



AP
57,6

Is Google enough? Comparison of an internet search engine with academic library resources

498

Jan Brophy and David Bawden

Department of Information Science, City University, London, UK

Received July 2005
Revised August 2005
Accepted August 2005

Abstract

Purpose – The purpose of the study was to compare an internet search engine, Google, with appropriate library databases and systems, in order to assess the relative value, strengths and weaknesses of the two sorts of system.

Design/methodology/approach – A case study approach was used, with detailed analysis and failure checking of results. The performance of the two systems was assessed in terms of coverage, unique records, precision, and quality and accessibility of results. A novel form of relevance assessment, based on the work of Saracevic and others was devised.

Findings – Google is superior for coverage and accessibility. Library systems are superior for quality of results. Precision is similar for both systems. Good coverage requires use of both, as both have many unique items. Improving the skills of the searcher is likely to give better results from the library systems, but not from Google.

Research limitations/implications – Only four case studies were included. These were limited to the kind of queries likely to be searched by university students. Library resources were limited to those in two UK academic libraries. Only the basic Google web search functionality was used, and only the top ten records examined.

Practical implications – The results offer guidance for those providing support and training for use of these retrieval systems, and also provide evidence for debates on the “Google phenomenon”.

Originality/value – This is one of the few studies which provide evidence on the relative performance of internet search engines and library databases, and the only one to conduct such in-depth case studies. The method for the assessment of relevance is novel.

Keywords Academic libraries, Search engines, Information retrieval

Paper type Research paper

Introduction

Without doubt, the internet, and specifically the world wide web, has transformed the information environment in the past decade, providing more rapid access to a greater volume of material than possible at any earlier time. Searching tools, though far from perfect, have played a major part in this transformation. One of these tools, the Google search engine, has become predominant, to the extent that “to Google” had become *de facto* a verb in the English language by mid-2003, despite the objections of the company (Quint, 2002; BBC, 2003).

Google is, therefore, representative of the variety of easy-to-use search engines, based on free-text searching of the content of public web pages. It is indeed their major representative, given the mission of the company “to make all the world’s information available” (Library Journal News, 2003). The extension of the “basic” Google search function into Google scholar (providing access to non-copyright academic material



(Tenopir, 2005)), Google print (searching the digitised full text of printed books, from publishers, booksellers or libraries, and allowing the viewing of a small extract of copyright material) (Fialkoff, 2005), and other ventures, suggests that this may not be a wild ambition.

While these engines have indisputably made much information searching quicker and more efficient, they have also led to the belief that all information is to be found there, and retrieved without undue effort: “library patrons expect to find it all in cyberspace . . . for the purposes of academic research, such expectations are unrealistic and even dangerous” (Lawrence and Miller, 2000, p. 1). In turn, this leads to a dismissal of any other sources of information, specifically of libraries and the formal information sources which they provide:

In less than a decade, Internet search engines have completely changed how people gather information. No longer must we run to a library to look up something; rather we can pull up documents with just a few clicks on a keyboard. Now . . . “Googling” has become synonymous with doing research (Mostafa, 2005).

Web search engines, and Google in particular, have created a generation of searchers who are choosing the simplicity of search engines on the open free web over the perceived complexity of library services. Libraries can no longer cater for “people who want fast, easy access to unlimited, full-text content using interfaces that require no critical thought or evaluation” (Bell, 2004). Fast and Campbell (2004) found that students “admired the organisation of (an) OPAC, but preferred to use the web in spite of its disorganised state”.

It is, of course, inevitable that convenient access to information which, while it may not be comprehensive or of the highest quality, is good enough will be alluring. This is a natural human impulse, codified by Zipf into his principle of least effort (Sole and Cancho, 2003) and by Simon in his concept of “satisficing” (Tennant, 2001; Agosto, 2002), not to mention the complaint of some in the library/information area that we live in a society “fuelled by a culture of instant gratification” (Stoffle *et al.*, 1996, p. 219).

According to Bell (2004), “Google has become the symbol of competition to the academic library”. He uses the term “infobesity” to compare the way students now search for information with the modern consumption of fast food. Originally coined by James Morris, the Dean of the School of Computer Science at Carnegie Mellon University, “infobesity” refers to the belief that searching Google for information provides a junk information diet. Bell believes that students often want to find something quickly and that they are generally not concerned about the quality. It is clear from a review of the numerous articles published on this subject, that there is a general belief in the library community that more “nutritious” information can be retrieved by using the specialised databases available in an academic library. There are debates as to the amount of information available through systems such as Google compared to the “hidden web” of library databases (Tenopir, 2004; Herring, 2001; Devine and Egger-Sider, 2004), as well as of concerns about the quality of material retrieved (perhaps uncritically) from a search engine (Herring, 2001; Tennant, 2001).

