Google Scholar's Coverage of the Engineering Literature: An Empirical Study

By

John J. Meier
Science Librarian
Physical and Mathematical Sciences Library
Pennsylvania State University
201 Davey Laboratory
University Park, PA 16802
meier@psu.edu

and

Thomas W. Conkling
Head, Engineering Library
Pennsylvania State University
325 Hammond Building
University Park, PA 16802
twc1@psu.edu

Title

Google Scholar's Coverage of the Engineering Literature: An Empirical Study

Abstract

Google Scholar's coverage of the engineering literature is analyzed by comparing its contents with those of Compendex, the premier engineering database. Records retrieved from Compendex were searched in Google Scholar, and a decade by decade comparison was done from the 1950s through 2007. The results show that the percentage of records appearing in Google Scholar increased over time, approaching a 90% matching rate for materials published after 1990.

Introduction

The introduction of a new web-based service or database can be a welcome event if the product represents an advance over what is currently available in terms of effectiveness, cost, or time savings, or if it shows promise for doing so. Interest in a new service is heightened when it comes from an organization with a stellar track record in the industry. Google has delivered a stream of innovative products for the web, beginning with its popular search engine and continuing with services such as Books Search, Docs and Spreadsheets, Image Search and Maps. For anyone involved in providing information services in an academic setting, the launch of Google Scholar in 2004 stirred a high level of interest. In particular, the promise of an easy to use bibliographic search tool, based on the familiar and popular Google search engine, held great appeal to academic librarians in the sciences and engineering. These individuals work with high-priced subject databases on a daily basis, and the potential offered by a free service is very intriguing. Academic librarians are also dealing with a student audience (both undergraduate and graduate) who have grown up with Google and use it as their "go to" search service.

Google Scholar has been available for several years now. Searches on Google Scholar consistently retrieve bibliographic records only, so it represents a definite improvement when searching for this type of information over the masses of results retrieved in standard Google searches. The next questions for academic information professionals to consider are the value of Google Scholar and its potential usefulness. How well does Google Scholar cover various subject disciplines? How does it compare with the standard databases in those fields? Can it be recommended to users as a trusted resource that will give them the information they are seeking?

This paper explores these questions for the field of engineering. It presents the results of a sampling study comparing Google Scholar's coverage to that provided by Compendex, the most highly regarded indexing database in the field. Subject searches were performed in Compendex in a number of major sub-disciplines and a random sample of the results were then searched in Google Scholar as a test of coverage. The study covers 1950 through 2007 by decade, to highlight differences over time. The results are presented and analyzed, and the implications for users are discussed.

Literature Review

It was obvious from the start that Google Scholar would have an impact on reference and information literacy[1] and on open access content[2]. Studies of Google Scholar have varied in their approaches, differing primarily over the analysis of the user interface functionality or of the content covered by the search engine. Of primary concern to many information professionals is the "discoverability" of information, so most research has examined Google Scholar's results against those of other databases.

Citation databases are vital tools for librarians and academics, and Google Scholar can also be utilized as a citation database. Jacso[3] found deflated, inflated, and phantom citation counts in results from Google Scholar searches in comparison to Web of Science and individual journals. Bauer[4] quantitatively compared citation counts from Web of Science, Scopus and Google Scholar in 1985 and 2000 for articles of *Journal of the American Society for Information Science and Technology*. Kousha[5] studied Web of Science and Google Scholar citation counts from 108 open access journals published in 2001.

The precision, recall, or relevance of search results from Google Scholar compared to scholarly databases has also been explored. Jacso[6] performed keyword searches in Google Scholar side by side with native interface of *Nature*, *Science* and the Astrophysics Data System and compared the number and quality of results. Notess tested Google Scholar against Scirus and counted total hits as well as behavior and coverage claims. Robinson[7] compared precision, recall and "Ability to Retrieve Top Ranked Pages" of Google Scholar with other search engines Ask.com and Yahoo! Brophy[8] did a case study approach with four subjects (education, environmental science, law and music) where Google results were compared with appropriate library databases on a number of qualitative measurements. Gardner[9] tested Google Scholar in the social sciences comparing the overlap between and relevancy of results from a single keyword query against PsycINFO, ERIC, and ISI SSCI. White[10] searched multiple subjects from the various databases to Google Scholar by comparing results of varied keyword searches.

