J Reid and M D Dunlop, "Evaluation of a Prototype Interface for Structured Document Retrieval", *Proceedings of HCI 2003*, Bath England, 2003 - Draft version – see proceedings for final version

Evaluation of a Prototype Interface for Structured Document Retrieval

Jane Reid[1] and Mark D Dunlop[2]

[1]*Computer Science, Queen Mary, University of London*
London E1 4NS, UK
Jane@dcs.qmul.ac.uk

[2]*Computer and Information Sciences, University of Strathclyde,*
Glasgow G1 1XH, UK
Mark.Dunlop@cis.strath.ac.uk

Document collections often display either internal structure, in the form of the logical arrangement of document components, or external structure, in the form of links between documents. Structured document retrieval systems aim to exploit this structural information to provide users with more effective access to structured documents. To do this, the associated interface must both represent this information explicitly and support users in their browsing behaviour. This paper describes the implementation and user-centred evaluation of a prototype interface, the RelevanceLinkBar interface. The results of the evaluation show that the RelevanceLinkBar interface supported users in their browsing behaviour, allowing them to find more relevant documents, and was strongly preferred over a standard results interface.

1. Introduction

Document collections often display structural characteristics. Structure can be found both within an individual document (e.g. a report may contain sections and subsections) and between documents (e.g. web documents may be connected by hyperlinks). Structured document retrieval (SDR) aims to combine structural and content information in order to improve retrieval effectiveness (e.g. Brin and Page 1998, Kotsakis 2002, Wilkinson 1994), cut down the amount of time and effort a user spends in identifying relevant information (e.g. Fuhr and Großjohann 2001, Roelleke 1999), and reduce time and disorientation caused by lack of proximity of related document components in results interfaces (e.g. Chiaramella et al. 1996).

Structural information can be exploited at several stages of the information retrieval process: firstly, the indexing stage, where document components are identified and indexed as separate, but related, units (Cleveland, Cleveland and Wise 1984, Tenopir and Ro 1990); secondly, the retrieval stage, using passage retrieval (e.g. Salton et al 1993), data modelling approaches (Burkowski 1992, Navarro and Baeza-Yates 1995), or aggregation-based approaches (e.g. Frisse 1988, Dunlop and Van Rijsbergen 1993, Lalmas and Moutogianni 2000, Roelleke et al 2002); and, thirdly, at the results presentation stage, using visualisation techniques such as TileBars (Hearst 1995), fisheye views (Furnas 1999) and expand/collapse operations (e.g. Hertzum and Frøkjær 1996), or grouping of related objects (e.g. Google's use of sub-lists in an otherwise traditional-style ranked document list; Northern Light Search's clustering interface).

The method that is investigated in this paper, however, focuses on exploiting users' natural browsing behaviour by employing the concept of *best entry points*. A best entry point (BEP) is a document component (or whole document) from which a user can obtain optimal access, by browsing, to relevant document components (Chiaramella et al 1996, Kazai et al 2001). The use of BEPs is thus intended to support the information-seeking behaviour of users, and enable them to gain more effective and efficient access to relevant information items.

Two methods of employing the concept of BEPs are currently being investigated. In the first approach, results presentation is explicitly focussed by presentation of BEPs, rather than relevant components. This approach is achieved through the use of *focussed retrieval*, which derives relevance scores for each document component based on the aggregation of the component itself and its structurally related components. This information may be used in conjunction with a set of heuristics to derive BEPs from a traditional ranked list of document components produced by an SDR system (Kazai et al 2002). Only these explicit BEPs are then presented to the user. In the second approach, standard relevant document components are presented to the user. However, the interface is designed to support users identify implicit BEPs within the results list

quickly and easily, i.e. document components from which the user can easily and quickly browse to other relevant document components.

This paper describes the implementation and initial evaluation of a prototype interface for supporting SDR, the RelevanceLinkBar (RLB) interface. The work discussed here involved the implementation of the RLB interface on the Web, using the Google interface as a basis. A user-centred experimental evaluation was then carried out to evaluate the potential ability of the interface to support explicit or implicit use of BEPs. The evaluation compared the effectiveness, efficiency and usability of the RLB interface with the standard Google interface.

