

Putting Google™ Scholar to the Test on Patent Research

by Andrea L. Hamilton

The effectiveness of Google™ Scholar as an initial starting point for legal researchers investigating patents, compared with the commercial databases Westlaw¹ and LexisNexis², may be a matter of interest to legal practitioners. Also of interest would be whether the results culled from Google Scholar are more current than those from Westlaw and LexisNexis. It may be tempting to turn to free resources for legal research; however, users should be prepared to evaluate whether the results are worth the potential cost savings. This article compares the Google Scholar search engine to that of Westlaw and LexisNexis specifically related to patent research, and reviews the accuracy of the results.

History of Google Scholar

The Google Scholar beta search tool went online in November 2004, presenting itself as a starting point for scholarly research. To facilitate content searches, the company collaborated specifically with publishers of scientific and academic materials such as technical reports, books, and peer-reviewed papers.³

Founded by computer science graduate students, Google has often courted “engineers, scientists and mathematicians [and] has made for a workforce attuned to the needs of the academic market.”⁴ The engineer who headed up the initiation of Google Scholar, Anurag Acharya, stated that this search tool was “one way we give back to the community.”⁵ He also wanted to address the matter of outdated material in libraries.⁶

Comparing Search Methods

One way to evaluate the effectiveness of Google Scholar for patent research is to conduct the same search in Google Scholar,

Westlaw, and LexisNexis. For example, a search for patents invented by Ricci L. Marzolf of New Hampton, Iowa,⁷ produced the same twenty-four results in Westlaw⁸ and LexisNexis⁹—a combination of applications, utility patents, and design patents occurring between 2001 and 2009. However, a search in Google Scholar¹⁰ using only the inventor’s first and last name retrieved thirty-six results; fourteen were relevant applications, utility patents, and design patents also picked up by the LexisNexis and Westlaw searches. However, because Google Scholar allows for patents to be searched only in conjunction with articles, the results also included twenty-one article and book references listed, as well as one patent unrelated to this inventor.¹¹

Search Results

An audit of the fourteen relevant Google Scholar patent results highlights some of the potential limitations of Google Scholar. Google Scholar failed to pick up ten patents and applications that were pertinent to this research and that were found by Westlaw and LexisNexis. This might indicate that Google Scholar uses an incomplete U.S. patent collection; however, because “Google Scholar doesn’t publish a list of its contributing databases and Web sites,”¹² and doesn’t state how often the database is updated, the researcher cannot know whether the results are complete or what content may be missing (for example, newer patents or non-U.S. patents) unless the Google results are compared to the Westlaw and LexisNexis results. Unfortunately, a review of the ten documents picked up by Westlaw and LexisNexis but missing from Google Scholar’s results does not necessarily provide insight as to Google Scholar’s patent coverage, because both the oldest¹³ and



About the Author

Andrea L. Hamilton is the Research Librarian at the Boulder office of Faegre & Benson LLP and a member of the Colorado Lawyers Committee—(303) 447-7733, ahamilton@faegre.com.

This Department, published quarterly, is sponsored by the Colorado Association of Law Librarians (CoALL) to assist attorneys with common problems in legal research. Readers interested in submitting research questions may send them to: CoALL, Legal Research Corner, at www.aallnet.org/chapter/coall; or to: Legal Research Corner, c/o Leona Martínez, Managing Editor of *The Colorado Lawyer*, at leonamartinez@cobar.org.

Members of CoALL will attempt to answer as many questions as possible, either individually or as part of this Department. The information provided in this space is for educational purposes only and is not intended as legal advice. No endorsement or recommendation is made of any product named in articles appearing in this Department. Department coordinators are CoALL members Andrea Hamilton, Madeline Cohen, Theresa (Tracy) Leming, and Esti Shay. For more information about CoALL, visit www.aallnet.org/chapter/coall.

one of the newest¹⁴ U.S.-granted patents invented by Ricci L. Marzolf were included in the results.

