Comparison of Retrieval Performance of Eleven Online Indexes Containing Information Related to Quaternary Research, an Interdisciplinary Science

Lura E. Joseph

Lura E. Joseph *is Geology and Digital Projects Librarian, University of Illinois at Urbana-Champaign (UIUC). She holds a Master of Science degree in geology and a Master of Library and Information Studies, both from the University of Oklahoma. Joseph worked as a petroleum geologist for fifteen years, as Physical Sciences Librarian at North Dakota State University for six-and-a-half years, and has been Geology Librarian at UIUC for more than five years.* *Submitted for review August 2, 2005; revised and accepted for publication February 28, 2006.*

Interdisciplinary research offers increasing information challenges for researchers and scholars as well as for librarians. Quaternary research is an example of a highly interdisciplinary area incorporating research ranging from geochemistry and microbiology to planetary science. This study compares retrieval performance of eleven online indexes that can be used for Quaternary research, and discusses three others. Recall, precision, and overlap and uniqueness were analyzed using search results (12,896 records) from the eleven databases for the publication year 2000. A broad search strategy was used in order to recover most of the relevant information from the databases for the whole discipline for one publication year in order to avoid problems encountered when using sampling and example searches. Implications for interdisciplinary research in general are discussed, and federated searching is suggested.

Reference & User Services Quarterly, vol. 47, no. 1, pp. 56–65

Given the exponential increase of information, staying current in a particular discipline, verifying a particular citation, or conducting an exhaustive search for information on a particular topic can be daunting to researchers and scholars. Helping the researchers and scholars with their information quests can be taxing for librarians and information specialists. The switch from print to online indexes is both a blessing and a curse. On one hand, the online indexes uncover much more information, but on the other, sorting through the mountain of information can be frustrating and time consuming. Quaternary research is an example of a highly interdisciplinary area whose researchers, scholars, and supporting librarians are faced with these problems.

QUATERNARY RESEARCH: ITS NATURE AND IMPORTANCE

Quaternary research is the study of the Quaternary, which is the period of time that spans approximately the last 2.6 million years of the Earth's geologic history.[1] The Quaternary geologic time period includes the Pleistocene, sometimes known as the Ice Age, and the Holocene, which is the geologic epoch in which we live.[2] The Pleistocene covers the time period spanning approximately two million years to ten

thousand years ago, and the Holocene includes the period spanning approximately ten thousand years ago to the present. Scientists currently disagree regarding the exact boundaries and dates for the beginning of the Quaternary, and whether it should be retained as a formal chronostratigraphic unit (a body of rock officially recognized as a unit based on the age of its boundaries).[3] In fact, the most recent International Commission on Stratigraphy removed the term "Quaternary" from the International Geologic Time Scale and included that interval of time in the Neogene Period.[4] The terminology is still being hotly debated among the geologic community, but no matter what the time period is called, or when the exact initial boundary is set, that geologic period of time will undoubtedly remain an extremely important area of research.

The study of the Quaternary time period is particularly interdisciplinary.[5] Figure 1 shows many of the areas of specialty within Quaternary research. Quaternary researchers study ice cores; ocean sediments; ocean circulation; lake sediments; cycles of the earth and sun; atmosphere; fossils and modern plants and animals, including invertebrates, vertebrates, insects, and pollen; and other subjects. Current topics of importance include atmospheric, ocean, and terrestrial interactions and the building of testable computer models; for example, models that can be used to predict climate change, sea level rise, or ocean circulation patterns.

Figure 1: Many of the Areas of Specialty in Quaternary Research

Anthropology and archaeology
Biology
Botany
Climatology
Ecology
Entomology
Geochronology
Geography
Geology
Geomorphology
Geophysics
Glaciology
Hydrogeology
Isotope geochemistry
Limnology
Molecular genetics
Oceanography and oceanology
Paleoceanography and paleoceanology
Paleoecology
Paleontology
Palynology
Planetary geology
Soil Science and pedology
Structural geology and neotectonics
Volcanology
Zoology

During the Quaternary, the Earth has experienced frequent, extreme, and often abrupt climate and environmental changes, including the advance and retreat of continental and mountain glaciers.[6] These climate and environmental changes have been accompanied by global changes in plant and animal communities and include numerous extinctions and the evolution and dispersal of humans. No matter what the causes, Earth's climate currently appears to be warming, and significant and rapid environmental changes could challenge modern civilization. Current major goals of Quaternary research include documenting past climate patterns at various time scales in order to understand current trends and to predict future climate patterns and environmental effects.

In 1928, a group of scientists who were conducting interdisciplinary research into environmental changes that occurred during the glacial ages founded the International Union for Quaternary Research (INQUA).[7] International congresses are held every four or five years, with attendance as high as one thousand participants. There are more than thirty-five member countries worldwide that also hold regional and local meetings. Out of a total membership of 18,445 within fifteen divisions, the Quaternary Geology and Geomorphology Division of the Geological Society of America (GSA) had 2,695 members (14.6 percent) in 2004, further demonstrating the importance of Quaternary research. There are a number of excellent peer-reviewed journals devoted to Quaternary research, and many Quaternary research articles appear in other highly respected general and specialized scientific journals. Other publications include newsletters and monographic series. See *Associations and Information Resources for Quaternary Research* (www.library.uiuc.edu/gex/bibs/QuaternaryInformationResources.html) for further information about the various Quaternary research societies and publications.

