A Comparative Study of the Effectiveness of Search Result Presentation on the Web

Hideo Joho and Joemon M. Jose

Department of Computing Science
University of Glasgow
17 Lilybank Gardens, Glasgow, G12 8QQ, UK
{hideo,jj}@dcs.gla.ac.uk

Abstract. Presentation of search results in Web-based information retrieval (IR) systems has been dominated by a textual form of information such as the title, snippet, URL, and/or file type of retrieved documents. On the other hand, document's visual aspects such as the layout, colour scheme, or presence of images have been studied in a limited context with regard to their effectiveness of search result presentation. This paper presents a comparative evaluation of textual and visual forms of document summaries as the *additional* document surrogate in the search result presentation. In our study, a sentence-based summarisation technique was used to create a textual document summary, and the thumbnail image of web pages was used to represent a visual summary. The experimental results suggest that both have the cases where the additional elements contributed to a positive effect not only in users' relevance assessment but also in query re/formulation. The results also suggest that the two forms of document summary are likely to have different contexts to facilitate user's search experience. Therefore, our study calls for further research on adaptive models of IR systems to make use of their advantages in appropriate contexts.

1. Introduction

The Internet has transformed into a main source of information for many and as a consequence web search engines have become an essential tool in our day to day life. Web search engines such as Google, Yahoo!, and MSN Search are processing millions of queries a day. The interaction paradigm of such engines follows more or less the same style assuming that this is the best for all users. However, recent user behaviour studies on commercial search engines challenge such assumptions [1]. In this paper, we conduct a comparative evaluative study assessing the effectiveness of various forms web search interfaces.

Most web search engines operate on a general principle of retrieval. Users' provide a set of query terms as a representative of their underlying information need. In response systems, after comparing the query to the documents in the collection, provide a list of potential documents which might contain useful information to satisfy the users' information need [2].

However, a number of issues are overlooked in such a simplistic view. The first one is that formulating a good query is proved to be cognitively challenging task for users [3]. Often queries are approximations of their underlying need and hence the

whole information seeking process is iterative in nature [4]. The second issue is in interpreting and assessing the relevance of documents in the returned list [5]. It has been shown that users of web search engines are reluctant to examine a large number of individual documents or even past the first page of the result list [1]. The users decision to view a document or not depend on the information in the document surrogates such as title, URLs and abstracts (often snippets extracted form the documents). The third issue is in the matching of the submitted query with the documents (or their indexes) with the intention of selecting a set of documents that contain information on the query. The first two issues make the development of information seeking interfaces a non-trivial task.

The major thrust of this paper is related to the first two issues. That is pertaining to user interaction which includes issues related to query formulation and judging the usefulness of each document in the list. In this paper we propose and evaluate a number of interfaces which facilitate the relevance (usefulness) judgement issues differently. The results of the experiments demonstrate the effectiveness of our proposed interfaces and demonstrate the inadequacy of current interfaces.

2. Background and motivation

The main purpose of search engines is to help people find information that is useful or relevant to completing a task. Search interfaces are the means through which users interact with search systems and control all aspects of their search.

The results from information seeking studies point to the fact that users look for information to complete a task. From a cognitive perspective, it has been termed that there is a gap in users' knowledge and the information is needed to fill this gap [3]. Finding relevant information may require running several queries, making judgments on the usefulness of documents returned, and reading many documents. Considering the importance of this task to many users, it is imperative to design interfaces that maximize the amount of information users can obtain during a search.

Submitted user queries are often an approximation of his/her underlying information need [6]. Since the system returns documents based on such queries, the usefulness of such documents are not certain. The documents in the result list might be not relevant, or partially relevant. Often documents contain partial information or redundant information from a previously seen document. In order to conduct an effective search it is imperative that users be able to make reasonably correct judgements about the documents in the result list.

Novel result visualisation techniques were proposed to address this problem [5]. Another techniques tried out are various summarisation techniques [7]. It has shown that the use of query-based summarisation techniques in improving the search effectiveness. Recently, new approaches to web page result presentation were tried. Most of these systems present the user with an unfamiliar, graphical interface that imposes an increased cognitive burden on the user or consider documents as finest level of granularity for result presentation [8].

