Interactive Okapi at TREC–6

M M Beaulieu*   M J Gatford*

January 1998

1 Experimental setting

A full description of the experimental system and conditions is given in Appendices A and B. Searchers filled in three types of questionnaires. The pre-session questionnaire established the user's experience and profile. In the post search questionnaires, searchers were asked questions regarding the topic, the search and the system used after undertaking each individual search. Finally in the post-session questionnaire, subjects were asked to provide an overview of the experiment. In addition to the questionnaires, searchers noted on a worksheet the different aspects of the topics they encountered whilst they undertook each search.

A total of eight subjects completed forty eight searches, that is three searches on each of the two systems, Okapi and ZPrise. The sessions were divided into two rounds of four searchers. Of the two groups who carried out the twenty-four searches on Okapi, Group A used the same interface as in TREC-5, but with incremental query expansion modified (Appendix A.3.2), and Group B searched a slightly different version which allowed the searcher to cancel the relevance feedback process or clear the query (Appendix A.4).

1.1 Searcher profiles and experience.

Subjects were post-graduate students at Masters level recruited from the Information Science Department at City. As such they had all undertaken some courses in information retrieval and at least in theory had been introduced to probabilistic methods and the principles of relevance feedback and query expansion. However in practice only three out of the eight declared they had some experience in searching ranked output IR systems and only two in systems with relevance feedback. Of those one subject had previously used another version of the Okapi system.

In terms of their overall online experience, half had had a year or less and half had between two and four years experience in searching systems with mouse-based interfaces. Searching on Web browsers was the most common for all subjects, followed by library catalogues. Half had little or no experience of searching CD-ROM systems, commercial online systems or full-text databases.

More than three-quarters of searchers claimed to have very little or no knowledge of the topics. Just under a quarter said that topics 339 on 'Alzheimer's Drug Treatment ' and 347 on 'Wildlife Extinction' were marginally familiar.

2 Search Results.

2.1 Systems And Rounds.

The overall results for aspectual precision and recall for both systems are given in Figure 1. The breakdown by round in Figure 2 shows slightly higher precision in round one and slightly higher recall in round two.

Comparative results (Figure 3) also show higher precision for both rounds for searches on ZPrise and higher recall for those on Okapi. Searches on Okapi achieved better precision in round one and only marginally better recall in round two.

*Centre for Interactive Systems Research, Department of Information Science, City University, Northampton Square, London EC1V 0HB, UK

System	Measure	Mean	Median	Variance	Range
Okapi	Recall	0.400	0.391	0.082	0.000 – 1.000
	Precision	0.706	0.833	0.098	0.000 – 1.000
ZPrise	Recall	0.381	0.261	0.107	0.000 – 1.000
	Precision	0.809	0.877	0.081	0.000 – 1.000

Figure 1: Recall And Precision By System

Round	Measure	Mean	Median	Variance	Range
1	Recall	0.385	0.317	0.102	0.000 – 1.000
	Precision	0.784	0.868	0.080	0.000 – 1.000
2	Recall	0.396	0.359	0.087	0.000 – 1.000
	Precision	0.731	0.833	0.103	0.000 – 1.000

Figure 2: Recall And Precision By Round.

2.2 Individual Topics.

Figure 4 presents results for individual topics for each system ranked by aspectual precision and recall as given in Appendix D. The ranking by recall for both systems is almost identical except for topics 307 'New Hydroelectric Projects' and 347 'Wildlife Extinction', which are in position 4 and 5 in Okapi and reverse order in ZPrise. Topic 322 'International Art Crimes' ranks the lowest in both measures for both systems. In terms of precision, the greatest discrepancies is between topics 339 'Alzheimer's Drug Treatment' and 303 'Hubble Telescope Achievements', which rank 1 and 5 in Okapi and 4 and 2 in ZPrise respectively.

2.3 Searchers.

A comparison of searcher performance on the different systems (Appendix D.1) indicates that searchers in round one performed equally well on both system with half getting better precision/recall on either Okapi or ZPrise. However in round two all four searchers achieved higher precision on ZPrise, whereas recall was comparable between the two systems.

There is no evidence to show that the options introduced in the Okapi interface in round two had any effect on search performance. The 'clear feedback' option was used only once on topic 322, and the 'clear query' option six times, five on topic 322 and once on 339. The results for these topics in Appendix D show that in both cases precision and recall were lower than in round one.

Round	System	Measure	Mean	Median	Variance	Range
1	Okapi	Recall	0.393	0.391	0.092	0.111 – 1.000
		Precision	0.756	0.868	0.083	0.200 – 1.000
	ZPrise	Recall	0.378	0.261	0.122	0.000 – 1.000
		Precision	0.811	0.917	0.083	0.000 – 1.000
2	Okapi	Recall	0.407	0.388	0.080	0.000 – 1.000
		Precision	0.655	0.829	0.117	0.000 – 1.000
	ZPrise	Recall	0.384	0.263	0.101	0.000 – 1.000
		Precision	0.807	0.877	0.086	0.000 – 1.000

Figure 3: Recall And Precision By Round And System.

	OK		ZP	
Rank	Precn.	Recall	Precn.	Recall
1	339	303	326	303
2	326	339	303	339
3	307	326	347	326
4	347	307	339	347
5	303	347	307	307
6	322	322	322	322

Figure 4: Results Of Individual Topics In Ranked Order.

3 Search Perceptions And Performance.

3.1 Searching task.

At the end of the experiment six searchers felt that they had had a good or very good understanding of the task involved and two had reservations. Only half the subjects noted meaningful aspects on the search worksheets and three left them blank. With regard to the typicality of the searching task, only one searcher expressed that the task was similar to the searching tasks normally performed whilst the rest felt it only had some similarity.

3.2 Search Difficulty.

Overall half of the searches were classed as easy, a quarter as marginally difficult and a quarter as difficult. However more searches were deemed to be easy on ZPrise than on Okapi, 58% compared to 38%. In particular the searches on topic 339 'Alzheimer's Drug Treatment' and 326 'Ferry Sinkings' were considered to be unanimously easy by the four searchers on ZPrise but only by one searcher on Okapi. The results show that both topics were ranked second and third in both systems respectively in terms of recall but precision and recall were higher on Okapi. The six topics searched across both systems were quite evenly distributed with a third covering less than 25% of aspects, a third between 25% and 50% and a third over 50%.

3.3 Search Satisfaction.

In spite of searchers' perception that more topics were easy to search on ZPrise, 50% of searches carried out on the Okapi system were deemed to have produced a satisfactory outcome compared to 37% on ZPrise. But an equal proportion (34%) were found to be not satisfactory for both systems. The greatest difference appeared in topic 303 'Hubble Telescope Achievements' and 347 'Wildlife Extinction', where searchers on the Okapi system were more satisfied with the search outcome. In both cases the results do not bear out the searchers' perception. Precision and recall for the two topics were higher for ZPrise.

3.4 Confidence In Coverage Of Aspects.

Searchers of the Okapi system expressed that they were slightly more confident that they had identified all the possible aspects for the topics searched than those using ZPrise, 33% compared to 29%. However overall across both systems, responses indicate that 31% were confident, 29% were marginally confident and 40% were not confident. Searchers were the least confident about topic 322 'International Art Crime' (7 searchers out of 8) and the most confident about topic 303 'Hubble Telescope Achievements'. The lack of confidence for topic 322 was confirmed by the poor results for both systems on both measures. On the other hand 303 achieved the highest recall in both systems particularly in ZPrise where precision and recall were both higher.

In addition it is worth noting that, although ZPrise searches on topic 339 'Alzheimer's Drug Treatment' were unanimously considered to be easy, the level of confidence was unanimously marginal. By contrast three out of four Okapi searchers found the topic marginally difficult but they were confident that they had covered all the aspects. The topic ranked second in both systems on recall but the results for precision were better on Okapi.

