

# Suggestions for a research framework for South Africa: how can we learn from web information seeking/searching studies?

Ina Fourie

Department of Information Science, University of Pretoria, Pretoria, 0002, South Africa  
fouriei@postino.up.ac.za

Received: 10<sup>th</sup> December 2002

Revised: 4<sup>th</sup> April 2003

*The growing interest in web information seeking/searching studies, and their importance for web site design, search engines, intranets, portals, etc. warrants a research agenda for web information seeking/searching studies in South Africa. A research framework directs research projects and should include: the rationale for and purpose of a study, descriptive information, links to related studies and traditional information retrieval, links to research paradigms and theoretical models, research methodology and methods, rationale for analysing the data, actual statistical analysis, interpretation and findings, and explicit reference to the practical impact of and need for further research. It should build on the importance of research in a particular context, and acknowledge the existing body of research literature. Suggestions are offered for practical and theoretical research projects relevant to the South African context. Although we can benefit from isolated projects, a research agenda, as suggested, may add greater impetus to our research efforts and understanding of the phenomenon under investigation.*

**Keywords:** web information seeking; web information searching; web studies; research framework; research agenda

## Introduction

The exponential growth in web sites, web users and access to web-based information resources is an indication of their growing importance. The World Wide Web (WWW) is used by a diverse group comprising scientists, researchers, academics, students, information technology (IT) specialists, members of the general public, schoolchildren and elderly citizens. A growing interest is reported for the last three groups (Bilal, 2000, 2001; Burwell, 2001) and in information seeking for everyday needs (Carey *et al.*, 2001) and decision-making (Choo, 2001, 2002). In South Africa there is also a growing interest in the web as an information source. Apart from its potential to meet academic and business information needs, it can also be an important resource filling the everyday information needs of South Africans. Spink and Cole (2001:303) note that:

[The] Internet is the driving force behind broadening LIS analysis of information seeking beyond work and school. The interactive potential of this hybrid information flow channel should bring the power of information use to many more sectors of society than is now the case, but it also forces researchers interested in these issues to take a wider, more integrative approach to studying information seeking and use that includes nonseeking behavior in its human information behavior perspective.

Web information seeking/searching studies can shed light on the design of user interfaces, search engines, navigational features, online help and intelligent agents, information architecture, content description and metadata, the teaching of information skills (e.g. to schoolchildren or senior citizens) and the refinement of our information retrieval (IR) research methodology. The findings may also be useful in web advertising, marketing and e-commerce (Ozmutlu *et al.*, 2002), and in drawing interest from designated target groups (e.g. getting teenagers interested in using the library). By understanding the variables (also called barriers or influencing factors) that affect web information seeking/searching we can improve all facets of web information spaces. Many such variables have been identified and tested in traditional IR (Wilson 1999b). More recently Hertzum *et al.* (2002), for example, reports on the notion of trust in the information seeking behaviour of software engineers with special reference to virtual agents and e-commerce. Spink and Xu (2000) conclude that 'continued research into Web user behaviour is needed to impact the development of new types of user interfaces and software agents to aid users in better Web searching'.

Information seeking is also important in the context of virtual or online learning environments (e.g. Sheremetov & Arenas, 2002; Yang, 2001). In the South African context this is particularly important since a number of tertiary institutions have already adopted virtual learning environments to supplement their teaching (e.g. the University of Pretoria), and because of the national emphasis on information literacy skills (addressed in the skills prescribed by the South African Quality Authority). It is therefore clear that not only libraries and information services but also society at large will benefit from a research programme for web information seeking/searching behaviour.

This article builds on a paper presented at the ProLISSA 2002 conference (24-25 October 2002, Pretoria, South Africa) (<http://www.dissanet.com>). The conference paper presents a literature review of web information seeking/  
SA Jnl Libs & Info Sci 2003, 69(2)

searching studies for the period 2000 to July 2002, suggestions for a research framework and research topics of interest to South Africa. To ensure extensive coverage the following databases were searched: *ERIC*, *Wilson Library Literature*, *Information Science Abstracts*, *Emerald*, *ScienceDirect*, and the *ISI* citation databases. *Library and Information Science Abstracts* could not be accessed. References in bibliographies were also followed up. This article will expand on the initial research framework, and will suggest a more detailed research agenda for South African web information seeking/searching studies. The suggestions will be linked to benefits to be gained from web information seeking/searching studies and lessons learned from the existing research corpus. In addition to the databases researched for the conference paper, *South African Studies* was searched to identify literature relevant to a South African research agenda.

This article will cover the following: clarification of key concepts, the need to link WWW information seeking studies to traditional Boolean IR studies, a brief overview of the status quo of web-information seeking/searching studies, benefits to be gained from web information seeking/searching studies, and suggestions for a research framework for South Africa.

### Clarification of concepts

Researchers are often accused of not qualifying their use of concepts. Among other things this makes it difficult to verify or compare research results. It is therefore essential to clarify key concepts such as information seeking, information searching, web spaces and web logging for the purposes of this article.

Information seeking and information searching both belong to the broader field of user studies. According to Wilson (1999b), information seeking and information searching models function in a complementary manner and address issues at various levels of information behaviour. For the purposes of this article we will adhere to Wilson's (1999b:361) explanation of information seeking/searching:

[I]nformation behaviour may be defined as the more general field of investigation ... with information-seeking behaviour being seen as a sub-set of the field, particularly concerned with the variety of methods people employ to discover, and gain access to information resources, and information searching behaviour being defined as a sub-set of information seeking, particularly concerned with the interactions between information user (with or without an intermediary) and computer-based information systems, of which information retrieval systems for textual data may be seen as one type.

