Can librarians trust resources found on Google Scholar? Yes… and no.

by Blog Admin   September 17, 2012

Many librarians are still unwilling to fully embrace Google Scholar as a resource. Michelle C. Hamilton, Margaret M. Janz and Alexandra Hauser investigate whether Google Scholar has the accuracy, authority and currency to be trustworthy enough for scholars.

One of the reasons science librarians prefer subscription-based databases (ie SciFinder or Web of Science) or those offered as a service of the government (ie PubMed) is the nonstandard behavior of Google Scholar when compared to those resources. Google Scholar is not a bibliographic index, it is a search engine. It does not sort results by relevance, it ranks them according to metadata. It crawls for papers from all across the web, not limiting its results to published documents. These differences make librarians skeptical of Google Scholar, but they do not speak to its actual performance.

In the information profession, there is a lot of talk about whether Google Scholar is a good resource for academics and, generally, librarians are hesitant to say that it is, especially if they work with science research. Librarians are a user-centered group, and understand that students like Google Scholar and use it often. But when science students ask about using Google Scholar for their research, many science librarians provide a response that advises students to be extra critical of search results found there or even direct them to different resources all together.

To determine if these surface deterrents were warranted, we of course had to thoroughly scrutinize Google Scholar. Its citation count accuracy and metrics practices were examined by reading the About Google Scholar web pages. The information gathered there provided many insights, but provided few specifics and raised some questions about what makes the search engine "scholarly." We took our research to the source and questioned Google Scholar via their contact page. We were aware of Google's notorious secrecy regarding the specifics of their metrics, so we limited our questions to those we regarded as non-proprietary: how does Google Scholar define what it considers "scholarly," what journals are crawled, and what metadata is considered for ranking search results. The "Google Scholar Team" responded and after a short exchange of email correspondence (much of which only directly quoted or linked to the About pages), we were told that they could no longer respond, writing, "Sorry, we aren't able to assist you in great detail, it's a free service." This like-it-or-lump-it attitude implies that users should just be happy that Google Scholar is free and should not ask any questions about how it works or where the information comes from.

Knowing the source of information is important for determining its accuracy and authority – something scientific research depends highly on. Previous studies comparing Google Scholar to SciFinder, BIOSIS Previews, PubMed and other databases concluded almost unanimously across disciplines that Google Scholar is a good supplementary resource to use in conjunction with those subject specific databases. Google Scholar retrieves citations for much of the same material found on subscription-based databases and includes results from nontraditional and open source documents not indexed by other resources. However, it's difficult to measure the value of these uniquely retrieved items because it is unknown whether or not they have endured the rigid and controlled evaluation system that defines a work as 'scholarly' in scientific research.

Perhaps the most important thing Google Scholar needs to do to minimize its criticisms among science

librarians is to make its definition of "scholarly" clear. In the sciences, scholarly research must adhere to empirical standards; its content must be verifiable. To meet these standards halfway, Google Scholar needs to provide more transparency about its inclusion process, even if those items retrieved remain questionable in nature. The recent appearance of their "Top Publications" list is a step in the right direction, but it is only one step.

Usability for Google Scholar could also be improved by making search limits more visible and allowing results to be sorted and reordered, rather than ranked.

Despite its faults, librarians should warm up to Google Scholar. They should play with it and test it alongside the old stand-bys as they do their own research and answer reference questions. Depending on the research topic, it can provide some relevant sources that might not be found outside of Google Scholar. Users should be advised to be critical of information found by any means, but perhaps be more vigilant with Google Scholar. For that reason, Google Scholar should be included in library instruction to familiarize users with its functionality as well as its limitations. Knowing how to get the most out of all resources is important for librarians; Google Scholar shouldn't be excluded simply because it's unconventional.

As MLS candidates at Indiana University Bloomington, Michelle C. Hamilton, Alexandra Hauser, Margaret M. Janz & Fiona Taggart, together with classmates Jerry Gray and Justin P. Peters examined Google Scholar as a resource for scientific research and how it should be considered by science librarians. Their article "Scholarish: The Value of Google Scholar to the Sciences," from Issues in Science and Technology Librarianship (Summer 2012), offers advice for science librarians when confronted with student questions regarding the use of Google Scholar.

Note: This article gives the views of the author(s), and not the position of the Impact of Social Sciences blog, nor of the London School of Economics.

About the authors:
Michelle Hamilton *has a Bachelor of Arts in History from Saint Mary's College and completed her Master of Library Science at Indiana University Bloomington in August 2012. She is especially interested in the aspects of reference, research, and instruction of information literacy among university students.*
Alexandra Hauser *has a Bachelor of Arts in History from Michigan State University and is currently pursuing a Master of Library Science at Indiana University Bloomington. Her research interests include information literacy issues and library assessment. She is also the co-author of the forthcoming article, "Beyond gatekeepers of knowledge: Scholarly communication practices of academic librarians and archivists at ARL institutions" to be published in College & Research Libraries.*
Margaret Janz *has a bachelor of science in information resources from the University of Wisconsin – Milwaukee and is currently pursuing a Master of Library Science with a specialization in chemical information at Indiana University Bloomington, where she works in the Chemistry and Life Sciences Libraries. She is particularly interested in issues relating to open science information, international information collaboration, and data management.*

Related posts:

1. Google Scholar Citations: a way for academics to compute citation metrics and track them over time

2. First impressions of Google Scholar Citations are good: it's easy to use and accurate

3. Wow – Google Scholar 'Updates' a big step forward in sifting through the scientific literature

4. Google Scholar Citations is now open to everyone. It shows great promise as a free, reliable way to track and compare academic impact over time.

5. Why 'Publish or Perish' has the edge over Google Scholar and Scopus when it comes to finding out how your work is used by other academics