In a “brute force” study by Mayr[11] each title from five separate journal lists (9,500 titles in all) were searched in Google Scholar. The first 100 hits recorded were analyzed to determine if the journal was present, what type of Google Scholar records existed, and which web servers (often publishers) provided the data. The study found the strongest results in percentage of titles covered and links in the journals from Thompson ISI’s Science Citation Index.

An increasing number of studies directly compare the coverage of Google Scholar to traditional library resources. In the most ambitious study to date Neuhaus[12] used 47 databases across multiple subjects and found Google Scholar to be particularly strong in science and medical databases. A sample of 50 records from the entire contents of each database were queried in Google Scholar to determine the percentage of coverage. Related studies of publication date and foreign language bias were performed in PsycINFO alone as was a study of upload frequency in BioMed Central and PubMed.

Jones[13] evaluated eight databases (BasicBIOSIS, ArticleFirst, ECO, ProQuest, WilsonWeb, SciFinder Scholar, HighWire, and MEDLINE) and Google Scholar as acceptable alternatives to the premiere biology abstracting service BIOSIS. A single search term was used to test the quantity of articles retrieved and results were compared for overlap in coverage since 1969, 1991, and 1996. Jones found significantly more results in Google Scholar for the past decade as well as more non-English journal articles. Recommendations are also made on which two of the eight other science databases can be used for an effective search.

Christianson[14] studied in detail the access to articles in the multidisciplinary field of ecology through Google Scholar. Seven levels of access both on and off campus were assigned to a random sampling of articles from 1945 to 2005 from journals in Cambridge Scientific Abstract's *Serials Source List for Ecology Abstracts*. Barriers to full text access were charted and a list of publishers was broken down in detail. Availability of citation or better results in Google Scholar is shown year by year, demonstrating a clear increase in recent decades.

Walters[15] also addressed a multidisciplinary topic from the field of social sciences, later-life migration. Building on his previous work in creating a core list of articles spanning 1990-2000, Walters compared Google Scholar against Academic Search Elite, AgeLine, ArticleFirst, GEOBASE, POPLINE, Social Sciences Abstracts, and Social Sciences Citation Index. The coverage of Google Scholar was found to be "comprehensive" despite the unsophisticated appearance of its records. Walters enumerates in detail the level of access, coverage by year, and results by publisher.

Most recently Levine-Clark[16] compared Google Scholar directly with the Chemical Abstract Service's SciFinder Scholar using one author, two topic, and two compound searches. The total quantitative results were compared over time with increasing granularity for recent years. The authors also broke down their results by type of resource, not limiting the study to journal articles. Google Scholar performed "better" for searches on chemistry topics while providing a "worthwhile substitute" for compounds and personal name searches.

Indexing Scope

The literature of engineering is fairly complex. Journal articles form the backbone of the literature, but conference proceedings and technical reports are also important vehicles for presenting research results and new advances. Industrial standards, patents, company catalogs, and software manuals are also useful to engineering students, researchers, and practicing engineers. Compendex is generally considered the premier database in engineering for the international coverage of journal articles and conference papers[17]. It has included a number of technical reports and other types of literature over the years, but its primary focus is on journals and conference papers. Compendex provides coverage of the literature back to 1884. The materials indexed by Compendex have been well documented over the years in the print volumes of *Engineering Index* and in the separate volume titled *Publications Indexed for Engineering*. Covered items are now listed online in the serial title indexes on Compendex. According to the Engineering Information Inc. website, Compendex covers 5,600 journals and conference proceedings, and 650,000 new records are added to the database each year. The indexed items come from materials published in 55 countries. There are now over 10 million records in the Compendex.

Google Scholar's indexing practices are more difficult to pin down. Google describes their sources as scholarly publishers, professional societies, and the Web. The list of partner publishers is not available, but fairly long lists have been compiled by researchers. Blackwell, Taylor & Francis, Springer, Cambridge, Wiley, Sage, Emerald, Nature Publishing[18] are known and some new publishers have opened up to Google indexing, in particular Elsevier's Science Direct content.[19] Notable Professional societies and non-profits include Association for Computing Machinery, IEEE, American Institute of Physics, Royal Society of Chemistry, BioMed Central, and the Public Library of Science. Google utilizes a proprietary technology to index articles from all sources that relies partially on the layout of a document for extracting metadata and citations.[20]

Google does give guidelines to publishers that indicates coverage of certain types of documents. It does not include trade journals or magazines unless they are proven "suitable primarily for a scholarly audience".[21] Google Scholar does not index monographs, instead directing publishers to the Google Book Search. However results from both Google Book Search and Google Patents are included in Google Scholar's results.

