

# Mental Models and Information Retrieval: What Can Search Queries Tell Us?

**Haidar Moukdad**

School of Library & Information Studies  
Dalhousie University  
Halifax, Canada  
moukdad@is.dal.ca

**Andrew Large**

Graduate School of Library & Information Studies  
McGill University  
Montreal, Canada

---

## ABSTRACT

When information seekers use an information retrieval system their strategy is based, at least in part, on the mental model they have constructed about this environment. A random sample was gathered of more than 2000 actual search queries submitted by users to one web search engine, WebCrawler, in two separate capture sessions. The results suggest that a high proportion of users do not employ advanced search features, and those who do frequently misunderstand them. Furthermore, a significant minority of users seem to have constructed a mental model of the Web that imbues it with the intelligence found in a reference librarian, for example, but not a retrieval system. The linguistic structure of many queries resembles a typical human-human communication model that is unlikely to produce satisfactory results in a human-computer communication environment such as that offered currently by the Web. Design of more intuitive systems is dependent upon a more complete understanding of user behavior at the intellectual and emotional as well as the technical levels.

## RÉSUMÉ

Quand les chercheurs de l'information utilisent un système de recherche documentaire leur stratégie est basée, du moins en partie, sur le modèle mental ils-elles ont construit concernant ce milieu. Un échantillon de la population pris au hasard a été attiré de plus de 2000 interrogations de recherche concrètes soumises par des utilisateurs au moteur de recherche, *WebCrawler*, dans deux sessions reprises séparées. Les résultats suggèrent qu'une proportion élevée des utilisateurs n'emploient pas des particularités de recherche avancées, et que ceux-celles qui les emploient les mal comprennent fréquemment. De plus, une minorité significative des utilisateurs semblait avoir construit un modèle mental du Web qui l'imprègne de l'intelligence qui se trouve chez un bibliothécaire des ouvrages de référence, par exemple, et non pas dans un système d'extraction de données. La structure linguistique de plusieurs interrogations ressemble un modèle typique de communication humain-humain qui a peu de chances de produire des résultats satisfaisants dans un environnement de communication humain-ordinateur tel qu'est offert actuellement par le Web. La conception des systèmes plus intuitifs dépend d'une compréhension plus complète du comportement de l'utilisateur au niveau intellectuel et affectif aussi bien qu'au niveau technique.

## INTRODUCTION

The success of an information retrieval system traditionally has been measured by the extent to which it is able to find for the user all the information required to meet the latter's needs while avoiding any information that is not required to meet this need. Such success is determined by the way in which the information is structured and organized in the database, and by the searching characteristics of the retrieval engine. But it is also dependent upon the ability of the user correctly to utilize these structures and

characteristics to the maximum effect. The ease with which users can accomplish this task in turn is predicated by the accuracy of the mental model that they have formed about the retrieval system.

The theory of mental models emerged from the human-computer interaction field as a metaphor for describing the location, function and structure of objects and phenomena in computer systems (Jonassen, n.d.). According to this theory, the facility which users apply to exploiting the functionality of computer systems depends on their conceptual models for describing the components and interactions of those systems. Mental models develop in the mind of the user, but they can vary, often significantly, from the conceptual model promoted by the designers because of varying prior knowledge, individual abilities and different beliefs about the purpose and functions of the system. Users' interactions with the system are guided by their mental models which influence their ability to predict system performance (Zhang & Chignell 2001). The quality of the interaction depends upon the functionality of the mental models users have formed of the system (Ring, Ellis & Reeves 1994). By investigating the mental models held by users of information retrieval system, we can better design systems that accommodate their behavior rather than trying to change that behavior to accommodate system design.

This paper seeks to illuminate the mental models held of one general-purpose, web-based search engine – WebCrawler – by a random sample of users. Using the actual queries submitted to WebCrawler during two sessions in March 2000, it provides a quantitative analysis of query characteristics and a qualitative evaluation of the queries themselves in an attempt to understand how these users view the web. It also offers conclusions on how the search engines might be adapted better to fit these models.