The WWW may be considered a new information space that differs from traditional information spaces such as traditional online databases, OPACs and in-house databases. According to Wang *et al.* (2000), the WWW as an information space consists of:

- Digital objects (including content, expression, relation, structure, hyperlinks)
- Activated objects (e.g. portions of the web that have been activated during interaction, such as home pages that are stored in the cache of a local computer)
- Web spaces (i.e. partitioned spaces of the web, such as search engines, directories or specific websites). These would also include home pages, intranets, portals and specific services offered (e.g. document delivery services)
- Organisational schemes (e.g. how each space is organised)
- Metadata (descriptive and representative data attached to an object and which provide access to the object)

The WWW can also accommodate traditional information "spaces" such as in-house web-enabled databases, databases available through commercial database and aggregator services such as Dialog, library catalogues and CD-ROM databases available through an intranet. Libraries and information services should especially monitor users' navigational behaviour in identifying and selecting information channels accessible from their web sites (e.g. databases, library catalogues, full-text journal services, suggested search engines, document delivery services and current awareness services). Apart from research on web-enabled library catalogues (Cooper, 2001; Cooper & Chen, 2001), no reports could be traced on "traditional web-enabled spaces". Any of the above-mentioned web information spaces should actually be covered in an extensive research agenda.

Web log analysis is used in most web information seeking/searching studies to collect and analyse information on users' actions and moves. Other terms that are found in the literature include web logging, web log file analysis, transaction log analysis, web tracking and log file analysis.

### Is it necessary to link to Boolean IR research?

There are different opinions on whether web information seeking/searching reflects totally new information behaviour, or whether it is similar to traditional IR. Although there is a great deal of overlap between traditional IR and the web, there are researchers such as Jansen *et al.* (2000) who are of the opinion that the Internet is being searched in a new way – perhaps users are even searching the web in ways that designers and IR researchers have not yet contemplated or assumed to exist. According to these authors, real-life Internet searching is changing information retrieval:

While Internet search engines are based on IR principles, Internet searching is very different from IR searching as traditionally practised and researched in online databases, CD-ROMs and online public access catalogs (OPACS). Internet IR is a different IR, with a number of implications that could portend changes in other areas of IR as well.

However, Jansen *et al.* (2000) also stress the fact that we need more comparisons between traditional IR and web search findings. This would include mostly Boolean based IR research such as reported by Borlund (2000).

A few obvious differences between the web and traditional IR include a difference in the user base, much wider accessibility and interest, a variety of digital formats (text, hypertext, sound, video, animation), different organisational schemes, different access methods, different document collections (e.g. games, videos, e-journals, e-books, music, images), different information needs and topics of interest (e.g. entertainment and recreation) (Spink *et al.*, 2001), and a lack of control (e.g. of vocabulary, content description, quality, scope, online guidelines). Although such differences, as well as opinions on the unique nature of the web, should be noted, it is of the utmost importance that our research efforts should build on what we have gained from traditional IR research. This point of view is supported by Jansen and Pooch (2001) and Vakkari (1999: 819) who remarks: 'A necessary condition for this understanding is to link results from information seeking studies to the body of knowledge by IR studies.'

Unfortunately current web information seeking/searching studies appear, on the whole, to lack reference to traditional IR research projects, theories and models of information behaviour. This is in spite of the many useful publications that could support the theory building of web information seeking/searching studies, e.g. Ingwersen (1992, 1999), Palmer (1999), Schamber (2000), Wilson (1999a, 1999b), Vakkari (1998, 1999, 2001), Belkin *et al.* (1995), Spink (1997), Kuhlthau (1991), Dervin (1999), Marchionini (1995), and Ellis and Haughan (1997). Some of these models have already been empirically tested in general and in the web environment (e.g. Kuhlthau & Tama, 2001; Byron & Young, 2000), but they should be used much more extensively.

In a recent joint project, Spink, Wilson, Ford, Ellis, Lam, Burton and Foster investigated mediated information seeking. In a series of articles they report on successive searching (Spink *et al.*, 2002c), uncertainty and its correlates (Wilson *et al.*, 2002), cognitive styles in information seeking (Ford *et al.*, 2002), user-intermediary interaction (Ellis *et al.*, 2002) and their theoretical framework and research design (Spink *et al.*, 2002b). These articles could also offer substantial theoretical support for web information seeking/searching studies.

### **Status quo of web information seeking/searching studies ( a brief overview)**

Before planning a research agenda it is essential to note past projects and progress. Since the early 1990s there has been a steady growth in research projects on web information seeking/searching. In 1993, Peters noted that there was still very little research on internet searching (which is in line with the fact that it was still the early days of the more general availability of the Internet). In 2001, Jansen and Pooch published a review of web-searching studies. They list 41 sources, of which six were published in 2000. Fourie (2002) covers a much more extensive list of publications for the period 2000 to July 2002. Since the publication of her paper a number of additional articles have appeared, namely Pu *et al.* (2002), Spink, Ozmutlu and Ozmutlu (2002), Pettigrew *et al.* (2002), Cooper (2002) and Yu and Roh (2002). Jansen *et al.* (2000) noted that there is a lack of *scholarly* research in particular.

Jansen and Pooch (2001) distinguish between primary web-searching studies (investigating web search engines as primary web IR systems) and secondary web-searching studies (i.e. studies providing a limited amount of web user data or that focus on a narrow aspect of web user searching). The latter includes studies of information seeking on a single web site (not a search engine), multiple web sites, searching using relevance feedback, multimedia searching, query terms, presentations on web searching, and information on the web about the web.

Fourie (2002) categorises web information seeking/searching studies according to their research foci. She distinguishes research reports according to the following: single search engines, single web sites (not search engines), comparisons between similar web sites, unspecified searching of multiple websites including search engines, analysis of aspects of web searching (e.g. the searching of specific entities or media types, search queries as focus, users' perceptions, web navigation, search strategies), influencing factors (e.g. gender differences, age groups, age differences, levels of experience, cross-country and cultural differences, trust, network competence, cognitive, affective and task influences, specific occupations, disciplines or domains, usability), innovative monitoring tools, special software and search support, theory building and the development of behavioural models, reviews of trends, and longitudinal studies.

