

**Using Lines from the Body of Document and Key Words
for Improving Display of Search Results
in Information Retrieval Systems**

[Offer Drori](#)

**Hebrew University of Jerusalem
SHAAM - Information Systems**

P.O. Box 10414 Jerusalem

ISRAEL 91103

Tel. (972)-2-5688439

Fax. (972)-2-5688681

E-mail: [offerd@ {cs.huji.ac.il, shaam.gov.il}](mailto:offerd@{cs.huji.ac.il, shaam.gov.il})

**Acknowledgements: I would like to thank Eliezer Lozinskii for his good suggestions
and to Avner Ben Hanoach for his help.**

ABSTRACT

This article displays the results of a comparative study between various user interfaces whose purpose is finding search results from textual databases. This study is one of a series which were carried out at the Hebrew University in Jerusalem where this study focuses on the affect of the display of elements from textual documents which were discovered as a result of the search activity and the effect on the behavior of the users of the system. The findings of the study show a clear advantage for the interface developed for the experiment which includes display of lines which fulfill the search criteria in combination with key words from the body of the document. The advantage of the interface is expressed in everything relating to ease for the user, his feeling of confidence during performing the task and his feeling with respect to the relevance of the information displayed for performing his task.

Keywords: Information display, search results, task relevancy

1. INTRODUCTION

1.1 Problem and Background

The operation of information retrieval in a text retrieval system includes, in most cases, a definition of the search strategy where the result of the search engine includes a list of existing documents which fulfils the search strategy. From the character of the search operation it is clear that in most of the cases the answer received will include a list of documents, the larger the database the number of documents fulfilling the search criteria will increase [4] [18]. A good example of a large unfocused information database is the Internet as it exists today. In order to find information on the Internet a number of search engines have been developed which perform, each by its own method, collection of material, from developing the words in the database and performing searches about them [2] [15]. Most of the search engines resemble one another in the way they display their information for the user. A list of documents fulfilling search criteria is displayed in a serial manner (while trying partially to grade the list) and in most cases includes the document title, its address on the Internet and often a number of first lines of the document, the date of revision, etc. In search engines existing today on the Internet most of the search operation produce lists of hundreds and thousands of documents while the average user looks at the first 10-20 results only [9]. In this situation, the information user may miss in many cases the information important to him and if he has strong motivation he must invest a large amount of time reading many documents which fulfill the search criteria but are not relevant to his search question.

The purpose of this paper is to deal with the display of the list of search results both in a short review of work which has been done in the past and focus on the results of a comparative study which examined various interfaces for performing the operations described above.

1.2 Attempts at a Solution in Previous Studies

There are two accepted methods today of displaying a list of search results from textual databases. One is a display of the list of document titles fulfilling the search criteria and the second a display of a number of the first sentences of the documents in addition to the document titles. The use of document titles assumes that the document title can show the contents of the document in the best way. Studies show that in reality this is not so. Peritz showed in her study that in journals in the field of library science and information, 21% of the cases included the article title which was not informative [11]. The idea of adding the first lines of the body of the text to the displayed title is intended to deal with this problem and put in the hands of the user additional information which could suggest the contents of the document and thereby permitting the user to make a decision regarding the relevance of the document. In various studies carried out in the field, additional methods were examined for displaying search results from textual databases. Display of the results in a number of hierarchical levels was carried out in the study of [8] using Venn diagrams. The findings of the study are that the browsers access is found quicker than with the Venn diagram. In Project Superbook the interface is designed as an electronic book. Three vertical windows are displayed in this interface, one window

providing the table of contents, the second a search definition and the third the document texts [6]. With the Scatter/Gather Browsing method 10 windows of information are displayed in parallel. The windows themselves constitute a cluster of collections in which there are a number of documents having common significant words. Under the words are displayed the titles of the first three documents. The user is given the possibility of broadening the information (scatter) or to reduce information (gather) by means of appropriate commands [12] [13]. This system has the advantage in everything relating to browsing a large unfamiliar database of documents. In the Butterfly user interface, results are displayed in a three dimensional manner where at the head of the butterfly items are displayed which identify the document, on the left wing are displayed the citations of the article and on the right wing a list of those citing the article [10]. In the system called TileBars, the user interface allow receiving simultaneously the relevant degree of the documents retrieved by a graphical display of a number of squares for each document and their color [7]. Another manner of meeting the challenge is by displaying a list of documents and especially a list of words or subjects requested which were implemented in a number of systems called incremental search. With this method the more the user presses an additional number of letters the system displays immediately the list of documents fulfilling the momentary search. Raskin in his study on the improvement in the use of the cursor maintains that in the incremental search a saving in time during search was obtained since the system continues in the search process while the user looks at the results obtained up to then [14].