The study reported here aimed to investigate, by detailed analysis of a small number of case studies, to what extent it is true that Google can supplant library databases, for typical queries of the sort likely to be researched by students. The study

was carried out as a Masters Degree dissertation, and full details may be found in the dissertation (Brophy, 2004).

Purpose and scope of the study

The purpose of this study was to compare the performance of Google to that of library database services in answering queries of the sort likely to be asked by students. In answering this question, it was hoped to gain an understanding of the optimal performance of both types of service, and of their relative advantages and disadvantages. By “library database services” we mean online catalogues and bibliographic databases, both general and specialised.

Although there have been numerous comparative evaluations of databases, library systems, and web search engines (for reviews of these studies see Brophy, 2004; Xie, 2004) relatively few have attempted this sort of direct comparison of two kinds of service.

One example is that of Xie (2004), who compared online database systems (Dialog and Factiva) with three different types of web search tool (search engine, directory and a meta-search engine). Students were asked to search the same two topics on each system and then required to give relevance scores, with precision of each system calculated on the total retrieved relevant documents (to a maximum of 20). Another is that of Fast and Campbell (2004), who examined the perceptions of students searching Google and a university library OPAC, using interviews, verbal reports and observations. A third is that of Griffiths and Brophy (2005), who report two studies of the use of Google and of various academic information resources, finding a predominant use of internet search engines. These examples illustrate the various tools which may be used to investigate a complex situation.

The scope of this project is restricted on the one hand to the basic Google web search function, and on the other to services provided by “typical” academic libraries in the UK. No attempt is made to deal with other internet search tools, or other Google services, nor the services provided by other library sectors.

The study involves four test queries, which are analysed in detail by a single expert searcher. While this number of queries is not enough to allow a claim that the results are applicable to all situations, the depth of qualitative analysis should allow a good understanding of the issues.

Methodology

Only the main points of the study methodology are described here: full details may be found in Brophy (2004).

A case study approach was adopted, with both quantitative and qualitative aspects. Quantitative results allowed an assessment of recall, precision, overlap and similar factors, while qualitative results allowed inclusion of ideas of quality and value of information (Bawden, 1990).

Using a small number of cases, each with a test query searched by the investigator, allowed a detailed study of the documents retrieved, and of the reasons for their not being retrieved from each service. This led to a detailed understanding of the reasons for the relative performance of the services.

The queries used were open-ended “research based” questions (Bilal, 2001), rather than closed reference style queries, since these allowed the performance of the services to be tested to best effect.

To avoid possible investigator bias in the choice of topics, four general subject areas were assigned by a librarian, unassociated with the study, as typifying appropriate domains. These were: environmental science, music, education, and law. Specific queries suitable for these subject domains were then chosen from a list of queries used in an online searching assignment. These were intended to mimic typical “real” student queries.

The test queries are shown in detail in Table I. It may be noted that the topics are as likely to be covered in web documents as in library database records, and that specialised vocabularies, while they may be available, are not required. In this sense they are a “fair test”, offering scope for both services to perform well.

The test queries were then searched on Google, and on databases and resources deemed suitable in an appropriate academic library. Different libraries were used because the library at City University London, at which the study was carried out, has little material on environmental sciences because of the subject profile of the university. The resources of Kingston University library were used for this query.

Each of the Library search sessions involved using an average of 14 queries to interrogate a variety of IR systems. Whilst some specialist subject databases such as RILM abstracts of Music Literature and Westlaw were used in only one task, a selection of more general sources (such as Factiva, Web of Knowledge, and Ingenta Journals) were used in all four tasks.

To create a more “natural” searching environment the researcher undertook all query formulation via direct interaction with each system. This allowed for goals and search strategies to change with every stage of the search process. As a result, no restriction was placed on the number of searches in each session. For practical reasons, however, the total session length for each task was restricted. All systems for a given query were tested in the same day, with each session lasting no longer than two hours. After this time period, the results of each search were saved and stored in order to allow evaluation to take place at a later stage. This was intended to avoid the possibility of pages being unavailable at a later stage of the evaluation, resulting in them being judged unfairly as inactive documents.

It was necessary to address four issues in analysing the results: the quality, relevance and accessibility of the documents retrieved, and the coverage of these documents in the systems tested.

Quality was assessed using Robinson’s (2000) framework, which considers:

Context:

- Relevance Is the resource suitable for its intended use/users?
- Authority Who is responsible for the resource? Are they qualified?
- Provenance How stable is the resource? What is its lifespan?
- Objectivity Does the resource provide a balanced, evaluated viewpoint?

Content:

- Currency Is the information up-to-date? Will it be updated regularly?
- Accuracy How accurate is the resource?
- Coverage Is the resource comprehensive? What is the subject coverage?