Section 2 describes the RLB interface in detail. Section 3 outlines the elements of the experimental design: the participants, tasks, experimental methodology and data collection methods. Section 4 presents the main results of the experiment in terms of the interface's effectiveness, efficiency and usability within the context of its aim to support the use of BEPs. Both quantitative and qualitative analysis of data was performed. We close with conclusions and further work in Section 5.

2. The RelevanceLinkBar Interface

In this section, we discuss both the generic properties of the RLB interface, and its implementation in the specific context of this experiment.

The RLB interface (see figure 1) is a prototype interface that employs a novel visualisation technique based on a standard ranked results list, but additionally providing the explicit representation of any links found within the document; the motivation behind RLB is similar to that of Hearst's TileBars interface (Hearst 1995). This information is provided, for each document surrogate, in the form of a bar of boxes, each of which represents an individual link. Each link box is coloured to represent the degree of relevance of the corresponding linked document. The degree of relevance of each link could be calculated using one of many possible criteria, e.g. the presence or absence of query terms in the linked document, or the appearance of the linked document in the ranked list. Each bar thus provides three pieces of information: firstly, the total number of links contained in a document; secondly, the degree of relevance of each of those links; and, thirdly, a graphical representation of the distribution of the links within the document, with relation to each other. The assumption behind the representation of the link distribution is that it will facilitate the identification of documents, or document components, that display a high concentration of relevant links. By positioning the mouse pointer over an individual link box, any available information about the linked document, e.g. document title or keywords, can be viewed.

The interface is intended to support users' information-seeking behaviour in two main ways. Firstly, it is intended to enable more efficient browsing by its explicit representation of contained links. Secondly, it is intended to improve the quality of document surrogates as predictors of document relevance, by providing

information that will allow users to assess quickly the potential usefulness of a document as a starting-point for further investigation. It is thus an ideal candidate interface to support effective, implicit identification of BEPs for an SDR system. The user has two main browsing strategies open to him: to click on the URL belonging to a document directly represented in the ranked results list, and browse from there, or to progress indirectly to a linked document by clicking on one of the RLB link boxes.

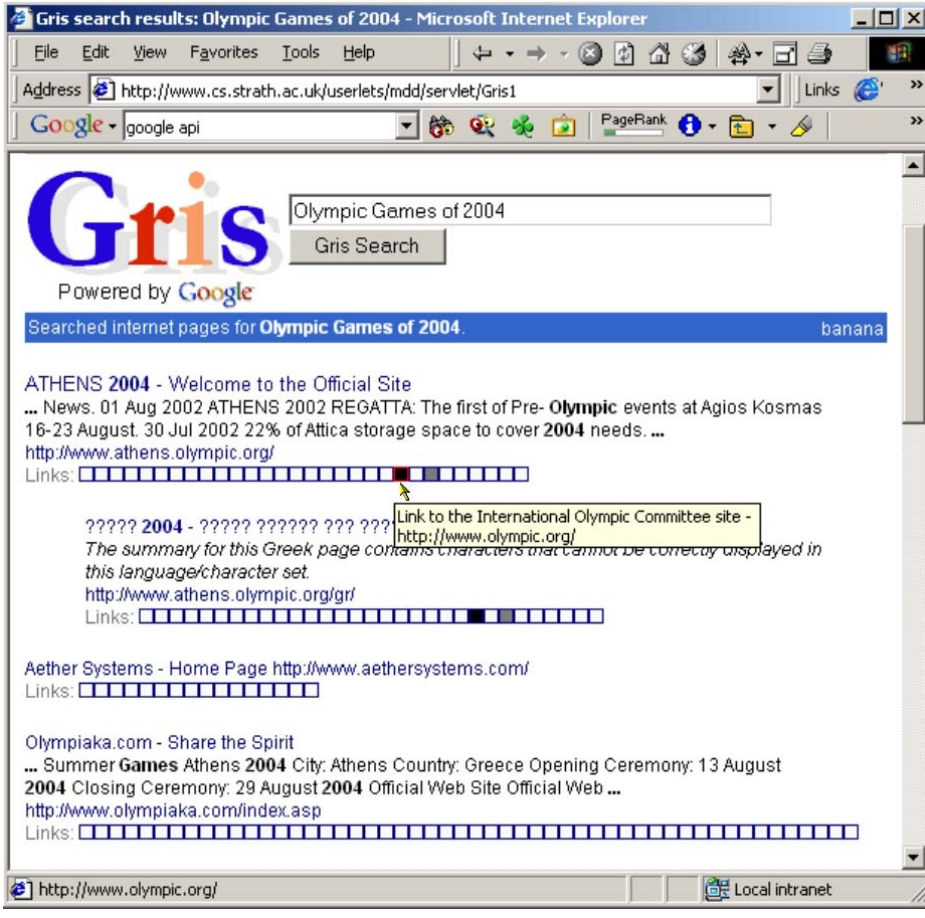

Figure 1. The RelevanceLinkBar Interface.

