Tutorials on Web Analytics: How to Conduct a Web Analytics Report on the Library Website

Le Yang, Metadata Librarian, Texas Tech University
Joy Marie Perrin, Digital Resources Library Librarian, Texas Tech University

Libraries, along with other information service institutions, have entered a digital era where we present resources electronically through various digital platforms. The library website is one of the most important digital platforms and functions as a community outreach tool as well as the main source of information to library patrons. Measuring the library website is a key factor in determining relative success or failure of library service. Web analytics are good tools to answer questions regarding evaluation of website performance. In this paper, we suggest step by step instructions on how to write a web analytics report to help gather, analyze, interpret, and discuss statistics of the library website. The instruction is facilitated by figures, charts, tables, and discussion of a real web analytics report on Texas Tech University Libraries website (Yang & Perrin, 2013).

Literature Review
In his article, Turner (2010) highlights that the library website is extremely important to the success of a library, because it not only provides access to resources, but also provides the users with a satisfactory and efficient service. Breeding (2008), emphasizes the performance of library website by pointing out the balance between physical facilities and virtual presence in the library has leaned more toward access through the web. As early as 2007, Rutgers digital librarian Fang points out in his website redevelopment project that as more and more digital content becomes available online, websites have become an essential component of the library service. Measuring the library website, therefore, becomes a key factor in determining relative success or failure of library service (Turner, p. 262). In addition, librarians, within the institutions, also want to demonstrate that digital resources made available through the website are constantly responding to trends and user needs (Marek, 2011a, p. 5).

Web analytics are tools to answer questions regarding evaluation of website performance. The implementation of web analytics originates from commercial website design, where they were used to track user seeking behaviors and to collect data. During the past decade, libraries have been successfully shifting to a business “return on investment” mindset, learning to justify library services and to demonstrate value of both public and private dollars (Marek, 2011a, pp. 5-6). Librarians have been using web analytics tools for evaluating web presence, digital interface,
and re-development of web projects. Among the web analytics tools being used, Google Analytics has become the one broadly favored and utilized.

Since mid-August 2006 when Google released Google Analytics to the public, many libraries have been using it to measure their library website performance and to analyze user behaviors. This is mostly because of Google Analytics’ free availability, tremendous features, and ease of use (Fang, 2007; Marek, 2011a). Unlike other web analytics tools using server log files, Google collects information by inserting simple Javascript codes into pages. The advantage of this method, as Turner (p.267) points out, is that Google Analytics can capture technical and demographic information that log files don’t normally provide such as browser, operating system, screen size, resolution, and so on. There are a few simple steps to set up before the tools can start gathering data from the website. In his book *Advanced Web Metrics with Google Analytics*, Brian Clifton (2012) provides detailed information about Google Analytics including step-by-step set up tutorials, inserting tracking codes to the websites, advanced features, and using built-in report tools. Fang (2007), Morgan (2010), and Marek (2011b) also provide tutorials with illustration about setting up Google Analytics in their articles.

How to conduct a web analytics report is another important issue that will need to be addressed. Kaushik (2007) stresses his 10/90 rule, as Marek (2011a, p. 6) quotes, that 10 percent of an organization’s cost for web analytics is for the tool itself and the setup, while 90 percent of the expense is for analyzing the data for actionable insights.

Borrowing from the e-commerce web analytics world, Turner (2010, pp.262-263), in his article, introduces Key Performance Indicators, including metrics of visits, bounce rate, conversion rate, average pages per visit, average time on site, and so on, which are defined by Google’s Conversion University, Turner (p.263) quoted, as “an essential or primary measure of performance.” However, Kaushik (2010) brings up a very good point of view, saying that the raw data only tells a portion of the story, which is “what.” He continues to state that one has to get into the mindset of the users to understand “why” and “how.”

Thus, other than conducting surveys and usability testing, the data retrieved from the web analytics presents only the factual statistics. Librarians have to find their own ways to comprehend such information through measureable indicators and interpret the data by using relevant Google Analytics metrics.

