



*Evidence Summary*

**More DOIs are Accessed Through Library Discovery Services than Through Google**

**A Review of:**

Wang, X., Cui, Y., & Xu, S. (2018). Evaluating the impact of web-scale discovery services on scholarly content seeking. *The Journal of Academic Librarianship*, 44(5), 545-552.  
<https://doi.org/10.1016/j.acalib.2018.05.010>

**Reviewed by:**

Judith Logan  
User Services Librarian  
University of Toronto Libraries  
Toronto, Ontario, Canada  
Email: [judith.logan@utoronto.ca](mailto:judith.logan@utoronto.ca)

**Received:** 4 Feb. 2019

**Accepted:** 12 Mar. 2019

■ 2019 Logan. This is an Open Access article distributed under the terms of the Creative Commons-Attribution-Noncommercial-Share Alike License 4.0 International (<http://creativecommons.org/licenses/by-nc-sa/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly attributed, not used for commercial purposes, and, if transformed, the resulting work is redistributed under the same or similar license to this one.

DOI: 10.18438/eblip29551

---

**Abstract**

**Objective** – To examine trends in digital object identifier (DOI) web referrals and explore the referring domains, especially those originating from web-scale discovery systems like ProQuest's Summon and Primo.

**Design** – Log analysis and web traffic analysis.

**Setting** – CrossRef, a web server that connects DOIs to the corresponding articles' landing pages.

**Subjects** – Web traffic that passed through CrossRef between 2011 and 2016.

**Methods** – The researchers collected data from CrossRef using a web tool called Chronograph. The data captured information about the websites users were on when they requested a DOI (called the *referrer*) and about the time and date of each request.

The researchers used time series analysis to discover longitudinal patterns in the data. Annual, monthly, and weekly trends were also examined with a seasonal adjustment model, a seasonal trend decomposition, and log transformation. They also isolated traffic from four institutions in Australia, Japan, Sweden, and the United States of America to determine

if overall seasonal patterns were reflected locally.

ProQuest websites were of particular interest to the researchers because they determined that it had the highest market share of discovery services. Much of the analysis focused on ProQuest's [serialsolutions.com](http://serialsolutions.com), [exlibrisgroup.com](http://exlibrisgroup.com), and [proquest.com](http://proquest.com) website domains.

**Main Results** – ProQuest servers sent over 25 million DOI referrals through CrossRef – more than either Web of Knowledge (n=24.47 million) or Google (n=15.38 million).

Referral traffic grew over the period with the sharpest growth rate occurring between 2011 and 2012. Of ProQuest's domains, [serialsolutions.com](http://serialsolutions.com) (Summon) had more traffic and more growth over the observation period than [exlibrisgroup.com](http://exlibrisgroup.com) (Primo).

In all of the years studied, the busiest months were September to November and January to March, while June to August and December were low points. Seasonal fluctuations were attributed to university vacation schedules as demonstrated in the traffic patterns of four ProQuest-subscribing institutions.

Weekly trend analysis showed that Monday to Thursday had consistently heavy referral traffic. Of the remaining days, the fewest referrals were observed on Saturdays.

**Conclusion** – DOI referrer traffic is closely tied to the university calendar. Library discovery products are used more frequently to access DOIs than Google.

### Commentary

The authors have introduced a novel method of examining scholarly resource usage. Log analysis was first adapted for libraries by Nicholas, Huntington, and Watkinson (2005) as a means of dissecting user interactions with an electronic resource or platform. Since then, many other researchers have used log analysis to better understand e-resource usage patterns (Tripathi & Jeevan, 2013). The study at hand is

similar in that it uses raw data to examine interactions with scholarly resources, but it also recalls web traffic analysis studies since web domain referrals are the primary focus. Web traffic studies are usually performed to provide libraries with actionable insights about their communities' behaviour using locally owned data sources like Google Analytics (Turner, 2010). The authors were able to perform a non-local analysis, however, thanks to CrossRef's statistical openness.

The study presented is tightly focused. The analyses center mainly on the referring web domain and the time and date when the referral occurred. This in itself is a rich source of data, and the authors have clearly taken pains to ensure that the temporal trends are presented accurately, though they failed to mention the application used to process the data once obtained from Chronograph. The authors acknowledge that time zones were not considered which could be a factor as referrals were not limited geographically. Future research could determine if time zones affect the patterns discovered in the present study.

The choice of delving into ProQuest domains is fruitful and well-considered. It allows the authors to make sense of the undoubtedly thousands of domains passing through CrossRef by selecting a highly visible suite of scholarly discovery products. Google domains were not included in the study's temporal analyses, so it might be interesting to compare these with ProQuest or other library products to see if user behaviour differs.

The findings are more interesting than actionable. The temporal analyses are strikingly similar to patterns in other scholarly resource or service usage studies suggesting strong external validity (Glynn, 2006). However, the study's methodology is its primary practical contribution. Researchers wishing to apply this methodology to other open scholarly resources may be limited by data availability. CrossRef has made their referral traffic publicly available through Chronograph and other nonprofit scholarly resources should follow suit. The Directory of Open Access Journals and the HathiTrust

come to mind as scholarly resources that would benefit from a similar study.

## References

Glynn, L. (2006). A critical appraisal tool for library and information research.

*Library Hi Tech*, 24(3), 387-399.

<https://doi.org/10.1108/07378830610692154>

Nicholas, D., Huntington, P., & Watkinson, A.

(2005). Scholarly journal usage: The results of deep log analysis. *Journal of Documentation*, 61(2), 248-280.

<https://doi.org/10.1108/00220410510585214>

Tripathi, M., & Jeevan, V. K. J. (2013). A selective review of research on e-resource usage in academic libraries.

*Library Review*, 62(3), 134-156.

<https://doi.org/10.1108/00242531311329473>

Turner, S. J. (2010). Website statistics 2.0: Using Google Analytics to measure library website effectiveness. *Technical Services Quarterly*, 27(3), 261-278.

<https://doi.org/10.1080/07317131003765910>