Combining Web Analytics and a Web-Based User Survey for LibGuides Assessment

Carmen Howard, MSLS*[a], Jung Mi Scoulas, PhD[b], Allan Berry, MSLIS, MA[c], Deborah L. Lauseng, AMLS[d]

[a]Senior Instructor and Regional Health Sciences Librarian, Library of the Health Sciences – Peoria, University of Illinois Chicago, Peoria, Illinois, https://www.orcid.org/0000-0002-3460-5417, choward4@uic.edu

[b]Associate Professor and Assessment Coordinator, University Library, University of Illinois Chicago, Chicago, Illinois, https://www.orcid.org/0000-0002-7536-295X, jscoul2@uic.edu

[c]Clinical Assistant Professor and Web Services Librarian, University Library, University of Illinois Chicago, Chicago, Illinois, https://www.orcid.org/0000-0002-1012-8601, aberry3@uic.edu

[d]Associate Professor and Regional Head Librarian, Library of the Health Sciences – Peoria, University of Illinois Chicago, Peoria, Illinois, https://www.orcid.org/0000-0001-9067-2932, dlauseng@uic.edu

Cite as: Howard C, Scoulas JM, Berry A, and Lauseng DL. Combining Web Analytics and a Web-Based User Survey for LibGuides Assessment. Hypothesis. 2024;36(2). doi:10.18060/27669

Abstract

Objective: The aim of this method is to assess an online library guide, using both qualitative and quantitative methods, with the specific objective of discovering who is using the guide and why, as well as understanding the impact of the guide on user learning outcomes.

Approach: This method combines the use of web/data analytics (quantitative) and a web-based survey (quantitative and qualitative) with an existing model for measuring learning outcomes. By combining these methods, a more robust assessment of an online guide can be achieved.

Data Type Used: The discussed online guide assessment used data from Google Analytics and from a web-based survey, created in Qualtrics, with learning outcome questions developed using the Project Outcome for Academic Libraries Toolkit.

Strengths: The combined analysis of the web analytics and survey data resulted in an expanded assessment of the library guide and a richer picture of guide users, their behavior, and their learning outcomes.

Limitations: Individually, each method used in this research has limitations. Web analytics are unable to determine users' motivation (why) for using the guide and whether they have learned (impact) from their behavior/interaction while using the guide. Surveys collect self-reported data and may be subject to recall bias. However, when the methods are combined, each helps to counteract the individual limitations.

Overview

Library guides, a type of library-created online resource that is frequently built using Springshare's LibGuides, are labor-intensive to maintain over time. For this reason, many libraries seek to assess their guides' use and effectiveness. Most of these assessments focus on quantitative data, whereas a few other studies have used qualitative methods, and very few have used a combination of methods.[1,2] While each method has its own strengths, Griffin and Taylor highlighted the need to incorporate qualitative data with quantitative usage data and advocated for "using multiple data sources and triangulating findings between data points".[1]

Analytics data provides a picture of actual user behavior through tracking and reporting traffic patterns with detailed aggregate statistics. Because it tracks every interaction, analytics data eliminates recall bias. However, interaction tracking is not always as complete and informative as might be expected. Although web analytics are objective in how data is gathered, the choice of what data to gather is not. Tracking movement on a website also cannot reveal user motivations, if they found what they were looking for, or whether they learned anything from the interaction.

Surveys provide a structured format for collecting information on various topics of interest, including demographics, behaviors, and attitudes. They have the potential to reach a range of patrons and ask them questions to better understand their motivations, perceptions, and experiences. Although they have many advantages, surveys also come with limitations. For

example, they rely on self-reported data provided by individuals who voluntarily participate, which may introduce biases.[3] Survey fatigue can occur due to the large volume of survey requests, potentially impacting response rates. In addition, the extensive nature of some surveys with multiple questions can discourage participants from dedicating the time necessary to complete them.

Our research team sought to assess the University of Illinois Chicago (UIC) Library's Evidence Based Medicine (EBM) LibGuides pages (our guide) and plan for the future of our guide using data (IRB Protocol #2019-0449). Given higher-than-anticipated usage statistics and multiple external requests to reuse content, our research team speculated that our guide was being used more broadly than initially intended. We aimed to understand user behavior, both what they did and why they did it, and to measure the impact of our guide on user learning experience and outcomes.[4] Additionally, we sought to provide evidence for transitioning our guide to a fully Open Educational Resource (OER).[4] Therefore, we developed an assessment approach employing two complementary methods used concurrently: quantitative usage data and a web-based user survey.

