Defense Technical Information Center

Indexing Quality

Baseline Analysis

Prime Contract MDA93-90-D-0022
Task Order 50

June 27, 1994

Submitted by

Jessica L. Milstead

The JELEM Company
P.O. Box 5063 • 17 Brooks Lane
Brookfield, CT 06804-5063

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
27 June 1994		

4. TITLE AND SUBTITLE
Defense Technical Information Center Indexing Quality Baseline Analysis

5a. CONTRACT NUMBER
MDA93-90-D-0022

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Jessica L. Milstead

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The JELEM Company
PO Box 5063
17 Brooks Lane
Brookfield, CT 06804-5063

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Defense Technical Information Center
Fort Belvoir, VA 22060-6218

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution Statement A - Approved for Public Release.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

When asked how well DTIC provides intellectual control over the subject content of its databases, nearly every end user, intermediary, indexer, or DTIC manager can offer an opinion of some sort on the quality of DTIC's indexing. These responses, however, typically have only two characteristics in common: High subjectivity, based on impressionistic data rather than on hard evidence. A consensus that, if it is to function in the future as an effective information provider, DTIC must find ways to modernize its control over the subject content of its databases. Based on this consensus, DTIC has undertaken to improve the quality of its indexing. The effort is seen as involving three phases: Determination of the baseline quality of the present indexing. Development of an indexing philosophy. Identification of access methods which can be successfully integrated with the Electronic Document System (EDS), which is presently under development. This report describes the first phase of this report, baselining of present indexing.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
Unclassified					
a. REPORT	b. ABSTRACT	c. THIS PAGE	UU		19b. TELEPHONE NUMBER *(include area code)*
U	U	U			

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

Management Summary

DTIC has undertaken a three-phase effort to improve the quality of its subject indexing. This report describes baselining of the present indexing, the first phase of the effort. Later phases are to include development of an indexing philosophy and identification of access methods which can be integrated with the Electronic Document System (EDS), which is presently under development.

In order to address the issue of indexing usability, the methodology selected was that of the recall/precision study. A sample of 19 queries was selected from the query files at DTIC. Queries were searched in both the Technical Report (TR) Bibliographic Database and Work Unit Information System (WUIS); retrieval was limited to unclassified citations.

The original design called for queries to be searched by two groups of expert searchers (searchers with extensive experience in searching the DROLS system): DTIC searchers who were instructed to search broadly with the goal of retrieving as many relevant documents as possible (data acquisition searchers); and searchers at user installations who were instructed to search as they would for a real user, insofar as possible (test searchers).

This design assumed that there would be significant differences in search strategies between broad, inclusive searching, and searching carried out under real-life constraints. Had this been the case, overlap of test searches with data acquisition searches would have been very high, but test searches would have retrieved fewer citations than data acquisition searches. It was intended that recall and precision would be calculated for test searchers only, because they were the ones to be searching under more "realistic" conditions.

In actual fact, neither the search strategies nor the overall characteristics of their retrieval differed significantly between the two groups of searchers. Data acquisition and test searchers retrieved similar numbers of hits for each query, and examination of their search strategies showed no clear differences between the two groups.

Therefore the retrieval of the two groups of searchers was amalgamated for analysis, and recall and precision were calculated for all searchers. The wide variation in search strategies described below means that the retrieval of all searchers combined is reasonably broad. Analyzing all searchers' results for recall and precision, rather than just those of the test searchers, provides a broader set of findings for interpretation, and should compensate for any effect of the loss of the distinction between the two search phases.

The variation in search strategies was great, not so much in actual choice of concepts to be searched as in refinements such as use of truncation, inclusion of IAC terms, and searching of words in titles or abstracts. Some of these variations are attributable to design of the indexing system, while others reflect features of the search system.

Hits were downloaded — with some difficulty — and processed into a database to permit identification of the fields in the downloads. Reports containing subject and narrative fields

were generated for judging of the relevance of the citations. If over 100 unique citations were retrieved for a query, a random sample of 100 citations was selected for judging.

Relevance judges were subject experts located at five of the DTIC-supported Information Analysis Centers (IACs). Each set of hits was evaluated by two experts located at the same IAC. Because concurrence between the judges was very low in numerous cases, recall and precision values were calculated both as if a relevant document were one judged as relevant by both, and as if it were one judged relevant by either judge.

By either criterion of relevance, the recall/precision values did not show the conventional inverse relationship between precision and recall. Overall, mean recall for each query was low to moderate, with great variation in precision from very low to very high. The relatively low mean recall is particularly striking when it is considered that the method of locating relevant documents was necessarily quite limited. The documents retrieved by the searches in the study formed the base from which relevant documents were selected; relevant documents in the database which were not retrieved could not enter into the evaluation.

Even though specifics cannot be determined at this time, it is reasonable to conclude that improvements to both indexing and the search engine are warranted. The difficulties encountered in downloading indicate that there are problems somewhere in the telecommunications system as well.

Phase Two will provide more definitive answers as to the causes of the retrieval inadequacies encountered, at least for the documents which were actually retrieved by at least one searcher. Those which remain unknown, of course, cannot be evaluated. In Phase Two, a sample of retrieved documents will be evaluated for each query. The reasons for failure of a given search to retrieve a relevant document, or for its retrieval of a nonrelevant document, will be recorded and categorized. It is anticipated that there will be two broad categories: indexing failures and search failures, with several subcategories such as vocabulary inadequacy (indexing) or failure to truncate (search).

A philosophy of subject indexing for DTIC will also be developed in Phase Two. This effort will include extensive consultations with DTIC units such as subject analysis as well as DTIC users.

Introduction

Background

When asked how well DTIC provides intellectual control over the subject content of its databases, nearly every end user, intermediary, indexer, or DTIC manager can offer an opinion of some sort on the quality of DTIC's indexing. These responses, however, typically have only two characteristics in common:

- High subjectivity, based on impressionistic data rather than on hard evidence.
- A consensus that, if it is to function in the future as an effective information provider, DTIC must find ways to modernize its control over the subject content of its databases.

Based on this consensus, DTIC has undertaken to improve the quality of its indexing. The effort is seen as involving three phases:

- Determination of the baseline quality of the present indexing.
- Development of an indexing philosophy.
- Identification of access methods which can be successfully integrated with the Electronic Document System (EDS), which is presently under development.

This report describes the first phase of this effort, baselining of the present indexing.

Present Indexing at DTIC

DTIC's databases are processed through a machine-aided indexing (MAI) program which selects candidate index terms from the DTIC Thesaurus, based on matching of words and phrases in titles and abstracts or full text. In the Technical Report (TR) Bibliographic Database, titles and abstracts are processed. Human editors review the candidate terms, and may add or delete terms to reflect the actual content of the document. Documents included in the Work Unit Information System (WUIS) are shorter, and the full-text record is entered into the system. Indexing for these records is mounted on the database without human review; i.e., it is automatic. In addition to being run through the MAI system, the words in the text of the records are also inverted. The Independent Research & Development (IR&D) Database is also indexed with MAI and no human review.

The system for selection of index terms has been well-reported in the literature (Klingbiel, 1973; 1985; Klingbiel & Rinker, 1976; Jacobs, 1990). This system was developed in the 1970's, and was updated to some extent in the early 1980's. In its current mainframe

version, it uses both phrase structure and transformational grammars to extract natural language phrases. These are then processed against a dictionary to produce descriptors. Processing is against titles and abstracts only in the TR Bibliographic Database, and against full text in WUIS and IR&D. Since the original system was developed, the state of the art of text processing has advanced significantly, in tandem with growth in computing power that permits more sophisticated processing.

In September 1992, a PC-based MAI system which is limited to word matching was implemented. The PC-based system represents a lower level of functionality, developed to permit indexing on PCs; it is applied only to the TR Bibliographic Database.

Some records in TR are originated and indexed by IACs, using a specialized vocabulary in addition to DTIC Thesaurus terms. Searching on these terms requires use of a special mnemonic because the terms are stored in a separate field from the DTIC-assigned terms. While it is also possible to limit a search by specific IAC(s), this capability was not relevant to this study, which was limited to subject indexing.

Other TR records are originated by members of DTIC's Shared Bibliographic Input Network (SBIN), using the same indexing policies as DTIC.

Plans for Indexing Development at DTIC

As indicated above, the purpose of the present study is to determine the baseline quality of the present DTIC indexing system. It is important to be aware of the problems which are present in order to make effective recommendations for improvement. Improvements will affect SBIN records as well.

Work is underway at DTIC on a system for electronic storage of documents, of which subject retrieval will be an important component. However, this system will not be fully operational much before the end of the decade, and it is desired to make improvements in the indexing system much sooner than that — and to make these improvements in a way that permits as much as possible of the effort to be carried over into the new system.

Methodology

Recall/Precision

The most important question to be addressed was the usability of the subject indexing DTIC produces, not its "quality" as an abstract concept. It was therefore important to select a methodology that would permit determination of how well the indexing aids the information retrieval process. The most-tested methodology that can meet this criterion reasonably well is the recall–precision study.

This method has the longest history — and therefore the best knowledge of both its strong points and its pitfalls. Such a study attempts to quantify two characteristics of retrieval: the proportion of the relevant records in a database which were actually retrieved by a search, and the proportion of the records retrieved which are actually relevant to the information need. A "relevant" record is one which bears on the query; the required closeness of match between record and query is usually a subjective judgment; a test may allow for degrees of relevance.

Recall is defined as:

$$\frac{R_{ret}}{R_{tot}} \times 100$$

Precision is defined as:

$$\frac{R_{ret}}{T_{ret}} \times 100$$

where R_{ret} is the number of relevant records retrieved in answer to a query, R_{tot} is the total number of documents in the system that are relevant to that query, and T_{ret} is the total number of records retrieved by the query.

