Exploring the unknown: a study of expert use of a digital library

Suzette Keith[1], Ann Blandford[2], Richard Butterworth[1], Bob Fields[1] and Yin Leng Theng[3]

[1] Interaction Design Centre, School of Computing Science, Middlesex University, Trent Park, Bramley Road, London, N14 4YZ, U.K.
{S.Keith, R.J.Butterworth, B.Fields}@mdx.ac.uk
[2] UCL Interaction Centre, University College London, 26 Bedford Way, London, WC1H 0AP, U.K.
A.Blandford@ucl.ac.uk
[3] Nanyang Technological University, 31 Nanyang Link, Singapore 637718.
TYLTheng@ntu.edu.sg

Abstract. This study investigates how expert intermediaries interact with a corporate digital library when searching specialised topics. We identify query reformulation as being key to successful search and therefore present a more detailed account of this stage of an information seeking process. We propose that query reformulation forms a significant, distinct step from initial query formulation or subsequent browsing and monitoring behaviour. We qualitatively analysed the interactions of expert intermediaries using existing models of information seeking; while those existing models are adequate for the observed interactions, query reformulation is downplayed within them. We define query reformulation as an iterative process by which the intermediary develops a fairly precise set of query results. The process described is reactive, not predictive, and makes subtle use of term frequency, and traditional information retrieval skills, together with prior knowledge of the system to make changes and examine results.

1 Introduction

The designers of any new system have expectations - explicit or implicit - about how that system will be used and who it will be used by. Digital libraries are no exception. A number of researchers have proposed models of the information seeking behaviour of users [1,3,9,11,12,13,15]. However, relatively few studies have investigated in detail how users actually interact with digital libraries, their goals, strategies, actions and reactions.

Borgman [6] comments that 'information systems continue to be difficult to learn and to use'. A study of novice users by Blandford *et al* [4] found that their information retrieval skills with existing digital libraries are weak: users lack strategies, and the feedback provided by libraries is currently inadequate for helping users develop their skills.

Exploring a digital library resource for complex or novel search is not an easy task. In Ellis's investigations he found that users relied on computerised searches 'to approach a new topic within an unknown field', but reported that both the scientists and engineers relied on expert intermediaries for this early stage of what he calls 'surveying' [9]. The users would continue to browse or monitor the results of these searches. This raises an important question as to how those expert intermediaries use a digital library to seek information:

- What strategies do library experts use?
- What features and tacit knowledge do they utilise?
- What claims can be made about how current designs aid researching a new topic?
- What are the implications for other classes of users?

In the study reported here, two people, who unusually were both developers and expert users of a corporate digital library, performed various searches on behalf of research engineers with a genuine information need. The interaction and commentary was recorded on video, transcribed and analysed qualitatively and in depth. The interactions of these experts at information seeking and of the system were compared to some established models of information seeking, as discussed below. The library they used is one developed at British Telecommunications plc, a large UK company with a substantial research centre.

1.1 The library and its users

The BT digital library project was initiated with the intention of supporting a paperless office environment in which researchers and engineers could search a number of resources and databases through an intranet facility. The system was developed and launched in 1995 and has since been subject to a number of evolutionary developments affecting both the interface and the content.

The library is staffed by qualified librarians who maintain, develop and use the digital library, and therefore have considerable expertise in the company business, information retrieval skills, and knowledge of the system and the resources. In contrast, the many casual and infrequent users experienced difficulties, resulting in ineffective searches.

Interviews with a representative range of users highlighted their complex, highly specific, and often novel information needs - typically supporting the start of new projects [9,10]. The study reported here identifies the key skills of experts in performing such searches, as a starting point for assessing how to help a broader spectrum of users develop those skills.

1.2 Features of the library

The digital library offers keyword and phrase searching, full-text access to selected journals and a number of browsable links. An 'information spaces' feature provides a monitoring service for new documents on specific themes or categories. The keyword search feature provides a simultaneous search and summary of a number of resources and databases using abstract, index and catalogue data. The two most comprehensive

and respected resources are INSPEC, a scientific and technical resource, and ABI/Inform which is more business orientated. Each resource can subsequently be searched in more detail, linking to abstracts and full text. The descriptor terms provided by the abstract and index services can be searched in a separate field, selected from the 'find similar' from individual abstracts or through a keyword browser feature which analyses the incidence of terms.