## **TRANSACTION LOG ANALYSIS ON THE WEB**

If the objective of this study is to investigate users' mental models of the web by examining the queries submitted to a typical web-based search engine, a set of real queries formulated by real users must be collected for analysis. One approach would be to observe users as they conduct their searches, whether under experimental or operational conditions, or to capture on videotape the searches as they are conducted. This latter technique has been employed by many researchers, including Large and his colleagues (Large, Beheshti & Breuleux 1998; Large, Beheshti & Moukdad 1999). Using such a technique, the searchers (and their individual characteristics), the objective of the search, and the results of the search are all known to the researchers. It is also possible to interview the searchers after the session is completed to probe their strategies and tactics as well as their mental models (Large & Beheshti 2000). The consequence of such a direct contact with the searchers, however, is that typically samples are small and are drawn from just one user community (in the above studies, for example, the searchers were all primary school students drawn from two classes in one school). Sasse (1991) argues that data on users' mental models often have been flawed because they are based on artificial and restrictive experimental scenarios and because the samples are too small.

There is also the possibility in such user studies that user behavior has been influenced by the presence of the research team.

In order to obtain much larger, more heterogeneous and therefore more representative samples, researchers on occasion have resorted to an alternative strategy: transaction logs. In such cases the actual queries as submitted to the search engine and stored by the information system itself are accessed for analysis. The searchers cannot be identified by their personal characteristics, and their level of satisfaction with the results of the search as well as the motives behind their search strategies cannot be ascertained. But in compensation for such methodological shortcomings, transaction logs provide the major benefit that very large query samples can be obtained from a cross-section of users in a non-intrusive environment (for a discussion of the strengths and weaknesses of transaction log analysis, see Kaske 1993 and Kurth 1993).

A number of important studies have utilized transaction log analysis to investigate information retrieval on the Web. Hoelscher (1998) looked at a sample of 16 million queries submitted to a German search engine, Fireball ([www.fireball.de](http://www.fireball.de)) in July 1998. Silverstein et al (1998) analyzed just under 1 billion queries submitted to AltaVista ([www.altavista.com](http://www.altavista.com)) over 43 days in 1998. Queries from the Excite search engine ([www.excite.com](http://www.excite.com)) have been analyzed in several studies: Xu (1999) looked at 30 billion queries collected between 1996 and 1999; and Spink and her colleagues (Jansen, Spink and Saracevic 2000; Spink et al 2001; Spink & Xu 2000) analyzed 51,473 queries and 1.2 million queries, both collected in March 1997. Smith, Ruocco and Jansen (1998) analyzed 851,770 queries also taken from Excite. Keily (1997) used smaller samples of 1000 queries from both WebCrawler ([www.webcrawler.com](http://www.webcrawler.com)) and Magellan ([www.mckinley.com/magellan/](http://www.mckinley.com/magellan/)). Jansen and Pooch (2000) provide a comparative table of the findings from the Hoelscher, Jansen et al, and Silverstein et al studies.

These various transaction log studies reveal a number of characteristics about the queries submitted to web search engines. Wolfram (2000) has made the point that "analysis of query-level regularities in submissions to search engines can provide a better understanding of user search characteristics that may not be evident in a term-level analysis." Several studies have found that the number of words per query is small, typically between 2 and 2.5 (Silverstein et al. 1998; Xu 1999; Jansen, Spink & Saracevic 2000). Web searchers are also unlikely to employ the search refinement features provided by the search engines such as Boolean operators, modifiers or phrase searching (Silverstein 1998; Spink et al. 2001). Furthermore, when such features are employed, users often make mistakes; Jansen, Spink and Saracevic (2000) found that in 32% of queries the AND Boolean operator was wrongly used, and this figure rose to 38% with the AND NOT operator. A high percentage of queries also incorporated searching techniques that the search engine (in this case Excite) did not support.

