intercom

HOME   EDITORIAL   **FEATURES**   SOCIETY PAGES   COLUMNS   DEPARTMENTS   RSS Feed

Improve Your Site with Simple Web Statistics

BY SUSAN YESHIN

In an ideal world, information architects would be able to watch over the shoulders of customers using our websites to validate design choices we made and make adjustments as needed. Often, we don't have the luxury of a user trial or the funds to purchase a sophisticated site analysis tool. Fortunately, you likely already have some valuable information regarding your site and its visitors available to you. Most Web servers collect some basic information, or Web statistics, on the pages they host and serve; servers store this information in a log file. If you can access the server log file, you can read it with one of many free log file readers, like AWStats (*http://awstats.sourceforge.net*), which will display the log file information in a meaningful, human-readable way. With some careful analysis, you can use the information provided to validate your design choices and to improve your visitors' experience.

Because Web statistics tools simply provide a visual representation of the server activity for your site, they are not a substitute for real feedback. You need to make some educated assumptions and track trends over time to avoid jumping to incorrect conclusions. In this article, I will walk through the information that you can get from many log file readers. For people new to Web statistics, you'll see how this information can help you improve your site. For those who have experience with Web statistics, perhaps you'll learn some new tricks or caveats.

Historical data

How many people have visited the site? This is what most people think of first when you talk about Web statistics. Some people use this information as proof that a website is a good one, but remember that the number of unique visitors to your site might not really be very accurate. There are a lot of things that might skew this number, such as internal users and testers, proxy servers, shared IP addresses, and browser cache. There are also people who landed on your site when looking for something else.

For these reasons, it's best to compare the number of visitors over time and use the data to validate changes. For example, if you recently submitted the site to search engines or optimized your site for search, an increase in visitors will validate your work. If you recently launched a new version of a product and new site, you can watch the number of visits to your new site increase as customers switch to the latest version. As your older site is visited less often, you can prioritize the new content.

The historical data also tells you what days and hours that visitors have been going to your site. Using this information, you can test the speed of your site in peak use hours to ensure that the pages are loading in an acceptable amount of time.

User data

Who is visiting your site and where are they? This information is extremely useful and I'd suggest that you use a statistics reporting tool that will give you these details. Knowing the top countries viewing your site can give you an idea of markets that you might not have considered. It can also help you make smart decisions about translating your site. Tracking the countries your visitors are from over time can also validate efforts of the marketing team.

As an information architect working on product documentation, I'm always eager for feedback but may not always have the opportunity to speak with customers directly due to funding and time constraints. Web statistics often contain the IP addresses of the people accessing your site. Using a free online tool, I can resolve the IP addresses and find out which companies are visiting my site most often. When you use this technique to discover who your key stakeholders are, you can arrange to contact them, or you can look for industry patterns in the top twenty visitors and target your content to these industries. For example, if your top

ARCHIVES

2012
January 2012
2011
2010
Earlier Issues

MENU

Home
Advertise
About Intercom
Subscribe
Email Intercom
Write for Intercom
Intercom Awards
Reprint Permissions

TABLE OF CONTENTS

customers are in the aerospace industry, consider creating scenario or sample information that uses the aerospace domain as its context.

Robots and spiders

Many programs will identify the robots and spiders that have crawled your site. These programs collect the information on your content that search engines use to give your site weight in the results. Some search engines allow you to register your site and to provide keywords. After you do search engine optimization (SEO) work on your site, use this data to ensure that your registration was successful and that your site is being crawled.

Visits and duration

This information is interesting, but is it useful? You can probably make some assumptions based on this information, but remember that they are only assumptions. If a percentage of people visited your site for fewer than thirty seconds, there's a good chance that those people did not mean to visit your site. However, if your site has great keywords and provides task-level content that can be scanned quickly, thirty seconds might be all that a customer needs to get the information needed to continue with a task.

If customers spend several minutes on your site, it could mean that they are intently reading your content, or it could mean that they didn't find what they were looking for but left your site open in the browser.

File types

The types of files accessed on your site could be misleading. Often it means that a file was loaded, but not necessarily that it was viewed. For example, if you added Flash animations to your tutorials, you might be excited to see that a good number of these files were accessed and assume that your tutorials are popular. But it only means that the page that contains the file was opened, as the animation plays automatically each time the page is accessed. You could use this statistic to validate a navigation change to highlight the tutorials, but you could also use page views for that.

The most and least popular content

What pages are people visiting, and which ones are they not? This is where your statistics get really interesting, especially if you have a large website. It's especially easy to make assumptions when looking at these numbers so make sure you consider all factors before making decisions about the future of your files.

