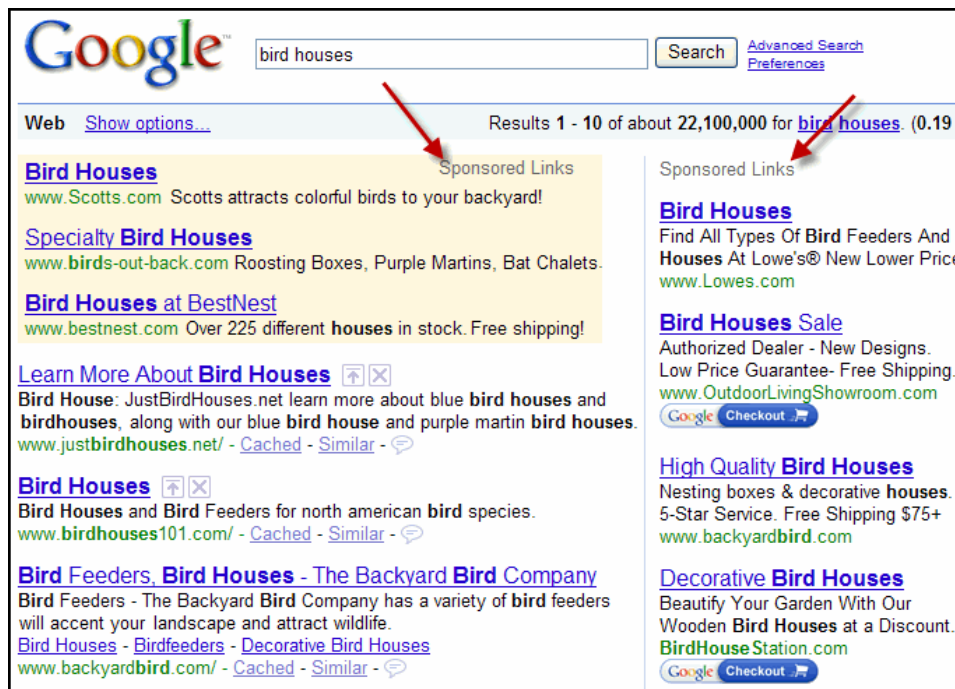


Figure 3. Paid search advertisements on Google [17].

PPC means that the advertiser only has to pay agreed amount when the consumer clicks the ad; showing it is free. Sponsored links which are displayed higher than others usually have higher cost per click, but search engines use also quality variable as a way to determine order of sponsored links. Google and Microsoft call the variable Quality Score while Yahoo calls it Quality Index. Quality Score includes such things as click

through rate (ratio of times ad is displayed to times ad is clicked), landing page quality, geographical considerations, ad relevance to the search and daily budget [18].

There are several ways to optimize paid search ads. Finding the best keywords to attract targeted consumers is very important and can be done with keywords research and analysis, similarly to SEO content optimization. Having a good landing page can boost sites' Quality Score as well as making it easier to convert the consumer. Since PPC ads involve a fee, it is also essential to pay attention to the ad's location and price. The cost of the PPC can be either flat-rate or bid-based.

With flat-rate PPC advertiser and publisher agree on a fixed amount that is paid with every click. The cost is determined by how many other websites are using same keywords or phrases to draw traffic and also by content. Content which is likely to attract more valuable consumers has higher price than content which attracts less valuable consumers. Lower rates are however usually negotiable with the publisher. Flat-rate PPC is commonly used on comparison search engines, such as eBay. They are usually compartmentalized into product or service categories, and because of that the consumers can be targeted to a very high degree, and because they are already farther along the buying process, there is a higher chance for conversion [16].

Bid-based PPC is a system where advertisers compete against each other in an advertising network, such as Google AdWords or Bing Ads. Advertisers inform publishers what is the maximum amount they are willing to pay in an auction for a given ad spot or a keyword, usually with the help of online tools. The auction is automated process which is run every time an ad is clicked or search is performed using auctioned keywords. When the ad is shown on the search result page, there are several factors that determine which bids are competing against each other. Some examples of these factors are the geo-location of the searcher and the day and time of the search [16]. The winning ads are then displayed and since there are usually a high number of them, their order is calculated with the formula $\text{bid} * \text{Quality Score}$. The exact weight of the Quality Score has not been revealed by any major search engine [19]. Every time the ad is clicked, the advertisers pay for the amount they had bid for.

Major advertising networks also have partnerships with third-party websites, which sign up to show contextual ads on their site for a portion of the ad revenue. Contextual ads are associated with the keywords of the website content of the third party and they

usually have a lower click-through rate than ads shown on search result pages and are therefore cheaper. There are systems that can automate the bid management in order to maximize success and achieve scale. They can be used by the advertisers themselves, but are more commonly used by advertising agencies who sell PPC bid management as service [16].

SEO usually takes some time before the site gets a higher rating with search results, whereas a PPC campaign can generate instant traffic to the site. The advertiser can decide what to show in the ad title, the ad description and the page where the ad leads, giving him or her greater control for targeting specific segments of consumers. Monitoring PPC campaigns allow advertisers to see which PPC ads and keywords work well, so they can drop the ones that don't contribute much and on the other hand reuse and optimize the well performing ones. PPC campaigns can also sometimes include real bargains, often when using a common keyword in a phrase, in which case the advertiser can generate traffic to the site with the fraction of the cost of any other form of paid advertising.

Luckily there are many web analytics tools to help improve the sites' performance as well as optimizing digital marketing. Some web analytics tools are products as licensed software, commonly the ones that use web log files for tracking. Most major web analytics tools are however available as Software as a Service (SaaS), meaning that the software and its associated data is located in a cloud. Prices vary a lot, from totally free to negotiable or fixed prices. In the next section we will focus on one of the major web analytics software.

4 IBM Coremetrics

4.1 Introducing Coremetrics

Coremetrics is a provider of digital marketing optimization solutions and web analytics and has been one of the key vendors in web analytics industry since the raise of web analytics. The solutions are designed for generating high return for online marketing investments. In 2010, IBM acquired Coremetrics, making it part of IBM's Smarter Commerce initiative. Dan Hobin writes as follows while referring to a report published by Josh Bernoff of Forrester:

The Information Age is evolving into a whole new age. With commoditization and the continuous disruption of technology, the Age of Information has given way to the Age of the Customer – an era where only “customer-obsessed companies will survive”. [20]

Like Dan Hobin writes, a new age has started. As the world has become more and more interconnected, instrumented and infused with intelligence, a new breed of customer has emerged: the empowered customer. They are empowered by the latest technology, increased business transparency and great amounts of information. They expect to engage with companies when and how they want, either in person, or using online and mobile means. And they want these methods to work together seamlessly. They can share their experiences with the whole world and have also access to the experiences of all the other empowered customers. It's not anymore enough for companies to merely react to customers' needs and preferences, they have to predict them. For this reason they need to understand the motivation behind customers' buying decisions [21]. IBM's Smarter Commerce initiative was made to deliver software and services to help companies transform their business processes to ones more suitable for the empowered customer.