A search session may include more than one query; Jansen Spink & Saracevic (2000) found that an average session included 2.84 queries, but when identical queries that had been re-input are ignored, this figure falls to 1.6. In fact, 67% of users had only 1 query per session, but 14% of users had three or more queries. Spink et al (2001) found

a mean of 4.86 queries per session. During a session, 41.6% of modifications made to earlier queries involved adding new terms, 25.9% subtracting terms, and 32.5% involved no change at all in search terms. In general they conclude that a high percentage of users do not modify queries to any extent, and when they do modify they change some terms but the total remains the same. Silversten et al (1998), however, found that 63.7% of all sessions consisted of only one query.

Several studies have looked at individual term use in queries. In one of the Excite studies already discussed above, out of 2,216,986 terms from over a million queries, 57.1% were only used once, and more than 78% of the terms were used three times or less (Jansen, Spink & Pfaff, 2000).

Jansen, Spink & Pfaff (2000) performed a lexical analysis on the first 511 queries from their total data set of more than a million queries gathered from Excite users. They concluded that generally “users do not apply the normal rules of English syntax in any coherent or consistent manner” and that there is “no ‘language’ to web queries.” They assigned the syntactic structures of the web queries to five categories. Adjective and noun queries comprised by far the most frequent category, accounting for 458 of the 511 queries. Only 14 queries were posed in complete and grammatically correct English sentences, almost all in the form of a question. Slightly fewer (11 queries) comprised verbal phrases, mostly gerunds. A fourth category comprised 13 queries that contained random strings of terms from several lexical groups. The final category – miscellaneous – represented 15 queries: these were URLs, e-mail addresses and grammatically incorrect phrases.

## **WEBCRAWLER**

WebCrawler began in early 1994 as a small single-user application developed at the University of Washington’s Department of Computer Science and Engineering. It was the first full-text search engine to appear on the Web, and quickly became an Internet favorite. In 1995, the engine was purchased by America Online and maintained as a commercial service until it was sold to Excite, Inc., the present owner, in 1996. WebCrawler and Excite are the flagship brands of the Excite Network, which also includes City.net and Magellan. General information on WebCrawler can be found on its Web site ([www.webcrawler.com](http://www.webcrawler.com)). Search capabilities and features relevant to our research are summarized below.

WebCrawler is designed to be simple and easy to use, and provides the user with a simple interface for search and retrieval. This might be the best example of an Internet application that caters to the needs of users whose mental models of search engines include simplicity and friendly interfaces. The search engine encourages users to use “natural language searching” without worrying about search syntax and complex search strategies or completely formed sentences. However, it provides a wide range of search features including Boolean operators and search modifiers. The searches need not be complete sentences. In simple searches, users describe what they are looking for with a series of words or a phrase. The default retrieval mechanism finds results that match any

or all of those words. However, WebCrawler assumes that the user is most interested in results that contain all of the entered words and gives those documents a higher relevance score so they appear at the top of the list.

Capitalization is disregarded in WebCrawler searches, and there is no stemming capability (no automatic stemming or truncation). Although there is no different interface for advanced searching, the engine allows users to enter complex search statements and more focused searches using the following search features:

- The Boolean operators (AND, OR and NOT): The operators must be entered in ALL CAPS
- AND is entered between two terms to retrieve pages containing both of them.
- OR is entered between terms to retrieve pages that include either of the terms or both
- NOT is entered between two terms to retrieve pages that include the first term but not the second.
- In queries of two terms or more, the modifier + (plus sign) is attached to the beginning of a term to indicate that the term must occur in a retrieved page.
- In queries of two terms or more, the modifier – (minus sign) is attached to the beginning of a term to indicate that the term must not occur in a retrieved page.
- ADJ is entered between two terms to retrieve pages that include the two terms in sequence. Using the modifier "..." to group terms has also the same effect, where these terms are treated as a phrase.
- Parentheses are used to create nested Boolean arguments, where terms and operators included within parentheses are treated as one term – e.g. Canada NOT (France OR Belgium) retrieves pages containing Canada but NOT France or Belgium.

