

Custom Interfaces for Advanced Queries in Search Engines¹

Mike Thelwall, Ray Binns, Gareth Harries, Theresa Page-Kennedy, Liz Price, David Wilkinson.

School of Computing and Information Technology, University of Wolverhampton, Wulfruna Street, Wolverhampton, WV1 1SB, UK.

E-mail: D.Wilkinson@wlv.ac.uk. Phone +44 1902 321452. Fax +44 1902 321491.

Abstract

Those seeking information from the Internet often start from a search engine, using either its organised directory structure or its text query facility. In response to the difficulty in identifying the most relevant pages for some information needs, many search engines offer Boolean text matching and some, including Google, AltaVista and HotBot, offer the facility to integrate additional information into a more advanced request. Amongst web users, however, it is known that the employment of complex enquiries is far from universal, with very short queries being the norm. It is demonstrated that the gap between the provision of advanced search facilities and their use can be bridged, for specific information needs, by the construction of a simple interface in the form of a web site that automatically formulates the necessary requests. It is argued that this kind of resource, perhaps employing additional knowledge domain specific information, is one that could be useful for web sites or portals of common interest groups. The approach is illustrated by a web site that enables a user to search the individual web sites of university level institutions in European Union-associated countries.

Introduction

Search engines are widely used by web surfers in order to find information. They work by crawling a portion of the web, indexing it in a large database, and then providing an interface, normally via a web page, for interrogating that database. Using search engines for retrieving information on the web is often difficult, perhaps partly because of the problem of formulating queries specific enough to find relevant pages. Some information needs, however, are intrinsically more difficult to satisfy due to problems of semantics. A user searching for “warehouses” may be surprised by the number of pages referring to software warehouses, rather than buildings, returned. A search for a more specific word such as “phenol” will, in contrast, unhampered by non-chemical uses for the term. There are many web sites that already provide sources of information for specific groups of web users, for example one for the chemical industry (www.sourcerer.co.uk) that sells access to its own database of UK chemical company products and services. As well as commercial initiatives, professional organisations, commercial common interest groups and other organisations with a web site aimed at a collection of people with shared characteristics may benefit from enhancing their web site with a facility for helping their target audience to search the web for topic-related information.

The difficulty of finding information on the web has been widely recognised, both directly¹⁻² and indirectly in the production of a large number of tools to aid the

¹ Thelwall, M. Binns, R. Harries, G. Page-Kennedy, T. Price E., and Wilkinson, D. (2001). Custom Interfaces for advanced queries in search engines, *ASLIB Proceedings*, 53(10), 413-422.

process³⁻⁶, including those aimed at common interest groups⁷⁻⁸. These tools include personal search engines⁹ which can automatically seek information relevant to the owner's web surfing history but cannot help novel tasks, and those designed to crawl a site in response to a query¹⁰, but this is a relatively time-consuming approach. Recent trends in search engine design have been driven by the need to identify the pages on the web most likely to satisfy any given information request. Modifications to the way in which search engines classify pages have also been proposed, for example using the Dewey Decimal scheme¹¹. One development has been an attempt to measure the perceived quality of information on a page by the number of links to it, with the rationale that a page that is the target of many links is more likely to contain useful information than one that none have links to¹². This development may also mean that information that is only infrequently needed will be harder to find.

Some search engines, including AltaVista, also offer advanced search facilities, enabling users to submit complex Boolean queries, but it is believed that relatively few searches actually include these facilities¹³, and the use of Boolean queries for information retrieval can be poor and error-prone^{14,15}. One reason for the lack of use of advanced queries has been suggested to be that they give only minor improvements¹⁶, although this study was based upon only fifteen random queries. Some requests can, however, only be answered with advanced syntax, for example site-specific searches. As an example of this, a user wishing to use AltaVista to find maths pages in the Wolverhampton University web site would need to submit the advanced query

maths AND host:.wlv.ac.uk

in order to avoid having the results from other sources swamping those from Wolverhampton. Many search engines now provide a web form based advanced search interface that allows users to specify additional criteria for their search without requiring them to create a Boolean expression, but this is intrinsically less powerful and flexible.

