September 2005

Technology Left Behind -- Google Scholar and the OpenURL

Cris Ferguson

Furman University, cferguson13@murraystate.edu

Follow this and additional works at: <https://docs.lib.purdue.edu/atg>

Part of the [Library and Information Science Commons](#)

Recommended Citation

Ferguson, Cris (2005) "Technology Left Behind -- Google Scholar and the OpenURL," *Against the Grain*: Vol. 17: Iss. 4, Article 32.
DOI: <https://doi.org/10.7771/2380-176X.4535>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

How to Order the “10 Reasons” Poster

“10 Reasons Why the Internet is No Substitute for a Library” (*American Libraries*, April, 2001) has been turned into a 4-color, 22" x 36" poster (ALA “READ” size poster) and is available for \$10 each (this **includes** shipping and handling). Bulk orders are available on request and begin at 16 or more orders for \$8 each. (South Carolina residents, add applicable sales tax.)

If interested, please respond to <herringm@winthrop.edu>. Check or money order is acceptable and should be made out to “Winthrop

University.” (You may also send a purchase order, or we will bill you.) They should be mailed to Winthrop University, Dacus Library, Attn: MY Herring, Rock Hill, SC, 29733.

Technology Left Behind — Google Scholar and the OpenURL

Column Editor: **Cris Ferguson** (Electronic Resources/Serials Librarian, James B. Duke Library, Furman University, 3300 Poinsett Highway, Greenville, SC 29613; Phone: 864-294-2713) <cris.ferguson@furman.edu>

Google Scholar

Last November, Google released **Google Scholar**, a new search tool focused on finding scholarly information on the Web. Unlike Google's basic search, **Google Scholar** searches a finite set of Web pages containing scholarly information and documents. **Google Scholar** analyzes the contents of a Web page and extracts citations, which are made available on the results screen, even if the document itself is not available online. “This means your search results may include citations of older works and seminal articles that appear only in books or other offline publications.” (<http://scholar.google.com/scholar/about.html>)

While Google has not released the details about what titles, publishers, and vendors are included in the database, searches include results from organizations such as the **American Physiological Society** and the **American Pediatric Society**. **Google Scholar** also searches electronic journal collections from publishers, like the **Oxford University Press** and **Springer**, as well as aggregated collections via services like **IngentaConnect**. Other resources cited in **Google Scholar** include open-access journals, like the *Journal of Biological Chemistry*, as well as subscription publications. One search even

included a hit on a paper provided on the Web page of the Chemistry Department at the **University of Wisconsin**. Overall, the database covers a broad spectrum of resources, but without specific indexing information, which Google does not provide at this time, it is difficult to evaluate the completeness of the database.

Google does not define anywhere on its Website what qualifies as scholarly and what does not. I ran a few searches in the database to see if I could get some idea of how well Google had pinpointed what is and is not “scholarly material.” I searched for **Leslie Marmon Silko**, a prominent Native American author of both fiction and non-fiction material. Without quotation marks around **Silko's** name the search retrieved 577 hits. The first ten hits were citations to books authored by **Silko** or anthologies to which she has contributed. The eleventh citation was to a book review in a journal in *Project Muse*.

I then conducted a search for the chef **Jamie Oliver**, author of a number of cookbooks and host of the popular cooking show, *The Naked Chef*, which airs on the Food network. I searched the phrase ‘**Jamie Oliver Naked Chef**’ without quotation marks or **Boolean** op-

erators. The search retrieved 27 results, most of which were citations to **Oliver's** cookbooks. However, there was an article from the *European Journal of Cultural Studies*, a Sage publication, on British television programming, and another article from the *British Food Journal* on celebrity endorsement. While I consider **Jamie Oliver's** cookbooks to be less scholarly than **Leslie Marmon Silko's** novels and essays, there is an argument to be made that the articles mentioning him are scholarly in nature.

Google Scholar sorts results in relevancy order, which is determined by “the full text of each article as well as the article's author, the publication in which the article appeared and how often it has been cited in scholarly literature.” (<http://scholar.google.com/scholar/about.html>) Unfortunately, there is no way to sort the results into any other order, such as by author or date, a feature that would certainly be useful for academics in any field.

Using the OpenURL

Almost immediately following the release of **Google Scholar**, librarians and information professionals began searching for ways to point their users from citations in **Google Scholar** to their libraries' resources.

continued on page 84

Within weeks of the release **Peter Binkley** at the **University of Alberta Libraries** wrote an extension for **Firefox** that adds **OpenURL** links to the **Google Scholar** results screen. **Binkley's** extension scrapes the **Google Scholar** results screen, parses a citation out of the data, and builds an **OpenURL** to a designated link resolver. (<http://www.ualberta.ca/~pbinkley/gso/>) **Art Rhyno** at the **Leddy Library** at the **University of Windsor** wrote a similar bookmarklet designed to "add a proxy address to the search results on a **Google Scholar** page of results." (<http://webvoy.uwindsor.ca:8087/artblog/librarycog/1100880268>) Coupled together these bookmarklets created the base for the **Windsor-Alberta-Georgia Google Scholar** Localization project (**WAG the Dog**), an open-source project available via **SourceForge**. (<http://gslocal.sourceforge.net/>) While the **WAG the Dog** bookmarklet is effective, it requires a great deal of user intervention, and a more seamless method of providing **OpenURL** links would be more effective.

