Online IR system evaluation: online databases versus Web search engines

Hong (Iris) Xie

The author

Hong (Iris) Xie is an Associate Professor in the School of Information Studies, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin, USA.

Keywords

Information retrieval, Internet, Online databases, Search engines

Abstract

This study evaluated two different types of online information retrieval (IR) systems: online databases and Web search engines, in terms of user generated criteria. It also compares four types of Web search engines: directories, search engines, meta-search engines, and specialized search engines. The results show that three elements are essential to users in the evaluation of online IR systems: interface design, system performance and collection coverage. While participants preferred the ease of use and intuitive interfaces of Web search engines, they also liked the credible and useful information offered by online databases. Based on the discussion of advantages and problems of online databases and Web search engines, implications of for the design of IR systems are further suggested.

Electronic access

The Emerald Research Register for this journal is available at
www.emeraldinsight.com/researchregister

The current issue and full text archive of this journal is available at
www.emeraldinsight.com/1468-4527.htm

Online Information Review
Volume 28 · Number 3 · 2004 · pp. 211–219
 · ISSN 1468-4527
DOI 10.1108/14684520410543652

Introduction and literature review

The emergence of the Internet brings Web search engines to millions of end-users. At the same time, it also facilitates access to online databases. Online database systems and Web search engines are interrelated with one another. On one hand, online database systems, such as Dialog, Lexis-Nexis are called original, or ultimate, search engines, and it is clear that the search engines of today owe much to these originals (Garman, 1999). In addition, Web search engines are entering a new era. They no longer just search old ordinary Web sites. Hock (2002) points out that information professionals should apply the same evaluation techniques to a Web search engine as they have to the traditional online databases since Web search engines offer more than Web pages now. At the same time, these original search engines are moving towards the Web. According to Xie (2003), almost every online database system has its Web version. Dialog recently has announced three new Web-oriented products: Dialog Portals, Dialog PowerPortal and WebTop, a Web search engine (O'Leary, 2000).

Jansen and Pooch (2000) compared searching studies of traditional information retrieval (IR), online public access catalog (OPAC) and Web studies, and they found both similarities and differences among the three. While the three have similarities in the use of advanced features and number of documents viewed, use of traditional IR systems have more in failure rates, session length, query length and use of Boolean operators. Web users show different search patterns compared with users searching traditional IR systems such as online databases (Silverstein *et al.*, 1998; Spink *et al.*, 2001). Wolfram and Xie (2002) identified two IR contexts which are related to online database systems and Web search engines: traditional IR and popular IR. Traditional IR is characterized by selective content inclusion from published and unpublished sources, more sophisticated search features, and is generally used for search topics of a non-personal nature; while popular IR which creates a context that permits easy user access to and use of a variety of full text information resources. The popular IR context has been criticized for lacking credibility in its content, sophistication in its resource organization and retrieval. Feldman (1998) compared the effectiveness of traditional online services to that of Web search engines by professional searchers, and concluded that a Web search engine is preferred for use but online services produced more relevant results.

Refereed article received 21 November 2003
Accepted for publication 28 November 2003

Online Information Review
Volume 28 · Number 3 · 2004 · 211–219

Chen *et al.* (2001) call for the need for both quantitative and qualitative data in evaluating IR systems. Retrieval effectiveness measured by precision and recall, and efficiency represented by time and effort expended are the most commonly used quantitative criteria. Subjects are encouraged to express their likes and dislikes concerning the system as well as to give reasons behind that as qualitative data. According to Sroka (2000), performance evaluation consists of the following criteria: precision, the coverage of the Web, and the overlap among retrieved documents. Main emphasis is given to the precision criterion since recall is nearly impossible to measure. Chu and Rosenthal (1996) evaluated three Web search engines in terms of their search capabilities and retrieval performance. Search capabilities consist of Boolean logic, truncation, field search, word and phrase search and retrieval performance that includes precision and response time. Precision was calculated based on the relevance judgments for the first ten matches. Ding and Marchionini (1996) compared three popular free Web search services based on their features (databases, indexing quality, functionality and usability) and search performances (precision, salience and relevance concentration). In this study, precision was measured based on the relevance judgments for the first 20 hits. According to Schwartz (1998), most studies take a practical approach to assess relevance only for the top ten or 20 of the ranked list, and the researchers not the users make the relevance judgements. Clarke and Willett (1997) created a method to evaluate the recall of the Web search engines in order to consider both recall and precision when evaluating the effectiveness of search engines. Spink (2002) developed a user-centered approach to evaluate a Web meta-search tool and that included effectiveness and usability.

