Subjective Relevance: Implications on Interface Design for Information Retrieval Systems

Shu-Shing Lee, Yin-Leng Theng, Dion Hoe-Lian Goh, and Schubert Shou-Boon Foo

Division of Information Studies,
School of Communication and Information,
Nanyang Technological University,
Singapore 637718
{ps7918592b, tyltheng, ashlgoh, assfoo}@ntu.edu.sg

Abstract. Information retrieval (IR) systems are traditionally developed using the objective relevance approach based on the "best match" principle assuming that users can specify their needs in queries and that the documents retrieved are relevant to them. This paper advocates a subjective relevance (SR) approach to value-add objective relevance and address its limitations by considering relevance in terms of users' needs and contexts. A pilot study was conducted to elicit features on SR from experts and novices. Elicited features were then analyzed using characteristics of SR types and stages in information seeking to inform the design of an IR interface supporting SR. The paper presents initial work towards the design and development of user-centered IR systems that prompt features supporting the four main types of SR.

1 Introduction

Traditional information retrieval (IR) systems are developed using the "best match" principle assuming that users can specify needs in queries [2]. Using this principle, the system retrieves documents "matching closely" to the query and regards these documents as relevant. Relevance is computed using a similarity measure between query terms and terms in documents without considering users' contexts [3, 14].

Hence, this objective relevance is limited somewhat as it does not consider users' needs, in particular, the possible distinction between experts' and novices' needs, and the contexts in which queries are submitted [3]. In other words, experts and novices have varying needs as experts have experience with IR systems and domain knowledge which allow them to search and judge relevance more effectively [4].

Subjective relevance (SR), alternatively, considers relevance from the perspective of users' knowledge and needs [8]. SR is defined as the usefulness of information objects for fulfillment of user's tasks [8]. Hence, one approach of addressing experts' and novices' needs is to possibly enhance objective relevance and tackle its limitations by considering relevance from the perspective of users' knowledge and contexts, an emerging research area in SR [8].

This paper presents initial work towards designing user-centered IR systems by investigating features prompting SR and exploring its implications towards interface

E.A. Fox et al. (Eds.): ICADL 2005, LNCS 3815, pp. 424 – 434, 2005.

design for experts' and novices' needs. The paper uses theories from SR and information seeking to provide rationale for designing an IR interface so that users are guided to find more relevant documents during their information seeking processes.

2 Related Studies

Different approaches have attempted to enhance objective relevance and address its limitations. Works by Chen and Kuo [5] and Gilbert and Zhong [7] have looked at facilitating query formulations by capturing users' personal interpretations of query terms and by accepting queries in natural language respectively. Another research area looks at collaborative browsing where users interact with each other to facilitate their browsing processes and retrieve more relevant documents. An example of this application is *Let's Browse* [9]. A third research area is collaborative filtering. This technique helps users retrieve documents for their needs by recommending documents based on users' past behaviors and behaviors of other users with similar profiles. Examples of such applications are *Fab* [1] and *GroupLens* [12].

The approach described in this paper differs from those described above. Firstly, a user-centered approach is employed by eliciting SR features from experts and novices. Secondly, concepts from SR [6] and information seeking [10] are used to provide theoretical underpinnings for understanding elicited features and informing interface design so that the designed IR interface supports experts' and novices' SR judgments during their information seeking processes.

3 Our Approach

In order to design a user-centered IR system with features supporting SR, we turn to two previous works on SR [6] and information seeking in electronic environments [10]. These works are described briefly in Sections 3.1 and 3.2 to provide theoretical underpinnings for investigating what features experts and novices need to support SR in information seeking. We then design a pilot study to elicit SR features from experts and novices. Methodology and findings of this study are presented in Sections 3.3 and 3.4 respectively. Elicited features are next analyzed using characteristics of SR [6] and stages in information seeking [10] to ensure features elicited and designed support users' SR evaluations and information seeking tasks. These analyses are presented in Sections 4 and 5 respectively. In Section 6, we present how analyses of elicited features in Sections 4 and 5 are used to inform interface design for an IR system supporting SR.