1.3 Existing information seeking models

This study of expert interactive information seeking can be related to extant models of the process. Ellis's model, developed using a grounded theory approach, identifies phases in the search process used by engineers and research scientists. The phases identified are surveying, chaining, monitoring, browsing, distinguishing, filtering, extracting and ending [9]. Starting a search is not easy; Ellis reported his users making use of expert intermediaries to deliver a literature search at the survey stage from which to chain or monitor.

Kuhlthau [12] presents a process model with a six-stage cycle of initiation, selection, exploration, formulation, collection and presentation covering the breadth of the information task resolving initial uncertainties through to applying the retrieved information. Belkin's [3] ASK model draws attention to the fact that the searcher often starts from a point of not knowing and uncertainty of the unknown. Marchionini [13] offers an eight stage process: recognise and accept an information problem; define and understand the problem; choose a search system; formulate a query; execute search; examine results; extract information; and reflect, iterate or stop. Recognising that this process is 'dynamic and action orientated', these stages fit into three phases of: understanding; planning and execution; and evaluation and use. These phases are echoed in Sutcliffe and Ennis' [13] process model which looks more specifically at the user-system interaction process and has four main activities: problem identification, need articulation, query formulation and results evaluation.

Bates and others [5, 14] draw attention to the progress of the search, and changes in the search goal as the researcher gathers more information. Like Ingwersen [11] she challenges the stable model of searching often used in information retrieval research where the measures of recall and relevance assume an unchanging search goal. Robertson's traditional model of information retrieval (reported in [1]) suggests that the search activity is broadly concerned with matching the information need and the query to the controlled vocabulary terms applied by the information provider. Index terms are valuable organisers of digital library collections bringing related 'about' information together, but they can cause difficulties for users who are not experienced in the classification system [7]. Bates [2] discusses Bradford's Law in relation to finding keywords where there is 'a small number of terms with a very high frequency and a large number of terms with a very low frequency'. This describes the difficulty reported by inexperienced users of either getting too many or too few results

Ingwersen [11] develops a cognitive theory that takes account of the information space, including text representations, and the system settings, including search language and indexing. He identifies important features of the users' real life information need arising from whether the user's concept of the problem is stable or

variable, and whether or not it is a well defined. Difficulties arise where the concept is stable but ill-defined, causing the user to become trapped in dead-ends.

Sutcliffe and Ennis [13] makes use of a taxonomy of information need, and four knowledge sources of the users – the domain, the resources, the device, and information retrieval skill – to predict the user's behaviour: for example, how a lack of skills and prior knowledge can result in premature abandonment.

In the study reported here, a process-based framework, based on a synthesis of extant models, is used to support a detailed analyse of the activities of the librarians who act as expert intermediaries. As will emerge, although their behaviour is largely consistent with these models, they exhibit a much more sophisticated query reformulation strategy, making considerable use of the controlled vocabulary and syntax changes to overcome the inherent difficulties of a searching an unknown topic. The experts only allow a modest level of flexibility in relation to the topics as they redefine the request into the language of the resource.

2 Method

As noted above, the aim of the study was to better understand how expert intermediaries find information in a digital library. Two qualified librarians who were frequent users and developers of the service acted as expert intermediaries.

Three research engineers submitted current research problems. They were interviewed by the investigators and agreed a written statement of each query. The three searches on "fault diagnosis and ADSL (Asynchronous Digital Subscriber Line)", "credit card fraud" and "affective design" were undertaken by both experts.

In this knowledge elicitation study, think-aloud protocols were used to capture the librarians' thoughts, observations and plans. The experts were observed by one of the investigative team and prompted only to elicit further knowledge. It was agreed that they would pursue each query for up to one hour with the intention of developing a search string that could be used by the originator for monitoring the topic specified. Although they had no opportunity to conduct a reference interview, it was agreed that they could make contact if they needed further information in order to proceed.

The interaction and think-aloud commentaries were recorded on video and transcribed for later analysis. The transcribed data was analysed primarily using qualitative methods to extract the dominant themes and strategies within the structure of the user-system interaction. The query originators were invited to comment on the usefulness of the results returned to them.

3 Summary of Activities

The expert intermediaries took between 47 and 63 minutes for each search, entering between 12 and 21 query strings and evaluating the results by browsing the summaries and abstracts. The two resources used most frequently were INSPEC and ABI: four out of the six searches were tried in both. One librarian also accessed another database (Intellact).

Having selected a resource, most of the activity revolved around the page that summarised the first 10 results within that resource. The experts rarely visited a second or third page of results; instead they browsed through the results and selected abstracts. They found descriptors shown with the abstracts and made use of the keyword browser feature to select descriptor terms, adding them to the query string. Only once did a librarian call for and read a full text document, although on another 2 occasions full text was called for but then aborted.