## **METHODOLOGY**

In 1999, WebCrawler started providing random samples of user queries on its site through Search Voyeur, a Java-based application that continuously displays, in random process, actual search queries by real users, but does not associate the queries with particular users.

The data used in this research were collected in two half-hour sessions on WebCrawler Search Voyeur on February 12, 2000. The first session occurred from 11:15 to 11:45 AM (EST) and the second from 7:45 to 8:15 PM (EST). Queries were captured at the approximate rate of 36 per minute, resulting in a total of 2067. This might be a relatively small sample by log transaction standards, but it is still a considerable volume of data when compared to traditional data capturing methodologies. Given the anonymous nature of this query-capturing method, it is impossible to determine the origins of the queries and, therefore, the geographic location of the users. That said, the evening session is likely to be dominated by North American users, especially those on the East Coast.

## DATA ANALYSIS

The captured queries were loaded into a Microsoft Access database. A query is defined as a single string of letters, numbers or other symbols submitted to WebCrawler from its search box. It may or may not include correct or incorrect/inapplicable search modifiers like Boolean operators. A term within a query is defined as any unbroken string of characters – letters, numbers or symbols.

Each query was then examined and tagged for the following characteristics:

- Use of individual WebCrawler search modifiers, whether applied correctly or incorrectly (three Boolean operators and +/- signs, capitalization of first letter in a term, quotation marks, parentheses)
- Incorrect use of any WebCrawler search modifier
- Number of individual terms in the query
- Linguistic category to which the query can be assigned
  - Key word(s) – including total number of keywords
  - Complete sentence including number of words in sentence
  - correctly/incorrectly formed
  - imperative mood
  - question without final question mark
  - question followed by question mark

There were many likely examples of spelling mistakes (for example, whht was the first comercial plane) but these were not counted because in some cases it was impossible to be certain whether or not a mistake actually had been committed (for example, the sole keyword in one query, Antartida, probably was a misspelling of the southern continent, but it is just possible that it is quite a different word.

Nor were non-applicable query modifiers counted (eg \”fiber optic internet service provider\” where the slash is not a modifier on WebCrawler).

## QUANTITATIVE ANALYSIS

Table 1 gives the basic statistics related to user queries and their search terms. Four statistics are provided:

1. The number of single-term queries (sum of queries where the user entered only one term as defined above).
2. The number of multi-term queries (sum of queries where the user entered two or more search terms)
3. Number of used terms (sum of the terms used in the queries)
4. Mean of terms (the mean number of terms used in the queries)

**Table 1. Queries and terms**

<b>Total number of queries</b>	<b>Single-term queries</b>	<b>Multi-term queries</b>	<b>Number of used terms</b>	<b>Mean of terms</b>
2067	491	1576	6962	3.37

In Table 2, we further break down the number of terms in queries to relate these numbers to the number of queries they represent and extract the percentage of occurrence.

**Table 2. Percentage of term count in queries**

<b>Term count</b>	<b>Query count</b>	<b>Percentage</b>
>13	16	0.77
13	9	0.44
12	9	0.44
11	28	1.35
10	41	1.99
9	42	2.03
8	65	3.14
7	82	3.97
6	76	3.67
5	84	4.06
4	150	7.25
3	367	17.75
2	607	29.37
1	491	23.75

The use of Boolean operators and search modifiers (search features) is quite popular among WebCrawler users. One or more of these features were used in 797 queries (38.56% of all queries). The search features were used correctly in 412 queries (51.69%). In addition, of the remaining queries, 474 (22.94% of all queries) included the use of capitalization, although this is ignored by the engine. Table 3 shows the use of search features and the percentage of their correct use (correct use indicates that the user has conformed to the rules for constructing queries as specified by WebCrawler).