Users' assessment of relevance of documents in the result list is based on the surrogates displayed (e.g., title, URL and snippets). It has been shown that such information is inadequate to provide effective search sessions [9]. For example, the quality of title information can vary mainly because of the casual approach to generating a title

at the time of web-page creation. The document snippets shown by many web search engines are fragments extracted from the whole document. Often such snippets are incomplete sentences extracted from the documents and as such inadequate to aid effective relevance judgements. The role of query-based sentences in assisting on making proper relevance assessments has been reported [9].

An aspect that affects information seeking process is the task at hand. For some tasks, it is important to know the genre of document. Often users may search for same information again and again. In this case, a thumbnail of the document would aid in judging the usefulness of the document [10].

In this work, we device three alternative forms of search result presentations. We use Google result presentation as a baseline. As additional surrogates we use query-based document summary called top ranking sentences, or TRS [11] and thumbnails of documents retrieved. While the previous study used the TRS as a replacement of Google snippet, in our interface, the TRS was used as additional information to the snippet. In addition, we make use of document thumbnails as a surrogate. We believe it is useful in judging the relevance and assessing the genre of a document. We augment information on the Google interface with thumbnails of documents. In the following session we will introduce these interfaces briefly.

The purpose of the experiment is to find the effectiveness of these interfaces in web information seeking tasks. We designed an experiment with real users, real tasks on the live Internet.

3. Interfaces for search result presentation

We augment web search system Google with 3 new interfaces. Our interfaces collect user queries and forward to the Google search system using Google API[1]. The result list from Google collected and processed. Information needed for new interfaces were created at this time. Like in web search result pages, user can peruse ten document records at a time. After this, they can either reformulate the query or peruse the next ten records.

As a baseline we use the Google interface. Three layouts of search results presentation were designed and compared to the baseline layout. All layouts were designed to show ten records per page as Google's default setting did. The rest of this section will present the layouts used in our experiment.

Layout 1: Baseline - The baseline layout was designed to provide an almost identical interface to the search result of Google. For each record, it had a title, snippet, URL, size, and the hyperlinks of cached page and similar page.

Layout 2: Baseline + TRS - The second layout integrated up to three top ranking sentences (TRS) into the baseline layout. The sentences were inserted below the snippet as a list, and background was highlighted to clarify the distinction between the TRS and snippet. The query terms were highlighted in bold in the same manner as it would have been in the title and snippet. There was some run-time overhead in generating TRS for retrieved documents. In order to minimise the difference of response time among the layouts, TRS was always created when a new query was submitted to the interface, but it was only displayed in Layout 2 and 4.

[1] http://www.google.com/apis/

Layout 3: Baseline + Thumbnail - The third layout integrated a thumbnail image of the web page screenshot into the baseline layout. The thumbnails were fetched from the Alexia's thumbnail archive[2]. The thumbnail was placed on the left side of the other document surrogates, and it was linked to the URL of the pages. The size of thumbnails was 112 (width) and 82 (height) pixels which was perhaps too small to read the texts, but we considered that it should be large enough to grasp the visual aspects of pages such as the layout, colour scheme, or the presence of images.

Layout 4: Baseline + TRS + Thumbnail - The last layout was the combination of Layout 2 and 3 (See Fig. 1). While this layout took up the largest space in the screen, it was designed to provide the largest amount of information per record among the four layouts.

Fig. 1. Search result with TRS and thumbnail (Layout 4)

As can be seen, we designed the four layouts so that the different variable between layouts remained to be a single element. This was due to our consideration for minimising the difference between layouts to evaluate the effectiveness of TRS and thumbnails in a systematic way. The next section will discuss the detail of our experiment based on the four layouts.

4. Experiments

A comparative user study was carried out to evaluate the effectiveness of the four search result layouts described above. This section will discuss the methodology adopted in our experiment.

A total of twenty-four people (6 female and 18 male) were recruited for our experiment. Most participants were the research students of the University of Glasgow. The rest was affiliated members of the University. The entry questionnaire established that the range of age varied from 20 to 37 with an average of 27.7. Their experience with search engines varied from 4.5 to 11 years with an average of 7.1 years. All participants carried out several searches every day, and 91.6% of them used Google most frequently.