These measures have some limitations. First and foremost, in a practical situation, it is impossible to judge every document in a database for relevance to every query in a test. Instead, approaches which seek to retrieve as many relevant records as possible are devised. These include searching the same query in multiple databases — if the document records are included in more than one database — or searching the queries very broadly or with multiple strategies, concentrating on retrieval of as many relevant records as possible, with little concern for the number of irrelevant records retrieved. Records retrieved by these devices are evaluated for relevance, and become the basis against which test searches are judged.

"Relevance" itself can be interpreted in several ways, but the one most appropriate here is topical relevance: a relevant record is one which bears on the query. Froelich (1994), in his paper introducing a special issue on relevance of the *Journal of the American Society for Information Science*, makes the point that topical relevance is the basis for other forms of relevance, such as user relevance (i.e., relevance to the user's information need at the moment of accessing the information). User relevance changes as more information is accessed. For example, if two documents happen to contain essentially the same information, both are topically relevant. However, only the first document seen may be relevant to the user; the second will be repeating information which the user already knows. This makes user relevance order-dependent, and would require that the information system have in-depth awareness of the user's knowledge state and its changes from moment to moment.

Since information retrieval systems are only partially successful today at producing high topical relevance, the first priority should be fulfillment of this basic criterion.

An exhaustive research project, supported by the National Science Foundation, to study information seeking and retrieving, was reported by Saracevic (1987, 1988). The retrieval effectiveness aspects of the Saracevic study provided the framework for the present study. Topics such as question typologies and searchers' cognitive traits which were covered in Saracevic's work are not germane to the analysis of indexing quality and were not included.

In the present study, real queries were searched exhaustively by expert searchers; then the same queries were searched in a more typical search situation. All of the retrieved citations were amalgamated and subject experts judged the topical relevance of either a sample or the entire set, depending on the size of the retrieved set. Given the impossibility of evaluating every document in a large collection for relevance, this strategy cannot be assumed to locate every relevant document in the system. However, it is the approach which has been most successful in previous research, when it was necessary to limit searching to a single database. (If it is possible to search for the same documents in multiple databases, a greater exhaustivity of retrieval can be attained.)

For this study, R_{tot} or total relevant records, was taken to be all the relevant records retrieved by all of the searches.

Scope and Limitations

The study was limited to unclassified documents, because it would not be feasible to attempt to judge relevance from sanitized data. It was determined that there is no difference in indexing policies between classified and unclassified documents. Both the TR Bibliographic Database and WUIS were included, but the IR&D Database was excluded to avoid problems of access to proprietary information.

In order to assure that real queries were the subject of the test, they were drawn from the files of search queries at DTIC. A limitation of this procedure as compared with an actual search situation is that it is not possible for the searcher to interact with the user to refine the query or determine if the retrieval is on target. While this limitation led to some concern on the part of the searchers, it also made it possible to focus on quality of indexing, rather than on searchers' ability to compensate for lack of quality.

It also was necessary to limit the topics of the queries to those for which expert judges were available to judge the relevance of the results. DTIC funds and manages a number of the DoD Information Analysis Centers (IACs), and it was determined that the best resource for relevance judging would be the staff of these IACs. Only queries which could be judged by experts at the IACs were selected for searching. A subset of the IACs at which particularly expert searchers were available was selected with the assistance of Dr. Forrest Frank of DTIC.

While 100-record samples of the citations retrieved by large searches were evaluated for relevance, rather than the entire set of the retrieval, appropriate statistical treatments were applied to show the confidence with which the results can be interpreted as applying to the entire set of retrieved citations. These are the confidence intervals shown in the various tables in this report. While the samples themselves ranged from 15 percent to 100 percent of the hits for a query, absolute sample size and variance in the results are more important to statistical inference than the percentage of the universe covered by a sample, so long as it is randomly selected.

Procedure

The following steps were followed in gathering and analyzing the data:

1. Selection from the DTIC query files of a set of queries to be searched.

2. Searching, divided into data acquisition and test phases.

3. Relevance judging.

4. Statistical analysis and inference.

Appendix 4 shows one search strategy, the number of hits, and calculation of precision and recall statistics for that search.

Query Selection

In August 1993, the DTIC query logs (queries phoned in by users for search by DTIC staff) were reviewed to make a tentative selection of queries which were within the scope of the IACs to be involved, with the goal of having 15–25 queries in the final set. Logs for September 1992, and for January, April, and July 1993, were sampled by looking at every tenth query for key words that indicated the subject of the search. Since the query log includes only a brief title, many searches were rejected at this point for lack of information.

The following criteria were established for queries:

- The query should be within the scope of one of the IACs selected for participation.
- A balance should be maintained, with a goal of 3–6 queries per IAC.
- The query should appear to have retrieved at least one relevant document.

The queries which from their titles seemed to be of possible value were then examined in the query files. Any queries which were not actually within the scope of the IACs being considered for participation, which did not appear to have relevant retrieval, or which were too vague, were rejected at this point. A total of 32 queries survived this first filtering.

All information about the query was recorded, but not the search strategy. From this information, a statement of the query was formulated which contained all the relevant terms. The strategy was omitted to avoid influencing either the search strategies developed for the study or the way in which the query was formulated for the searchers.

The 32 queries were tentatively organized according to IAC, and were then reviewed with Dr. Forrest Frank for suitability. He reassigned some of the queries to different IACs, and selected 19 which were most appropriate for searching. These 19 queries became the query set for the study. A list of the queries, and of the IACs to which they were assigned for relevance judging, may be found in Appendix 1.

Searching

Information gathered

It was important to select an appropriate set of fields for downloading. The study concerned indexing quality, and it was not desirable for a judge's determination of relevance to be influenced by such factors as author's affiliation. Still, it might be useful to be able to identify the source of a document at a later date. Therefore, fields that would contribute to a judgment of relevance on the basis of subject, plus fields identifying the corporate source and the data, were downloaded. Downloaded fields are as follows:

TR Database

1 — DTIC Accession Number
5 — Corporate Author
6 — Title
11 — Report Date
23 — Descriptors
25 — Identifiers
27 — Abstract
44 — IAC Subject Terms

WUIS Database

an — DTIC Accession Number
rd — Report Date
ti — Title
de — Descriptors
poa — Performing Organization Activity Name
ran — Responsible Organization Activity Name
kw — Key Words
app — Approach
obj — Objectives
prg — Progress

In the TR Bibliographic Database, Corporate Author and Report Date were not provided to the relevance judges; in WUIS, the Performing Organization, Responsible Organization, and Report Date were not provided. This information was omitted in order to prevent any possible influence of non-subject data on the relevance judgments.

Because the purpose of this baseline study was to determine the present quality of indexing, it was decided to accept the default search limitation to the past 10 years, rather than searching the full database. After a number of searches had been completed, it was learned that this default applies only to the TR Bibliographic Database and that the default for the WUIS Database is no time limitation. Therefore, the searches of the two databases covered different time spans. This difference was not seen as presenting a problem because the two databases were to be analyzed separately in any case.

Data acquisition searches

These searches were conducted in-house by DTIC staff, who were instructed to develop broad strategies designed to retrieve as many relevant records as possible, even if this meant a substantial number of irrelevant retrievals. A copy of the introductory materials provided to searchers may be found in Appendix 2. The searchers received a copy of "Introduction to the Study" and "Conducting Searches: Data Acquisition Phase." A total of six searchers participated in this phase.

In order to remove the influence of search order from the results, the order of the queries was randomized for each searcher. The goal was to have each query searched by at least two searchers; due to the varying success rates of searching and the effects of randomization, queries actually were searched by from two to four searchers each. Assignment of queries to individual searchers was random, with no relation to the subject content of the searches.

TR and WUIS were searched consecutively for each query, since in many cases the same strategy was applicable to both. The searchers input and revised their strategies, and when they were satisfied with the retrieved set, downloaded the records including the specified fields.

Dialup access was used so that records could be downloaded to floppy disks for analysis. A problem with loss of data during transmission was encountered with downloading. When the message "data lost toward terminal" was encountered, a gap that might include several records was invariably discovered somewhere earlier in the download. This happened quite consistently with almost every search that retrieved a significant number of hits. The problem was solved — without ever tracing the actual cause — when it was noted that searchers who made a habit of storing their results in a user file and then downloading from that file, instead of displaying search results directly, never had data losses during transmission. The user file step was included in all future searches, and downloading was successful, without loss of data.

In order to acquire the missing data, defective searches were rerun using the original strategy. After verifying that the strategy was unchanged, the new retrieved set was substituted for the defective set. Unfortunately, the Descriptor field was omitted from some downloads of these rerun searches; these strategies were run a third time, downloading only Accession number and Descriptor fields in order to save time. The retrieved records were merged with the earlier records so that the record contained full information.

Even aside from telecommunication problems, downloading was found to be a very slow process, reducing significantly the number of searches that could be executed even by searchers who used two terminals simultaneously. The principal investigator was present during search strategy formulation and running of the original searches, but not during the reruns.

The use of dialup access also caused problems in search modification; searchers who were accustomed to being able to modify and refine searches in progress on a dedicated terminal had to rekey the search from the beginning on the dialup system in order to make even minor changes.

Test Searches

The test searches were carried out by experienced searchers at defense installations. Two searchers searched each query. As with the data acquisition searches, the order of searching was randomized, and no two searchers searched exactly the same queries in the same order. They were given a copy of the introduction to the study, and the instructions for the test phase, reproduced in Appendix 2.

Originally it was thought that seven searchers would be used for this phase, but it was found that the results from six were adequate to assure that each query was searched twice; therefore only the results from these six were used.

The original goal of the test searches was to determine the quality of retrieval under conditions as close as possible to those prevailing in a real situation. It was assumed that this would mean a more focused search, with some care taken to minimize irrelevant retrievals, even at the price of less than maximum recall, because time pressures would be greater and users would not want to scan an immense quantity of output. In reality, the searchers were in research situations where the concern was to find as much relevant material as possible. Their searches were as broad as the data acquisition searches. Furthermore, the strategies were different enough that overlap was not very high; that is, they tended to retrieve different citations.