One alternative to standard and advanced searches is to search for information by category in the directory section of a search engine. If a suitable category exists, then this approach should yield only sites that are relevant to the information need. If a particular piece of information is being sought that may or may not be in any of the category sites, then Google offers the facility to search inside the sites in any category, giving a similar effect to an advanced search with AltaVista that specifies the host. Google even allows searching in every site in a whole category, a very powerful feature. As an example of this, a search for 'Metallurgist' in the Google UK university category would return (on several pages) one or two matches from each university which had a web site containing this word. It will be demonstrated, however, that the interface presented can be customised to include additional useful information and to make certain types of searching easier.

The vacuum formed by the gap between the requirement for complex queries and their lack of use has allowed tools that provide a simple substitute interface to be developed. Two examples use semantic disambiguation techniques to automatically create advanced queries which are, in general, more successful than the unprocessed originals¹⁷⁻¹⁸. This approach essentially gives indirect access to more powerful searches to those who would not otherwise use it. It is aimed at the general information needs, but it is also possible to create a simple interface targeted at specific types of queries, for example site-specific searches. Such a tool could be provided, for example, by an interest group that wishes to create an information resource web site and expects that its user base would have a need to search a

common set of web sites. Many such sites already maintain a list of useful links, and this would be a logical extension to that. Indeed, some sites already provide specialist search facilities¹⁹ but this represents a level of technological complexity and expense that would rule it out for many groups.

In order to test the concept of using automatically generated advanced queries to create a portal for a common interest group, a web site at <http://www.scit.wlv.ac.uk/~cm1993/eusearch/> was created to aid academics to find research partners in unknown disciplines and in universities throughout Europe. Such a site can have two potential advantages: making the formulation of appropriate queries easier; and being able to include additional, target-user specific information. Both of these aspects will be discussed and the site subjected to an initial evaluation with a tool from the discipline of human-computer interaction that is used by developers for theoretical testing of interfaces.

Site-Specific Searching

The potential pitfalls from using standard approaches will be illustrated with an example of a similar piece of information to be extracted from a collection of sites. The example is to extract the email address or other contact information for five academics from different universities in the UK with a research interest in information systems. In order to solve the problem, a search engine may be useful if used as part of an appropriate strategy. Ideally, the user would enter the search text “information systems research UK” and the search engine would return a list of links to the web pages of all information systems researchers in the UK. In tests, AltaVista, HotBot, Yahoo! and Google returned a variety of pages related to information systems, mainly to national organisations and web sites. Only Hotbot returned a directly useful link, to the Management and Information Systems Research Group in the University of York. In practice, therefore, the lists of pages returned containing all or most of the terms: “information”; “systems”; “research”; “UK” provided an imperfect solution to the problem. A fundamental information retrieval problem here is that there may be many pages that only contain the information that they belong to a UK university implicitly from being part of its web site, rather than explicitly by containing the words “university” and “UK” in the text of the page. The query text may also be elsewhere on the site, may be scattered throughout the site, or may even not be present at all. The directory structure of search engines can be used to narrow down searches before entering the search terms, in an attempt to circumvent this problem. A search from the Yahoo!Directory category for UK Colleges and Universities for “information systems research” did not yield any academic or research group pages because the search does not cover the majority of the web sites. Google does attempt to search a large part of each academic site in its category search, and so it was also used. A Google search of the category for UK Colleges and Universities for “information systems research” did yield a large number of useful links. Google returned only one or two matches per site, which enabled it to return several universities per page. It does not allow selection of the order in which the universities will appear, however. Google could also be used to search each individual site by selecting their category first.

A third approach is to find a listing of UK university websites and then to use the advanced search facility of an engine like AltaVista to search each domain specifically. An example of this would be

host:warwick.ac.uk AND information systems research

This is clearly a more complex strategy, requiring the knowledge of advanced syntax. It is essentially equivalent to the search offered by Google by first selecting the university from a category and then choosing the option to search within the selected web site. This type of search will form the basis of the custom interface discussed below, and would be particularly useful for any categories not covered by Google or by any other search engine offering the same category search functionality. This approach may also be applicable to other less clear-cut uses of advanced syntax when information about the context of the query can be inferred by the presence of the user at a special interest host site. As a simple example of this, a fan or club site for a football team may have an interface which automatically adds "AND NOT soccer" to all user queries.