In most of the previous studies, the evaluation criteria are pre-determined. This study intends to evaluate online IR systems, in particular, Web search engines versus online databases from the users perspectives. To be specific, this study addressed the following questions:

- Which type of online information retrieval systems provides higher precision: online databases or Web search engines? Which type of Web search engines provides higher precision?
- What are the benefits of using directories, search engines, meta-search engines, specialized search engines, and online databases?
- What are the problems of using directories, search engines, meta-search engines and specialized search engines, and online databases?

Methodology

A total of 21 undergraduate students who major in Information Resource participated in the study. They were asked to fill in a pre-questionnaire regarding their experiences in using computers and different types of online IR systems. Based on the pre-questionnaire, the average participant's computer skills was 3.57, in which 1 = no knowledge or skills, 2 = beginner, 3 = intermediate, 4 = advanced and 5 = expert. In addition, they were asked to indicate on a 1-5 scale regarding their frequency use of online IR systems, in which 1 = never use, 2 = rarely use, 3 = occasionally use, 4 = often use and 5 = use daily. These subjects were frequent users of directories ($\bar{x} = 3.53$) and search engines ($\bar{x} = 4$), but not frequent users of meta-search engines ($\bar{x} = 2.5$). They have no experience with online databases Dialog and Factiva. Table I presents subjects' frequency use of variety of online IR systems.

The researcher chose two types of topics that were popularly searched and available across different types of Web search engines: a health-related and a business-related topic. Yahoo, Google and MetaCrawler were selected to represent three types of search engines. Dialog and Factiva were chosen to represent one comprehensive and one specialized online database system. The participants were asked to search one health related topic on an online database system (Dialog), a directory (Yahoo), a search engine (Google) and a meta-search engine (MetaCrawler), and a specialized search engine of their own choices. Examples are: "I want to travel to the Amazon rainforest: what diseases could I contract what can cause it and how I prevent it?"; "Treatment for depression, preferably other than anti-depressant medication" etc. They were also asked to search on a business-related topic on an online database system (Factiva), a directory (Yahoo), a search engine (Google) and a meta-search engine (MetaCrawler), and a specialized search engine of their own choices. Examples are: "Why gas prices change and how it affects us?"; "The message behind Gap's 'For every generation' marketing campaign" etc.

Table I Frequency of use of online IR systems

Types of online IR systems	Frequency of use (mean)	Standard deviation
Directories	3.53	1.4
Search engines	4	1.17
Meta-search engies	2.5	1.3
Dialog	1.11	0.32
Factiva	1.14	0.35

Online Information Review
Volume 28 · Number 3 · 2004 · 211–219

After searching two topics on four types of Web search engines and two types of online databases, participants were surveyed about the benefits and problems of using four types of search engines and two online database systems. Open-ended questions were used to elicit information regarding what the participants liked the most and disliked the most about directories, search engines, meta-search engines, specialized search engines and Dialog command search, Dialog guided search, and Factiva. Qualitative methods were used to analyze the benefits and problems of using different types of online IR systems while quantitative methods were employed to analyze the precision of each type of online IR system.

Results and discussion

The results of this study were summarized to answer the research questions proposed in the first section of this paper.