Searchers were asked to proceed as if the search were a real one requested by one of their users. The only difference was to be that user feedback was unavailable. All of the searchers were somewhat uncomfortable with this limitation, because they correctly regarded user feedback as fundamental to the search process. With the explanation that this limitation was unavoidable in the study, most were able to proceed without further

discussion. When searchers posed the problem of user differences, particularly in willingness to accept a large retrieved set, they were asked to assume a user preference for moderate retrieval, i.e., somewhere between "comprehensive" and "a few good citations."

The principal investigator was present while searches were conducted by four of the searchers. DROLS was down when the site visit was made to work with a fifth searcher; he prepared the strategies at that time and ran them later. The sixth searcher found it possible to carry out only a few searches, as time permitted, so that it was not possible to be present while he was working.

Most users did not have downloading capability. Downloading as searches were performed was possible in only two cases. Even these installations had difficulty because they rarely or never used downloading. Other available installations had no downloading facilities at all.

The procedure was modified to permit gathering the crucial input — their strategies — from the other test searchers. Dedicated or dialup access was used, whichever was most convenient for a particular facility. The searcher developed and modified the strategy until s/he was satisfied that the retrieval was what was desired, but did not print or download the full list of hits. The strategy was recorded, and was executed by a trained searcher at DTIC, using dialup access and downloading the hits. The principal investigator verified that the strategy that was actually searched was identical to the strategy developed by the test searcher.

Search strategies for each query are listed in Appendix 7. Data acquisition searchers are coded A–F; test searchers are coded 1–6.

Data processing

The raw search output was first run through word processing. For each search, the strategy was separated from the retrieved citations. Then a listing of the accession numbers was generated, and a count performed to assure that if the accession number count did not match the citation count in the retrieved set, the reason could be determined. The reasons included notes such as "unannounceable category" or "document not available." No subject data were available for such hits, and they were excluded from further analysis. Similarly, Referrals were excluded from the analysis, because these are not "documents" and the records do not contain a significant amount of substantive information.

System messages, line noise, etc., were deleted from the citation file, which was then sent to a data processing contractor who read the data into a database, permitting all retrieved fields to be identified. The records for all searches of a given query were merged, retaining an indication of the searcher(s) who had retrieved each citation. In addition to all of the fields in the retrieved record, each record included the query number, the database, and the searcher who retrieved it. Searcher codes were alphabetic for data acquisition searchers and numeric for test searchers, making the two groups easily distinguishable.

A deduplication routine, using the accession number as the key, was run for each query, and the duplicates stored in a separate file. The file without duplicates was used to generate the report for the relevance judges. Tables 1 and 2 show the number of hits available for further analysis (i.e., excluding records for unavailable documents and referrals) for data acquisition and test searches, total hits, and mean (average) number of hits for each searcher as well as the total unique (deduped) hits for each query.

Query No.	No. of searchers		No. of hits		Total hits	Mean hits	Unique hits
	D.A.	Test	D.A.	Test			
1	3	2	471	102	623	125	445
2	2	2	7	70	19	5	18
3	2	2	89	84	173	43	154
1	2	2	570	346	916	229	425
5	4	2	127	11	138	23	130
2	2	2	159	143	302	20	179
2	2	2	101	70	171	43	112
2	3	2	231	390	621	124	560
2	2	2	264	251	916	129	369
16	3	1	174	7	171	45	176
14	3	2	11	9	20	4	9
12	3	1	102	42	144	36	112
13	3	2	296	78	374	75	186
14	4	2	128	9	138	20	112
16	3	2	26	11	37	7	26
16	3	2	34	22	56	11	26
14	3	2	257	78	336	67	265
18	2	1	504	90	594	198	351
19	3	2	537	87	624	125	350

Table 1. Number of Hits by Query: TR Bibliographic Database

Query No.	No. of searchers		Number of hits		Total hits	Mean hits	Unique hits
	D.A.	Test	D.A.	Test			
1	3	2	749	266	1,015	203	682
2	3	2	17	108	125	25	108
3	2	2	28	242	270	68	207
4	2	2	446	308	754	189	321
5	4	2	127	27	154	26	126
6	3	2	208	131	339	68	137
7	4	2	178	63	241	40	122
8	3	2	180	238	418	84	363
9	3	2	186	283	469	94	287
10	3	1	196	41	237	59	222
11	3	2	10	6	16	3	6
12	3	2	74	46	120	24	93
13	3	2	260	204	464	93	225
14	5	1	86	31	117	20	100
15	3	2	37	65	102	20	60
16	3	1	96	27	123	31	93
17	3	2	67	96	163	33	114
18	2	1	137	12	149	50	125
19	3	2	80	57	137	27	78

Table 2. Number of Hits by Query: WUIS

After the retrieval results were deduped, a random sample was taken of the hits for queries which retrieved more than 100 unique citations. One hundred citations per query was considered to be the maximum that it would be appropriate to expect a relevance judge to review in a relatively short time — particularly since most judges would be reviewing the results of four or five queries in each of the two databases.

The retrieval pattern

Most searchers, as indicated above, searched rather broadly. One data acquisition searcher was uncomfortable with this practice, and refined his searches extensively in an attempt to minimize retrieval of irrelevant records. He typically retrieved only a few citations, and for a number of queries retrieved none. In two cases a test searcher's strategy for a query was so broad that it produced over 1000 hits, and no downloading was attempted.

Since there was no consistent pattern of difference between the data acquisition and test searches, the distinction between them was abandoned for purposes of data analysis, even though a record of the phase in which a search was performed was maintained.

Relevance Judging

Searching is frequently carried out by generalists for specialists. It is important that the relevance judges be qualified to determine the relevance of a particular citation to a query, and a reasonable knowledge of the subject field is required for this. For this reason, subject-specialized personnel at a number of DTIC-sponsored IACs were called on to serve as judges. As noted above, Dr. Forrest Frank provided guidance in IAC selection. Five IACs — CSERIAC, GACIAC, IRIA, MTIAC, and SURVIAC, were involved, with two judges at each IAC. The queries are listed by the IAC judging them in Appendix 1. The judges were selected by the director of the relevant IAC.

Each judge was sent a package containing a letter expressing thanks for participating in the project and briefly explaining the task, plus a copy of the "Introduction to the Study" and "Relevance Judging," found in Appendix 2. These materials were on top of a set of envelopes, each envelope containing a cover sheet with the query number and query statement, plus the citations to be judged. Judges were given 2 weeks to complete the task; most completed it in a week to 10 days.

The director of one IAC asked to have both sets of materials sent to him; a staff member called to confirm that the judges were expected to work independently. Another judge called with a request for more information on a particular query. When this individual was assured that the queries were real ones that had been received at DTIC, and it was explained that the principal investigator had no more knowledge of the query than had been provided (and had to be careful not to interpret for fear of contaminating the results), this seemed to resolve the the problem.

Statistical analysis

The relevance judgments and codes identifying the judges were added to the citation records. "Relevant" and "Partially relevant" judgments were combined for all analyses. Measures of concurrence between judges were calculated (Appendix 5).

Table 3 shows the lambda values representing the extent of concurrence between the judges. Lambda is a statistical measure, varying from 0 to 1, which is used to determine the extent to which a dependent variable can be predicted from the value of an independent variable. In this case, the variables are the two judgements, and the lambda value shows how well it is possible to predict the judgment of one judge if the judgment of the other is known.

Overall, the concurrence of the judges was not high, with a pooled lambda (for all queries) of 0.21 for TR and 0.29 for WUIS.

There is an interesting variation by pairs of judges which is only partly evident from the lambda values above. The same pair of judges evaluated queries 1–4. These judges' level of agreement on relevance was fairly high, but one judge clearly imposed a higher standard for relevance, so that while one judge considered nonrelevant citations that the second judge considered relevant, the reverse was not the case — the second judge's nonrelevant documents were also nonrelevant for the first judge.

Agreement between the pair of judges who evaluated queries 5–9, on the other hand, was quite high, with a number of instances of complete agreement (lambda=1.00). Agreement between the judges who evaluated queries 10–14 was moderate to low, while agreement between those who evaluated 15–18 and 19 was extremely low, with many lambdas of 0.00.

It should be noted, however, that a lambda of 0.00 does not imply complete disagreement between judges. For example, on TR query 18, the judges agreed on 61 of 83 documents or 73 percent, but the lambda was 0.00; on WU query 18 they agreed on 78 of 96, or 81 percent, but the lambda was 0.10. This situation arises because lambda gives the incremental accuracy of predicting one judge's opinion from his/her own history, as opposed to basing it on the other judge's data, and a certain amount of agreement is to be expected by chance.

Query	TR	WU
2	0.11	0.07
2	0.22	0.07
3	0.50	0.20
4	0.15	0.00
5	1.00	0.75
6	1.00	1.00
7	1.00	0.88
8	1.00	0.94
9	0.91	0.33
16	0.50	0.33
17	0.50	0.33
17	0.28	0.17
13	0.16	0.43
14	0.36	0.00
16	0.07	0.00
16	0.07	0.00
17	0.07	0.00
18	0.00	0.10
19	0.06	0.00
Pooled	0.21	0.29

Table 3. Lambda values of relevance judge concurrence

Recall and Precision Measures

Recall and precision ratios were calculated for each searcher and each query in both the TR Bibliographic Database and WUIS (Appendix 6), and amalgamated ratios were then calculated for all searchers of each query (Appendix 6 and Tables 4 and 5).

Since the measures of concurrence were rather low for most queries, the analyses of precision and recall were carried out twice using different criteria for determining a relevant document:

- one which both judges concurred in evaluating as relevant or partially relevant
- one which either judge had evaluated as relevant or partially relevant

The first of these methods is stricter than the second. Appendix 6 shows precision and recall ratios by searcher for each query. Weighted means are given by searcher and by query, with 95 percent confidence limits. As could be expected, precision is higher — sometimes much higher — using the more generous criterion. Recall, on the other hand, tends to be lower. Tables 4 and 5 summarize the precision and recall ratios for TR and WUIS, respectively. Appendix 7 gives the total number of hits in the sample for each searcher and each query, with the number of these which were relevant by each of the two criteria.