[2] The Alexa archive (www.alexa.com) did not always contain the thumbnail of the web pages retrieved during our experiment. Our understanding was that a missing thumbnail was replaced by a parent site when it was available. Otherwise it showed the Alexa's logo image to indicate the absence of thumbnails. In our experiment, a missing thumbnail was treated as a similar case to a dead link on the web.

Participants were asked to carry out four search tasks in the experiment. The tasks were designed based on the simulated work task approach [12]. The simulated work task described a task as a form of short scenario. The scenario explained the contexts and motivation of the search with the sufficient information about the relevance of pages. The details of the tasks used in our experiment were as follows.

Task 1: Background search task - This task asked participants to find general background information on a topic. In our experiment, participants were asked to find the pages which provide the information about the recent change of student populations. Task 1 and the following Task 2 were originally used and replied by [13].

Task 2: Decision making task - This task asked participants to make a decision about a topic. In our experiment, participants were asked to find the best Hi-Fi speakers available in a target price. Participants were encouraged to compare the speakers' details in the decision making.

Task 3: Known item search task - This task asked participants to find the information about a topic which was previously known by the searcher. In out experiment, participants were asked to find the current whereabouts of a person who assumed to be a previous colleague of the searcher.

Task 4: Topic distillation task - This task asked participants to find a list of key resources for a topic. The definition of key resources was based on the instruction of the Web Track of TREC[3]. The main criteria for being a key resource was that the website was principally devoted to the topic. In our experiment, participants were asked to find the key resources for designer handbags.

One of our intentions behind the selection of these tasks was to investigate a different level of documents' textual and visual elements that were likely to be significant to complete the tasks. For example, Task 1 was likely to involve more textual information than visual while Task 4 was likely to involve visual aspects of documents in a greater degree than other tasks. Task 2 and 3 were supposed to involve both aspects in a similar degree.

The user study was carried out in the following manner. At arrival time participants were asked to read an information sheet which described the guideline for the participation and goal of the experiment. Upon the agreement of participation, participants were asked to fill in an entry questionnaire to indicate their age, sex, and search experience. Then they were presented with a training topic and explained the nature of simulated-work task. They were given approximately 10 minutes to familiarise with the search interfaces and task activity. During the training session, the four layouts were introduced to participants and the questions regarding the interface and tasks were answered.

During the tasks, participants were asked to bookmark the pages they thought relevant. However, no explicit instruction was given to participants regarding the number of bookmarks required to complete the tasks. We asked participants to bookmark pages to ensure their engagement to search tasks. Participants were given up to 15 minutes to complete a task, but allowed to end it when they felt they completed the tasks.

After the first task was completed, participants were asked to fill in a post-search questionnaire to provide subjective assessments about their search. Then a new task was given to them and change of layout was informed. The same procedure was re-

[3] http://es.csiro.au/TRECWeb/guidelines_2004.html

peated four times. The presentation order of topic and layout was rotated according to a Latin-Square arrangement to reduce bias from participants performing the same tasks with the same system in the same order.

After the completion of four tasks, participants were asked to fill in an exit questionnaire to indicate their overall preference of layouts, followed by an open-end interview to capture their feedback and comments of the result presentation and experiment.

5. Experimental results

This section presents the results of our experiment. A total of 96 search sessions were performed by participants and analysed in our investigation. Due to the nature of study concerning search results presentation, both quantitative and qualitative data were equally important to our study. The quantitative data such as participants' interactions with the interface were based on the system logs recorded during the experiment. The qualitative data such as participants' subjective assessments of search were established by the post-search questionnaires and exit interview. We used a 7 point scale to capture participants' assessments where a positive assessment was represented by a low score in the analysis.

The results of our experiment were analysed from four perspectives as follows: user interaction, relevance assessment, contribution of layout features, and finally, layout preference. In this section, the discussion about the results is often based on the comparison to our *baseline layout* (Layout 1). For simplicity, we sometimes use the term *summary layouts* to refer to Layout 2, 3, and 4. In addition, one of the aspects we were interested in this investigation was participants' decision of which document to visit from search results. We refer such decisions as to an *initial relevance assessment* in this section.

