Google Scholar

Google Scholar

Web address	scholar.google.com [1]
Type of site	bibliographic database
Registration	Optional
Owner	Google
Current status	Active

Google Scholar is a freely accessible web search engine that indexes the full text of scholarly literature across an array of publishing formats and disciplines. Released in beta in November 2004, the Google Scholar index includes most peer-reviewed online journals of Europe and America's largest scholarly publishers, plus scholarly books and other non-peer reviewed journals. It is similar in function to the freely available Scirus from Elsevier, CiteSeerX, and getCITED. It is also similar to the subscription-based tools, Elsevier's Scopus and Thomson ISI's Web of Science. Its advertising slogan – "Stand on the shoulders of giants" – is a nod to the scholars who have contributed to their fields over the centuries, providing the foundation for new intellectual achievements.

History

Google Scholar arose out of a discussion between Alex Verstak and Anurag Acharya, both of whom were then working on building Google's main web index.

In 2006, in response to release of Microsoft's Windows Live Academic Search, a potential competitor for Google Scholar, a citation importing feature was implemented using bibliography managers (such as RefWorks, RefMan, EndNote, and BibTeX). Similar features are also part of other search engines, such as CiteSeer and Scirus.

In 2007, Acharya announced that Google Scholar had started a program to digitize and host journal articles in agreement with their publishers, an effort separate from Google Books, whose scans of older journals do not include the metadata required for identifying specific articles in specific issues.

In 2011, Google removed Scholar from the toolbars on its search pages, making it both less easily accessible and invisible to users who were not already aware of its existence.

In 2012, an individual Google Scholar page feature was added. Individuals, logging on through a Google account with a bona fide address usually linked to an academic institution, can now create their own page giving their fields of interest and citations. Google Scholar automatically calculates and displays the individual's total citation count, h-index, and i10-index. Examples are here and here. Top citations in a field of interest can also be accessed, e.g. here.

Features and specifications

Google Scholar allows users to search for digital or physical copies of articles, whether online or in libraries.[2] It indexes "full-text journal articles, technical reports, preprints, theses, books, and other documents, including selected Web pages that are deemed to be 'scholarly.'" Because many of Google Scholar's search results link to commercial journal articles, most people will be able to access only an abstract and the citation details of an article, and have to pay a fee to access the entire article. The most relevant results for the searched keywords will be listed first, in order of the author's ranking, the number of references that are linked to it and their relevance to other scholarly literature, and the ranking of the publication that the journal appears in.

Using its "group of" feature, it shows the available links to journal articles. In the 2005 version, this feature provided a link to both subscription-access versions of an article and to free full-text versions of articles; for most of 2006, it provided links to only the publishers' versions. Since December 2006, it has provided links to both published versions and major open access repositories, but still does not cover those posted on individual faculty web pages;[*citation needed*] access to such self-archived non-subscription versions is now provided by a link to Google, where one can find such open access articles.

Through its "cited by" feature, Google Scholar provides access to abstracts of articles that have cited the article being viewed.[3] It is this feature in particular that provides the citation indexing previously only found in Scopus and Web of Knowledge. Through its "Related articles" feature, Google Scholar presents a list of closely related articles, ranked primarily by how similar these articles are to the original result, but also taking into account the relevance of each paper.[4]

As of July 2013[5], Google Scholar is not yet available to the Google AJAX API.

Google Scholar's legal database of US cases is extensive. Users can search and read published opinions of US state appellate and supreme court cases since 1950, US federal district, appellate, tax and bankruptcy courts since 1923 and US Supreme Court cases since 1791. Google Scholar embeds clickable citation links within the case and the How Cited tab allows lawyers to research prior case law and the subsequent citations to the court decision. The Google Scholar Legal Content Star Paginator extension inserts Westlaw and LexisNexis style page numbers in line with the text of the case.