Precision of Web search engines and online databases

All the participants were required to judge the relevance of the top 20 hits of the search results on one health related and one business related topic. The precision of each online IR system was calculated based on total retrieved relevant documents/total retrieved documents. If retrieved documents were more than 20, only the first 20 were evaluated and calculated since the results of all Web search engines are generated based on best match.

Tables II and III present the precision of online IR systems on health- and business-related topics except specialized search engines since participants selected different specialized search engines for their searches. On the health related topics, Google ($\bar{x} = 65\%$) and Yahoo ($\bar{x} = 64\%$) offered the average highest precision while Dialog command ($\bar{x} = 49\%$) and guide mode ($\bar{x} = 51\%$) had the average lowest precision. On the business related topics, Google ($\bar{x} = 57\%$) and Factiva ($\bar{x} = 55\%$) produced higher average precision than Yahoo ($\bar{x} = 50\%$) and MetaCrawler ($\bar{x} = 47\%$).

Table II Precision of online IR systems on health related topic

Online IR systems	Precision (mean %)	Standard deviation
Yahoo	64	0.35
Google	65	0.33
Metacrawler	56	0.32
Dialog command mode	49	0.40
Dialog guided mode	51	0.41

Table III Precision of online IR systems on business related topic

Online IR systems	Precision (mean %)	Standard deviation
Yahoo	50	0.30
Google	57	0.31
Metacrawler	47	0.33
Factiva	55	0.38

Precision is one of the most important criteria to evaluate an IR system performance. The results indicated that it was difficult to conclude whether Web search engines or online databases provided higher precision. While Dialog's precision was lower, Factiva's precision was higher comparing with some of the Web search engines. In addition, precision was topic related on some online IR systems, for example, Yahoo yielded higher precision on health-related topics than business-related topics.

As discussed in the methodology section, the participants of this study are expert users of Google and Yahoo and novice users of Dialog and Factiva. While some participants were good at using Dialog, others had a difficult time using the system. That contributes partially to the low precision of Dialog. The following sections continue the discussion of the issue.

Benefits of using Web search engines and online databases

Each type of online IR system has its advantages and disadvantages. Based on the analysis of open-ended questions, participants liked to use directories, such as Yahoo, for searching and browsing, to be more specific, for the following circumstances:

- When they had problems in getting started; for example one subject said, "If you are having trouble finding out where to start, Yahoo is a good choice".
- When they were not clear about a topic or when they need to browse a topic: "Yahoo is a helpful search engine to use when you are not completely sure what you are searching for. I am a Google user, but when I tried my search on Yahoo, I was happy with my results. I clicked through the hierarchy easily and quickly found what I was looking for, retrieving more results than I had initially expected", one subject explained. Another echoed, "I enjoy Yahoo because I could browse for a topic".
- When they looked for search words to use: "Directories could be a benefit for beginners and for people looking for other search words to use", commented one subject.
- When they needed relevant results: "I agree that an advantage to directories is for subject

Online Information Review
Volume 28 · Number 3 · 2004 · 211–219

searching. It is an interesting thing however, that I got the most relevant results from Yahoo out of all the other types of searches".

Google is one of the most popular search engines. It was highly praised for:

- Its simplicity and ease of use. One subject well described it: "Google is my absolute favorite. I normally use it for everything. I like it because it is simple. You can search for Web, images, groups, directory or news, simply by clicking the corresponding tab. Simplicity is priceless". "What I like about Google is that it is very simple to use. The searcher doesn't have to think about query formulation or reformulation. The query can just be entered in natural language with no proximity operators", another one added.
- Its quickness. One subject claimed: "It is the quickest way to get results. I find what I need about 90 percent of the time".
- Its default operator. "It is a good idea to use AND as the default operator. From the search that I did above, almost all of my results were relevant, but those that were not completely relevant were missing a search term", commented one of the subjects.
- Its volume. One subject explained: "I like Google because for what it lacks in accuracy it gives you in sheer volume of content and this way you can stumble upon one of the esoteric sites and get the answer you need".