AP 57,6	<i>Test query 1</i>	
	Domain	Law
502	Topic	Internet piracy
	Task	The internet has changed the way people obtain music. Does file sharing hurt the music industry, and can record companies take legal action against illegal down-loaders? Can any additional measures be taken to deter potential “pirates”?
	Description	Documents should identify information regarding the current trend for downloading music from the internet. Documents should resolve any number of the following questions: How has the internet changed the way people obtain music? Is the music industry losing revenue? Is downloading music illegal? Is so, what legislation is in place? Does the current legal protection need to be changed? What measures can be taken immediately to deter internet “pirates”? Information is required which covers the legal aspects of the situation and possible changes which could make file-sharing an acceptable practise to all involved
	Narrative	Relevant documents will be suitable for an information science student studying a “Law Module”. They should provide information/discussion of the primary concepts detailed below. Documents should be in English and be suitable for a student whose main subject is not law. Documents aimed at law professionals will not be suitable
	Concepts	Intellectual property Piracy or pirate Illegal downloading File sharing Legislation Legal implications Napster Music industry or music recording industry Record companies Digital copyright Peer-to-peer file sharing or P2P file sharing Copyright infringement Intellectual property Online music
	<i>Test query 2</i>	
	Domain	Environmental Science
	Topic	Endangered Animals
	Task	How many species of animals are in danger of extinction? What could/should be done? What are the scientific/ethical issues?
	Description	Documents should resolve any number of the following questions: How do you define an endangered animal? How many are endangered? Why are they threatened? What conservation steps are being taken and what more can be done? What legislation exists? What are the scientific and ethical issues?

Table I.
Test queries

(continued)

Narrative	Relevant documents will be suitable for an undergraduate Environmental Science student. They should provide information regarding the concepts listed below. The student has a general understanding of environmental and biological terms but is by no means a specialist in this area. As such a variety of documents covering general background issues through to more focused analysis (which is not too technically demanding) will be required. In addition this student has not previously studied any law so information regarding legislation should be of a suitable level. Documents regarding endangered plants, insects, crustaceans and fish are not considered relevant. Book reviews are not considered relevant unless they contain a significant amount of data
Concepts	Endangered species or threatened species Animals or mammals or wildlife Population Conservation or environmental management Environmental legislation Human impact or human activities effects or Environmental impact Environmental monitoring Biological diversity or biodiversity or ecological balance Habitat protection or loss of key habitat Ethics or moral concepts or moral values Scientific information or environmental information
<i>Test query 3</i>	
Domain	Education
Topic	Children's reading habits
Task	Has the success of Harry Potter encouraged children's reading in general? If so, is this likely to continue? If not, why not?
Description	Documents should identify information regarding the impact of J.K. Rowling's books on children's reading. Information is required which covers the popularity of the Harry Potter books with children, changes in the reading habits of children, the causes of any such changes and any action being taken towards improving the situation
Narrative	Relevant documents will be suitable for a student studying Educational Theory. They should provide information/discussion of the primary concepts detailed below. Documents about teaching "methods" (teacher's notes, etc.), book/film reviews are not relevant. Documents should be in English and be aimed at undergraduate level studies
Concepts	Harry Potter Children's reading habits Rowling, Joanne K. Children's books Harry Potter books Children's literature Media habits of children Good reading habits Reading fundamental Effect of internet/computer games/television on children's reading English literature twentieth century English literature twenty-first century
<i>Test query 4</i>	
Domain	Music
Topic	Opera

(continued)

Table I.

AP 57,6	Task	Describe the literary, musical and historical background to Gounod's opera, "Faust". How does this particular work fit into its historical context? What recordings of the work are available?
	Description	Documents should identify documents which provide some answer to any of the following questions? What literature is the opera based on? What is the cultural, social and historical background to the opera? What period of history was Gounod writing in? How does this work fit into its historical context? What recordings of this opera are available? Which is considered to be the best?
504	Narrative	Relevant documents will be suitable for a first year Music student. General musical terminology is understood but the student has little knowledge of opera and its history. Details of individual productions are not relevant unless historically important or relate to specific recordings. Similarly singer biographies and university syllabus' are not considered relevant
	Concepts	Faust or Faustus Gounod Opera Literature Goethe Marlow or Marlowe Recordings Historical context Composer Musical period Musical style Opera (nineteenth) century French lyric opera Structure and style

Table I.

Relevance was assessed by the investigator. For each query within the four search sessions using Google, only the top ten hits were judged for relevance, in keeping with the level of evaluation in previous studies of this type (Schwartz, 1998).

Each retrieved document was judged for relevance using the three levels used by Chu and Rosenthal (1996): relevant, partially relevant and not relevant (Table II). The following novel relevance framework – based upon the work of Saracevic (1996), Barry and Schamber (1998) and Greisdorf (2003) – was adopted as the method for judging the degree to which each document was considered relevant:

- A document is considered topical if it matches the subject matter of any aspect of the query.
- A document is considered pertinent if it can be considered to be informative, that is to say if it contains substantive information, rather than just, for example, a list of resources.
- A positive mark for utility means that the document is considered useful for satisfying the information need.