Figures 1 through 4 plot the weighted mean precision vs. the weighted mean recall for each query, by both criteria. Perhaps the most striking thing about these ratios is that there is no evidence of the conventional inverse relationship between precision and recall for these queries.

Figure 1 shows the ratios for the TR Bibliographic Database, based on concurrence of judges. About half of the searches (nine) cluster between recall levels of .27 and .49 with precision of .28 to .43. Another nine searches are scattered at a higher precision level of .54 to .93, with a scattering of recall from .18 to .64. Finally, there is a single query with recall of .60 and precision of .30.

Figure 2, showing precision and recall based on either judge, does not show much clustering. Recall ranges from .15 to .57 and precision from .30 to .93.

Figures 3 and 4 provide similar information for WUIS. Figure 3 is similar to Figure 1, with nine searches more or less clustered at recall levels .27 to .63 and precision of .15 to .34. The remaining searches are much more scattered in precision, from .42 to .91, but with similar recall of .22 to .65.

In Figure 4, precision and recall based on either judge, 17 of the 19 values are concentrated at low to moderate recall levels of .13 to .38 with a range of precision of .22 to .91. The two outliers on recall, at .65 and .59, are also among the highest precision values at .80 and .95, respectively.

Query Number	Relevance based on concurrence of judges		Relevance based on either judge	
	Precision	Recall	Precision	Recall
2	.64	.31	.30	.26
2	.64	.22	.88	.22
3	.28	.22	.80	.18
4	.28	.62	.88	.57
8	.73	.45	.73	.18
6	.93	.64	.93	.42
7	.93	.38	.48	.38
8	.88	.22	.88	.22
9	.37	.31	.30	.31
10	.33	.27	.37	.19
11	.37	.60	.30	.10
18	.31	.60	.41	.18
4	.39	.45	.88	.31
17	.37	.23	.48	.18
15	.39	.38	.54	.25
16	.88	.60	.73	.31
17	.73	.38	.66	.21
17	.64	.64	.72	.51
19	.41	.49	.55	.34
Pooled	.56	.38	.65	.29

Table 4. Precision and Recall Ratios: TR Bibliographic Database

Query Number	Relevance based on concurrence of judges		Relevance based on either judge	
	Precision	Recall	Precision	Recall
1	.66	.38	.68	.28
2	.62	.25	.66	.18
3	.32	.39	.37	.22
4	.78	.59	.95	.59
5	.81	.22	.82	.21
6	.80	.65	.80	.65
7	.91	.36	.91	.35
8	.82	.24	.83	.24
9	.34	.36	.36	.36
10	.17	.38	.22	.15
11	.31	.63	.44	.38
12	.23	.33	.33	.20
13	.51	.42	.51	.32
14	.20	.27	.29	.12
15	.25	.56	.33	.16
16	.15	.43	.32	.13
17	.42	.47	.54	.18
18	.67	.40	.73	.34
19	.29	.38	.45	.27
Pooled	.53	.39	.60	.27