Together with the methods mentioned above and on the assumption that human knowledge in the main is textual knowledge this writer has designed and investigated a third method of displaying a lists of documents from databases. The method includes the display of a number of lines within the relevant paragraph fulfilling the search criteria together with the titles of the documents [3]. The point of departure in this method is that most databases today do not undergo the a process of keying in advance and thus most of the documents are not scientific documents which have undergone an accepted control process. In the light of these assumption the information which should be passed on to the user is the most relevant which is possible to find within the document based on the search criteria in the search question and not on the basis of a fixed place in the document. For these reasons it may be supposed that the most relevant location in the body of the document is the paragraph or lines which fulfill the search criteria and which cause the document to enter the list of answers as a result of the search question. A study of this subject carried out at the Hebrew University in Jerusalem found that an advantage exists in the display of lines fulfilling the search criteria from the body of the document over the first lines of the document [5]. In this study in addition to the title of the document and the relevant lines which fulfill the search criteria are added the key words in the document. A description of the study which examined this assumption and its results is the principal subject of this article.

1.3 Questions in this Study

The following questions of the study discussed in this paper are:

1. What is the most important information to display in the list of search results?
2. In comparison with various methods of displaying information in a list of search results which is the preferred method according to the user.

2. METHODOLOGY

2.1 Experiment and Sample Environment

The experiment was carried out testing a number of defined problems in an academic environment using a sample of 51 end users. Three questions required the type of answers from the participant which required performing search tasks in a database in the English language built for the experiment. The search tasks used three types of user interfaces. All participants were experienced in using all the interfaces. Group members were similar (students studying for an MBA at the School of Business Administration and an MSc in information science). Their experience in working with computer systems and with search engines on the Internet were similar. The average age of the participants was 31 years.

2.2 Types of Information and Methods of Retrieval

Three types of interfaces were used where the study compared the manner of displaying search results which were obtained by searching in the textual database. Two of them were similar to the usual type of search engines on the Internet (without the addition of displaying keywords as carried out in this study): The first, method A provided a list of document titles plus key words. The second, method B provided a list of document titles plus first lines of document plus key words. The third, method C was developed specifically for the experiment included document titles, relevant lines fulfilling search criteria plus key words in the article. A number of subjects was compared with the various methods. Those examined were; time for performing the search tasks for each interface, feelings of users of the different methods as to ease of performing the work, satisfaction, feeling of confidence in each of the methods and the degree of help the method gives in obtaining relevant information to perform the task. Information gathered was based on two sources. Objective information, as to search speed and finding the correct answer, was measured by computer. Subjective information was obtained from a computer feedback sheet which all participants were required to fill out at the end of the experiment. The feedback sheet asked questions regarding feeling, satisfaction, etc. Table 1 provides a list of criteria examined in the experiment with their evaluation.

Table 1 List of Criteria and Their Values

Subject Tested	Table of Values for Answer
Feeling of ease in searching	1-10 1 - Not at all easy 10- Very easy
Want the search engine interface to be of a certain method	1-10 1- Do not want at all 10- Want very much
Feeling of confidence during solution of question	1-4 1 - Lack of confidence 10 - Very confident
Degree of relevance of information to the search question	1-4 1 - Not relevant 10 - Very relevant
Degree of misleading of information to the search question	1 -4 1- Not misleading 10- Very misleading
Contribution of key words to effectiveness of search	1-10 1 - Low contribution 10- Very high contribution
Time for finding information	Measured in seconds
Correctness of answer	Answer given by user examined

2.3 Methods of Analyzing Data

A statistical analysis of the data was carried out using a number of tools: Pearson Test, Wilcoxon test for pairs and McNemar for symmetry of the matrix. Most of the results obtained were absolutely high and showed a clear relation between the variables tested. The final results related to the comparison of data between the various methods. Since three methods were examined in the study and for each method an opinion of the same user is given, it was possible to perform a comparative analysis between pairs of methods and to determine the relations between them. An analysis of the data referred at the first stage to the correctness of the answer and afterwards to the time it took the user to arrive at it. In addition, answers to questions on the feedback sheets were analyzed.