Coremetrics became the integrated web analytics solution for IBM WebSphere Commerce and IBM WebSphere Portal. In this thesis the focus is on the WebSphere Commerce. It is a software platform framework for electronic commerce, also known as e-commerce. E-commerce is the buying and selling of services or products through the Internet or other electronic measures. WebSphere Commerce includes sales, marketing, customer and order processing functionality and offers the ability to do business with consumers, directly with other businesses and indirectly through channel partners. WebSphere Commerce has three main components: a database, Java application server, and a webserver. Web pages in WebSphere Commerce are done with Java-Server Page (JSP) technology, which can create dynamic web pages using HTML, XML and other document types [22].

Through the fusion of customer profiles, digital marketing execution and testing, and web analytics, a solution called IBM Coremetrics Digital Marketing Optimization Suite was born. It offers hundreds of tools and features such as:

- Cross-channel real time analytics and reporting for evaluating marketing effectiveness

- Personalized customer experience based on their lifetime interaction with the business, delivering highly targeted ads and recommendations
- Comparative benchmarks and intelligence about the performance of peers and competitors
- PPC management tool for monitoring PPC campaigns or automated bid management

All the tools and features are available as part of products in the cloud-based Coremetrics application which is accessed with a web browser. The application has an easy to use interface and enables the optimization of web presence and digital marketing channels with flexible and intuitive analytics dashboards, benchmarking and reports. Figure 4 has an example view of the Coremetrics Web Analytics application.

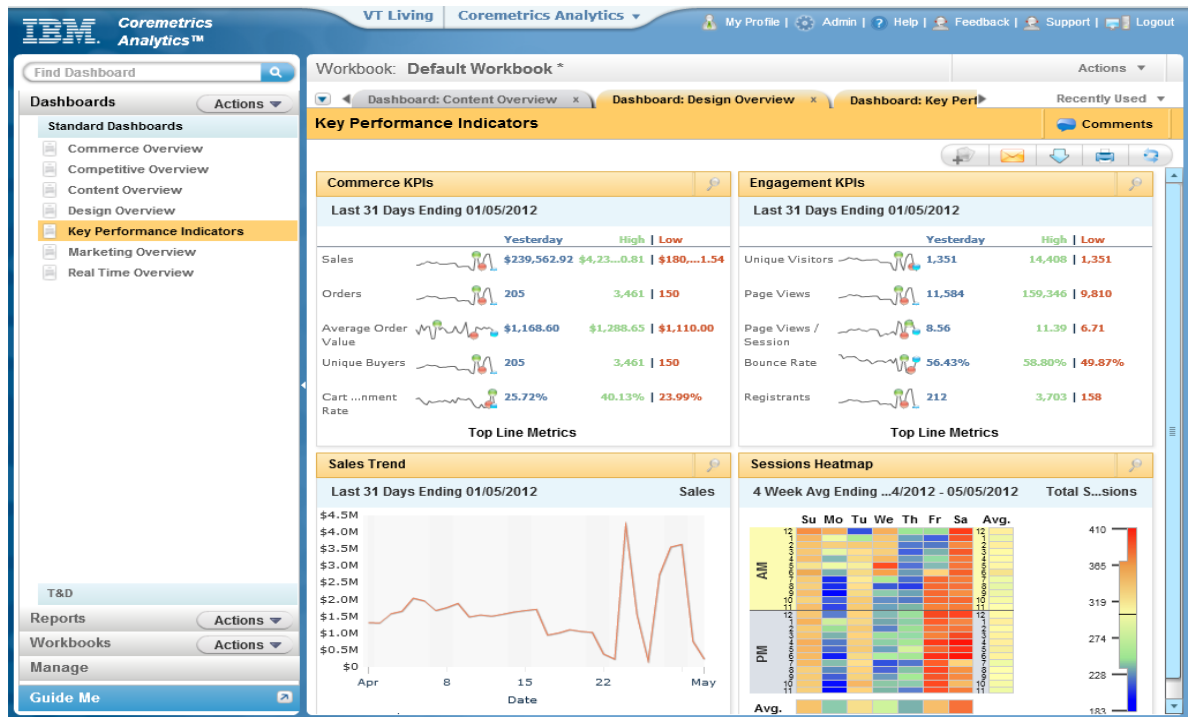


Figure 4. Dashboards of Key Performance Indicators in Coremetrics Analytics [23].

In the Figure 4 can be seen several dashboards for Key Performance Indicators (KPI). Dashboards provide a business snapshot and each dashboard can be modified, for example by changing the timeline. Dashboards can also be created easily from scratch if the standard dashboards provided by Coremetrics are not enough.

4.2 Implementing Coremetrics

Coremetrics is a paid web analytics tool, with the price being negotiable and depending on things like monthly server calls and enabled features. The first thing to do is to establish a contract with Coremetrics for collecting data and generating and hosting analytics reports. After that it is time to integrate the store using WebSphere Commerce with Coremetrics. In this section we will take very detailed look at how this integration is done. The following information and instructions are taken from the WebSphere Commerce version 7 Information Center [24].

Before beginning the integration, the store should be using WebSphere Commerce version 7 Feature Pack 3 or higher. Coremetrics uses WebSphere Commerce analytics tags to gather data. Feature Pack 2 uses custom Coremetrics library while with Feature Pack 3 the tags have been updated to work with the standard Coremetrics library, making it possible to take advantage of new analytics features as they become available.

4.2.1 Configuring communication between the store and Coremetrics

The configuration details are located in a XML file called `biConfig.xml`. Enabling the store to communicate with Coremetrics requires copying the sample `biConfig.xml`, placing it in the right directory and modifying some values in it. The values that must be changed are `storeId`, Coremetrics client ID's and the domain name of the stores' WebSphere Commerce Server. `StoreId` represents the stores' entity ID in the WebSphere Commerce database. If multiple stores have the same Coremetrics client ID, they can be enabled at the same time by either separating `storeId`'s with commas or giving a range for the `storeId` values. Coremetrics provides its' clients with two ID's, one for sending data to the Coremetrics production server and another one for sending data to the Coremetrics test server.