Most of the participants had not used meta-search engines before, however, they did like the meta-search engine for a couple of reasons. Most important, meta-search engines provided more options. One subject discussed these options: "I really like the way you can pick and choose which search engines you want to search in the advanced search options". They thought: "Using the meta-search engine is very helpful because it gives you a wide variety of sources". Another relevant benefit was effectiveness: "MetaCrawler is good because it searches a wide variety of search engines. The results are also very effective in the search of information. That is if you get relevant results", one of the subjects cautiously commented. One of the features in MetaCrawler considered quite good was "that you can e-mail the results".

There are a variety of specialized search engines available and sponsored by different organizations, companies and individuals. According to participants: "They are helpful when you are looking for certain types of results. If I am looking for something in particular this is the type of search that I would turn to". In addition to a specific information search, they thought that specialized search engines provided more relevant results than other types of search engines. Here are the quotes from two subjects:

> While I only got four good hits out of 20, they were all extremely relevant for the topic at hand. I would strongly suggest a specialized search engine for use in retrieving information.

> All the hits that I retrieved were absolutely what I was looking for and the number of hits was easily manageable. The trouble with using specialized search engines though is finding them.

Participants in general were annoyed by too many results from search engines, and "specialized search engines are limited with its results". Some of the specialized search engines were also praised for its legitimate results. One subject argued: "because the sources is specialized and 'legitimate' I can trust that information I have is true and reliable".

Dialog has two modes: command and guided mode. Command mode was favored by those who liked to have user control in the search process. One subject wrote his preference:

> If I had to pick between the two kinds of searching, I think that I'd pick the command search because it requires more precise queries and exact commands and allows for easy narrowing, thus returning better results, yet it also tends to return more results if not specific in query.

One way of user control was being able to use the options that offered by the system: "I like Dialog command search because there were many options to narrow and /or broaden my queries. However it is definitely a system that should be used on a regular basis in order to keep your memory fresh with all the search possibilities", commented one of the subjects. Another form of user control was being able to choose databases:

> I do like searching this way because you get to pick your databases you search. This is very helpful, even more helpful than specialized search engines, when you are searching for specialized information.

Effectiveness was the main reason given by those who liked to use the command mode. One subject claimed:

> I like the command search better than the guided search. I think it is effective in commands to find your relevant results. I don't like the tediousness of the guided search. I guess since I am a programmer, I like and relate better to the command search.

Another one added:

> It is easier to type in exactly what I am looking for in commands rather than filling in the boxes for the guided search.

Participants were also happy about being able to obtain precise results. Just as one subject explained:

Online Information Review
Volume 28 · Number 3 · 2004 · 211–219

> What I like about dialog is the ability to retrieve very precise hits. With all the proximity operators available and specialized search features it is much easier to retrieve very precise and detailed information.

Ease of use was the main reason for those favored guided mode. One subject expressed her opinion:

> The guided search was easy to use and anybody could use it. I especially like the drop-down boxes you could fill with your information and get results.

Another subject compared the guided mode with a directory:

> The advantage of this mode is like a directory, except the credibility of these results is extremely better.

Most of the participants liked Factiva for its ease of use. Here is a typical comment:

> I find it easy to search. I also like how you are able to change the category you are searching under, for example, if I wanted to search for articles that just pertained to the United States, I could change the region area to "United States".

To be more specific, participants praised Factiva's intuitive, self-explanatory, well-organized and balanced design, and many options. One subject wrote:

> Factiva is easy to use and very intuitive, although it could be improved by including more information about non-business.

Another one echoed:

> I like the interface, especially the source browser. It is relatively self-explanatory. I like this even more than I like Dialog. I will definitely use this system in the future for research.

Here is one comment after the subject compared Google with Factiva:

> Google has the simplest and most easy to use interface by any novice user. However, if you are looking for a system that has more flash and looks, but is organized well, then Factiva works for me.

One subject commented highly of the many options available in Factiva:

> I enjoy using this system more than Yahoo, Dialog and some search engines because you can choose all the search criteria, for example, what region, what type of publication etc.