It is possible that a document could be rated “Y” for utility, even though it is rated “N” for topicality. This would be an example of serendipity: something not matching the

question posed, but useful, for some unanticipated reason, in the supposed user context. (This did not occur in this study.)

Accessibility. Documents from Google were considered accessible if the link to them was active. Documents from other services were considered accessible if full-text access was available either instantly, on demand through a service such as Athens, or if the resource was easily accessible in hard-copy format. The latter was taken to mean that it was available in a library available to the researcher, and within reasonable travelling distance.

Coverage was assessed by carrying out a failure analysis, as a way of identifying relevant documents which are present in a system but have not (for whatever reason) been retrieved, augmenting the list of retrieved items, following the method of Robinson *et al.* (2000).

“Recall” was not assessed, due to the large and unstructured nature of the search space, which makes traditional recall measures infeasible. Recall measurements were substituted by measures of the coverage of useful items in each system.

Results

A summary of results only is given here. A fuller treatment is given in Brophy (2004).

A total of 723 documents were retrieved over all the searches for the four test queries. Of these, 237 from Google, and 163 from library systems, were judged relevant according to the framework above.

The results are summarised below under the criteria: retrieval; precision; quality; overall quality; accessibility; coverage; failure analysis; and uniqueness.

Figure 1 shows the percentage of retrieved documents within each of three relevance categories.

Figure 2 shows the precision scores (calculated in the usual way, as number of relevant retrieved items in ratio to total retrieved items) for the two different systems in the four respective subject domains. The results indicate that neither system consistently provides a more precise set of results. In three out of four cases the results were similar, suggesting that the top ranked results from a Google search have similar relevance to results obtained using Library systems.

Figure 3 shows a summary of the results of the Quality assessments conducted on the results of all four tasks. Overall, 52 per cent of Google results were found to be good quality whilst library systems recorded a figure of 84 per cent (Figure 4). This seems to indicate that if good quality results are required, it is better to use library systems. However, only a very small percentage of top-ranked Google results (4 per cent) were

Document number	Topicality	Pertinence	Utility	Relevant	Partially relevant	Not relevant
1	Y	Y	Y	X		
2	Y	Y	N		X	
3	Y	N	N			X
4	Y	N	Y		X	
5	N	N	N			X
6	N	N	Y			X
7	N	Y	Y		X	
8	N	Y	N			X