Our research team initially reviewed the usage statistics integrated into LibGuides; however, we found the data limited to high-level analysis, such as page views, monthly trends of use, etc. Given our interest in a deeper understanding of users and their use patterns, we determined that Google Analytics (GA) would provide a more robust data set. We also realized that our questions could not be fully answered by analytics and that we would need to obtain users' input through a survey. Our interest in measuring educational outcomes led us to the Project Outcome for Academic Libraries Toolkit, hereafter referred to as the Project Outcome Model.[5] The Project Outcome Model is a free online toolkit developed to measure the impact of library programs and services. We adapted questions from the Project Outcome Model into our survey to gain insights into user knowledge and confidence.

The combination of two methods, web/data analytics and a web-based survey, provides a more robust approach in the assessment of online guides. This combination gave our research team a richer understanding of our users, their behavior, and their perceptions of learning outcomes. For librarians interested in expanding their evaluation of library guides or other webpages, this article focuses on our process and the strengths and limitations of each method and tool we used, as well as their combination. A summary of those strengths and limitations can be found in Table 1 following the discussion. For a brief review of the literature on evaluating library guides and a discussion of our research findings, please see our team's previously published work.[4]

Example with Discussion

UIC's EBM (Evidence Based Medicine) Guide (https://researchguides.uic.edu/ebm) is a subject guide that introduces the basic concepts in EBM with pages such as "Levels of Evidence" and "Asking Clinical Questions." Additionally, our guide covers higher level concepts like "Clinical Filters" and "Appraisal." Our guide was initially created for students in the UIC College of Medicine and for non-UIC students in an online course for librarians. However, integrated tracking statistics indicated that it was being used by far more people than this limited pool of intended students. Our research team needed to analyze user behavior to determine who was using our guide and why, so we looked at what methods

would assist us in this specific need. We also found that some pages were used at significantly higher rates, which led us to develop survey questions about individual page use and to use analytics data to look specifically at page traffic.

At the beginning of our research project, we considered a few different methods and tools. As the project progressed, we discovered several strengths and limitations of the methods and tools that we chose. See Table 1 for a summary.

Quantitative Method - Web/Data Analytics

Our research team chose to look at web analytics data because we believed it could answer questions about how our users interacted with our guide - the pages they visited, how long they stayed, and how they were interacting with the pages. We also knew that it could inform us about the devices used, how users found our guide, and their general geographic location. Analytics data can be obtained from various tools. Our research team considered and/or used the following tools described below.

Springshare's integrated "Guide Tracking"

Our research team initially used the integrated statistics from Springshare's LibGuides primarily because this is the software used to develop our guide. While this was helpful in seeing counts of views per page and monthly trends in overall use, the data is only available at a high level. Although Springshare does have additional tracking features, we needed a more comprehensive, granular, look at how the site was used.

Google Analytics (GA)

Our research team decided to move to GA because GA was currently being used by the library and is widely regarded as an industry standard.[6,7] It is a clear market leader[8,9] and considered a required skill for entry-level marketing analytics professionals.[10] GA met our need to see how users interacted with our guide. During our research, we used the latest version available of GA called "Universal Analytics" (UA). Google has subsequently released a major update to the GA system, named "Google Analytics 4" (GA4).

GA provides rich data for user behavior analysis. We were specifically interested in the following quantitative information from GA:

- how users were finding our guide
- how they moved between guide pages
- how they used specific pages (most-visited pages and average time on a page)
- their general geographic location, such as country and city (we did not collect specific location data, such as street or ip addresses)
- the type of device they were using to access the page.

The first issue we encountered involved URLs and GA reporting. We configured GA to only return data about our guide itself. Our guide consists of 10 total URLs: three unique URLs for the home page and one unique URL for each of seven additional pages. We elected to remove URLs with third party parameters for simplicity; for example, Facebook or Google Translate, would sometimes add proprietary URL parameters like "translate.googleusercontent.com/translate_c?depth=1&hl=ar&prev=search&rurl=translate.google.com&sl=en&sp=nmt4&u=" to the beginning of our URLs and for their own purposes of user tracking. Accounting for these additional URLs significantly increased the complexity of our report data. Most of these modified URLs, while presumably representing legitimate user requests, were each only accessed during a single session. Therefore, we explicitly excluded these proprietary URLs by filtering out all traffic except to our expected 10 URLs. These 10 URLs covered 99.6% of all unique users from the full GA pool during the research time-period. No excluded URL represented more than 0.05% of our guide's unique users. The added simplicity seemed worth the small cost in comprehensiveness.