For this work, the RLB interface was constructed on top of the Google ranked results list interface. The RLB was implemented as a Java servlet querying Google (initially through HTML parsing of results, and now using the Google

API). The experiment reported in this paper used two versions of the Google interface:

- *Plain* - a slightly reduced version of the standard Google interface;
- *RLB* – identical to plain except for the addition of relevance link bars.

Both versions used the same servlet for querying and were coded, as far as possible, to have comparable search times. Once the top 100 Google search results had been retrieved, each entry in the top 10 was post-processed as follows: (1) The target page was loaded by the Java servlet; (2) The page was parsed to extract its list of links; (3) Each of these links was then annotated with the position of its target page in the original top 100 search results. Once annotated links had been produced for the top 10 search results (or a preset time limit had expired), the servlet returned the results as a plain or RLB page, depending on the interface version. For the RLB interface, each annotated link was translated into a box on the RelevanceLinkBar, with the darkness of the box being determined by the closeness of the link destination to the top of the Google Top100. Although this prototype servlet implementation slows searching to some degree, any future implementation would be based at a search engine site, where much of the information is stored locally. The RLB link bar itself is composed by repetition of four small images, and uses plain HTML. The impact of the RLB on query execution time would, therefore, be negligible.

An initial technical evaluation of the RLB interface was carried out, with the aim of verifying the claim that it offers additional and distinct information from a standard ranked results interface. A sample of 19 queries from MetaSpy.com was submitted to Google, and 100 documents were requested in the results set. RLBs were then calculated for the resulting ranked document list, according to the method described above. The data were examined for a correlation between document ranking and the percentage of relevant links contained within each document. There was no noticeable correspondence between these two measures, confirming that the concept of RLBs offers additional information not available from a standard ranked document list. The next stage was to perform a user-centred experimental evaluation in order to examine the characteristics of the RLB interface in more detail.

3. Experimental Design

The overall purpose of the experiment was to evaluate the ability of the RLB interface to support users in the use of BEPs. This involved a comparison of the RLB interface with the plain ranked results interface, in terms of effectiveness, efficiency, usability and overall user satisfaction. The following sub-sections discuss in more detail the participants, tasks, experimental methodology and data collection methods.

3.1 Participants

Twelve participants were recruited, and a pre-questionnaire (Questionnaire A) was issued to all participants, in order to elicit information on personal and educational background, domain knowledge, collection knowledge and general information-seeking preferences and strategies. It was also used to collect descriptions of four real information needs per participant, two of which were later selected by the experimenter as the basis for experimental tasks for that participant (section 3.2).

Of the 12 participants, 5 were male and 7 female, and the average age was 25 (youngest 22, oldest 36). Eleven of the 12 were students and 1 was a desktop publishing employee. The students were mostly from a computer science background (7 participants), with some from civil engineering (2) and information technology (2). On a five-point scale, 11 of the 12 participants described both their experience with computers and their experience with Internet tools as excellent (point 1 on the scale) or good (point 2). Seven of the 12 used the Web on a daily basis, and all participants used it at least once a week. Ten of the 12 participants had used Google; other previously used search sites were Yahoo and Lycos.

Participants were also asked some general questions about their usual information-seeking preferences and strategies. The first question in this category was intended to establish the participants' criteria for a successful search. Four participants stated that they preferred to be presented with only highly relevant documents, while seven aimed to find most of the relevant documents available. This indicates that the participants were (unusually) biased towards a recall-oriented search, rather than a precision-oriented search. The majority of participants consciously employed both query-based and browsing strategies in the course of their information-seeking, with only 1 participant claiming to use querying exclusively, and only 2 claiming to use browsing exclusively. Half the participants stated that, when faced with a ranked document results list, they examined the documents sequentially, according to the ranked order, while 5 stated they examined the documents selectively. One participant stated that he combined these two strategies. Participants were also asked for reasons why they might consider a document to be non-relevant: common reasons were the language and age of the document, and the quality of the information contained within it.