Table 5. Precision and Recall Ratios: WUIS

Precision and Recall Ratios
TR Bibliographic Database: Concurrence of Judges

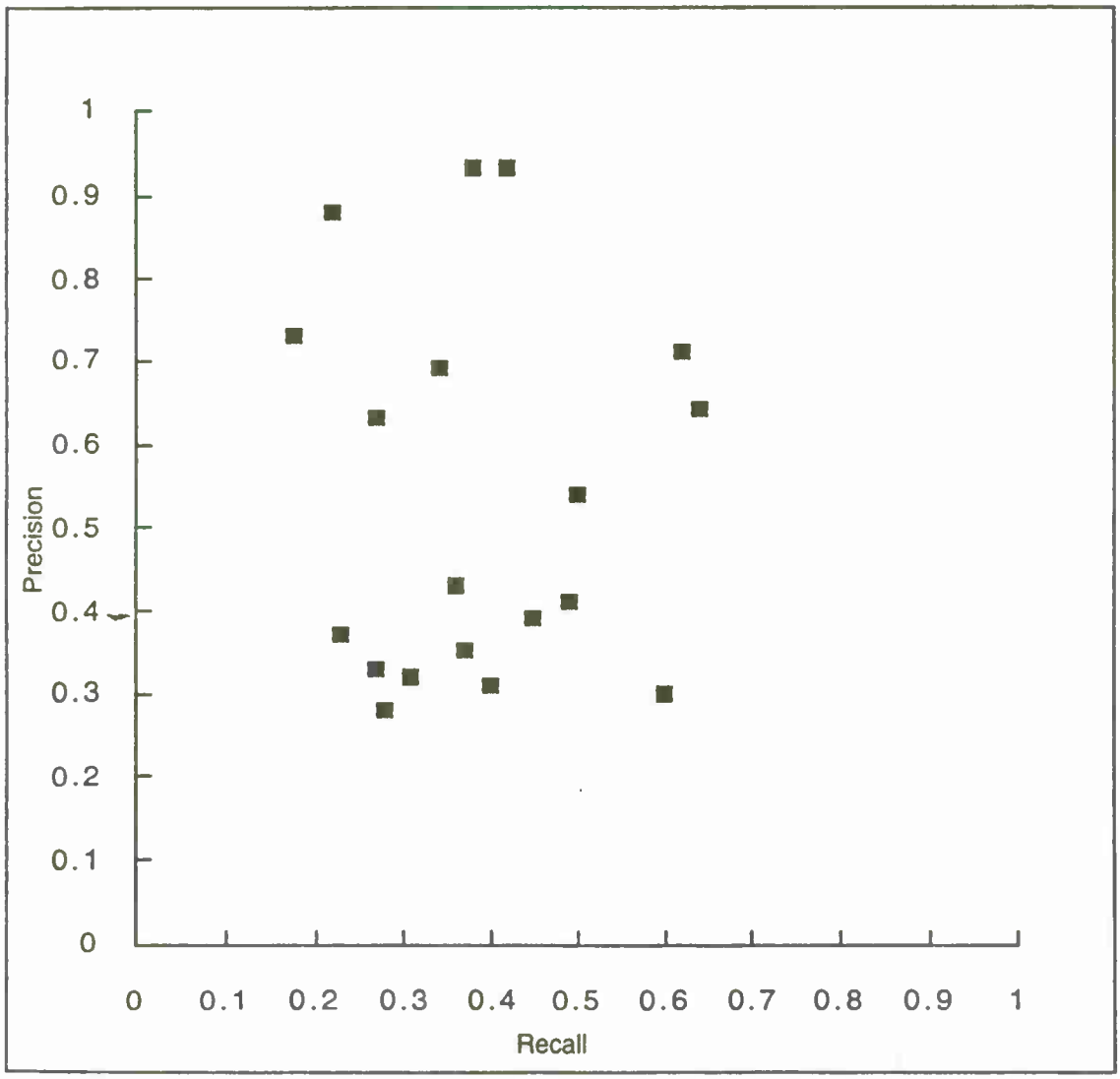

Figure 1

Precision and Recall Ratios
TR Bibliographic Database: Either Judge

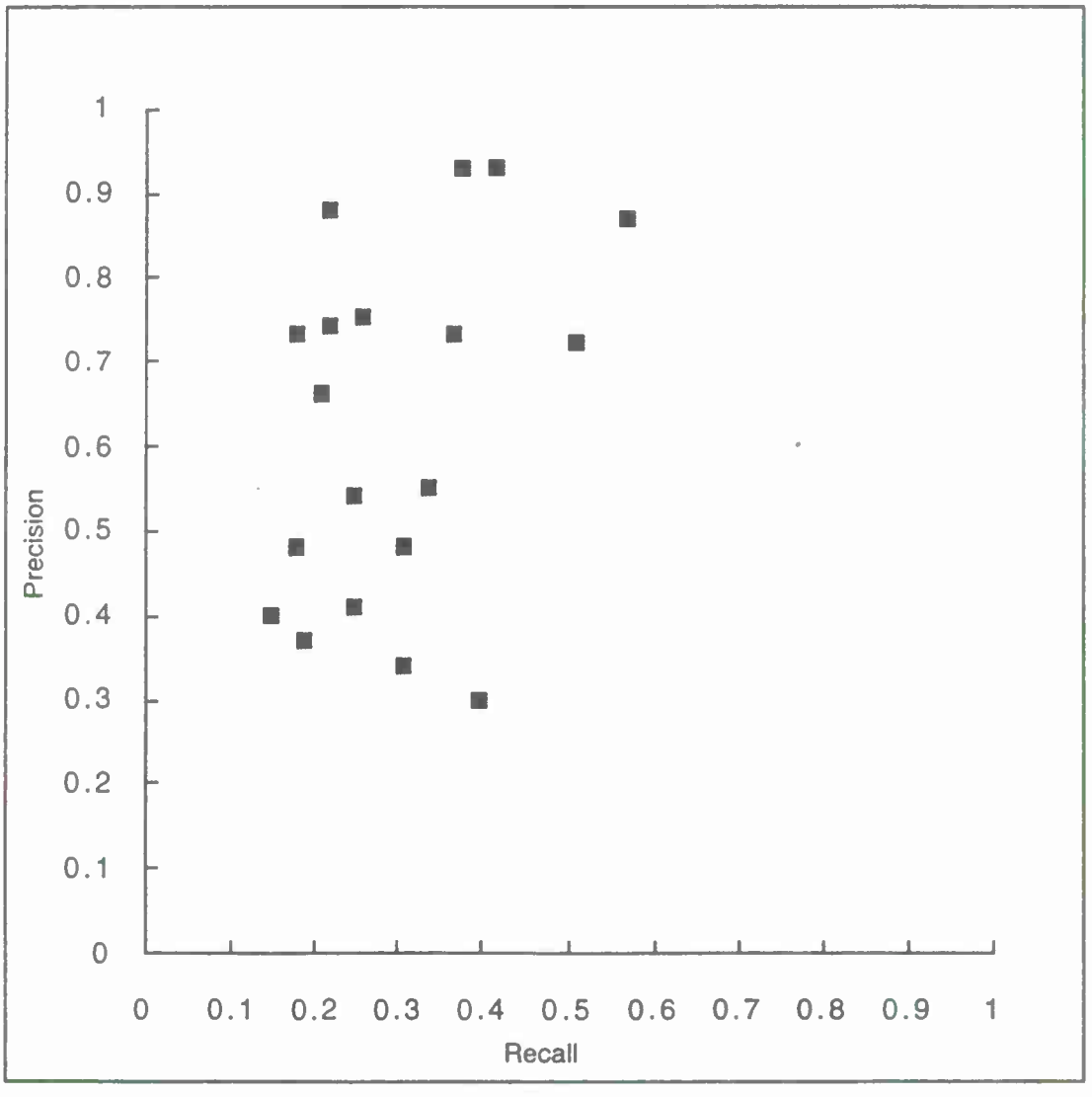

Figure 2

Precision and Recall Ratios
WUIS: Concurrence of Judges

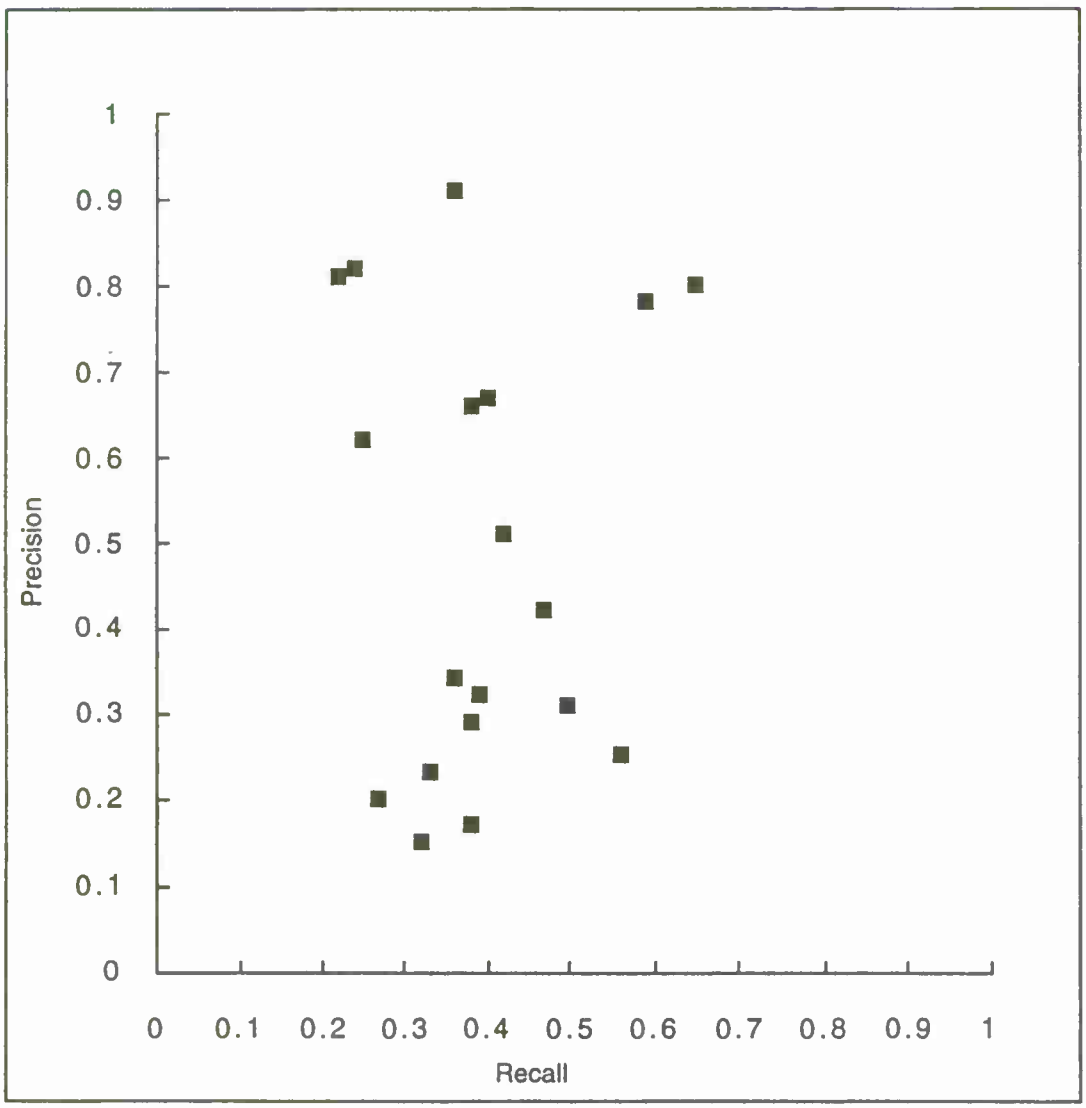

Figure 3

Precision and Recall Ratios
WUIS: Either Judge

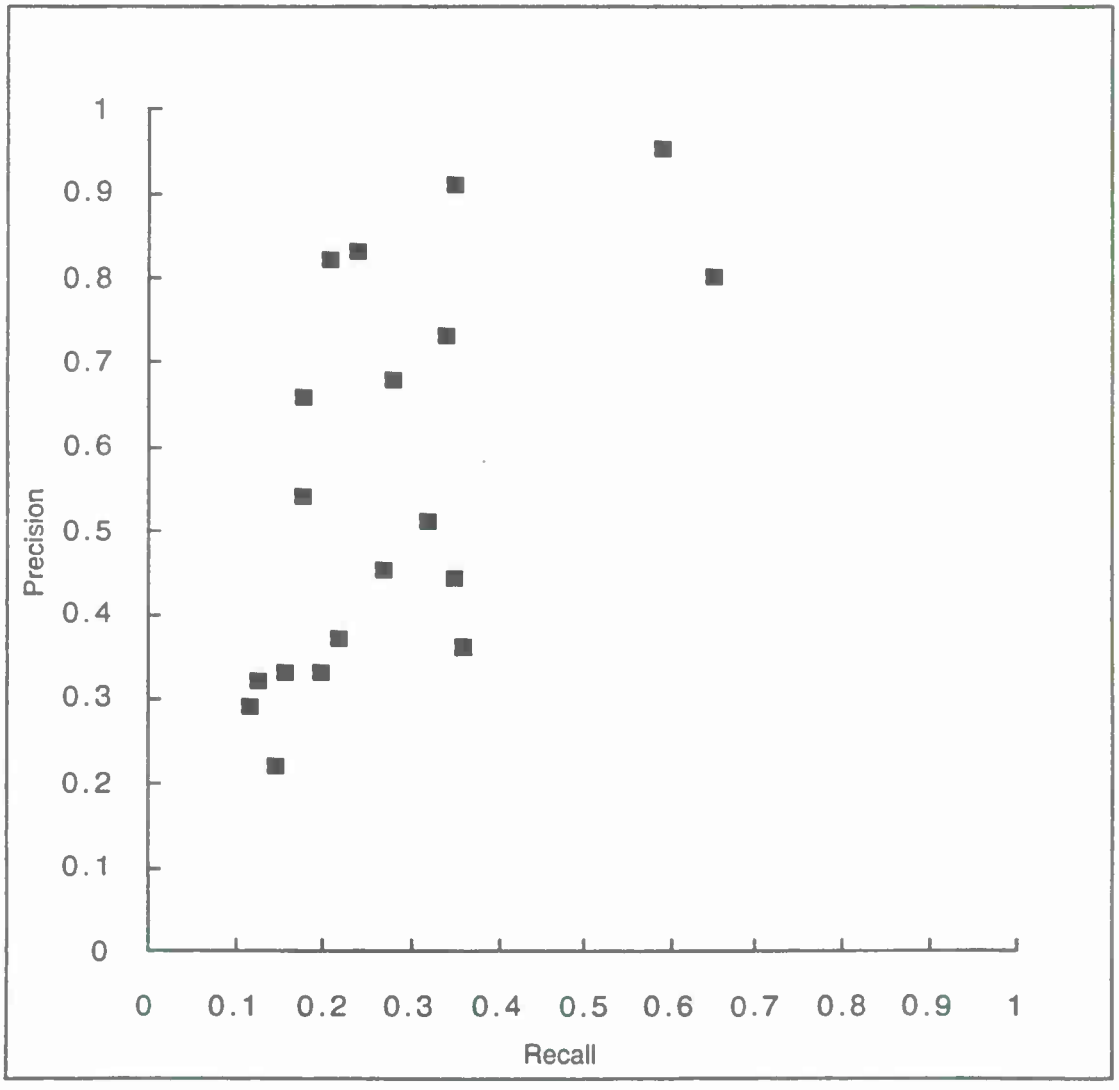

Figure 4

It is interesting to examine an example of the factors affecting precision and recall by different searchers on the same query. Query 6 in the TR Bibliographic Database was selected as an illustration. Recall and precision by concurrence of judges and by either judge were identical. Of 95 unique hits, 91 were judged relevant.