For most differentials presented in this section, the Kruskal Wallis Test was applied to the data to establish statistical significance of the results. When a statistical significance was found between groups, Dunn's post hoc test was applied to determine the significant pairs. Due to the size of cases and arguably large variance of layouts' performance across the tasks, we did not find a statistical significance in many cases. We only report it when the significance was found in this section.

5.1 User interaction

Table 1 shows participants' interactions with the four layouts evaluated in our experiment. The second column shows the average number of queries submitted to the interface per search session. The third column shows the number of words used in the queries. The fourth column shows the number of result pages viewed during the tasks. The fifth column shows the number of pages viewed per iteration. The sixth column shows the number of retrieved records clicked per result page. The seventh column shows the number of bookmarked URLs. The eighth column shows the time taken to complete the tasks. The numbers in 2nd to 5th rows are a mean value across 24 sessions, and the standard deviation of the value is shown in the brackets.

Table 1. User interaction

Layout	Iteration	Query Length	Page	Page / Iteration	Click / Page	Bookmark	Time (min)
1	5.6 (3.9)	3.6 (2.6)	7.7 (5.7)	1.5 (0.7)	2.2 (1.6)	4.0 (2.8)	13.5 (2.8)
2	8.5 (6.3)	3.4 (1.6)	10.1 (6.4)	1.3 (0.6)	1.6 (1.3)	3.4 (2.2)	13.3 (3.0)
3	7.4 (4.7)	3.9 (1.5)	10.1 (5.7)	1.9 (1.5)	1.3 (0.8)	3.6 (3.3)	13.6 (2.5)
4	7.6 (5.6)	3.2 (1.4)	10.0 (6.1)	1.7 (1.1)	1.6 (1.4)	4.0 (3.7)	13.7 (2.5)
Total	7.3 (5.2)	3.5 (1.7)	9.6 (6.0)	1.6 (1.0)	1.7 (1.3)	3.7 (3.0)	13.6 (2.6)

n=24 (Layout 1-4), n=96 (Total)

One of the noticeable differences in Table 1 is the number of iterations. Participants tended to submit more queries to the interface in the summary layouts compared to the baseline layout. We also examined the presence of phrases in the queries. The number of queries that had at least one phrase was six in Layout 1, while 31, 21, and 30 queries contained at least one phrase in Layout 2, 3, and 4, respectively. This suggests that participants were more engaged in query re/formulation with the summary layouts compared to the baseline layout.

More iteration in the summary layouts led to a greater number of result pages viewed by participants to find relevant documents. However, the click per page ratio shown in the sixth column of Table 1 suggests that participants tended to click fewer records in the summary layouts. Given that an underlying search engine was identical across the layouts, this suggests that participants made more relevance judgements based on the document surrogates in the summary layouts, thus, they did not have to visit the retrieved pages as much as they did with the baseline layout. This suggests that the additional information offered by TRS or thumbnails appears to facilitate participants' relevance assessments.

A statistical significance was found in the query length between Layout 1 and Layout 3. No statistical significance was found for other differentials.

Figure 1 shows the distribution of click-through documents' ranking positions. Two trends can be found for the summary layouts in this figure. Firstly, more clicks were found in the top ranking positions which can be due to the larger number of iterations. Secondly, the clicks were stretched across a wider range of the ranking position compared to the baseline layout. This suggests that the TRS and thumbnails can contribute to an increasing level of exhaustively in relevance assessments.