3.5 Searching Time.

A much higher percentage of Okapi searchers considered that they had enough time to carry out an effective search than those using the ZPrise system, 45% as opposed to 25%. For 17% of searches users claimed to have marginally enough time compared to 33% for ZPrise. For both systems topics 322 ‘International Art Crime’ and 347 ‘Wildlife Extinction’ appear to have been the most problematic. As already mentioned topic 322 produced poor results overall, whereas 347 ranked 4th and 5th on recall and 3rd and 4th on precision in ZPrise and Okapi respectively.

4 System Perception.

4.1 Ease Of Use.

Overall searchers appear to have perceived little difference in the ease of use of the two systems. Both systems were deemed to be difficult to use in only 8% of the searches. Okapi was classed as easy to use for 63% of the searches as compared to 59% for ZPrise.

4.2 Learnability.

Okapi was deemed to be easy to learn for 75% of the searches undertaken and marginally easy for 25%. In the case of ZPrise, the system was considered to be easy to learn for 46% of searches, marginally easy for 50% and not easy for 4%. Topic 303 ‘Hubble Telescope Achievements’ showed the greatest discrepancy, with all four Okapi searchers but only one ZPrise searcher claiming the system to be easy to learn. Although Okapi searchers expressed a high degree of satisfaction with the search outcome for this topic, the results for ZPrise as already noted were better.

4.3 Understanding.

For 71% of searches on Okapi, searchers divulged that they found the system easy to understand compared to 54% on ZPrise. The extent in which the relevance feedback process is made more visible in the Okapi system, could account for this difference in user perception.

5 Searching Behaviour.

Data on searching behaviour and the search process is presented in Appendix C.

5.1 Query Terms And Iterations.

As would be expected there was little difference in the number of initial query terms used in either systems but the mean for Okapi in Trec5 was almost double that of Trec6, 7.06 as opposed to 3.16. Equally there is little difference between the final query terms for Okapi (4.84) and ZPrise (3.79). The modification in the incremental query expansion facility in Trec6 would account for the fewer terms in the final query compared to the 18.21 mean in Trec5.

User defined terms for all iterations were also comparable for both systems, 6 for Okapi and 6.92 for ZPrise. Because of the incremental query expansion, Okapi searchers introduced less than half the number of terms in the query after the first iteration than in ZPrise (2.33, 5.54). They also removed query terms. The current implementation of incremental query expansion led to substantially fewer terms being removed than in Trec5, 2.54 compared to 25.22.

Although ZPrise searchers did not appear to have generated more query terms, they did undertake slightly more iterations than in Okapi, 4.67 compared to 3.38.

5.2 Documents ‘Viewed’ And ‘Seen’.

Searchers in Okapi demonstrated similar behaviour in scrolling through hitlists as in Trec5, covering a mean of 51.92 items compared to 60.58 previously. They equally examined the same number of full records, (14.15, 14.13) but fewer than in ZPrise (20.96).

TREC	Minimum number of relevant documents	RSV Must Not Be Less Than two-thirds the average RSV of
5	2	the terms in the last working query formulation
6	3	all terms that have been in the working query

Figure 5: TREC 5 and 6 Incremental Query Expansion Conditions

5.3 Relevance Judgements.

Although ZPrise searchers examined more full records on average, they made fewer positive relevance judgements 4.67 as opposed to 5.83 in Okapi. Relevance judgements in Okapi were also predominantly made on the full record (4.58) rather than on the best passage only (1.25).

6 Summary Of Results And Conclusions.

6.1 System Performance.

Okapi with its relevance feedback and query expansion facility clearly favoured recall. But since only one searcher overode the facility as a last resort in one search only, there is no evidence to show that this option could affect search performance. Searchers tended to rely on the system and intervened in a minimalist way. No diagnostic analysis has yet been undertaken to determine the effectiveness of the incremental query expansion.

The conditions specified in Appendix A.3.2 (2b) for members of the set of relevance feedback (RF) terms to be included in the set of candidate terms for the working query were modified from those used in TREC 5. The conditions, summarised in figure 5, were altered because the TREC 5 conditions often produced large sets of candidate terms (possibly several hundred) which sometimes resulted in:

1. apparently (from the user's viewpoint) erratic changes to the working query, and/or
2. the user removing large numbers of terms from the working query while searching for suitable RF terms.

The modified threshold conditions resulted in a reduced number of candidate terms, thereby making fewer and less marked changes to the working query. However, if we compare the number of user-entered and RF terms used in queries (figure 6) we see that, with the exception of three searches – (p24, t303i), (p11, t326i) and (p23, t326i) – very few RF terms were used by searchers. In 15 of the searches none were used at all. This may have been because user's saw them but removed them, or, more likely, few RF terms were determined by Okapi as suitable candidate terms. Clearly we need to carry out further experiments to:

1. establish more suitable conditions to provide greater recall and precision.
2. determine the usefulness of relevance feedback in a task of this nature.

6.2 Searcher Characteristics.

The sample of searchers appeared to be homogeneous and fairly representative of end-users. They all had a degree of experience with Web based systems, basic general knowledge and some awareness of advanced retrieval principles. The test systems were sufficiently different so that no individual subject had an unfair advantage for searching.

6.3 The Topics.

The sample of topics also seemed to have been evenly distributed between difficult, marginally difficult and easy searches. Topic 303 'Hubble Telescope Achievements' ranked first in terms of recall in both systems, and in terms of precision topic 326 'Ferry sinkings' in ZPrise and topic 339 'Alzheimer drug Treatment' in Okapi ranked first. There is no direct evidence to show that the number of aspects associated with any one topic determined the level

Topic	Searcher	User	RF	Topic	Searcher	User	RF
303	p12	2	0	326	p11	6	9
	p14	8	0		p13	3	4
	p22	6	0		p21	1	3
	p24	5	7		p23	3	15
307	p11	4	1	339	p12	3	0
	p13	2	0		p14	8	0
	p21	2	1		p22	4	0
	p23	2	0		p24	8	0
322	p11	10	1	347	p12	2	0
	p13	7	0		p14	5	3
	p21	10	0		p22	4	0
	p23	6	0		p24	6	0

Figure 6: Okapi: Number of User-Entered and RF Terms By Topic/Searcher

of difficulty. Whilst searchers could largely predict the performance of the easiest and most difficult topics, i.e. 322 'International Art Crime' they were more uncertain about those in between. On the whole there appeared to be no strong correspondence between searchers' perception about the topics and the system and the actual results.

6.4 Searchers' System Preferences.

Searchers found more searches to be easier on ZPrise but overall both systems were considered to be equally easy to use. Nevertheless searchers were more satisfied and confident about search outcomes with Okapi. The apparent discrepancies and inconsistencies between users' perception of the system and the searches would seem to indicate that searchers can differentiate between the procedural and more conceptual aspects of using a system or carrying out a search. However, there is clearly a tension between the two. It could be argued that the more interactive environment in Okapi offered the user the possibility for greater control in the searching process. Even though there was more to learn, the functionality was more apparent and made learnability easier. However, when asked which system they preferred to use, five out of the eight subjects opted for ZPrise.

References

[1] Robertson S E and Sparck Jones K. Relevance weighting of search terms. *Journal of the American Society for Information Science 27* May–June 1976 p129–146.

[2] Robertson S E. On term selection for query expansion. *Journal of Documentation 46* Dec 1990 p359–364.

[3] Robertson S E *et al.* Okapi at TREC–2. In: [6]. p21–34.

[4] Robertson S E *et al.* Okapi at TREC–3. In: [7]. p109–126.

[5] Robertson S E *et al.* Okapi at TREC–5. In: [9].