Searcher A: %collision avoid iacs=%collision avoid and %warning sys iacs=%warning sys	4 hits P=.25 (of 4 hits, 1 was relevant) R=.01 (of 91 relevant in sample, 1 retrieved)
Searcher F: iacs=%collision avoid iacs=%aircraft collision av iacs=%airborne collision collision avoidance %collision avoidance sys %aircraft collision av %airborne collision av %aircraft cas(collis	86 hits P=.93 (of 86 hits, 82 were relevant) R=.90 (of 91 relevant in sample, 82 retrieved)
Searcher 1: %collision avoidance %collision warning and %aircraft %airplane	29 hits P=.93 (of 29 hits, 27 were relevant) R=.30 (of 91 relevant in sample, 27 retrieved)
Searcher 6: collision avoidance and ?60%system	45 hits P=.93 (of 45 hits, 42 were relevant) R=.46 (of 91 relevant in sample, 42 retrieved)

Searcher F achieved both high recall and high precision on this search, apparently as a result of searching on both DTIC and IAC terms, and of using key terms beginning with the words "aircraft..." and "airborne..." that are not in the thesaurus, but may have been assigned as identifiers. Use of "AND" logic reduced recall by the other searchers. There are other factors, however, such as use or nonuse of truncation (the "%" sign) that may also have affected recall.

The small number of searcher A's hits that were included in the sample implies that one cannot draw firm conclusions about the low precision and recall of this search; precision of the other searchers was uniformly high.

In Phase Two of this study, individual citations will be examined to determine why relevant citations were not retrieved or nonrelevant citations were retrieved by each search of a

query. This retrieval failure analysis will provide specific information about the factors in indexing that facilitate or hinder retrieval. Originally it was intended to conduct a preliminary failure analysis as part of Phase One. The delays and difficulties encountered in data gathering made this impractical, and all failure analysis was rescheduled for Phase Two.

Conclusions

The queries which were searched for this study were relatively simple ones. They were limited to subject information, and usually were on relatively broad topics. Yet the variation in search strategies was great, not so much in actual choice of concepts to be searched as in refinements such as use of truncation, inclusion of IAC terms, and searching of words in titles or abstracts.

The design of the indexing scheme is implicated in some of this variation. For instance, the need for explicit inclusion of IAC-assigned terms in a search can lead to some loss of information, because some IAC-originated documents have no terms in the DTIC-assigned indexing field (23), and searchers may neglect to include the IAC term field (44) in a search. Similarly, hierarchy searching can only be as good as the hierarchies themselves. On the other hand, truncation capabilities and the necessity of using a separate step (the @qsrtab@ or @srtab@ command) to search abstracts are features of the search system.

Since this study was intended to evaluate indexing, not searching or searchers, the overall average precision/recall ratios for each query are more significant than the ratios for individual searchers. The most interesting point about these ratios is their failure to show the conventional inverse relationship between recall and precision.

Another striking point is that, even though the method used to determine the base of relevant documents was necessarily quite limited, none of the queries had a very high mean recall. The relevant documents in this study are a subset of those that were retrieved by the searches carried out for the study; there was no feasible way to determine how large a proportion this subset is of all the documents in the database which are relevant to a query. That is, we do not know how many relevant documents were not retrieved by any of the searchers. Despite this limitation, which would bias the results toward higher apparent recall than was actually the case, the mean recall was not extremely high for any of the queries.

However, inspection of the search strategies shows that searchers, even though they frequently searched on the same terms, used a number of different capabilities of the search system — but did not use them all, even when asked to search comprehensively. The number of different permutations of hierarchy, truncation, and different subject term fields, plus the unavailability of narrative text fields in the TR Bibliographic Database, except for separate qualification searching, make it difficult to devise the optimum strategy.

Even though specifics cannot be determined at this time, it is reasonable to conclude that improvements to both indexing and the search engine are warranted. The difficulties

encountered in downloading indicate that there are problems somewhere in the telecommunications system as well. One improvement in searcher training can also be suggested: that descriptor searches regularly take account of field 44 (IAC-assigned terms), to assure that IAC-originated documents are retrieved whenever appropriate.

Phase Two will provide more definitive answers as to the causes of the retrieval inadequacies encountered, at least for the documents which were actually retrieved by at least one searcher. Those which remain unknown, of course, cannot be evaluated. In Phase Two, a sample of retrieved documents will be evaluated for each query. The reasons for failure of a given search to retrieve a relevant document, or for its retrieval of a nonrelevant document, will be recorded and categorized. It is anticipated that there will be two broad categories — indexing failures and search failures — with subcategories such as vocabulary inadequacy (indexing) or failure to truncate (search).

While these analyses will contribute to the long-run goal of developing the new system, they should also suggest some concrete improvements to indexing and searching that could be included in training programs for use of the present system.

Also in Phase Two, a philosophy of subject indexing will be developed. This effort will involve extensive consultation with units such as subject analysis within DTIC, as well as with users of the DTIC databases.

References

Froehlich, Thomas J. "Relevance Reconsidered -- Towards an Agenda for the 21st Century: Introduction to Special Topic Issue on Relevance Research." *Journal of the American Society for Information Science* 45(3):124-135, April 1994.

Jacobs, Richard. Telephone conversation, July 10, 1990.

Klingbiel, Paul H. "Machine-Aided Indexing of Technical Literature." *Information Processing and Management* 9:79-84, February 1973.

Klingbiel, Paul H. "Phrase Structure Rewrite Systems in Information Retrieval." *Information Processing and Management* 21(2):113-126, 1985.

Klingbiel, Paul H.; Rinker, C.C. "Evaluation of Machine-Aided Indexing." *Information Processing and Management* 12(6):351-366, 1976.

Saracevic, Tefko, et al. "Experiments on the Cognitive Aspects of Information Seeking and Retrieving." Final Report for NSF Grant IST-8505411. Springfield, VA: National Technical Information Service, 1987; Washington, DC: Educational Research Information Center, 1987. (PB87-157699; ED 28153).

Saracevic, Tefko, et al. "A Study of Information Seeking and Retrieving." *Journal of the American Society for Information Science* 39(3):161-216, May 1988.

Appendix 1
Queries in the Sample

CSERIAC

1. Flight control and instrumentation: including displays and related topics.

2. +Gz acceleration: electrophysiological effects of cardiac arrhythmias and dysrhythmias on subject undergoing +Gz acceleration.

3. Technologies to mitigate physiological effects of fatigue: Non-pharmacological and non-invasive methods to aid in sleep promotion or induction. Specifically auditory or deep muscle relaxation strategies for sleep or reduction in environmental distractions.

4. Helmet mounted displays, including Kaiser-Agile Eye, Polhemus-Magnatrak, General Dynamics-Falcon Eye, Electro optics Industry-Knight's Eye, and Elbit's-Dash

GACIAC

5. Passive ranging from passive sensors in an airborne environment

6. Collision avoidance systems

7. Neural networks in automatic target recognition

8. Navigation and guidance, target detection, range and position finding

9. Navigation detection countermeasures, including radar countermeasures, optical detection & detectors, infrared detection & detectors.

IRIA

10. Infrared surveillance, wide area surveillance, clutter, ground targets, camouflage, weighted-difference algorithms, dual band IR, color ratios, multi-band IR

11. Infrared detectors, especially for drug and narcotics detection

12. Visual models: visual detection and acquisition of target of military vehicles (ground); camouflage (visual) of military vehicles; color contrast; motion detection (visual detection).

13. Night vision, thermal imagery, electro optical, FLIR, forward looking infrared, sighting devices, day/night sight, night vision.

14. Beach reconnaissance: Minefield detection by means of imaging, including automatic target recognition.

MTIAC

15. Fastener coating - IVD: coatings, aluminum, IVD aluminum coatings; ion vapor deposition; coatings, metal; corrosion protection; fastener.

16. Grenade Assembly: M77 grenade assembly;weapons systems assembly process; safety pin removal process in grenades for weapons systems.

17. Neural networks in manufacturing: Neural networks in composites manufacturing; including reports dealing with neural net algorithms, learning techniques, implementation issues, integration of neural nets within expert systems & different application areas, especially in manufacturing.

18. Work measurement, process improvement, etc. How to measure work or processes, etc.

SURVIAC

19. Radar warning receivers and their use by helicopters and fixed wing aircraft.

Appendix 2
Introductory Materials

Introduction to the Study

The purpose of this study is to determine the baseline quality of DTIC indexing by studying retrieval from the DTIC databases. The findings will be used to make any improvements to the indexing system which may be warranted. The study is designed as a measurement of recall and precision,[1] and will consist of two phases: a data acquisition phase and a test phase.

In the data acquisition phase, a representative group of questions will be selected from the DTIC query log. A group of experienced DTIC searchers will search the questions exhaustively, attempting to locate as many relevant citations as possible, without considering how many irrelevant citations may be retrieved. The 10-year default will be accepted, so that records from the past 10 years will be retrieved. This will limit retrieval to the relatively recent past, while assuring that a reasonably broad range of material is retrieved.

The purpose of this phase is to approach as closely as possible the retrieval of all records in the databases which are relevant to a given query and for which an unclassified document is retrievable. Since it is not practical to evaluate every record in the database for relevance to each query, this approach of exhaustive searching will make it possible to have a large base for comparison with retrieval in the test phase. The relevant records retrieved in the data acquisition phase, plus any additional ones found in the test phase, will be treated as if they were all of the relevant records to be found in the database. This procedure is typical of those followed in studies of recall and precision on large databases.

In the test phase, experienced search intermediaries from the DTIC user community will perform more typical searches on the same questions, attempting to produce results that might be useful for the requester — i.e., with a reasonable balance between recall and precision in retrieval.

Neither group of searchers will carry out any post-processing or evaluation of individual citations for relevance; the entire final retrieved set will be submitted for relevance evaluation. Subject-matter experts based at DTIC IAC's will evaluate the relevance to the query of each record retrieved, permitting precision and recall ratios to be determined for the test searches.

The failures of recall (relevant records not retrieved) and precision (irrelevant records retrieved) in the test phase will be analyzed to determine the reasons for the failures. Those

---

[1] Precision is the percentage of documents retrieved which is actually relevant to the query. Recall is the percentage of the total relevant documents in a collection which is actually retrieved by a query.

failures which represent indexing problems will then become the subject of recommendations for improvement of the indexing system in the next phase of the effort.