Fourie (2002) also analyses web information seeking/searching studies according to trends that could be identified by making a comparison with the earlier study by Jansen and Pooch (2001). The changes and trends concern the following and are reported in more detail by Fourie (2002): metrics, experimental situation, focus and nature of studies, research methods, shifts in research paradigms, study samples, duration of research projects, findings and conclusions, calls for the

refinement of research methodology, references to and comparisons with traditional IR, web spaces and web objects studied, conceptualisations, target groups, types of behaviour monitored (e.g. cognitive, affective and physical behaviour), influencing factors monitored, and intervening measures taken (e.g. software to monitor or support the search process). Fourie (2002) notices that there are still very few joint research projects (exceptions are the work for the Excite project: Jansen *et al.*, 2000; Spink & Xu, 2000; Spink *et al.*, 2001). Results from a large-scale, international collaborative research project have, however, been recently reported by Spink *et al.* (2002b, 2002c), Ellis *et al.* (2002), Ford *et al.* (2002) and Wilson *et al.* (2002). Although the project does not focus on web information seeking/searching *per se*, it does offer valuable insights.

South African publications on web information seeking/searching

In the international literature research reports concerning web information seeking/searching also appear in journals for computer science, informatics, education, instructional design and the natural sciences. An extensive literature search was therefore conducted on *South African Studies* to identify South African publications relevant to web information seeking/searching. Only a study by Icheva (1998) was identified. No publications dealing with web logging or web tracking, a key method for collecting quantitative data for web information seeking/searching studies, could, for example, be found. There therefore appears to be a need for an accelerated research programme in South Africa.

### **What can we gain from web information seeking/searching studies?**

Web information seeking/searching studies should offer solutions to praxis, and should help us to build a sound theory. Models of information seeking/searching behaviour can, for example, offer a framework for thinking about a problem, describe activities and/or the causes and consequences of such activities, offer a source of hypotheses to be tested (e.g. relationships between stages in information seeking behaviour), and help to identify gaps in existing research (derived from Wilson, 1999b). The research projects reported in the subject literature highlight a number of benefits that can be gained from such studies. Only a small selection of the benefits to be gained will be briefly considered.

Improved search engines or portals for designated target groups (e.g. children, elderly citizens, minority groups)  
The research by Bilal (2000, 2001, 2002), Bilal and Kirby (2002) and Large *et al.* (2002a, 2002b) shows that children need special support for their web information seeking/searching. Bilal, for example, found that Yahoo!igans! (a search engine designed for children aged 7 - 14) is not meeting the needs of their intended target group. The children used incorrect search syntax, misspelled keywords and made extensive use of looped searches and hyperlinks that lead to disorientation (Bilal, 2002:110). She suggests that Yahoo!igans! introduce a spell-checker, context-sensitive help, search and browsing instructions and examples and a natural language interface among other things. Large *et al.* (2002a) used children's comments on search engines (also called portals) for younger users to make suggestions for the improved design of such portals. They covered Ask Jeeves for Kids, KidsClick, Lycos Zone and Yahoo!igans! They found that children find entertainment opportunities and lots of colour attractive. They also prefer larger fonts and animation, the vocabulary should be suitable for the intended target group and subject categories should be available along with the keywords (Large *et al.*, 2002a:89-91).

These studies were all conducted with children in the United States and Canada, and one would expect that South African children would also show a need for search engines that acknowledge their unique physical, affective and cognitive behaviour in searching the web. No South African search engines or portals for children could be identified.

In the United States elderly citizens show great interest in searching the web (Burwell, 2001). It might be interesting to verify if there is a similar trend in South Africa, and whether any cultural differences are reflected in dedicated portals. Although research has been reported on the information seeking behaviour of minority groups, this has not been specifically linked to the web. This might be another research focus of interest to South Africans – especially if we can focus on possible cultural differences in order to design appropriate portals.

Improved search features

Research projects on single search engines and comparisons of search engines have shed light on important search features that could be added (e.g. buttons for image searching to the Excite search engine) (Goodrum & Spink, 2001; Jansen *et al.*, 2000; Spink & Xu, 2000). An analysis of the search queries people submit to search engines could also shed light on their use of keywords, combination of keywords, use of search operators and the length of queries (Moukdad & Large, 2001; Ford *et al.*, 2002; Ross & Wolfram, 2000). Other studies that focused on search queries include the work of White and Iivonen (2001), Iivonen and White (2001), Jansen (2000), Lucas and Topi (2002), Dennis *et al.* (2002), Radev *et al.* (2002), Pu *et al.* (2002), and Jansen *et al.*, (2000). Xie (2000, 2002), for instance, believes that in order to design IR systems to support various information seeking strategies across a variety of contexts, it is important to explore what leads to the selection of different types of information seeking strategies. This type of study is essential for businesses,

service organisations, companies with intranets and educational services. South African search engines could certainly benefit from such studies.

#### Improvement of website navigational features

This aspect is strongly linked to the improvement of search features. All website owners/masters can benefit from a better understanding of how people navigate their websites. Such studies have been reported on by Rozic-Hristovski *et al.* (2002) among others. There is also a need for studies of software that can support web navigation or the monitoring of web navigation (e.g. as reported by Danielson, 2002). Libraries and information services that make their catalogues available through the web could find such studies particularly beneficial. Cooper and Chen (2001), for example, report on a study to predict when a person's search on the web-based catalogue would be perceived to be relevant.

Comparisons between similar websites (e.g. financial institutions, consumer sites) can also offer valuable insights. Nicholas *et al.* (2002) report on a comparison between two consumer health websites.

#### Encouraging and improving the quality of information seeking in a working environment

Effective information searching will become increasingly important in knowledge organisations (Choo, 2001; Ellis & Haughan, 1997). Choo *et al.* (1999, 2000a, 2000b, 2000c) report on how knowledge workers use the web to seek information. They developed a behavioural model of information seeking/searching that allows for four modes of searching based on the work of Arguiles, namely undirected viewing, conditioned viewing, informal searches and formal searches. It also allows for six activities (starting, chaining, browsing, differentiating, monitoring and extracting) identified in earlier studies by Ellis, Hall and Cox (1989) and Ellis and Haughan (1997). Based on their findings Choo *et al.*, (2002c: 186-187) suggest that portals should be designed to support undirected, serendipitous viewing, users should be trained to evaluate information and to use web notification services, and good pre-packaged search strategies should be provided. Their findings also have implications for intranet design, which should promote information and knowledge sharing. The value of studies concerning people's work or task-related information seeking/searching behaviour is also pointed out by Vakkari (1999, 2001), Bryström (2002) and Bryström and Järvelin (1995). Their studies do not focus on the web *per se*, but offer valuable insights into issues to consider.