Besides studying the use behaviors, libraries can also conduct analysis on the web content itself. Some libraries choose to start the evaluation of website performance with measuring their homepages, because they believe that homepages should function as gateways for web visitors. In a blog article about a seminar of University of Oklahoma Library, Clow (2009) notes that in order to understand how the library homepage functions, their librarians conducted an analysis on the homepage using the feature “In-Page Analytics.” By doing that, their librarians were able
to generate reports outlining popular pages and density clicks based on the data “In-Page Analytics” provides.

Why Do a Web Analytics Report?
Google Analytics provides a number of built in reports, charts, and tables, that are of helps to some extent, but visualizing the data in customized formats can be more appealing and understandable to stakeholders. Terms that are used in built in Google Analytics reports can also be confusing unless you have looked up every single term carefully, thus explicit languages and straightforward interpretation make more sense to stakeholders in a web analytic report. In addition to these, library websites are usually consisted of front pages and lots of different sources linked to it, some of which might not be tracked with the main web page’s Google Analytics account.

While these obstacles can be overcome with some investigation and messing with the reports, they still present a problem when trying to express statistics to administrators or coworkers who don’t have the time to become familiar with web analytics on their own.

In order to address such issues, a well written web analytics report can summarize information, reformat data into visuals that make it easier to understand, explain the implications of the data, and provide definitions of terms used in context. The report can also provide information about what is and is not tracked so that stakeholders can become more familiar with what the stats can and cannot present.

How to Do a Web Analytics Report?
Step 1: Chose Time Period
The first part of doing a web report is deciding what time period the report should cover. In the Texas Tech example report we decided to work on a whole year’s worth of data. For academic libraries, what is useful about using a whole year’s worth of data is that it takes all the different ups-and-downs of the year into consideration so you can get a good sense of the website’s health as a whole. Picking smaller time periods can be more helpful if you need to follow a particular problem and see the change over time, or if you need more immediate feedbacks about the site’s performance. Texas Tech University Libraries also have a website committee that does a paired down report every quarter.

A whole year’s worth of data is recommended for web reports conducted the first time, because it helps establish a baseline for all statistics.

Step 2: Decide What to Report on
What is included in the report depends on what kinds of statistics stakeholders are interested in seeing. You can decide to do a report on all found statistics, or can choose specific statistics that
are representative. In Texas Tech University Library web report, for example, we decided to focus on the home page, audience, traffic source, and content statistics.

In our example report, we chose to start with the main web page and do an evaluation of the links on the front page and their usage. We then talked about the “Audience” section where we broke down demographics of country and territory, and language, and talked about audience “behavior” and technology. We also focused on Traffic Sources (such as direct traffic, referral traffic, and search engine traffic). Only then did we talk about stats about our content pages where we discussed the top 10 pages used and their specific statistics for a year. Conducting an overview report like this can help get a general sense of how the website is performing over time.

If stakeholders are more interested in assessment, then you might consider creating a service to URL document. If the organization is interested in using Google Analytics to assess how different services are doing, then the first step is to find the central URL or URLs for that service so that you can display the statistics accordingly. In this map, you can also clearly state what the stats of the URLs are actually tracking. For example, you can say that a certain group of URLs are only tracking people looking up information about a service, not people using that service. An example of a service map is included in Appendix 1.

As websites change, the URLs sometimes change, and this document is a good place to make note of old URLs and new URLs, and the effective date that the URL changed. That way, if the web analytics report is service focused, you can make sure to include old URLs if time period includes them, and can also point out if a service’s stats have been moved to a separate system.

Something to include in this document is how to search for those pages within Google Analytics. Looking at the content pages, it gives you the option to search for strings of text. If a service is associated with a particular string of text, you can include that string in the document so that you do not need to remember it every time and other people will understand how the stats can be gathered. These strings can be any words or characters that are unique to the service that appears in all the URLs associated with it. Some examples in Appendix 1 are our catalog, which is associated with “/F” and our personal librarian services, which is associated with the string “/personal librarians/”.

**Step 3: The Home Page – In-Page Analytics**

When dealing with website analytics, it is sometimes easier for administrators and coworkers to start with something familiar. Starting the web analytics report with the home page is a good way to ease in to the statistics.

