Geoscience Information Society Proceedings, v. 39, 2008, p. 63-66

GEOREF, ISI WEB OF KNOWLEDGE, GOOGLE SCHOLAR - WHAT IS THE FUTURE FOR ABSTRACTING AND INDEXING SERVICES IN THE GEOSCIENCES?

Mary W. Scott
Geology Library
Ohio State University
scott.36@osu.edu

Abstract — Access to previous results of research is basic to all research. Recent articles on information resources in high-energy physics and engineering have raised questions about the relevance of commercial abstracting and indexing services in those fields. Do the same questions apply to the geosciences? What is the first choice for students and researchers for searching today? Preliminary results from a survey of faculty and students suggest that GeoRef is not the first place they look.

INTRODUCTION

In April 2008 two papers were published, *Information Resources in High-Energy Physics* and *Google Scholar's Coverage of the Engineering Literature: An Empirical Study*. Both these papers discuss the role of the Internet in improving access to scholarly information. In the field of high-energy physics a survey of about 10% of researchers revealed almost no use of commercial services. Their primary sources for information were the community-based services such as the arXiv and SPIRES systems. These services have made the content available for harvesting which benefits users of Google or Google Scholar. (Gentil-Beccot et al., 2008) The second paper compared Google Scholar's coverage to Compendex, a commercial service. The conclusion of this study was that Google Scholar is a "useful new tool for accessing the engineering literature published in the last ten to fifteen years." (Meier and Conkling, 2008)

At the time these two papers were published and discussion on listservs about them started, I was looking at a serial budget problem, reviewing undergraduate research project posters, and editing entries for the Bibliography of Ohio Geology, a new online database the Ohio Geological Survey is developing. These combined events caused me to wonder about geoscience databases:

- What were my faculty and students using?
- Was there anything I could cancel?
- What was the future of commercial databases?
- What about the niche databases such as Ohio Geology?

There are several papers reporting on different aspects geoscience databases. Two of them are comparisons of Google and GeoRef. (Tahirkheli, 2003; Musser and Fletcher, 2008) The most recent by Linda Musser is being presented as a poster at this meeting. Others have dealt with content analysis by comparing serial lists. (Scott, 2004; Scott, 2003) Lura Joseph compared retrieval performance of several databases for Quaternary research. (Joseph, 2007)

SURVEY

A survey of the faculty and graduate students in my department revealed that most were using the ISI Web of Knowledge℠-Web of Science® database as their first choice. I also asked a colleague at another university in Ohio to do the same survey. Those returns indicated a preference for GeoRef. The survey was simple: only a list of databases and a request for them to indicate their first choice.

The databases included in the survey are listed below. (For descriptions from the various database information pages see Appendix A.)

Academic Search Complete
American Geophysical Union Digital Library
Arctic & Antarctic Regions
BIOSIS Previews
Chemical Abstracts/ SciFinder Scholar
Compendex.
EJC-OhioLINK Electronic Journal Center
Environment Complete
Environmental Sciences and Pollution
GEOBASE
GeoRef database
GeoScience World
Google Scholar
IEEE Xplore
ISI Science Citation Index
Science Direct
Scirus

SURVEY RESULTS

Responses were fairly good considering it was summer and geologists are not typically around. I had a 33% return from the faculty and about 50% from the graduate students for my department and about a 50% return from the faculty at the other school. The combined totals are summarized here (see Figure 1).

GeoRef and ISI are equal but there are 5 others that were first choices for 15 people. This was a very small sample group but often collection managers in libraries need to make quick decisions about serial cuts with very little data. So I decided to look at this small sample to see what I could learn and to see if a larger survey was warranted.

ANALYSIS

The references cited in recently published papers by faculty from both schools were searched in GeoRef and ISI databases and the results analyzed. (A list of the papers is attached as Appendix B.) A total of 828 references were searched. Of these 655 were found in GeoRef and 506 in ISI. There were 80 references that did not appear in either. While 245 were unique to GeoRef and 82 were unique to ISI.

Database	Number of Positive Responses
GeoRef®	13
ISI Web of Knowledge℠-	13
Electronic Journal Center (OhioLINK)	6
Google®	5
Academic Search Premier	1
IEEE	1
SciFinder Scholar	1
GEOBASE®	1

Figure 1 Survey Results

References from two papers, a total of 118 references, were also searched in Google. A distinction was made between finding the full text of the article or just a citation for it. Of the 118 references, 67 links to the full text were found, and 30 links to a citation for the article. These two articles had 8 references that were not in either GeoRef or ISI. All 8 of these were found by Google, but Google did not have 12 of the references that were in GeoRef or ISI.

