



Virginia Commonwealth University
VCU Scholars Compass

VCU Libraries Faculty and Staff Publications

VCU Libraries

2010

Beyond Description: Converting Web Site Usage Statistics into Concrete Site Improvement Ideas


Julie Arendt

Virginia Commonwealth University, jaarendt@vcu.edu

Cassie Wagner

Southern Illinois University Carbondale, cwagner@lib.siu.edu

Follow this and additional works at: http://scholarscompass.vcu.edu/libraries_pubs

 Part of the [Graphics and Human Computer Interfaces Commons](#), and the [Library and Information Science Commons](#)

Copyright Julie Arendt and Cassie Wagner

Recommended Citation

Arendt, J. & Wagner, C. (2010). Beyond description: Converting Web site usage statistics into concrete site improvement ideas. *Journal of Web Librarianship*, 4(1), 37-54. doi: 10.1080/19322900903547414

This Article is brought to you for free and open access by the VCU Libraries at VCU Scholars Compass. It has been accepted for inclusion in VCU Libraries Faculty and Staff Publications by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.

Beyond Description:

Converting Web Site Usage Statistics into Concrete Site Improvement Ideas

Abstract

Web site usage statistics are a widely-used tool for Web site development, but libraries are still learning how to use them successfully. This case study summarizes how a Morris Library at Southern Illinois University Carbondale implemented Google Analytics on its Web site and used the reports to inform a site redesign. As the main campus library at a research university with approximately twenty thousand undergraduate and graduate students, the library site included resources from multiple library departments on a single site. In planning the redesign, Morris Library's Virtual Library Group combined usage reports with information from other sources such as usability tests and user comments. The Virtual Library Group faced barriers to interpreting and applying the usage statistics in the site redesign, including some that were specific to the library's implementation of the Google Analytics tool and some that were limitations inherent in any Web usage statistics. Some key barriers in applying the usage statistics to a redesign included the need to sift through data that did not have implications for the site redesign, the need to interpret the implications of usage numbers for the site redesign, and the need to balance competing interests within the library. Nevertheless, the usage statistics enabled the Virtual Library Group to make better decisions by providing a source of factual information about the site use rather than relying on staff members' opinions and conjectures.

Keywords: usage statistics, Web site design, academic libraries, Web site development, Web analytics, Google Analytics, user-centered design, Web site evaluation

Introduction

Like design standards, heuristics, focus groups, and usability tests, Web site usage statistics can be used to aid developers in improving Web sites. Site usage statistics provide information such as what pages users view, how users get to a site, and what links are followed within a site. They also track information such as a user's browser, operating system, and the Internet Protocol (IP) addresses of users. Usage statistics have some advantages over other site evaluation tools because they monitor how users actually work with a site, rather than what they say they would do. In a formal testing setting, such as a usability test, users may change their answers to what they think will please the observer or rationalize their behavior (Nielsen 2001). There are also questions about how accurately the tasks set for usability testing simulate the way users approach research on the Web (Huntington and Nicholas 2006). Because usage statistics only present information about what users do, they do not provide explanations of why users do particular things or how that information could be used to improve a Web site (Wiggins 2007).

Different methods of user research can be applied at the points when they are needed in an iterative process of site development and user research (Kuniavsky 2003). For example, contextual inquiry or focus groups might occur early in the development process to determine user needs. Based on those needs, prototypes would be developed, tested with usability tests, refined based on the usability tests, and retested. Once initial site development is completed, user feedback would be used to gain information about what parts of the site are problematic for some users, and usage statistics would be used to examine if the site is being used as intended.

Guidance about how to apply site usage statistics, especially for businesses, is plentiful in Web development literature. However, the literature on applying usage statistics specifically to

library Web site improvement is sparse compared to other tools. One possible reason that library Web development literature emphasizes other methods is that the guidance for applying Web site usage statistics often has limited applicability in a library context. For example, commercial Web analytics methods apply site usage statistics to provide actionable performance indicators in a business context to achieve business goals (Burby and Atchison 2007, Sen et al. 2006, Waisberg and Kaushik 2009). A positive performance indicator for a business might be more time spent on the site, a larger proportion of users making a purchase, or a decreased bounce rate (Waisberg and Kaushik 2009). A library, on the other hand, may want to make it easier to for users to access library resources hosted outside the library's site, which would cause users to leave the site more quickly. Some work has been done to employ site usage statistics in libraries to measure conversion rates of site visitors performing desired activities such as renewing a book online or suggesting materials for purchase, but this kind of measurement is complicated because many library resources reside outside the library Web site server such as in the library's integrated library system or sites owned by database vendors (Whang 2007). Case studies that examine the basic usage statistics of library Web sites without attempting to develop performance indicators are more readily available.