In addition to the mandatory changes, some additional configuration can be made, such as enabling or disabling debugging, defining whether to use the standard or custom Coremetrics library, specifying export mode for updating Coremetrics-based customer segments, and whether to use e-mail address for identifying customers in Coremetrics instead of the customers' member ID in the database. Customer segments are individuals who are grouped together based on common indicators relevant to marketing, such as age or gender. Coremetrics-based new customer segments can be either

appended to existing customer segments or they can replace them. Using e-mail as a means to identify customer can be used for example to correlate a customer filling in a product inquiry form on an external site to the same customer registering with the store. E-mail identification can also help to track multiple orders made by the same guest shopper as guest shoppers get a different unique member ID every time they make a purchase, while their e-mail is likely to stay the same. Using e-mail for identification should however be used only when e-mail is a mandatory information when registering to the store or making purchases. The biConfig.xml can also contain common JavaScript functions and statements in its <header> and <footer> elements, which are then placed on every page that has page view tag.

4.2.2 Category Definition File

Coremetrics requires Category Definition File (CDF) for adequate classification of content and products. It can be generated using a CDFGenerator command-line utility, which produces a comma-separated values (CSV) file in the Coremetrics defined CDF format. CDF file describes the product catalog hierarchy. It can also capture page content information if relevant data is added manually to the generated output file. Each line in the file defines a category in the hierarchy and has four values: Coremetrics Client ID, Category ID, Category name, Parent Category ID. For top level categories that don't have a parent category, the Parent Category ID will be left empty. If a category has multiple parent categories, Coremetrics will issue warnings. In a case like this, the warnings can just be ignored and Coremetrics will reject automatically the additional parent categories. Example code 4 shows the CDF file format.

```
99999999,101,MENS,
99999999,102,SALE,
99999999,123,MENS SALE,101
99999999,123,MENS SALE,102
```

Example code 4. Example of CDF file with additional parent categories [24].

In the example code 4 the last two lines are a unique category having two different parent categories, which would invoke warnings. In addition to Coremetrics automatically rejecting additional duplicate records, they can be removed manually from the output file or if the additional parent categories are really important, Coremetrics can provide other implementation alternatives. The CDF file needs to be uploaded to Coremetrics before any page tagging is done.

4.2.3 Tagging store pages for Coremetrics

WebSphere Commerce comes with several starter stores: fully functional sample stores that can be used to as a base for customized online stores. If a store is based on one of the starter stores, all the default stores' JSP pages are already tagged with WebSphere Commerce analytics tags and can be just enabled from the stores' management tool called Management Center. If the store or parts of it is custom-made, most analytics tags must be added manually to the pages. WebSphere Commerce analytics tags generate Coremetrics JavaScript data tags, which send data to Coremetrics. Analytics tags start with "cm" prefix and Coremetrics library needs to be included before using them.

The only analytics tags that can be created automatically are `<cm:pageview />` and `<cm:output />`. By running the auto tagging utility all the JSP pages are tagged with both of them. The page view analytics tag generates a Coremetrics tag: `cmCreatePageViewTag(pageID, categoryID, searchString, searchResults, attributes, extraFields)`. It collects information about viewed pages and passes the following data to Coremetrics: page name, page category, search keywords if the viewed page is search results page, search results count if the viewed page is search results page, store identifier and optionally extra parameters for custom reports in Coremetrics Analytics. Many of the analytics tags support the extra parameters. The page view tag must be the last tag within `<body>` element. The other automatically created tag, `<cm:output />`, runs the JavaScript functions and statements defined in the `biConfig.xml` `<header>` and `<footer>` elements.

The order tag, `<cm:order />`, collects information about the customer's order in two data groups, one for the items purchased, and other one for the order details. For each item purchased some data is passed to Coremetrics. This data includes for example product's name and part number, quantity purchased, base and subtotal prices, customer and order identifiers and currency used. The order details data group has information like order's identifier and subtotal, shipping charges and also customer's identifier, state, city and ZIP code. The order tag should be placed inside order confirmation page.

The shopping cart tag, `<cm:cart />`, collects data for each item in shopping cart, similar to order tags' items purchased data group, but leaving out data about order or customer. It should be included in the shopping cart page.

Other basic tags include `<cm:product />` which collects information about products that customers have viewed and is included in the product page, `<cm:registration />` which collects customer's registration information and is included in the registration page, and `<cm:campurl />` which is a campaign tag and collects information such as impressions (times that an ad has been viewed) and click-through rates about marketing campaigns that use WebSphere Commerce's e-Marketing Spots.

There is also the conversion event tag, `<cm:conversion />`, which collects information about events that are not directly related to product purchases, but may have affected the buying decision. Such events might be things like using a product comparison tool, using a store locator, adding items to wish list, subscribing to newsletter or downloading documents. The conversion tag does not generate any data by itself. Instead the tag parameters must be used in order to pass in the data that one would like to collect. The required and optional parameters for the conversion tag are:

- `eventId`. A required parameter, unique identifier for the type of conversion event.
- `category`. An optional parameter allowing grouping of events into defined category.
- `actionType`. A required parameter expressing whether the conversion event was initiated or successfully completed. There are two valid values, 1 and 2. 1 means that the event was initiated, while 2 means that the event was completed successfully.
- `points`. An optional parameter indicating the value of the conversion event. For example initiating low-value event might be worth 5 points and successfully completing high-value event might be worth 50 points.
- `extraparms`. An optional parameter for providing additional information for custom reports.
- `returnAsJSON`. An optional parameter to use when customer interactions do not cause full page refresh. JavaScript Object Notation (JSON) is a lightweight data-interchange format suited for JavaScript. A value can be true in which case a JSON object is returned and false in which case the analytics data tag is generated.

The conversion event tags should be placed in pages which load after a specific conversion event has been triggered.

Lastly there is the element tag, `<cm:element />`, which can track interaction with content within a page, such as interacting with flash movie, using faceted search or inte-

racting with Asynchronous JavaScript and XML (AJAX) components. AJAX techniques allow updating parts of the page without reloading it. Like the conversion tag, the element tag also needs tag parameters in order to generate data. The element tag has many optional parameters, like page identifier and location of the element in the page, but only one required parameter: `elementId`. `ElementId` specifies an identifier for the interacted element and it should represent a recognizable name for the elements, such as “Faceted search”.

While many analytics tags have the option to return them as JSON, the element tag is the one that really needs it. In cases where a new page isn’t loaded or a full page reload doesn’t happen, the `returnAsJSON` parameter should be set to true and the element tag placed in the same page that has the interactive element. Otherwise it can be used similarly to conversion event tag and placed in the loaded page.