THE RESEARCH QUESTION

This study compared the retrieval performance of eleven online indexes that can be used to find information pertaining to Quaternary research. The main purpose was to determine which databases contain the largest number of relevant references (recall and precision), and how much overlap exists between the databases. This information can be used to help researchers and librarians determine which databases are most appropriate for Quaternary research, which might safely be ignored (and possibly cancelled by librarians faced with budget cuts), and which should be included in metasearch (federated search) tools. Results also can be used to indicate which databases could best substitute for others that might not be available, and how much information would be missed if a particular database was not searched.

LITERATURE REVIEW

Research focusing on interdisciplinary, cross-disciplinary, and multidisciplinary information has continued in recent years. Ackerson discussed the challenges of research and teaching in cross-disciplinary areas, focusing on engineering, and including obstacles to intellectual and physical access to information.[8]

Other studies related to interdisciplinary research include database comparisons. These comparisons generally fall into two categories: descriptive comparisons, such as the number of journals indexed and the size, range, and subjects covered; and performance measures, such as recall, precision, and

novelty versus overlap. Starr reviewed the literature and recommended that future database comparisons should emphasize analytic research and study reasons for search results differences rather than concentrating on descriptive comparisons or performance measures.[9] Her criticism of the study of performance measures was that most results are based on a small number of sample searches and the studies do not use random samples. Also, the studies typically use only a few professionals as searchers, and therefore results cannot be generalized to other user groups. This study avoids those problems because it looks at the results of a search strategy encompassing the entire interdisciplinary area rather than sample searches within the area, and the search terms used are not unique to any particular user group within the discipline. The reasons for observed differences in search results between the databases, such as database differences, are not the focus of this study.

In the past ten years, database comparison research has continued. For example, a comparison of two business literature databases by Rey and Fereres resulted in the suggestion of simultaneously searching both databases.[10] McCain used a database filtering approach by combining citation, indexing, and productivity analysis to identify core and important noncore journals supporting multidisciplinary research and development in the area of biotechnology.[11] Hooper-Lane and Atkins examined the journal coverage, uniqueness versus overlap, and currency of three multidisciplinary databases in relation to marine science, and concluded that the multidisciplinary databases complemented the subject-specific databases and significantly improved recall and currency of returned results.[12]

Cavanagh compared the retrieval performance of five multidisciplinary tables of contents (TOC) databases with three specialized biological databases in terms of precision and recall.[13] The TOC databases had higher precision but lower recall than subject-specific databases. The chosen databases, as a group, identified only 75 percent of the articles known from independent sources.

Brown, Edwards, and LaSee-Willemssen conducted a "deep analysis of indexing" for two education indexes.[14] They found that one of the indexes covered twice the number of journals but only half the number of articles. Based on indexing policies, even so-called cover-to-cover indexing may be incomplete. McDonald, Taylor, and Adams analyzed coverage and overlap of databases indexing psychiatry journals at the journal rather than article level.[15] They identified 213 abstracting and indexing services and 977 psychiatry journals. Four databases accounted for 90 percent of all the psychiatry journals that were indexed; more than four hundred journals were not indexed at all. Thirty-five percent of the indexed journals were included in only one of the top four databases, emphasizing the need to search multiple databases. Sutton and Foulke compared eight databases (three subject-specific, two multidisciplinary, and three general academic) at the journal level for coverage of literature related to anthropology.[16] They found that coverage was not comprehensive, even by subject-specific databases.

Read and Smith compared three databases for library and information science using a subject profile technique, and found maximum overlap was only 21 percent, indicating the need to search more than one database if comprehensive results are desired.[17] Jatkevicius compared BIOSIS Previews, BasicBIOSIS, and Cambridge Scientific Abstracts' Biological Sciences Collection using the same forty-two searches for each database and analyzing recall rather than precision. BIOSIS Previews retrieved more than 67 percent of overall hits from the forty-two searches and was judged best for graduate biology students and faculty.[18]

Brettle and Long compared retrieval from six databases related to rehabilitation of people with severe mental illness, and found that 42 percent of the papers were found only in one database, signifying the importance of searching multiple databases.[19] Tellman compared retrieval of information of interest to historians from two online European journal indexes and found that significant numbers of articles would be missed if using only one of the indexes.[20]

Janke compared journal coverage and currency of indexing for Current Contents Connect (CCC) and PubMed for the subjects of agriculture, biology and environmental studies, clinical medicine, and life sciences, and found that, of the 801 journals jointly indexed by CCC and PubMed, for 15.5 percent CCC was less current, for 29.6 percent CCC was more current, and for 53.7 percent currency of indexing was the same.[21] Kristick compared Meteorological and Geoastrophysical Abstracts with eight other databases at the source title list level, and found the highest level of overlap was with Environmental Sciences and Pollution Management (ESPM)(59 percent), followed by GeoRef (53 percent). The combination of ESPM, GeoRef, and Chemical Abstracts yielded an overlap of 80.3 percent.[22] Walters and Wilder evaluated the effectiveness of twelve databases that index the literature pertaining to the multidisciplinary field of "later-life migration."[23] An article level analysis was used. They found that four multidisciplinary databases each provide better index coverage than any of the specialized subject indexes. There was a relatively low degree of overlap among the twelve databases, therefore multiple databases should be searched.