Literature Review

Many, but not all (e.g. Fang 2007, Whang 2007), case studies of Web site usage statistics on library Web sites provide descriptive information about how a library's site is used but do not reach beyond the description of user behavior to provide guidance for site improvements. A few reports in the library literature offer suggestions of ways that usage statistics could be used to improve a library's Web site.

Basic Reports

The most basic usage report is the number of visitors to the site. Usage statistics may undercount or overcount the absolute number of visitors because of technical details such as page caching and visits by Web crawlers (Dowling 2001, Cohen 2003a). Even though the absolute numbers of visitors and page views are too problematic to be accepted at face value (Dowling 2001, Cohen 2003a), usage statistics provide a broad overview of a site's users, what they do, and how site use changes over time. Basic reports also can summarize the percentage of visitors from on campus or off-campus IP addresses, the time of year the site is visited most often, and the browsers used (e.g. Asunka et al. 2009, Black 2009, and Coombs 2005). One limitation of this basic information is that it confirms expectations about a site rather than providing indications of where changes should be made. When the basic information violates expectations, it could have implications for Web site design or services. For example, if a residential college found that the majority of its library's Web site visits came from off-campus IP addresses, this information would be surprising. Further examination of what parts of the site off-campus users accessed would be needed to deduce why the site had so many off-campus users and possibly what these off-campus users' needs were.

Basic site usage statistics reports also can reveal the technical shortcomings of a site. Data such as the hit rate, time to serve documents, and peak hits can be used to measure server load and to justify server upgrades (Cohen 2003b). Similarly, aborted file transfers can be examined to determine if there is a way to improve the transfers (Mariner 2002). Error messages, such as "file not found" errors, can be examined to correct the related problems (Dowling 2001, Mariner 2002, Cohen 2003b, Breeding 2005). The browsers, features, and tools (e.g. JavaScript, screen resolution, connection speed) that users employ can help a Web designer determine which

browsers should be used for site testing and which advanced features would be appropriate for users' technology (Dowling 2001, Mariner 2002, Cohen 2003b, Fang 2007).

Most Popular Content

Site usage statistics reports typically list which pages on a site are visited most frequently. Frequently-visited pages are often prominent on the site's home page. This pattern follows the “law of surfing,” which predicts that users will not navigate deeply into a site to find information (Huntington et al. 2007). High numbers of visits to pages deep in a site may indicate that users want easy access to these pages and that navigation to these pages should be streamlined. To do this, it may be necessary to drop links to rarely-used resources from the home page, to rearrange the layout to put heavily used items on top, or to cluster resources differently (Ghaphery 2005, Fang 2007). Not everyone agrees with the interpretation that frequently-used pages should be given the most prominence. If a rarely-used page is important, the low use could be interpreted as a sign that the page needs greater prominence on the site (Cohen 2003b).

Navigation Summary

Navigation summaries in site usage statistics reports provide an overview of how users navigate through a site, including how they arrive at a page, what they click on a particular page, and how often they exit the site from each page. The listing of the most-frequently exited pages can be examined to determine if they lead to off-site resources or if users leave for other reasons, such as poor design (Guenther 2001). Navigation summaries also list the most frequent landing pages, pages at which users enter the site. If a top landing page is not easy to access from other parts of the site, it could suggest that the site needs improvement to make it easier to navigate to

the page (Guenther 2001). Fine-grained navigation details can track users' paths to examine how users move from page to page within the site and whether they navigate the site as expected (Breeding 2005).

Keywords

Site usage statistics also typically provide information about the search engines and keywords that lead users to the site. For sites with an internal site search feature, some site usage statistics also include reports on the keywords that were used in the site search. Search logs for site searches can be used to infer what visitors are trying to find (Breeding 2005, Guenther 2001). These lists of keywords can be used to ensure the words that are important to users are incorporated into the site to optimize search results. Improving the sitemap (Black 2009) and incorporating keywords into the content (Cohen 2003b) and the metadata (Cohen 2003b, Guenther 2001) are some ways to ensure this.

Context & Methods

Morris Library at Southern Illinois University Carbondale is the main campus library and serves a research university with approximately twenty thousand undergraduate and graduate students. Large portions of the library building were closed from mid-2005 to early 2009 because the library underwent a major renovation. During that time the library kept over seventy computers for patron use and maintained some basic services, but much of the library collection was stored off site. The library's Web site's last major redesign was completed in fall 2005. In that redesign, the library implemented Plone, an open source content management system. It integrated at least three different page layouts and color schemes into a single format. With the

2005 design, the site gained a site-wide navigation scheme. Almost all pages on the site maintained a set of links, shown in Figure 1, that were repeated from the home page. These site-wide links gave the site a more unified look and feel.