To sum up:

> Factiva is a good system. It mixes the command search nicely with the guided search.

Finally Factiva was chosen for its unique information. According to one of the subjects:

> It is easy to use, and you receive unique information. The information is a lot different than the information you get from search engines although good information comes from there as well.

Problems of using Web search engines and online databases

Participants also identified problems in using these online IR systems. For directories, major problems were related to the presentation of the categories. One subject discussed his confusion regarding the categories:

> You may not know what category to look for. For example, if you have a topic that could be categorized under several different categories, how do you know which to choose, if you choose the wrong one, you have to start all over again.

Another one added:

> Sometimes the categories are very vague or it is hard to identify what category to choose because there are so many. Also you may not necessarily choose the category that is best for your query and this is a problem with directories.

One subject further explained the reason:

> Directories are difficult to navigate unless you happen to think just like the person who indexed and organized the directory. I find my thinking is often quite different and therefore I find directories very frustrating to use.

To most of the participants, they were clear:

> [. . .] the directory is not that great for narrow searching. It is much better suited for broad subject matter.

Furthermore, they were not satisfied that navigation was required:

> If I knew exactly what I am looking for, I had to figure out where it would be located in the category. I could not jump right in to the specific site. Navigation is required.

Another problem with Yahoo and other search engines is dead links. Here is one typical comment:

> I did not like the fact that some of the results were dead links. I may have had a few more relevant results if all the links worked.

The only problem in using Google mentioned by participants is irrelevant results. One subject expressed her disappointment:

> When I first started using the Internet this would happen to me all the time since I did not know effective ways of searching. No matter how many times you may refine you query to get more accurate results, I believe that you will always retrieval irrelevant results.

Another one compared search engines with directories, and concluded:

> As far as things that I don't like about search engines is the fact that it is just grabbing anything and giving it to you. Where a directory knows all the information that it is searching, the search engines just grabs anything it thinks pertains to your subject. You will get a lot of results that are pointless.

The problems of using meta-search engines are related to the fact that meta-search engines allow

Online Information Review
Volume 28 · Number 3 · 2004 · 211–219

users to search multiple search engines. Participants complained about the duplicated results they received from MetaCrawler. One subject wrote:

> One disadvantage I found with meta-search engines is how sometimes that same results will be listed twice or two different pages from the same site.

Another explained:

> With the Metasearch engine, I found the same results a few times probably due to the fact that different search engines have some of the same information listed.

In addition to duplicated results, inconsistent results is another problem, one subject expressed his frustration:

> What I don't like about the Meta-search engine is each time you search for something new results will come up. One group of results will better than the next. You will never get the same results.

Moreover participants were surprised to find that "it's updating the results quicker than I could copy and paste them; the results would change order or change period within the same search".

Since most of the participants were not expert users of meta-search engines, they disliked the interface. One subject described his experience:

> The interface seemed complicated. I was not sure how to read the results when they were listed.

Some of them thought that:

> [...] MetaCrawler seems to be for advanced users because there are many tabs to click on including audio, directories, multimedia, shopping and others. Google has some of these tabs but they seem be less inconspicuous.

Almost every participant considered it was difficult to find a specialized search engine. "I think I'll just stick with a normal search engine or a meta-search engine to find information since it seems like more of a chore to find an engine", one subject expressed her frustration. Some of them preferred: "Using Google or MetaCrawler to search for multiple keywords retrieves more results without the headache or trying to figure out how to use a specialized search engine".

Participants considered Dialog command mode as not designed for beginner. One subject claimed that "the command usage is horrible for a beginner user". Database selection was the main concern. One subject discussed it in detail:

> One thing that I don't like about Dialog is the fact that you have to read "bluesheets" or scan through subject topics in order to find a database to search. Even if you use Dialog a lot, it seems that you could never really memorize all the databases available, along with their numbers.

Query formulation with different languages was another issue:

> The annoying part is having to reformulate your query because they use a different language than the other ways we learned.