The research overwhelmingly indicates the inadequacy of using a single, or even several, databases to find comprehensive information, especially in interdisciplinary areas. Bar and Finkler advocated using a multidatabase, multidisciplinary search approach using databases from a single vendor.[24] Today it is possible to use the same multidatabase, multidisciplinary approach using multiple vendors. This procedure of using metasearch (also known as federated or broadcast search) technologies has been described by Mischo and Schlembach, and Luther; National Information Standards Organization (NISO) has developed standards for these tools.[25] When

creating metasearch tools, it is useful to know which databases to include for particular areas. This type of study can indicate which databases are most important to include. If federated searching is not possible, then the information can be used by researchers to select the most useful databases to search and by librarians to determine which databases to purchase or retain.

THE RESEARCH PLAN

Extensive association with Quaternary researchers while working as a librarian at North Dakota State University indicated that the most frequently used online indexes are GeoRef, Web of Science or Current Contents, and GEOBASE. An informal survey of an electronic list monitored by many of the Quaternary researchers supported these conclusions and revealed seven others that are consulted on a less regular basis. One other database of potential value was added to this analysis. The eleven indexes included in the comparison study are AGRICOLA, AGRIS, Aquatic Science and Fisheries Abstracts 3, Biological and Agricultural Index, Biological Abstracts, CAB Abstracts, Current Contents, GEOBASE, GeoRef, Water Resources Abstracts, and Zoological Record. For each of the eleven online indexes, the study compared the number of search returns, the number and percentage of relevant hits, and overlap versus uniqueness of content for the entire interdisciplinary area of Quaternary research. The appendix lists and describes the databases included in this study. Web of Science (WOS) was not analyzed because, at the time of the search, it was not possible to download the complete set of records retrieved by the search or to further break the search into parts (it is now possible to output five hundred records at a time). ISI, the producer of both WOS and Current Contents, gave the assurance that Current Contents can be considered as representative of both databases; however, due to factors such as differences in entry points (searchable fields) and update frequency, this is questionable.[26]

This research is practical in nature and focused on results. Rather than considering the databases themselves (scope, journals indexed, frequency of update, and so on), this study was devoted to examining the results of a broad search that could be conducted by individuals at the University of Illinois, Urbana-Champaign (UIUC), and that would retrieve most of the information related to Quaternary research. A more realistic search would likely include the broader search but focus it with one or two more specific terms.

METHOD

The search string "Quaternary or Pleistocene or Holocene" was used to search each of the eleven databases, and the search was limited to the publication year 2000. The publication year 2000 was selected because it was fairly recent; however, enough time had passed, hopefully, to allow indexing of most publications. All databases were searched during a limited time range (May 2–3, 2002) in order to have a realistic, comparative time snapshot of the content of the various databases. Material from nonjournal literature, such as from special publications, memoirs, conference proceedings and transactions, and government publications, are important resources in Quaternary research, and were therefore included.

In order to manage the large amount of data, the decision was made to populate an Access database with the search results. In order to populate the database, it was first necessary to parse the search results from each of the indexes. A grant was obtained from the UIUC Library Research and Publication Committee to hire a graduate assistant for the project. The assistant wrote computer programs to parse the search results from each of the eleven databases, populated an Access database with the results, and created a search engine to query the data. The parsing programs and search engine can be modified for future database comparison studies by library faculty at UIUC. Duplicates within indexes were found and removed from the database, resulting in a total of 12,896 records. Each of the records was examined by the author and designated either as "relevant" or "not relevant" to Quaternary research. The author is a geologist, and has worked with Quaternary researchers for a number of years, and is therefore qualified to determine relevancy. Relevance was interpreted in the broadest sense. In other words, if the article had anything to do with the geologic time intervals Quaternary, Pleistocene, or Holocene, the material was considered to be relevant. The term "quaternary" yielded varying numbers of irrelevant returns, depending on the databases. The term "quaternary" means "consisting of four things or parts; a set of four (things); belonging to the fourth order or rank" (Oxford English Dictionary). A number of scientific disciplines, such as chemistry and biochemistry, use it to refer to topics other than the geologic time period. After the database was populated and relevancy was noted, the database was queried to determine the number and percentage of relevant records from each online index. All subsequent queries were performed only on relevant records. Queries were conducted to determine the number of unique relevant records from each of the eleven commercial online indexes and also the overlap or uniqueness of the results.

RESULTS

Table 1 shows a total of 12,896 records, of which 11,168 (86.60 percent) are relevant. If all duplicates between indexes are removed, there are 6,459 unique relevant records. There are 4,023 relevant records that were contained in only one particular database. Table 1 also shows, from each index, the total number of records, number of relevant records, percent of relevant records, and the number of unique relevant records.