Crowley *et al.* (2002) used focus groups to collect information on users' perceptions of a library's web page. Such studies are essential for improving site quality and accessibility.

#### Finding improved methods to support people's information seeking/searching efforts

A growing interest in ways to support users in their search efforts is noted in more recent research projects. Some of these are the use of constantly visible site maps (Danielson, 2002), self-organising maps (Smith & Alan, 2002) or special software for search support (Dempsey *et al.*, 2000), automated document clustering (Roussinov & Chen, 2001), and information visualisation techniques (Heo & Hirtle, 2001).

#### Developing web information spaces allowing for cultural differences

The possible impact of cultural differences is still very under-researched. Iivonen and White (2001) and White and Iivonen (2001) investigated the initial choice of search strategies and whether there are any cultural differences. They draw a comparison between Finnish and American searchers.

Although not in the context of web information seeking, Xie (2000) argues the need to design IR systems for supporting various information-seeking strategies across a variety of contexts. This requires an exploration of different types of information-seeking strategies and cognitive behaviour. With its rich cultural variety, South Africa could make a valuable contribution to the study of the possible impact of cultural variables. It may for example be interesting to note whether non-native English speakers use the Web as much as English-speaking communities, as well as how they use the web (e.g. choice of search terms, criteria for selection of web resources), how the web is used by communities with a strong oral tradition, and whether the web is equally accessible to all members of the South African community (with acknowledgement to an anonymous referee for some of the suggestions).

#### Supporting collaborative information seeking/searching

There is a growing interest in collaborative information seeking/searching. Fidel *et al.* (2000) and Large *et al.* (2002b) report on gender differences between grade 6 boys and girls in collaborative web searching. Collaborative learning environments are also being mentioned more frequently (Zhang, 2002; Zafeiriou *et al.*, 2001). Due to the potential of the web as an electronic learning environment, it would be worthwhile to pursue studies on collaborative information seeking/searching. The author is currently working on such a project.

### Improved information literacy and information research skills

Information literacy programmes have become very important as is clear from the numerous articles and conference papers on the topic. Information literacy is also linked to programmes on information research skills (Fourie & Van Niekerk, 2001).

Research by Bilal (2000, 2001, 2002) has shown that it is essential to understand people's information seeking/searching behaviour on a physical, cognitive and affective level. Kuhlthau's model of the information search process (ISP model) has shed light on information searcher's affective behaviour (e.g. feelings, emotions and anxiety) in particular (Kuhlthau, 1991, 1993). The importance of the information seeker's cognitive state and cognitive behaviour is stressed by Ingwersen (1992); the research by Vakkari (1999, 2001) and Byström and Järvelin (1995) stresses the impact of task complexity. Savolainen (2002) argues that network competence will affect information seeking on the web.

Against this background it appears essential to improve our understanding of people's cognitive, affective and physical web information seeking/searching behaviour, so that we can adapt information literacy and information research skills programmes to allow people to operate effectively in terms of their search skills and information usage in environments typical of knowledge and learning organisations and also in collaborative work/learning. Cognitive studies can especially be important for research on different cultural groups, and people from formerly disadvantaged communities to monitor for example the impact of a digital divide.

### Differentiating between expert and novice searchers

Traditional IR studies have shown that search experience can impact on search behaviour and the quality of search results. Lucas and Topi (2002), Hölscher and Strube (2000) and Lazonder *et al.* (2000) compared expert and novice searchers in order to identify the types of knowledge and knowledge structures required. Kim and Allen (2002) report on the differences in users' cognition and search tasks, while Palmquist and Kim (2000) investigated cognitive style and experience as predictors of web search performance. If we understand the impact of experience, we can accommodate it in our training programmes, the design of web spaces, interface design, and even our organisational structures (such as metadata and taxonomies) for such web spaces.

### Theory and model building

There are unfortunately still too few research efforts focussing on building a theory of web information seeking/searching behaviour. A need for such efforts can be linked to calls for the improvement of a theory of information behaviour and the improvement of our research methodology (e.g. Smith & Alan, 2002). In an editorial introduction to a special issue of *Information Processing and Management*, focusing on information seeking in context (ISIC), Kuhlthau and Vakkari (1999) stress the need for methodology leading to new approaches to research methods that address broader aspects of users' lives than formerly recognised: 'This is the early stage of what is developing into a significant body of research that holds the potential for making a major contribution to the knowledge base of information science.'

Choo (2001), Choo *et al.* (1999, 2000a, 2000b, 2000c) and Wang *et al.* (2000) have made some progress in contributing to such a theory. Choo and his associates have developed a behavioural model of web information seeking behaviour. It is strongly embedded in traditional IR research findings and models (e.g. the work of Kuhlthau, Wilson, Allen and Ellis). Wang and his associates made an effort to offer a holistic model for user-web interaction. We should certainly aim to add to such research efforts.

### Suggestions for a South African web studies research agenda

In preceding sections the value of web information seeking/searching studies and its status quo have been briefly considered. It is clear that substantial web information seeking/searching studies are necessary to refine our knowledge of web information spaces, their design and maintenance and training-related issues. The same applies to the theoretical basis. To support such research efforts, we need a research agenda (working according to a plan to get things done) with a research framework. Although we can benefit from isolated projects, a research agenda may add greater impetus to our research efforts and understanding of the phenomenon under investigation.