Another one criticized the unique commands used by Dialog command mode.

To the participants, Dialog guided mode had the same problem as directories. One of the subjects talked about the category confusion:

> The problem I have with this type of search is the same that I have with directories because of the category issue.

Time consuming was another issue of Dialog guided mode: "Dialog guided search may take a good deal of time to search. I would not choose this method if I am in a hurry", another one commented.

A majority of the participants liked Factiva. Only the problem of narrow coverage was identified; for example, one subject complained:

> This system does not come in much use unless it is more of a business topic you are looking for which would be one of the disadvantages to this system.

Discussion: online databases versus Web search engines

In this study, participants identified three elements that were essential in evaluating online IR systems: system performance, interface design, and content coverage.

Based on the results it is difficult to claim which type of online IR systems has higher precision. While Google was doing great, Factiva also yielded higher precision than some of the Web search engines. User experience in using the system, the system design, the design of the interface and the volume of information covered in the system all had impact on the system performance. Among all the tested online IR systems, MetaCrawler was one of the systems that had the low precision, and it was also one of the systems that participants were not familiar with. Dialog offered low precision partly because participants were novice users and many of them considered the system difficult to use. Although Dialog offered fewer relevant results, it provided highly useful results. Participants pointed out that although it is difficult to use Dialog, it did produce excellent results. Here is one subject's experience:

> I found Dialog Command mode my favorite. Although it takes time to get more accurate results, the results you get are more useful than the ones from a search engine or directory.

Utility criteria, such as how useful the results are in helping users to solve their problems; and, whether are users satisfied with the results, also need to be

Online Information Review
Volume 28 · Number 3 · 2004 · 211–219

taken into consideration in evaluating an IR system. According to participants, Factiva's ease to use interface and comparatively narrow coverage contributed to its high precision.Interestingly, participants' search topics also influenced a system's precision, especially on directories. Yahoo's precision changed according to the search topics. It was easier for participants to identify categories in health-related topics than business-related topics. Some topics that involved more than one category were more difficult to search than topics that only focused on one category.

Comparatively speaking, online databases have more sophisticated interfaces than Web search engines. Most participants liked the simple, self-explanatory and intuitive interfaces of Google, and complained about the complicated rules and codes in using online databases, especially Dialog's command mode. One subject wrote about his experience:

> Search engines give you simple ways to understand how to search using plain English unlike online databases we've covered that using annoying codes that do not work. I usually have to end up modifying the query either with Boolean operators or with some kind of operators.

Simultaneously, some participants, especially those who were good at programming, loved the control and efficiency they had in using Dialog command mode. One subject discussed how online databases facilitate his query formulation and reformulation:

> Dialog and Factiva generally allow for better qualified query formulation and reformulation. They also outperform search engines with display formats and usefulness of results.

In general, participants were happy that they "could use natural language for query formulations and don't have to pick a database or publication to search in Google". At the same time, their "frustrations with Dialog and Factiva were in finding the correct databases with the correct search parameters. Having an error in either one can lead to zero results". That again raised the issue that system design needs to support both ease of use and user control. To most of the participants, Factiva's interface was well designed. Here are two typical quotes about database selection and query formulation which illustrate the importance of ease-of-use:

> Factiva is the database I like the best. Factiva's databases are easy to access and select. The fact that databases are broken down to where you can select a business name or newspaper makes it unique. Dialog's databases can be difficult at times and can lead you in circles. The directories don't break down into categories like Factiva.

> I like Factiva because it gives you the key words at the bottom which allows me to easily refine my search. Factiva makes it extremely easy to reformulate and change your query. Plus, it isn't too strict with commands, so you could use natural language searching.

Many participants noted that their preferences might change as their proficiency improved. When they become expert users of a specific system, they might start to like it. Basically, to the participants, Web search engines were for novice users while online databases were for experienced users. Just as one subject wrote:

> Dialog's command mode can be one of the most challenging IR systems. However, the results are very good. A novice user will not be able to jump in and start on this system.