2.4 Course of the Experiment

All of the participants in the experiment received a sheet where three search problems were defined. For each problem there was one document in the database in which there was the answer to the problem. The user operated the search engine at the site of the experiment and received a list of answers which fulfilled the search criteria

3. RESULTS

The results gathered in regard to tests on each subject examined are as follows:

3.1 Feeling of Ease in Search

This category tested how much the participant in the experiment felt ease in the search process, The average designates the difference between method C and B and C and A where the range was between 1-10 (Table2).

Table 2 Ease - Difference between Method C and Methods A and B.

Methods	N	Mean	Std. Error	P (for paired T-Test)
C vs. B	51	1.07	0.23	<0.0001
C vs. A	51	1.33	0.33	<0.0001

3.2 Desire that the Search Engines Would Work with the Proposed Method

This category examined the desire of the participant in the experiment that the search engines on the Internet would work in one of the proposed methods (methods A and B are the most common on the network and method C not). The average designates the difference between method C and B and between C and A where the range was between 1-10 (Table 3).

Table 3 Search Engines- Difference between Method C and Methods A and B

Methods	N	Mean	Std. Error	P (for paired T-Test)
C vs. B	51	1.50	0.18	<0.0001
C vs. A	51	2.39	0.29	<0.0001

3.3 Feeling of Confidence of User During Solution of Query

This category tested the feeling of confidence of the user with regard to the information delivered to him with the various methods during solution of the problem.

Table 4 presents a comparison between methods B and C and Table 5 presents a comparison between method A and C. Table 6 displays a comparison between method A and B. In the McNemar test the values of the oblique axis designate symmetry (that is there is no difference between the methods). The values in the lower left triangle designate the preference for method C while the values in the upper right triangle designate the preference for method B (Table 4). The data in the tables are in percentages. The values which were reported were from 1 to 4 where: 1-lack of confidence and 4- Very confident.

Table 4. Confidence – Symmetry of Matrix in Method C vs. Method B

		Method B			
		1	2	3	4
Method C	1	0	1.96	0	0
	2	0	5.88	1.96	3.92
	3	0	0	13.73	3.92
	4	1.96	1.96	60.78	3.92

Values found existing $P < 0.001$

Table 5 Confidence – Symmetry of Matrix in Method C vs. Method A
Method A

		1	2	3	4
Method C	1	0	0	1.96	0
	2	0	7.84	1.96	1.96
	3	1.96	3.92	5.88	5.88
	4	1.96	58.82	3.92	3.92

(P<0.001)

Table 6 Confidence – Symmetry of Matrix in Method B vs. Method A
Method B

		1	2	3	4
Method A	1	0	1.96	1.96	0
	2	1.96	5.88	62.75	0
	3	0	1.96	7.84	3.92
	4	0	0	3.92	7.84

(P<0.001)

3.4 Degree of Relevance of Information Accompanying the Titles (in the List of Results) to the Search Query

This category tested the degree of relevance of the information accompanying the document title (in Method B the top lines of the document and in method C lines from the paragraph which fulfilled the search criteria) for the search query. Use was made of the McNemar test here also. The data in the table are in percentages. The values examined are from 1 to 4 where: 1- information is not relevant and 4- information is very relevant.

Table 7 Relevance – Symmetry of Matrix Method C vs. Method B
Method B

		1	2	3	4
Method C	1	1.96	0	1.96	0
	2	0	5.88	3.92	0
	3	0	3.92	11.76	1.96
	4	0	3.92	60.78	3.92

(P<0.001)

3.5 Contribution of Key Words to Search Effectiveness

This category examined to what degree key words participate in the sensitive experiment for search effectiveness. The average designates the difference between method C and B and between method C to A when the range is between 1-10 (Table 8)

Table 8 Key Words - Difference between Method C and Methods A and B.