In an earlier conference paper Fourie (2002) suggests a number of both practical and theoretical research topics. These include information-seeking behaviour on specific websites; information seeking/searching related to entertainment and online shopping; comparisons between similar websites; the testing of intervening measures such as training efforts, intelligent agents, filtering mechanisms, adaptive IR systems for a specific South African audience; the effectiveness of metadata and indexing practices; searching on textual as well as non-textual collections (e.g. multimedia, video, images, audio); the effectiveness of site architecture and its impact on information-seeking/searching behaviour; web-based environmental scanning and the use of current awareness services; the impact (if any) of cultural diversity (the variety of cultures in South Africa should make this a rich field for future research); the cognitive processing required by users (e.g. from different cultures); surrogate information that can support relevance feedback; the usefulness of query

reformulation, query modification and relevance feedback; studies focusing on the various decisions web users have to make (e.g. their coping skills and search tactics, or the eventual use made of information); analysis of the different metrics used (e.g. search time, accuracy and search efficiency); refinement of research methodologies; failure analysis and linking it to training programmes; and web information seeking/searching in collaborative learning environments.

When selecting a research topic, researchers should take note of their own “philosophies”, personalities, experience and interests, because these will have an impact on their choice of research methods and research paradigms. Ford (1999) links research approaches to the different information processing styles of individuals.

When formulating a research agenda and framework we should build on traditional IR research experience. We should also learn from past mistakes that have hindered progress in IR research. Ellis (1990:viii), for example, remarks: 'In some respects contemporary information retrieval research comprises a plethora of different approaches and techniques, each with its own adherents and opponents, experts and neophytes. The first encounter with the literature can be baffling and confusing.'

#### A research agenda

A research agenda should serve as a plan of what to do. The following are suggestions for a South African effort. It is based on the preceding sections and the conference paper by Fourie (2002). It attempts to stimulate South African based research efforts that will improve local web information spaces and training practices, and that will provide us with the research experience that is essential in designing effective intranets, portals, collaborative learning environments, etc. The agenda entails the following:

- Literature reviews to establish the status quo of web information seeking/searching studies. Attempts have been made by Jansen and Pooch (2001) and Fourie (2002). These should be supplemented with alerting services and frequent updates.
- A research framework. Attempts have been made by Jansen and Pooch (2001) and Fourie (2002). Spink *et al.* (2002b) also offer useful suggestions. Such frameworks should be revised and improved on an ongoing basis as more experience is gained with web information seeking/searching studies, and as we build our theoretical base.
- An analysis of existing web information seeking/searching models. If necessary new theoretical models can be proposed to be tested empirically with regard to South African target groups. (Fourie is currently working on a conference paper in this regard. The paper will be delivered at the 6<sup>th</sup> Conference of the ISKO Spanish Chapter, 5-7 May 2003.)
- An extensive overview of traditional IR research and findings to support web information seeking/searching studies. It is essential to realise that although web information seeking/searching may be totally different from traditional IR, there is much to be learned from earlier theories, or at least by identifying theories worth testing in a web environment.
- A review and reconsideration of the methodologies used in web information seeking/searching studies. These include quantitative and qualitative methodologies. Although much has been published on web tracking for example, there are still many problems. We need to refine the methods as well as the tracking software. Both established and new methods to gather qualitative data should also be considered, so that we can present 'information rich' results. Wang (1999) offers an excellent review article on methodologies and methods for user behaviour research – the umbrella field for information seeking/searching.
- A study of the impact and suitability of the different research paradigms, namely the systems-oriented paradigm, user-centred paradigm, cognitive paradigm and the socio-cognitive paradigm. This should be mapped to the purposes of reported research projects, and their potential for future projects of interest. This could serve as a guideline to researchers when selecting research methods.
- A detailed survey of web information seeking/searching and IR literature to identify other disciplines that may support our research efforts (e.g. motivational theory, Phenomenology).
- A mapping of the purposes and foci of reported research projects (including the identification of web information spaces e.g. a library catalogue, organisational intranet, subject portal, specific search engines) and of target groups (e.g. elderly citizens, school children, engineers, lawyers). Both the web spaces and target groups should be linked to the purpose of the research project (e.g. to collect data on students' physical, cognitive and affective behaviour in order to improve information literacy training programmes and portal interfaces, or the inclusion of web information seeking in school curricula).
- An inventory of variables (e.g. influencing factors such as culture, age, gender) that may affect information seeking/searching. These should be systematically tested in the South African context.
- A list of projects that lend themselves to longitudinal and comparative studies (e.g. comparisons between children or students from different backgrounds, or the changes in the information seeking/searching behaviour of students over a three year period).
- A list of projects that lend themselves to international, collaborative research efforts.
- A collection of research data to support comparative research projects, or projects where data can be analysed from different perspectives (e.g. the research reported by Spink *et al.* [2002a, 2002b]; Wilson *et al.* [2002], Ellis *et al.* [2002],

Ford *et al.* [2002]).

A research framework

A research framework that serves as a point of departure for research should support a research agenda. Such a framework should help researchers to plan research and to write up their research reports. It should support comparative studies and the development of a theoretical basis and relevant research models. It should also offer guidelines on the content of research reports. Jansen and Pooch (2001:242), for instance, believe that the framework they propose will support comparative and contrasting studies: 'Researchers utilising common terminology and following a similar framework whenever possible will improve the ability to make comparisons among research results.' In their suggestions for a research framework they allow for descriptive information, analysis presentation and statistical analysis.

Fourie (2002) suggests a preliminary framework that allows for the following (more detail can be found in Fourie, 2002):

- Rationale and purpose (these should be clearly spelled out and linked to decisions on research methods and the nature of the data to be collected).
- Descriptive information (e.g. the background information on the web information space being investigated; specific features of the web space as an IR system; demarcation of the study and the rationale for the demarcation; size of the study in terms of the number of searchers and the queries collected; description of the target group; influencing factors considered; the experimental setting; nature of the queries [real-life or controlled, fact related, research tasks, language of the queries, language of the document collection]).
- Links to related studies (especially comparative projects and traditional IR research).
- Links to research paradigms, theoretical models and information seeking/searching models.
- Research methodology and methods (e.g. type of methodology, research methods, type of data collected, research instruments, measuring units, assumptions, relevance judgements and their interpretation, guidelines given to participants.)
- Analysis of data (methods and rationale).
- Actual statistical analysis.
- Interpretation of data and findings.
- Explicit correlations between findings and suggestions for practical implementation and further research.