The references cited that were in GeoRef and not indexed in ISI included primarily monographs or chapters in monographs. Many of these were special publications of societies. The second largest group was references in journals not covered by ISI. References to articles in foreign publications were another large group. Also included in the list were references to government publications, the USGS as well as state geological surveys; abstracts; guidebooks, conference proceedings; theses and dissertations; and maps. This list includes document types that ISI does not index.

The references cited that were in ISI and not in GeoRef were in a variety of journals but could be grouped into broad subjects of physics, biology, environment, remote sensing, and chemistry --

subjects or journals normally included in GeoRef.

The subjects of the papers included paleontology, physics, remote sensing, climate change, geochemistry, groundwater, stratigraphy, tectonics, sedimentology, marine biology, and ecology. The range of subjects indicates the interdisciplinary nature of geoscience research today. In turn this requires access to multiple databases. This can be a challenge to support with today's library budgets. It can also be a challenge for library bibliographic instruction. Many researchers do not know about all the various options, or are they aware of what is included or not included in any given database. They also do not want to take the time to search more than one or at the most two databases.

CONCLUSIONS

1. There is not one database that provides everything. Even Google does not index everything. So I will need to continue to provide access to multiple databases, at least for a few more years. The future for geoscience databases should include some expansion of subject coverage to reflect the wider definition of geoscience research. As more commercial publishers and also societies provide open access to at least the table of contents and maybe abstracts for their journals, Google Scholar and other web search engines might reach a point of becoming the primary database. I say "might" because we are a long way from this happening. The active journals may soon be there, but there are a lot of dead journals that contain a great deal of important information. Databases such as GeoRef provide access to some of this information. Federated or multi-database searching is one option, but often the list of databases that can be included in a search is not comprehensive. As nice as the idea of federated searching is, there are problems such as duplicate records and indexing differences between databases. Development work is still needed to make this a good option.

2. Are GeoRef, ISI and other commercial databases still important in the geosciences? Yes, I believe they are. The questions are how many do we need? and which ones should we support? From this survey I feel I need to continue support for GeoRef and ISI on my campus. But the other databases on my list may need more review.

3. I need to do a better job of educating my faculty and students about the various databases, what they index, and when they might need to use more then one. This does not mean I will change their preferences, but they will be more knowledgeable when they make their decision.

4. The quality of a database depends in part on feedback from librarians. If you discover errors, you should report them.

5. Would a larger survey be useful? I don't think so. I am not sure that we would learn anything different.

6. The activity of searching all the references from the various faculty papers was interesting. It provided a different view of their research and has given me some leads for collection development. I found—especially when I was searching in Google—a lot of new online free resources which I can have added to the library catalog. These were not necessarily resources the faculty had referenced: I just spotted them in the list of results and took time to look at them. I also got some ideas for areas that I should expand on the Geology Library web page.

7. The last question is about niche databases such as state or specific subject bibliographies: Should we continue to support and develop these? I have been involved in two of them, the first one for North Dakota, and now one for Ohio. I know they include historical material and regionally published material that is probably not in GeoRef or any other commercial database. They also can provide some more local or specialized indexing that a large database is not going to provide. However, since they are labor intensive to develop and maintain, I suspect their future is limited.

REFERENCES

Gentil-Beccot, Anne, Annette Holtkamp, Heath B. O'Connell; and Travis C. Brooks. Information Resources in High-Energy Physics. 2008 [cited May 29, 2008 2008]. Available from http://arxiv.org/abs/0804.2701.

Joseph, Lura. 2007. Comparison of retrieval performance of eleven online indexes containing information related to Quaternary research, an interdisciplinary science: Reference and User Services

Quarterly,v.47 (1), p.56-65.

Meier, John J., and Thomas W. Conkling. 2008. Google Scholar's coverage of the engineering literature: an empirical study: Journal of Academic Librarianship, v.34 (3), p.196-201.

Musser, Linda R., and L. B. Fletcher. 2008. Comparison of GeoRef and Google Scholar: Geological Society of America, Abstracts with Programs, v. 40 (6), p.469.

Scott, Mary W. 2004. Searching for current international geoscience information: Geoscience Information Society Proceedings, v. 35, p. 9-12.

Scott, Mary W. 2003. Status of bibliographic control of pre-1900 geoscience literature: Geoscience Information Society Proceedings, v. 34, p. 105-108.

Tahirkheli, Sharon N. 2003. Google vs. GeoRef: effective information retrieval: Geological Society of America, Abstracts with Programs, v. 35, (6) p. 359.

APPENDIX A

DESCRIPTIONS OF THE DATABASES INCLUDED IN THE SURVEY

(All information directly provided by the database producers)

Academic Search Complete is the world's most valuable and comprehensive scholarly, multi-disciplinary full-text database, with more than 5,990 full-text periodicals, including more than 5,030 peer-reviewed journals. In addition to full text, this database offers indexing and abstracts for more than 9,990 journals and a total of more than 10,400 publications including monographs, reports, conference proceedings, etc. The database features PDF content going back as far as 1887, with the majority of full text titles in native (searchable) PDF format. Searchable cited references are provided for more than 1,000 journals. There are 259 journal titles listed under the subject of geology.