[PLACE FIGURE 1 HERE]

Despite all of the improvements included in the 2005 redesign, the site design, like all site designs, could still be improved. Some of the improvements, such as expansion of the Special Collections section of the site, were made incrementally over the next few years. Some improvements, such as a more visually appealing template for the entire site, could not be made piecemeal. In addition, in the years after 2005, the library developed new services. Although links to the new services were added to the site where they seemed appropriate, their positioning, given the site's existing architecture, was not always ideal. For example, the library added group study rooms during the building renovation, but the information about reserving a group study room was placed deep in the site in a section for library policies because the existing architecture did not provide a better position for it.

In spring 2007, Morris Library's Web Development Librarian began leading the Virtual Library Group in discussions to plan for a major redesign of the library's Web site. The Virtual Library Group, which included staff members from various library departments, existed to facilitate the multi-department collaboration required for the library's online projects. Virtual Library Group worked to ensure that the new site would fit the needs of both internal and external stakeholders.

In preparing for this redesign, the Virtual Library Group reviewed comments and complaints from library staff and patrons about the existing site. These included concerns that the existing site template was visually unappealing, that the font was too small, and that some

library departments were hard to find. The Virtual Library Group also conducted usability tests using paper prototypes of different possible designs. The library also installed Google Analytics (Google 2009a) on its Web site to collect usage statistics about its Web site. All of these sources of information contributed to decisions that were made about the new site design.

During the previous redesign in 2005, the library queried the data from its server logs to produce a Web site usage report. Because of obstacles specific to Morris Library's Web hosting and maintenance structure, installation of server-side usage statistics software could not easily be done, so the library instead produced the report in-house through scripted log queries. That report showed the times of day that the site received the most traffic and showed that IP addresses within the library were the single largest source of visits to the library's Web site. The amount of staff effort required to produce this report was large compared to the amount of information it provided that could be applied to the redesign. Members of the Virtual Library Group hoped that Google Analytics would provide more useful information with less staff effort than the previous report and that using this information could lead to improvements in the site.

Google Analytics is a service that tracks Web site usage data and produce site usage statistics and reports. Currently Google Analytics is free for any site that has up to five million page views per month. It uses a cookie in conjunction with JavaScript to collect data as users navigate through a Web site (Google 2009c). The data is hosted and processed on Google's computers, and the reports are accessible via the library's account on the Google Analytics Web site. This remote storage of user data raises concerns for some libraries. At Morris Library, the limited information about users that was collected by Google to provide this service was accepted.

Other possible concerns that some libraries may have regarding Google Analytics include sections of its terms of service (Google 2009b) that permit Google to change the fees, the service itself, or the terms of service. In particular, the terms of service (Google 2009b) leave it up to the user to periodically check the terms of service for changes and state that the new terms may be binding if, “You continue to use the Service after Google has posted updates to the Agreement or to any policy governing the Service.” At Morris Library, it was decided that the terms of service available at the time Google Analytics code and thus far during its use have been acceptable.

Google Analytics code was added to the library’s Web site template, so it automatically populated all of Morris Library’s Web site pages. Google Analytics was installed in a matter of minutes. Cost, including the amount of staff time needed to install and manage the tool was the driving factor for choosing Google Analytics.

Because a new design was already in development, the committee decided not to spend the limited time of the Web Development Librarian on extensive site alterations to facilitate data collection by Google Analytics. For example, the default installation of Google Analytics only tracked clicks on links that led to other pages with the library’s site. To track links to off-site resources such as library databases or the library catalog, changes to each of those outbound links would need to be made on the library’s site. The Web Development Librarian instead worked on usability testing and on developing the new site design. Even without site modifications, such as those to track links that took users away from the library's site, Google Analytics produced clearer and more detailed reports than previous server log queries that had taken hours to develop and generate.

Although the Virtual Library Group used multiple methods to gather information for the site redesign, this paper emphasizes the specific information collected from the Web site usage

statistics and how that information was applied. For ease of discussion, the results presented in this report came from one year, beginning May 1, 2008 and ending April 30, 2009.

Findings

Basic Reports

The basic reports included information about the number of site visitors and about when the site was accessed. From May 1, 2008 to April 30, 2009, Google Analytics recorded 986,755 visits to the library's Web site.

A great deal of the information provided by the basic reports was unsurprising and had limited utility for the site redesign. For example, the site was used every hour of the day, but use was much lower from midnight to eight AM than during the rest of the day. The use was lower during the summer and during academic breaks than when classes were in session. However, two pieces of information from the basic reports were useful for the redesign. First, more than 95% of visitors had 1024 x 768 or better screen resolution. This information suggested that page did not need to be designed for lower resolutions, as it had been in the 2005 redesign.

Another illuminating result from the basic reports was that the site had a high bounce rate. In 61% of the visits, the user viewed only one page of the site. Two factors may have contributed to this phenomenon. First, the library's home page opened by default when patrons started a browser on a library computer, so it was possible that a large portion of the bounces came from patrons who were using library computers but not library resources. Second, because the home page had a direct link to the library catalog, it was possible that a large portion of the users went directly to the catalog. Google Analytics reports were produced that excluded visits

where the user immediately bounced from the home page. In that report, the number of visits dropped from over 986,000 to approximately 452,000, and the percentage of visitors with on-campus IP addresses dropped from 71% to approximately 56%. This information implied that many of the bounces may have come from in-library users. Because filters to separate in-library IPs from non-library IPs and code to track direct clicks to the catalog were not part of the library's installation of Google Analytics, it was not possible to know exactly how much each source contributed to the high bounce rate.