The framework should be revised as more experience is gained in web information seeking/searching studies.

## Conclusion

Web information-seeking/searching studies are still at an early stage. Although some progress has been made, there are still a number of issues that need to be clarified. One of these is the link between traditional IR research and web information-seeking/searching studies, and the value of traditional information-seeking/searching models. In the South African context there are many possibilities for research projects. A major contribution could be made with regard to the impact of cultural diversity and cognitive behaviours. In order to improve web information spaces and web searching in praxis, and to contribute to an information seeking/searching body of theory, we need to work according to a research agenda and a research framework. It is hoped that a framework such as the one proposed in this article will support web information seeking/search studies, the building of models of information seeking/searching and the development of a substantial theory.

## References

- Belkin, N.J. *et al.* 1995. Cases, scripts and information-seeking strategies: on the design of interactive information retrieval systems. *Expert Systems with Applications*, 9:379-395.
- Bilal, D. 2000. Children's use of the Yahoo!igans! Web search engine. Part 1. Cognitive, physical, and affective behaviors on fact-based search tasks. *Journal of the American Society for Information Science*, 51(7):646-665.
- Bilal, D. 2001. Children's use of the Yahoo!igans! Web search engine. Part 2. Cognitive and physical behaviors on complex research tasks. *Journal of the American Society for Information Science*, 52(2):118-136.
- Bilal, D. 2002. Perspectives on children's navigation of the World Wide Web: does the type of search task make a difference. *Online Information Review*, 26(2):108-117.
- Bilal, D. & Kirby, J. 2002. Differences and similarities in information seeking: children and adults as Web users. *Information Processing and Management*, 38:649-670.
- Burwell, L.A. 2001. Too old to surf? No way! An Internet course for seniors. *American Libraries*, 32(10):40.
- Byron, S.M. & Young, J.I. 2000. Information seeking in a virtual learning environment. *Research Strategies*, 17(4):257-267.
- Byström, K. 2002. Information and information sources in tasks of varying complexity. *Journal of the American Society for Information Science and Technology*, 53(7): 581-591.
- Byström, K. & Järvelin, K. 1995. Task complexity affects information seeking and use. *Information Processing & Management*, 31(2):191-213.
- Carey, R.F., McKechnie, L.E.F. & McKenzie, P.J. 2001. Gaining access to everyday life information seeking. *Library and Information Science Research*, 23:319-334.

- Choo, C.W. 2001. Environmental scanning as information seeking and organisational learning. *Information Research*, 7(1). [Online]. Available: <http://InformationR.net/ir/7-1/paper112.html>.
- Choo, C.W. 2002. *Information management for the intelligent organisation*, 3rd ed. Medford, NJ: ASIS&T.
- Choo, C.W., Detlor, B. & Turnbull, D. 1999. Information seeking on the web: an integrated model of browsing and searching. ASIS Annual Meeting. [Online]. Available: <http://choo.fis.utoronto.ca/fis/respub/asis99/>
- Choo, C.W., Detlor, B. & Turnbull, D. 2000a. Information seeking on the Web: an integrated model of browsing and searching. *First Monday*, 5(2). [Online]. Available: [http://firstmonday.org/issues/issuess\\_2/choo/index.html](http://firstmonday.org/issues/issuess_2/choo/index.html).
- Choo, C.W., Detlor, B. & Turnbull, D. 2000b. Working on the Web: an empirical model of web use. Paper presented at the HICSS 33 (Hawaii International Conference on Systems Science), 4-7 January 2000. [Online]. Available: <http://choo.fis.utoronto.ca/fis/respub/HICSS/default.html/>
- Choo, C.W., Detlor, B. & Turnbull, D. 2000c. *Web work: information seeking and knowledge work on the world wide web*. Dordrecht: Kluwer Academic.
- Cooper, L.Z. 2002. A case study of information-seeking behaviour in 7-year-old children in a semistructured situation. *Journal of the American Society for Information Science and Technology*, 53(11):904-922.
- Cooper, M.D. 2001. Usage patterns of a Web-based library catalog. *Journal of the American Society for Information Science and Technology*, 52(2):137-148.
- Cooper, M.D. & Chen, H.M. 2001. Predicting the relevance of a library catalog search. *Journal of the American Society for Information Science and Technology*, 52(10):813-827.
- Crowley, G.H., Leffel, R., Ramirez, D., Hart, J.L. & Armstrong, T.S. 2002. User perceptions of the library's web pages: a focus group study at Texas A&M University. *Journal of Academic Librarianship*, 28(4):205-210.
- Danielson, D.R. 2002. Web navigation and the behavioral effects of constantly visible site maps. *Interacting with Computers*. (In press)
- Dempsey, B.J., Vreeland, R.C., Sumner, R.G. Jr & Yang, K. 2000. Design and empirical evaluation of search software for legal professionals on the WWW. *Information Processing and Management*, 36:253-273.
- Dennis, S., Bruza, P. & McArthur, R. 2002. Web searching: a process-oriented experimental study of three interactive search paradigms. *Journal of the American Society for Information Science and Technology*, 53(2):120-133.
- Dervin, B. 1999. On studying information seeking methodologically: the implications of connecting metatheory to method. *Information Processing and Management*, 34:727-750.
- Ellis, D. 1990. *New horizons in information retrieval*. London: The Library Association.
- Ellis, D. & Haughan, M. 1997. Modeling the information seeking patterns of engineers and research scientists in the industrial environment. *Journal of Documentation*, 53(4):384-403.
- Ellis, D. et al. 2002. Information seeking and mediated searching. Part 5. User-intermediary interaction. *Journal of the American Society for Information Science and Technology*, 53(11):883-893.
- Fidel, R. et al. 2000. Collaborative information retrieval (CIR). *The New Review of Information Behaviour Research*, 1:235-247.
- Ford, N. 1999. The growth of understanding in Information Science: towards a developmental model. *Journal of the American Society for Information Science*, 50(12):1141-1152.
- Ford, N., Miller, D. & Moss, N. 2002. Web search strategies and retrieval effectiveness: an empirical study. *Journal of Documentation*, 58(1):30-48.
- Ford, N. et al. 2002. Information seeking and mediated searching. Part 4: Cognitive styles in information seeking. *Journal of the American Society for Information Science and Technology*, 53(9):728-735.
- Fourie, I. 2002. A review of web information-seeking/searching studies (2000-2002): implications for research in the South African context. In Bothma, T. & Kaniki, A. (Eds), *ProLISSA: Progress in Library and Information Science in Southern Africa. Proceedings of 2<sup>nd</sup> biennial DISSAnet Conference, 24-25 October, 2002*. Pretoria: Infuse:49-76.
- Fourie, I. & Van Niekerk, D. 2001. Follow-up to the use of portfolio assessment for a module in Research Information Skills: an analysis of its value. *Education for information*, 19(2):107-126.
- Goodrum, A. & Spink, A. 2001. Image searching on the Excite web search engines. *Information Processing and Management*, 37:295-311.
- Heo, M. & Hirtle, S.C. 2001. An empirical comparison of visualization tools to assist information retrieval on the web. *Journal of the American Society for Information Science and Technology*, 52(8):666-675.
- Hertzum, M., Andersen, H.H.K., Andersen, V. & Hansen, C.B. 2002. Trust in information sources: seeking information from people, documents, and virtual agents. *Interacting with Computers*. (In press)
- Hölscher, C. & Strube, G. 2000. Web search behavior of Internet experts and newbies. *Computer Networks*, 33:337-346.
- Icheva, S.A. 1998. *Authoring and information seeking on the word-wide web: an experimental study*. M (Information Studies) dissertation. Pietermaritzburg: University of Natal.
- Iivonen, M. & White, M.D. 2001. The choice of initial web search strategies; a comparison between Finnish and American searchers. *Journal of Documentation*, 57(4):465-491.
- Ingwersen, P. 1992. *Information retrieval interaction*. London: Taylor Graham.
- Ingwersen, P. 1999. Cognitive information retrieval. *Annual Review of Information Science and Technology*, 34:3-52.
- Jansen, B.J. 2000. The effect of query complexity on Web searching results. *Information Research: Electronic Networking Applications and Policy*, 6(1). [Online]. Available: <http://informationr.net/ir/6-1/paper87.html>].
- Jansen, B.J., Goodrum, A. & Spink, A. 2000. Searching for multimedia: analysis of audio, video and image queries. *World Wide Web*, 3:249-254.
- Jansen, B.J. & Pooch, U. 2001. A review of web searching studies and a framework for future research. *Journal of the American Society for Information Science and Technology*, 52(3):235-246.
- Jansen, B.J., Spink, A. & Saracevic, T. 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.