3.2 Tasks

After a short period exploring each interface, each participant was allocated 4 tasks in total, 2 for completion during the first stage of the experiment and two during the second stage. Half of these tasks (one for each stage) were based on information needs gathered from the participants themselves in Questionnaire A.

The other tasks were simulated, i.e. generated by the experimenter. This design was adopted in order to provide some tasks for which the participants should be highly motivated to find the answers, and some for which results could be compared across participants. The simulated tasks, which were chosen to be of general interest, i.e. not requiring any specialist knowledge or understanding, were:

1. I would like to find information on Greek philosophy and philosophers. I would like to read about the philosophers of ancient Greece (e.g. Aristotle and Plato) and their work. This is a general interest of mine.
   Search statement: Greek philosophy and Greek philosophers
2. I have to write a report on the Seven Wonders of the World. I have been asked to give a good description of each (e.g. where they are located, the history of them, why they are considered wonders, etc).
   Search statement: Seven Wonders of the World

3.3 Experimental Methodology

The experiment was divided into two main stages: a usability evaluation and a functionality evaluation. Both interfaces were used in both stages of the experiment. We wished to assess both the participants' first impressions of the interface and their opinions after they had used the interface for a period of time. It was, therefore, decided to run the usability evaluation first, in order to collect users' first impressions. Any confusion or remaining queries could then be answered before the second stage of the experiment, the functionality evaluation. This would avoid any bias in our functionality results due to lack of knowledge about the interface.

A within-subjects design was followed, with each participant undertaking both experimental conditions in both stages of the experiment. Each participant was allocated 4 tasks in total, two for completion in the first stage of the experiment (one real and one simulated), and two for the second stage (one real and one simulated). The participants were assigned to experimental conditions so that half of them performed a real task followed by a simulated task in the first stage, then a simulated task first followed by a real task in the second stage. The other 6 participants performed the real and simulated tasks in the opposite order. Different sets of tasks were used for stages 1 and 2 of the experiment, in order to avoid learning effects. Since the participants all had previous experience of using both Web search engines and the RLB interface itself (from practice sessions), it was not deemed necessary to alternate the order of the experimental conditions themselves, so the participants all used the RLB interface first, and the plain interface second.

3.4 Data Collection

Background information about the participants was collected by a pre-questionnaire (Questionnaire A, see section 3.1). There were 2 further questionnaires: a usability questionnaire (B), and a functionality questionnaire (C). In addition, two recording forms were employed: a relevance recording form (Recording Form A) used by the participants, and an observation form (Recording Form B) used by the experimenter. Table 1 shows a summary of the experimental methodology and data collection methods.

	User Tasks	Participant	Experimenter
Stage 1		Questionnaire A *(background)*	
	Usability task 1 Usability task 2		
		Questionnaire B *(usability)*	
Stage 2	Functionality task 1	Recording Form A *(relevance)*	Recording Form B *(observation)*
		Questionnaire C part 1 *(functionality)*	
	Functionality task 2	Recording Form A *(relevance)*	Recording Form B *(observation)*
		Questionnaire C part 2 *(functionality)*	
		Questionnaire C part 3 *(preference)*	

Table 1. Summary of experimental methodology.

The usability questionnaire (B) was intended to elicit opinions on the learnability, ease of use, and good and bad points of the interface, together with suggestions for improvements and a preference for one of the two systems.

On the relevance recording form (Recording Form A), participants recorded their order of assessment of the documents in the results list and their assessment of the relevance of each individual document (relevant, partially relevant or non-relevant). They also recorded their desired ordering of the documents, i.e. the order in which they would like to have viewed the documents originally.

At the same time, the experimenter recorded her observations (Recording Form B) of the number of visited links, the number of steps a participant took to fully explore an individual document, the number of unreliable links encountered and the total amount of time spent on each document.

After each task in stage 2, the participants filled in part of a functionality questionnaire (Questionnaire C), which elicited information regarding their satisfaction with the results for that particular task.

Finally, the user was asked to restate an overall preference for one of the two interfaces (in Questionnaire C).

