

Optimal results presentation for dynamic search (a position paper)

Frances C. Johnson
The Information Research Institute
Manchester Metropolitan University
f.Johnson@mmu.ac.uk

Adaptive information retrieval may use feedback to capture the context of the user's task and aspects of relevance that are hard to express in a query. However many users may find it difficult to express even the topic of interest in an initial query. The cognitive view of search is of a dynamic, problem solving activity with the user modifying the information interest as new information is retrieved. To an extent, current interactive retrieval systems support (with direct access) the query modification expected when users are engaged in understanding and developing an information interest. This paper focuses on the role and requirements of results presentation, especially summaries, in this model of interactive search and retrieval. Results presentation is an important part of a retrieval system enabling, in response, the user's query modification and system query calibration for a closer match to the relevant documents. The relevance to this workshop on adaptive retrieval lies with the interest in explicit feedback facilitating the users' evolving query. The question posed is whether there is an optimal presentation of the retrieved documents to support the process of learning and query clarification during search.

For any given task it is likely that different users will hold a different view of the information required and adopt different strategies for obtaining it. This may be partly explained with reference to Kuhlthau's [1] model of information searching as a task process with various stages at which the understanding of the task changes. Each stage is characterised by a subtask and it is the associated thoughts and feelings that can influence the actions taken to advance the process. Those with greater knowledge of the information task may be at a very clear and focused stage and able to identify keywords and formulate an effective query. Those with less knowledge may be identified as being at the earlier vague and confusing stage and are more likely to browse to learn about the topic.

The key to successful search may lie in the system's ability to keep the user focused on understanding the information interest and to progress, even flow, through the stages. With this view the role of results presentation goes beyond simply indicating content of the retrieved items and hopefully why they were retrieved in response to the query. The user is further looking (from the retrieved items) for ways to conceptualise the query and use in manipulating the search. This is a very important function of the system and further research is needed to explore the optimal presentation, that is the type of information to represent content and its visualisation through understanding the interaction between results presentation and the user's search process.

Few studies have evaluated the effectiveness of different search results presentations. Dumais et al [2] found that a combination of 'clues' improved performance and suggested that the category names help users focus in on areas of interest and the page titles help to disambiguate the category names. White et al [3] found query biased summaries were more effective than general summaries in assisting users gauge document relevance. Tombros & Sanderson [4] had similar findings and attributed this to fact that they indicated the context within which potentially ambiguous query terms were used.

The context in which the query terms appear clearly helps the user in their task and can be determined in the processing of texts. Information retrieval techniques based on term frequency distributions identify representative terms in a document for use in calculating query-document similarity and interdocument similarities for clustering. Generating document summaries are usually based on these statistical techniques, typically to extract sentences and generally to good effect [5].

Thus it is possible to present to the user summary representations indicating the key topic(s) - what the document is about – and the semantic relation held (if any) between the query terms as they appear together in the document. Furthermore, the terms with which these key terms co-occur could be shown or used to extract further sentences with a view to showing the aspects of the document/query topic(s). Representation of term distribution in the document or its structure may further indicate the meaning of the key terms in the document. It is possible to speculate that these snippets of information assist the user as they learn about the terminology and the concepts of their information interest; and, as they identify key words and formulate search expressions and tactics to manipulate the query; as well as, judging the relevance or utility of the retrieved results. Whether there is an optimal presentation, as defined above, must be addressed in further research involving users at various stages of search as they interact with results representations varying in content, size, form and structure.

Whilst this is at an early stage as a proposal it is based on decades of research on users' search interactions and research on surrogate representations. Only recently, however, has there been an interest in finding synergy between the research areas of information seeking and information retrieval. The notion that summaries serve quite specific purposes is not novel, but the proposal here aims to add a new dimension to the development and evaluation of representations specifically tailored to the users' task of formulating search. Challenging issues and questions remain for its effective implementation. These are not dissimilar to those that face the development of any information retrieval system that focuses on the users, tasks and contexts.

References:

1. Kuhlthau, C.C.: Seeking Meaning. Norwood, N.J: Ablex (1993)
2. Dumais, S., Cutrell, E., Chen, H.: Optimizing search by showing results in context. In: Proceedings of the SIGCHI conference on Human Factors in Computing Systems, Seattle (2001) 277-284
3. White, R.W., Jose, J.M., Ruthven.I.: A task-oriented study on the influencing effects of query-biased summarisation in web searching. Information Processing and Management. 39(5) (2003) 707-733
4. Tombros, A., Sanderson. M.: Advantages of query biased summaries in information retrieval. In: Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. (1998) 2-10
5. Paice, C.D.: Constructing literature abstracts by computer: techniques and prospects. Information Processing and Management. 26(1) (1990) 171-186.