The first tool you can use is In-Page Analytics on the home page. It shows the percentage of clicks that happen for any URL that is liked. The report looks like the example in Figure 1.
While it clearly shows the percentages of clicks for each link, it is hard to quickly interpret the numbers. There is an option for color coding the links, but with only three colors it’s hard to make use of it especially for website stakeholders. Moreover, numbers will not be calculated in total for those links having drop-down menus. In order to get statistics for individual links in the drop-down menu, you need to move the cursor over the links and find data for each one.

This highlights one of the major problems with Google Analytics. It does not easily let you compare different pages or links together. You can view stats for groups of pages and see their numerical values next to them, but the pages are easier to compare when you put the numbers in a chart like we have done in Figure 2.

In order to provide a direct view on comparison of different links, we used MS Excel to record both click through count and percentage for each link on the navigation bar. For those having drop-down menus, we added up the counts and recalculated the percentage for the whole group. We then generate a chart (Figure 2) having two groups of data together for a direct reflection.
As you can see in Figure 2, some links on our left navigation bar are barely used while others are used all the time. With this chart, we might consider moving some of the more heavily used links up the page and maybe even group them together so that they are easier to get to while the lower used pages might be best moved off the front page and onto a secondary page. Something to point out is that our “Document Delivery” page links out to another system for our interlibrary loan service. In this report it shows zero clicks, but this is because we are not gathering statistics for that URL with this account.

Starting the web analytics with the home page is a good way to explain what pages you can and cannot collect statistics for, and also introduce the concept. You may even want to present this information visually, or just with a listing of services and links that can’t be tracked.

**Step 4: Summarizing Specific Stats**

*Summarizing by Month to Show Change*

While Google Analytics can show data by month, it will typically only display the total visits by month in any given stats. For example, when you look at “New vs. Returning” visits, you get a
single chart with the total visits per month and then a report of the number of new visits and returning visits for the whole time. The chart provided in Google Analytics is represented in Figure 3.

![Figure 3: New vs. Returning Visits per month from Google Analytics](image)

To make this report more useful and reader friendly, you can change your date range to each month and get the numbers for each month, and put those numbers in a customized chart that now compares the new visits to the returning visits visually, as seen in Figure 4, which is a lot easier to read.

![Figure 4: New vs. Returning By month](image)

Interpreting the data, you can see that during the summer months of May, June, July, and August we have fewer new visits and more returning visits, which makes sense because we have a fewer
number of people on campus and those people tend not go to school for the first time in the summer session.

Providing some stats in customized charts like this can help stakeholders and administration understand how the statistics behave.

Representing Percentage of Data in Pie Charts & Tables

With some of the Google Analytic stats, it can be better represented as pie charts and include details in tables to facilitate the interpretation. For example, in order to show how the homepage is the “portal” for the whole website, it’s easier to see when compare the number of visits on the home page to the total of visits. In Figure 5, we see that the TTU Library’s homepage gets a large percentage of use. And one of the reasons behind this, as we know from the context, is because on the library computers, the library home page is automatically loaded.

![Figure 5: Pageviews Percentage of Homepage vs. Site Total](image)

In order to provide a more precise analysis on this, details of comparison between the homepage and the site total can be facilitated with a table, including statistics for each metrics. Figure 6 provides a direct reflection of data on different metrics between the homepage and the site total.

<table>
<thead>
<tr>
<th>Pageviews</th>
<th>Unique Pageviews</th>
<th>Avg. Time on Page</th>
<th>Avg. PG Load Time</th>
<th>Bounce Rate</th>
<th>% Exit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homepage</td>
<td>1,525,133</td>
<td>1,070,815</td>
<td>0:04:48</td>
<td>1.89 sec</td>
<td>68.62%</td>
</tr>
<tr>
<td>Site Total</td>
<td>4,078,881</td>
<td>2,890,525</td>
<td>0:03:33</td>
<td>1.56 sec</td>
<td>63.42%</td>
</tr>
</tbody>
</table>

![Figure 6: Homepage vs. Site Total](image)

For the Audience section of Google Analytics, some of the “behavior” statistics, such as Frequency, Recency, or Engagement, can be hard to compare with just numbers. For example, in Google Analytics report, the Frequency statistic provides only number of visits and is
illustrated with blue bars only without percentage shown. For the convenience of the stakeholders and administration, we gathered the data from Google Analytics and calculated the percentage, then included both of them into a table in MS Excel.