4. Results And Analysis

The results are presented under 5 main headings, each section corresponding to one of the questionnaires or forms. Where appropriate, results are given across all tasks, and across simulated tasks only. In order to check that participants were equally motivated when performing simulated tasks as real tasks, the amount of time the participants spent on each type of task was analysed. For real tasks, the total amount of time spent was 104 minutes 4 seconds (mean 8 minutes 40 seconds per task). For simulated tasks, the total amount of time spent was 86 minutes 28 seconds (mean 7 minutes 12 seconds per task). This difference was found to be non-significant at $p<=0.10$ using a parametric T-test, thus showing that participants treated both types of task with equal seriousness.

4.1 Usability Analysis

Participants were firstly asked to rate the RLB interface on a 3-point scale for 2 questions: how straightforward and easy to use it was, and how reliable the links were. The results are shown in Table 2.

Ease of use of RLB	
Straightforward	4
Ambiguous	7
Totally confusing	1

Link reliability	
Reliable	9
Relatively reliable	3
Completely unreliable	0

Table 2: Usability results.

Participants were then asked to state up to 5 good points and 5 bad points of each interface. Although a broad variety of opinions were displayed here, some of the common views expressed were as follows:

- The plain interface was described as simple to use (5 participants), providing sufficient information (4) and a familiar interface (3). However, 3 participants pointed out that browsing is necessary to find relevant information.
- The RLB interface was praised for its indication of relevant links (7 participants), economy of space (2), time-saving support for browsing (2), and the additional information provided by the link bar (2). However, 2 participants stated that the use of colour was not a good indicator of relevance and 2 stated that the pop-up boxes for individual links did not provide enough information.

It is clear from the participants' stated interface preferences that the value of the additional information provided by the RLB interface, together with its ability

to allow users faster access to relevant information, outweigh the disadvantages of the interface.

4.2 Relevance Assessment Analysis

Analysis of the participants' relevance assessments was based on data gathered from Recording Form A, which was filled in by the participants as they completed each of the two tasks in the second stage of the experiment. The analysis focussed on comparison of the number of relevant, partially relevant and non-relevant documents found using the RLB interface vs. the plain interface. This analysis was performed across all tasks, and then across the simulated tasks only. The statistical test used was the parametric t-test (related for within-subject, used across all tasks, and unrelated for between-subjects, used across simulated tasks only).

Table 3 shows the mean number of documents found per relevance category over all searches with a given interface and, in parentheses, over only the simulated tasks. Significant results are shown in *emphasis* ($p<=0.10$).

	RLB	Plain
Relevant documents found	5.5 (*4.83*)	4.83 (*3.50*)
Partially relevant docs found	2.42 (2.50)	2.33 (2.00)
Non-relevant found	*1.66 (2.00)*	*2.58 (4.33)*

Table 3: Mean documents found per relevance category.

In summary, participants found significantly less non-relevant documents with the RLB interface than the plain interface. The results also indicate that participants found more relevant and partially relevant documents with the RLB interface; however, this finding was only significant for relevant documents found during simulated tasks.

4.3 Ranking Correlation Analysis

Participants were asked to state two rankings on Recording Form A: firstly, the order in which they assessed the documents, and, secondly, the order in which they finally decided that they would like to have originally seen the documents. The system's ranking of the retrieved documents was also logged. Correlations between these rankings were then investigated by means of Spearman's Rank Correlation with the following aims:

- The system ranking / order of assessment correlation was analysed in combination with the participants' stated preference for judging document rankings sequentially or selectively (section 4.1), in order to identify which interface provides better support for selective examination of documents.

- The order of assessment / final ranking correlation examined which interface provides better support for participants in identifying a good order of assessment of documents.
- The final ranking / system ranking correlation examined which interface better matched the participants' retrospective evaluation of document relevance.

None of these correlations were found to be significant, indicating no difference in performance between the two interface variants ($p<=0.025$). Overall, participants were slightly more likely to disagree with the system ranking of the RLB interface than the plain interface, which may indicate that the display of links provided by the RLB encouraged participants to seek an alternative order of assessment. Participants may have been inhibited from making a more consistent attempt to identify implicit BEPs by the small number of retrieved documents and the strong influence of presentation order in this context (Purgailis Parker and Johnson 1990). This supposition is supported by the sequential assessment of documents, over both interface variants, by several participants who had previously stated a preference for selective assessment.