GeoRef contained the largest number of total records, followed by Current Contents, GEOBASE, and Biological Abstracts, with the remaining databases accounting for the rest. GeoRef also contained the largest number of relevant records, followed by GEOBASE, Current Contents, Biological Abstracts,

Zoological Record and the rest. Note that, based on relevant records, GEOBASE switched position with Current Contents for second place, while all the other databases maintained their position. There is quite a bit of rearrangement to the list when sorted by percent relevant: GEOBASE holds first place for highest percent of relevant records (99.51 percent), followed by GeoRef (99.29 percent), Zoological Record (97.81 percent), Water Resources Abstracts (97.06 percent), and the others in the list. Current Contents falls to ninth place, with 65.41 percent relevant. It is not surprising that GEOBASE and GeoRef contain a higher percentage of relevant records, as they are primarily geological databases and have a lower percentage of records related to chemistry or biochemistry, subjects that use "quaternary" for other than the geologic time period. GeoRef contained the highest number of unique records (2,889), followed by Current Contents, GEOBASE, Zoological Record, and the others. It is not surprising that GeoRef contains the highest number of unique records, as the database contained the highest number of total records and it indexes many nonjournal publications, such as government reports, book chapters, theses and dissertations, and abstracts. Nor is it surprising that Current Contents and GEOBASE ranked high in unique records, as they also contained a large number of total records. It is important to note that each database contained some unique records, information that would have been overlooked had a given database not been searched.

Table 2 shows the overlap between databases. Examining the overlap between databases may indicate possible substitute databases if a particular index is not available to the user. For example, table 2 indicates that Current Contents or Biological Abstracts might be the best substitutes for AGRICOLA, based on number of relevant records in common. Examination of table 2 indicates that for seven of the databases, Current Contents appears to be the best substitute in terms of number of relevant records in common; for three databases, GEOBASE appears to be the best substitute; and for one database each, Biological Abstracts and GeoRef appear to be best. For second best substitutes, GEOBASE is indicated for five databases, Current Contents and GeoRef are indicated for two each, and Biological Abstracts would be the second best logical substitute for one database. In all cases, Current Contents, GEOBASE, Biological Abstracts, and GeoRef are the first, second, or third best substitutes on the basis of number of relevant records in common (overlap). Please note that, in all cases, considerable information is lost by substituting with any one index.

It is useful to know the combined number of nonduplicate relevant records from the two databases with the highest number of relevant records—GeoRef and GEOBASE—and the percent of total nonduplicate relevant records this represents. In other words, if a researcher searched these two indexes, what percentage of the available relevant information available from the eleven databases would be recovered? The formula "GeoRef relevant records + GEOBASE relevant records – duplicates" was used for this calculation (4,592 + 2,251 – 1,418) (see figure 2). GeoRef and GEOBASE recovered a total of 5,425 nonduplicate individual records from relevant records possible, or 83.99 percent of the 6,459 total nonduplicate records.

The same sort of analysis was conducted with the addition of Current Contents, using the formula: "GeoRef relevant records + GEOBASE relevant records + Current Contents relevant records – (GeoRef/GEOBASE duplicates + GEOBASE/Current Contents duplicates + GeoRef/Current Contents duplicates) + GeoRef/GEOBASE/Current

Table 1. Results of Searches

Database	Total Records	Relevant Records	% Relevant	Unique Relevant Records
AGRICOLA	69	34	49.28	8
AGRIS	56	33	58.93	10
ASFA 3	33	29	87.88	3
Bio. & Ag. Index	71	51	71.83	13
Biological Abstr.	1,229	866	70.46	130
CAB Abstracts	291	235	80.76	50
Current Contents	3,359	2,197	65.41	325
GEOBASE	2,262	2,251	99.51	316
GeoRef	4,625	4,592	99.29	2,889
Water Res. Abs.	170	165	97.06	7
Zoological Rec.	731	715	97.81	272
Total	**12,896**	**11,168**	**86.60**	**4,023**

Table 2. Overlap and Summary of Databases

	AGRICOLA	AGRIS	ASFA3	Bio. Ag. Index	Biol. Abs.	CAB Abs.	Cur. Cont.	GEOBASE	GeoRef	Water Res.	Zoo. Rec.
	Relevant 34	Relevant 33	Relevant 29	Relevant 51	Relevant 866	Relevant 235	Relevant 2,197	Relevant 2,251	Relevant 4,592	Relevant 165	Relevant 715
	Unique 8	Unique 10	Unique 3	Unique 13	Unique 130	Unique 50	Unique 325	Unique 316	Unique 2,889	Unique 7	Unique 272
	Relevant Records in Common										
AGRICOLA	-	0	1	5	22	14	22	19	7	3	3
AGRIS	0	-	0	0	17	2	13	10	18	0	2
ASFA3	1	0	-	2	14	6	24	23	10	13	2
Bio. Ag. Index	5	0	2	-	27	12	37	31	17	0	18
Biol. Abs.	22 [1]	17 [2]	14 [3]	27 [3]	-	106 [3]	555 [3]	464 [3]	321 [3]	41	245 [3]
CAB Abs.	14	2	6	12	106	-	163	144	82	18	13
Cur. Cont.	22 [1]	13 [3]	24 [1]	37 [1]	555 [1]	163 [1]	-	1,518 [1]	1,257 [2]	147 [2]	288 [1]
GEOBASE	19 [3]	10	23 [2]	31 [2]	464 [2]	144 [2]	1,518 [1]	-	1,418 [1]	149 [1]	258 [2]
GeoRef	7	18 [1]	10	17	321 [3]	82	1,257 [2]	1,418 [2]	-	101 [3]	237
Water Res.	3	0	13	0	41	18	147	149	101	-	6
Zoo. Rec.	3	2	2	18	245	13	288	258	237	6	-

[1] Best substitute; [2] second best substitute; [3] third best (based on # of relevant records in common)

Contents duplicates" [4592 + 2251 + 2197 – (1418 + 1518 + 1257) + (1075)] (see figure 3). GeoRef, GEOBASE, and Current Contents recovered a total of 5,922 nonduplicate relevant records, or 91.69 percent of the 6,459 total nonduplicate records.