Dissertations and technical reports are other formats specifically included. Publishers are only required to supply abstracts, but are encouraged to allow indexing of their full text. From a technical standpoint, Google considers each file indexed as a separate article and can handle PDF files with OCR, optical character recognition.

Methodology

This study attempts to determine the relative coverage of the engineering literature provided by Google Scholar. In other words, how many of the publications indexed by Compendex are also "discoverable" in Google Scholar. We are defining a discoverable record as one that shows up as a normal record in Google Scholar search results or as a citation. Citations can appear in two ways in a Google Scholar search. Many are noted as "[citation]" when they are taken directly from the reference list of a publication. Cited articles that weren't clearly identified as such by the Google software can appear as fragments in the retrieved search records, or only appear in the linked source of indexing. They may also appear as "snippets of the context of the full text matching the query".[22] For the purpose of this study, the presence of a record in any of these formats qualifies as a match or "hit".

The initial research plan was to retrieve random records from across the entire Compendex database, but the accession numbering patterns of the records eliminated this option for practical purposes. It was decided instead to perform eight subject searches back to 1950 in representative areas and select a random sample of 20 records from each decade from each group of search results. The sampled references from each discipline were then searched in Google Scholar. Searches were done in aeronautical, civil, computer, electrical, environmental, industrial, mechanical, and nuclear engineering. Common topics were chosen that spanned the decades under study (the 1950s through the 2000s). Search result totals were kept to around 2000 or less per decade by progressively adding additional keywords, shown in italics below. This was done to keep the sampling process more manageable.

Subject search keywords:

aeronautical: aerodynamics AND wing *AND angle*

civil: buildings AND reinforced concrete *AND dynamics*

computer: computer AND memory *AND storage AND density*

electrical: microwave AND communications *AND antenna*

environmental: water AND supply AND quality

industrial: automobile AND assembly

mechanical: machinery AND vibration *AND analysis*

nuclear: nuclear AND reactor AND stability

Once a subject search was done for a decade and the size of the results list was known, the Web site Randomizer.org (http://www.randomizer.org/) was used to generate a string of random numbers. Records from the search results with numbers corresponding to the random numbers were retrieved and searched in Google Scholar. The selected references were thoroughly searched in Google Scholar. The most productive strategy was to search for title word strings, enclosed in quotes. Partial title strings were also useful, as were author/title keyword and author/author searches. Despite our attempts at precision retrieval, some searches in Google Scholar turned up dozens of potentially interesting records, which then had to be reviewed for relevance. Often, the target references were located by following the title link of promising search result records. One difficulty encountered in Google Scholar was the presence of misspelled words and incorrect letter spacing that might have resulted from poor scanning or optical character recognition processing.

Results

The results of the Google Scholar searches are shown in Table 1. The percentages reflect the number of sampled Compendex records that were also found in Google Scholar. Some overall patterns are apparent in these figures - in general, the number of matches in almost all of the subject categories increases with the more recent decades. Coverage during the 1950s range from 5% in mechanical and industrial engineering to 70% in nuclear, with 33% being the average hit rate for the decade. By the 2000s, the coverage for the sample records range from 70% in industrial engineering to 100% for civil and electrical engineering. Discipline-wide, the coverage peaks at 89% in the 1990s, with a slight drop to 88% in the 2000s. The data indicates that the relative coverage of the sampled records is fairly strong in all subject areas for the two most recent decades, ranging from 70% to several 100% match rates in those years.

[TABLE 1 HERE]

The strongest results appear in the aeronautical and nuclear engineering areas, which is likely due to the presence of large numbers of bibliographic records on the Web for these subjects. There are two publicly available databases that cover the literature in these areas -- the NASA Technical Report Server and the Department of Energy's Energy Citations Database. The NASA database concentrates on technical reports, but also includes conference papers and journal articles. Energy Citations Database covers the nuclear literature broadly, indexing all formats of publications. The records in both of these databases are available to Google Scholar for indexing, thus contributing to enhanced coverage of these subject areas.