[6] *The Second Text REtrieval Conference (TREC–2).* Edited by D K Harman. Gaithersburg, MD: NIST, 1994.

[7] *Overview of the Third Text REtrieval Conference (TREC–3).* Edited by D K Harman. Gaithersburg, MD: NIST, April 1995.

[8] *Overview of the Fourth Text REtrieval Conference (TREC–4).* Edited by D K Harman. Gaithersburg, MD: NIST, November 1996.

[9] [TREC–5 proceedings.] NIST. To be published 1997.

A Okapi Interactive System Description

A.1 Introduction

A.1.1 The Okapi Interactive Interface.

The Okapi interface is an adaptation of that used in TREC 5 [5, Appendix A]. Figures 7 and 8 show screen dumps of the running system. The two major differences implemented for TREC 6 were:

1. A modification to the incremental query expansion conditions (see Section A.3.2).
2. Additional query manipulation facilities in the second round of searches (see Section A.4).

A.2 The Structure Of The GUI.

The interface is composed of: (1) a main window (figure 7) divided into six areas, and (2) a pop-up window in which full documents are displayed.

A.2.1 Main Window

1. **Query Entry Box**

   A text entry widget for user input of query terms. See Section A.3.1.

2. **Working Query**

   A scrollable, ranked list of the terms in the current working query. See Section A.3.2.

3. **Removed Terms**

   A scrollable list of any terms removed by the user from the working query, displayed in removal order. See Section A.3.7

4. **Hitlist**

   A scrollable, ranked hitlist for the current iteration. See Section A.3.4.

5. **Pool of Positive Relevance Judgments**

   A scrollable, ranked list of positive user relevance judgments. See Section A.3.6

6. **Function Buttons**

   At the bottom of the window are either two or three context-sensitive buttons.

   Search See Section A.3.3.

   Exit See Section A.3.9.

   Query Options (Round 2 only). See Section A.4.

A.2.2 Full Document

A pop-up text window for the display of a full record selected from the hitlist. At the bottom of the window are three buttons for making relevance judgments. These are described in Section A.3.5.

A.3 User Interaction

A.3.1 User Input Of Query Terms

Users may enter one or more words into the query entry box followed by an optional "phrase" operator (a '+' sign) as the last non-space character in the line. All the terms entered will be stemmed and looked-up in the database; stopwords and non-indexed terms will be discarded. If the operator is:

1. **None**: Each non-stopped index term will be a single term in the query.

2. **Phrase (+)**: If all of the non-stopwords are index terms an Okapi "phrase" (which constitutes a single term) will be formed as follows.

   (a) Two sets will be formed:

   A: A phrase in term input order, possibly with intervening stopwords: n(A) postings, weight w(A).

   S: A "within same sentence" occurrence of the terms in any order: n(S) postings, weight w(S).

   (b) These are combined into one set — a single query term — using the bestmatch operator BM25 ([4, Section 3.2]) where appropriate, according to the following rules.

Sets Generated	Sets Used	Displayed Operator
1. n(A) = n(S) = 0	Both discarded	
2. n(A) = 0, n(S) > 0	S(S), n(S), w(S)	(S)
3. n(A) > 0, n(A) = n(S)	S(A), n(A), w(A)	None
4. 0 < n(A) < n(S)	S(A), n(A), w(A) and S(S)-S(A), n(S)-n(A), w(S)	(B)

The weight calculated for each term is a Robertson/Sparck-Jones F4 predictive weight, with halves [1]. The weighting function allows each user entered term to be assigned a "loaded" weight by assuming the existence of a set of "mythical rels" of which a fixed number contain the term. The number of "mythical rels" is called "bigrload" of which "rload" contain the term. The values used for "rload" and "bigrload" were 4 and 5 respectively.

A.3.2 The Working Query

The entire set of user-entered and system-generated terms (the "termset") will be referred to by the letter Q. The number of members of Q is n(Q). Using similar terminology there will, at any time during a search, exist the following subsets of Q.

U: n(U) user-entered terms.

E: n(E) non-user terms extracted during relevance feedback.

C: n(C) candidate terms, those members of Q that satisfy the query threshold conditions.

W: n(W) members of the working query.

After each change to Q the working query will be re-generated in the following three stages.

1. Q is sorted by the two keys:

   (a) USER_TYPE: U (user) or E (extracted), descending.

   (b) RSV: (Robertson Selection Value [2]) descending

2. C is formed from Q by taking:

   (a) All members of U.

   (b) Members of E that occur in at least three relevant documents and have an RSV >= two-thirds the average RSV of all terms that have been in W.

3. W is formed from the top N members of C (N <= MAX_TERMS, a system defined limit), i.e. all user-entered terms plus and the top {MAX_TERMS - n(U)} members of E. If n(U) >= MAX_TERMS then W will be made up of members of U only. In both rounds MAX_TERMS defaulted to 20. In the second round users were able to increase its value to 30 or 40 if they wished, although no users did so.

NOTE: The conditions specified in A.3.2 (2b) were modified from those used in TREC 5. In TREC 5 the non-U members of C had to occur in at least two relevant documents and have an RSV >= two-thirds the average RSV of the terms in the *LAST WORKING QUERY FORMULATION*. These conditions (sometimes) produced large sets of candidate terms (possibly several hundred) which either caused some erratic behaviour (from the user's viewpoint) in changes to the working query, and/or resulted in the user removing large numbers of terms from W while searching for suitable query terms. The modified threshold conditions resulted in a reduced value of n(C), thereby making fewer and less marked changes to the working query.

Terms are displayed in the working query window in descending order of RSV.

A.3.3 Searching The Database.

Each search, performed by clicking the "Search" button, marks the next iteration. The members of W are combined using a best match operation (bm25). Passage retrieval [3, 4] is applied to the document set generated with parameters p_unit = 4, p_step = 2, k1 = 1.6 and b = 0.7. This will result in the system finding two passages for each document.

1. The full document: weight = w(F), length = l(F).
2. A sub-passage of the document: weight = w(P), length = l(P).

There are two cases to consider.

1. l(F) = l(P) :: The passage and the full document are the same and w(F) = w(P).
2. l(F) > l(P) :: The passage is distinct from the full document; the weight assigned to the document will be the greater of w(F) and w(P).

A.3.4 Hitlist Generation

Searching the database will result in a hitlist, generated from the ranked document set, displayed in the hitlist window — A.2.1(4). The hitlist is made up of the top H ranked documents (H <= 50). Each document must satisfy the following conditions.

1. It is not a member of the current set of relevant documents (i.e. it must not have already been seen in full by the searcher.
2. l(P) (or l(F) if there is no passage) must be less than DOC_THRESHOLD characters in length. DOC_THRESHOLD defaults to 10K.

An entry for each document consists of:

- A header line.

  ⟨record_no⟩ ⟨docid⟩ ⟨normalised_weight⟩[⟨passage_length⟩/] ⟨document_length⟩

  The ⟨normalised_weight⟩ is the system weight mapped onto the range 1..1000. ⟨document_length⟩ and ⟨passage_length⟩ are given to the nearest page, where a page is taken to be 2000 characters.
- A system generated title, made up from approximately the first 150 characters from the start of the document.
- Query term occurrence information

  A count of the occurrences of the stems of each query term within the document. The stem of a query term may occur in different source forms within the document. The first source in the document for each stem will be shown.

The hitlist entry at the top of the window is the first unseen document in the list.

A.3.5 Showing Documents

Double-clicking anywhere in a document's hitlist entry will display the full document in a pop-up, scrollable text window. Query terms are highlighted in green. The passage (if there is one) is highlighted in light grey. The line of the document displayed to the user at the top of the window is dependent upon the values of w(F|P) and l(F|P) (see Section A.3.3) i.e.