Search engine category listings normally contain a brief description of each site suggested to the user, often with the text drawn from the pages targeted. Sometimes this information will not be enough to enable searchers to decide which sites to choose. They may, however, require specific information that is not provided, such as geographic location, product types or company sizes in order to make the choice. If such a need were to be identified as common to a number of people, then it would make sense for it to be provided by the search engine at the point of need, as part of the company listing in the directory structure. No freely available commercial search engine, however, has done this yet. This is perhaps a reflection of the difficulty in ascertaining the information that would be most pertinent in any given category, and over a sufficient range of categories to make a marketable additional facility. This does, therefore, make available an alternative strategy for an interested third party: to provide their own interface to a search engine, which can include the necessary additional information. This is technically possible because search engine queries are HTML constructs, and so can be formed in any web page and passed on to the search engine. In fact many search engines provide the HTML for a search box themselves and encourage visitors to add it to their sites. The type of third party that may benefit from this approach would be an interest group that identifies common information retrieval needs which search engines are not able to easily satisfy. The relationship here would be a symbiotic one, because the queries submitted would be generating hits on the search engine and, therefore, producing advertising revenue. In fact AltaVista, as part of its affiliate network, will pay users to add the HTML for a search query box to the search and so may be willing to pay the portal for this facility. There are already many portal sites that attempt to collate information for their user base and provide aid in web navigation. One method has already been suggested to automate this process in the face of its time-consuming nature²⁰, but the approach described here is dealing with a related problem through a completely different technique.

A Test Case: The Academic Partner Web Site

As a test case, a web portal to designed to enable the speedy selection and searching of university web sites from 31 European Union-associated countries was created and will be examined. It also provides pertinent information to aid site selection and automates the task of constructing Boolean queries. University web sites tend to be well covered by AltaVista²¹ and for this reason seem to be a good target for this approach. A similar web site has been created before for the UK to enable industry to have access to information about university research²², but this project involved the construction and maintenance of a database, a relatively expensive method. The web

site was constructed both to fill an information need and to provide a demonstration of the ability of third parties to provide an added value service to specific interest groups, based upon existing commercial search engines. The target group for the site was university academics, and the need was of finding other academics in highly rated research institutions in the EU for specific disciplines, potentially ones unknown to the searcher. This is the kind of need that can arise, for example, when putting together a consortium to bid for EU funding under a research budget such as the Fifth Framework Programme²³. This source of funding encourages both multinational and multidisciplinary co-operation, leading individuals and teams wishing to submit a bid to seek out others in unknown disciplines and countries. Web sites are often not designed in a way that aids all users²⁴, and finding information on academic web sites is known to be difficult²⁵, so this is a task that would benefit from help. In order to identify through a web search suitable academics the following pieces of information would be useful.

1. The names and domain names of university level institutions in the target countries.
2. An indication of the relative research standing of each university in its own country.

The academic partner web site attempts to provide both these pieces of information in an environment seamlessly linked to AltaVista so that the universities can be searched individually. It also provides a link to AltaVista's translation page to help searches in countries where the language is unfamiliar to the user.

There are many web sites that give lists of university home pages in various countries, including Yahoo! and Google. In order to use search engine advanced syntax, however, it is necessary to know the domain names that host the majority of web pages for any given university. The team working on the academic partner web site collated lists of university web sites in EU-associated countries from various sources and used AltaVista to count the pages hosted on each one. Any university with few pages was followed up, often resulting in the discovery of a change of domain name, or a second site hosting more pages than the official site. As a result, a more useful list of domain names was produced. This was a time consuming operation, with over 700 sites to check.

The second requirement, an assessment of the relative research standings of universities, is much more difficult to satisfy. In the UK there is an official research assessment exercise performed by the government, which can be used to construct useful general research quality indicators²⁶. Such exercises are, however, not universal internationally. An alternative measurement has been created by Ingwersen²⁷, the Web Impact Factor, which is based upon counting links to a site and dividing by the number of pages hosted in the site. It has been shown that this measurement correlates strongly with the accepted research ratings in the UK²¹. For example, counting the number of UK university web pages outside (say) Oxford University that contain a link to any Oxford University web page, and dividing this by the total number of web pages in the Oxford site gives a figure that it is useful to compare with other similar figures. It was decided to provide official research ratings for universities in countries where they were available, but to provide this calculation, accompanied by an appropriate disclaimer warning of their potential lack of reliability, when they were not.