Table II.
Framework for judging
relevance

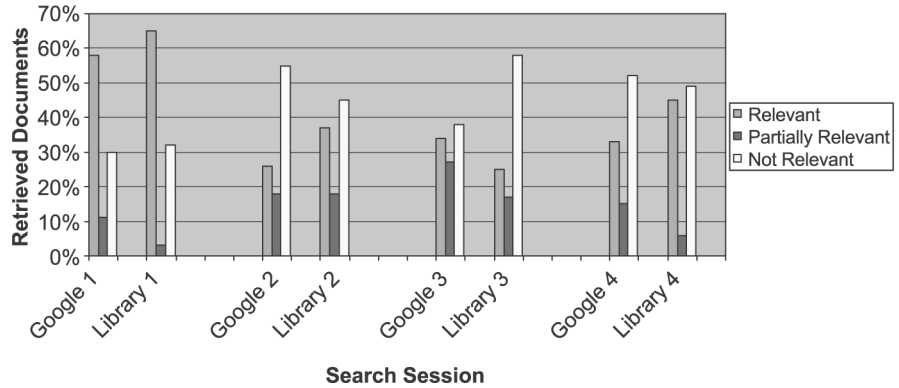


Figure 1.
Retrieved documents by
relevance

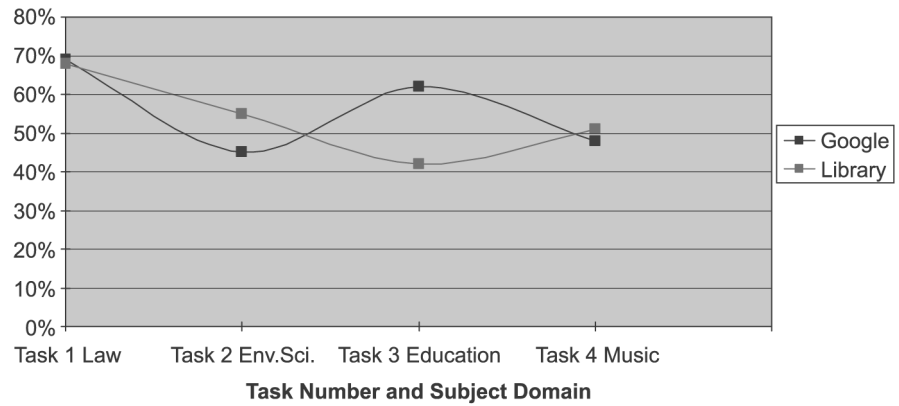


Figure 2.
Precision

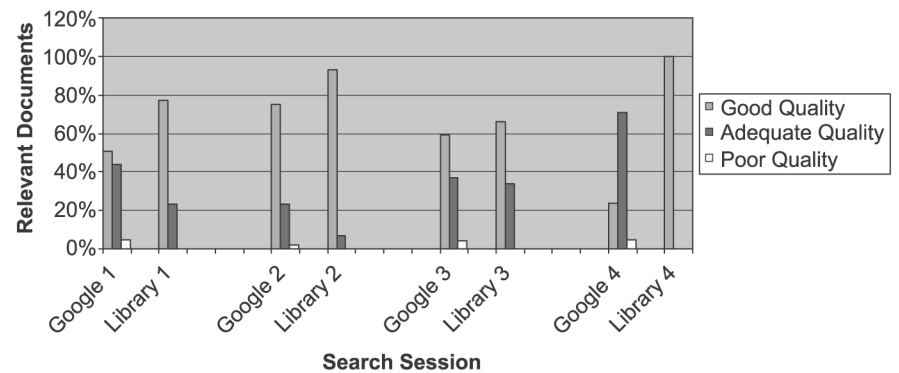


Figure 3.
Relevant documents by
quality

poor quality. Only two broken links were found when evaluating the four Google document result sets, indicating the system is updated frequently in order to remove any bad links. Google should not, therefore, be dismissed as providing consistently poor quality results.

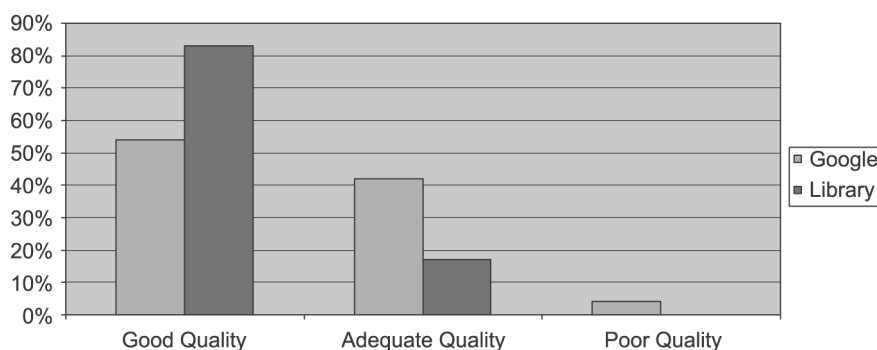


Figure 4.
Overall quality of relevant documents

The presumed ability of search engines like Google to provide immediate access to full-text material is confirmed in the results for accessibility, which consistently shows Google providing immediate access to over 90 per cent (over 97 per cent in three out of four tasks) of its relevant results (Figure 5).

The results for the library systems were more varied than for Google, with 65 per cent of the total results being found to be immediately accessible. In search sessions 2-4, a significant number of documents had to be sourced via the British Library or via an Inter-Library Loan Service (ILL). This would undoubtedly be off-putting to many students due to the time and costs involved in accessing the information. Whilst the numbers are low (less than 10 per cent in three out of four cases), relevant good quality documents were ultimately found which were unavailable even in the British Library collections. Overall, almost 35 per cent of the relevant library documents were found to require either an ILL or be inaccessible altogether. Virtually all Google documents were immediately accessible or not accessible at all,

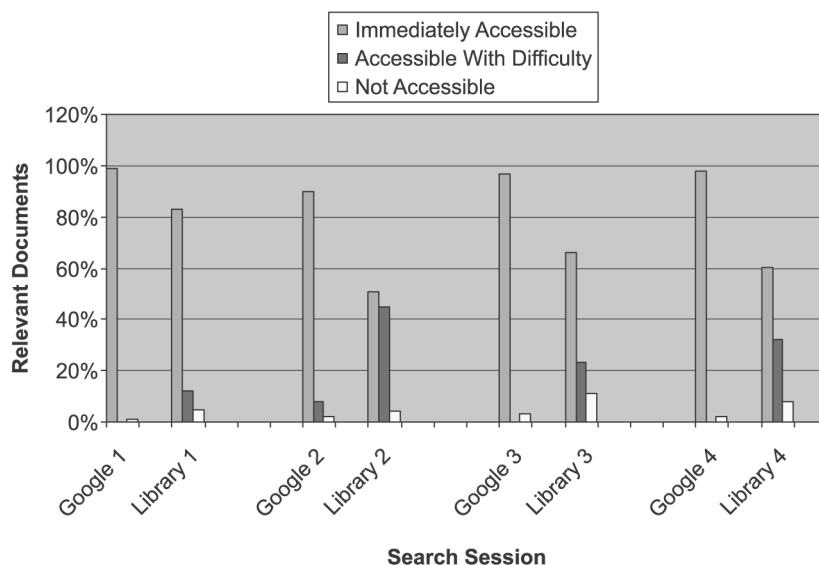


Figure 5.
Relevant documents by accessibility

due to broken links: a small number were accessible with difficulty, in that the link address had changed, and had to be identified by searching.