Methods	N	Mean	Std. Error	P (for paired T-Test)
C vs. B	51	1.09	0.21	0.0001
C vs. A	51	0.80	0.27	0.0053

3. 6 Time for Finding Correct Answer in Various Tasks

In the attached table are the times in seconds for finding correct answers after incorrect answers were sifted out. Relevant data was collected with regard to the method for each search task. It was not possible to make a comparison of the various tasks because of the great difference between them.

Table 9 Times for Finding Answer with Various Methods according to Different Tasks

Task (name of the task)	N	Method A	Method B	Method C	Average Time for Task (sec.)
1 (F)	21	116	124	84	108
2 (N)	22	50	81	105	79
3 (A)	24	149	146	144	146

4. DISCUSSION

The results and their significance are discussed according to the results as described in the previous section results. A separate analysis of data was made for each criterion.

4.1 Ease in Search

Table 2 compares method C and each of the other methods. The table shows that method C is preferable to method B by 11% (two degrees of freedom) and that method C is preferable to method A by 13% (1.5 degrees of freedom).

4.2 Desire that the Search Engines Work in the Proposed Method

Table 3 compares method C and two other methods. This table shows that method C is preferable to method B by 15% (1.5 degrees of freedom) and that method C is preferable to method A by 24% (2.5 degrees of freedom).

4.3 Feeling of Confidence of User during Answer to Question

Table 4 compares method C and method B. It shows that 65% of the participants felt very confident during operation of the search based on

method C against only 8% who felt the same degree of confidence in using method B ($P < 0.001$). It can be seen from Table 5 that 65% of the users of method C felt very confident during the search operation against 8% with similar feelings as those using method A ($P < 0.001$). Table 6 shows that 69% of the users of method B felt confidence during the search operation against only 6% of the users of method A only who felt confidence ($P < 0.001$).

4.4 Degree of Relevance of Information Accompanying Title of the Search Query

Table 7 compares method C and method B as to the degree of information relevance with the title in the list of results of the search query. It can be seen that 65% of the participants feel that the information displayed to them in the list with method C (lines from the paragraph fulfilling the search criteria) is very relevant to the search query. In practice, this enables finding the answer to the question. Only 2% of the participants feel that the information displayed to them by method B (first line from the document) is relevant to the search query ($P < 0.001$) in contrast to method C.

4.5 Contribution of Key Words to Search Effectiveness

Table 8 shows a comparison between method C to the other two methods. We see from the table that C has precedence over method B by 11% (1 degree of freedom) and that method C has precedence over method A at 8% (1 degree of freedom).

4.6 Time in Finding Correct Answer to Different Tasks

Table 9 displays information which summarizes the average time for doing different tasks using different methods. Tasks are graded by their level of difficulty. In analyzing average times for finding the right answer, one can see the degree of task difficulty. Analyzing times shows that task 2 was the easiest (79 seconds on the average), task 1 was intermediate in degree of difficulty (108 seconds on the average) and task 3 was the hardest (146 seconds on the average). The relative advantage of each of the methods is shown. Method A was the fastest in the easiest task. In tasks of intermediate difficulty, method C leads. Method C was the fastest in the most difficult (though the difference between method B and A was not large).

5. CONCLUSIONS

A number of conclusions can be drawn from the results of this study and by analyzing the different categories. These will be discussed according to the questions asked at the beginning of this paper.

5.1 First Question

In the first question, "What is the most important information to display in a list of search results?" it can be seen that together with the document title that is used by all methods, a clear preference for displaying lines of the document fulfilling the search criteria and not the first lines of

the document as takes place with most search engines and information retrieval systems. Hence, the addition of key words together with document titles contributes to search effectiveness in the eyes of the user. In method C there is an advantage in the feeling of ease for the user. Feeling of confidence increases during search time, displayed of relevant information is better (the lines fulfilling search criteria). In everything relating to the speed of finding information we saw that the information display in this system brought a benefit during the time of the search when the search task was intermediate. In simple tasks other methods were found to be faster. For complex tasks there in no advantage found for this method over method B.