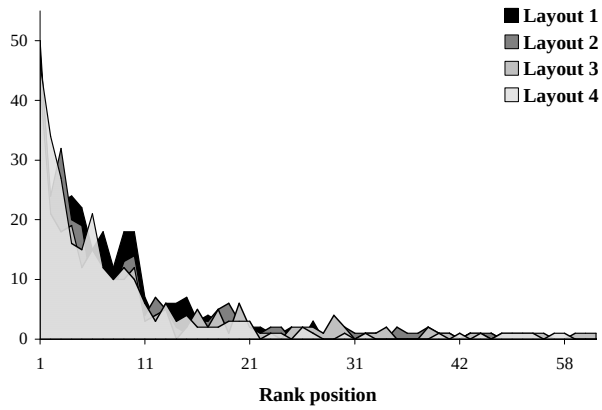

Fig. 2. Rank positions of click-through pages (n=1039)

5.2 Relevance assessments

The previous results indicated that the initial relevance assessment might be more focused and exhaustive when TRS and/or thumbnails were added to the document surrogates. Table 2 presents participants' perception of relevance assessment from three perspectives. The second column is the score given to the question regarding the ease of finding relevant documents from search results. The third column is the score regarding the ease of finding the documents which contained new information. Finally, the third column is the score regarding how often the documents contained the contents they expected to find in the full text. A positive assessment is represented by a low score (i.e., Score 1 = Very easy or very often, 7 = Not at all). The numbers are the average of 24 sessions, and the standard deviation of the mean value is shown in the brackets.

Table 2. Relevance assessment

Layout	Ease of finding	New information	Contents prediction
1	3.6 (1.7)	3.9 (1.6)	3.3 (1.5)
2	3.6 (2.1)	3.1 (1.4)	4.1 (1.7)
3	2.8 (1.5)	3.3 (1.4)	3.8 (1.9)
4	2.5 (1.6)	3.2 (1.1)	3.1 (1.8)
Total	3.2 (1.8)	3.3 (1.4)	3.6 (1.8)

n=24 (Layout 1-4), n=96 (Total)

As can be seen, a similar or more positive average score was found in Layout 2, 3, and 4 compared to Layout 1 with regard to the ease of finding relevant documents as well as of finding new information. While Layout 4 was given the best score among them in the ease of finding relevant documents, Layout 2 was given a better score than the others in the ease of finding new information. This suggests that both a tex-

tual and visual presentation of document's summary had the cases where user's relevance assessments were facilitated by them.

A slightly contradicting result was found in the expectation of document's contents. While Layout 4 was given a better score than Layout 1, participants tended to give a lower score to the other two layouts. It is not clear why participants found Layout 2 and 3 less predictable for document's contents, but there might be an unfamiliarity factor of TRS or thumbnails causing confusion in user's contents prediction.

5.3 Contribution of layout features

The previous sections highlighted the advantages and disadvantages of adding new elements to the baseline presentation. The overall results suggested that Layout 4 were likely to offer a better support in user's information seeking process than Layout 1. However, the difference between Layout 2 and 3 was less clear in several aspects. This section compares TRS and thumbnails by analysing the contribution of layout features in initial relevance assessments.

The document surrogates are some of the primary sources for the searchers to decide which documents to view from the search results. A typical document surrogate in search engines consists of the title, snippet, URL, size, and/or file type. In our experiment, after each task, participants were asked to indicate to what extent each of the layout features contributed to their decisions of viewing documents from the search results. Like the previous section, a 7 point scale was used for the assessments. The result is shown in Table 3 where a stronger contribution is represented by a low score. The numbers are the average, and the standard deviation of the mean value is shown in the brackets. Note that the sample size differs across the layout features.

Table 3. Contribution of layout features

Task	Title	Snippet	TRS	Thumb.	URL	Size	Type
1	1.6 (1.1)	2.4 (1.7)	2.1 (1.2)	4.4 (2.2)	4.3 (1.9)	6.7 (0.8)	5.8 (1.7)
2	1.9 (1.5)	2.7 (1.8)	2.7 (1.4)	4.8 (2.0)	3.5 (2.3)	6.5 (1.0)	6.4 (1.2)
3	2.1 (1.6)	2.0 (1.3)	2.6 (1.7)	4.3 (1.8)	4.4 (2.0)	6.8 (0.5)	5.8 (1.7)
4	1.8 (1.2)	2.5 (1.5)	3.3 (1.7)	2.3 (1.5)	3.0 (2.0)	6.4 (1.3)	5.8 (1.7)
Total	1.8 (1.4)	2.4 (1.6)	2.6 (1.5)	4.0 (2.1)	3.8 (2.1)	6.6 (0.9)	5.9 (1.6)

n=12 (TRS and Thumb. in Task 1-4), n=24 (the rest in Task 1-4), n=48 (TRS and Thumb. in Total), n=96 (the rest in Total)