The web site home page provides the user with a search text box similar to that of a search engine, but with the additional option to select a country upon which to focus. Submitting this page does not return a set of results but a table of

information about the universities in the country chosen. This includes the name of each university and their (normally estimated) research ratings. The table also contains a link for each university that is an appropriately formulated advanced AltaVista query to search the specific university web site for the terms entered on the previous page. A single click is enough to submit the query to AltaVista and load it into a new browser window. For example if the term “mathematics” had been entered, and Spain selected on the first page then clicking on the University of Seville link would be equivalent to submitting to AltaVista in a new browser window the query: mathematics AND host:seville.es. It was decided to use AltaVista instead of an equivalent Google category search, simply because the AltaVista approach is slightly more generic, with Google needing to have an appropriate category set up in order to use its site search feature.

University list for Spain

Rating	University	University-based search
0.038	Alicante	Search Alicante for maths
4/0	Almería	Search Almería for maths
0.343	Andalucia	Search Andalucia for maths

FIG. 1 An extract from a site page after a maths query has been submitted for Spain. It shows ratings, links to homepages and links to trigger AltaVista advanced site-specific searches for maths

The aim of the web site, in summary, was to aid both the selection of universities and the subsequent searching within the chosen sites.

Interface Analysis and Discussion

The web site will be compared with an advanced search engine and a directory based search service, in the context of seeking unknown academics in other countries. The basis for comparison will be twofold: how likely is it that the method will be successful in meeting the user’s need; and how easy the interface is to use. This second point of comparison will be attacked in a relatively abstract fashion, the interest being in the approach used rather than the specifics of the actual implementation.

One way in which the academic partner site is more satisfactory than the comparators is in the provision of the additional, task specific information relating to university research ratings. If this is critical information for the task, then it may be the case that the partner site is the only realistic web search-based option. It may, however, be possible to combine the other search strategies with alternative methods of identifying research standings, for example consultation with peers, other web sites, or official offices for international affairs. These options would, however, add to the time required to complete the task.

The academic partner site is less satisfactory than the comparators in transferability of skills. A different web site, no matter how simple, requires some time to learn and

master. A user already familiar with standard search and directory services would, therefore, be forced to undergo new learning if they were to use the partner site, time that would be producing less transferable skills than if a generic search site were being used. It would also, of course, be necessary to find the site to start with, a factor that is difficult to account for in comparative studies.

Comparison with AltaVista

Searching by university web site in AltaVista using the advanced search syntax amounts to manually entering the queries for each university, rather than having them automatically created by the site. For those unwilling to learn the necessary syntax, or who do not know of its existence, the partner site makes this type of search possible. Its advantages over AltaVista for those who could use either method are in the provision of the domain name information for the universities in the chosen countries and in the automatic creation of the queries for each institution. The partner site creates a page of clickable links to the search results pages, so that the query terms only need to be entered once, and the domain names and advanced syntax do not need to be entered. To activate successive searches all that is needed is to click the query links in turn, closing the new browser window created by each link once the search has been finished. One simple way to compare two methods is to count the number of keystrokes and other similar actions that are required to complete the actions, this modelling the difference in time for an expert user to execute the tasks without error²⁸. This does not, however, include any attempt at quantification of the difficulty or length of decision-making processes that using either method would entail. Table 1 shows an analysis of the actions needed to do this, compared with those for a user entering the queries manually, referring to another web site in a separate browser window for the domain names, and using the browser 'back' button intelligently to minimise the total number of actions needed. The task is the execution, but not following up, of queries for the first five listed Spanish universities, with the query 'Maths'.

Partner site

Starting with partner site loaded:

- type "maths", select Spain from the country list and click on the submit button.

Next, for each of the five universities:

- click on the search link to activate the search in a new AltaVista window. Next click on a link in the AltaVista results page to simulate finding the information and then close the window.