Additionally, our research team learned the importance of understanding the way that GA data points are calculated. Initially, we anticipated using "bounce rate" and "session duration" to inform our understanding of user behaviors. However, upon further investigation into our guide's specific data, all UIC guides data, and a deeper review of GA processing, we determined that these metrics did not reveal much about the use of educational subject guides.

When a user visits a single page without further interaction, GA considers this a "bounce."[11] Our guide had a large proportion of "bounces" (86% of sessions). For example, consider a student completing a course assignment. They might spend 10 minutes using a guide page, afterward closing their browser without further navigation. Most librarians would call this a success. However, since "session duration" is calculated by time-stamping each action that a user takes (for example, clicking a link), GA counts this as a "bounce", with a "session duration" of zero seconds.[11, 12] Sessions can also end automatically after an arbitrary time expires without any actions.[12]

Google documentation acknowledges that a high bounce rate is only problematic if you are trying to direct people to another page,[11] such as when your home page is only a waypoint on the way to a product page. Thus, while "bounces" were considered an important metric in UA, they were later de-emphasized in GA4 in favor of other engagement metrics. In the case of our guide, even the home page is an educational page, not meant to direct users to subsequent pages, so high "bounce rates" were not a concern for us. For these reasons, our research team found "average time on page" to be more informative in evaluating the success of our guide than either "session duration" or "bounce rate."

As we learned more about GA, we belatedly realized additional features could have been activated prior to our research study, which may have provided even deeper insights. One example is "scroll depth," or how far a given user scrolls down a page, which can be a proxy for user engagement. However, tracking data for "scroll depth" in UA required activation before data collection; at the outset of our project, we did not consider it relevant. Later, it seemed problematic to alter our data collection midstream. GA4 provides this feature automatically, which should provide additional rich data for future research.

Quantitative & Qualitative Method - Web Survey

While GA offers insights into general user patterns, it does not provide information on user motivation and reasoning or assess the impact of their usage on learning outcomes. So, in addition to the GA data, our research team developed a survey to gather information from users regarding their usage of our guide and the extent to which they achieved their learning objectives. In the survey, we collected various information, including demographics, user behavior, future referral intentions, user perceptions, satisfaction levels, and user learning outcomes in terms of knowledge and confidence. Multiple survey tools and models are available. Our research team adapted the Project Outcome Model to assess learning outcomes by using a survey administered through Qualtrics software.

Project Outcome Model

To measure users' learning outcomes, we utilized two questions from the Project Outcome Model.[5] This commonly used model is based on a series of self-assessment learning statements and uses a Likert scale. We asked questions related to the user's change in knowledge and confidence. For example, we adapted this question: "After using the EBM Guide, I feel more confident about my ability in each of the following areas."[4,5] The Project Outcome Model allowed us to gain insights into users' learning without asking them to take a pre- and post-test of their knowledge. However, we acknowledge that survey responses were based on self-perception and that the question formatting used in the Project Outcome Model has been noted[4] to have the potential of acquiescence bias, or the "tendency for the interviewee to agree with the questioner."[13]

Qualtrics

Qualtrics is a web survey platform used to create and distribute surveys. Additional web survey tools are noted in Table 1. Our research team chose Qualtrics as our survey tool due to its many useful features and its availability at our institution. It enables researchers to gather, analyze, and report data effectively, with features to enhance data security.[14] However, it is important to note that there are financial costs associated with using advanced features and conducting surveys with larger sample sizes. This may pose accessibility and affordability challenges for survey designers operating on limited budgets. Additionally, certain features of Qualtrics may have a learning curve, particularly when it comes to coding scales and setting up surveys as originally intended. This learning process is crucial for ensuring accurate data analysis and utilizing the platform to its fullest potential.