As the number of databases is increased, the formula becomes more complicated. Suffice it to say, if Zoological Record's 272 unique relevant records are added, the total of nonduplicate records rises to at least 6,194, 95.9 percent of the total 6,459 nonduplicate records (not considering any records that are duplicated by indexes other than GeoRef, GEOBASE, and Current Contents). Further, if Biological Abstracts is added to the mix, those 130 unique relevant records bring the total to at least 6,324, which is 97.91 percent of the total 6,459 nonduplicate records. Therefore, five indexes account for at least 97.91 percent of the nonduplicate records in this study. However, the 2 percent or fewer records that are not included in these five databases may contain information of value to the information seeker. Without searching all the possible databases, this information is unknown and essentially lost.

As a follow-up, the searches were run again for ten of the eleven data-

Figure 2: GeoRef and GEOBASE Relevant Records

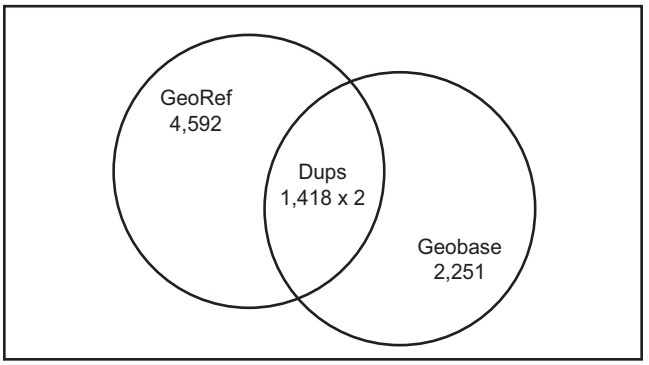
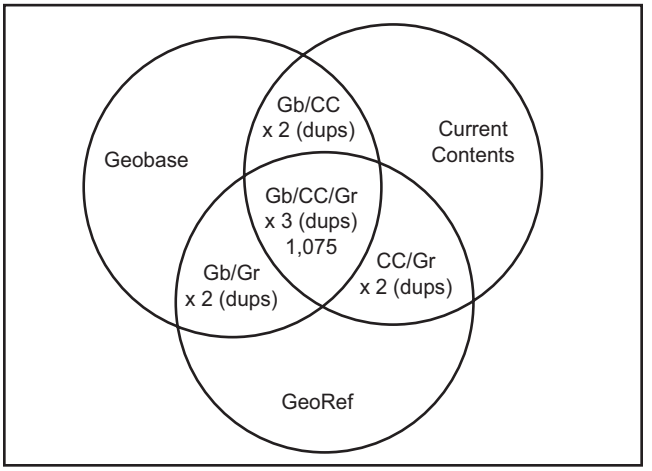


Figure 3: GeoRef, GEOBASE, and Current Contents Relevant Records



bases (UIUC no longer subscribes to Biological and Agricultural Index) on March 11, 2005, nearly three years after the initial searches that generated the data for this study. The same search terms were used, and again limited to the 2000 publication year. Table 3 compares the results of the 2002 versus 2005 searches. In one case (AGRICOLA), there were fewer records, possibly due to removing duplicates, but in many cases there was substantially more material.

There are other databases, possibly not used often by Quaternary researchers, that might have significant amounts of relevant information. Examples are Compendex, INSPEC, and NTIS. These three databases were searched after the comparison study to determine the amount of relevant information that each contains. The same search strategies were used as for the comparison databases. It was not possible to compare overlap with the comparison databases in the study because of the time lag between searches; however, the results are still informative.

Compendex via Engineering Village 2 was searched using the terms "Quaternary or Pleistocene or Holocene" for the publication year 2000. The search was conducted on March 11, 2005, and found 1,013 records. The same search was run several hours later in order to check for relevance, and 1,017 records were found. Note that all of this material was published in 2000, approximately five years prior to the search date, and yet material was still being added to the index, literally by the hour. Of the 1,017 records, 458 were relevant. There were 57 pairs of duplicates, so there were 401 nonduplicate relevant records, a significant amount of information. The amount of unique information is unknown, due to the above reasons, but there is a high probability that Compendex contains unique information.

When the search was conducted using Inspec, no duplicates were recovered. A total of 708 records were retrieved, of which 360 (50.85 percent) were relevant. A search of NTIS resulted in thirty-nine total records (no duplicates), of which twenty-nine (74.36 percent) were relevant. These supplemental searches of databases that are probably less used by Quaternary researchers further illustrate the utility of searching multiple databases, the likelihood that information is being overlooked, and the value of federated searching.

DISCUSSION

This study supports previous research, covered in the literature review, that emphasizes the need to search multiple databases if researchers are seeking comprehensive information. In addition, the results illustrate the need for using federated search or metasearch tools. Federated searching provides the ability to simultaneously search multiple databases, sources, platforms, and vendors and to retrieve collective results.[27] All of the databases considered in this paper should be included in federated searching for Quaternary information, if possible, and other databases should be tested for possible inclusion.