The coverage of the systems over the four tasks is shown in Figure 6, which displays the percentage of useful items which are indexed by each system for each of the four tasks. Google's coverage is better in each case. It is clear from these results that there is a degree of overlap in the documents obtained from the two systems. In three out of four cases the total useful coverage can be seen to be greater than 100 per cent, demonstrating that a proportion of documents in each case are indexed by both systems.

In order to understand the reason certain documents were not retrieved, though present in the system, the search queries used in each session were assessed (Figure 7). For Google, 31 per cent of the additional documents could have been found by improving the terms used when constructing the search queries. However, for almost 70 per cent of documents, Google required the specific title of the document to be queried. For the library, a more balanced result was apparent with 50 per cent of documents being sourced on improving the search query and 50 per cent requiring very specific query terms. This result shows a large proportion of results which would have been extremely difficult to retrieve, irrespective of the skills of the searcher. It also suggests that the improvement of searching skills, by training programmes, for example, is more appropriate to library systems rather than Google.

Figure 8 shows the proportion of relevant documents which are only indexed by one system: "uniques".

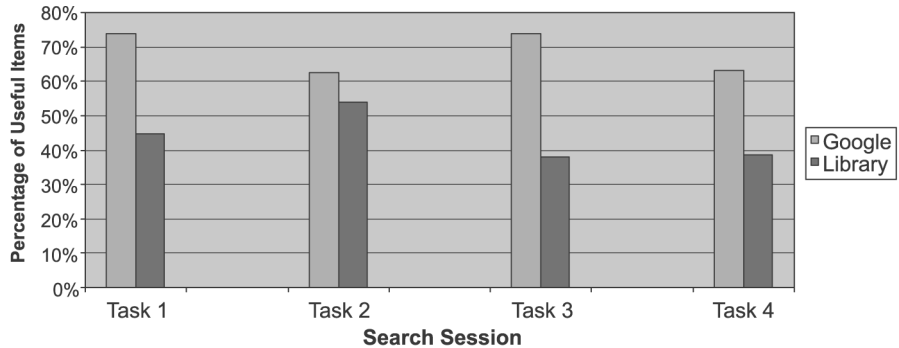


Figure 6.
Coverage

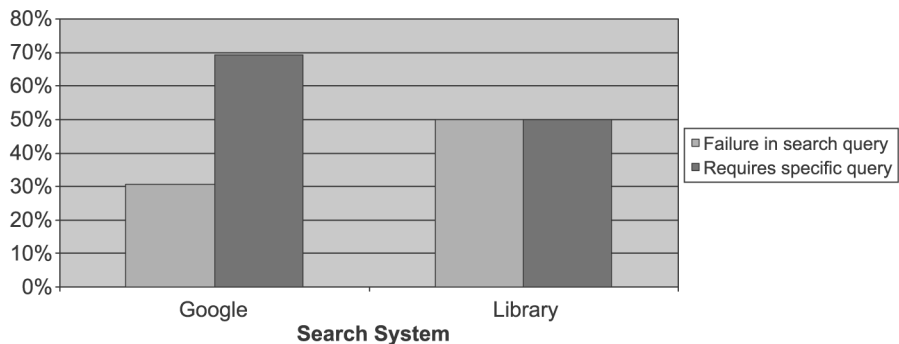


Figure 7.
Failure to retrieve relevant documents

Shown as a percentage, the results are fairly high across all eight search sessions, with Task 1 (library) the only session to record less than 60 per cent of its results as unique. Overall, Google provided a higher proportion of unique results than library systems, indicating that more library material is indexed by Google than vice versa.

The main data from these results are summarised in tabular form in Tables III-V. These show figures for quality, accessibility and retrieval performance (precision, coverage, and uniques), respectively.

Conclusions

The results of this study may be summarised by saying that the two kinds of resource – search engines and library databases – seem to be complementary, as assessed by their performance over this small set of test queries.