4.3 Using Coremetrics Web Analytics

When the implementation is done and the store is able to communicate with Coremetrics, it is time to start using it. Coremetrics Web Analytics can be accessed by navigating to <https://welcome.coremetrics.com>, where a simple login page can be found. The client is prompted to enter his or her client ID, user name and password. If the client logs in for the very first time, Coremetrics Web Analytics will display a welcoming screen offering a product tour, getting started user guide and other valuable resources.

4.3.1 Application overview

As seen in Figure 4 (on p. 19) which displayed an example view of the Coremetrics Web Analytics, the application is divided to three parts: header, left navigation pane and workbook pane. From header the client can access any other client IDs that are linked to his or her user ID, as well as switch between obtained Coremetrics products and learn more about the ones not obtained. Header also has options for customizing the profile, section for system administrators, for getting help or support, and a logout button. The left navigation pane launches dashboards, reports, workbooks and manage options or finds specific parts of them using a search. The Left navigation pane also has a guide me button to give best practice assistance related to viewed report. Selec-

tions made in the left navigation menu can be seen and accessed in the workbook pane.

4.3.2 Dashboards

Flexible dashboards provide a combination of real time metrics, key performance indicators, funnels, top performers, benchmarks and much more in a consolidated view for quick research. They can be used as a basis for more detailed analysis or to promote a discussion about business performance. Coremetrics Web Analytics comes with a set of standard dashboards, giving insight into areas like Commerce, Competitive Benchmarking, Content, Design, Key Performance Indicators, Marketing and Real Time Information. If one finds standard dashboards lacking, new ones can be created, or existing ones edited or deleted.

4.3.3 Reports

Reports are divided into many types. The basic report view provides a report with chart and table of a single time period or optionally of a compare period. Trend view provides a visual understanding for movement across multiple time periods. Some reports also have views for heat map or insight. Figure 5 has an example view of the basic report type.

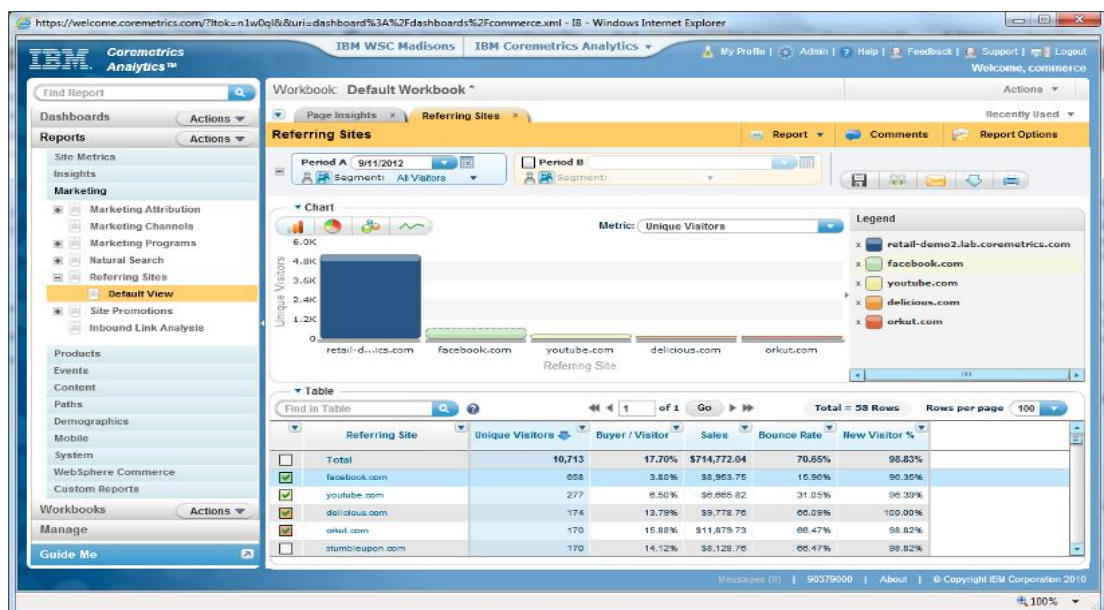


Figure 5. Example view of the Coremetrics Web Analytics basic report type [25].

In the Figure 5 can be seen two reports, with the Referring Sites basic report view being active. Up to 10 reports can be open simultaneously and switching between them can be done using tabs. The report type can be changed from the report header's dropdown. Report header has also buttons for comments and report options. Comments can be used to discuss reports, explain something or just to leave reminders to oneself. Report options allow a number of changes to be made to the reports, such as time periods, comparison settings, and metrics, filters or segments used. Reports also have calendar, which can be used to quickly change the timeline, and buttons for saving, emailing, sharing, printing or downloading the report. Reports can be saved as image, or in Excel or CSV formats.

Tables can be used to gain additional information, as well as to manage what is shown in the charts. Charts used in reports come in many varieties. Most reports can be visualized by using bar charts, pie charts, bubble charts and trend charts. Bar charts shows information using rectangular bars with lengths being proportional to the values they represent. Pie charts are circular charts which are divided into sections, each section's size illustrating proportion. Bubble charts require at least two visible metrics in the report and are usually used to analyze three metrics at once. Trend charts can be viewed in standard trend to understand movement over multiple time periods, but also in cumulative view, which sums up each successive data point to provide a cumulative aggregation through the trend period. Trend charts have also the option to add and view annotations for important events that may have affected the trend.

Some reports also have options for heat map, insights, clickstream and funnel visualizations. Heat maps are available only for the Top Line Metrics reports. They show activity on the site with colors, red being the most active time and blue the least. With heat maps it is easy to plan for example outages or when to run campaigns. Insight reports may help to see things that would have been otherwise overlooked. Top Line Insights allow client to trend four chosen metrics simultaneously to discover which dates are outliers for chosen metrics. Marketing, Product and Page Insights allow a brief look into what programs, products and pages are helping with objectives and what the distributions look like. Clickstream reports analyze the popular paths that the visitors take. It can be used to notice trends in visitor behavior as well as to conclude which paths are most effective and lead to conversions. Funnel reports visualize the visitors leaving the web site through a series of steps on the site, like how many visitors browsed something, initiated an event, or completed an event. It can be used to meas-

ure the success of visitors completing events such as registering or checkout, and to identify the pages that cause visitors to leave.

4.3.4 Workbooks and Manage

Workbooks are collections of reports. As up to ten reports are launched in to different tabs, they can be saved as a workbook for quick access in the future. Other options for workbooks include creating a new workbook, deleting existing workbook, caching the workbook so it loads faster and an option to quickly close all the tabs in workbook pane.