4.4 Functionality Analysis

Participants were asked to rate their experience of each of the interfaces after they had performed the task using that interface. They were asked to comment on their satisfaction with the results, what contribution the results had made to the resolution of the problem, and whether it was worth the time spent. Participants were asked, at the end of the experiment, to express a preference for one interface. This preference was then compared with the preference expressed after stage 1 of the experiment (usability evaluation) to see if their opinions changed with greater exposure and experience.

Table 4 shows overall satisfaction with the results and how much users felt the results contributed to resolution of the problem. Eleven of the participants using the RLB interface and all participants using the plain interface stated that the time they had spent on their searches had been worthwhile. Ten of the 12 participants stated a final preference for the RLB interface over the plain interface.

One participant, performing a real task using a very general search statement on the RLB interface, experienced considerable problems, which she attributed to the interface; this dissatisfaction is reflected in the results presented here. However, it was later determined that the problem lay with the task, and would, therefore, have been replicated if repeated on the plain interface.

Overall satisfaction		
	RLB	Plain
Very Satisfied	7	4

Contribution to problem resolution		
	RLB	Plain
Substantial	7	6
Good	4	5
Little	0	1

Satisfied	4	8
Dissatisfied	1	0

Very little	1	0
Nothing	0	0

Table 4: Results from functionality analysis

In summary, the RLB interface showed slightly higher levels of overall satisfaction and satisfaction with the individual search results, and was strongly preferred over the plain interface.

4.5 Observation Measure Analysis

The experimenter observed 3 main measures in the course of the experiment: the number of links the participants visited during the session, the number of steps involved in each task, and the time spent in evaluation of the retrieved documents. Again, the parametric T-test was used to test for statistical significance.

Table 5 shows the results – none of which were statistically significant ($p<=0.10$). The, albeit non-significant, differences can be partially explained by the observation that some participants first followed the main links from the retrieved list (ignoring the RLB bar) and browsed to other links from within these documents. They then returned to the retrieved list and followed links from the RLB, resulting in another visit to the same pages. This appeared to be a way of checking the accuracy and reliability of the RLB interface, so this effect could be reduced by further experience with the interface.

	RLB	Plain
Links visited	18.67 (18.83)	14.75 (14.83)
Total steps	37.83 (38.67)	29.58 (29.67)
Time spent (min:sec)	8:50 (9:29)	7:02 (7:51)

Table 5: Observational results

In summary, although participants did spend more time and effort on the RLB interface, this difference was not significant, and the participants appeared to judge any additional effort worthwhile.

5. Conclusions And Further Work

This paper has introduced the RelevanceLinkBar (RLB) interface for supporting structured document retrieval, described a prototype implementation of the interface, and presented the results of an initial user-centred evaluation. The results of the evaluation show that users found more relevant, and less non-

relevant, documents when using the RLB interface compared to a standard web search interface.

The evaluation failed to prove that the RLB interface is suitable for the implicit identification of BEPs. This was demonstrated by the lack of correspondence between the order of document assessment and the final user ranking, showing that the interface did not significantly support effective identification of BEPs by the participants. This result requires further investigation using a larger scale experiment with more than 10 retrieved documents per query, as the small size of the retrieved set may have artificially discouraged participants from scanning the results list selectively in order to find best entry points.

The results did show, however, that the interface provided good support for browsing, as evidenced by the increased use of links and the qualitative feedback elicited from the participants. We can conclude, therefore, that the interface should prove effective when used in conjunction with explicit representation of BEPs. A further experiment to test, directly, the validity of this conclusion should be carried out.

Finally, although the RLB interface was preferred to the standard interface by 10 of the 12 participants, improvements are required in order to ensure that the interface is both reliable and usable. More information in the pop-up boxes for individual links in the RLB would be useful, e.g. document title or brief surrogate. The graded relevance links did not always appear to support users effectively in their identification of relevant documents, so further investigation of the use of degrees of relevance for links representation is recommended. Other variants of the RLB, e.g. showing only relevant (or partially relevant) links, should be implemented and evaluated, in order to assess what combination of information best supports users in their information seeking behaviour. In addition, in order to support both effective browsing behaviour and effective navigation, the RLB could continue to be shown throughout the examination of the main documents from the retrieved list and the documents linked from those. This could be achieved by presenting the RLB for each of the main documents in a separate window, while linked documents are being examined. Further experiments to evaluate all the above variations will be required.