This quantitative data gives a high level view of the intense activity in which these experts engaged; the more detailed qualitative analysis below examines these activities in relation to a framework for describing the process of information seeking.

3.1 Example interaction: Librarian B searching on 'fault diagnosis'

For illustrative purposes, we present in detail the interaction summary for Librarian B working on the fault diagnosis problem, which illustrates the strategy of exploring the terms, expanding the search and refining the results.

The first search was defined by an experienced senior telecommunications engineer involved in developing tools and applications to improve the ADSL service. He was looking for a more effective method for diagnosing faults. This problem was relatively new, but he had made some searches already.

As shown in Table 1, Librarian B began by dividing the researcher's key phrase into three separate terms: 'fault', 'diagnosis' and 'ADSL' which she reformulated several times before repeating the search in a different resource.

Table 1. Example of query reformulation

Query term	Number of hits
1. fault, diagnosis, ADSL	>7000
2. Diagnos*, fault, ADSL	>20000
3. Diagnos*, fault, +ADSL	644
4. +fault, +ADSL	6
5. +fault, +de=subscriber	76
6. +de=fault, +de=subscriber	32

After launching the search to all resources, she chose INSPEC, which contained the largest results set. However, her choice was not just on size; she commented: "I know intuitively that he would be happier with the technology area".

Examining the results list and the abstracts, she found that there were no matches for one of the three terms and concluded: "So I don't think this term is very good".

Keeping the same terms, she used truncation, a change of order and the '+' syntax for one term to create a set in which all three terms were included. After reading through the summaries of the first 10 results carefully she commented: "They don't look so awful".

She viewed the first ten abstracts, commenting on the appropriateness of the generic descriptor terms being used, then linked to the keyword browser to view the

analysis of the descriptor terms. Her assessment of this was: “So I am fairly happy about getting towards ADSL words and now I am looking for the diagnosis words”.

Including ‘+’ on both terms reduced the set to only 6. This caused her to change her strategy to use descriptors noticed earlier in the search: “I know there are those two terms for ADSL. I’m broadening out ADSL a bit now into subscriber lines”.

Query 5 used the generic descriptor ‘+de=subscriber’ and increased the set. She read through the first page of results before linking to the keyword browser.

Changing the term ‘fault’ into a descriptor using ‘+de=‘ resulted in a tight search of 32 items. She worked through these results for over 5 minutes, looking at all 4 pages of results and selecting to view half the abstracts.

She noted that the generic term was causing the inclusion of some aspects of the technology that were irrelevant, commenting that “it is very difficult to get a tight search that is exactly what you want”. An attempt to exclude the unwanted concepts using the exclude syntax ‘-’ resulted in no matches.

She repeated the search in another resource, ABI. Getting no matches here provided the stimulus to split the concepts and hunt for descriptors for ADSL separately, explaining “we’ll do it in sections”. She saved this search result by opening a new window to hunt for terms for fault diagnosis. She added multiple ‘OR’ terms found in the abstracts and the keyword browser to build a set of 95,000 results. She cut the set by copying and pasting back the generic descriptor for ADSL saved in the previous window and, after reviewing the results, deleted all the fault-related terms except ‘de=testing’. She reviewed and saved this query string which the originator later commented were clearly in the right area, and supported his requirement to monitor activity.

4 Analysis of Results

Both expert intermediaries were highly motivated to complete the search tasks and were persistent in trying to resolve the complex problems set by the originators of the searches. The think-aloud protocols showed them to be reactive to the results of the preceding query.

One overriding finding is how much time and effort they devoted to reformulating the queries. After an initial exploratory phase, the librarians engaged in a cycle of evaluation and reformulation to expand and refine the results. Precision was an important requirement of the search, but the novel elements demanded a comprehensive search of the major resources. During the reformulations the librarians utilised a number of different strategies, but remained generally close to the original specified query. The searches remained within the phase that Ellis identifies as surveying [9], while the information need according to Ingwersen’s model was stable [11]. Based on the process models discussed earlier, we have analysed the search processes observed according to the phases: problem definition; source selection; query formulation; results examination; query reformulation; and results presentation.