The bottom row of the table suggests that participants often found the title of retrieved documents the strongest factor in deciding which document to view from the search result. This echoes the finding of [13]. When we compare TRS and thumbnail to Google's snippet, TRS was given a stronger score in Task 1 while the thumbnail was given a stronger score in Task 4. This suggests that the effectiveness of TRS and thumbnail can vary across the tasks. Also it indicates that the benefits of TRS and thumbnails might be mutually exclusive. In other words, the thumbnails might be useful where TRS are less effective, and vice versa.

We were also interested in the correlation of the layout features contribution. Table 4 shows Spearman correlation coefficient of seven layout features measured in our

experiment. The correlations that are statistically significant ($p < .05$) are highlighted in italic in the table. As can be seen, TRS was found to have a positive correlation with both the title and snippet of the retrieved documents. On the other hand, the thumbnails had a small but significant negative correlation with the snippet and a positive correlation with URLs. The negative correlation with the snippet again suggests that the usefulness of textual information might be mutually exclusive to the visual information in user's initial relevance assessments. Also, thumbnails' positive correlation with URLs indicates that the thumbnails can be more influential when the genre or category of web pages is an important factor in the search tasks.

Table 4. Correlation of layout features contribution

	Title	**Snippet**	**TRS**	**Thumb.**	**URL**	**Size**	**Type**
TRS	*.410*	*.314*	1.000	-.175	-.202	.010	.147
Thumb.	.210	*-.265*	-.175	1.000	*.284*	*.247*	.051

n=48

5.4 Layout preference and participants' comments

Upon the completion of four tasks, participants were asked to rank the four layouts in their order of preference. In the exit open-ended interview, participants had an opportunity to provide any comments and feedback regarding the layouts and overall experiment. We asked participants to rank the layouts based on the search experience with the given tasks. The most preferred layout was given Score 1 and the least was given Score 4. The counts of participant's ranking are presented in Table 5, along with the average ranking. As can be seen, nearly half of participants preferred Layout 4 over the other layouts. Layout 1 was give the largest votes as the least preferred layouts. There appeared to be a slight preference towards Layout 2 compared to Layout 3.

Table 5. Layout preference

Preference	**Layout 1**	**Layout 2**	**Layout 3**	**Layout 4**
1 (Most)	4	6	3	11
2	6	8	8	2
3	2	7	10	5
4 (Least)	12	3	3	6
Average rank	2.92	2.29	2.54	2.25

6. Implications

The results of our study have several implications for the design of search interface on the web. First of all, adding the new elements that are designed to support user's information seeking activity are likely to increase the level of interaction with a search interface. Our results suggest that query re/formulation and initial relevance assessments are likely to be facilitated by adding TRS and thumbnails to the result presentation. Given that many search engine users are reluctant to offer their effort in

search [1], it is encouraging to see the cases where the proposed presentation of search results can contribute to the enhancement of search experience. Our results indicate that the additional information might have a positive effect for increasing the number of iterations. Participants often found it easier to find relevant documents and new information when TRS and thumbnails were added to the document surrogate. This suggests that the current search engine's result presentation is not necessarily optimised and there is a room for improving the presentation.

Our study also provided additional insight into the nature of textual and visual forms of documents' summary. Previous study shows that, for example, TRS can be useful for supporting users of interactive IR systems [7], and the effectiveness of thumbnails can vary across the types of search tasks [14]. Our experiment with the four types of layouts allowed us to compare the effectiveness of these two forms of additional information in a systematic way. Moreover, TRS and thumbnails were evaluated as an additional element as opposed to a replacement of some of layout features used in the current search engines. Our results suggest that the textual and visual presentation of documents' summary is likely to offer additional information in a different context. Therefore, the effectiveness of TRS and thumbnails is often task dependent, but also their usefulness can be mutually exclusive in the search tasks. The overall positive performance of Layout 4 appears to be due to the fact that it could offer a support in a wider range of tasks than Layout 2 or 3.