AltaVista direct use

Starting with AltaVista in one window, and in another window a list of domain names for Spanish universities:

- type "maths AND host:.ua.es", then click on the submit button. Next click on a link in the AltaVista results page to simulate finding the information and then click on the back button to get back to the AltaVista results page, which contains the original search syntax and can be used as the basis for the next university search. (21 keystrokes, 2 links, 1 back)

For the next university, www.ual.es:

- swap windows to identify the next domain name, eg www.ual.es, swap back to the AltaVista page, click between the 'a' and '.' of ua.es to insert 'l' and click to submit the new search (this is the quickest way to modify the query). Next click

on a link in the AltaVista results page to simulate finding the information and then click on the back button to get back to the AltaVista results page. (1 keystroke, 1 click, 2 links, 1 back)

The previous step must then be repeated for www.uia.es, www.uab.es and www.ub.es.

TABLE 1. A comparison of the partner site and AltaVista in terms of user actions to search five Spanish universities

Action	Partner	AltaVista
Press Key	5	28
Click browser back button	0	5
Select option button	1	0
Click web page button or link	11	10
Click to close browser window	5	0
Click to swap windows	0	8
Click in text or click and drag to select query text (to be deleted or overtyped)	0	4

From table 1, the number of actions in AltaVista is much higher. It should be noted that these actions include entering correctly formatted syntax, a cognitively more challenging exercise, even when practised, than clicking a mouse button.

An experienced user may attempt to recover unsuccessful searches by combining search terms with additional advanced syntax in order to get a more precise query, and could do so in the first search results page returned by AltaVista, whichever approach was being used, although this adds an extra step to the process.

Comparison with Google Web Directory

The comparison that will be made here is with the searching of individual sites by Google, rather than using the national university directory global search option to search all Spanish university sites with one query. This latter type of search would require fewer actions, but is excluded on the grounds that it does not allow the pre-selection of sites: the results appear in an order uncontrolled by the user.

The partner site interface does have a slight measurable advantage over the Google directory search in terms of the number of actions required to operate it. This stems from the fact that for each query, the university must be selected from the national university list directory and then the keywords must be typed in. This gives two click actions and one typing action per university, whereas the partner site only requires one typing action and one click for all universities, in addition to a single click for each individual one. The difference between the two will be minimised if the query term is copied and posted rather than retyped from scratch in each Google university entry. Table 2 shows a comparison between the two for five Spanish sites. This is, in fact, only a theoretical comparison because, at the time of testing, Google only had a separate category for six of the Spanish universities, three in the English language category and three in Spanish. In fact, none of these were from the test set and so the study is conducted under the assumption that such categories will be added, or that the same functionality will be added to other search engines. Google does have a single very large page listing many higher education institutions in alphabetical order, but this did not give the necessary national identity information to make it useful, and would still have been unwieldy if it had.

Google

Starting with the Spanish Google Web Directory category for European universities (the fairest possible starting point since Google has separate language categories):

- Select the country España and click submit, then click on the University of Alicante category, type “maths”, select the word and type control-C to copy it (for later use) and click on the submit button. Then click on a link to simulate a search and click on the back button twice to return to the España page.

Next, for each of the four remaining universities:

- click on the university category, click on the query box and type control-C (to paste the query), and click on the submit button. Then click on a link to simulate a search and click on the back button twice to return to the España page.

TABLE 2. A comparison of the partner site and Google Directory in terms of user actions to search five Spanish universities

Action	Partner	Google
Press Key	5	10
Click browser back button	0	10
Select option button	1	0
Click web page button or link	11	11
Click to close browser window	5	0
Click to select a text box	0	5

If the user wishes to search universities in more than one country, then in the partner site this can be achieved in three actions: pressing the browser ‘back’ button from the current country query screen; clicking the name of the new country and then clicking the submit button. It is not much more difficult in the Google Directory, since it has a category for education and a subcategory for Europe. The multi-lingual variations, would, however, be a slight confusing factor.