American Geophysical Union Digital Library is a comprehensive collection of more than 100 years of Earth and space science research. The library contains more than 90,300 articles from all the journals published by AGU and will eventually include an additional ~25,000 articles from books and the weekly newspaper for AGU. The library starts with volume 1, issue 1 of *Terrestrial Magnetism* published in 1896.

Arctic & Antarctic Regions (AAR) is the world's largest collection of international polar databases. With over 1 million records from 1800 to the present, Arctic & Antarctic Regions covers a wide variety of sources from multiple disciplines. Many sources are indexed only in Arctic & Antarctic Regions making it the best resource for research on cold regions anywhere, from temperate regions with cold winters to the Himalayas of Tibet.

BIOSIS Previews, including Biological Abstracts, contains references to journal articles, reports, literature reviews, conference papers, patents, and book synopses in the life sciences. Coverage includes 5500 journals and 1500 international meetings in agriculture, biology, biotechnology, environment, wildlife, ecology, agriculture, forestry and the health sciences. The journal list includes 29 titles under Geology and 28 under Geosciences interdisciplinary.

Chemical Abstracts/ SciFinder Scholar is the largest and most comprehensive database of chemical literature in the world. It covers not only the core areas of chemistry, but also chemistry related sciences such as biotechnology, agricultural chemistry, toxicology and environmental science among others. SciFinder Scholar is an interface to four databases: Chemical Abstracts Plus, the Registry File, CASREACT and now MEDline.

Compendex is a comprehensive interdisciplinary engineering database, the electronic equivalent of the print Engineering Index. Compendex covers the entire spectrum of engineering, in depth, with abstracts from over 2,600 international journals, conference papers and proceedings, and technical reports.

EJC-OhioLINK Electronic Journal Center – Electronic Journals subscribed to by OhioLINK, it includes about 130 geology journals. The EJC provides a search by author, title, subject, keyword of all the journals or subject subsets of journals or a single journal.

Environment Complete offers deep coverage in applicable areas of agriculture, ecosystem ecology, energy, renewable energy sources, natural resources, marine & freshwater science, geography, pollution & waste management, environmental technology, environmental law, public policy, social impacts, urban planning, and more. *Environment Complete* contains more than 1,957,000 records from more than 1,700 domestic and international titles going back to the 1940s (including 1,125 active core titles). The database also contains full text for more than 680 journals and 120 monographs. There are 130 titles in the journal list under the subject Geology.

Environmental Sciences and Pollution Management provides unparalleled and comprehensive coverage of the environmental sciences. Abstracts and citations are drawn from over 6000 serials including scientific journals, conference proceedings, reports, monographs, books and government publications.

GEOBASE is a bibliographic database of the global literature in earth science, ecology, geography and marine science. The range of sources abstracted make this tool appropriate for searching multidisciplinary topics such as environmental or geographical studies and other areas that cross traditional subject boundaries.

GeoRef **database**, established by the American Geological Institute in 1966, provides access to the geoscience literature of the world. GeoRef is the most comprehensive database in the geosciences and continues to grow by more than 90,000 references a year. The database contains over 2.9 million references to geoscience journal articles, books, maps, conference papers, reports and theses.

GeoScience World - A comprehensive Internet resource for research and communications in the geosciences, built on a core database aggregation of over 40 peer-reviewed journals indexed, linked, and interoperable with the GeoRef index.

Google Named for the mathematical term "googol," Google is widely recognized as the "world's best search engine" because it is fast, accurate and easy to use. Google's breakthrough technology and continued innovation serve the company's mission of "organizing the world's information and making it universally accessible and useful."

Google Scholar finds scholarly literature (peer-reviewed papers, theses, preprints, abstracts, technical reports) from a wide variety of academic publishers, professional societies, preprint repositories and universities and across the web. Google Scholar also automatically analyzes and extracts citations and presents them as separate results, even if the documents they refer to are not online, so search results may include citations of older works and seminal articles that appear only in books or other offline publications.

IEEE Xplore provides full text access to IEEE & IEE journal articles and conference papers from 1988 to present; current IEEE standards; selected IEEE pre-1988 content; and IEEE periodicals cover-to-cover beginning in 2004.

INSPEC scans papers from approximately 4,200 journals, 1,000 conferences, and other publications, adding over 250,000 records each year. INSPEC is an excellent source of information on: Computing, Control Technology, Electronics, Electrical Engineering, Information Technology, Physics.