Most popular content

The usage statistics reports that provided a ranked list of how frequently pages on the site were viewed could be used in two ways: to identify pages that were heavily used that may need greater prominence on the site and to identify prominent pages that were not heavily used.

Google Analytics provided two separate reports of the most frequently-visited pages, a “Top Content” list and a “Content by Title” listing. The “Top Content” report provided a rank-order list of the most-frequently visited URLs on the library’s site. The “Content by Title” report provided a rank-order list of the most-frequently visited page titles on the library’s site. The library’s content management system operated in such a way that a single page on the site might have more than one URL, but it was rare that two different pages on the site had the same page title. The Virtual Library Group relied on the “Content by Title” report to determine which pages were visited most frequently. The pages that were linked directly from the home page had titles that were identical to their link labels, so it was easy to identify which links led to heavily used pages.

The library's home page was by far the most-visited page, even after bounces from the home page were excluded from the count. Generally the other heavily-used pages on the site

were ones that had direct site-wide navigation links to them. This information, while affirming decisions made in the previous design, was not helpful for suggesting changes to the home page. At best, it supported decisions already made by the Virtual Library Group about what parts of the site to focus on for usability testing.

In one instance, high use of a page contradicted Virtual Library Group members' expectations, but it did not lead to a change in the site design. The sixth most-visited page was "Online Courses," which led to the Blackboard learning management system. Because Blackboard was not a traditional library function, some members of the Virtual Library Group believed it did not belong on the library's home page or in its site-wide navigation links. Blackboard's high ranking informed the decision to keep a link to Blackboard at least on the library's home page.

Several pages that received prominence on the library's site did not rank in the twenty most-used pages. This information did not always result in a decision to make the pages less prominent. Removing the links to "Distance Learning" and to "Library Downloads" from the library's home page could be justified because neither was among the twenty most-viewed pages and neither led to a section of the site intended for large groups of library users. Decisions to remove some of the other less-viewed links from the home page or the site-wide navigation were more difficult. "Sitemap" appeared at number fifty-five of the most-viewed pages, but Web design recommendations (e.g. Nielsen 2008) encouraged sites to make a sitemap readily available. "Faculty Resources" and "Giving" were two other pages that had links on the home page and had low usage. Removing either of these links from the home page was blocked by resistance from within the library. Multiple library departments and groups, including Reserves, Instructional Support Services, and liaison librarians provided content for faculty through the

“Faculty Resources” page. They wanted to make these resources easy for faculty to find. Similarly, the dean of the library wanted it to be simple for potential donors to locate information about giving.

Several parts of the site that did not have direct site-wide navigation links were among the twenty most-viewed pages. Despite their relatively heavy use, many of these pages were rejected for inclusion as direct links. The Virtual Library Group assumed that “Staff Directory” and “Library Accounting,” the eighteenth and nineteenth most visited pages, respectively, received their high rankings from heavy use by library staff and were not especially important for patrons. “Graduate Assistantship Opportunities,” a page that listed graduate assistantships vacancies in the library, was the sixteenth most-viewed page. The Virtual Library Group did not believe it was necessary to add a direct link to this page because the library only occasionally had graduate assistantships available. However, because the employment section of the site was heavily used, the Virtual Library Group included a task on its usability tests to see if the employment pages were easy to locate.

On the list of frequently visited pages one page stood out as possibly needing greater prominence on the site. Despite not having a direct link from the library's home page, the Special Collections Research Center was among the twenty most-viewed sections of the site. Even before the collection of site usage statistics, the Special Collections Research Center staff lobbied the Virtual Library Group to add a direct link to their section of the site. Its high use bolstered the case that many visitors wanted easy access to this part of the site. The Virtual Library Group tentatively decided not to include a link to the Special Collections Research Center on the home page at this time, but planned to revisit the issue going forward.

Navigation Summary

The Google Analytics report that listed the top landing pages where users entered the site did not provide much additional information. Even after bounces from the library's home page were excluded, the home page still was the landing page for three fourths of the site visits. The next most frequent landing pages, "Databases / Find Articles" and the "eJournal Finder," each accounted for less than 5% of the arrivals to the site. Other pages in the list of top the top ten landing pages included frequently visited pages such as "Online Courses" and "Books." The top ten list of landing pages also included pages that the Virtual Library Group did not think were of long-term importance to most library patrons. An obituary of a library faculty member and a library accounting page for library staff were two of the top ten landing pages, but each accounted for less than 1% of the arrivals to the site.