1. l(F) = l(P): Don't highlight the passage. The line of the document at the top of the window is the first line containing a query term.
2. l(F) > l(P): Highlight the passage. The line of the document at the top of the window is:
   (a) If w(P) >= w(F):: the first line of the passage.
   (b) If w(P) < w(F):: the first line containing a query term.

At the bottom of the text window are three buttons to allow users to make a relevance judgment.

1. **Full Document**: The document contains one or more aspects. Extract terms from the entire document.
2. **Passage Only**: The document contains one or more aspects. Extract terms from the highlighted passage only.
3. **Not Relevant**: The document does not contain any aspects.

Searchers must make a relevance judgment before they may go onto to any other part of the search process. For a document shown from the current hitlist a relevance judgment can be altered at any time until a new search is made. The change, (F -> P), (P -> F), ([F | P] -> N), or (N -> [F | P]) will modify the set of extracted terms appropriately.

A.3.6 Relevance Judgments Pool

The ranked hitlist information for all documents currently judged as relevant. Any member of the current relevance judgments pool that exists in a document set for a new iteration has its weight set to its value in the latest document set; the display order is adjusted accordingly.

A.3.7 Removing Terms

Any term may be removed from the working query by double clicking on its entry in the query window. Removed terms are displayed in the removed terms window (A.2.1(3)). If n(C) > MAX_TERMS (Section A.3.2), as terms are removed, other terms may be promoted to take their place.

A.3.8 Reinstating Removed Terms

A removed term may be reinstated in the query by double-clicking on its entry in the removed_terms window, although, as the working query changes, (i) its rank position may be different from that when it was removed, and (ii) in the case of extracted terms, it may not go in at all if n(C) >= MAX_TERMS.

A.3.9 Quitting

Quitting the search is achieved by clicking once on the "Exit" button.

A.4 Changes To The Interface For The Second Round Of Searches.

The functionality of the interface used in the first round of searches, with the modifications to the incremental expansion of the working query as described in Section A.3.2, was identical to that used in TREC 5. For the second round additional functionality was added as a direct result of observing users in Round 1 (a) attempting to change track during some searches (particularly 322i) and (b) obtaining during the course of a search a set of candidate terms (C) greater than 20 — the default value of MAX_TERMS — in size.

The additional functionality was implemented by installing a third function button (see Section A.2.1(6)) on the Main Window. Clicking on this button initiates a pop-up menu containing the entries:

1. Clear Relevance Feedback.
2. Clear Working Query.
3. Set Working Query Size

A.4.1 Clear Relevance Feedback

1. The set of terms added automatically to Q by the system after each positive relevance judgement are "removed" from Q (the entire set of user-entered and system-generated terms). The terms in fact remain in Q but their values of little r (the number of relevant documents they occur in) are set to zero and they are "flagged" to indicate that they should not be members of the set of candidate terms (C), unless subsequently:
   (a) re-input by the user, or
   (b) satisfy the threshold conditions based upon new relevance judgements (Section A.3.2)
2. The current set of relevance judgements would appear "greyed-out" in the relspool window (A.2.1(5)) to indicate that they were made before the relevance feedback was cleared. As before, these documents are not allowed to be members of any subsequent hitlist.

A.4.2 Clear Working Query

In addition to the "removal" of relevance feedback terms (Section A.4.1), all user-entered terms are similarly "removed".

A.4.3 Set Working Query Size

The default maximum size of the working query (MAX_TERMS) is 20 terms. In general, as a result of the modified threshold conditions for membership of the set of candidate terms (see Section A.3.2), it was found that n(C) did not often exceed this value. However, if n(C) is > 20 the only way the user can see the extra terms in the set is by removing one or more of the top 20 terms so that lower terms are promoted.

This option allows the user to increase the maximum working query size to 30 or 40 terms thus:

1. Enabling the user to see more than the top 20 terms from C without having to remove terms from W,
2. Allowing the user to increase the number of terms used to search the database if relevance feedback is producing enough good terms to warrant this.

During the second round of searches the additional functionality was seldom used (see Section C.7). The "clear working query" and "clear relevance feedback" facilities were only used a limited number of times, mostly on topic 322i which all users found difficult. The "Set Working Query" option was not used.

B Experimental Conditions

B.1 Searcher characteristics

Eight searchers, all MSc students in the Department Of Information Science, were divided into two groups: Group 1 and Group 2. Group 1 consisted of one male and three females; Group 2 consisted of three males and one female. Their ages ranged from late 20s to early 40s. The searchers were end-users who had:

1. No specialist knowledge of any of the subject domains covered by the topics, and
2. Only theoretical knowledge of relevance feedback IR systems.

B.2 Task Description / Training.

The task for each searcher consisted of:

1. Training.

   Each searcher conducted a tutorial session for each IR system separately from their search session. This was conducted in the week preceding the search session.

2. Search Session.

   (a) Filling in a pre-session questionnaire.

   (b) Reading a set of introductory instructions which outlined the nature of the task.

   (c) Performing six searches on the assigned topics (in the order 326i, 322i, 307i, 347i, 303i and 339i), three on Okapi (OK) and three on the control IR system, ZPrise (ZP). The searcher-system-topic combinations were:

Searcher	326i	322i	307i	347i	303i	339i
1	OK	OK	OK	ZP	ZP	ZP
2	ZP	ZP	ZP	OK	OK	OK
3	OK	OK	OK	ZP	ZP	ZP
4	ZP	ZP	ZP	OK	OK	OK

   After each search a post-search questionnaire was completed.

   (d) Filling in a post-session questionnaire.

   Each search was to take a maximum of 20 minutes which include reading time for the topic as well as actual searching time.

C Search Process

Unless otherwise stated the column 'Type' entries **N**, **+** and **A** refer to:

N terms defined with no adjacency operator
+ terms defined with an adjacency operator
A all terms defined

C.1 Clock Time

Times are given to the nearest tenth of a minute.

System	Mean	Median	Variance	Range
Okapi	19.8	20.5	3.20	15.3–21.8
ZPrise	20.3	20.7	2.61	16.1–22.5

C.2 Number of User Defined Terms

C.2.1 At All Iterations

System	Type	Mean	Median	Variance	Range
Okapi	N	4.04	3.0	13.17	0–16
	+	1.96	1.0	6.82	0– 9
	A	6.00	4.5	24.61	1–25
ZPrise		6.92	6.0	15.99	2-17

C.2.2 After The First Iteration

System	Type	Mean	Median	Variance	Range
Okapi	N	1.08	0.0	5.99	0–10
	+	1.25	0.0	5.15	0– 8
	A	2.33	0.5 5	14.84	0–17
ZPrise		5.54	5.5	20.26	0-15

C.2.3 "Phrases" Defined By Searchers (Okapi only)

Phrases generated:	47
Phrases used:	38

C.3 Number of Terms Used In Queries

C.3.1 Initial Query

System	Type	Mean	Median	Variance	Range
Okapi	N	2.58	3.0	3.21	0–8
	+	0.58	0.0	0.51	0–2
	A	3.16	3.0	2.67	2–8
ZPrise		3.58	3.0	3.21	2- 8

C.3.2 Final Query

System	Type	Mean	Median	Variance	Range
Okapi	N	3.63	3.0	13.03	0–13
	+	1.21	1.0	2.52	0– 7
	A	4.84	4.0	12.93	1–15
ZPrise		3.79	3.5	4.61	1-10

C.4 Number of Iterations

An iteration, i.e. a new query formulation, was taken to be marked by each 'search' command.

System	Mean	Median	Variance	Range
Okapi	3.38	2.0	7.11	1- 9
ZPrise	4.67	4.0	14.93	1-17

Okapi note: Expansion was performed incrementally. No data was kept to indicate how many times and when each working query was altered by the inclusion/exclusion of extracted terms.