3.1 Subjective Relevance Types

Relevance is a relation between the user and an information object [13]. Since objective relevance using recall and precision measures does not consider users' needs, we examine SR which considers relevance from the perspective of users' changing knowledge and information needs [8]. Four SR types are discussed in [6] and are briefly described in this section. This work is selected because it describes characteristics of the four SR types in a comprehensive manner.

- Topical relevance: This relevance is achieved if the topic covered by the assessed information object is "about" the topic specified in the query.
- Pertinence relevance: This relevance is measured based on a relation between a user's knowledge state and retrieved information objects as interpreted by the user.
- Situational relevance: This relevance is determined based on whether the user can use retrieved information objects to address a particular situation/task.
- Motivational relevance: This relevance is assessed based on whether the user can use retrieved information objects in ways that are accepted by the community.

The discussion on SR seems to indicate two important components: 1) user's information seeking behavior and skills and 2) user's domain knowledge. User's information seeking behavior and skills are important because they provide a means for the user to retrieve information objects and SR evaluations to occur. User's domain knowledge is also important in SR because it may affect how the user evaluates a document. For example, a document may be appropriate for a user's task but due to a lack of domain knowledge, the document may be deemed as irrelevant. Since user's information seeking behavior is a key component for evaluating documents towards task completion, we will next review an established information seeking model and discuss how information seeking may support task completion.

3.2 Information Seeking in Electronic Environments

Here, we examine a well-established model for information seeking in electronic environments [10]. This model is selected because it explicitly describes the information seeking stages users may go through while using an IR system.

Machionini's [10] model describes eight stages in information seeking and its transitions. The eight stages are: 1) recognizing and accepting an information problem; 2) defining and understanding the problem; 3) choosing a search system; 4) formulating a query; 5) executing search; 6) examining results; 7) extracting information; and 8) reflection/iteration/stopping. Transitions between these stages (depicted as 1-8) are presented in Figure 1. The default and high probability transitions are presented as solid arrows and dotted arrows respectively.

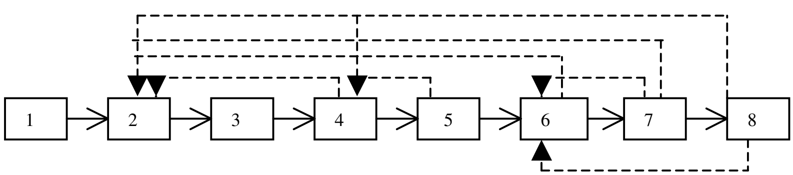

Fig. 1. Transitions and Stages in Information Seeking

To further understand how users may carry out different information seeking stages in an IR system to complete tasks, the eight information seeking stages in Marchionini's [10] model are viewed using Norman's [11] generic model of user interaction in interactive systems. Norman's [11] model describes seven activities that users go through while interacting with a system to complete tasks. Using this model, it may be inferred that Marchionini's [10] information seeking stages can be divided into three phases in terms of user-system interactions proposed by Norman: 1) before execution of an action; 2) during execution of an action; and 3) evaluation of action. Marchionini's Stages 1-4 in information seeking can be mapped to Norman's Phase 1, with Stage 5 in information seeking referring to Phase 2, and Stages 6-8 in information seeking referring to Phase 3.

3.3 Pilot Study

Theoretical frameworks for SR and information seeking, described briefly in Sections 3.1 and 3.2 respectively, seemed to indicate that one possible way to value-add objective relevance and address its limitations could be to design a user-centered IR system supporting users' SR evaluations in information seeking. Hence, a pilot study was carried out to elicit SR features from experts and novices.

Selected Groups: Profiles of Subjects

Eight students (6 Masters and 2 PhD students) from the School of Communication and Information, Nanyang Technological University, were selected as subjects. They were divided into four groups, namely Groups A-D, with 2 subjects in each group. Subjects' level of domain knowledge was determined based on the nature of the task while level of information seeking skills was determined based on whether they had taken a module on "information sources and searching" in their postgraduate studies. Profiles of subjects are shown in Figure 2. In this figure, domain knowledge is depicted as DK and information seeking skills is depicted as IS.