**Table 3. Use of search features**

Feature	Number of queries	Percentage (all queries)	Correct use	Percentage
Boolean Operator(s) only	162	7.84	58	35.80
Modifier(s) only	593	28.7	329	55.48
Mixed use	42	2.03	25	59.52
Total	797	38.56	412	51.69

Table 4 includes statistics on the use of fully formed sentences and the tone expressed in them. A sentence is considered complete when stands on its own as a meaningful grammatical unit, and it is considered correct if it does not have any grammatical mistakes. The imperative tone of a sentence is detected when the user asks the engine to do something, e.g. "find me" or "tell me".

A question is any sentence that is formed to convey an interrogative tone in English whether it ends with a question mark or not. If the user entered a question mark character (?) at the end of the query, the query is classified a question mark sentence.

**Table 4. Use of sentences**

Complete sentence	Imperative tone	Correct sentence	Question	Question mark
421	12	415	414	269

## THE QUERIES

The advanced search features on WebCrawler are much less prominent than on many other web search engines. Instead of offering them as an alternative interface to the regular search interface on the opening screen, they must be entered in the regular search box but only after they have been "discovered" by first selecting Help, hidden at the foot of the opening page, and then by choosing Advanced Searching. Nevertheless, a significant number of queries do incorporate a variety of advance features, that are both correctly and incorrectly applied. The most common misunderstanding was the inclusion of a redundant + sign (to indicate that the term must always be present in retrieved hits) with just a single keyword (although, of course, this will not adversely affect the search).

WebCrawler itself is not entirely clear on the model it is presenting to users. The size of its search box certainly does not encourage long queries, whether strings of keywords or complete sentences. Its instructions to users under "Basics of WebCrawler Searching", however, begin by informing users that "natural language searching" is supported "so that users can type their searches in plain English without worrying about mastering complex search syntax". It then goes on to add, "you just need to be able to describe what you're looking for with a series of words or a phrase". Perhaps any users



bothering to consult the Help sections can be forgiven for some confusion in their understanding of WebCrawler's searching expectations.

Users with knowledge of the size and heterogeneous nature of the Web's resources nevertheless might hesitate before entering just one single keyword as a query. Yet a significant minority of queries do fall into this category. The likelihood of enormous and largely irrelevant recall is increased when the single keyword is broad and/or ambiguous in meaning. Examples from the query set are terms such as travel, bees, electricity and gardening, all of which alone constituted queries. At the other extreme, though much less common, were the queries that defied the limited size of the search box by extending beyond 10 keywords (for example, one query included the following terms: electronics, computers, audio, video, music, health, professional, stocks, fashion, games, travel, business...).

Multiple term entry, excluding complete sentences, can be either a series of individual keywords with no syntactic relationships between them (for example, +frame +drum +technique) or, more commonly, phrases. The phrase queries can be divided into two kinds. There are many examples where the first term in the phrase modifies the second or subsequent term. A common form would be found in names of people or places, where the last term is modified by the earlier term(s) as in martin luther king, troy michigan, or where an adjective precedes a noun (crochet patterns, or irish genealogy forums). The second type of phrase is distinguished by its inclusion of link words as well as nouns and adjectives. Examples are address of darva conga, sites in cuba, how to make a bibliography. In a few cases, more than one phrase is input as a single query: for example, school inspections Parbold Douglas Church of England Primary School.

In the above cases, despite the fact that some of the queries will prove more successful than others in meeting the information need, the mental model held by the users appears to be of a retrieval system that expects individual words, or at least short phrases on which to search. The users have grasped the requirement to formulate their needs in a small number of key terms. The most surprising finding from this examination of sample Excite queries, however, is the large percentage that comprise complete sentences, often well formed and even including correct punctuation. The sentences take a number of grammatical forms. Most commonly, they are framed as questions (with or without a final question mark):

What is space?

Where can I buy western wedding clothes on line

Are there any resorts at surfers paradise australia

How does speech recognition technology work?

What are the effects of smoking and drinking during pregnancy?