Conclusions

The study has clearly identified that a web site offering a specialist interface can be measurably superior in some aspects of usability to general search engine strategies. This finding, essentially that less actions need to be performed on the bespoke site for multiple searches, is in addition to the extra information provided, and makes a strong case for the efficacy of such a solution. It would be rash to declare, however, that the approach will be efficacious for all users or types of query. Many different models have been applied to information retrieval²⁹, highlighting the fact that different information needs may benefit from a variety of approaches. The claim that interest groups may be able to aid searches by users of their site by automatically constructing advanced queries with additional terms to clarify context is given credence by research has shown that semantic disambiguation techniques^{17,18} and the opportunity to focus a search after submitting keywords by choosing potentially relevant domains can be fruitful³⁰. In support of this, one search engine has developed apparently similar logic to automatically suggest additions to a search to clarify any ambiguity, constructing a complex combined query³¹. It may well be the case, then, that the approach described here is not only useful for site specific searching, as demonstrated, but also for subject-related searching. Further research is

needed with end users in order to ascertain whether the approach is genuinely viable and useful.

The biggest disadvantage for any given site constructed for a specific user need in the way that the academic partner site was, from a computer-human interaction point of view, is its lack of universality. This is apparent from a task-centred point of view in the need to first identify the appropriate site when its services would be required. It is also evident in the fact that the skills used to execute its searches would be less transferable than if a standard search engine strategy were to be used. The analysis has also not proved that the site is actually effective in practice because it has not included experiments with real users. The point was, however, to demonstrate that the strategy was capable of showing significant gains, at least in some areas, over existing tools. The move towards increasing automation of tasks, as provided here, is the type of facility that is a progressive trend in history³², and so, in general terms, can be seen as an almost inevitable development. If the results of future testing with end users are positive then the technique may be useful for any interest group whose membership is expected to use specific types of searches that would benefit from advanced search syntax and/or additional supporting information.

References

1. Schwartz, C. 'Web Search Engines.' *Journal of the American Society for Information Science* 49(11), 1998, 973-982.
2. Lawrence, S. and Giles, C. L. 'Searching the Web: General and Scientific Information Access.' *IEEE Communications Magazine* 37(1), 1999 116-122.
3. Ching, C. and Chia, H. 'WebYacht: a concept-based search tool for WWW.' *International Journal on Artificial Intelligence Tools (Architectures, Languages, Algorithms)* 8(2), 1999 137-56.
4. Bharat, K. 'SearchPad: explicit capture of search context to support Web search.' *Computer Networks* 33(1-6), 2000, 493-501.
5. Wang J., Jin, X., Yang X. and Zhang, F. 'Distributed and cooperative information retrieval on the World Wide Web.' *Journal of Computer Science and Technology* 15(6), 2000, 611-18.
6. Ying, D., Chowdhury, G. G., Foo, S. and Weizhong, Q. 'Bibliometric information retrieval system (BIRS): a web search interface utilising bibliometric research results.' *Journal of the American Society for Information Science* 51(13), 2000. 1190-204.
7. Terveen, L., Hill, W. and Amento, B. 'Collaborative Filtering To Locate, Comprehend, and Organize Collections of Web Sites.' *SIGART Bulletin* 9(3-4), 1998, 10-17.
8. Prinz, W. and Grather, W. 'Das Social Web Cockpit' *Der GMD Spiegel* 1/2, 2000, 48-51.
9. Chen, H., Chun, Y. and Ramsey, M. 'A smart Itsy Bitsy Spider for the Web.' *Journal of the American Society for Information Science* 49(7), 1998, 604-618.
10. Hersovici, M., Jacovi, M., Maarek, Y. S., Pelleg, D., Shtalhaim, M. and Ur, S. 'The shark-search algorithm - An application: tailored web site mapping.' *Computer Networks and ISDN Systems* 30(1-7), 1998, 317-26.
11. Jackson, M. S. and Burden, J. P. H. 'WWLib-TNG - New Directions in Search Engine Technology.' *IEE Informatics Colloquium: Lost in the Web - Navigation on the Internet*, November 1999, (pp. 10/1-10/8).