In the Manage area client can manage outbound distributions, marketing programs and channels, installations and report options. Report options simply allow management of comments, annotations and segments.

Distribution allows the management of scheduled Action Ready Reports, e-mails, alerts, and the Coremetrics report application programming interface (API). Action Ready Reports provide vital information about the web business in an attractive and accessible format. E-mail section can be used to manage scheduled dashboard and report emails. Alerts are e-mail notifications sent when a certain point has been reached for a chosen metric or data type. The Coremetrics report API offers a programmatic access to all reports and can be used for example to deliver reports in-house or to build custom widgets that use the data from reports.

Marketing allows the management of invalid marketing programs report, marketing channels and tracking codes. Invalid marketing programs report helps to track if any campaign tags have improperly formatted parameters or if they are missing elements. Marketing channels are the highest level of marketing source categorization and they can be created, edited or deleted. Tracking codes can be used to measure the effectiveness of both onsite and off-site links. Coremetrics Web Analytics comes with a tracking code generator.

Installations allows for programmatic access to reports from Excel by installing and using the Excel API. Another thing that can be installed is the Coremetrics Tools Browser Plug-in. The tools can be used for example to overlay site traffic and conversion data over the web site. Installations have also tools for managing mobile access

options. The mobile content features access to Top Line Metrics, key reports and some real time metrics.

5 Comparing Coremetrics to other web analytics tools

This section looks at how Coremetrics differs from its competitors. Coremetrics is compared with three other major web analytics tools: Omniture, Webtrends and Google Analytics. All four of them are hosted, web based applications that depend on JavaScript tagging. This section also takes a look at some recent reviews and research about web analytics tools.

5.1 Coremetrics and Omniture

Omniture's history is similar to that of Coremetrics. Omniture has also been one of the key vendors in web analytics industry since the beginning and in the 2009 it was acquired by Adobe. Omniture's web analytics tool is called SiteCatalyst and it is also a paid tool.

Implementing SiteCatalyst can be quite hard especially for someone who is new to web analytics. Implementing some basic recording of site traffic, referrals and visitor information is pretty easy. Using the SiteCatalyst tool the client must first create a report suite, which will store the collected data. When it is created, the client needs to fill out some attributes. After that the client can use Code Manager found within SiteCatalyst to generate the JavaScript code needed on the site. The generated code needs to be saved in a file named s_code.js and then included on every page on the site in order for SiteCatalyst to start collecting information.

The problem comes however when wanting to really harness the power of SiteCatalyst. The client needs to review the key objectives for the site and then identify the KPIs that will tell if the goals are being met. Identified KPIs are used to create a Solution Design Document, which will serve as a roadmap for implementing the code to the site. In other words, SiteCatalyst implementation requires a lot of planning and analysis, whereas Coremetrics provides a lot of information, including KPIs, out of the box [26].

The abilities of the two web analytics tools don't differ that much in the end, although SiteCatalyst does deliver some advanced video tracking. Both provide tons of features and information. Aside the fact that SiteCatalyst might require a bit more commitment, only major differences are with visitor tracking and customer support. Coremetrics LIVE profiles (Lifetime Individual Visitor Experience) tracks and collects data of the web site visitors over the long term at a level no other web analytics tool can provide [27]. SiteCatalyst is not however lacking with visitor details and can provide lots of meaningful data, just not at the level of Coremetrics.

Another difference is in customer support. SiteCatalyst offers a telephone support only to main contacts within an organization, and an online knowledge base. With Coremetrics when an agreement is made, Coremetrics provides clients with an account manager and a dedicated client executive with no additional cost. Their job is to get to know the clients business and provide them with help whenever needed. In addition to that, Coremetrics also has advanced training courses and webinars to provide clients with deeper understanding on Coremetrics's applications and services. Coremetrics also offers an annual conference called Client Summit, which provides a great opportunity for clients to learn about new and innovative ways to take advantage of Coremetrics.

5.2 Coremetrics and Webtrends

Webtrends is one of the oldest web analytics companies. It was founded in 1993 and has set the standard for innovation and value ever since. It is still a privately owned company. Like Coremetrics Web Analytics and SiteCatalyst, Webtrends Analytics is a commercial web analytics tool.

Webtrends offers implementation service to help their clients with the implementation. Their consultants will help the clients to discover key objectives and identify KPIs and then construct the data collection and report design. They then partner up with their clients' technical staff in order to implement, test and deploy the solution based on the design [28].

Webtrends Analytics provides information and reports on things that most clients do not probably even know that they need. Like with Coremetrics, Webtrends customer support is of very high quality when a Care Plan is purchased, providing telephone, email,

knowledge base and online chat supports. Webtrends Analytics visitor tracking is on par with SiteCatalyst's and can be used to create targeted advertising and marketing campaigns, taking into account even the likes and dislikes of the visitor. Webtrends Analytics is however lacking in one area: file exporting. File exporting in Webtrends Analytics supports CSV, XML and Excel formats. Portable Document Format (PDF) is also available but not by default. Webtrends Analytics is however missing the common file formats that its competitors support: HTML and Word documents.

5.3 Coremetrics and Google Analytics

Of the reviewed web analytics tools, Google Analytics is the only one available for free. Being free of charge is also the biggest advantage of it, making it highly popular. In 2005 Google acquired a web analytics company called Urchin. Urchin had software for analyzing web server logs together with an online version of the product called Urchin on Demand. Google Analytics was based on the Urchin on Demand. Google kept releasing new versions of the Urchin software while developing Google Analytics, but as the Urchin software was overshadowed by Google Analytics, Google finally discontinued the Urchin software in March 2012 [29].

Implementing Google Analytics is simple. After setting up an account, Google provides the code, which is then just included in every page as the last part before closing <head> tag.

For a free tool, Google Analytics provides enormous amounts of data and reports, which can be further improved with custom variables, events and tracking. It does however lack in certain areas. Google Analytics has a very easy-to-use user interface, offering some basic selection of metrics and conversion events, but more advanced advertisers might find themselves limited with Google's selection. Google Analytics data can be accessed usually after 24 hours, unlike real-time statistics from paid web analytics tools. Cookies in Google Analytics last only 30 days, for example cookies in SiteCatalyst exist 15 years. Google Analytics has also performance issues on very high traffic sites (tens of millions of views) and may stop collecting data after certain amount of hits. It supports up to 10 million views per month which should be enough for most sites, but also starts to sample data when data sets get large, meaning it will only analyze a subset of data. Google Analytics is also very strict when it comes to visitor track-

ing, anything that can be used to track individual user is prohibited. Google also owns all the data collected with Google Analytics. Google Analytics can be upgraded to premium account which fixes most of the issues with the free version for a flat annual \$150 000 fee [30].