- Kim, K-S. & Allen, B. 2002. Cognitive and task influences on Web searching behavior. *Journal of the American Society for Information Science and Technology*, 53(2):109-119.
- Kuhlthau, C.C. 1991. Inside the search process: information seeking from the users' perspective. *Journal of the American Society for Information Science*, 42(5):361-371.
- Kuhlthau, C.C. 1993. A principle of uncertainty for information seeking. *Journal of Documentation*, 49(4):339-355.
- Kuhlthau, C.C. & Tama, S.L. 2001. Information search process of lawyers: a call for "just for me" information services. *Journal of Documentation*, 57(1):25-43.
- Kuhlthau, C.C. & Vakkari, P. 1999. Information seeking in context (ISIC): editorial. *Information Processing and Management*, 35:723-725.
- Large, A., Beheshti, J. & Rahman, T. 2002a. Design criteria for children's web portals: the users speak out. *Journal of the American Society for Information Science and Technology*, 53(2):79-94.
- Large, A., Beheshti, J. & Rahman, T. 2002b. Gender differences in collaborative Web searching behavior: an elementary school study. *Information Processing and Management*, 38:427-443.
- Lazonder, A.W., Biemans, H.J.A. & Wopereis, I.G.J.H. 2000. Differences between novice and experienced users in searching information on the World Wide Web. *Journal of the American Society for Information Science*, 51(6):576-581.
- Lucas, W. & Topi, H. 2002. Form and function: the impact of query term and operator usage on Web search results. *Journal of the American Society for Information Science and Technology*, 53(2):95-108.
- Marchionini, G. 1995. *Information seeking in electronic environments*. Cambridge, UK: Cambridge University Press.
- McGillis, L. & Toms, E.G. 2001. Usability of the academic library web site: implications for design. *College and Research Libraries*, 62(4):355-367.
- Moukdad, H. & Large, A. 2001. Users' perceptions of the Web as revealed by transaction log analysis. *Online Information Review*, 25(6):349-358.
- Nicholas, D., Huntington, P. & Williams, P. 2002. Evaluating metrics for comparing the use of web sites: a case study of two consumer health web sites. *Journal of Information Science*, 28(1):63-75.
- Ozmutlu, S., Spink, A. & Ozmutlu, H.C. 2002. Multimedia web searching trends: 1997-2001. *Information Processing and Management*. (In press)
- Palmer, C.L. 1999. Aligning studies of information seeking and use with domain analysis. *Journal of the American Society for Information Science*, 50(12):1139-1140.
- Palmquist, R.A. 2001. An overview of usability for the study of users' web-based information retrieval behavior. *Journal of Education for Library and Information Science*, 42(2):123-136.
- Palmquist, R.A. & Kim, K.S. 2000. Cognitive style and on-line database search experience as predictors of web search performance. *Journal of the American Society for Information Science*, 51(6):558-566.
- Peters, T. 1993. The history and development of transaction log analysis. *Library Hi Tech*, 42(1):41-66.
- Pettigrew, K.E., Durance, J.C. & Unruh, T.U. 2002. Facilitating community information seeking using the Internet: findings from three public library-community network systems. *Journal of the American Society for Information Science and Technology*, 53(11):894-903.
- Pu, H.T., Chuang, S-L. & Yang, C. 2002. Subject categorization of query terms for exploring Web users' search interests. *Journal of the American Society for Information Science and Technology*, 53(8):617-630.
- Radev, D.R., Libner, K. & Fan, W. 2002. Getting answers to natural language questions. *Journal of the American Society for Information Science and Technology*, 53(5):359-364.
- Ross, N. & Wolfram, D. 2000. End user searching on the Internet: an analysis of term pair topics submitted to the Excite search engine. *Journal of the American Society for Information Science*, 51(10):949-958.
- Roussinov, D.G. & Chen, H. 2001. Information navigation on the web by clustering and summarizing query results. *Information Processing and Management*, 37(6):789-816.
- Rozic-Hristovski, A., Hristovski, D. & Todorovski, L. 2002. Users' information-seeking behavior on a medical library website. *Journal of the Medical Library Association*, 90(2):210-217.
- Savolainen, R. 2002. Network competence and information seeking on the Internet. *Journal of Documentation*, 58(2):211-226.
- Schamber, L. 2000. Time-line interviews and inductive content analysis: their effectiveness for exploring cognitive behaviors. *Journal of the American Society for Information Science*, 51(8):734-744.
- Sheremetov, L. & Arenas, A.G. 2002. EVA: an interactive Web-based collaborative learning environment. *Computers and Education*, 39:161-182.
- Smith, K.A. & Alan, N.G. 2002. Web page clustering using a self-organizing map of user navigation patterns. *Decision Support Systems*. (In press)
- Spink, A. 1997. Study of interactive feedback during mediated information retrieval. *Journal of the American Society for Information Science*, 48(5):382-394.
- Spink, A. & Cole, C. 2001. Introduction to the special issue: everyday life information-seeking research. *Library and Information Science Research*, 23:301-304.
- Spink, A., Ozmutlu, H.C. & Ozmutlu, S. 2002. Multitasking information seeking and searching processing. *Journal of the American Society for Information Science and Technology*, 53(8):639-652.
- Spink, A., Wolfram, D., Jansen, M.B.J. & Saracevic, T. 2001. Searching the web: the public and their queries. *Journal of the American Society for Information Science and Technology*, 52(3):226-234.
- Spink, A. & Xu, J.L. 2000. Selected results from a large study of web searching: the Excite study. *Internet Research: Electronic Networking Applications and Policy*, 6(1). [Online]. Available: <http://informationr.net/ir/6-1/paper90.html>.
- Spink, A. et al. 2002a. From e-sex to e-commerce: web search changes. *IEEE Computer*, 35(3):107-109.
- Spink, A. et al. 2002b. Information-seeking and mediated searching. Part I: Theoretical framework and research design. *Journal of the American Society for Information Science and Technology*, 53(9):695-703.