The listing of site exit pages similarly did not yield ideas for site improvements. Most of the top ten exit pages were pages intended to guide users to resources outside the library's Web site, such as databases and electronic journals.

The "Navigation Summary" report in Google Analytics made it possible to examine the paths that users followed to enter a page, to navigate from that page to another on the site, and to exit the site. Because the Virtual Library did not have the resources to examine the "Navigation Summary" for every page, it conducted this fine-grained navigation analysis for just a few parts of the site. A couple sections of the site that library staff thought users may have found confusing received this fine-grained analysis. Reference and instruction librarians believed that students had difficulty distinguishing between the intent of the "Databases / Find Articles" link and the "eJournal Finder" link. Additional justification for focusing on these pages for fine-grained analysis was that these pages were frequently visited and were among the top landing pages.

As intended, users generally navigated away from the site after clicking on “Databases / Find Articles.” The “eJournal Finder” had a clear pattern of failure. If users followed the intended path, they would have exited the site directly from the “eJournal Finder” page to go to a specific electronic journal or would have continued on to the “Citation Linker” page and left the site to access a specific article in an electronic journal. Of users who entered the “eJournal Finder,” 50% exited the library's site from that page, as anticipated. A substantial fraction, 18%, of the “eJournal Finder” users clicked on the “Databases / Find Articles” link from the “eJournal Finder” page. The Virtual Library Group interpreted this information to mean that the functions of the “eJournal Finder” and of the “Databases / Find Articles” sections of the site were frequently confused. Less than 4% of the users who entered the “eJournal Finder” clicked on “Citation Linker.” The Virtual Library Group interpreted this information to mean that the path to Citation Linker was not clear. The Virtual Library Group decided that the layout of the “eJournal Finder” page, shown in Figure 2, made it difficult for users to notice the “Citation Linker.” The Virtual Library Group decided to add better cues make the “eJournal Finder's” purpose clearer and to make the “Citation Linker” more obvious.

[PLACE FIGURE 2 HERE]

Fine-grained analysis was not always so informative. Although “About the Library” was part of the site-wide navigation scheme, members of the Virtual Library Group had not expected it to be a heavily used section of the site. It, however, was the eighth most-viewed page. Examining the “Navigation Summary” report for an explanation of “About the Library's” frequent use was not revealing. Given that the page consisted of links to other pages on the library's site, it was not surprising that most (93%) of the visits to “About the Library” continued on to another page within the site. There was no single path that users generally followed after

they accessed the “About the Library” page. The staff directory, the most-used link from this page, only received 22% of the clicks from the “About the Library” page. The detailed navigation examination simply offered no clear direction for links on the “About the Library” page that might warrant greater prominence on the site.

Keywords

The Google Analytics report also provided lists of keywords that users entered into search engines to find the site. This report did not include the site search within the library’s site, but it did include keywords from site searches within the university’s main site. Unfortunately, the keywords from the university site search were reported as Google searches and could not be distinguished from searches made using a Google search from google.com.

Variations of the library's name dominated the top rankings of keywords that led users to the site. The rest of the report therefore provided useful information for only a few keywords. Over 56% of the search engine traffic to the site came from variations on the library's name. Nine of the top ten keywords were variations on the library's name, as were fourteen of the twenty most-used keywords. Furthermore, by the bottom of the top twenty list, the keywords seemed to be in a "long tail" of the keywords that each brought relatively few visitors. For example "dental" was the eighteenth most-frequent keyword leading to the site. This keyword brought 687 visits, or less than one half of one percent of all visits initiated by Web searches. Nevertheless, useful information for the redesign could still be gleaned from the report on keywords.

"ISS" was the fifteenth most common keyword. ISS was the abbreviation for Instructional Support Services, the unit of the library that administered Blackboard and performed many other instructional support functions for faculty. As shown in Figure 3,

Instructional Support Services could be found on the library's Web site by navigating through the "For Faculty" page. Because "ISS" was a popular keyword for accessing the site and the "For Faculty" section of the library's site received little traffic, the usage statistics report suggested that replacing the "For Faculty" link with an Instructional Support Services link would be a reasonable change to the site. This idea was unsatisfactory for librarians because traditional library services for faculty would be buried in the site by comparison. Another possible design would be to have two links for faculty, one for Instructional Support Services and another for traditional library services. Because the redesign priority was to emphasize undergraduates, this possibility seemed unsatisfactory. The Virtual Library Group decided to keep Instructional Support Services under a "For Faculty" link but to give greater prominence to Instructional Support Services on the "For Faculty" page.