	DK Expert	DK Novice
IS Expert	Group D (2 subjects)	Group B (2 subjects)
IS Novice	Group A (2 subjects)	Group C (2 subjects)

Fig. 2. Profile of Subjects

Methodology

The study was conducted in two sessions, Session 1 for Groups A and B and Session 2 for Groups C and D. This was done because different tasks were used in each session to distinguish subjects' levels of domain knowledge. In each session, subjects were first briefed on the study's objective and the different types of SR. After that, they were given 20 minutes to complete a task. The purpose of this task was to set a context to get subjects thinking about what design features could help them assess a document's relevance. A form was constructed for subjects to note down features that were useful for their tasks, and if not, how these features could be improved.

In Session 1, subjects' task was to gather information for a discussion on "The social impact of the Internet" using two IR systems: Communication Abstracts (a subscription-based database); and ACM Digital Library (a digital library).

In Session 2, subjects' task was to gather information for a discussion on "The different types of information seeking models" using three IR systems: Emerald

Fulltext (a subscription-based database); Library and Information Science Abstracts (a subscription-based database); and ACM Digital library (a digital library). These IR systems were selected as they were commonly used by the subjects.

After completing the task in each session, subjects brainstormed features that they thought were useful for assessing a document's relevance. They were also asked to indicate which features were most important amongst the suggested ones.

3.4 Findings

The study elicited a list of features supporting SR from experts and novices grouped according to their domain knowledge and information seeking skills. A total of 23 and 33 features were elicited from Sessions 1 and 2 respectively. To facilitate analysis, elicited features from both sessions were consolidated by removing duplicates to arrive at a final list of 52 features. To illustrate, we present some features from this list in Table 1. Features elicited from Session 1, Groups A and B were coded with symbols, S1A and S1B, respectively. Features elicited from Session 2, Groups C and D were coded with symbols, S2C and S2D, respectively. Symbols, (++) and (+), were coded next to each feature to indicate whether it was "very important" (++) or "nice to have" (+) to support SR judgments respectively.

Table 1. Example of SR features elicited from the pilot study

Features	Example of SR features elicited from the pilot study
1	Provide recommendations of documents and related topics based on queries users submitted (S2D/++)
2	Rank retrieved documents by relevance (S1B/+)
3	Provide tutorials / search examples (S1A/++; S1B/++)
4	Provide search options, for example, search by title, author, abstract, etc. (S1A/++; S1B/++; S2D/++)
5	Provide abstract of documents retrieved in results list (S1B/++; S2D/++)
6	Provide direct download of documents in PDF format (S1A/++)
7	Provide selected references used in documents (S1A/++; S1B/+)
8	Provide collaborative features (S1B/+)

4 Analyzing Features Using Characteristics of SR

In this section, we describe how we verified if elicited features (see Table 1) supported the four SR types. To achieve this, features elicited from the study were coded to characteristics of the four SR types as described in [6]. Analysis was done based on whether an elicited feature helped users achieve characteristics of a particular SR type. If the feature supported characteristics of a particular SR type, it was coded to that SR type. The coding process was done for all 52 features elicited.

Due to space constraints, we are unable to show coding and rationale for all 52 elicited features. As an illustration, we will describe how elicited features from Table 1 were coded to the four SR types. This coding is presented in Table 2.

- Features in Table 2, Rows 1 and 2, were coded to *topical relevance* as they might provide users with access to other documents and topics that could be similar to topics specified in the query. Moreover, ranking of retrieved documents might also provide an indication of similarity between topics in retrieved documents and topics specified in the query.

- Features in Table 2, Rows 3 and 4, were coded to *pertinence relevance* as they might guide users in query formulations which could be useful for novices.
- Features in Table 2, Rows 5 and 6, were coded to *situational relevance* as document's abstract and full text might provide users with access to document's contents for evaluation towards task completion.
- Features in Table 2, Rows 7 and 8, were coded to *motivational relevance* as references might provide an indication of whether reputable sources had been used to develop document's contents. Moreover, collaborative features might also facilitate discussions to help determine if a document is favored by the community.

5 Analyzing Features Using an Information Seeking Model

In order to understand how elicited features might support SR in users' information seeking processes, we turned to an established model of information seeking by Marchionini [10] to code elicited features to stages in information seeking.