5.2 Second Question

The second question was " Comparing methods of displaying information in the list of search results, which method is preferred by the user?". Three methods were tested for displaying the list of search results derived from the textual database. They were tested in a number of categories. These were; Method A - display of the list of results including document titles plus key words of documents (without additional information from within the document). Method B - Display of the list of results including document titles plus key words (as in method A) and in addition, the first lines of each document. Method C - display of the list of results including document titles plus key words (as in method A) and in addition the lines fulfilling search criteria from the body of the text.

The following are conclusions from the study..

1. Method C is preferred over method B and over method A in every thing relating to the feeling of ease in searching.
2. Participants in the experiment would prefer that search engines on the Internet were done by method C more than methods B and A.
3. During solution of the answer the experiment users felt very confident in using method C over methods B and A. In an accompanying test it was found that with method B the participant feels confidence compared to method A. The significance of this fact is that when only two methods are involved, method B and A, there is a preference for method B although it is less powerful than the difference between method C and method B. Answers obtained between method C to B, related to the very great confidence in contrast to the answer received in comparing method B to A which referred to confidence only).
4. The degree of relevance of information accompanying the title (in the list results) to the search query would have been greater in method C than in method B. The meaning of this is that the lines which fulfill the search criteria in the body of the document are significantly more relevant than the first lines of the document and

this in contrast to the accepted method today in most of the search engines on the Internet.

5. Contribution of Key Words to Search Effectiveness

The purpose of this category is to verify to what degree the addition of key words to the information displayed contributes to search effectiveness. The addition of key words is possible by a number of means where the easiest and fastest is based on a list of key words found in the document itself (as part of the document text or additional information found in some of the documents as meta tags). Another way is the use of existing algorithms to produce key words based on an analysis of the text by various methods [16]. Findings show that the addition of key words improves search effectiveness by 10% with respect to method C as against the two other methods.

6 Times in finding correct answer in different tasks.

Results here were more ambiguous because time was more complex. Each search task was concluded at a different time and the average times for finding answers were different for different methods so that a different method was preferable for each task. An analysis was made of the character of the tasks from the standpoint of difficulty to find an explanation for time differences. Details of this analysis are found in Table 10. Averages showed significant differences during the time of carrying out tasks: 79 seconds (task 2), 108 seconds (task 1), 146 seconds (task 3) and this with respect to the average time in performing each one of the tasks. The length of time for carrying out the task attests to the degree of difficulty since the database and number of documents were identical for every task. An analysis was made regarding the percentage of wrong answers for each task to verify the subject of query difficulty. It was found that in task 3, 18.18% wrong answers were obtained vs. task 1 where 8.79% were obtained and in task 2, 6.19% wrong answers were obtained. Assuming that the number of wrong answers show task difficulties, the conclusion as to the degree of difficulty in the tasks was able to provide more encouragement.

Table 10: Degree of Difficulty of the Task According to Times and According to Percentage of Wrong Answers

Task	Average time for task (sec.)	Degree of difficulty according to time	Percentage of wrong answers	Degree of difficulty according to wrong answers
1	108	Intermediate	8.79	Intermediate
2	79	Easy	6.19	Easy
3	146	Intermediate - hard	18.18	Hard

In the light of this difference it was found that the various methods are suitable to tasks of different degrees of difficulty where method A is good for easy tasks, and method C is good for tasks of intermediate difficulty - and each as noted with respect to the components of time.

5.3 General

As we saw in this study and in the previous one [5], the first lines of the document do not necessarily reflect the contents of the document. We also saw that in most cases there are other lines in the document which reflect the contents and these are the ones which fulfill the search criteria. In addition, we learned from this study about the contribution of key words for understanding the contents of the document. As we saw in method C a combination of key words and relevant lines gives the best search results. The relevant lines help the specific task and key words help in understanding the text in general. It may be inferred that that key words make possible the first filtering of the document which are not within the general search field and the relevant lines aid in refining the final search. The conclusion is a clear instruction to writers of articles, developers of databases, etc. to add key words to the writing material. If the material does not contain key words it is possible to produce them for every written text by using existing algorithms [16] [17] [19].