One of our conclusions, therefore, might be that it is *safer* to show both the textual and visual summaries of documents in the result presentation. It might offer some searchers a greater degree of control in the selection of useful information to carry out searches. However, it is also likely that user's cognitive load will be increased when more elements are added to the search interface. Therefore, we suggest that the search interface should be able to offer a right form of document's summary in an appropriate context or task. Consequently, this study calls for more research on the understanding of users' search contexts and adaptive technique to capture their needs in an appropriate context. This study presented the cases where such advance can be used to improve several aspects of search experience.

7. Conclusion

This paper presented a user study investigating the effectiveness of search result presentation on the web. Both the textual and visual forms of document's summary were evaluated as additional information that can be integrated into the current search engine interface. Our evaluation investigated a wider range of aspects of information seeking behaviour than those previously carried out. Our results presented the cases where the additional elements of result presentation were likely to have a positive effect not only in the relevance assessment but also in query re/formulation. Therefore, it was suggested that the typical result presentation used in the major search engine was not necessarily optimised and had a room for improving searching experience. The textual and visual forms of document summaries were likely to have different contexts to facilitate user's search experience. However, we will need more progress on adoptive models of information retrieval systems to make use of their advantages in an appropriate context.

8. Acknowledgements

The authors thank to the participants of our experiment for their time and valuable feedback. This work was funded by the EPSRC (Grant ref: EP/C004108/1). Any opinions, findings, and conclusions described here are the authors and do not necessarily reflect those of the sponsor.

9. References

1. Jansen, B.J., et al., *Real Life Information Retrieval: A Study of User Queries on the Web.* ACM SIGIR Forum: A Publication of the Special Interest Group on Information Retrieval, 1998. **32**(1): p. 5-17.
2. Ingwersen, P. and K. Järvelin, *The Turn - Integration of Information Seeking and Retrieval in Context.* The Information Retrieval Series. Vol. 18. 2005: Springer.
3. Belkin, N.J., R.N. Oddy, and H.M. Brooks, *ASK for information retrieval: Part I. Background and theory.* Journal of Documentation, 1982. **38**(2): p. 61-71.
4. Belew, R., *Finding Out About- Search Engine Technology from a Cognitive Perspective*. 2000: Cambrdige University Press.
5. Hearst, M.A. and J.O. Pederson. *Re-examining the Cluster Hypothesis: Scatter/Gather on Retrieval Results*. in *Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1996. Zurich, Switzerland: ACM.
6. Taylor, R.S., *Question negotiation and information seeking in libraries.* College and Research Libraries, 1968. **29**: p. 178-194.
7. White, R., J.M. Jose, and I. Ruthven, *Using Top-Ranking Sentences to Facilitate Effective Information Access.* Journal of the American Society for Information Science and Technology, 2005. **56**(10): p. 1113-1125.
8. Chen, H. and S. Dumais. *Bringing Order to the Web: Automatically Categorizing Search Results*. in *Proceedings of the CHI 2000 Conference on Human factors in computing systems*. 2000. The Hague Netherlands: ACM.
9. White, R., J.M. Jose, and I. Ruthven, *A task oriented-study on the influencing effects of query-biased summarisation in web searching.* Information Processing and Management, 2003. **9**(5): p. 707-733.
10. Dziadosz, S. and R. Chandrasekar. *Do thumbnail previews help users make better relevance decisions about web search results?* in *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*. 2002. Tampere, Finland: ACM.
11. Tombros, A. and M. Sanderson. *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. Melbourne, Australia: ACM.
12. Borlund, P., *Experimental components for the evaluation of interactive information retrieval systems.* Journal of Documentation, 2000. **56**(1): p. 71-90.
13. Tombros, A., I. Ruthven, and J.M. Jose, *How Users Assess Web Pages for Information Seeking.* Journal of the American Society for Information Science and Technology, 2005. **56**(4): p. 327-344.
14. Woodruff, A., et al., *A comparison of the use of text summaries, plain thumbnails, and enhanced thumbnails for Web search tasks.* Journal of the American Society for Information Science and Technology, 2002. **53**(2): p. 172 - 185.