Another one added:

> If I were someone who did a lot of searching within a specific field I would want to subscribe to a service like Factiva or Dialog. These services allow for experienced users to streamline their searching by allowing for specific categories, fields, subjects, publication dates, authors etc. if you know the system well enough to manipulate it effectively.

Therefore the design of an online IR system needs to consider the needs of both novice and expert users.

In terms of resources covered in the systems, online databases have more advantages than Web search engines, mainly because the collections are humanly selected in online databases while Web search engines mostly grab what are available from the Web. Online databases were praised for their credibility while Web search engines results were not considered to have the same authority. One subject well discussed it:

> I do feel that in a professional work environment Dialog and Factiva have an upper hand since all sources can be traced to a specific author and publisher. They are more likely to be a credible source than John Doe's Web page.

Furthermore, online databases also provide unique information that is not only unavailable from Web sites but also more in-depth. Just as one subject commented:

> Factiva and Dialog give you information that you cannot usually find on a search engine or directory. I found some information on business or corporations that you cannot find on the Web.

Another one added:

> It is easier to find relevant results using a searching engine than using an online database system, but Dialog and Factiva provide more in-depth information.

At the same time, some participants did complain that the coverage in online databases was not for the general public because they were too technical. One subject explained why she preferred Web search engines to online databases:

Online Information Review
Volume 28 · Number 3 · 2004 · 211–219

> Another reason I prefer Google and Yahoo was that the Web sites that were returned were for everyday people. I could understand the medical talk, where is Dialog, everything was science based and extremely technical.

Driven by least-efforts principle, participants liked to access full-text documents instead of just abstracts. This is one of the reasons that some of them did not think highly of Dialog. "I dislike about this is that some of the results would not display the whole story, just a short abstract. This problem is also apparent with Dialog searching", complained by one subject.

Each type of online IR systems has its advantages and problems. Table IV summarizes the problems and benefits of using online databases and Web search engines. Most importantly, it suggests how these online IR systems can learn from each other and improve themselves.

Overall we need to design online IR systems that facilitate both novice and experienced users to effectively retrieve information. To be more specific, the interface needs to be intuitive to attract novice users. Simultaneously, the system also needs to provide more specific or advanced search options in an unobtrusive way. Therefore, system design can support ease of use as well as user control to satisfy diverse user needs. Another essential component is its performance. Users prefer systems that offer high precision. Furthermore, utility criteria of an IR system are more important than relevant criteria since users conduct searches in order to solve their problems. Not every relevant document plays the same role in helping them fulfill their tasks. Users are more satisfied if they are allowed to rank the usefulness of the top results, and are able to find documents like this or document not like this. In terms of content coverage, an ideal IR system needs to cover information for professionals as well as for everyday people, and it needs to find some ways to offer credible information. More important, more and more users like to access full-text documents so they do not have to make more efforts to find those documents themselves.

Table IV Comparison of online databases and Web search engines

	Online databases	Web search engines
System performance	Varied; Factiva is better than Dialog	Varied; Google and Yahoo are better than MetaCrawler
Interface design	Complicated with more specific field search options, but not easy; Factiva interface is better designed to facilitate database selection and query reformulation	Intuitive and simple, but you cannot conduct sophisticated searches
Collection/coverage	Credible published and unpublished articles, but not all of them are full-text	Full-text Web pages that have no authority control

Conclusion

This study investigated the advantages and disadvantages in using different types of online IR system. Each type of the system has its unique benefits and problems. Directories are useful for those users who do not know how to start or how to formulate their queries, but they need to have similar judgments as indexers. Moreover, their topics cannot be very complicated because that might involve searching in several categories. Web search engines are easy to use, but they have to filter through a large volume of irrelevant results. Meta-search engines have more options and can be effective, but often produce unpredictable results. Specialized search engines provide excellent results, but they are difficult to identify. Dialog's command mode is liked by users who prefer to have more control, but it is too difficult for novice users. Dialog's guided mode is much easier to use, but it has the same problem as directories. Factiva is praised for its ease of use, but its narrow coverage is the problem.