By contrast, the coverage of the sampled records in the earlier decades was the weakest in industrial, mechanical, and civil engineering. An analysis of the materials not found in Google Scholar (Table 2) reveals that the trade journals accounted for the largest category of missed items in the 1950s, 1960s, and 1970s. The 1950s coverage in mechanical and industrial engineering was particularly weak. This weakness may be attributable to our selection of search topics as well as to Google Scholar indexing practices. Trade journals are an important information format for the practicing engineer, supplying news in the field on developments, new techniques, and solutions to problems. However, a review of our sample records for mechanical engineering in the 1950s shows that 95% (19/20) are for trade journals. This proportion is clearly not representative of the literature of the 1950s as an informal scan of *Engineering Index* (the print equivalent of Compendex) contents for the period shows. The keyword combination of "machinery" and "vibration" probably contributed to the disparity by retrieving an inordinate amount of articles describing fixes to common vibration problems encountered in industry. A review of our retrieved sample records in industrial and civil shows a similar records pattern leaning towards news and solution-oriented trade journal articles for the earlier decades. These factors combined with Google Scholar’s focus on scholarly rather than trade publications probably led to the low matching rates in these categories.

[TABLE 2 HERE]

Conference papers made up a large portion of the missed references for the 1980s through the 2000s. This could be a potential problem area for research-level users since conferences are a well respected venue for the exchange of research findings. The number of missed scholarly journal articles varied throughout the years, forming an average of 21.7% of these materials. Technical reports made up the smallest portion of materials not retrieved at 3.4%. This is not unexpected since this format is not heavily indexed by Compendex. For all types of materials, those in a language other than English constituted about 11.3% of the missed items across the study.

Analysis

The main purpose of this study was to determine the relative coverage of the engineering literature that is being provided by Google Scholar, based on a sampling of records retrieved from Compendex by test subject searches. Taken in their entirety, the results show that Google Scholar is capturing the records of many of the same publications that can be found in Compendex, particularly in the more recent decades (Figure 1). Weaker coverage in the earlier decades is likely due to the smaller numbers of records from that period existing in digital form on the Web that can be indexed by Google Scholar. We expect that our subject search-based methodology would provide similar results in most of the subject areas in Compendex, based on the trends of increasing coverage that were evident in each of the diverse areas tested in the study.

[FIGURE 1 HERE]

Content-based studies of Google Scholar have shown varying degrees of coverage, with the more promising results usually coming in the sciences. In the life sciences, a comparison of BIOSIS records that were also indexed by Google Scholar[23] showed that Google Scholar covered an increasing percentage of the records for the more recent years. That study indicated that the highest coverage (39%) occurred from 1991 to the present, while our study gave a figure of 89% for 1990 to 2007. A recent study by Walters[24] found that Google Scholar provided excellent coverage (93%) of a select set of core references in a particular topic in sociology published between 1990 and 2000, even though only the first 30 hits from Google Scholar were examined. Interestingly, Google Scholar coverage was found to outperform that of seven other well-known databases covering the subject area. Christianson's study[25] of the presence of sample articles from core ecology journals from 1945 to 2005 found that 77% could be considered indexed by Google Scholar. This is a slightly higher percentage than that found in our study which focuses approximately on the same period. However, the Christianson study was limited to journal articles only while this study looks at all publication formats. A 2004 study[26] of a random sampling of the entire contents of multiple databases found that 76% of the sample references in the sciences and medicine could also be located in Google Scholar. The coverage in humanities, education, business, and the social sciences was somewhat lower. A recent study by Levine-Clark[27] tested Google Scholar as a subject resource for chemistry and compared it to SciFinder Scholar. The study looked at the coverage of 702 records retrieved as a result of five searches done in both indexes. It found that SciFinder Scholar did a better job indexing pre-1989 papers while Google Scholar performed better for papers published after 1990.