C.5 Documents "Viewed" (hitlist) and "Seen" (full text)

C.5.1 Viewed

"Viewing" a document consisted of seeing its entry in the hitlist but not the full document. The figures, for Okapi only, represent the percentage distance scrolled through the hitlist by the searcher. There was no information available from the ZPrise log files.

System	Mean	Median	Variance	Range
Okapi	51.92	50.0	611.30	14–98

C.5.2 Seen.

"Seeing" a document consisted of showing the full record. In the hitlist for any iteration Okapi does not show documents that were included in the hitlist(s) of any previous iteration(s) (see Section A.3.4. This is not the case with ZPrise where the same document may have occurred in several hitlists. Two sets of figures are shown: (a) all shows, including repeated shows, and (b) distinct shows. All includes some documents that were shown more than once with a view to (potentially) changing the relevance judgement.

System		Mean	Median	Variance	Range
Okapi	Distinct	13.58	13.5	31.38	5-30
	All	14.25	14.5	29.33	7-30
ZPrise	Distinct	15.92	15.0	83.73	2-32
	All	20.96	20.0	165.43	3-52

C.6 Relevance Judgments

Documents may have been judged more than once during a session, either altering the judgement or leaving it unchanged. In an Okapi session this may only happen during a single iteration due to the conditions specified in Section A.3.4. In a ZPrise session, since individual documents may occur in several hitlists this may occur at any time during the session.

Judging a document relevant implies that it contains one or more aspects. The types of relevance judgement for Okapi (F, P and N) are described in A.3.5. ZPrise has two judgements: R, relevant, means that the document contains one or more aspects. U, undecided, means that the document does not contain any aspects.

Two sets of figures are given, one for the final judgement and one for the total judgements on each document.

1. Final judgement only.

System	Rel	Mean	Median	Variance	Range
Okapi	F	4.58	4.0	11.56	1-11
	P	1.25	1.0	2.37	0- 6
	N	7.96	6.0	23.78	3-24
	All	13.79	13.5	29.91	6-30
ZPrise	R	4.67	2.0	34.41	0-26
	U	0.08	0.0	0.08	0- 1
	All	4.75	2.0	33.67	1-26

2. All Judgements.

The figures for "All Judgements" include the following re-judgements of documents.

System	(F->N) or (R->U)	(N->F) or (U->R)	Unchanged	Total
Okapi	2	4	9	15
ZPrise	1	1	2	4

System	Rel	Mean	Median	Variance	Range
Okapi	F	4.67	4.0	11.80	1-11
	P	1.25	1.0	2.37	0- 6
	N	8.33	6.5	23.10	3-24
	All	14.25	14.5	29.33	7-30
ZPrise	R	4.79	2.5	34.95	0-26
	U	0.13	0.0	0.11	0- 1
	All	4.92	3.0	34.43	0-26

C.7 Use of System Features

The figures in parentheses for Okapi are the equivalent values for Trec 5.

System	Command		Mean	Median	Variance	Range
Okapi	Define	N	4.04 (4.96)	3.0 (3.5)	13.17 (21.58)	0-16 (0-16)
		+	1.96 (3.83)	1.0 (3.0)	6.82 (9.80)	0-9 (0-13)
		A	6.00 (8.79)	4.5 (8.0)	24.61 (24.96)	1-25 (2-18)
	Search		3.38 (3.63)	2.0 (3.0)	7.11 (1.98)	1- 9 (1- 7)
	Show		14.15 (14.13)	14.5 (13.5)	29.33 (15.42)	7-30 (9- 21)
	Remove		2.54 (25.22)	0.5 (13.5)	12.00 (1119.18)	0-11 (0-116)
	Restore		0.38 (0.85)	0.0 (0.0)	1.64 (3.40)	0- 6 (0- 8)
	Clear_rf		0.08	0.0	0.08	* 0- 1
	Clear_query		0.58	0.0	2.08	** 0- 5

* Topic 322i

* Topic 322i (6 times, 5 by one searcher)
Topic 339i once.

System	Command	Mean	Median	Variance	Range
ZPrise	Define	6.92	6.0	15.99	2-17
	Search	4.67	4.0	14.93	1-17
	Show	15.92	15.0	83.73	2-32

C.8 Number of User Errors

Data on user errors were not collected.

C.9 Topic 326 - Ferry Sinkings.

C.9.1 Search Narrative

The initial query terms entered were chosen from the topic specification itself. Eight documents were viewed from the initial hitlist but only four were deemed to meet the criteria set out by the narrative. One was chosen after subsequent re-examination. As some of the documents selected from the first hitlist referred to the sinking of the ferry 'Herald of Free Enterprise' without giving the number of deaths involved the term 'Herald of Free Enterprise' was added for the second iteration. Four of the five documents viewed from the second hitlist were judged to be relevant. The term 'Herald of Free Enterprise' was then removed and the term 'Manila ferry' added for a third iteration. This was in order to find other documents that dealt more specifically with the sinking of this ferry, but this failed to retrieve any new documents that were relevant.

As the query terms used in each search is reflected in the weighted list of documents retrieved the continued inclusion of the search term '100 deaths' may not have been particularly useful for the second and third iteration and on reflection should, perhaps, have been removed from the term set. In retrospect the use of the search phrase '100 deaths' may have contributed to some relevant documents having a lower ranking on the hitlist for all three iterations. The display of extracted query terms did not prove to be particularly useful and the inclusion of the search terms 'ferry' and 'sinks' as well as the search phrase 'ferry sinking' may not have been very helpful either. In all cases, but one, in making positive relevance judgements, the searcher chose to select the full document rather than just the highlighted passage. Due to the time constraints the searcher did not review the final list of documents selected as relevant in order to weed out those which may have duplicated reports of any particular ferry sinking.

The searcher was not very happy with the way in which relevance judgements had to be made that were, upon subsequent iterations, irretrievable. The searcher would have also preferred that on iterations involving a newly defined search term or phrase that there would have to be positive selection of existing query terms to be included in the new search instead of the automatic inclusion that did occur. The need for selection or de-selection of query terms required for each new iteration was found to be cumbersome.

The search itself was not particularly difficult. The searcher did experience some difficulty in introducing a search term that would help select reports of a ferry sinking that involved the loss of more than 100 lives and this may have led to some problems later in the search. It may have been better to simplify the initial search by restricting the query to 'ferry sinking' or 'ferry' and 'sinks'.

C.9.2 Breakdown of command usage.

Define (N)	0
Define (+)	4
Search	3
Show	16
Remove	1

D City Recall-Precision Results

D.1 Results Grouped By Round, Searcher And System.

D.1.1 Recall

Round	Searcher	System	Mean	Median	Variance	Range
1	p11	Okapi	0.315	0.391	0.032	0.111 – 0.444
	p12	Okapi	0.672	0.900	0.235	0.115 – 1.000
	p13	Okapi	0.353	0.391	0.051	0.111 – 0.556
	p14	Okapi	0.232	0.154	0.021	0.143 – 0.400
	p11	ZPrise	0.631	0.700	0.167	0.192 – 1.000
	p12	ZPrise	0.140	0.111	0.005	0.087 – 0.222
	p13	ZPrise	0.600	0.500	0.130	0.300 – 1.000
	p14	ZPrise	0.140	0.087	0.030	0.000 – 0.333
2	p21	Okapi	0.206	0.174	0.050	0.000 – 0.444
	p22	Okapi	0.544	0.400	0.163	0.231 – 1.000
	p23	Okapi	0.278	0.333	0.022	0.111 – 0.391
	p24	Okapi	0.600	0.700	0.035	0.385 – 0.714
	p21	ZPrise	0.464	0.600	0.115	0.077 – 0.714
	p22	ZPrise	0.169	0.174	0.003	0.111 – 0.222
	p23	ZPrise	0.618	0.700	0.184	0.154 – 1.000
	p24	ZPrise	0.287	0.304	0.078	0.000 – 0.556

D.1.2 Precision.