A large amount of information is being generated each year in the interdisciplinary area of Quaternary research. As mentioned in the introduction, an informal survey of an electronic Quaternary list and extensive association with Quaternary researchers indicate that most Quaternary researchers are using GeoRef, GEOBASE, and Current Contents (also Web of Science) as their primary tools for literature searches. This study indicates that they are using the most important databases. The informal query of researchers indicated that Biological Abstracts is probably used less frequently, but this study indicates that it is an important resource. All of the databases in the comparison study yielded unique records.

It is disturbing to consider all the information that may be hidden in the literature, unknown to researchers, and it is likewise intriguing to think about the possible outcomes if obscure lines of research and information could readily come together. Interdisciplinary meetings, such as those held by the

Table 3. Comparison of Results of Searches Conducted in 2002 and 2005 for PY2000

	Total Records 2002	Total Records 2005	Difference	% Difference
AGRICOLA	69	58	-11	-15.94
AGRIS	56	86	+30	+53.57
ASFA 3	33	48	+15	+45.45
Bio. & Ag. Index	71	X (Unknown)	Y (Unknown)	Unknown
Biological Abstracts	1,229	1,272	+43	+3.50
CAB Abstracts	291	400	+109	+37.46
Current Contents	3,359	3,360	+1	Negligible
GEOBASE	2,262	2,311	+49	+2.17
GeoRef	4,625	8,627	+4,002	+86.53
Water Resources Abs	170	235	+65	+38.24
Zoological Record	731	869	+138	+18.88
Total	**12,896**	**17,266 + X**	**4,441 + Y**	**Unknown**

various Quaternary societies, are one answer, but individual researchers may not be aware of or associated with these societies. Librarians and other information specialists may be able to help by being aware of who is studying what, by noticing information, and by connecting people and information sources.

In addition to indicating the best databases for Quaternary research and the amount of unique material contained in the various databases, this study also found a considerable and varying amount of lag time in indexing material. More research is needed to determine the extent to which indexing services are unable to keep up with the literature. The lag time between when material is published and when it enters the indexes is a moving target and has serious implications for researchers, scholars, and information professionals. The fact that indexing services prioritize certain titles and types of material for indexing should be considered by the end user when searching for certain types of information. It is not known how much material never makes it into any database.

The addition of four records to Compendex in a matter of hours underscores the necessity for the end user to continually rerun searches for older publication years in order to discover all of the material relevant to a topic. One cannot assume that one or even several searches over time of a particular publication year will yield comprehensive results, as a substantial amount of material is being added to databases even five years later. More research is needed to understand the longitudinal changes to databases regarding recall, precision, and overlap.

Denno pointed out indexing problems encountered by interdisciplinary researchers, problems such as differences in terminology and in orientation across disciplines.[28] Gerhard, Jacobson, and Williamson studied indexing adequacy of journals in the interdisciplinary area of women's studies and found that most of the journals were not sufficiently indexed.[29] It is possible that important information is being missed by indexers who are unfamiliar with Quaternary research. There may be gaps in the inclusion of non-English material; material that applies to Quaternary research but does not contain the words "Quaternary" or "Pleistocene" or "Holocene"; chapters in books; gray literature, such as field trip guide books; and very old information.

There are a number of factors that were not considered by this study, but of which researchers and scholars should be aware. The indexes vary in the types of sources included (journals, book chapters, dissertations, conference proceedings), the titles of journals included, the priority of title indexing (how soon a title is indexed), the depth of indexing (cover-to-cover versus partial indexing), the frequency of updates, and so forth. Also, titles may be added or dropped from indexing over time. This will affect the amount and type of information retrieved by a search. In addition, the vendor used will affect information retrieval due to varying frequency of updates and use of different search engines. The institution's subscription also may affect search results due to the number of years included in the subscription. In addition, longitudinal studies of databases will be affected by an institution changing years covered, dropping databases, or changing vendors.

The following comments about the research process may be helpful to others contemplating this sort of research. Computer analysis makes this sort of study possible because of the ability to process large amounts of data; however, programming the parsing is still time consuming, and discovering true duplicates is especially problematic. Determining the relevance of each record also is a time-consuming process.

As stated in the foregoing, the following research is needed: 1) additional science databases should be studied to determine relative value for Quaternary research; 2) changes in the retrieval of information from Web of Science make this sort of research possible, and this should be done; 3) longitudinal research to determine lag time in indexing and to determine changes to databases regarding recall, precision, and overlap should be conducted to compare the most important databases for Quaternary research; 4) recall, precision, and overlap of indexes should be studied for the subdisciplines within Quaternary research; and 5) strategies for finding older information related to Quaternary research should be examined.

CONCLUSIONS

This study indicates that, of the eleven indexes studied, the most important for Quaternary research are GeoRef, GEOBASE, Current Contents, Zoological Record, and Biological Abstracts. (Web of Science also is likely very useful, but retrieval problems prevented inclusion in this study). These five indexes together accounted for at least 97.9 percent of the nonduplicate relevant records in this study. However, each of the other six indexes in the comparison study contained unique, relevant information. In order to conduct a comprehensive search of the Quaternary research literature, as many science indexes as possible should be consulted. Currently, the most efficient method of doing this type of comprehensive search involves the use of federated search tools.

Acknowledgments: The author gratefully acknowledges and thanks Christopher Hamb, who created the programming for this project; the UIUC Library Research and Publication Committee, which provided financial support for the completion of this research; and Allan C. Ashworth, Geology Department at North Dakota State University, whose questions about federated searching for Quaternary research inspired this study.