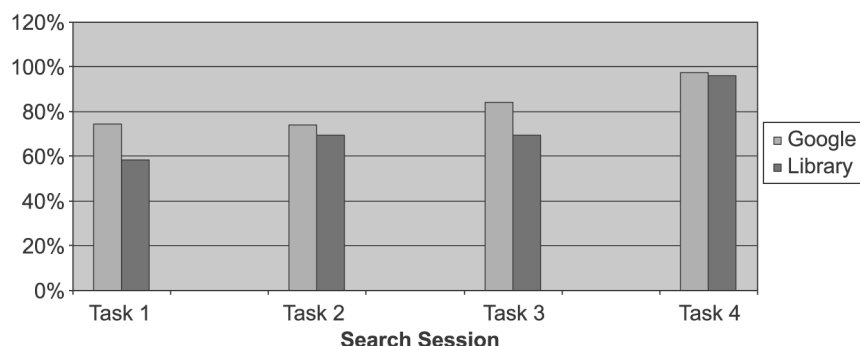


Figure 8.
Proportion of results which are unique in each search session

Quality level	Google (mean) (per cent)	Library (mean) (per cent)
Good quality	52	84
Adequate quality	44	16
Poor quality	4	0

Table III.
Overall quality of relevant material (mean)

Accessibility level	Google (mean) (per cent)	Library (mean) (per cent)
Immediately accessible	96	65
Accessible with difficulty	2	28
Not accessible	2	7

Table IV.
Overall accessibility of relevant material (mean)

	Google (mean) (per cent)	Library (mean) (per cent)
Precision	56	54
Coverage	69	44
Unique	83	73

Table V.
Overall retrieval performance (mean)

Google appears clearly superior in both coverage and accessibility. Google also gives more unique items, though the difference is small, and both types of resource are needed if good coverage is required.

The library databases are superior for quality of results, though Google's performance here is not unacceptable, when only the top ten results are considered. Precision is similar for both systems, again with the caveat that only the top ten Google results are considered. The precision is not particularly impressive in either case, being less than 60 per cent.

Intriguingly, improving the skills of the searcher is likely to give better results from the library systems, but not from Google. This has implications for user awareness and training programmes. It may be seen as a worrying factor, given the tendency identified above, and emphasised in the findings of Griffiths and Brophy (2005) for ease of use, and by implication lack of need for training, to be the major factor in choice of source. If systems like Google are indeed usable to best effect without training to improve competence, then it is highly likely that they will be preferred.

In comparing the systems, in terms of advantages and disadvantages, the conclusions can be summarised as follows:

Google

- a high proportion of relevant documents retrieved;
- an ability to retrieve a fairly precise set of documents;
- a high proportion of adequate or good quality results;
- a high proportion of unique documents; and
- no problems with accessibility.

Library databases

- a moderate proportion of relevant documents retrieved;
- an ability to retrieve a fairly precise set of documents;
- a high proportion of good quality results;
- a high proportion of unique documents; and
- some problems with accessibility.

The main discriminating factors seem to be quality (favouring library systems) and accessibility (favouring Google). Coverage also favours Google, although both systems are needed to achieve anything approaching comprehensive recall.

These conclusions seem to reflect the issues raised at the start, and reflect those identified by Griffiths and Brophy (2005). Accessibility is likely (rightly or wrongly) to be favoured over quality as a determinant of choice by the student users considered here. Lack of comprehensiveness in retrieval is unlikely to be a strong motivator for these users to use any retrieval systems in addition to an internet search engine. Nor is the prospect of undertaking extra training to make better use of library databases likely to be attractive, when this is not useful for Google.

It may be that, in the medium term, such issues will disappear, as search engines like Google – and the more advanced ones which will succeed it (Mostafa, 2005) – acquire more “academic” content, and as library system interfaces take on the look-and-feel of search engines. In effect, the systems will merge, hopefully

encapsulating the best features of both. The difficulties of adapting library systems in this way should not be underestimated (Fast and Campbell, 2004): they have, after all, been designed with the basic assumption of users having what Borgman (1996) terms “a rich conceptual framework for information retrieval”.

In the shorter term, there may be a role for information specialists, in their capacity as facilitators of information literacy, in helping their users to appreciate the limitations of all available systems, and suggest strategies to overcome them. This could also include some understanding of the nature of academic information and its structuring. While it is perhaps unreasonable to expect a typical student to gain Borgman’s conceptual framework in its entirety, some realistic understanding of the world of information may be a realistic aim.