From a searcher's perspective, the nature of the records located in Google Scholar might impact search results. Many of the matching records located in Google Scholar were information "thin"- they were basic bibliographic records or fragments of basic records. Quite a few of the Google Scholar records lacked the abstracts and assigned subject indexing terms that are commonly included in commercial databases such as Compendex. A keyword-based search should retrieve more records from keyword-rich databases than from one with bibliographic records alone. However, title words are generally descriptive of the major thrust of a paper and can be an effective retrieval mechanism. The fact that Google Scholar indexes portions of the full text of papers that it can reach might help ameliorate this also. The thin records may cause a serious problem in the identification and acquisition of the actual papers. Fragmented records in Google Scholar could prove challenging for many searchers to locate, especially if they don't have access to the standard databases in the subject area for verification purposes.

It appears that Google Scholar is doing an effective job covering non-English language materials. As noted earlier, 11.3% of all the missed materials were in a language other than English. To measure the language content of the subject areas that were studied, we searched the first two base keywords or keyword phrases in each of our subject areas in Compendex and compared the "all language" results to the "English" only results. For all publications appearing between 1950 and 2007, the range of non-English language content in Compendex varied between 10.8% and 28.8% for the disciplines. The average amount of non-English materials was 20.5% for those years. In this study, only 11.3% of the missed papers in Google Scholar were non-English. This might be interpreted that Google Scholar is doing reasonably well at covering the non-English language engineering literature.

Conclusions

The results of this study showed a close correspondence between the materials indexed by Google Scholar with those included in the Compendex database from the 1990s to the present. There may be differences in the richness of the record content between the services, but this might not be a barrier to positive search results for users at a certain level. Undergraduate engineering studies revolve around acquiring skills in mathematics, concepts in engineering and the physical sciences, and problem solving – literature searching is not a required part of many courses. Google Scholar seems well-suited to undergraduate assignments and research projects that call for the identification of a modest group of recent publications. Google Scholar can be promoted to students as a gateway tool for accessing the engineering literature, in tandem with presenting Compendex as a comprehensive backup to be consulted if they can't locate enough relevant information. Students are already familiar with the Google search interface and the simplicity that it offers. The jump from Google to Google Scholar may also be conceptually easier to grasp than the jump from Google to Compendex, INSPEC, and similar databases.

The easy-to-use interface of Google Scholar also has appeal for individuals with more advanced information needs such as graduate students, faculty, and research staff. However, these users might benefit more by doing their primary searching on Compendex and similar databases with a known scope of indexed materials and better retrospective coverage. These users could find that supplemental searches done on Google Scholar would uncover useful references falling outside of the normal coverage of the commercial tools. A repeat of the subject searches from this study in Google Scholar (using the same base keyword combinations) retrieved results sets that were larger for Google Scholar for each subject and decade than those retrieved in Compendex, except for the 1950s aerodynamics search. An analysis of the unique records retrieved would provide a further understanding of the overall content and scope of Google Scholar coverage in engineering. The citation indexing capabilities of Google Scholar will also be a powerful draw to researchers.

Acquiring experience with Google Scholar may serve the students well after they graduate since not all engineering firms have the resources to supply their employees with access to commercial databases. For engineers in many companies, Google Scholar may offer the best way to review the publishing activity in their fields. A recent study by Napp[28] showed that access to commercial engineering or science databases is rare at most engineering design firms. The same study pointed out that upper-level management at these firms expect that their engineers will be able to find the information that they need on their own, agreeing with the findings of an earlier paper by Rodrigues.[29] Freely available Web services such as Google Scholar may prove to be a valuable resource for professional engineers in this situation.

Google Scholar is a useful new tool for accessing the engineering literature published in the last ten to fifteen years. As with Google, it provides a simple interface and quick results. Its value will increase with the number of publishers that partner with it. Scholars would also benefit if the partnerships were publicized. The service has room for improvement in the unfinished and incomplete nature of some of the records it returns, but it does provide users with another valuable resource for exploring the scholarly activity in engineering.

Bibliography

1. Martin Kesselman & Sarah Barbara Watstein, "Google Scholar and Libraries: Point/Counterpoint," *Reference Services Review* 33/4 (2005): 380-387.

2. Frederick J. Friend, "Google Scholar: Potentially Good for Users of Academic Information," *Journal of Electronic Publishing* 9 (Winter 2006).