While multiple searchers will be searching each question, the purpose of the study is not to examine searcher performance. For most of the analysis, the work of all searchers for a given question in each phase will be amalgamated; even when retrieval of individual searchers is compared, the purpose will be to isolate searcher differences as a variable for statistical analysis.

Conducting Searches
Data Acquisition Phase

The purpose of this phase is to gather data comprising as many as possible of the relevant documents to be found in the database, even at the price of retrieving a large number of irrelevant citations. These relevant documents will serve as the basis for determining the recall ratios of more typical searches, conducted during the test task. The indexing system is being evaluated in this study, not searcher performance.

Please conduct the searches in the order designated. These are all queries which have been submitted to DTIC and searched by DTIC searchers in the recent past. All of the available information about the query has been provided, but not the search strategy which was used in the original search. Since the original search was conducted under uncontrolled conditions, it is not relevant to the study, and it should not be permitted to influence the design of your strategy.

This is not a test of searching speed, and you may not finish all the searches. The order of the searches has been randomized, so that each searcher searches the queries in a different order, but each query will be searched at least three times if each searcher searches 10 or more queries. Ignore the number beside each query; it is there for coding purposes only.

For each query, devise a broad search strategy on the general topic, designed for maximum recall. You may refine and reformulate the strategy as much as you wish, until you are convinced that you are probably retrieving as many as possible of the documents in the database which are relevant to the query. Search both the Technical Report (TR) and Work Unit (WUIS) databases,[1] modifying the strategy as appropriate to maximize relevant retrieval from each. Limit retrieval to unclassified limited documents — i.e., documents for which an unclassified abstract is available, and use the 10-year default to retrieve documents only from the past 10 years.

When the search is complete, print your results in reverse order by date (i.e., latest date first), together with the log of the search. Also, download the results to floppy disk. No post-processing of the results should be carried out; we are interested in the output of the search system, rather than in human ability to compensate for its precision failures. List on the query sheet any additional information you think may be helpful. In particular, since this is a test of subject indexing quality, please give your reasons for any use of full text in your search. For example, did you use full text because the concept was too new to be in the thesaurus, or because it was narrower than any available thesaurus term?

---

[1] Depending on time limitations, in the afternoon of the second day you may be asked to concentrate on the TR database, rather than searching both TR and WUIS.

Conducting Searches
Test Phase

The purpose of this phase is to retrieve from the database a set of documents that would be useful to the person posing a query. The indexing system is being evaluated in this study, not searcher performance.

You will be assigned one or more queries to be searched. Please conduct the searches in the order designated. These are all queries which have been submitted to DTIC and searched by DTIC searchers in the recent past. You will be provided all of the available information about the query, but will not be provided the search strategy which was used in the original search. Since the original search was conducted under uncontrolled conditions, it is not relevant to the study, and it should not be permitted to influence the design of your strategy.

For the queries assigned to you, please devise a search strategy designed to produce what seems to you like a reasonable balance between precision and recall. That is, you should attempt to retrieve as many documents as possible, without burdening the user with an unreasonable number of irrelevant documents. You may refine and reformulate the strategy as much as you wish, until you have spent what seems to you to be the amount of time you would spend on a typical search of this nature. Search both the Technical Report (TR) and Work Unit (WUIS) databases, modifying the strategy as appropriate to maximize relevant retrieval from each. Limit retrieval to documents for which an unclassified record is available, and use the 10-year default to retrieve documents only from the past 10 years.

Since there has been a lapse of time since the data acquisition searches were run, you will need to exclude items added to the database after the date of the data acquisition search for a given query. The sheet for each query indicates the cutoff date for that query.

When the search is complete, print your results in reverse order by date (i.e., latest date first), together with the log of the search. Also, download the results to a floppy. No post-processing of the results should be carried out; we are interested in the output of the search system, rather than in human ability to compensate for its precision failures. List on the query sheet any additional information you think may be helpful. In particular, since this is a test of subject indexing quality, please give your reasons for any use of full text in your search. For example, did you use full text because the concept was too new to be in the thesaurus, or because it was narrower than any available thesaurus term?

Relevance Judging

Determination of the relevance to the query of each citation retrieved by that query is required to permit determination of recall and precision ratios. The total number of citations retrieved for a query is sometimes very large. For judging purposes, an upper limit of 100 citations for each version (TR and WUIS) of each query has been established. When the total retrieval was above this number, a sample of 100 citations has been selected for relevance judging. The purpose of this limitation is to keep the effort required of the judges within reasonable bounds.

Please review each citation and abstract and make a decision about its relevance to the query, on a four-point scale: Relevant / Partially or probably relevant / Not relevant / Not determinable.

Relevant: A citation which bears directly on the query; one which seems very likely to contain significant information.

Partially or probably relevant: A citation which contains useful information which is not central to the query, or one which seems likely to contain useful information.

Nonrelevant: A citation which seems very unlikely to contain information of value to the query.

Not determinable: A citation for which you cannot determine from the information available whether it is likely to be relevant.

Your initials should be placed in the "Judge" blank at the top of the form.

Appendix 3

Search Strategies[1]

01 TR

Searcher A

@str@
iacs=%flight control
%flight control
$flight control
and
$flight instruments
%flight instrumen
iacs=flight instrumen
end

Searcher C

@str@
$*flight control systems
%flight display
flight instruments
and
instrumentation
$flight instruments
end

Searcher E

@str@
%flight control
$flight control systems
and
%instrument
end

Searcher 3

@str@
*flight control systems
?60flig
*?00flight
and
*flight control sytems
?60control
*?00control
and
*instrumentation
?60%instrument
*?00%instrument
end

Searcher 5

@str@
$%flight control
$%aircraft control
and
$%instrument
$%display
end

[1] When the "@swuwps@" command was used to search WUIS with the same strategy used for TR, the command is shown, but the strategy is repeated for the convenience of the reader of this report.

Searcher A

@swu@
%flight control
de=$flight control
and
de=$flight instruments
%flight instrumen
end

Searcher C

@swu@
$flight control systems
flight instruments
%flight display
and
instrumentation
$flight instruments
end

Searcher E

@swu@
%flight control
de=flight control systems
sub=flightcontrol
and
%instrument
end

Searcher 3

@swuups@
*flight control systems
?60flig
*?00flight
and
*flight control sytems
?60control
*?00control
and
*instrumentation
?60%instrument
*?00%instrument
end

Searcher 5

@swu@
sub=$%[illegible]ht control
sub=$%aircraft control
and
sub=$%instrument
sub=$%display
end

02 TR

Searcher A

@str@
arrhythmia
%arrhythmia
and
electrophysiology
end

@srtab@
acceleration
gz
end

Searcher C

@str@
+gz
%gz
end

@qsrtab@
acceleration
end

Searcher 1

@str@
$acceleration
and
%cardiac
%heart
and
%effect
?60%effect
?00%effect
end

Searcher 5

@str@
?60+gz
$%+gz
$%acceleration
and
$%cardiac
$%electrophysiology
end

02 WU

Searcher A

@swu@
arrhythmia
arrhythmias
kw=dysrhythmia during acceleration
and
de=electrophysiology
sub=electrophysiology
nar=electrophysiology
end

Searcher C

@swu@
%+gz
%gz
and
%arrhythmia
%cardiac arrhythmia
%dysrhythmia
end

Searcher D

@swu@
%arrhythmi
%dysrhythmi
and
cardi
heart
hearts
end

@srtab@
acceleration
accelerating
accelerate
accelerates
accelerated
accelerations
end

Searcher 1

@swu@
de=cardiac
de=heart
kw=%cardiac
de=%cardiac
kw=heart
%cardiac
%heart
and
$acceleration
and
de=%effect
kw=%effect
%effect
end

Searcher 5

@swu@
sub=$%+gz
sub=$%acceleration
and
sub=$%cardiac
sub=$%electrophysiology
end

Searcher C

@str@
fatigue
fatigue(physiology)
and
sleep
sleep deprivation
end

Searcher F — 3 strategies combined in user file

@str@
fatigue(physiology)
and
%sleep promot
sleep induction
muscle relaxation
deep muscle relaxation
muscle relaxants
sleepability
relaxation(physiology)
end

@str@
?60relaxation
?60%sleep
and
?60%promot
?60%induce
end

@str@
fatigue(physiology)
?60%fatigue
and
%sleep promo
%sleep induc
%muscle relax
relaxation(physiology)
end

Searcher 2

@str@
$sleep
$sleep deprivation
and
$relaxation
$muscles
$hearing
%bioacoust
$sound generators
%noise mask
?60%masking
?00%masking

Searcher 4

@str@
fatigue(physiology)
$sleep deprivation
and
$relaxation
%environmental disturbance
disturbance
disturbances
?60relaxation
?60%disturbance
?60%nonpharma
?60%non-pharm
?60%noninvasive
?60%non-invasive
$environments
?60%environment
end

Searcher C

@swu@
fatigue
fatigue(physiology)
and
slp
sleep deprivation
sleep disturbances
sleep loss
sleep patterns
sleep disorders
%sleep cycle
end

Searcher F — 3 strategies combined in user file

@swu@
fatigue(physiology)
and
%sleep promot
sleep induction
muscle relaxation
deep muscle relaxation
muscle relaxants
sleepability
relaxation(physiology)
end

@swu@
ti=relaxation
ti=%sleep
and
ti=%promot
ti=%induc
end

@swu@
fatigue(physiology)
ti=%fatigue
and
%sleep promo
%sleep induc
%muscle relax
relaxation(physiology)
end

Searcher 2

@swu@
$fatigue(physiology)
$fatigue (physiology)
and
$physiological effects
%stress
%physiolog

Searcher 4

@swu@
fatigue(physiology)
$sleep deprivation
and
$relaxation
%environmental disturbance
disturbance
disturbances
ti=relaxation
ti=%disturbance
ti+%nonpharma
ti=%nonpharma
ti=%non-pharm
ti=%noninvasive
ti=%non-invasive
$environments
ti=%environment
end