ISI Science Citation Index indexes 5,300 major journals across 164 scientific disciplines and contains searchable, full-length, English-language author abstracts for approximately 70 percent of the articles in the database.

Science Direct is the index to the Elsevier online journals.

Scirus is the most comprehensive scientific research tool on the web. With over 450 million scientific items indexed at last count, it allows researchers to search for not only journal content but also scientists' homepages, courseware, pre-print server material, patents and institutional repository and website information.

APPENDIX B

PAPERS USED IN THE STUDY

Bevis, Michael, Douglas Alsdrof, Eric Kendrick, Luiz Paulo Fortes, Bruce Forsberg, Robert Smalley Jr., and Janet Becker. 2005. Seasonal fluctuations in the mass of the Amazon River system and Earth's elastic response: *Geophysical Research Letters, v. 32*, L16308, doi:10.1029/2005GL023491.

Bonelli, Jr, James R., Carlton E. Brett, Arnold I. Miller, and J. B. Bennington. 2006. Testing for faunal stability across a regional biotic transition: quantifying stasis and variation among recurring coral-rich biofacies in the Middle Devonian Appalachian Basin: *Paleobiology, v.32* (1), p.20-37.

Carey, Anne E., Christopher B. Gardner, Steven T. Goldsmith, W. Berry Lyons, and D. Murray Hicks. 2005. Organic carbon yields from small, mountainous rivers, New Zealand: *Geophysical Research Letters, v. 32*, L15404, doi:10.1029/2005GL023159.

Chocyk-Jaminski, Marzena, and Craig Dietsch. 2002. Geochemistry and tectonic setting of metabasic rocks of the Gneiss Dome Belt, SW New England Appalachians: *Physics and Chemistry of the Earth, Parts A/B/C, v. 27* (1-3), p.149-167.

Freedman, Vicky L., and Motomu Ibaraki. 2003. Coupled reactive mass transport and fluid flow: issues in model verification: *Advances in Water Resources, v.26* (1), p.117-127.

Kamrer, Thomas W., and William I. Ausich. 2006. The "Age of Crinoids": a Mississippian biodiversity spike coincident with widespread carbonate ramps: *Palaios, v. 21* (3), p.238-248.

Kim, Jeongkon, Franklin W. Schwartz, Leslie Smith, and Motomu Ibaraki. 2004. Complex dispersion in simple fractured media: *Water Resources Research, v.40*, W05102, doi:10.1029/2003WR002631.

Krause, R. A.,Jr., and D. L. Meyer. 2004. Sequence stratigraphy and depositional dynamics of carbonate buildups and associated facies from the Lower Mississippian Fort Payne Formation of southern Kentucky, U.S.A.: *Journal of Sedimentary Research, v. 74* (6), p.831-844.

Lyons, W. B., Kathleen A. Welch, Anne E. Carey, Peter T. Doran, Diana H. Wall, Ross A. Virginia, Andrew G. Fountain, Bea M. Csatho, and Catherine M. Tremper. 2005. Groundwater seeps in Taylor Valley Antarctica; an example of a subsurface melt event: *Annals of Glaciology, v.40*, p.200-206.

Moreno, Patricio I., George L. Jacobson Jr., Thomas V. Lowell, and George H. Denton. 2001. Interhemispheric climate links revealed by a late-glacial cooling episode in southern Chile: *Nature (London), v. 409* (6822), p.804-808.

Mortimer, Estelle, Isabelle Coutand, Lindsay Schoenbohm, Edward R. Sobel, José Sosa Gomez, and Manfred R. Strecker. 2007. Fragmentation of a foreland basin in response to out-of-sequence basement uplifts and structural reactivation; El Cajon-Campo del Arenal Basin, NW Argentina: *Geological Society of America Bulletin, v. 119* (5-6), p.637-653.

Munk, LeeAnn, Gunter Faure, Douglas E. Pride, and Jerry M. Bigham. 2002. Sorption of trace metals to an aluminum precipitate in a stream receiving acid rock-drainage; Snake River, Summit County, Colorado: *Applied Geochemistry, v. 17* (4), p. 421-430.

Panero, Wendy R., and Lars P. Stixrude. 2004. Hydrogen incorporation in stishovite at high pressure and symmetric hydrogen bonding in δ-AlOOH: *Earth and Planetary Science Letters, v. 221* (1-4), p. 421-431.

Saltzman, Matthew R. 2005. Phosphorus, nitrogen, and the redox evolution of the Paleozoic oceans: *Geology, v. 33*(7), p.573-576.

Turgeon, Steven C., Robert A. Creaser, and Thomas J. Algeo. 2007. Re–Os depositional ages and seawater Os estimates for the Frasnian–Famennian boundary: Implications for weathering rates, land plant evolution, and extinction mechanisms: *Earth and Planetary Science Letters, v. 261* (3-4), p.649-661.