According to both Westlaw and LexisNexis, the most recently granted patent invented by Ricci L. Marzolf is U.S. Patent Number 7,611,173, “Latch System Kit and Method of Making a Latch System.” Although a quick scan by patent name would seem to indicate that Google Scholar did pick up this result, closer inspection of the document pinpoints another potential limitation of this search tool: the document included is the published application,¹⁵ not the granted patent. It is a relevant document, but not the most recent version. Google Scholar also does not clearly indicate whether there is any previous or subsequent history to a patent application publication. In this case, for example, links provided in the Google results list directing users to “all versions” of a patent document actually link to two versions of the same patent application publication that were pulled from two different sources.¹⁶ Moreover, the links to “all versions” do not include those posted by the U.S. Patent and Trademark Office, the federal agency that grants U.S. patents.¹⁷

In contrast, both Westlaw and LexisNexis provided the most recently granted patent in the returned results (the published application also was included in the results). Additionally, a researcher may KeyCite (Westlaw) or Shepardize (LexisNexis) patents to review the full history of a particular patent, providing context to the document that is being viewed in full. A search in Google Scholar for U.S. Patent Number 7,611,173 yielded no results, which explains why it did not appear in the list of search results conducted by the inventor’s name.

Misspelled Searches

Further investigation of the “all versions” links provided for other patent documents related to a search produced surprising results. Google Scholar provides a link to four versions of U.S. Application Number 2005/0,211,535, invented in part by Ricci L. Marzolf. Again, two of these results are the same published patent application pulled from two different sources. The other two versions are the granted patent,¹⁸ which did not turn up in the results lists for the Westlaw, LexisNexis, or Google Scholar searches. The reason: the inventor’s name is misspelled as Ricci L. “Mazolf.” Thus, Google Scholar might have made the connection between the published application and the granted patent using the application number,¹⁹ providing a value-added service to the researcher. A Westlaw KeyCite report did not list the granted patent as part of the history portion.

Additional Search Options

Google Scholar does not offer sorting capabilities for a researcher to isolate the relevant patent-related documents from the total hits; results are ranked by relevancy according to Google’s proprietary algorithm.²⁰ For a more refined patent search, Google offers a separate Google Patents beta interface²¹ with an Advanced Patent Search option.²² The Google Patents and Advanced Patent Search also can be accessed via a top menu that appears when viewing an application, utility patent, or design patent in full. The Advanced Patent Search allows for, among other fields, a query by inventor. A search for “ricci marzolf” and “rl marzolf” yielded no results, and a search for “marzolf” yielded one patent invented by Norbert F. Marzolf.²³ A more general search—not limited to a

particular field—for “ricci marzolf”²⁴ turned up fifteen documents, including ten results that were part of the fourteen results from the previous search, as well as the one previously found patent not related to this inventor.²⁵ Interestingly, three relevant patents that were not included among the fourteen results from the previous search are included in this set of results.²⁶ Also of interest, the remaining result is the patent that was not picked up in previous searches because of the inventor’s misspelled name.

It is difficult to account for the discrepancy between these two sets of inventor results because Google Scholar does not provide a list of contributing sources, and it is sometimes difficult to determine the source of a document, even when viewing it in full.²⁷ This is a specialized search tool and it is possible that Google Patents incorporates more patent-specific proprietary data. As stated in the *Canadian Medical Association Journal*: “Through agreements with publishers, Scholar accesses the ‘invisible’ or ‘deep’ Web, that is, commercial Web sites the automated ‘spiders’ used by search engines such as Google cannot access.”²⁸

Conclusion

So, how effective is Google Scholar as an initial starting point for legal researchers, not only for patent research, but for potential areas of interest such as caselaw and scholarly articles? A prominent legal blogger commented: “Both Lexis and Westlaw think their analysis and assurance of accuracy and completeness make them immune to the Google challenge.”²⁹ It may be true that Westlaw and LexisNexis contain a significant amount of information and commentary in multiple areas of law and news, but it can be demonstrated that their content is not without flaws.

Google Scholar, however, also has its own set of shortcomings. The collections being searched are incomplete, the documents are supplied by undisclosed sources, and Google does not disclose when and how often new content is added. The documents found may not be the most recent version—something to be kept in mind when conducting case law research—and any previous or subsequent history of the document may not be clear. Also, because Google Scholar’s relevancy rankings “emphasize pages that are cited more often, this creates a bias toward older literature,”³⁰ especially applicable when conducting article research.