Coding of elicited features to stages in Machionini's [10] model was done by analyzing whether elicited features supported characteristics of a particular stage. If an elicited feature supported characteristics of a stage, it was coded to that particular stage. The coding was done for all 52 features elicited from the pilot study.

To illustrate, we will describe how elicited features from Table 1 were coded to stages in Marchionini's [10] model. Features in Table 2, Rows 3 and 4, were coded to

Table 2. Coding of features elicited from the pilot study

Row	Example of features elicited from the pilot study	SR type	Information seeking stages	Phases in task completion
1	Provide recommendations of documents and related topics based on queries users submitted (S2D/++)	Topical relevance	Stage 6	Evaluating action
2	Rank retrieved documents by relevance (S1B/+)	Topical relevance	Stage 6	Evaluating action
3	Provide tutorials / search examples (S1A/++; S1B/++)	Pertinence relevance	Stage 4	Before executing action
4	Provide search options, for example, search by title, author, abstract, etc. (S1A/++; S1B/++; S2D/++)	Pertinence relevance	Stage 4	Before and during execution
5	Provide abstract of documents retrieved in results list (S1B/++; S2D/++)	Situational relevance	Stage 6	Evaluating action
6	Provide direct download of documents in PDF format (S1A/++)	Situational relevance	Stage 7	Evaluating action
7	Provide selected references used in documents (S1A/++; S1B/+)	Motivational relevance	Stage 6	Evaluating action
8	Provide collaborative features (S1B/+)	Motivational relevance	Stage 6	Evaluating action

Stage 4 (formulate a query) as they might guide users in query formulation. Features in Table 2, Rows 1, 2, 5, 7, and 8, were coded to Stage 6 (examine results) as they provided users with information about documents retrieved to facilitate relevance evaluation. Table 2, Row 6, was coded to Stage 7 (extract information) because it allowed users to access document full text for information extraction.

To further ensure that features helped users complete information seeking tasks, elicited features in Table 2 were viewed from the perspective of interaction phases described in Norman's model of interaction [11] (see Section 3.2). Using this perspective, elicited features in Table 2 seemed to support all three phases, hence, indicating that these features might be useful for task completion.

6 Interactive Interfaces for Supporting SR

We next describe how analyses of elicited features using characteristics of SR types (see Section 4) and information seeking stages (see Section 5) were used to inform the design of an IR system that might support users' SR evaluations during information seeking. As an example, we will use elicited SR features and coding in Table 2.

Coding of SR features in Table 2 was used to inform the design of three interactive interfaces: 1) basic search page supporting "before execution" phase; and 2) results list page with 3) document record page supporting "after execution" phase. The overall design aimed to enhance objective relevance and support users' contexts by explicitly supporting user's SR judgments during information seeking. This was achieved by designing a user interaction flow and tips section in each of the three interactive interfaces (see Figures 2-4) to respectively indicate different stages that users had gone through in the system and to show users how to use features to support their SR evaluations.

Users' level of information seeking skills and domain knowledge were also considered in the overall design. The search interface selection page had remarks next to selection options for basic and advanced search pages to help users choose the right interface for their needs. Moreover, users had to indicate their level of domain knowledge in the search page so that retrieved documents in the results list page were appropriate for their contexts.

The subsequent paragraphs in this section describe in detail how elicited SR features and coding in Table 2 were used to inform the design of three interactive interfaces (basic search page, results list page, and document record page).

Elicited SR features in Table 2 coded to Stage 4 in information seeking included: 1) provide tutorials / search examples and 2) provide search options, for example, search by title, author, abstract, etc. SR features, mentioned in this paragraph, might be designed in a search page to possibly support retrieval of documents for relevance evaluation. These features were gathered from experts and novices in information seeking skills. Thus, it was inferred that SR features elicited from experts (information seeking skills) could be designed in an advanced search page while SR features elicited from novices (information seeking skills) could be designed in a basic search page. In this example, we took SR features as those elicited from novices to design a basic search page for query formulation and execution.

Marchionini's [10] information seeking model suggested that query execution led to the results list page (see Section 3.2). Similarly, executing a query in the designed basic search page led to the results list page. Figure 3 shows the designed basic search page for novices and its features.