4.1 Problem definition

Much of this early part of the process of problem definition belonged to the people who had originated the problems, providing context, concepts and suitable keywords, using their domain knowledge. The librarians' searches remained very close to the problems specified by the users, with little of the goal shifting seen when users search for themselves [1, 3, 5, 14]. However they avoided the 'dead-end' trap of a stable but ill-defined information need indicated by Ingwersen, by permitting some variability in order to improve the definition of the problem [11] – for example, in exploring the different contexts in which credit card fraud occurs rather than just the one specified by the originator, and through examining the effect of multiple keywords.

4.2 Source selection

An important feature of this digital library is that it can simultaneously search a number of resources. This repositions the phase of resource selection after, instead of before, submitting the first query. Sutcliffe and Ennis [13] suggest that knowledge of the resource is an important factor contributing to a successful search. Both experts were very familiar with the resources available and showed a clear preference for the coverage and power of ABI and INSPEC, only referring once to the news service, Intellact, which has limited search refinement capabilities. Librarian B developed all searches within INSPEC, because she preferred the 'harder vocabulary' of the technical database, even though during two of the searches she commented that the content of the other resource would be more relevant.

4.3 Query Formulation

The experts were familiar with the terminology both of the business and as used in the resources. They chose their initial query terms after a careful examination of the information given by the originator and entered up to three words individually or as phrases. In two searches, the librarians picked out terms, such as 'MOTO' and 'cnp', because they were unfamiliar, and to find out "if they were any use". The pattern of interaction was to start with a simple exploratory search and subsequently adapt on the basis of the results.

4.4 Results Examination

The experts' strategies were highly reactive to the feedback offered. The results page showed the number of results found, and gave the title, brief summary and the incidence of the query terms for each result. The results were ranked according to the number of terms matched, with descriptor field terms ranked higher than others.

The overall pattern was to scroll quickly through the first page of results, scanning the headings and brief details. If at least some of the results seemed broadly relevant, they read through them more carefully, following up some abstracts, and opening the

keyword browser. They would comment on the relevance of the material or inclusion of non-relevant concepts, the terminology used in the text and descriptors. If the results were poor, they chose different terms from the originators description.

The experts referred to the number of hits and term frequency information shown with the results, particularly to analyse unexpected results – for example, where only some terms were matched, or if automatic stemming included variations of the terms required; e.g. 'affective design' produced results for the phrase 'affects design'.

Overall, particularly in the early stages of the search, they seemed to be more concerned with finding 'good' terms than 'good' results. In the later stages they evaluated the effects of changes made to terms, syntax or field to reduce the quantity of results, improving precision to match the needs of the originator.

4.5 Query reformulation

The librarians reformulated the query terms many times for each search, changing the terms, using descriptors and making syntax changes. These are analysed in more detail in relation to the activities observed.

Changing terms

Starting with the terms given by the originator, the experts added and changed terms after viewing the results. Terms were found in the text of the relevant summaries and abstracts, and in the keyword browser. In one example, Librarian B found 'testing' and a number of broad terms including 'expert systems' to describe testing type activities instead of the more specific 'fault diagnosis' given by the originator.

Similarly, in the search for affective design, Librarian A found the term 'artefact' within the full text, and then used it with the terms 'product' and 'design': " I'm wanting to see what will turn up in general, because I'm looking for a bit of help with the terminology..."

Reference to an expert provided a further source of terms; in one search, the librarian stopped to discuss terms with the search originator.

This strategy was more reminiscent of 'orienteering' [14]. Each of these examples allowed some movement to take place, expanding on the original concept and exploring the effectiveness of the terms.

Descriptors

The strongest support for the librarians' strategies came from the descriptor terms provided by the abstract and index services. The search interface of this digital library did not offer a thesaurus search, but terms could be found in abstracts and the keyword browser. Occasionally they relied on memory or guesswork, revealing that they had a highly developed knowledge of the terminology of the resources.

The experts' effectiveness was dependent on the quality and precision of the controlled vocabulary developed by the index provider. Both tolerantly expressed concerns about the relationship between the use of a generic term such as 'digital subscriber' and the specifics of ADSL. As librarian B said: "I think they're using digital subscriber which is not technically correct". However, she accepted the term

from the keyword browser, commenting, “that looks like it is going to be this one here”. The other expert was discouraged from using the generic term following a discussion with the originator, who offered more specific alternatives, but she remained dissatisfied with the results.

The search for affective design proved much harder because no suitable descriptor terms could be found. Both commented about the ‘newness’ and ‘soft terminology’ being used by the authors, and discussed the inconsistencies that occur before the terminology stabilises sufficiently to become an accepted descriptor term. The ‘find similar’ feature associated with a relevant document showed only one descriptor that was evaluated as not in the relevant domain. In another search, the ‘find similar’ descriptors led to a successful series of searches from an abstract marked by the search originator as the result of an on-line document request feature.