This past spring, **Google** began pursuing **OpenURL** linking on its own, launching a pilot program testing **OpenURL** linking within **Google Scholar** in conjunction with several link resolver vendors, including **Ex Libris**, **Serials Solutions**, and **1Cate**. Initially, **OpenURLs** from **Google Scholar** included only **Digital Object Identifiers (DOIs)** and **PubMed IDs**, leading to a high rate of failure when retrieving the full text. Since then, **Google** has begun transmitting additional pieces of metadata, like title, author, volume, and issue information, and the resolution rate has improved.

There are distinct advantages to setting up **Google** as a source in a library's link resolver, such as the ability to push services out to users in a way that has never been permitted before and at the same time brand those services to the library. **Google's** implementation of the **OpenURL** allows an integration of a Web search engine and the library's resources that has not previously been possible. However, **Google's** policies and procedures for setting up the **OpenURL** links have created a stir and are hotly debated in the library and information science world.

Before **Google** will turn on **OpenURL** linking for a library, it requires the library to submit its electronic full text holdings. Submitting the holdings themselves is a relatively simple process, as **Google** has expedited processes for this with many link resolver vendors. It is how **Google** has put those holdings to use that has created controversy. Using the holdings information **Google** determines if the electronic full text is available via the library's subscriptions, and, if it is, the link to the resolver appears prominently in the citation in a large font.

If the electronic full text is not available, the link to the resolver still appears but it is in a smaller font face and buried below the citation. Libraries have control over the wording of those links, but they cannot use a button or some other image to represent the link.

Many libraries are reluctant to share their holdings with **Google**, feeling that the policy is against the principle of the **OpenURL**. Other issues of concern are the lack of information about what **Scholar** is indexing, the completeness of the **OpenURLs** **Google** is sending, and the placement of the two **OpenURL** links.

Conclusion

Google Scholar is not bad, but it's not great either. It is what it is, a Web search engine focused on scholarly materials. Based upon the few simple searches I conducted, the materials cited in the database seemed either scholarly themselves or at least related to the scholarly materials in some way. But we are still talking about a search engine, and I still want my patrons to approach it with the same caution that they would any other search engine for all the same reasons (incomplete indexing, lower quality metadata, unreliable linking, "scholarlyness" is subjective, Web-crawlers can only do so much, etc.).

However, my gut feeling is that no matter what I tell the students I see at the reference desk they are going to hear the words "**Google**" and "**Scholar**" and think they have found the cure for what ails them. One of the reasons setting up **Google Scholar** as a source in our link resolver is appealing is the fact that once users perform a search we can at least point them back to our full text electronic resources, print subscriptions, and document delivery services. On the other hand, I don't particularly want to have to maintain our electronic subscriptions in another location. Nor do I want **Google** to bury the link to the link resolver if the electronic full text is not available through our subscrip-

tions. The **OpenURL** link should appear in the same place every time. While we have not set up our link resolver in **Google Scholar** at this time, it is easy to foresee that we will should the need seem sufficient enough or if our users specifically request to be able to link directly to our electronic resources from **Google Scholar**.

This conflict between libraries and **Google** over the best way to implement the **OpenURL** in **Google Scholar** is a demonstration in the fundamental differences in a library's philosophy and **Google's** philosophy. **Google's** main priority is to get the user to the electronic full text, and their policies on the submission of holdings information and the format of the **OpenURL** links reflect that. While as librarians, our main priority is to find users the appropriate copy whatever its format, print, electronic, microform, or a document ordered through a document delivery service.

If you are interested in further reading about **Google Scholar** and the **OpenURL**, these issues are explored in greater detail, including interviews with librarians, link resolver vendors, and **Google Scholar**, in an article by myself and **Jill Grogg**, Electronic Resources Librarian at the **University of Alabama** in an upcoming issue of *Searcher Magazine*.

Other Articles of Interest

Grogg, Jill E. and Christine L. Ferguson. "OpenURL Linking in **Google Scholar**." *Searcher Magazine* [forthcoming].

Notess, Greg R. "Scholarly Web Searching: **Google Scholar** and **Scirus**." *Online* 29, issue 4. (July/August 2005): 39-41.

Tenopir, Carol. "Google in the Academic Library." *Library Journal* 130, issue 2 (February 1, 2005): 32.

Wlekinski, Joann M. "Studying **Google Scholar**: Wall to Wall Coverage?" *Online* 29, issue 3 (May/June 2005): 22 (5).

Comments sought on the draft COUNTER Code of Practice for Books and Reference Works

by **Peter T. Shepherd** (Project Director, COUNTER)
<pt_shepherd@hotmail.com>

The draft of Release 1 of the new **COUNTER Code of Practice** for online books and reference works was published in January 2005. This marks the first expansion of **COUNTER's** coverage beyond journals and databases. The **Code of Practice** for online books and reference works has been developed with input from a task force of librarians and publishers with expert knowledge of these products and is the first attempt to introduce a compre-

hensive industry standard for the recording and reporting of online book usage data. Its overall format and structure are consistent with the existing **COUNTER Code of Practice** for journals and databases: only the content of the usage reports has been changed and the set of definitions of terms expanded. The specifications for report delivery, data processing, auditing, and compliance are identical to those already prescribed in the **Code of Practice** for journals and databases.

continued on page 85