How do I install/download direct x6.1

Where can I find computer software that shows me how to design floor plans for a house?

What should I name my new puppy?

What is 3 times 5?

How do you use a suppository?

Here we have questions posed to WebCrawler very much as they might be presented to a reference librarian. Indeed, the mental model of the reference encounter in a library or even the caller to a local radio advice show seems a more suitable one than an encounter with a keyword-based remote search engine.

This sense of a personal information encounter between the user and WebCrawler is heightened in cases where the user appears to assume that WebCrawler knows something about the user's personal situation – especially in terms of place and time. Examples include:

Where can I find a list of names of people buried in a cemetery

Where can I register to vote?

Where can I buy a barbacue grill

Where is the dog show

Where can I find a neural therapist?

How high is the smith tower?

What do the numbers inside the recycling symbol on the plastic containers mean?

What time will the sunset occur today in california?

Where can I find the most up to date information on Mikhail Lomonosov

In at least one example, however, the user does try to help WebCrawler in pinpointing time and place:

What is ironic about the death of the Regent Morton? Time period is “the during the age of Elizabeth.

A number of queries suggest that the user sees the Web as a national (and typically United States) rather than an international information resource:

What are the results of the primarys?

What is McCain's voting record on abortion

Not all complete-sentence queries are framed as questions. Although less frequent, some queries are issued as commands using the imperative mood – here WebCrawler is perhaps likened to a trusty retainer (as in the Jeeves metaphor used by one search engine):

Find tv program “friends”

Find me the INS

Give me a flag of peru , large one

Yet another variant is the user who expresses to the search engine a note of urgency or even desperation:

I need info on martyrs

I need names of hotels in amsterdam holland

I need ideas for a model to do of ancient Rome

Or the user who thinks it more appropriate to coax information from WebCrawler, or otherwise to soften the hardness of the encounter:

Tell me about coeds in the fifties  
Do you know where I can find walkthroughs for gameboy games?  
Can I have some info on places I south america

## CONCLUSIONS

The purpose of this paper was to explore the mental models formed of the web in general, and search engines in particular, by examining a sample of queries submitted to one such search engine. Based on the ways in which WebCrawler searches as well as its instructions and suggestions to its users, an accurate mental model would see the Web as an enormous mass of information on every conceivable topic that is best searched by inputting a number of precise keywords linked by operators. This model is indeed followed by some users. In many cases, however, it is replaced by very different models. At one extreme is the query that comprises one very general term, like history, that is unlikely to prove helpful whatever the user's actual information need. This is a variant on the traditional reference question that has been posed broadly even though the user in fact has a rather precise information need. At the other extreme are found the large number of queries that do not use keywords but rather are formed as complete sentences – usually questions, but on occasions commands or supplications. Here the mental model is almost anthropomorphic – WebCrawler is approached rather as a human expert who is being asked to find information.

Only one major web-based search engine – Ask Jeeves – encourages complete sentence queries, yet if our queries are in any way typical, this may be a more accurate reflection of many users' mental models than the much more common keyword-based search engines. There seems to be a need either to re-engineer users' mental models or the search engine algorithms.

The extensive use either of single keywords or complete sentences also casts doubts on the efficacy of relevance feedback and sophisticated ranking techniques. There seems little point in attempting to refine searches that from the outset are doomed to trouble. If the problem lies with the initial query then the solution must be found at this first step in the search process and not further down the line. Natural language processing of sentences has proven extremely difficult, especially in broad subject domains, but it remains the most obvious way to bring user and system models into line.

## ACKNOWLEDGEMENTS:

This research was supported by a grant from the Faculty of Graduate Studies, Dalhousie University.