By doing this, the report about the number of people/IP addresses that visited the site and how often they showed up (and how long they stayed) is much easier to understand as it is presented with percentages and with a bar graph as shown in Figure 7.

<table>
<thead>
<tr>
<th>Count of Visits</th>
<th>Visits 1,474,821</th>
<th>Pageviews 4,078,831</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>771,148</td>
<td>1,753,950</td>
</tr>
<tr>
<td>2</td>
<td>143,791</td>
<td>406,466</td>
</tr>
<tr>
<td>3</td>
<td>61,928</td>
<td>206,350</td>
</tr>
<tr>
<td>4</td>
<td>39,370</td>
<td>137,072</td>
</tr>
<tr>
<td>5</td>
<td>28,984</td>
<td>102,849</td>
</tr>
<tr>
<td>6</td>
<td>23,277</td>
<td>82,543</td>
</tr>
<tr>
<td>7</td>
<td>33,449</td>
<td>87,930</td>
</tr>
<tr>
<td>8</td>
<td>19,733</td>
<td>63,274</td>
</tr>
<tr>
<td>9-14</td>
<td>68,444</td>
<td>244,313</td>
</tr>
<tr>
<td>15-25</td>
<td>63,056</td>
<td>232,311</td>
</tr>
<tr>
<td>26-50</td>
<td>66,288</td>
<td>243,959</td>
</tr>
<tr>
<td>51-100</td>
<td>53,958</td>
<td>207,525</td>
</tr>
<tr>
<td>101-200</td>
<td>42,049</td>
<td>147,293</td>
</tr>
<tr>
<td>201+</td>
<td>59,346</td>
<td>162,996</td>
</tr>
</tbody>
</table>

Figure 7: Frequency of Visits

**Conclusion: Presenting a Web Report**

The important thing about a web report is giving people the background for any recommendations you have. In our example report, we talked about each section and made recommendations for future usability studies or recommendations for future development of the site. We also included those recommendations at the beginning of the report in an executive summary. If they wonder why we made those recommendations, they could read that section of the report and get a good sense of the data behind the recommendation.

Most people might not want to read past the recommendations, but having the data to back up a recommendation makes a difference when people are of differing opinions about what should and should not happen on the website. What is great about a well written report is that the terms used can be explained, the data can be presented and then analyzed, and context can be added.
Reference


Appendix 1: Website Service to URL Document (Example)
The following is a sample of a Service to URL Document

- Catalog
  - http://iris.ttu.edu/F (search for “/F”)
- Federated search (Not Tracked with Main Website)- Tracked with its own Google Analytics site.
- EBooks- (Not Tracked with Main Website)
- Circulation- All of these URLs are tracked with the main library website can either represent people looking up information about services or going to form pages to use services. Circulation services that are not tracked are noted.
  - Faculty Delivery and Pickup: http://library.ttu.edu/services/circulation/fac_delivery_pickup.php
  - Loan Periods: http://library.ttu.edu/services/circulation/loan_periods.php
  - Study Carrel checkout: http://library.ttu.edu/services/circulation/carrels.php
  - Information for Appealing Library Charges: http://library.ttu.edu/services/circulation/appealing_lib_charges.php
  - Requesting the purchase of an item:
  - …
  - Room reservations: best tracked in reservation system
- Course Reserves- (Not Tracked with Main Website)
- Document Delivery/ILLiad
  - Before 20 April 25, 2012- Not tracked
  - April 25, 2012 to now- tracked by searching for “/docdel”
    - Example URL: http://library.ttu.edu/docdel/
- Personal Librarians- search for “/personal librarians/”
  - Example URL: http://library.ttu.edu/personal_librarians/accounting/articles.php
- Technology- These URLs only represent people looking up information about technology services in the library. Use of that technology is better represented with tracking programs.
  - http://library.ttu.edu/services/technology/3dlab/index.php
  - http://library.ttu.edu/services/technology/dms/index.php
  - …