5.4 Comparison conclusion

Choosing the right web analytics tool is highly situational. All of them provide data that can be used to improve the site as well as optimize digital marketing. Based on what we have learned from the reviewed web analytics tools, we would choose as follows:

- An enterprise-level, complex e-commerce site using WebSphere Commerce: IBM Coremetrics Web Analytics.
- An enterprise-level, complex e-commerce site not using WebSphere Commerce: Omniture SiteCatalyst.
- Smaller site, not so complex: Google Analytics.

For complex, high revenue sites that want answers for increasing return on investment (ROI), powerful and customizable analytics tools are needed, as well as a dedicated analyst who knows how to really operate them. Coremetrics Web Analytics is one of the most powerful tools, and if the store is built upon WebSphere Commerce, the choice is simple. If the store is not using WebSphere Commerce and the organization is not interested to integrate their store with it, Omniture SiteCatalyst while being a bit hard to use, is a high-performing tool in the right hands. Those who are new to web analytics and cannot take advantage of the more powerful tools, and those whose sites do not require all the power that commercial web analytics tools can provide, Google Analytics is an excellent option.

5.5 What others thing

While there are many reviews about web analytics tools, most of them are outdated since the tools continually evolve. One of the most recent reviews was published by TopTenReviews, which is a site dedicated for reviews and research [27].

TopTenReviews compared ten commercial web analytics tools with each other and gave scores to them based on features, file exporting, traffic stats, referrals, report stat intervals, events, visitor details and help/support. The featured products were: Coremetrics Web Analytics, Omniture SiteCatalyst, Webtrends Analytics, Unica Affinium NetInsight, HitsLink Platinum, VisiStat 7, OneStat platinum, Clicky Super Pro, GoStats Professional, and NextSTAT.

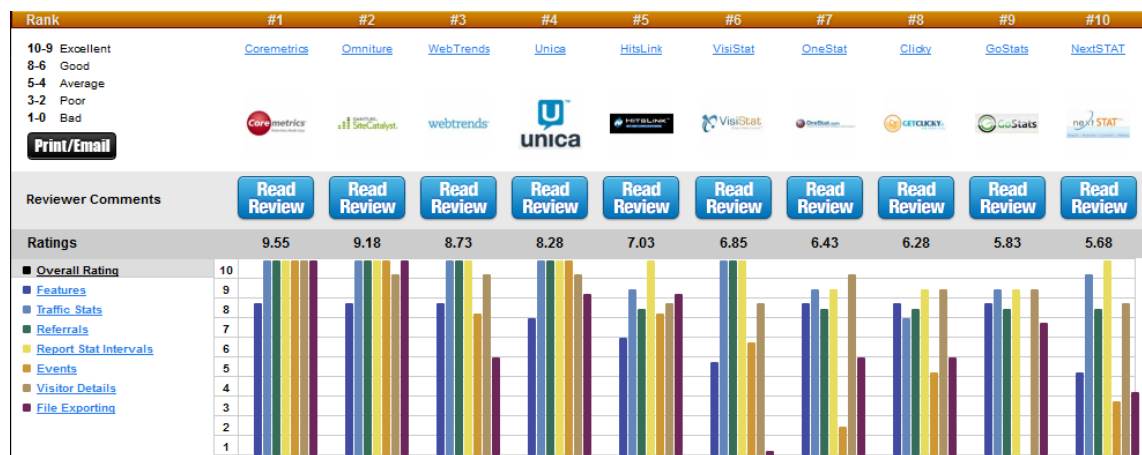


Figure 6. TopTenReviews's web analytics tools comparison final results [27].

Figure 6 shows the final results for the product comparison. The four main competitors were Coremetrics, Omniture, Unica and Webtrends, with the rest staying behind. Coremetrics scored 9.55 points out of 10. Omniture came very close, scoring 9.18 points. The third was Webtrends with 8.73 points and fourth Unica with 8.28 points. The competition between Coremetrics and Omniture was tight, but in the end Coremetrics took away the victory with its advanced visitor tracking and customer service. In the summary of Coremetrics review TopTenReviews mention that “Coremetrics has combined versatility and simplicity to create what is undoubtedly the strongest web analytics and campaign monitoring service available on the market today”.

An independent technology and marketing research company named Forrester Research released a report about web analytics in October 2011 [31]. According to the Forrester's marketing technology adoption survey, 84 percent of businesses were using web analytics at the time. It also noted that almost all vendors provide fundamental web analytics capabilities. Forrester evaluated the strengths and weaknesses of the top web analytics vendors. The chosen vendors were selected based upon criteria of companies having over \$10 million in revenues and a significant base of enterprise-

class clients. Selected top vendors were Adobe, AT Internet, comScore, Google, IBM, Webtrends and Yahoo. Figure 7 shows the results of the Forrester's evaluation.