[PLACE FIGURE 3 HERE]

The absence of some keywords from the top twenty list also was useful. As noted previously, Special Collections was not directly linked from the library's home page, but it received a relatively large number of visitors. Keywords relevant to Special Collections, such as "Special Collections," "Special Collections Research Center," or "SCRC" brought few visitors. The keyword "special collections" ranked 115th, bringing just 109 visitors to the site. If Special Collections was not easily found by navigating through the library's site and visitors generally were not using search engines to look for "special collections," how were they finding it? Keywords describing holdings in Special Collections' holdings such as "Sherman Theatre," "Julia Carter Ingersoll," "Everett W. Hall," "Anna Cemetery," "Marjorie Lawrence," "freedom of the press," "Mordecai Gorelik blacklist," and "Camp Dubois' Anna" all were among the hundred most-used keywords and were used more frequently than "special collections." The low ranking

of "Special Collections" weakened the case that Special Collections should have a direct link in the site-wide navigation links, at least not one labeled "Special Collections."

Lessons Learned & Discussion

The information the Virtual Library Group gleaned from the Google Analytics reports at times strengthened the case for particular changes to the site, at times weakened the case and at times presented contradictory information about the site usage. For Morris Library, Google Analytics was easy to install and produced easy-to-read reports about site usage. Unlike the library's previous experience with collecting usage statistics from server logs, this collection of usage statistics required little time or money. Although Morris Library had its own particular reasons that usage statistics were not very useful in 2005, other libraries may have experienced similar challenges in compiling, interpreting, and using this data. Free and inexpensive tools released in the past few years, such as Google Analytics, have made it easier to produce usage reports.

The usage statistics reports provided useful information that assisted in Morris Library's redesign. For members of the library's Virtual Library Group, evidence from the site usage helped to persuade resistant members of the group of the need or lack of need for particular changes in the site. It provided facts to assist in the decision making process, rather than relying on staff members' opinions and conjecture alone. Knowing that Blackboard received heavy use from the library's home page reduced the debate about whether to keep the link or to drop it. Observing the navigation patterns in the "eJournal Finder" informed decisions about how to redesign that page. Knowing the most used sections of the site aided the priority-setting for

usability testing. The usage statistics data helped the Virtual Library Group to make informed decisions.

The factual information from the Google Analytics reports sometimes made the course of action obvious to the Virtual Library Group, but often the implications of the usage statistics were unclear. Although the numbers seem concrete, the analysis required leaps of inference that may or may not have been justified. For example, the Virtual Library Group interpreted the relatively high level of search traffic for Instructional Support Services to mean that the site navigation to Instructional Support Services was inadequate. Another reasonable interpretation was that faculty did not think of Instructional Support Services as part of the library, so they searched for this department rather than going to the library's site to find it. Similarly, the Virtual Library Group interpreted 4% to be a low percentage of visitors who went from the "eJournal Finder" to the "Citation Linker" and that the source of this low number was poor visibility of the "Citation Linker" on the "eJournal Finder" page. Another possible interpretation was that only a tiny proportion of visitors had a need for the "Citation Linker." If the Virtual Library Group had accepted these other interpretations, the implications for site design would have been quite different. Sometimes the usage statistics simply did not provide enough information to definitively answer some questions about the site. Similarly, the usage statistics showed conflicting patterns regarding Special Collections. The discussion between the Special Collections representative in the Virtual Library Group and the rest of the group was probably more important in the decision about where links to Special Collections should go than the information from the use. Professional judgment and internal politics also played parts in the decision making process.

Sometimes internal needs were more powerful than the usage data. At times, the Virtual Library Group faced a conflict between the competing goals of making it easier for users to do what they already did and of guiding users toward resources that members of the Virtual Library Group felt were underused. Analyzing usage statistics did not resolve all controversies about how the Web site ought to be arranged, but it supported certain design decisions more than others.

Some of the limitations of the Web site usage statistics were inherent to the statistics themselves. For example, the raw hit numbers to the Morris Library's site were not a good indicator of how much the site was used to access library resources because of the use of the library as a computer lab. Although that example was particular to the Morris Library, other libraries likely would have their own issues and peculiarities that would limit their abilities to take parts of their usage statistics at face value. Even if the usage statistics provided completely accurate numbers about site usage, they still would be limited in what they could measure. Usage statistics could not differentiate among the different groups of visitors: students, faculty, staff, and others, that were important to the library. Although usage statistics could track the proportion of on-campus and off-campus users, they could not indicate how many of the users were affiliated with the university. The usage statistics also could only report what site visitors did; they could not explain why users followed particular paths through the site. Although the Virtual Library Group made some inferences about the reasons users followed particular paths, the statistics by no means guaranteed that the Virtual Library Group made correct inferences.

Another limitation of the usage statistics was the enormous quantity of information provided in the Google Analytics reports. Google Analytics created easy-to-understand reports much more easily than those generated by in-house staff using raw server logs, but it still took

some effort to identify which pieces of information would be useful in the site redesign. The reports provided a lot of information that was interesting but not useful. Knowing that use decreased sharply during academic breaks or that the library's name was the keyword that led the most people to the site did not help the Virtual Library Group make decisions for site improvements. To sift through all these reports, the Virtual Library Group used simple rules of thumb such as looking for things in the reports that members found surprising, important, or confusing. It might have been possible to pull other useful information by examining the reports more thoroughly, but it would have taken a lot more time and effort to sift through the usage reports and pull out the pieces of information that would be useful for a site redesign.