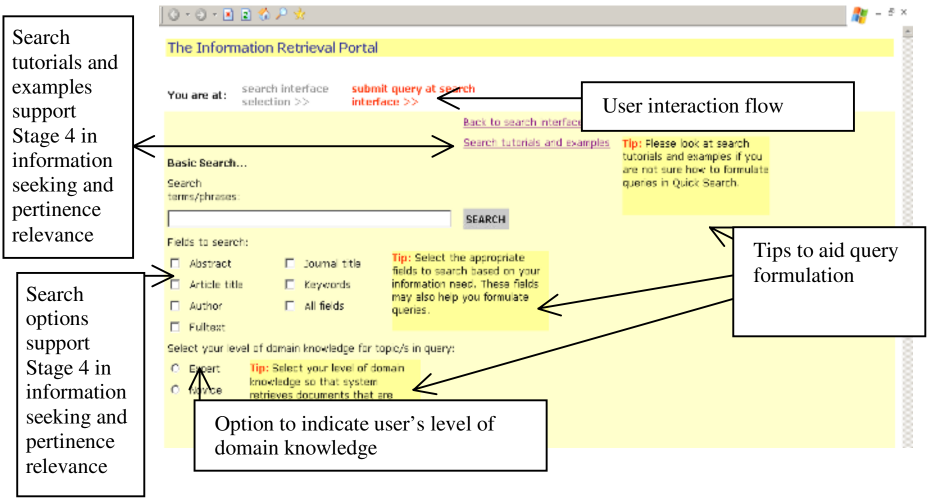

Fig. 3. Basic Search Page for Novices ("Before Execution" Phase)

Elicited SR features in Table 2 coded to Stage 6 in information seeking could be designed in the results list page to possibly support evaluation of documents retrieved. The designed SR features included: 1) provide recommendations of documents and related topics based on queries users submitted; 2) rank retrieved documents by relevance; 3) provide abstract of documents retrieved in results list; 4) provide selected references used in documents; and 5) provide collaborative features. These SR features were elicited from experts and novices in domain knowledge. Thus, it was inferred that SR features elicited from domain experts might be designed in a results list page for domain experts while SR features elicited from domain novices might be designed in a results list page for domain novices. Designed SR features 1, 3, and 4 (mentioned in this paragraph) were elicited from domain experts while designed SR features 2 and 5 (mentioned in this paragraph) were elicited from domain novices. Thus, it was inferred that the designed results list page was for domain experts as most of the designed features were elicited from experts. Although designed SR features 2 and 5 were elicited from domain novices, they were also designed in the results list page. This was because the designed SR feature 2 provided ranking of retrieved documents which was inferred as a common feature in IR systems while the designed SR feature 5 allowed users to discuss issues with other users so that characteristics of motivational relevance could be supported.

Marchionini's [10] information seeking model indicated that Stage 6 in information seeking led to Stages 4 and 7 (see Section 3.2). Hence, it was inferred