Table 1. Strengths and limitations of methods and tools

Product (Platform)	**Analytics**		**Survey**	
	Springshare LibGuides Analytics	**Google Analytics**	**Project Outcome Model**	**Qualtrics (Platform)**
Research questions	This data did not answer our research questions but provided an overview of guide usage.	1. Who are the users? (demographics such as geographic location and device used.) 2. How are users interacting with our guide?	1. What learning outcomes are achieved from the users' perspectives?	1. Who are the users? (detailed demographics such as educational level and profession.) 2. What are the users' motivations?
Method: Strengths	1. Comprehensive data 2. Actual use patterns/behaviors 3. Automated data after set-up 4. Eliminates recall bias		1. Qualitative data 2. Motivational data 3. Question structure flexibility allowing integration of existing questions from the Project Outcome Model or other models5.	
Tool: Strengths	1. Easy to use 2. Native to the webtool/publishing platform 3. Ability to compare with our other guides	1. Cost-effective 2. Detailed statistics regarding traffic patterns and use 3. Industry standard	1. Provides existing questions related to learning outcomes	1. Comprehensive options 2. Analysis and reporting features 3. Enhanced data security features
Method: Limitations	1. Unable to understand users' motivation (why) and whether they have learned (impact) 2. Unable to collect certain demographic info (i.e. - educational level) 3. Possible user privacy concerns		1. Self-reported data which may be skewed due to self-selected participation 2. May be subject to recall bias 3. Survey fatigue	
Tool: Limitations	1. Limited data points	1. Data analysis can be complicated with proprietary URLs 2. Data point calculations (for example, "bounce rate") may not be intuitive	1. May be subject to acquiescence bias	1. Financial costs to access advanced features 2. May have learning curve for sophisticated features
Methods: Potential alternatives	1. Quantitative Survey		1. Focus Groups 2. Semi-structured Interviews 3. Usability Testing	

Tools: Potential alternatives		1. Adobe Analytics 2. Amplitude 3. Cloudflare Web Analytics 4. Matomo Analytics 5. Mixpanel 6. Other options are available	1. Several models are available depending on your research topic	1. Email distribution 2. Google forms 3. LibWizard 4. Polling 5. Survey Monkey 6. Other options are available

Conclusions

Our research team found that the use of both quantitative and qualitative data provided a richer picture of our users, their behavior, and their learning outcomes. We suggest that GA data and survey data each have their own strengths. When combined, these methods are complementary; each helps to counteract the limitations inherent in the other. By combining methods to include both web/data analytics and a web-based survey that incorporates the Project Outcome Model, we were able to assess LibGuides pages more robustly and develop an expanded methodological assessment approach.

The strength of the combined method can be seen in a variety of ways. First, GA kept the survey shorter. Because GA could easily answer some of our questions, we did not have to ask users about their navigation of the site or about the devices they used. Keeping surveys shorter helps to address survey fatigue.

Second, an anonymous web-based survey can provide insights to many complex research questions like understanding user motivation and outcomes. For example, although GA could tell us if a user was coming to our site from an educational site, like a course management system, GA cannot provide further insight into educational uses or outcomes. Using the Project Outcome Model to measure changes in user confidence and their perception of increased knowledge has provided a clearer understanding of user learning. The learning outcome data, which showed an increase in self-reported knowledge and confidence, was used to justify expanded work on our guide. For detailed results and a discussion of our findings, please see our previously published article.[4]

Finally, the combination of the two methods removed the uncertainty of some GA data points and increased our confidence overall that the survey participants were representative of all users. Our research team included questions in the survey to look for data agreement or lack of agreement with data gathered by GA. For example, in terms of geographic distribution, a brief examination of GA data in preparation for our research revealed that most users were not coming from the Chicago area. This could be an indication that they might not be UIC affiliates, although many UIC students and faculty do attend and teach courses remotely. This data point led us to include a question on the survey about UIC affiliation. When we ran the survey and looked at concurrent GA data, we confirmed that most users were not associated with UIC. Another example, where we included a question to look for data agreement but did not find agreement, was in comparing how users came to our guide. GA reported much higher

levels of discovery by an organic search than users self-reported in the survey. Although it was extra work and our results were somewhat mixed, we felt this effort to look for data agreement was worth the investment for our overall increased confidence in the results.[4]

Our assessment was enhanced by carefully examining the strengths and limitations of each method and tool used to address our research questions. Further, our assessment was strengthened through a combination of methods and tools. The most important characteristic in this type of assessment is the combination of quantitative data with a qualitative component. Our research team proposes that this combined method could be used by others to evaluate not just LibGuides pages but other types of web resources. Alternative tools can be used in implementing this proposed method. For example, if your institution/organization uses primarily Google Forms to gather feedback from your users, this tool could be used instead of Qualtrics. If you want to explore challenges or behaviors for non-users or frequent users in depth, conducting focus groups could be more appropriate. Librarians and researchers simply need to examine the purpose of their project and the most appropriate quantitative and qualitative methods and tools available to address their questions.