Acknowledgment

The main experimental work described in this paper was carried out at Queen Mary, University of London during summer 2001 by Zoi Gkaranatsi, as fulfilment of the project component of the MSc in Advanced Methods in Computer Science. The project was partly supported by The Royal Society international exchange programme.

References

Brin, S. and Page, L. The Anatomy of a Large-scale Hypertextual Web Search Engine. In 7th WWW Conference (Brisbane, 1998).

Burkowski, F. J. Retrieval Activities in a Database Consisting of Heterogeneous Collections of Structured Texts. In 15th ACM SIGIR (Copenhagen, 1992) 112-125.

Chiaramella, Y., Mulhem, P. and Fourel, F. A Model for Multimedia Information Retrieval, Technical Report Fermi ESPRIT BRA 8134, University of Glasgow (1996).

Cleveland, D. B., Cleveland, A. D. and Wise, O. B. Less than full text indexing using a non-Boolean searching model. JASIS 35(1984) 19-28.

Dunlop, M.D., and Van Rijsbergen, C.J., "Hypermedia and free text retrieval", Information Processing & Management, 29(3), May 1993.

Frisse, M. Searching for Information in a Hypertext Medical Handbook. Communications of the ACM 31(1988) 880-886.

Fuhr, N. and Großjohann, K. A Query Language for Information Retrieval in XML Documents. In 24th ACM-SIGIR (New Orleans, 2001) 172-180.

Furnas, G. W. The fisheye view: A new look at structured files in Card, S. K., Mackinlay, J. D. and Shneiderman, B. (eds.), Readings in Information Visualization: Using Vision to Think, 312-330, Morgan Kaufmann, 1999 (Reprinted from: The Fisheye View: A New Look at Structured Files, Bell Laboratories Technical Memorandum #81-11221-9, October 12, 1981).

Hearst, M. A. TileBars: Visualisation of Term Distribution Information in Full Text Information Access. In Proceedings of the ACM Conference on Human Factors in Computing Systems (1995) 59-66.

Hertzum, M. and Frøkjær, E. Browsing and querying in online documentation: A study of user interfaces and the interaction process. In ACM Transactions on Computer-Human Interaction, 3(1996)136-161.

Kazai, G., Lalmas, M. and Roelleke, T. A Model for the Representation and Focussed Retrieval of Structured Documents based on Fuzzy Aggregation. In String Processing and Information Retrieval, SPIRE 2001 (Laguna De San Rafael, 2001).

Kazai, G., Lalmas, M. and Roelleke, T. Focussed Structured Document Retrieval. In String Processing and Information Retrieval, SPIRE 2002 (Lisbon, 2002).

Kotsakis, E. Structured Information Retrieval in XML documents. In Proceedings of the 17th ACM Symposium on Applied Computing, SAC'02 (Madrid 2002).

Lalmas, M. and Moutogianni, E. A Dempster-Shafer Indexing for the Focussed Retrieval of a Hierarchically Structured Document Space: Implementation and Experiments on a Web Museum Collection. In 6th RIAO Conference on Content-Based Multimedia Information Access (Paris 2000).

Navarro, G. and Baeza-Yates, R. A Language for Queries on Structure and Content of Textual Databases. In 18th ACM-SIGIR (Seattle, 1995) 93-101.

Purgailis Parker, L. M. and Johnson, R. E. Does Order of Presentation Affect Users' Judgement of Documents? JASIS 41(1990) 493-494.

Roelleke, T. POOL: Probabilistic Object-Oriented Logical Representation and Retrieval of Complex Objects - A Model for Hypermedia Retrieval, Ph.D. Thesis, University of Dortmund, Verlag-Shaker (1999).

Roelleke, T., Lalmas, M., Kazai, G., Ruthven, I. and Quicker, S. The Accessibility Dimension for Structured Document Retrieval. In 24th European Colloquium on Information Retrieval Research, ECIR02 (Glasgow, 2002).

Salton, G., Allan, J. and Buckley, C. Approaches to Passage Retrieval in Full Text Information Systems. In 16th ACM SIGIR (Pittsburgh, 1993) 49-58.

Tenopir, C. and Ro, J.S. Full Text Databases. Greenwood Press, 1990.