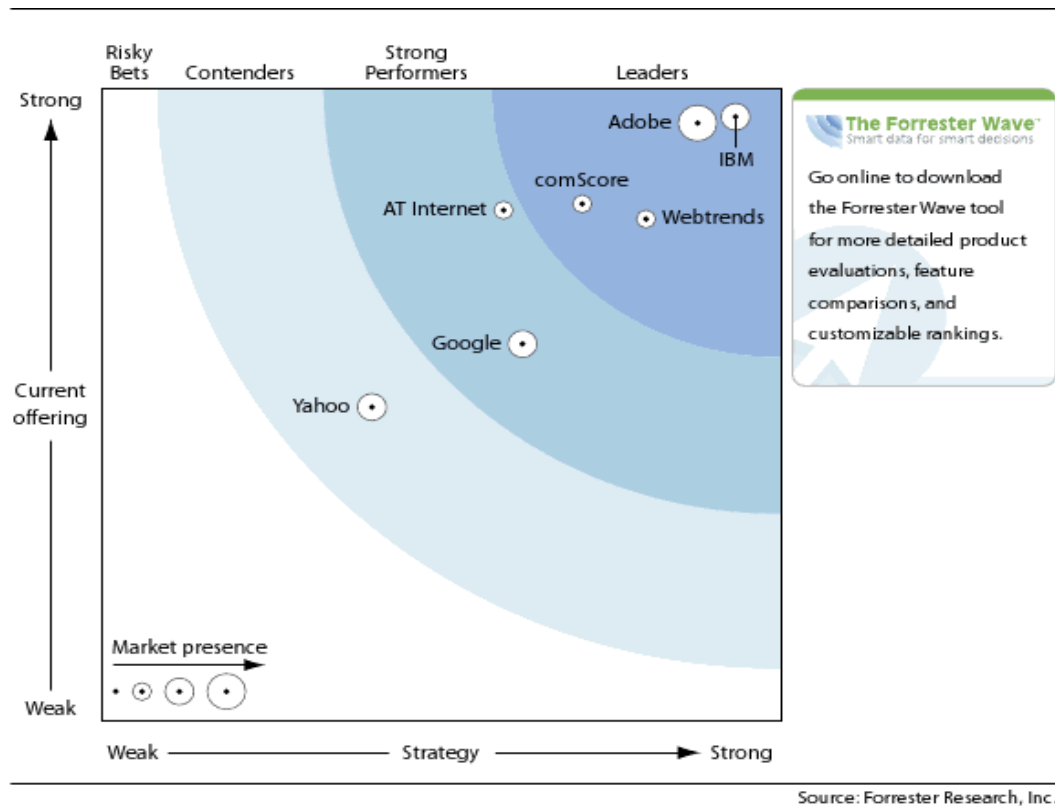


Figure 7. Forrester's evaluation of web analytics vendors [31].

As can be seen in the Figure 7, two main competitors were Adobe, vendor for Omniture SiteCatalyst, and IBM, vendor for Coremetrics Web Analytics. Vendors were evaluated against 80 criteria, which were grouped into three main categories: current offering, strategy and market presence. In the main categories IBM scored the highest with current offering and strategy, with Adobe being close behind, and in market presence Adobe took a clear win thanks to its high score in installed base.

6 Coremetrics success stories

In this section we will look into a couple case studies of how IBM Coremetrics Web Analytics helped its clients to succeed.

6.1 Boden increases trouser sales by 160 percent with IBM Coremetrics

Boden is a clothing company based in London. They sell clothing and accessories through telephone, mail order and online in the UK, US and Germany. Boden has also a shop in London.

The case study is about Boden evaluating new features hosted on its web store, and improving the overall customer experience. With the help of Coremetrics Web Analytics, two features were found that were more popular in customer initiatives than others, and they were prioritized for additional investment. These features were called Shop by Trouser Fit tool and New Arrivals section. Boden wanted to investigate how these new features had affected positively on the sales.

With the help of Coremetrics LIVE profiles Boden could see if the visitors using new features were likely to make a purchase. Boden also compared the offline and online product sales several weeks before and after the launch of the tools and was able to get a complete picture of their multi-channel sales in addition to identifying trousers that customers viewed and bought. The analysis showed that the more popular trousers were featured on the lower parts of the page, giving Boden a reason to redesign their site a bit. By placing the popular trousers higher on the page, they received more views and conversions.

The Shop by Trouser Fit tool ended up to raise trouser sales by 160 percent in the following weeks, helping Boden to sell more trousers online than through telephone or physical store, which it had never done before. Boden also learned that the conversion rate for visitors using the Shop by Trouser Fit tool was higher than those who browsed the store normally.

Boden also learned that the New Arrivals section had increased sales by 16 percent and that the visitors used the section as a base to explore products in other categories, as well as returning frequently for comparisons against new products. Visitors continued to display this behavior even after new arrivals stopped being relevant, informing Boden about the potential value of fresh content on these pages [32].

6.2 Ace Mart reduces online marketing spend by 25 percent to improve ROI

Ace Mart is a restaurant supply retailer and is committed to building and maintaining one of the best online stores. It allocated a monthly budget for online marketing, including things like paid search and online text banners. Ace Mart wanted to maximize their ROI from online marketing spend, but the tools they had available did not give any insight into how their online marketing was performing. In addition, Ace Mart received analysis from their marketing partners which only showed the click through rates, but not how they affected sales.

Ace Mart decided to implement Coremetrics feature called Marketing Management Center in order to track and analyze their ROI from online campaigns in a single interface. They were immediately able to recognize and drop underperforming campaigns, resulting in a 25 percent decrease in the monthly online marketing budget and thus increasing ROI.

While all the online marketing campaigns drove some traffic to Ace Mart's site, it was with the help of Marketing Management Center that Ace Mart was able to see which marketing campaigns resulted in sales, using the saved money to improve top performing campaigns. Coremetrics also revealed that 50 percent of the PPC keywords didn't convert to sales, and paid inclusion marketing program was not delivering enough ROI, leading to 75 percent decrease in paid inclusion marketing spent. Ace Mart also learned that the conversion rate for their online text banner advertising was only 32 percent, which made them to eliminate the program altogether [33].

7 Conclusion

The number of the Internet users has increased by 566.4 percent in twelve years, totaling currently over 2.4 billion users [10]. The Internet will only grow larger and so will the electronic business with it. With the continually improving technology, allowing access to the Internet anywhere and anytime, there will probably come a day when online store sales will overtake physical store sales. For a specific web store to be able to stand out in the sea of countless web stores some extremes measures are needed. Web analytics provide tools not only for improving the site itself but also for improving digital marketing to attract new consumers.

Even though most web sites use web analytics tools, not all utilize them right. What good is information of what happened, if there is no information of why it happened? Information can also be quite difficult to understand, especially with complex sites, and the web analytics tool should be powerful enough to capture it and then present it in as accessible format as possible.

Appealing customers has also become increasingly difficult. Customers have become digital, sharing experiences with others and having the knowledge of almost everything about the products and brands. They value trust and transparency, and companies have to adjust to that, usually by taking part in the social media. In order to win customers over, retailers have to anticipate their preferences and behavior. The best web analytics tools help to achieve this.

While the ways companies engage and convert customers will undoubtedly take new forms, we see no immediate changes in the web analytics industry. Some companies have been popular since the beginning of web analytics and will remain so in the future, even if bigger companies acquire them and their names might change. They have been continually evolving and will keep doing so. Even the “newcomer” Google Analytics is based on the software of the old and respected web analytics company Urchin.

To conclude, it is easy to understand the importance of both web analytics and digital marketing optimization to the business world which is increasingly revolving around the Internet. With a focus on the IBM Corematics Web Analytics Tool, this thesis looked at how web analytics tools work, and how to utilize them properly, and as such, will be helpful for any web business wishing to collect specific information on their customers in order to improve their performance.