References and Notes

1. The information in this section closely follows the introductory section in a chapter written by the author: Lura Joseph, "Quaternary Research," in *Literature Search Strategies for Interdisciplinary Research: A Sourcebook for Scientists and Engineers*, ed. Linda G. Ackerson, 31–45 (Lanham, Md.: Scarecrow, 2007).
2. Scott A. Elias, "Introduction," in *Encyclopedia of Quaternary Science*, ed. Scott A. Elias, xi–xii (New York: Elsevier, 2007).
3. J. J. Clague, "Revision of the Geological Time Scale: Implications for the 'Quaternary,'" e-mail communication via the

Quaternary@morgan.ucs.mun.ca list, June 16, 2004.
4. J. G. Ogg, "Status of Divisions of the International Geologic Time Scale," *Lethaia* 37, no. 1 (2004): 184–99; F. M. Gradstein and J. G. Ogg, "Geologic Time Scale 2004—Why, How, and Where Next!," *Lethaia* 37, no. 2 (2004): 176–81. Also, *Quaternary Perspectives* 14, no. 2 (2004) contains several updates and comments related to the debate.
5. S. C. Porter, *Quaternary Science*, www.inqua.tcd.ie/about.html#quaternary (accessed Mar. 31, 2007); Elias, "Introduction."
6. Elias, "Introduction."
7. Porter, *Quaternary Science*.
8. Linda G. Ackerson, "Challenges for Engineering Libraries: Supporting Research and Teaching in a Cross-Disciplinary Environment," *Science & Technology Libraries* 21, no. 1/2 (2001): 43–52.
9. Susan Starr, "Database Comparison—Directions for Future Research," in *National Online Meeting 1983: Proceedings of the Fourth National Online Meeting New York, Apr. 12–14, 1983*, ed. Martha E. Williams and Thomas H. Hogan, 531–38 (Medford, N.J.: Learned Information, 1983).
10. Jesus Rey and Raquel Fereres, "A Comparison of Two Databases for the Retrieval of Business Literature: ABI/INFORM and Management Contents," *New Review of Information and Library Research* 1 (1995): 161–80.
11. Katherine W. McCain, "Biotechnology in Context: A Database-Filtering Approach to Identifying Core and Productive Non-Core Journals Supporting Multidisciplinary R&D," *Journal of the American Society for Information Science* 46, no. 4 (1995): 306–17.
12. Christopher Hooper-Lane and David P. Atkins, "Is ISGA Sufficient? The Role of Inexpensive Multidisciplinary Databases in the Marine Science Information Center," *International Association of Aquatic and Marine Science Libraries and Information Centers Conference Series* 22 (1997): 31–38.
13. Anthony K. Cavanagh, "A Comparison of the Retrieval Performance of Multi-Disciplinary Table-of-Contents Databases with Conventional Specialized Databases," *Australian Academic and Research Libraries* 28, no. 2 (1997): 142–54.
14. M. Suzanne Brown, Jana S. Edwards, and Jeneen LaSee-Willemssen, "A New Comparison of the Current Index to Journals in Education and the Education Index: A Deep Analysis of Indexing," *The Journal of Academic Librarianship* 25, no. 3 (1999): 216–22.
15. Steve McDonald, Libby Taylor, and Clive Adams, "Searching the Right Database: A Comparison of Four Databases for Psychiatry Journals," *Health Libraries Review* 16, no. 3 (1999): 151–56.
16. Ellen Sutton and Lori Foulke, "Coverage of Anthropology by Major Electronic Indexes: A Comparison," *Reference Services Review* 27, no. 2 (1999): 134–57.
17. Eleanor F. Read and R. Craig Smith, "Searching for Library and Information Science Literature; A Comparison of Coverage in Three Databases," *Library Computing* 19, no. 1/2 (2000): 118–26.
18. James Jatkevicius, "Biological Sciences Databases in Academic Libraries: Comparing Cambridge Scientific Abstracts with BIOLOGICAL ABSTRACTS," *EContent* (Feb./Mar. 2000): 55–59.
19. Alison J. Brettle and Andrew F. Long, "Comparison of Bibliographic Databases for Information on the Rehabilitation of People with Severe Mental Illness," *Bulletin of the Medical Library Association* 89, no. 4 (2001): 353–62.
20. Jennalyn W. Tellman, "A Comparison of the Usefulness of IBZ and FRANCIS for Historical Research," *Reference & User Services Quarterly* 41, no. 1 (2001): 56–66.
21. Robert G. Janke, "Current Contents Connect and PubMed—A Comparison of Content and Currency," *Health Information and Libraries Journal* 19, no. 4 (2002): 230–32.
22. Laurel L. Kristick, "Database Reviews & Reports: Analysis of Serials Indexed in Meteorological and Geoastrophysical Abstracts (MGA)," *Issues in Science & Technology Librarianship* 36 (2002), www.istl.org/02-fall/databases.html (accessed July 14, 2005).
23. William H. Walters and Esther I. Wilder, "Bibliographic Index Coverage of a Multidisciplinary Field," *Journal of the American Society for Information Science and Technology* 54, no.14 (2003): 1305–12.
24. J. Bar and M. Finkler, "A Multi-Database Multidisciplinary Search Approach for Tomorrow's Online," in *Online Information 87* (Proceedings of the Eleventh International Online Information Meeting, London, 8–10 December 1987), 147–55.
25. William H. Mischo and Mary C. Schlembach, "Metasearch Technologies: Definitions, Issues, Reference Applications (PowerPoint)," in *Session 2441: Federated Searching* (ASEE 2004 National Conference, June 22, 2004). Available at http://eld.lib.ucdavis.edu/conf/04/asee04metas.ppt (accessed Feb. 17, 2005); Judy Luther, "Meta Search Engines—The Librarian's Answer to Google?" *The Charleston ADVISOR* 4, no. 4 (2003), www.charlestonco.com/features.cfm?id=123&type=np (accessed Mar. 7, 2005); Judy Luther, "Trumping Google? Metasearching's Promise," *Library Journal* 128, no. 16 (2003): 36–39; National Information Standards Organization, "NISO Metasearch Initiative," 2004, www.niso.org/committees/MS_initiative.html (accessed Mar. 7, 2005).
26. Nancy J. Butkovich, Helen F. Smith, and Claire E. Hoffman, "Database Reviews and Reports: A Comparison of Updating Frequency between Web of Science and Current Contents Connect," *Issues in Science & Technology Librarianship* 39 (2004),www.istl.org/04-winter/databases.html (accessed December 12, 2005).
27. Mischo and Schlembach, "Metasearch Technologies"; Luther, "Meta Search Engines"; Luther, "Trumping Google?"; National Information Standards Organization, "NISO Metasearch Initiative."
28. Deborah W. Denno, "Indexing of Interdisciplinary Literature on Crime, Violence, and Mental Disorder," *Journal of Information Science* 13, no. 2 (1987): 117–22.
29. Kristan H. Gerhard, Trudi E. Jacobson, and Susan G. Williamson, "Indexing Adequacy and Interdisciplinary Journals; The Case of Women's Studies," *College & Research Libraries* 54, no. 2 (1993): 125–35.