- Spink, A. et al. 2002c. Information-seeking and mediated searching. Part 3: Successive searching. *Journal of the American Society for Information Science and Technology*, 53(9):716-727.
- Tan, B., Foo, S. & Hui, S.C. 2001. Web information monitoring: an analysis of web page updates. *Online Information Review*, 25(1):6-18.
- Vakkari, P. 1998. Growth of theories on information seeking: an analysis of growth of a theoretical research program on the relation between task complexity and information seeking. *Information Processing and Management*, 34(2/3):361-382.
- Vakkari, P. 1999. Task complexity, problem structure and information actions: integrating studies on information seeking and retrieval. *Information Processing and Management*, 35:819-837.
- Vakkari, P. 2001. A theory of the task-based information retrieval process: a summary and generalization of a longitudinal study. *Journal of Documentation*, 57(1):44-60.
- Wang, P. 1999. Methodologies and methods for user behavioral research, in: *Annual Review of Information Science and Technology*, 34. Williams, M.E. (ed). Medford, NJ: Information Today: 53-99.
- Wang, P., Hawk, W.B. & Tenopir, C. 2000. Users' interaction with World Wide Web resources: an exploratory study: using a holistic approach. *Information Processing and Management*, 36(2):229-251.
- Wathen, C.N. & Burkell, J. 2002. Believe it or not: factors influencing credibility on the Web. *Journal of the American Society for Information Science and Technology*, 53(2):134-144.
- White, M.D. & Iivonen, M. 2001. Questions as a factor in web search strategy. *Information Processing and Management*, 37:721-740.
- Wilson, T.D. 1999a. Exploring models of information behaviour: the "uncertainty" project. *Information Processing and Management*, 35:839-849.
- Wilson, T.D. 1999b. Models in information behaviour research. *Journal of Documentation*, 55(3):249-270.
- Wilson, T.D. et al. 2002. Information seeking and mediated searching. Part 2: Uncertainty and its correlates. *Journal of the American Society for Information Science and Technology*, 53(9):704-715.
- Wolfram, D., Spink, A., Jansen, B.J. & Saracevic, T. 2001. Vox populi: the public searching of the web. *Journal of the American Society for Information Science and Technology*, 52(12):1073-1074.
- Yu, B.-M. & Roh, S.-Z. 2002. The effects of menu design on information-seeking performance and user's attitude on the World Wide Web. *Journal of the American Society for Information Science and Technology*, 53(11):923-933.
- Xie, H. 2000. Shifts of interactive intentions and information-seeking strategies in interactive information retrieval. *Journal of the American Society for Information Science*, 51(9):841-857.
- Xie, H. 2002. Patterns between interactive intention and information-seeking strategies. *Information Processing and Management*, 38:55-77.
- Yang, S.C. 2001. Language learning on the World Wide Web: an investigation of EFL learners' attitudes and perceptions. *Journal of Educational Computing Research*, 24(2):155-181.
- Yang, C.C. & Chung, A. 2002. A personal agent for Chinese financial news on the Web. *Journal of the American Society for Information Science and Technology*, 53(2):186-196.
- Zafeiriou, G., Bpatista, J.M. & Ford, N. 2001. Using student's perceptions of participation in collaborative learning activities in the design of online learning environments. *Education for Information*, 19:83-106.
- Zhang, X. 2002. Collaborative relevance judgment: a group consensus method for evaluating user search performance. *Journal of the American Society for Information Science and Technology*, 53(3):220-231.
- Zhang, X. & Chignell, M. 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.