Resources

Web Content

- Project Outcome for Academic Librarians Toolkit: https://acrl.projectoutcome.org/about
- Google. Analytics Help. Glossary (GA4): https://support.google.com/analytics/topic/9355633?hl=en&ref_topic=9143232&sjid=-95027 3908224713767-NA
- The Ultimate Google Analytics Glossary (2023 edition): https://www.lovesdata.com/blog/google-analytics-glossary

Further Reading

- Connaway LS, Radford ML. Survey research and the questionnaire. In: *Research methods in library and information science*. 7th ed. Santa Barbara, Calif.: Libraries Unlimited; 2021. p. 129-168.
- Farney T. Getting the best Google Analytics data for your library. Libr Technol Rep. [Internet] 2016;52(7):5-8. Available from: https://doi.org/10.5860/ltr.52n7
- Watson CA. Can an understanding of data use in academic libraries assist in public service efforts? Public Serv Q. [Internet] 2021;17(1):59-64. Available from: https://doi.org/10.1080/15228959.2020.1860861

CRediT

Carmen Howard: Conceptualization, Methodology, Investigation, Resources, Writing - Original Draft, Writing - Review & Editing **Jung Mi Scoulas:** Conceptualization, Methodology, Formal analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing **Allan Berry:** Conceptualization, Methodology, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing **Deborah L. Lauseng:** Conceptualization, Methodology, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing

References

1. Griffin M, Taylor TI. Employing analytics to guide a data-driven review of LibGuides. J Web Librariansh. [Internet] 2018 [cited 2023 Jul 26] 12(3):147-159. Available from: https://doi.org/10.1080/19322909.2018.1487191

2. Jackson R, Stacy-Bates KK. The Enduring landscape of online subject research guides. Ref & User Serv Q. [Internet] 2016 Spring [cited 2024 Jan 23] 55(3): 219-225. Available from: https://doi.org/10.5860/rusq.55n3.219

3.McNabb DE. Response error. In: Nonsampling Error in Social Surveys. SAGE Publications, Inc.; 2014:114-134. Available from: https://doi.org/10.4135/9781483352923

4. Lauseng DL, Howard C, Scoulas JM, Berry A. Assessing online library guide use and open educational resource (OER) potential: an evidence-based decision-making approach. J Web Librariansh. [Internet] 2021 [cited 2023 Jul 26] 15(3):128-153. Available from: https://doi.org/10.1080/19322909.2021.1935396

5. Performance Measurement Task Force/Public Library Association (PLA), Project Outcome for Academic Libraries Task Force/Association of College & Research Libraries (ACRL). Project outcome for academic libraries toolkit [Internet]. Chicago; American Library Association; 2019 [cited 2023 Jul 26]. Available from: https://acrl.projectoutcome.org/about

6. Packer J. Google Analytics alternatives: a guide to navigating the world of options beyond Google [Internet]. Quantable, LLC; 2023 Feb 13 [cited 2023 Jul 26]. Available from: http://arks.princeton.edu/ark:/88435/dsp013n204232m

7. Jansen BJ, Jung S, Salminen J. Measuring user interactions with websites: a comparison of two industry standard analytics approaches using data of 86 websites. PLOS ONE. [Internet]. 2022 [cited 2024 Jan 25] 17(5): 3. Available from: https://doi.org/10.1371/journal.pone.0268212

8. Datanyze. Market share of leading web analytics technologies worldwide in 2023 [Internet]. In: Statista. 2023 Nov 27 [cited 2024 Jan 26]. Available from: https://www.statista.com/statistics/1258557/web-analytics-market-share-technology-worldwide/

9. 6Sense. Google Analytics market share [Internet]. 2024 [cited 2024 Jan 24]. Available

from: https://6sense.com/tech/analytics/google-analytics-market-share

10. McGuirk M. Performing web analytics with Google Analytics 4: A platform review. J Mark Anal, (2023). 11(4), 854-868. Available from: https://doi.org/10.1057/s41270-023-00244-4

11. Google. Bounce rate [Internet]. Mountain View, CA: Google, LLC, 2023 [cited 2023 Jul 26]. Available from: https://web.archive.org/web/20230705160536/https://support.google.com/analytics/answer/-1009409?hl=en

12. Google. How a web session is defined in Universal Analytics [Internet]. Mountain View, CA: Google, LLC, 2023 [cited 2024 Jan 23]. Available from: https://web.archive.org/web/20231226204029/https://support.google.com/analytics/answer/-2731565?hl=en

13. Upton G, Cook I. Acquiescence bias. In: A Dictionary of Statistics. Oxford University Press; 2014 [cited 2024 Jan 23]. Available from: https://www.oxfordreference.com/view/10.1093/acref/9780199679188.001.0001/acref-9780199679188-e-1783

14. Qualtrics. Qualtrics security and compliance [Internet]. Provo, UT: Qualtrics; 2023 [cited 2023 Jul 26]. Available from: https://www.qualtrics.com/platform/security/