If you provide different types of information on your site, can you see patterns with your most popular files? Are people using your site for an overview, to take a tutorial or complete a course, to complete a task, or are they there trying to troubleshoot? Knowing why people are visiting your site will help you organize your navigation to better satisfy the goals of your customers.

If your Web statistics don't provide you with the path customers took through your site, you can compare the number of hits on multi-page tutorials or tasks to see if your customers followed through or if they only visited the first page or two. If customers are not following through, consider chunking your information differently so the key information is on the pages that are read.

If your site covers a large number of topics, you can use the statistics to prioritize areas of work. Remember that statistics are more valuable if you track them over time. While a large number of hits on a troubleshooting or error message file can signal that there is room for improvement in your product, a small number of hits on a file can signal that the information is unused and not necessary, or it could signal that the topic is not easily found. Before using the statistics as a justification for removing information, ensure that the topics contain the appropriate keywords, appear in the navigation, and are included in the related links of other files.

If the most popular page viewed on your site is the home page, have you optimized it so customers can easily find important information that you want to highlight? Are the pages that your home page links to more popular than other pages? If customers arrive on your site from a search result, will they understand where they have landed, and have you provided them with a good reason to stay on your site? You might want to ensure that you have the name of your site and a link to the home page readily available on every page, as well as related links to overview content.

Use the numbers to validate changes to your navigation or even the value of the navigation in general. Are customers finding new content or do you have to do a better job of highlighting it? If you recently reorganized your navigation but there was no change to the most and least popular files, compare those files to the topics

POPULAR ARTICLES

SUBSCRIBE

Subscribe to **RSS** or enter your email to receive newsletters for news, articles, and updates about what's new.

for which people are searching. If search is how customers are finding your information, consider making it easier for them by optimizing your content and providing a powerful internal search option.

Operating systems and browsers

The information relating to the operating systems and browsers used by your viewers can be used for testing, but will also help you make a decision about whether or not it is time to focus on optimizing your site for small-screen devices.

Referrers

It's interesting to see who provides a link to your site. However, it is more useful to think about who is not pointing to your site, but could. Adding a link to your information from a heavily used site like Wikipedia could significantly increase traffic to your site, and Wikipedia's search results can rank higher than a direct link to your site even when you enter a very direct search phrase. If you look at the search results for the topics for which you expect your customers to search, consider whether you could contact the site owners of the top ten results and have a link added to your site.

Keywords and phrases

The keywords and phrases are the search terms that people entered that generated a link to your site. Keep in mind that they might not have been looking for your site or the type of information you provide. Use the keywords to see trends in search behavior. If people are searching for information on a certain technology, for example, perhaps you can prioritize that technology. Also consider the pages to which the search engines are linking. If it is better for your users to land on your home page or overview pages rather than task-assistance pages, you need to add more keywords and optimize those pages for search.

Look at the terminology used and ensure that you have keywords that match the words that people are actually using to conduct their searches. Consider also updating your index entries with alternate terminology to help ensure that your visitors find the topics they need.

You can try to make some assumptions based on the types of information for which people are searching. For example, if they are searching for error message IDs from a software product, you might see a pattern with the most common error messages. Using this information, you can work on improving the user interface so customers don't end up with the error, or improve the in-product error message with better recovery information so already frustrated users don't have to leave the product and search for help on the Web.

404 Errors

There's not much to say about the list of 404 errors that most Web statistics programs provide to you. You might find that some of the errors listed are bogus, caused by the robots crawling pages that customers will never see, but all of the results should be examined and you should ensure that all valid errors are fixed.

As an information architect, you might want to use Web statistics to identify potential issues with the structure of the content, but there is a lot more value to be gained. Consider how the numbers can be useful to all roles on a development team. Editors will want to look at the search keywords to ensure that the information is indexed properly and that the writers are using terminology that is consistent with that of the customers. Documentation team leads can take cues from the statistics to help them create content plans and focus on the most used types of documentation. Product development and documentation managers can assign resources to the key technologies for which your customers are searching. The test team can update their test plans and test the browsers and versions that customers are actually using. Writers can improve their documentation by fixing errors, adding keywords and related links, and can better prioritize their work. Usability teams can focus on areas of the product that customers are having problems using.

Remember not to let your imagination provide false context around the numbers, and to keep track of your results over time before making generalizations. Analyzing the basic Web statistics are a great start to better understanding the habits of your customers and validating the improvements that you make to your site.

Susan Yeshin *is an information architect at IBM. She has 15 years of experience working with software documentation for a variety of products, and a special interest in metrics and measurable improvement.*

Leave a Reply

Name (required)

Mail (will not be published) (required)

Intercom

Website

CATEGORIES

Editorial
Features
Society Pages
Columns
Departments

SOCIAL MEDIA

RSS Feeds
Twitter
LinkedIn
YouTube