The expert intermediaries reported looking for repetitions of descriptors in a set of abstracts that were only partially relevant; this strategy was further supported by the keyword browser’s analysis of descriptors. Some descriptors were applied to very large sets of data, but this was only apparent from the results of selecting a term; others could prove to be sub-sets that made no difference to the overall size. Identifying more specialised terms which related only to small sets, and combining terms using the must have ‘+de=‘ syntax within the descriptor field, were used to reduce the results set and improve precision.

In the fault diagnosis search described earlier in detail, 12 of the 20 queries included descriptors. The best solution used two terms that were in themselves generic but when combined, ‘+de=subscriber, +de=testing’, gave rise to a tightly specified result set.

Syntax

The expert intermediaries were very skilled in using the search syntax to their advantage. An important effect was the result of deciding whether to treat the search request as a phrase (exact matching) or to search individual terms. Early in the search process, the experts used the default ‘OR’ to test the response of the resource to combinations of words and phrases.

Adding more terms could improve the ranking of the results, displaying matches to all terms above partial matching. More effective in identifying overlaps and conjunctions and eliminating irrelevant results were the use of ‘+’, the ‘must have’ syntax, and the field marker ‘de=‘ to direct the search to the descriptors.

Getting rid of non-relevant items was less easy and, as demonstrated using the ‘-’ syntax, tended to remove too much.

4.6 Results presentation

The best query strings or URLs were saved to send to the originator for further evaluation or to set up the monitoring feature. Neither expert regarded the search tasks as complete, and expressed a need to have some response from the researcher in order to refine the search further.

5 Towards a revised model of search for expert intermediaries

The previous section discusses the activities of the experts using a process-oriented framework based on earlier models; more detailed qualitative analysis of the interaction and commentary revealed a more sophisticated cyclic process that was needed in order to develop the search. The expert intermediaries were presented with a 'difficult' search problem presented by domain experts working on novel or innovative topics. Having passed the problem to the expert intermediaries, the information need became, in Ingwersen's terms, a conceptually stable need – the expert intermediaries lacked the authority to make any significant change to the topic. The experts treated the search as ill-defined, but set about improving this, interactively testing terms by searching and thus overcoming a situation that had led the less experienced users to a dead-end [11].

The experts used all available information for assessing relevance, including content, quantity, and change from previous results.

The emerging pattern of the development of the query was:

- Opening cycle – exploration of the response of the resource to the query:
    - Explore terms
    - Identify concepts
    - Identify descriptors
- Intermediate cycle – expanding the search:
    - Add multiple or broader terms, and descriptor terms
    - Change resource
    - Evaluate completeness of search to the required concepts
- Final cycle – refine and accept results:
    - Add specific terms, make syntactic changes and combine descriptor terms
    - Evaluate relevance and precision within the resource

The opening cycle tested whether or not the problem was well defined and explored the terms and possible descriptors. The terminology offered by the originators was validated, and descriptor terms identified from the results.

In the intermediate cycle, the results were evaluated on many levels, browsing document summaries and abstracts for terminology and content relevance. The experts applied traditional information retrieval strategies to explore, build and test the search. The descriptor terms found with the abstracts or through the keyword browser were tested for effect. Most often the subsequent reformulations added to the result set, maximising recall of relevant items. The results were examined for relevance and poor results rejected in favour of the more successful alternatives. If necessary to overcome small or irrelevant results, the information need was allowed to become less stable in order to divide the concepts and consider alternative perspectives.

In the final cycle the function of the keyword browser and abstracts changed to become a source of specific or narrower terms that could be applied to create a small and well-defined collection of results. Terms were combined to find not only hierarchically more specific results but also more innovative intersections between two or more well-defined concepts. When experts stopped, it was because they were

satisfied or needed further instruction from the originator: the information need was at least temporarily returned to a stable and well-defined state.

Changing the resource had the effect of testing the strength of the concepts from a different perspective and added to the exploration of the information space. Four of the searches were repeated in the second abstract and index service, gaining a different perspective from the scientific and technical resource to the more application orientated business perspective.

This cycle of activity very delicately explored and renegotiated the space between the users' conceptual needs and the information space defined by the resource.