APPENDIX. DESCRIPTION OF DATABASES

AGRICOLA
Provider: National Technical Information Service; National Agricultural Library
Coverage: 1970 to present; worldwide.
Updates: Monthly
Vendor used: SilverPlatter
Description: http://grc.ntis.gov/agricola.htm

AGRIS
Provider: AGRIS Co-ordinating Centre, the Food and Agriculture Organization (FAO) of the United Nations.
Coverage: 1975 to present
Updates: Monthly
Vendor used: SilverPlatter
Description: www.fao.org/agris

Aquatic Science and Fisheries Abstracts: ASFA 3; Aquatic Pollution and Environmental Quality
Provider: Cambridge Scientific Abstracts
Coverage: 1971 to present
Updates: Monthly
Vendor used: Cambridge Scientific Abstracts
Description: www.csa.com/factsheets/aquclust-set-c.php

Biological and Agricultural Index
Provider: H. W.Wilson
Coverage: 1983 to present
Updates: Monthly on disk; daily on Web
Vendor used: H. W. Wilson
Description: www.hwwilson.com/databases/bioag.htm

Biological Abstracts
Provider: Biological Abstracts/Thomson Scientific
Coverage: 1969 to present
Updates: Biweekly
Vendor used: SilverPlatter
Description: www.biosis.org/products/ba

CAB Abstracts
Provider: CAB International
Coverage: 1972 to present
Updates: Monthly
Vendor used: SilverPlatter
Description: www.cabi.org

Compendex
Provider: Engineering Information/Elsevier
Coverage: 1970 to present
Updates: Weekly
Vendor used: Engineering Village 2
Description: www.ei.org/databases/compendex.html

Current Contents
Provider: Institute for Scientific Information (ISI)/Thomson Scientific
Coverage: 1993 to present
Updates: Weekly
Vendor used: Ovid
Description: www.isinet.com

GEOBASE
Provider: Elsevier Science
Coverage: 1980 to present
Updates: Semimonthly
Vendor used: FirstSearch
Description: www.oclc.org/support/documentation/firstsearch/databases/dbdetails/details/GEOBASE.htm

GeoRef
Provider: American Geological Institute
Coverage: 1785 to present, North America; 1933 to present, worldwide
Updates: Biweekly
Vendor used: SilverPlatter
Description: www.agiweb.org/georef/index.html

INSPEC
Provider: Institution of Electrical Engineers (IEE)
Coverage: 1969 to present; the backfile of Science Abstracts from 1898 to 1968 is now also available as the INSPEC Archive.
Updates: Weekly
Vendor used: Engineering Information/Elsevier
Description: www.iee.org/Publish/INSPEC

NTIS
Provider: National Technical Information Service (NTIS), U.S. Department of Commerce
Coverage: 1964 to present; some citations back to 1899.
Updates: weekly
Vendor used: Engineering Information/Elsevier
Description: www.ntis.gov

Water Resources Abstracts
Provider: Cambridge Scientific Abstracts
Coverage: 1968 to present
Updates: Monthly
Vendor used: Cambridge Scientific Abstracts
Description: www.csa.com/factsheets/water-resources-set-c.php

Zoological Record
Provider: BIOSIS
Coverage: 1978 to present
Updates: Monthly
Vendor used: SilverPlatter
Description: www.ovid.com/site/catalog/DataBase/200.jsp?top=2&mid=3&bottom=7&subsection=10