Other limitations in the usage statistics were caused by the Virtual Library Group's choice to use the most basic installation of the Google Analytics tool. Morris Library's implementation of Google Analytics did not track clicks to off-site resources. For example, it would have been useful to have tracked clicks to external content because so much of the library's Web site served as a portal to vendor resources. It would have been useful to know which of those resources were sought most often, so the site could have given greater priority to those links. Although this information would have been useful, the library made a tradeoff between time spent on gathering and analyzing the analytics data and time spent on other projects for the library.

One of the reasons that the Virtual Library Group chose not to take much time to customize the site's usage statistics data collection was to leave more staff time to gather other types of data about the site. The Virtual Library Group decided to combine usage statistics with results from paper prototype usability tests, feedback from users, and staff discussions to lead to a series of recommendations for site changes. Each method yields different types of information

about how users interact with the Web site. All of the sources complemented each other. For example, reference librarians' concerns about the clarity of the "eJournal Finder" and of "Databases/Find Articles" led closer examination of the usage statistics and usability tests of these parts of the site. These, in turn, confirmed some of the librarians' concerns and led to recommendations for changes to the labeling and layout of those pages.

Because of the limitations and the possibility of a misinterpretation, it would be dangerous to base all design decisions on site usage statistics. Complementary data sources are necessary to corroborate or challenge the conclusions that might be drawn from usage statistics alone. This process is especially useful when approaching major design changes, such as a switch from an organizational structure based on library departments to one oriented toward specific research tasks.

Web site usage statistics served as a helpful but limited tool for the site redesign. For example, the usage statistics allowed the Virtual Library Group to examine whether a part of the site that reference librarians thought users found confusing was being used as intended. The statistics demonstrated that parts of the site, such as Instructional Support Services and Blackboard, which librarians may not have considered important, received fairly heavy use. The usage statistics provided facts to inform their decisions about site changes, but the facts were subject to interpretation. The usage statistics could not answer every question that the Virtual Library Group would have liked them to answer. Sometimes information gleaned from other sources could be combined with the usage statistics to answer the questions. Other times, the Virtual Library Group still had to rely on its members' judgments to make decisions for the site redesign.

References

- Asunka, Stephen, Chae Hui Soo, Brian Hughes, and Gary Natriello. 2009. Understanding academic information seeking habits through analysis of Web server log files: The case of the Teachers College Library Website. *Journal of Academic Librarianship* 35(1): 33-45.
- Black, Elizabeth L. 2009. Web analytics: A picture of the academic library Web site user. *Journal of Web Librarianship* 3(1): 3-14.
- Breeding, Marshall. 2005. Analyzing Web server logs to improve a site's usage. *Computers in Libraries* 25(9): 26-29.
- Burby, Jason, and Shane Atchison. 2007. *Actionable Web analytics: Using data to make smart business decisions*. Indianapolis, IN: Wiley.
- Cohen, Laura B. 2003a. A two-tiered model for analyzing library Website usage statistics, part 1: Web server logs. *portal: Libraries & the Academy* 3(2): 315-26.
- . 2003b. A two-tiered model for analyzing library Website usage statistics, part 2: Log file analysis. *portal: Libraries & the Academy* 3(3): 517-26.
- Coombs, Karen A. 2005. Using Web server logs to track users through the electronic forest. *Computers in Libraries* 25(1): 16-20.
- Dowling, Thomas. 2001. Lies, damned lies, and Web logs, in "Net Connect," special issue *Library Journal* 126(7): 34-35.
- Fang, Wei. 2007. Using Google Analytics for improving library Website content and design: A case study. *Library Philosophy and Practice* 9(3): 1-17.
- Ghaphery, Jimmy. 2005. Too quick? Log analysis of quick links from an academic library Website. *OCLC Systems & Services* 21(3): 148-55.

