Chapter 1.9
A Review of Methodologies for Analyzing Websites

Danielle Booth
Pennsylvania State University, USA

Bernard J. Jansen
Pennsylvania State University, USA

ABSTRACT
This chapter is an overview of the process of Web analytics for Websites. It outlines how visitor information such as number of visitors and visit duration can be collected using log files and page tagging. This information is then combined to create meaningful key performance indicators that are tailored not only to the business goals of the company running the Website but also to the goals and content of the Website. Finally, this chapter presents several analytic tools and explains how to choose the right tool for the needs of the Website. The ultimate goal of this chapter is to provide methods for increasing revenue and customer satisfaction through careful analysis of visitor interaction with a Website.

INTRODUCTION
Web analytics is the measure of visitor behavior on a Website. However, what kind of information is available from Website visitors, and what can be learned from studying such information? By collecting various Web analytics metrics, such as number of visits, visitors, and visit duration, one can develop key performance indicators (KPIs) — a versatile analytic model that measures several metrics against each other to define visitor trends. KPIs use these dynamic numbers to get an in-depth picture of visitor behavior on a site. This information allows businesses to align their Websites’ goals with their business goals for the purpose of identifying areas of improvement, promoting popular parts of the site, testing new site functionality, and ultimately increasing revenue. This chapter covers the most common metrics, different methods for gathering metrics, how to utilize key performance indicators, best key practices, and choosing the right Web analytics tool.

The first section addresses metrics, information that can be collected from visitors on a Website. It covers types of metrics based on what kind of data is collected as well as specific metrics and how they can be utilized. The following section discusses
the two main methods for gathering visitor information -- log files and page tagging. For each method, this section covers the advantages and disadvantages, types of supported information, and examples for data format. Following this is a section on how to choose the key performance indicators (KPIs). This includes outlining several business strategies for integrating Web analytics with the rest of an organization as well as identifying the type of Website and listing several specific KPIs for each site type. The following section provides the overall process and advice for Web analytics integration, and the final section deals with what to look for when choosing analytics tools as well as a comparison of several specific tools. Finally, the conclusion discusses the future of Web analytics.

**METRICS**

In order to understand the benefits of Website analysis, one must first understand metrics – the different kinds of available user information. Although the metrics may seem basic, once collected, they can be used to analyze Web traffic and improve a Website to better meet its overall goals. According to Panalysis (http://www.panalysis.com/), an Australian Web analytics company, these metrics generally fall into one of four categories: site usage, referrers (or how visitors arrived at your site), site content analysis, and quality assurance. Table 1 shows examples of types of metrics that might be found in these categories.

Although the type and overall number of metrics varies with different analytics vendors, there is still a common set of basic metrics common to most. Table 2 outlines eight widespread types of

<table>
<thead>
<tr>
<th>Metric</th>
<th>Description</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visitor Type</td>
<td>Who is accessing the Website (returning, unique, etc.)</td>
<td>Site Usage</td>
</tr>
<tr>
<td>Visit Length</td>
<td>The total amount of time a visitor spends on the Website</td>
<td>Site Usage</td>
</tr>
<tr>
<td>Demographics and System Statistics</td>
<td>The physical location and information of the system used to access the Website</td>
<td>Site Usage</td>
</tr>
<tr>
<td>Internal Search Information</td>
<td>Information on keywords and results pages viewed using a search engine embedded in the Website</td>
<td>Site Usage</td>
</tr>
<tr>
<td>Visitor Path</td>
<td>The route a visitor uses to navigate through the Website</td>
<td>Site Content Analysis</td>
</tr>
<tr>
<td>Top Pages</td>
<td>The pages that receive the most traffic</td>
<td>Site Content Analysis</td>
</tr>
<tr>
<td>Referring URL and Keyword Analysis</td>
<td>Which sites have directed traffic to the Website and which keywords visitors are using to find the Website</td>
<td>Referrers</td>
</tr>
<tr>
<td>Errors</td>
<td>Any errors that occurred while attempting to retrieve the page</td>
<td>Quality Assurance</td>
</tr>
</tbody>
</table>
Related Content

**Evaluating Social Interaction and Support Methods over Time**
[www.igi-global.com/chapter/evaluating-social-interaction-and-support-methods-over-time/137398?camid=4v1a](www.igi-global.com/chapter/evaluating-social-interaction-and-support-methods-over-time/137398?camid=4v1a)

**WSBen: A Web Services Discovery and Composition Benchmark Toolkit**
[www.igi-global.com/chapter/wsben-web-services-discovery-composition/37660?camid=4v1a](www.igi-global.com/chapter/wsben-web-services-discovery-composition/37660?camid=4v1a)

**Probabilistic Models for Social Media Mining**
[www.igi-global.com/article/probabilistic-models-social-media-mining/52803?camid=4v1a](www.igi-global.com/article/probabilistic-models-social-media-mining/52803?camid=4v1a)

**Comparative Analysis of Ontology Ranking Algorithms**
[www.igi-global.com/article/comparative-analysis-ontology-ranking-algorithms/72992?camid=4v1a](www.igi-global.com/article/comparative-analysis-ontology-ranking-algorithms/72992?camid=4v1a)