References

- 1 Kaushik, Avinash. 2007. Web Analytics: An Hour a Day. Indianapolis, IN, USA: Wiley Publishing.
- 2 Wikipedia. 2012. Web counter [online document]. <http://en.wikipedia.org/wiki/Web_counter>. 8 May 2012. Accessed 19 October 2012.
- 3 Wikipedia. 2012. Web crawler [online document]. <http://en.wikipedia.org/wiki/Web_crawler>. 16 October 2012. Accessed 19 October 2012.
- 4 Wikipedia. 2012. Web analytics [online document]. <http://en.wikipedia.org/wiki/Web_analytics>. 22 October 2012. Accessed 26 October 2012.
- 5 Wikipedia. 2012. Web bug [online document]. <http://en.wikipedia.org/wiki/Web_bug>. 20 October 2012. Accessed 2 November 2012.
- 6 DIY themes. 2012. Add Google Analytics Code to Your Site [online document]. <<http://diythemes.com/thesis/rtfm/how-to-add-google-analytics>>. Accessed 8 November 2012.
- 7 Wikipedia. 2012. Packet analyzer [online document]. <http://en.wikipedia.org/wiki/Packet_analyzer>. 7 November 2012. Accessed 11 November 2012.
- 8 Wikipedia. 2012. Digital marketing [online document]. <http://en.wikipedia.org/wiki/Digital_marketing>. 13 November 2012. Accessed 14 November 2012.
- 9 Yahoo! voices. 2007. Push and Pull Digital Marketing – What You Need to Know [online document]. <<http://voices.yahoo.com/push-pull-digital-marketing-know-668615.html>>. 25 November 2007. Accessed 14 November 2012.
- 10 Internet World stats. 2012. Internet Usage Statistics [online document]. <<http://www.internetworldstats.com/stats.htm>>. 30 June 2012. Accessed 16 November 2012.
- 11 Ryan, Damian and Jones, Calving. 2012. Understanding Digital Marketing: Marketing Strategies for Engaging the Digital Generation. Philadelphia, PA, USA: Kogan Page.

- 12 Wikipedia. 2012. Meta element [online document]. <http://en.wikipedia.org/wiki/Meta_element>. 30 October 2012. Accessed 19 November 2012.
- 13 Google. 2010. Search Engine Optimization Starter Guide [online document]. <http://static.googleusercontent.com/external_content/untrusted_dlcp/www.google.com/fi/webmasters/docs/search-engine-optimization-starter-guide.pdf>. 30 September 2010. Accessed 19 November 2012.
- 14 Wikipedia. 2012. Sitemaps [online document]. <<http://en.wikipedia.org/wiki/Sitemaps>>. 3 September 2012. Accessed 21 November 2012.
- 15 Ymarketing. 2011. 5 Online Marketing Optimization Strategies to Insure Your Reputation [online document]. <<http://www.ymarketing.com/blog/bid/52534/5-Online-Marketing-Optimization-Strategies-to-Insure-Your-Reputation>>. 9 March 2011. Accessed 24 November 2012.
- 16 Wikipedia. 2012. Pay per click [online document]. <http://en.wikipedia.org/wiki/Pay_per_click>. 22 November 2012. Accessed 24 November 2012.
- 17 WordStream. 2012. Google Ads: What Are Google Ads? [online document]. <<http://www.wordstream.com/google-ads>>. Accessed 24 November 2012.
- 18 Developer Resource. 2009. What is PPC? [online document]. <<http://www.developer-resource.com/what-is-ppc.htm>>. Accessed 24 November 2012.
- 19 Wikipedia. 2012. Quality Score [online document]. <http://en.wikipedia.org/wiki/Quality_Score>. 8 August 2012. Accessed 28 November 2012.
- 20 Dan Hobin. 2012. Welcome to the Age of the Customer: The Information Age is Over [online document]. <<http://getg5.com/blog/digital-strategy/age-customer-information-age>>. 28 March 2012. Accessed 4 December 2012.
- 21 IBM. 2012. Smarter Commerce [online document]. <http://www.ibm.com/smarterplanet/global/files/us_en_us_commerce_brochure_final_new.pdf>. February 2012. Accessed 4 December 2012.

- 22 Wikipedia. 2012. IBM WebSphere Commerce [online document].
<http://en.wikipedia.org/wiki/IBM_WebSphere_Commerce>. 14 November 2012. Accessed 5 December 2012.
- 23 Client Spectrum. 2012. IBM Coremetrics [online document].
<www.clientspectrum.com/products/ibm-coremetrics>. Accessed 6 December 2012.
- 24 IBM. 2012. WebSphere Commerce Version 7 Information Center [online document].
<<http://pic.dhe.ibm.com/infocenter/wchelp/v7r0m0/topic/com.ibm.commerce.Coremetrics.doc/concepts/cmtmain.htm>>. 3 December 2012. Accessed 6 December 2012.
- 25 Bob's Blog. 2012. Why inbound marketing is becoming key to retailers [online document].
<<http://blog.balfes.net/tag/coremetrics/>>. 13 September 2012. Accessed 16 December 2012.
- 26 Steven Rose. 2012. Getting started with Adobe Omniture SiteCatalyst and Google Analytics [online document].
<<http://www.cardinalpath.com/getting-started-with-adobe-omniture-sitecatalyst-and-google-analytics>>. 29 August 2012. Accessed 21 December 2012.
- 27 TopTenReviews. 2012. Web Analytics Review 2013 [online document].
<<http://web-analytics-review.toptenreviews.com/index.html>>. Accessed 21 December 2012.
- 28 Webtrends. 2012. Implementation services [online document].
<<http://webtrends.com/solutions/services/implementation-services>>. Accessed 28 December 2012.
- 29 Google. 2012. Urchin from Google [online document].
<<http://www.google.com/urchin>>. Accessed 3 January 2013.
- 30 Google. 2013. Google Analytics Premium Feature List [online document].
<<http://www.google.com/analytics/premium/features.html>>. Accessed 6 January 2013.
- 31 Ebookbrowse. 2011. The Forrester Wave: Web Analytics, Q4 2011 [online document].
<<http://ebookbrowse.com/forrester-wave-web-analytics-q4-2011-pdf-d173322308>>. 6 October 2011. Accessed 8 January 2013.

- 32 IBM. 2012. Boden increases trouser sales by 160 percent with IBM Coremetrics [online document].
<http://www-01.ibm.com/software/success/cssdb.nsf/CS/DJOY-8QJA4N?OpenDocument&Site=default&cty=en_us>. 15 January 2012. Accessed 10 December 2012.
- 33 IBM. 2011. Ace Mart reduces online marketing spend by 25 percent to improve ROI. [online document].
<http://www-01.ibm.com/software/success/cssdb.nsf/CS/DJOY-8LZ4JA?OpenDocument&Site=default&cty=en_us>. 23 September 2011. Accessed 11 December 2012.