- Google. 2009a. Google Analytics. <http://www.google.com/analytics/> (accessed July 1, 2009).
- . 2009b. Google Analytics terms of service. <http://www.google.com/analytics/tos.html> (accessed November 24, 2009).
- . 2009c. How does Google Analytics work? <http://www.google.com/support/googleanalytics/bin/answer.py?hl=en&answer=55539> (accessed November 23, 2009).
- Guenther, Kim. 2001. Know thy remote users. *Computers in Libraries* 21(4): 52.
- Huntington, Paul, and David Nicholas. 2006. Improving the relevance of web menus using search logs: A BBCi case study. *Aslib Proceedings: New Information Perspectives* 58(1/2): 118-128.
- Huntington, Paul, David Nicholas, and Hamid R. Jamali. 2007. Site navigation and its impact on the content viewed by the virtual scholar: A deep log analysis. *Journal of Information Science* 33(5): 598-610.
- Kuniavsky, Mike. 2003. *Observing the user experience: A practitioner's guide*. San Francisco: Morgan Kaufmann Publishers.
- Mariner, Vince. 2002. Logging usability, in "Net Connect," special issue *Library Journal* 127(1): 30-31.
- Nielsen, Jakob. 2001. First rule of usability? Don't listen to users. Jakob Nielsen's Alertbox. <http://www.useit.com/alertbox/20010805.html> (accessed October 16, 2009).
- . 2008. Site map usability. Jakob Nielsen's Alertbox. <http://www.useit.com/alertbox/sitemaps.html> (accessed October 2, 2009).
- Sen, Arun, Peter A. Dacin, and Christos Pattichis. 2006. Current trends in Web data analysis. *Communications of the ACM* 49(11): 85-91.

- Waisberg, Daniel, and Avinash Kaushik. 2009. Web analytics 2.0: Empowering customer centricity. *The Original Search Engine Marketing Journal* 2(1), http://www.semj.org/documents/webanalytics2.0_SEMJvol2.pdf (accessed December 1, 2009).
- Wang, Michael. 2007. Measuring the success of the academic library Website using banner advertisements and Web conversion rates: A case study. *Journal of Web Librarianship* 1(1): 93-108.
- Wiggins, Andrea. 2007. Data-driven design: Using Web analytics to validate heuristics. *Bulletin of the American Society for Information Science and Technology* 33(5): 20-24.

Morris Library

at Southern Illinois University Carbondale



small text | normal text | large text

- [home](#)
 - [about the library](#)
 - [event space](#)
 - [giving](#)
 - [hours](#)
 - [help](#)
 - [contact us](#)
 - [sitemap](#)
- you are here: [home](#) » [ask-a-librarian](#)

quick links

SIUCat search

- [My Account](#)
- [Books \(Print & Online\)](#)
- [Databases / Find Articles](#)
- [eJournal Finder](#)
- [Other Library Materials](#)
- [Course Reserves](#)
- [Request Storage Materials](#)
- [Interlibrary Loan](#)
- [Distance Learning](#)
- [Online Courses](#)
- [Faculty Resources](#)
- [Downloads](#)



questions?



Ask-A-Librarian

Comments: [Morris Library Web Administrator](#)

Copyright © 2007, Board of Trustees, Southern Illinois University Carbondale

[Privacy Policy](#) | [EOE Policy](#)

Last Updated Sunday, November 22, 2009

Figure 1: Several links, shown here, were repeated on virtually every page of the site.

questions?

 Ask-A-Librarian

- quick links
- SIUCat search
- My Account
 - Books (Print & Online)
 - Databases / Find Articles
 - eJournal Finder
 - Other Library Materials
 - Course Reserves
 - Request Storage Materials
 - Interlibrary Loan
 - Distance Learning
 - Online Courses
 - Faculty Resources
 - Downloads

eJournal Finder

Use [Citation Linker](#) if you are searching for a specific article or citation.

Find e-Journal

Find by Title Find by Subject Advanced Search

0-9 A B C D E F G H I J K L M N O P Q R S T U V W X Y Z Others

Title: Starts with Contains

Powered by SFX

Figure 2: The “Citation Linker” could be accessed from the “eJournal Finder” page.

Faculty Resources — Morris Library

[Interlibrary Loan](#)
Information about the services provided by Interlibrary Loan and the forms to request materials not available at SIUC

Make Materials Available to Students

[Blackboard](#)
Information about the Blackboard course management system, support, and training

[Reserves and Electronic Reserves](#)
Information about placing print and non-print items on reserve to ensure students enrolled in faculty members' classes have ready access to the materials. Faculty may request print materials be placed on reserve using our convenient online form, and learn how to submit non-print items to the Reserves Department.

Services for Faculty in the Classroom

[Instructional Support Services](#)
ISS provides a wide variety of technology services for faculty to enhance teaching and learning

[Instructional Technology](#)
Instructional Technology provides classroom technology support for Lawson Hall and auditoriums and classrooms across campus

[Liaison Program](#)
Liaison librarians who are specialists in your subject area can teach research skills and library usage to your students.

[Turnitin Plagiarism Prevention System](#)
Turnitin (turn-it-in) allows instructors and students to check research papers for originality.

Services for Students in the Library

[Instruction](#)
Information about workshops and seminars that teach library and technology skills

[Information Services](#)
Information Services includes Circulation and Interlibrary Loan, Geographic Information Systems (GIS), Government Information (Documents), Maps, Periodicals and Special Formats, Reference (Information Desk) and Reserves

Figure 3: Instructional Support Services was one of the services listed on the “For Faculty” page.