3. Peter Jacso, "Deflated, Inflated and Phantom Citation Counts," *Online Information Review* 30/3 (2006): 297-309.

4. Kathleen Bauer & Nisa Bakkalbasi, "An Examination of Citation Counts in a New Scholarly Communication Environment," *D-Lib Magazine* 11 (September 2005).

5. Kayvan Kousha & Mike Thelwall, "Google Scholar Citations and Google Web/URL Citations: A Multi-Discipline Exploratory Analysis," *Journal of the American Society for Information Science and Technology* 58 (May 2007): 1055-1065.

6. Peter Jacso, "Google Scholar: The Pros and the Cons," *Online Information Review* 29/2 (2005): 208-214.

7. Mary L. Robinson & Judith Wusteman, "Putting Google Scholar to the Test: a Preliminary Study," *Program: electronic library and information systems* 41/1 (2007): 71-80.

8. Jan Brophy & David Bawden, "Is Google Enough? Comparison of an Internet Search Engine with Academic Library Resources," *Aslib Proceedings* 57/6 (2005): 498-512.

9. Susan Gardner & Susanna Eng, "Gaga Over Google? Scholar in the Social Sciences," *Library Hi Tech News* 22 (September-October 2005): 42-45.

10. Bruce White, "Examining the Claims of Google Scholar as a Serious Information Source," *New Zealand Libraries* 50 (October 2006): 11-24.

11. Philipp Mayr & Anne-Kathrin Walter, "An Exploratory Study of Google Scholar," *Online Information Review* 31/6 (2007): forthcoming

12. C. Neuhaus, E. Neuhaus, A. Asher & C. Wrede, "The Depth and Breadth of Google Scholar: An Empirical Study," *portal: Libraries and the Academy* 6 (April 2006): 127-141.

13. D. Yvonne Jones, "Biology Article Retrieval from Various Databases: Making Good Choices with Limited Resources," *Issues in Science & Technology Librarianship* 44 (Fall 2005).

14. Marilyn Christianson, "Ecology Articles in Google Scholar: Levels of Access to Articles in Core Journals," *Issues in Science & Technology Librarianship* 49 (Winter 2007).

15. William Walters, "Google Scholar Coverage of a Multidisciplinary Field," *Information Processing and Management* 43 (2007): 1121-1132.

16. Michael Levine-Clark & Joseph Kraus, "Finding Chemistry Information Using Google Scholar: A Comparison with Chemical Abstracts Service," *Science & Technology Libraries* 27/4 (2007): 3-17.

17. Bonnie A. Osif, *Using the Engineering Literature* (London, UK: Routledge, 2006), p.21.

18. White, "Examining the Claims," p. 14.

19. Peter Brantley, "Science Direct-ly into Google," O'Reilly Radar , posted on July 3, 2007, http://radar.oreilly.com/archives/2007/07/science_directl.html

20. Google, "Support for Scholarly Publishers," http://scholar.google.com/intl/en/scholar/publisher.html#tech6.

21. Ibid., http://scholar.google.com/intl/en/scholar/publisher,html#faq5.

22. Jasco, "Google Scholar: The Pros and the Cons," p.209.

23. Jones, "Biology Article Retrieval from Various Databases," p.6.

24. Walters, "Google Scholar Coverage of a Multidisciplinary Field," p.1125.

25. Christianson, "Ecology Articles in Google Scholar."

26. Neuhaus, "The Depth and Breadth of Google Scholar," p.132.

27. Levine-Clark, "Finding Chemistry Information Using Google Scholar," p.12.

28. John B. Napp, "Survey of Library Services at Engineering News Record's Top 500 Design Firms: Implications for Engineering Education," *Journal of Engineering Education* 93(July 2004): 247-254.

29. Ronald J. Rodrigues, "Industry Expectations of the New Engineer," *Science & Technology Libraries* 19(2001): 179-188.