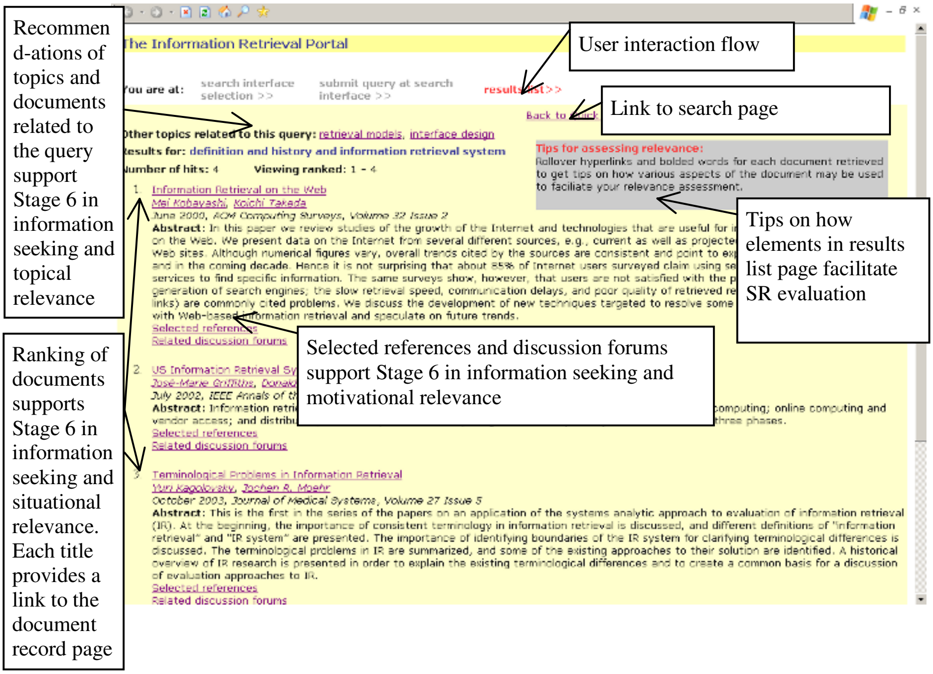

Fig. 4. Results List Page for Experts ("After Execution" Phase)

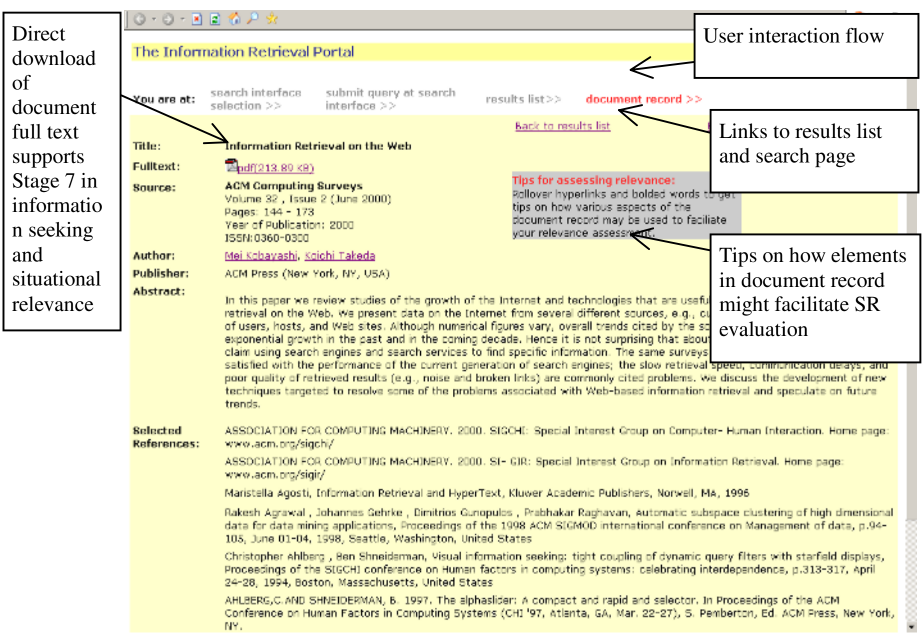

Fig. 5. Document Record Page for Experts and Novices ("After Execution" Phase)

that the designed results list page could have a link back to the basic search page and each document title in the results list page could lead to a document record page providing users with more details about the document. Figure 4 highlights some features implemented in the designed results list page for experts.

The SR feature in Table 2 coded to Stage 7 in information seeking was: provide direct download of documents in PDF format. This feature could be designed in the document record page as it helped experts and novices (information seeking skills and domain knowledge) accessed document full text for information extraction to occur. Detailed document information was also designed in the document record page so users could cite document's source when its contents were used for task completion.

Marchionini's [10] information seeking model indicated transitions to Stages 4 and 6 from Stage 7 (see Section 3.2). Hence, links back to the basic search page and results list page were designed in the document record page. Figure 5 presents the designed document record page for experts and novices.

This is only an illustration of how elicited SR features from Table 2 could inform interface design. Certainly, designing the initial interface for an IR system supporting SR would involve using all SR features elicited from the pilot study.

7 Conclusion and On-Going Work

In this paper, we have described an approach using concepts from SR and information seeking to design a user-centered IR interface supporting SR. A pilot study was conducted to gather features to support experts' and novices' SR judgments. Elicited features were then analyzed using characteristics of SR types and stages in information seeking to inform the design of an IR interface supporting SR.