This study identified three important components in evaluating online IR systems. Interface design, system performance and content coverage were considered essential for the design of an ideal IR system. Although participants were more familiar with Web search engines, they did find benefits of using online databases. Based on the results of this study, the precision varied within Web search engines and online databases. Both Goolge and Factiva have high precision. While participants preferred the ease of use and intuitive interfaces of Web search engines, they also liked the credible and useful information provided by online databases.

This study suggests that:

- online IR system design needs to support ease of use and user control for both novice and expert users;
- online IR system design needs to support relevant as well as utility criteria for users' evaluation of results; and
- online IR system design needs to cover information for both professionals and laymen.

Further research needs to continue examining how to design online IR system to satisfy these requirements.

Online Information Review
Volume 28 · Number 3 · 2004 · 211–219

References

Chen, H., Fan, H., Chau, M. and Zeng, D. (2001), "MetaSpider: meta-searching and categorization on the Web", *Journal of American Society for Information Science and Technology*, Vol. 52 No. 13, pp. 1134-47.

Chu, H. and Rosenthal, M. (1996), "Search engines for the World Wide Web: a comparative study and evaluation methodology", in Hardin, S. (Ed.), *Proceedings of the 59th ASIS Annual Meeting: Global Complexity: Information, Chaos, and Control*, Information Today Inc., Medford, NJ, pp. 127-35.

Clarke, S.J. and Willett, P. (1997), "Estimating the recall performance of Web search engines", *ASLIB Proceedings*, Vol. 49 No. 7, pp. 184-9.

Ding, W. and Marchionini, G. (1996), "A comparative study of Web search service performance", in Hardin, S. (Ed.), *Proceedings of the 59th ASIS Annual Meeting: Global Complexity: Information, Chaos, and Control*, Information Today Inc., Medford, NJ, pp. 136-42.

Feldman, S. (1998), "The Internet search-off", *Searcher*, Vol. 6 No. 2, pp. 28-35.

Garman, N. (1999), "The ultimate, original search engine", *Online*, Vol. 23 No. 3, p. 6.

Hock, R. (2002), "A new era of search engines: not just Web pages anymore", *Online*, Vol. 36 No. 5, pp. 20-7.

Jansen, B.J. and Pooch, U. (2000), "A review of Web searching studies and a framework for future research", *Journal of American Society for Information Science and Technology*, Vol. 52 No. 3, pp. 235-46.

O'Leary, M. (2000), "Dialog's new tools for Web-age knowledge workers", *Online*, Vol. 24 No. 3, pp. 91-2.

Schwartz, C. (1998), "Web search engines", *Journal of the American Society for Information Science*, Vol. 49 No. 11, pp. 973-82.

Silverstein, C., Henzinger, M., Marais, H. and Moricz, M. (1998), "Analysis of a very large Alta Vista query log", *SIGIR Forum*, Vol. 33 No. 1, pp. 6-12.

Spink, A. (2002), "A user-centered approach to evaluating human interaction with Web search engines: an exploratory study", *Information Processing and Management*, Vol. 38 No. 3, pp. 401-26.

Spink, A., Wolfram, D., Jansen, B.J. and Saracevic, T. (2001), "Searching the Web: the public and their queries", *Journal of the American Society for Information Science*, Vol. 52 No. 3, pp. 226-34.

Sroka, M. (2000), "Web search engines for Polish information retrieval: questions of search capabilities and retrieval performance", *International Information & Library Review*, Vol. 32 No. 2, pp. 87-98.

Wolfram, D. and Xie, H. (2002), "Traditional IR for Web users: a context for general audience digital libraries", *Information Processing & Management*, Vol. 38 No. 5, pp. 627-48.

Xie, H. (2003), "Supporting ease-of-use and user control: desired features and structure of Web-based online IR systems", *Information Processing & Management*, Vol. 39 No. 6, pp. 899-922.