6. IMPLEMENTATION AND FUTURE STUDIES

Textual information databases are being used more and more. The quantity of information which is not keyed or catalogued has increased because of questions of cost and lack of centralization and control of their contents. This is especially true for the Internet. This paper has shown the possibility of defining the most important information to display when searching produces large amounts of results. The proposed method C provides a significant advantage in a number of categories regarding system users. It is not difficult to implement the ideas set forth in this paper with present search engines used on the Web and in dedicated systems as found with information suppliers and internal organization systems, etc. Information components displayed for the user can be decided at almost

every level of textual retrieval systems and conventional systems having lists of documents supplied by criteria.

A number of directions are available in future studies: finding information required for display in the list of results provides a view of document environment in addition to information in the body of the document. This study will need to find the types of information and also examine their effect on the behavior of the user in a manner similar to this study.

REFERENCES

- [1] D. Blair and M. Maron. An evaluation of retrieval effectiveness for a full-text document retrieval system. *Communications of the ACM*, New York: ACM, 289-299. (1985)
- [2] G. Chowdhury. The Internet and information retrieval research: A brief review. *Journal of Documentation*, **55** (2) 209-225 (1999).
- [3] O. Drori. The user interface in text retrieval systems. *SIGCHI Bulletin, A Quarterly Publication of the ACM Special Interest Group on Computer-Human Interaction*, New York: ACM, **30** (3) 26-29. (1998)
- [4] O. Drori. Integration of text retrieval technology into formatted, (conventional) information systems. *Software Engineering Notes (SEN)*, New York: ACM, **24** (1) 78-80. (1999)
- [5] O. Drori. The Benefits of Displaying Additional internal Document Information on Textual database Search Results Lists, *Proceedings of ECDL2000 - Fourth European Conference on Research and Advanced Technology for Digital Libraries (September 2000, lisbon, Portugal), Lecture Notes in Computer Science, No. 1923*, Berlin: Springer Verlag, 2000, pp. 69-82. (2000)
- [6] D.E. Egan et al. Behavioral evaluation and analysis of a hypertext browser *CHI '89 Proceedings*, New York: ACM, 205-210. (1989)
- [7] M.A. Hearst. TileBars: Visualization of term distribution information in full text information access. *CHI '95 Proceedings*, New York: ACM, 59-66. (1995)
- [8] M. Hertzum and E. Frokjaer. Browsing and querying in online documentation: A study of user interface and the interaction process. *ACM Transaction on Computer-Human Interaction*, New York: ACM, **3** (2) 136-161. (1996)
- [9] S. Kirsh. . Infoseek's experiences searching the Internet. *SIGIR Forum*, New York: ACM, **32** (2) 3 (1998)
- [10] J.D.Mackinlay, R. Rao and S.K.Card. An organic user interface for searching citation links. *CHI '95 Proceedings*, New York: ACM, 67-73. (1995)
- [11] B. Peritz. On the informativeness of titles. *International Classification*, Frankfurt: INDEKS, 11 (2) 87-89. (1984)
- [12] P. Pirolli et al. Scatter/gather browsing communicates the topic

- structure of a very large text collection. *CHI '96 Proceedings*, New York: ACM 213-220. (1996)
- [13] J. Pitkow and P. Pirolli. Life, death, and lawfulness on the electronic frontier. *CHI '97 Proceedings*, New York: ACM, 213-220. (1996)
- [14] J. Raskin. Systemic implications of leap and an improved two-part cursor: A case study. *CHI '89 Proceedings*, New York: ACM, 167-170. (1989)
- [15] B. Regan. Reasons why finding information on the Web is so difficult. *IdeaBytes*, New York: Giga information Group Inc. (1998)
- [16] C. Rijsbergen. *Information Retrieval*. London, England (1979)
<http://www.dcs.gla.ac.uk/Keith/Preface.htm>
- [17] G. Salton. *Automatic Text Processing - The transformation, Analysis, and Retrieval of Information by Computer*. Reading, Massachusetts: Addison-Wesley, . 232-236, 384-385. (1989)
- [18] A. Sievert. Full-text information retrieval: Introduction. *Journal of The American Society for Information Science*, New York, Wiley, **47** (4), 261-262 (1996)
- [19] J. Sparck. *Automatic Keyword Classification for Information Retrieval*, London, England (1971)