6 Conclusions

The expert intermediaries' goal was to deliver useful search strings and results that could be used for monitoring. These expert intermediaries were committed to, and confident in their progress towards this goal, as confirmed by two of the search originators. The results of the third search suggested that the resources being used were not appropriate to the domain. In each case confidence was derived from a systematic and iterative process, exploring terms and concepts, and careful examination of the results. These information retrieval skills were supported by an understanding of the system they had designed and experience with the resources.

Extant models of information seeking do not provide adequate accounts of the level of reformulation observed and the role played by sophisticated information retrieval strategies. It is clearly central to effective and efficient information retrieval, and a key component of librarians' information seeking skill. The overall strategy appears to be domain independent and transferable between resources.

In accordance with Ingwersens' model, the experts remained close to the original search request, but were able to explore beyond the boundaries set by the originators using prior knowledge of the business domain and knowledge gained during the search. The searches appeared originally well defined by the originators who were domain experts. However from their prior experiences of the structure and terminology of the resources and the results of the initial exploration of the terms, the intermediaries were able to improve the results by redefining the terms and applying descriptor terms. This improvement was possible because of the quality of the resource indexing and supported by a descriptor term analysis feature that they had developed. The experts in this current study were able to progress difficult searches described by Ellis as surveys, typically carried out at the start of new projects. The experts' experience of the resources and the system enabled them to make modifications where less experienced users would have abandoned the search.

There was little interaction that could be categorized as chaining, berry-picking or orienteering [1,5,14]. This is possibly an important distinguishing feature of the behavior of expert intermediaries from end-users but is also a necessary starting point. The sub-goals of exploring, expanding and refining the terms used are not in themselves difficult to understand, and the feedback provided supported the all-important identification of appropriate terminology - especially the descriptor terms provided in this and many other abstract services.

This work has focused on the detailed qualitative analysis of the interaction and behaviour of two users of one particular digital library. Nevertheless, it provides a starting point for further study of the skills needed for effective and efficient use of digital libraries and for novel design paradigms that will enable less experienced users to achieve greater success when working within specialised digital libraries.

Acknowledgements

This work is funded by EPSRC Grant GR/N37858. We are grateful to the British Telecommunications plc staff (librarians and users) who took part in this study, and to George Buchanan for comments on an earlier version of this paper.

References

1. Bates, M J (1989) The design of browsing and berrypicking techniques for the on-line interface. On-line Review 13 (5) 407-424
2. Bates, M J (1998)Indexing and access for digital libraries and the internet: Human, database, and domain factors. J Am. Society for Information Science 49 (13) 1185-1205
3. Belkin, N J (1980) Anomalous states of knowledge as a basis for information retrieval. Canadian Journal of Information Science. 5. 133-134
4. Blandford, A., Stelmaszewska, H. & Bryan-Kinns, N. (2001a) Use of multiple digital libraries: a case study. In Proc. JCDL 2001. 179-188. ACM Press.
5. Blandford, A. & Stelmaszewska, H. (2001b) Shooting the information rapids. In Vanderdonckt, Blandford & Derycke (Eds.) IHM-HCI2001 Vol. II (short paper). 51-54.
6. Borgman, C. (2000) From Gutenberg to the global information infrastructure. MIT Press.
7. Chen, H.& Dhar, V. (1990) Online query refinement on information retrieval systems: a process model of searcher/system interactions. Proc 13th Int conf R&D in information retrieval 115-133 ACM Press
8. Doubleday, A., Ryan, M., Springett, M. & Sutcliffe, A. (1997) A comparison of usability techniques for evaluating design. In Proc. Designing interactive systems : processes, practices, methods, and techniques 1997. ACM press. 101 –110
9. Ellis, D. & Haugan, M (1997) Modelling the information seeking patterns of engineers and research scientists in an industrial environment. J Documentation 53 (4) 384-403
10. Hirsh, S. (1999) Information seeking at different stages of the R&D process. Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval 1999. ACM Press. 285-286
11. Ingwersen, P. (1996) Cognitive perspectives of information retrieval interaction: elements of a cognitive IR theory. J Documentation 52 (1) 3-50
12. Kuhlthau, C.(1988) Longitudinal case studies of the information search process of users in libraries. Library and information science research 10 (3) 257-304
13. Marchionini, G (1995) Information seeking in electronic environments. Cambridge University Press
14. O'Day, V. L., & Jeffries, R. (1993). Orienteering in an Information Landscape: How Information Seekers Get From Here to There. In Proc. InterCHI '93, pp. 438-445.
15. Sutcliffe, A. & Ennis, M. (1998) Towards a cognitive theory of information retrieval. Interacting with computers 10. 321-351