Round	Searcher	System	Mean	Median	Variance	Range
1	p11	Okapi	0.807	0.846	0.016	0.667 – 0.909
	p12	Okapi	1.000	1.000	0.000	1.000 – 1.000
	p13	Okapi	0.641	0.833	0.146	0.200 – 0.889
	p14	Okapi	0.578	0.400	0.135	0.333 – 1.000
	p11	ZPrise	0.861	0.833	0.016	0.750 – 1.000
	p12	ZPrise	1.000	1.000	0.000	1.000 – 1.000
	p13	ZPrise	0.826	0.812	0.028	0.667 – 1.000
	p14	ZPrise	0.556	0.667	0.259	0.000 – 1.000
2	p21	Okapi	0.611	0.833	0.287	0.000 – 1.000
	p22	Okapi	0.861	0.833	0.016	0.750 – 1.000
	p23	Okapi	0.675	0.824	0.177	0.200 – 1.000
	p24	Okapi	0.472	0.500	0.044	0.250 – 0.667
	p21	ZPrise	1.000	1.000	0.000	1.000 – 1.000
	p22	ZPrise	0.933	1.000	0.013	0.800 – 1.000
	p23	ZPrise	0.722	0.833	0.037	0.500 – 0.833
	p24	ZPrise	0.573	0.800	0.250	0.000 – 0.920

D.2 Results Grouped by “System”, “Round” and “Topic”

D.2.1 Recall.

System	Round	Topic	Mean	Median	Variance	Range
Okapi	1	303	0.572	0.572	0.367	0.143 – 1.000
		307	0.391	0.391	0.000	0.391 – 0.391
		322	0.111	0.111	0.000	0.111 – 0.111
		326	0.500	0.500	0.006	0.444 – 0.556
		339	0.650	0.650	0.125	0.400 – 0.900
		347	0.135	0.135	0.001	0.115 – 0.154
	2	303	0.857	0.857	0.041	0.714 – 1.000
		307	0.283	0.283	0.024	0.174 – 0.391
		322	0.056	0.056	0.006	0.000 – 0.111
		326	0.389	0.389	0.006	0.333 – 0.444
		339	0.550	0.550	0.045	0.400 – 0.700
		347	0.308	0.308	0.012	0.231 – 0.385
ZPrise	1	303	1.000	1.000	0.000	1.000 – 1.000
		307	0.087	0.087	0.000	0.087 – 0.087
		322	0.056	0.056	0.006	0.000 – 0.111
		326	0.278	0.278	0.006	0.222 – 0.333
		339	0.500	0.500	0.080	0.300 – 0.700
		347	0.346	0.346	0.047	0.192 – 0.500
	2	303	0.857	0.857	0.041	0.714 – 1.000
		307	0.239	0.239	0.008	0.174 – 0.304
		322	0.056	0.056	0.006	0.000 – 0.111
		326	0.389	0.389	0.056	0.222 – 0.556
		339	0.650	0.650	0.005	0.600 – 0.700
		347	0.116	0.116	0.003	0.077 – 0.154

D.2.2 Precision.

System	Round	Topic	Mean	Median	Variance	Range
Okapi	1	303	0.667	0.667	0.222	0.333 – 1.000
		307	0.840	0.840	0.000	0.833 – 0.846
		322	0.434	0.434	0.109	0.200 – 0.667
		326	0.899	0.899	0.000	0.889 – 0.909
		339	1.000	1.000	0.000	1.000 – 1.000
		347	0.700	0.700	0.180	0.400 – 1.000
	2	303	0.500	0.500	0.125	0.250 – 0.750
		307	0.912	0.912	0.015	0.824 – 1.000
		322	0.100	0.100	0.020	0.000 – 0.200
		326	0.917	0.917	0.014	0.833 – 1.000
		339	0.834	0.834	0.055	0.667 – 1.000
		347	0.667	0.667	0.055	0.500 – 0.833

System	Round	Topic	Mean	Median	Variance	Range
ZPrise	1	303	1.000	1.000	0.000	1.000 – 1.000
		307	0.834	0.834	0.055	0.667 – 1.000
		322	0.500	0.500	0.500	0.000 – 1.000
		326	1.000	1.000	0.000	1.000 – 1.000
		339	0.750	0.750	0.014	0.667 – 0.833
		347	0.781	0.781	0.002	0.750 – 0.812
	2	303	0.750	0.750	0.125	0.500 – 1.000
		307	0.800	0.800	0.000	0.800 – 0.800
		322	0.500	0.500	0.500	0.000 – 1.000
		326	0.960	0.960	0.003	0.920 – 1.000
		339	0.917	0.917	0.014	0.833 – 1.000
		347	0.917	0.917	0.014	0.833 – 1.000

E Search Evaluation: Questionnaire Results

E.1 Pre-Session Questionnaire.

1. What is the highest degree you have or expect to obtain?

Degree:	Bachelors	Masters	PhD
	0	8	0

2. What is your age?

Range:	<21	21-30	31-40	41-50	51-60	¿60
	0	5	3	0	0	0

3. What is your gender (M = male, F = female)?

Gender:	M	F
	4	4

4. Have you participated in previous TREC searching studies (Y/N)?

Answer:	Y	N
	0	8

5. Have you experience of using Okapi or ZPrise?

System:	Okapi	ZPrise
	1	0

6. How many years have you been doing online searching?

Years:	0	1	2	3	4	5	6	7
	1	3	1	2	1	0	0	0

7. How much experience have you had of the following (1=none, 3=some, 5=a great deal)

	1	2	3	4	5
Searching on computerised library catalogues	1	1	3	3	0
Searching on CD ROM systems	3	1	3	1	0
(e.g. Infotrac, Grolier) Searching on commercial online systems (e.g. Dialog, Lexis, BRS Afterdark)	3	1	2	1	1
Searching on WWW browsers (e.g. Netscape)	0	0	0	6	2
Searching on other systems	7	1	0	0	0
Searching full-text databases	3	2	1	2	0
Searching in ranked output IR systems	2	3	2	1	0
Searching in IR systems that provide relevance feedback	3	3	2	0	0
Using a mouse-based interface	0	0	2	2	4

E.2 Exit Questionnaire.

Rate the following on a scale of 1-5 (1=Not at all, 5=Completely)

1. To what extent were you able to understand the nature of the task?

Rating:	1	2	3	4	5
	0	0	2	5	1

2. To what extent did you find this task similar to other searching tasks that you typically perform?

Rating:	1	2	3	4	5
	0	2	5	1	0

3. How different did you find the following systems from each other?

Rating:	1	2	3	4	5
	0	0	4	4	0

4. Please rank the systems (1 = easiest).

	Okapi	ZPrise
Easiest to use:	4	4
Easiest to learn to use:	4	4
Liked best:	3	5

E.3 Post-Search Questionnaires.

The following questions were asked after each search session (i.e. after performing a search on one topic on one system).

1. How familiar are you with the topic?
2. How difficult was it to do the search?
3. How satisfied are you with your search results?
4. How confident are you that you identified all of the possible aspects for this topic?
5. Did you have enough time to do an effective search?
6. How easy was it to use this IR system?
7. How easy was it to learn to use this IR system?
8. How well did you understand how to use this IR system?

E.3.1 All Answers. (Percentages out of 48)

	1	2	3	4	5
1	60	19	19	2	0
2	25	23	29	19	4
3	19	15	23	38	5
4	25	15	29	27	4
5	17	23	25	23	12
6	0	8	31	46	15
7	0	2	38	42	18
8	0	0	38	60	2

E.3.2 Answers By System (Percentages out of 24).