Ranking algorithm

While most academic databases and search engines allow users to select one factor (e.g. relevance, citation counts, or publication date) to rank results, Google Scholar ranks results with a combined ranking algorithm in a "way researchers do, weighing the full text of each article, the author, the publication in which the article appears, and how often the piece has been cited in other scholarly literature". Research has shown that Google Scholar puts high weight especially on citation counts[6] and words included in a document's title. As a consequence the first search results are often highly cited articles.

Limitations and criticism

Access to publications — Use of Google Scholar does not automatically grant searchers access to publications available through subscription sites. The full texts of articles in Google Scholar are not necessarily available freely to all searchers, though searchers with access through an institution such as a research laboratory or university may be able to access select articles freely. Some articles found through Google Scholar are hosted by sites that allow searchers to subscribe or purchase the full text of their articles.

Quality — Some searchers consider Google Scholar of comparable quality and utility to commercial databases. The reviews recognize that its "cited by" feature in particular poses serious competition to Scopus and ISI Web of Knowledge, although, in a study limited to the biomedical field, the citation information in Google Scholar has been found to be sometimes inadequate, and less often updated. Another important issue is that the relative coverage of

Google Scholar varies by discipline compared to other general databases.

Vulnerability to Advertising - A growing problem with Google Scholar is the ability of vendors to include advertising and product brochures among the results of scholarly works. For example a search for 'high accuracy electrometers' will result in a product brochure for Keithly electrometers from a vendor called 'TestForce' who is blatantly compromising Google Scholar for marketing purposes. This brochure is 'clearly' not a scholarly work.

Secrecy about coverage — A significant problem with Google Scholar is the secrecy about its coverage.Wikipedia:Manual of Style (words to watch)#Editorializing Some publishers do not allow it to crawl their journals. Elsevier journals were not included before mid-2007, when Elsevier began to make most of its ScienceDirect content available to Google Scholar and Google's web search. As of February 2008 the absentees still included the most recent years of the American Chemical Society journals. Google Scholar does not publish a list of scientific journals crawled, and the frequency of its updates is unknown. It is therefore impossible to know how current or exhaustive searches are in Google Scholar. Nonetheless, it allows easy access to published articles without the difficulties encountered in some of the most expensive commercial databases.

Matthew effect — Google Scholar puts high weight on citation counts in its ranking algorithm and therefore is being criticised for strengthening the Matthew effect; as highly cited papers appear in top positions they gain more citations while new papers hardly appear in top positions and therefore get less attention by the users of Google Scholar and hence fewer citations.

Incorrect field detection — Google Scholar has problems identifying publications on the arXiv preprint server correctly. Interpunctuation characters in titles produce wrong search results, and authors are assigned to wrong papers, which leads to erroneous additional search results. Some search results are even given without any comprehensible reason.

Vulnerability to spam — Google Scholar is vulnerable to spam.[7] Researchers from the University of California, Berkeley and Otto-von-Guericke University Magdeburg demonstrated that citation counts on Google Scholar can be manipulated and complete non-sense articles created with SCIgen were indexed from Google Scholar. They concluded that citation counts from Google Scholar should only be used with care especially when used to calculate performance metrics such as the h-index or impact factor. Google Scholar started computing an h-index in 2012 with the advent of individual Scholar pages. Several downstream packages like *Harzing's Publish or Perish* also use its data. The practicality of manipulating h-index calculators by spoofing Google Scholar was demonstrated in 2010 by Cyril Labbe from Joseph Fourier University, who managed to rank "Ike Antkare" ahead of Albert Einstein by means of a large set of SCIgen-produced documents citing each other (effectively an academic link farm).

Google Scholar is also not able to shepardize case law, as Westlaw and Lexis can.