Table 1
Compendex Record Matches in Google Scholar (by decade)

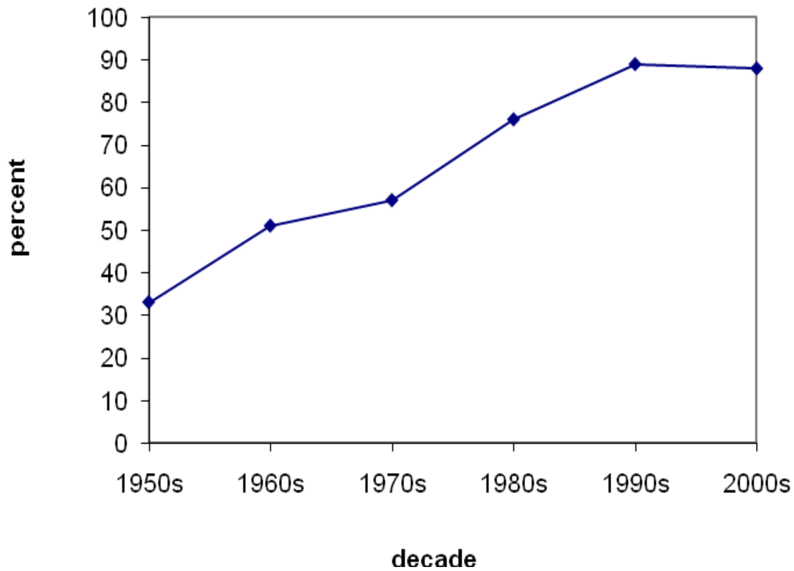
	1950s	1960s	1970s	1980s	1990s	2000s	Average
Aeronautical	50% (10/20)	75% (15/20)	95% (19/20)	95% (19/20)	100% (20/20)	90% (18/20)	84% (101/120)
Civil	15% (3/20)	25% (5/20)	40% (8/20)	65% (13/20)	90% (18/20)	100% (20/20)	56% (67/120)
Computer	45% (9/20)	75% (15/20)	55% (11/20)	65% (13/20)	100% (20/20)	85% (17/20)	71% (85/120)
Electrical	35% (7/20)	35% (7/20)	60% (12/20)	80% (16/20)	100% (20/20)	100% (20/20)	68% (82/120)
Environmental	40% (8/20)	50% (10/20)	55% (11/20)	80% (16/20)	70% (14/20)	95% (19/20)	65% (78/120)
Industrial	5% (1/20)	25% (5/20)	35% (7/20)	65% (13/20)	90% (18/20)	70% (14/20)	48% (58/120)
Mechanical	5% (1/20)	30% (6/20)	35% (7/20)	75% (15/20)	80% (16/20)	80% (16/20)	51% (61/120)
Nuclear	70% (14/20)	90% (18/20)	80% (16/20)	80% (16/20)	85% (17/20)	85% (17/20)	82% (98/120)
Overall	33% (53/160)	51% (81/160)	57% (91/160)	76% (121/160)	89% (143/160)	88% (141/160)	66% (630/960)

Table 2

Characteristics of Unmatched Records

	1950s	1960s	1970s	1980s	1990s	2000s	Average
Trade Journals	71.9% (77/107)	47.4% (37/78)	45.7% (32/70)	23.7% (9/38)	18.8% (3/16)	50.0% (9/18)	51.1% (167/327)
Scholarly Journals	16.8% (18/107)	28.2% (22/78)	25.7% (18/70)	15.8% (6/38)	37.5% (6/16)	5.6% (1/18)	21.7% (71/327)
Technical Reports	5.6% (6/107)	10.3% (8/78)	4.3% (3/70)	2.6% (1/38)	0% (0/16)	0% (0/18)	3.4% (11/327)
Conference Papers	2.8% (3/107)	11.5% (9/78)	21.4% (15/70)	52.6% (20/38)	43.8% (7/16)	44.4% (8/18)	18.9% (62/327)
Other (books, etc)	2.8% (3/107)	2.6% (2/78)	2.9% (2/70)	5.3% (2/38)	0% (0/16)	0% (0/18)	4.9% (16/327)
Non-English (All Types)	7.5% (8/107)	14.1% (11/78)	18.6% (13/70)	2.6% (1/38)	18.8% (3/16)	5.6% (1/18)	11.3% (37/327)

Figure 1
Google Scholar Coverage in Engineering



Decade	**% Covered**
1950s	33% (53/160)
1960s	51% (81/160)
1970s	57% (91/160)
1980s	76% (121/160)
1990s	89% (143/160)
2000s	88% (141/160)