System	Qu	1	2	3	4	5
Okapi	1	70	8	22	0	0
	2	25	13	36	22	4
	3	21	13	16	42	8
	4	25	13	29	29	4
	5	17	21	17	21	24
	6	0	8	29	46	17
	7	0	0	25	58	17
	8	0	0	29	67	4

System	Qu	1	2	3	4	5
ZPrise	1	50	29	17	4	0
	2	25	33	21	17	4
	3	17	17	29	33	4
	4	25	17	29	25	4
	5	17	25	33	25	0
	6	0	8	33	46	13
	7	0	4	50	25	21
	8	0	0	46	54	0

E.3.3 Answers By Topic (Percentages out of 8).

Topic	Qu	1	2	3	4	5	Topic	Qu	1	2	3	4	5
303	1	50.0	37.5	12.5	0.0	0.0	326	1	62.5	25.0	12.5	0.0	0.0
	2	12.5	12.5	62.5	12.5	0.0		2	37.5	25.0	37.5	0.0	0.0
	3	12.5	12.5	37.5	37.5	0.0		3	12.5	0.0	37.5	37.5	12.5
	4	12.5	25.0	12.5	50.0	0.0		4	12.5	12.5	50.0	25.0	0.0
	5	12.5	0.0	37.5	37.5	12.5		5	12.5	12.5	50.0	12.5	12.5
	6	0.0	0.0	62.5	25.0	12.5		6	0.0	0.0	25.0	50.0	25.0
	7	0.0	0.0	37.5	50.0	12.5		7	0.0	0.0	25.0	50.0	25.0
	8	0.0	0.0	50.0	50.0	0.0		8	0.0	0.0	25.0	62.5	12.5
307	1	87.5	0.0	12.5	0.0	0.0	339	1	37.5	12.5	37.5	12.5	0.0
	2	37.5	25.0	12.5	25.0	0.0		2	25.0	37.5	37.5	0.0	0.0
	3	0.0	12.5	37.5	50.0	0.0		3	0.0	12.5	12.5	62.5	12.5
	4	12.5	25.0	37.5	12.5	12.5		4	0.0	0.0	62.5	37.5	0.0
	5	12.5	37.5	12.5	25.0	12.5		5	0.0	25.0	37.5	25.0	12.5
	6	0.0	0.0	25.0	62.5	12.5		6	0.0	0.0	37.5	50.0	12.5
	7	0.0	0.0	37.5	37.5	25.0		7	0.0	0.0	37.5	50.0	12.5
	8	0.0	0.0	25.0	75.0	0.0		8	0.0	0.0	37.5	62.5	0.0
322	1	87.5	12.5	0.0	0.0	0.0	347	1	37.5	25.0	37.5	0.0	0.0
	2	12.5	0.0	12.5	50.0	25.0		2	25.0	37.5	12.5	25.0	0.0
	3	62.5	25.0	12.5	0.0	0.0		3	25.0	25.0	0.0	37.5	12.5
	4	75.0	12.5	0.0	12.5	0.0		4	37.5	12.5	12.5	25.0	12.5
	5	25.0	37.5	12.5	12.5	12.5		5	37.5	25.0	0.0	25.0	12.5
	6	0.0	25.0	12.5	50.0	12.5		6	0.0	25.0	25.0	37.5	12.5
	7	0.0	12.5	37.5	25.0	25.0		7	0.0	0.0	50.0	37.5	12.5
	8	0.0	0.0	37.5	62.5	0.0		8	0.0	0.0	50.0	50.0	0.0

E.3.4 Answers By System–Topic (Percentages out of 4).

System	Topic	Qu	1	2	3	4	5	Topic	Qu	1	2	3	4	5
Okapi	303	1	75	25	0	0	0	326	1	75	0	25	0	0
		2	25	0	50	25	0		2	0	25	75	0	0
		3	0	0	50	50	0		3	25	0	50	25	0
		4	0	25	0	75	0		4	25	25	50	0	0
		5	0	0	25	50	25		5	25	25	25	0	25
		6	0	0	50	25	25		6	0	0	50	25	25
		7	0	0	0	75	25		7	0	0	25	50	25
		8	0	0	50	50	0		8	0	0	25	50	25
Okapi	307	1	100	0	0	0	0	339	1	50	0	50	0	0
		2	50	25	25	0	0		2	25	0	75	0	0
		3	0	25	0	75	0		3	0	25	0	50	25
		4	0	25	75	0	0		4	0	0	25	75	0
		5	0	50	0	25	25		5	0	0	50	25	25
		6	0	0	25	75	0		6	0	0	25	50	25
		7	0	0	25	75	0		7	0	0	25	50	25
		8	0	0	0	100	0		8	0	0	25	75	0

System	Topic	Qu	1	2	3	4	5	Topic	Qu	1	2	3	4	5
Okapi	322	1	100	0	0	0	0	347	1	25	25	50	0	0
		2	0	0	0	75	25		2	50	25	0	25	0
		3	75	25	0	0	0		3	25	0	0	50	25
		4	100	0	0	0	0		4	25	0	25	25	25
		5	50	25	0	0	25		5	25	25	0	25	25
		6	0	25	25	50	0		6	0	25	0	50	25
		7	0	0	50	50	0		7	0	0	25	50	25
		8	0	0	25	75	0		8	0	0	50	50	0
ZPrise	303	1	25	50	25	0	0	326	1	50	50	0	0	0
		2	0	25	75	0	0		2	75	25	0	0	0
		3	25	25	25	25	0		3	0	0	25	50	25
		4	25	25	25	25	0		4	0	0	50	50	0
		5	25	0	50	25	0		5	0	0	75	25	0
		6	0	0	75	25	0		6	0	0	0	75	25
		7	0	0	75	25	0		7	0	0	25	50	25
		8	0	0	50	50	0		8	0	0	25	75	0
ZPrise	307	1	75	0	25	0	0	339	1	25	25	25	25	0
		2	25	25	0	50	0		2	25	75	0	0	0
		3	0	0	75	25	0		3	0	0	25	75	0
		4	25	25	0	25	25		4	0	0	100	0	0
		5	25	25	25	25	0		5	0	50	25	25	0
		6	0	0	25	50	25		6	0	0	50	50	0
		7	0	0	50	0	50		7	0	0	50	50	0
		8	0	0	50	50	0		8	0	0	50	50	0
ZPrise	322	1	75	25	0	0	0	347	1	50	25	25	0	0
		2	25	0	25	25	25		2	0	50	25	25	0
		3	50	25	25	0	0		3	25	50	0	25	0
		4	50	25	0	25	0		4	50	25	0	25	0
		5	0	50	25	25	0		5	50	25	0	25	0
		6	0	25	0	50	25		6	0	25	50	25	0
		7	0	25	25	0	50		7	0	0	75	25	0
		8	0	0	50	50	0		8	0	0	50	50	0