References

- Agusto, D.E. (2002), “Bounded rationality and satisficing in young people’s web-based decision-making”, *Journal of the American Society for Information Science and Technology*, Vol. 53 No. 1, pp. 16-27.
- Barry, C.L. and Schamber, L. (1998), “Users’ criteria for relevance evaluation: a cross-situational comparison”, *Information Processing and Management*, Vol. 34 Nos 2/3, pp. 219-36.
- Bawden, D. (1990), *User-Oriented Evaluation of Information Systems and Services*, Gower, Aldershot.
- BBC (2003), “Google calls in the language police”, *BBC News*, 20 June, available at: <http://news.bbc.co.uk/1/hi/3006486.stm> (accessed 3 March 2005).
- Bell, S. (2004), “The infodiet: how libraries can offer an appetizing alternative to Google”, *The Chronicle of Higher Education*, Vol. 50 No. 24, p. B15, available at: <http://chronicle.com/prm/weekly/v50/i24/24b01501.htm> (accessed 3 March 2005).
- Bilal, D. (2001), “Children’s use of the yahooligans! web search engine: II. Cognitive and physical behaviours on research tasks”, *Journal of the American Society for Information Science and Technology*, Vol. 52 No. 2, pp. 118-36.
- Borgman, C.L. (1996), “Why are online catalogs still hard to use?”, *Journal of the American Society for Information Science*, Vol. 47 No. 7, pp. 493-503.
- Brophy, J. (2004), “Is Google good enough? A study comparing a major search engine with academic library research”, unpublished MSc dissertation, Department of Information Science, City University London, London, available at: www.city.ac.uk/library
- Chu, H. and Rosenthal, M. (1996), “Search engines for the world wide web: a comparative study and evaluation methodology”, *ASIS 1996 Annual Conference Proceedings*, pp. 19-24.
- Devine, J. and Egger-Sider, F. (2004), “Beyond Google: the invisible web in the academic library”, *Journal of Academic Librarianship*, Vol. 30 No. 4, pp. 265-9.
- Fast, K.V. and Campbell, D.G. (2004), “I still like Google: university student perceptions of searching OPACs and the web”, *Proceedings of the 67th ASIS&T Annual Meeting*, pp. 138-46.
- Fialkoff, F. (2005), “Access by Google”, *Library Journal*, 15 January, available at: www.libraryjournal.com/article/CA491137 (accessed 3 August 2005).
- Greisdorf, H. (2003), “Relevance thresholds: a multi-stage predictive model of how users evaluate information”, *Information Processing and Management*, Vol. 39, pp. 403-23.
- Griffiths, J.R. and Brophy, P. (2005), “Student searching behaviour and the web: use of academic resources and Google”, *Library Trends*, Vol. 53 No. 4, pp. 539-54.

-
- Herring, M.Y. (2001), "10 reasons why the internet is no substitute for a library", *American Libraries*, April, pp. 76-8, available at: www.ala.org/ala/online/selectedarticles/10reasonswhy.htm (accessed 5 March 2005).
- Lawrence, H. and Miller, W. (2000), *Academic Research on the Internet: Options for Scholars and Libraries*, Haworth Information Press, New York, NY.
- Library Journal News* (2003), "Google guy: reference librarians can keep jobs", *Library Journal News*, 13 May, available at: www.libraryjournal.com/article/CA2999471 (accessed 3 March 2005).
- Mostafa, J. (2005), "Seeking better web searches", *Scientific American*, Vol. 292 No. 2, pp. 51-7.
- Quint, B. (2002), "Searcher's voice: Google: (v) . . .", *The Searcher*, Vol. 10 No. 2, pp. 6-9, available at: www.infotoday.com/searcher/feb02/voice.htm (accessed 5 March 2005).
- Robinson, L. (2000), "A strategic approach to research using internet tools and resources", *Aslib Proceedings*, Vol. 52 No. 1, pp. 11-19.
- Robinson, L., McIlwaine, I., Copestake, P. and Anderson, C. (2000), "Comparative evaluation of the performance of online databases in answering toxicological queries", *International Journal of Information Management*, Vol. 20 No. 1, pp. 79-87.
- Saracevic, T. (1996), "Relevance reconsidered", in Ingwersen, P. and Pors, N.O. (Eds), *Information Science: Integration in Perspectives*, The Royal School of Librarianship, Copenhagen, pp. 201-18, available at: www.scils.rutgers.edu/~tefko/CoLIS2_1996.doc (accessed 11 April 2005).
- Schwartz, C. (1998), "Web search engines", *Journal of the American Society for Information Science*, Vol. 49 No. 11, pp. 973-82.
- Sole, R.V. and Cancho, R.F. (2003), "Least effort and the origins of scaling in human language", *Proceedings of the National Academy of Sciences (USA)*, Vol. 100 No. 3, pp. 788-91, available at: www.pnas.org/cgi/reprint/0335980100v1.pdf (accessed 5 March 2005).
- Stoffle, C., Renaud, R. and Veldof, J. (1996), "Choosing our futures", *College and Research Libraries*, Vol. 57 No. 3, pp. 213-25.
- Tennant, R. (2001), "Digital libraries – the convenience catastrophe", *Library Journal*, 15 December, available at: www.libraryjournal.com/article/CA185367 (accessed 14 May 2005).
- Tenopir, C. (2004), "Is Google the competition", *The Library Journal*, 1 April, available at: www.libraryjournal.com/article/CA405423 (accessed 3 March 2005).
- Tenopir, C. (2005), "Google in the academic library", *The Library Journal*, 1 February, available at: www.libraryjournal.com/article/CA498868 (accessed 3 March 2005).
- Xie, H. (Iris) (2004), "Online IR system evaluation: online databases versus web search engines", *Online Information Review*, Vol. 28 No. 3, pp. 211-9.