## REFERENCES

- Hoelscher, C. 1998. How Internet experts search for information on the Web. Quoted in Jansen, B.J. & Pooch, U. 2000, Web user studies: a review and framework for future work. *Journal of the American Society for Information Science and Technology* 52 (3), 235-246

- Jansen, B.J. & Pooch, U. 2000. Web user studies: a review and framework for future work. *Journal of the American Society for Information Science and Technology* 52 (3), 235-246
- Jansen, Bernard J.; Spink, Amanda & Pfaff, Anthony. 2000. Linguistic aspects of web queries. *ASIS 2000: Proceedings of the 63<sup>rd</sup> ASIS Annual Meeting, Volume 37, Chicago, November 2000*. Medford: Information Today, 169-176
- Jansen, Bernard J; Spink, Amanda & Saracevic, Tefko. 2000. Real life, real users and real needs: a study and analysis of user queries on the web. *Information Processing and Management* 36 (2), 207-227
- Jonassen, David H. n.d. *Operationalizing mental models: strategies for assessing mental models to support meaningful learning and design-supportive learning environments*. (<http://www-csc/95.indiana.edu/csc195/jonassen.html>)
- Kaske, N. 1993. Research methodologies and transaction log analysis: issues, questions and a proposed model. *Library Hi-Tech* 11 (2), 79-86
- Keily, L. 1997. Improving resource discovery on the Internet: the user perspective. *Proceedings of the 21<sup>st</sup> International Online Information Meeting*. Oxford: Learned Information, 205-212
- Kurth, M. 1993. The limits and limitations of transaction log analysis. *Library Hi-Tech* 11 (2), 78-86
- Large, Andrew & Beheshti, Jamshid. 2000. The web as a classroom resource: reactions from the users. *Journal of the American Society for Information Science* 51 (12), 1069-1080
- Large, Andrew; Beheshti, Jamshid & Breuleux, Alain. 1998. Information seeking in a multimedia environment by primary school students. *Library and Information Science Research* 20 (4), 343-376
- Large, Andrew; Beheshti, Jamshid & Moukdad, Haidar. 1999. Information seeking on the web: navigational skills of grade-six primary school students. *Proceedings of the 62<sup>nd</sup> Annual Meeting of the American Society for Information Science, Washington DC, 31 October-4 November 1999*. Medford: Information Today, 84-97
- Ring, Geoff; Ellis, Rod & Reeves, Thomas C. 1994. *Mental models research and human computer interface design*. (<http://cleo.murdoch.edu.au/gen/aset/confs/iims/94/qz/ring2.html>)
- Sasse, M.A. 1991. How to t(r)ap users' mental models. In M.J. Tauber & D. Ackerman, eds. *Mental Models and Human-Computer Interaction*. Amsterdam: Elsevier, vol. 2
- Silverstein, Craig; Henzinger, Monika; Marais, Hannes & Moricz, Michael. 1998. Analysis of a very large AltaVista query log. *SRC Technical Note 1998-014* (<http://www.research.digital.com/SRC/>)
- Smith, T., Ruocco, A. & Jansen, B. 1998. Digital video in education. *Proceedings of the 30<sup>th</sup> ACM SIGCSE Technical Symposium on Computer Science Education*, 122-126
- Spink, Amanda; Wolfram, Dietmar, Jansen, B.J. & Saracevic, Tefko. 2001. Searching the Web: the public and their queries. *Journal of the American Society for Information Science and Technology* 52 (3), 226-234
- Spink, Amanda & Xu, Jack L. 2000. Selected results from a large study of web searching: the Excite study. *Information Research* 6 (1) (<http://www.shef.ac.uk/~is/publications/infres/paper90.html>)
- Wolfram, Dietmar. 2000. A query-level examination of end user searching behaviour on the Excite search engine. *CAIS 2000: Dimensions of a Global Information Science. Canadian Association for Information Science Proceedings of the 28<sup>th</sup> Annual Conference* (<http://www.slis.Ualberta.ca/cais2000/wolfram.htm>)
- Zhang, Xiangmin & Chignell, Mark. 2001. Assessment of the effects of user characteristics on mental models of information retrieval systems. *Journal of the American Society for Information Science and Technology* 52 (6), 445-459