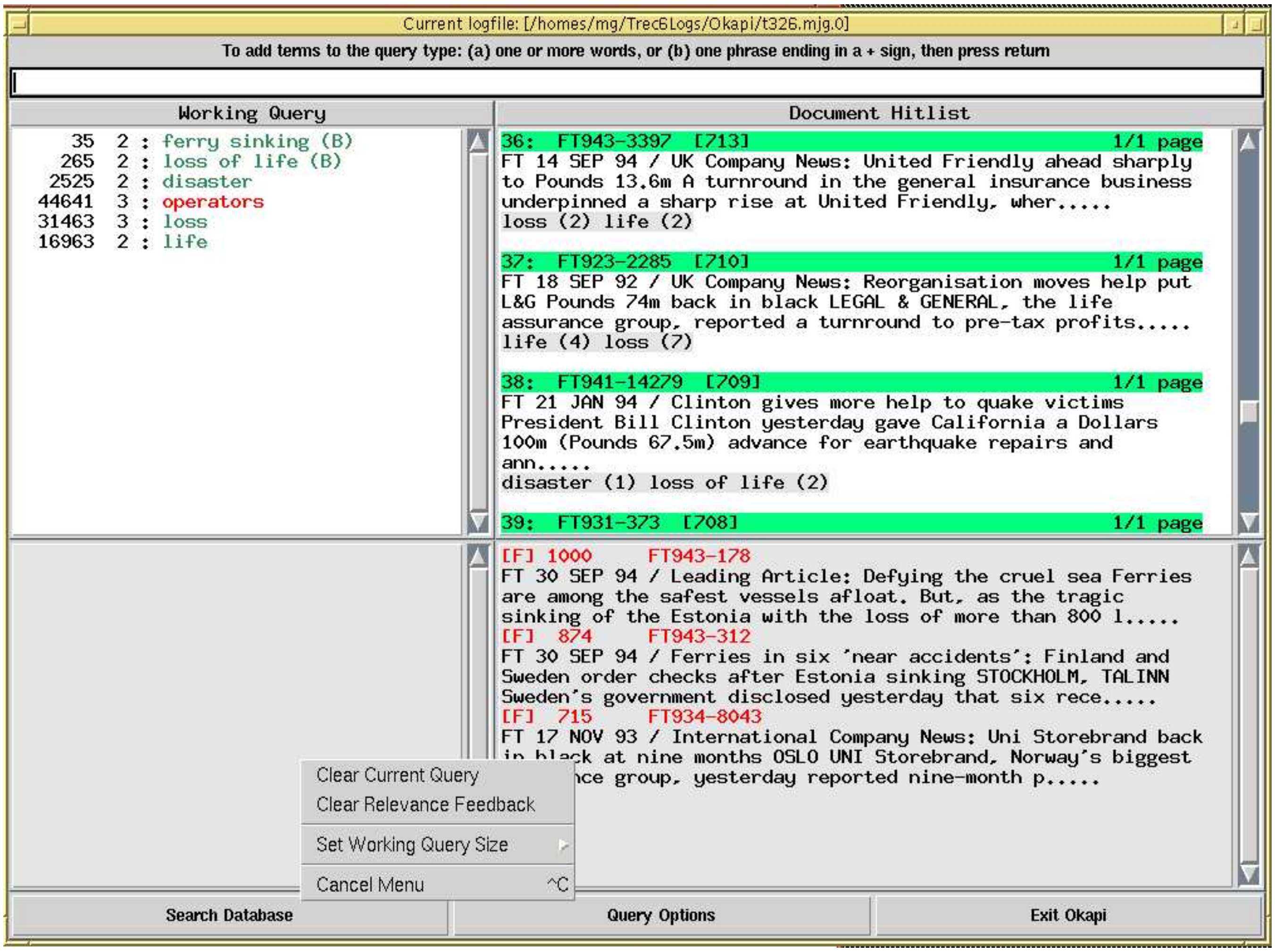
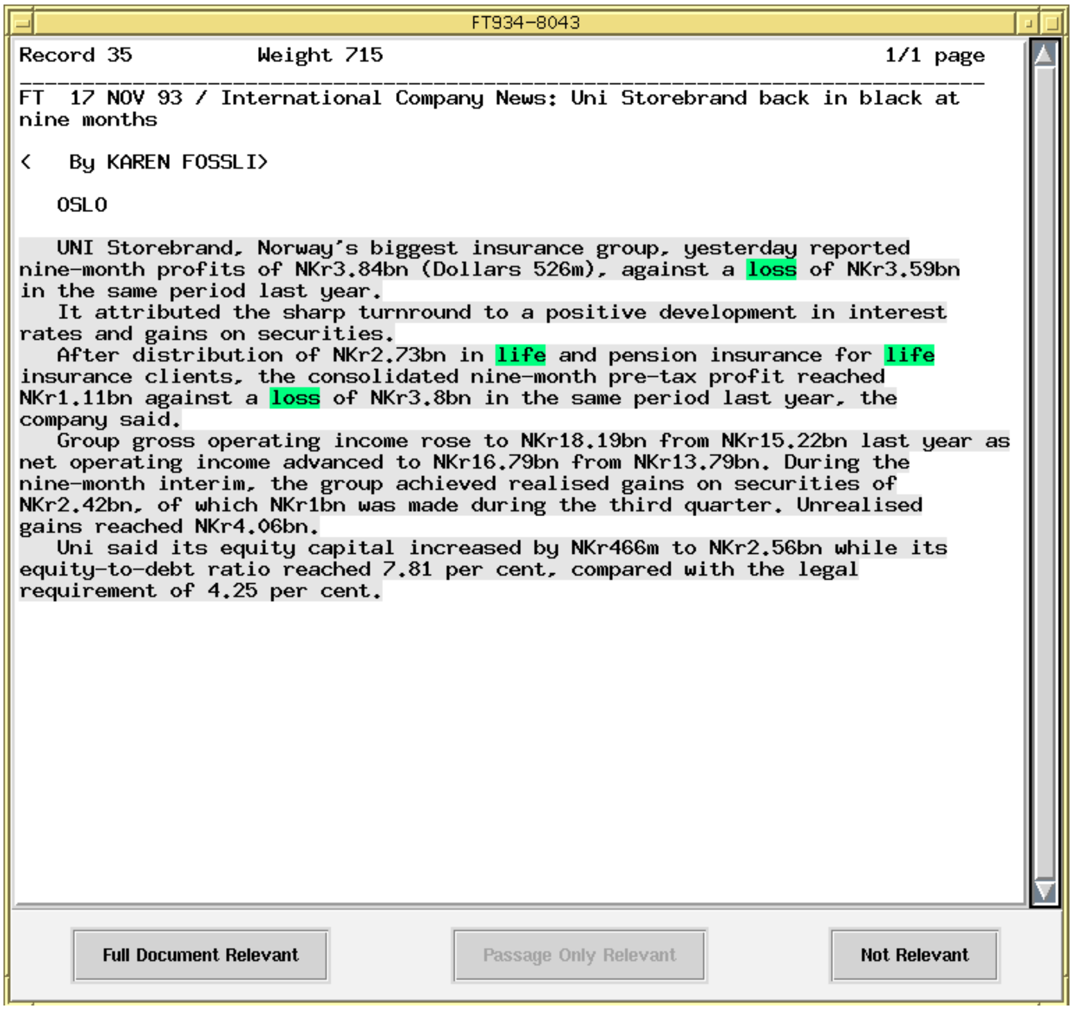


Figure 7: Interactive interface: Main Search Screen



FT934-8043

Record 35 Weight 715 1/1 page

FT 17 NOV 93 / International Company News: Uni Storebrand back in black at nine months

< By KAREN FOSSLI>

OSLO

UNI Storebrand, Norway's biggest insurance group, yesterday reported nine-month profits of NKr3.84bn (Dollars 526m), against a loss of NKr3.59bn in the same period last year.

It attributed the sharp turnround to a positive development in interest rates and gains on securities.

After distribution of NKr2.73bn in life and pension insurance for life insurance clients, the consolidated nine-month pre-tax profit reached NKr1.11bn against a loss of NKr3.8bn in the same period last year, the company said.

Group gross operating income rose to NKr18.19bn from NKr15.22bn last year as net operating income advanced to NKr16.79bn from NKr13.79bn. During the nine-month interim, the group achieved realised gains on securities of NKr2.42bn, of which NKr1bn was made during the third quarter. Unrealised gains reached NKr4.06bn.

Uni said its equity capital increased by NKr466m to NKr2.56bn while its equity-to-debt ratio reached 7.81 per cent, compared with the legal requirement of 4.25 per cent.

Figure 8: Interactive interface: Full Record Display