Furthermore, commercial databases such as Westlaw and LexisNexis offer enhancements that are not available on Google, such as field restrictors,³¹ proximity searching, the ability to sort and filter results, along with twenty-four-hour customer service. “The bottom line is that if you have a casual need for scholarly information, use Google Scholar. But for serious endeavors, don’t,”³² or use it as a starting place to “provide a small but representative selection of scholarly research on most subjects”³³ so that you may be more efficient with your time and money when using the costly commercial databases.

Notes

1. Westlaw® is available at www.westlaw.com.
2. LexisNexis® is available at www.lexis.com.
3. Markoff, “Google Plans New Service For Scientists and Scholars,” *The New York Times* C6 (Nov. 18, 2004).
4. Gaither, “Google Designs an Engine for Eggheads,” *Los Angeles Times* 1 (Dec. 27, 2004).
5. *Id.*
6. Markoff, *supra* note 3.

7. This name was selected for its uniqueness and because the inventor's body of work is fairly small and recent.

8. The following search query was run in the U.S. Patents Full-Text (US-PAT) database: "INV(RICCI /3 MARZOLF)."

9. The following search query was run in the Utility, Design and Plant Patents database: "INVENTOR(RICCI /3 MARZOLF)."

10. The following search query was run at scholar.google.com: "ricci marzolf (Articles, includes patents)."

11. U.S. Patent No. 5,887,577, "Apparatus for Propelling a Projectile," was picked up because Marzolf was listed in References Cited and because the primary examiner was John A. Ricci.

12. O'Leary, "Google Scholar: What's in It for You?" 35 *Information Today* 22(7) (July 1, 2005).

13. U.S. Patent No. D435,779, "Operator For a Closure Element" (date of patent: Jan. 2, 2001).

14. U.S. Patent No. 7,338,097, "Latch Assembly For a Movable Closure Element" (date of patent: March 4, 2008).

15. U.S. Publication No. 2008/0018116 A1, "Latch System Kit and Method of Making a Latch System" (Jan. 24, 2008).

16. See www.google.com/patents and www.freepatentsonline.com.

17. The U.S. Patent and Trademark Office (USPTO), "The USPTO: Who We Are" (Dec. 30, 2009), available at www.uspto.gov/about/index.jsp.

18. U.S. Patent No. 7,234,329, "Lock System for Movable Closure Element" (date of patent: June 26, 2007).

19. Application No. 10/811,692.

20. "Google Scholar aims to rank documents the way researchers do, weighing the full text of each document, where it was published, who it was written by, as well as how often and how recently it has been cited in

other scholarly literature." See www.scholar.google.com/intl/en/scholar/about.html. However, it is unclear if this ranking system applies to patents, as well.

21. See www.google.com/patents.

22. See www.google.com/advanced_patent_search.

23. U.S. Patent No. 2,627,141, "Egg Scouring Machine" (date of patent: Feb. 3, 1953).

24. This search was run at www.google.com/patents.

25. U.S. Patent No. 5,887,577, "Apparatus for Propelling a Projectile" (date of patent: March 30, 1999).

26. U.S. Patent No. 7,097,216, "Latch Assembly For Movable Closure Element" (date of patent: Aug. 29, 2006); U.S. Patent No. 6,942,259, "Latch Assembly" (date of patent: Sept. 13, 2005); U.S. Patent No. D506,377 (date of patent: June 21, 2005).

27. The researcher might assume—whether correctly or not—that the USPTO is not the direct source for the search results because, when selecting to view a result in full, Google Patents provides a separate link to "View patent at USPTO."

28. Henderson, "Google Scholar: A source for clinicians?" *Canadian Medical Association Journal* (June 7, 2005).

29. Mystal, "Google Does Evil to LexisNexis and Westlaw?" *Deal-Breaker: Above the Law* (Nov. 24, 2009).

30. Henderson, *supra* note 28.

31. See www.scholar.google.com/advanced_scholar_search and www.google.com/advanced_patent_search. These links provide templates for field searching, but their potential ineffectiveness has been demonstrated.

32. O'Leary, *supra* note 12.

33. *Id.* ■