12. Brin, S. and Page, L. 'The Anatomy of a large scale hypertextual web search engine.' *Computer Networks and ISDN Systems* 30(1-7), 1998, 107-117.
13. Spink, A. and Xu, J. L. 'Selected results from a large study of Web searching: the Excite study.' *Information Research* 2000. Available: <http://www.shef.ac.uk/~is/publications/infres/paper90.html>. Accessed on January 30, 2001.
14. Sutcliffe, A. G., Ennis, M. and Watkinson, S. J. 'Empirical studies of end-user information.' *Journal of the American Society for Information Science* 51(13), 2000, 1211-31.
15. Jones, S. McInnes, S. and Staveley, M. S. 'A graphical user interface for Boolean query specification.' *International Journal on Digital Libraries* 2(2-3), 1999, 207-23.
16. Jansen, B. J. 'The effect of query complexity on Web searching results.' *Information Research* 2000. Available: <http://www.shef.ac.uk/~is/publications/infres/paper87.html>. Accessed on January 30, 2001.
17. Stibel, J. 'Getting the User to Ask the Right Question and Receive the Right Answer: A Cognitive and Linguistic Science Approach to Searching the Internet.' *Proceedings of the twenty-first National Online Meeting*, May 16-18, 2000. Information Today, Inc. New York, NY, (pp. 425-429).
18. Moldovan, D. and Mihalcea, R. 'Using WordNet and Lexical Operators to Improve Internet Searches.' *IEEE Internet Computing* 4(1), 2000, 34-43.
19. Thede, E. (2000). 'Intelligent Searching: Flavors of Web-based Text Retrieval.' *PC-AI*, 14(4), 35-37.
20. McCallum, A. K., Nigam, K., Rennie J and Seymore, K. 'Automating the construction of Internet portals with machine learning.' *Information Retrieval* 3(2), 2000, 127-63.
21. Thelwall, M. (2002, to appear). 'Extracting macroscopic information from web links.' *Journal of the American Society for Information Science and Technology*.
22. Gomes, M. P. S. F., Vaux, J. H., Ezingear, J-N., Grieve, R. J., Race, P. and Woolgar, S.W. 'Promoting exploitation of university research by SMEs. On evaluating technology dissemination on the world wide Web.' *Industry and Higher Education* 11(1), 1997, 21-27.
23. Europa, 'Fifth framework programme of the European Community for research, technological development and demonstration activities (1998-2002).' 2001. Available: <http://europa.eu.int/comm/research/fp5.html>. Accessed on January 10, 2001.
24. Ellis, D. Ford, M. and Turner, J. 'In search of the unknown user: indexing, hypertext and the World Wide Web.' *Journal of Documentation*, 54(1), 1998, 28-47.
25. Gullikson, S., Blades, R., Bragdon, M., McKibbon, S., Sparling, M. and Toms, E. G. 'The impact of information architecture on academic web site usability.' *The Electronic Library* 17(5), 1999, 293-304.
26. Elkin, J. and Law, D. 'The 1996 Research Assessment Exercise: the Library and Information Management Panel.' *Journal of Librarianship and Information Science* 29(3), 1997, 131-141.
27. Ingwersen, P. 'Web Impact Factors.' *Journal of Documentation* 54(2), 1998, 236-243.
28. Card, S, Moran, T. and Newell, A. *The Psychology of Human-Computer Interaction*, 1983. London: Erlbaum.

29. Beaulieu, M. 'Interaction in Information Searching and Retrieval.' *Journal of Documentation* 56(4), 2000, 431-439.
30. Ben-Shaul, I. Herscovici, M. Jacovi, M. Maarek, Y. S. Pelleg, D. Shtalhaim, M. Soroka, V. and Ur, S. 'Adding support for dynamic and focused search with Fetuccino.' *Computer Networks* ,31, 1999, 1653-1665.
31. Gallagher, M. 'The Changing Landscape of Business Research: How Internet-Based Services are Empowering End-Users and Focusing Information Professionals.' *Proceedings of the twenty-first National Online Meeting*, May 16-18, 2000. Information Today, Inc. New York, NY, (pp. 107-118).
32. Bhavnani, S. K. and John, B. E. 'Delegation and circumvention: two faces of efficiency.' *CHI 98 Conference Proceedings*, 1998. Addison Wesley: New York, NY, (pp.273-80).