References

[1] http://scholar.google.com

[2] Google Scholar Library Links (http://scholar.google.com/intl/en/scholar/librarylinks.html)

[3] Google Scholar Help (http://scholar.google.co.uk/intl/en/scholar/help.html)

[4] Official Google Blog: Exploring the scholarly neighborhood (http://googleblog.blogspot.com/2006/08/exploring-scholarly-neighborhood.html)

[5] http://en.wikipedia.org/w/index.php?title=Google_Scholar&action=edit

[6] Jöran Beel and Bela Gipp. Google Scholar's Ranking Algorithm: An Introductory Overview. In Birger Larsen and Jacqueline Leta, editors, Proceedings of the 12th International Conference on Scientometrics and Informetrics (ISSI'09), volume 1, pages 230–241, Rio de Janeiro (Brazil), July 2009. International Society for Scientometrics and Informetrics. ISSN 2175-1935. (http://www.sciplore.org/publications/2009-Google_Scholar's_Ranking_Algorithm_--_An_Introductory_Overview_--_preprint.pdf)

[7] Joeran Beel and Bela Gipp. On the Robustness of Google Scholar Against Spam. In *Proceedings of the 21st ACM Conference on Hyptertext and Hypermedia*. ACM, June 2010. Preprint PDF (http://www.sciplore.org/publications/2010-On_the_Robustness_of_Google_Scholar_against_Spam--preprint.pdf)

External links

- Google Scholar WebSite (http://scholar.google.com)
- Google Scholar Blog (http://googlescholar.blogspot.com)

Article Sources and Contributors

Google Scholar *Source*: http://en.wikipedia.org/w/index.php?oldid=571105085 *Contributors*: 2601:D:380:B5:691F:ECF1:ABF0:5803, 478jjjz, A5b, Adamkvonende, Alan Liefting, Angrytoast, Arthena, AuthorAuthor, BDD, Bejnar, Bender235, Beta M, Bobo192, Brandon, Briancricks, Bug27, Byrgenwulf, CIreland, Chansonh, Cheerocracy, Cherubino, Chris Capoccia, Chris the speller, Cnilep, CommonsDelinker, Compfreak7, DGG, David Eppstein, Derekwill500, Dicarlo, Djavidan, Dmfigaro, DocWatson42, Dr.alf, Dudes15r, Duncan.Hull, EdRicardo, Epbr123, Eric Kvaalen, Esperant, Esvwiki, FayssalF, Fi99ig, Flatterworld, Flyguy649, Fmccown, Fmenczer, Foxjwill, Fraggle81, Francisco Couto, Francs2000, Gfoley4, Gwernol, HaeB, Headbomb, Hgilbert, Higalva, Hroðulf, I am not a dog, Ilovelibrary, Iohannes Animosus, Ipeirotis, Ivo, JavOs, Jonathan Schuster, Jreiss186, Kaelswanson, Kanzure, Karnesky, Kesla, Km1498, Lightmouse, Luogang.china, Mack2, Massreg, Materialscientist, Maurice Carbonaro, Mervyn, Miym, Mporcheron, MrOllie, Mucus, NTox, Najoj, Neo-Jay, Nohat, Nurg, Okyea, Ontoraul, Paul Siebert, Peterdx, Pickhorn, Popopp, RDBrown, Rabarberski, Rainmannn, Rajah, Rich Farmbrough, Richard Arthur Norton (1958-), Richard001, Riecker608, Rjwilmsi, RoyBoy, Ruud Koot, ScholarBug, SchreiberBike, Shlg2, Sift&Winnow, Tellyaddict, Tijfo098, Tisane, UBeR, Valeria.depaiva, VxOsx, Wdustbuster, Wikip rhyre, Wile E. Heresiarch, Wjejskenewr, Xicouto, Z10x, Zenohockey, Zollerriia, Zvika, 140 anonymous edits

Image Sources, Licenses and Contributors

File:Google_Scholar_logo.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Google_Scholar_logo.png *License*: Public Domain *Contributors*: Original uploader was PNG crusade bot at en.wikipedia. Later version(s) were uploaded by Logan, Beao, GageSkidmore at en.wikipedia.

License