The designed IR interface presented here seemed to include existing features in current IR systems. Reasons for this included: 1) the small number of participating subjects and 2) subjects brainstormed SR features after they had completed a task, hence, causing them to suggest features that they had found useful whilst using example IR systems. This is an initial study and certainly more subjects should be recruited in future studies to brainstorm novel SR features.

In contrast to other studies addressing limitations in objective relevance [e.g. 1, 5], our approach described in this paper is unique in several aspects. Firstly, a user-centered approach incorporating theoretical frameworks from SR [6] and information seeking [10] were used to elicit and validate features supporting experts' and novices' SR judgments in information seeking. Secondly, the designed IR system explicitly addressed experts' and novices' needs by guiding users select the appropriate search page for their information seeking skills. Moreover, users' level of domain knowledge were also considered as they had to specify their level of domain knowledge in the search page so that the results list page included appropriate details to support their SR judgments. Thirdly, the designed IR system explicitly embraced users' SR judgments in information seeking by providing a user interaction flow to indicate pages that users' had gone through in the system. In addition, tips were designed to explicitly indicate how designed features could be used to support SR and help users find relevant documents for their tasks.

Findings presented here are preliminary and part of on-going research to elicit features and explore a framework using concepts from SR and information seeking to inform the design of an IR system that supports users in finding relevant documents for their needs. Elicited SR features and the approach used to inform interface design need to be further refined and tested before they can emerge as principles for designing user-centered IR systems. Future work could focus on validating elicited SR features through a quantitative study to ensure that the designed IR interface meets experts' and novices' needs.

References

1. Balabanovic, M. and Shoham, Y. (1997). Fab: Content-based, collaborative recommendation. *Communications of the ACM, 40*(3), 66-72.
2. Belkin, N. J., Oddy, R. N., & Brooks, H. (1982). ASK for information retrieval: Part I. background and theory. *The Journal of Documentation, 38*(2), 61-71.
3. Case, D. O. (2002). Looking for information: A survey of research on information seeking, needs, and behaviour. California, USA: Elsevier Science.
4. Chen, H., Houston, A. L., Sewell, R. R., & Schatz, B. R. (1998). Internet browsing and searching: Use evaluations of category map and concept space techniques. *Journal of the American Society for Information Science, 49*(7), 582-603.
5. Chen, P.-M. and Kuo, F.-C. (2000). An information retrieval system based on a user profile. *The Journal of Systems and Software, 54*, 3-8.
6. Cosijin, E., and Ingwersen, P. (2000). Dimensions of relevance. *Information Processing and Management, 63,* 533-550.
7. Gibert, J. E. and Zhong, Y. (2003). Speech user interfaces for information retrieval. *Proceedings of the 12th International Conference on Information and Knowledge Management,* New Orleans, USA, 77-82.
8. Ingwersen, P. and Borlund, P. (1996). Information transfer viewed as interactive cognitive processes. In Ingwersen, P. and Pors, N. O. (Eds.). *Information Science: Integration in Perspective* (pp. 219-232). Copenhagen, Denmark: Royal School of Librarianship.
9. Lieberman, H. (1995). An agent for web browsing. *Proceedings of the International Joint Conference on Artifical intelligence*, Montreal, 924-929.
10. Marchionini, G. (1995). *Information seeking in electronic environments* .Cambridge, U.K.: Cambridge University Press.
11. Norman, D. A. (1988). *The psychology of everyday things*. New York: Basic Books.
12. Resnick, P., Iacovou, N., Mitesh, S., Bergstron, P., and Riedl, J. (1994). GroupLens: an open architecture for collaborative filtering of Netnews. *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, Chapel Hill, USA, 175-186.
13. Saracevic, T. (1996). Relevance reconsidered '96. In Ingwersen, P. and Pors, N. O. (Eds.). *Information Science: Integration in Perspective* (pp. 201-218). Copenhagen, Denmark: Royal School of Librarianship.
14. Tang, R. and Soloman, P. (1998). Toward an understanding of the dynamics of relevance judgment: An analysis of one person's search behavior. *Information Process and Management, 34,* 237-256.