Searcher E

@str@
?00%helmet mounted display
end

Searcher F

@str@
%helmet mounted displ
%head up displ
%heads up displ
%HUD(head
%HUDS(head
iacs=helmet mounted
iacs=helmet mounted display
iacs=%helmet mounted d
iacs=%helmet-mounted di
iacs=%hud(head
iacs=%huds(head
end

Searcher 3

helmet mounted displays
?00helmet
?60helmet
and
helmet mounted displays
?00mount
?60mount
and
helmet mounted displays
?00%display
?60%display
end

Searcher 4

@str@
%helmet mounted display
%head mounted
%head-mounted
%hmd%
%hmd(head
%hmd(helmet
%agile eye
%falcon eye
end

Searcher E

@swu@
de=%helmet mounted displays
hmd
sub=helmet mounted display
kw=helmet mounted display
end

Searcher F

@swu@
%helmet mounted displ
%head up displ
%heads up displ
%hud(head
%huds(head
iacs=helmet mounted
iacs=helmet mounted display
iacs=%helmet mounted d
iacs=%helmet mounted di
iacs=%hud(head
iacs=%huds(head
end

Searcher 3

@swu@
helmet mounted displays
?00helmet
?60helmet
and
helmet mounted displays
?00mount
?60mount
and
helmet mounted displays
?00%display
?60%display
end

Searcher 4

@swu@
%helmet mounted display
%head mounted
%head-mounted
%hmd%
%hmd(head
%hmd(helmet
%agile eye
%falcon eye
end

Searcher A

@str@
$passive systems
and
$range finding
end

@srtab@
air
airborne
aerospace
end

Searcher B

@str@
$detectors
%passive ranging
%passive sensor
and
$aerospace environments
%airborne environment
end

@srtab@
passive sensors
ranging
airborne
passive sensor
end

Searcher C

@str@
passive systems
%passive
and
$detectors
%sensors
and
range finding
$
and
airborne
end

Searcher E

@str@
Passive ranging
end

Searcher 3

@str@
?00%sensor
?60%sensor
and
airborne
?60airborne
end

@srtab@
passive
end

Searcher 5

@str@
$%passive ranging
$%passive sensor
and
$%airborne
$%aircraft
end

Searcher A

@swu@
%passive system
de=passive systems
and
$range finding
%range find
end

@srtab@
air
airborne
aerospace
end

Searcher B

@swu@
$detectors
%passive ranging
%passive sensor
and
$aerospace environments
%airborne environment
end

@srtab@
passive sensors
ranging
airborne
passive sensor
end

Searcher C

@swu@
passive systems
%passive
and
$detectors
%sensors
and
$range finding
and
airborne
end

Searcher E

%passive ranging
%passive sensors
end

Searcher 3

@swu@
kw=passive
sub=passive
%passive ranging
and
%passive ranging
kw=ranging
sub=ranging
and
kw=airborne
sub=airborne
%airborne
end

Searcher 5

@swu@
sub=$%passive ranging
sub=$%passive sensor
and
sub=$%airborne
sub=$%aircraft
end

Searcher A

@str@
%collision avoid
iacs=%collision avoid
and
%warning sys
iacs=%warning sys
end

Searcher B

@str@
collision avoidance systems
collision avoidance
and
collision
collision avoidance
end

Searcher F — 3 searches combined in user file

@str@
?60avoidance
and?
60%collision
and
?60system
?60systems
end

@str@
collision avoidance
%collision avoidance sys
%aircraft collision av
%airborne collision av
%aircraft cas(collis
end

@str@
?60cas
end

@qsrtab@
collision
collisions
collision avoidance
collide
collided
aircraft collision
aircraft collisions
airborne collision
airborne collisions
end

Searcher 1

@str@
%collision avoidance
%collision warning
and
%aircraft
%airplane
end

Searcher 6

@str@
collision avoidance
and
?60%system
end

Searcher A

@swu@
%collision avoid
and
%warning sys
%proximity warn
end

Searcher B

@swu@
collision avoidance systems
collision avoidance
avoidance
and
collision
collision avoidance
end

Searcher F — 3 searches combined in user file

@swu@
%collision avoid
%aircraft collision av
%airborne collision ab
%airborne collision av
%CAS(collision
end

@swu@
ti=avoidance
ti=avoiding
ti=avoid
ti=avoided
and
ti=collision
ti=collide
ti=collides
ti=collided
ti=colliding
and
ti=aircraft
ti=airborne
ti=helicopter
ti=helicopters
ti=gunship
ti=gunships
ti=tanker
ti=tankers
ti=refuel
ti=refueling
end

@swu@
ti=cas
sub=cas
sub=collision avoidance
end

Searcher 1

@swu@
%collision avoidance
%collision warning
%aircraft guidance system
de=%collision avoidance
kw=%collision avoidance
de=%collision warning
kw=%collision warning
de=%aircraft guidance system
kw=%aircraft guidance system
end

Searcher 6

@swu@
collision avoidance
end

07 TR

Searcher A

@str@
%neural net
iacs=%neural net
and
%target recog
%target det
%automatic target recog
%automatic target det
iacs=%target recog
iacs=%target det
iacs=%automatic target recog
iacs=%automatic target det
end

Searcher E

%neural network
and
$target recognition
end

Searcher 5

@str@
$%automatic target
and
$%neural network
$%artificial intelligence
end

Searcher 6

@str@
neural nets
and
target recognition
end

Searcher A

@swu@
%neural net
and
%target recog
%target de
%automatic target recog
%automatic target de
end

Searcher B

@swu@
$networks
neural nets
and
$networks
neural nets
and
target recognition
automatic target recognition
end

Searcher D — 2 searches combined in user file

@swu@
neural
and
net
nets
%network
end

@srtab@
neural net
neural nets
neural networking
and
atr
automatic target
end

@swu@
%neural net
and
%atr(auto
%automatic target recog
end

Searcher E

%neural network
and
$target recognition
end

Searcher 5

@swu@
sub=$%automatic target
and
sub=$%neural network
sub=$%artificial intelligence
end

Searcher 6

@swuups@
neural nets
and
target recognition
end

Searcher B

@str@
*navigation
*target detection
*position finding
*range finding
and
*guidance
*target detection
*position finding
*range finding
end

@srtab@
navigation and guidance
target detection
range detection
position finding
range finding
end

Searcher C

@str@
$*targets
target detection
and
$*detection
target detection
and
$range finding
$position finding
and
$navigation
$guidance
end

Searcher D — 2 strategies combined in file

@str@
$range(distance)
and
$position(location)
AND
$guidance
and
$navigation
and
$targets
%target detect
and
%target detect
$detection
end

@str@
iacs=%guid
and
iacs=%navigat
and
iacs=%target
and
iacs=%detect
end

@srtab@
position
positioning
and
guidance
guided
guiding
end

Searcher 4

@str@
$navigation
$guidance
and
target detection
and
$range finding
$position finding
end

Searcher 5

@str@
$%target detection
and
$%navigation
$%guidance
$%range
$%position finding
end

Searcher B

@swu@
navigation
target detection
position finding
range finding
and
guidance
end

@quftab@
navigation and guidance
target detection
range finding
position finding
end

Searcher C — 2 strategies combined in user file

$*targets
target detection
and
$*detection
target detection
and
$range finding
$position finding
and
$navigation
$guidance
end

@swu@
$targets
target detection
and
$detection
target detection
and
$range finding
$position finding
and
$navigation
$guidance
end

Searcher D — 2 strategies combined in user file

@swu@
$range(distance)
and
$position(location)
and
$navigation
and
$guidance
and
$targets
%target detect
and
%target detect
$detection
end

@swu@
range
ranging
and
positioning
end

@srtab@
target
targets
and
navigation
navigating
and
guidance
guiding
end

Searcher 4

@swuups@
$navigation
$guidance
and
target detection
and
$range finding
$position finding
end

Searcher 5

@swu@
sub=$%target detection
and
sub=$%navigation
sub=$%guidance
sub=$%range
sub=$%position finding
end

Searcher A

@str@
radar countermeasures
iacs=radar countermeasures
$optical countermeasures
iacs=optical countermeasures
and
%optical detect
$optical detection
iacs=optical detection
iacs=optical detectors
end

Searcher C

@str@
$navigation
and
$detection
$detectors
$optical detectors
and
$countermeasures
$electronic countermeasures
end

Searcher E

@str@
$navigation
and
$radar countermeasures
$optical detection
$optical detectors
end

Searcher 2 — 2 strategies combined in user file

@str@
$navigation
$navigation aids
$iff systems
and
$detection
$detectors
and $countermeasures
end

@str@
$navigation
$navigation aids
$iff systems
and
$radar
$sonar
and
$electronic warfare
$infrared countermeasures
end

Searcher 4 — 2 strategies combined in user file

@str@
$navigation
and
$detection
and
$countermeasures
end

@str@
radar countermeasures
$optical detection
$optical detectors
%detection countermeasure
and
$navigai
$navigation
?60navigation
end

Searcher A

@swu@
nar=navigation
and
nar=%detect
and
nar=%countermeasure
end

Searcher C

@swu@
$navigation
and
$detection
$detectors
$optical detectors
and
$countermeasures
$electronic countermeasures
end

Searcher E

@swu@
$navigation
and
$radar countermeasures
$optical detection
$optical detectors
end

@qsrtab@
navigation
radar countermeasures
end

Searcher 2 — 2 strategies combined in user file

@str@
$navigation
$navigation aids
$iff systems
and
$detection
$detectors
and $countermeasures
end

@str@
$navigation
$navigation aids
$iff systems
and
$radar
$sonar
and
$electronic warfare
$infrared countermeasures
end

Searcher 4 — 2 strategies combined in user file

@swu@
$navigation
and
$detection
and
$countermeasures
end

@swu@
radar countermeasures
$optical detection
$optical detectors
%detection countermeasure
and
$navigation
ti=navigation
end

Searcher D — 2 strategies combined in user file

@str@
%weighted difference alg
%color ratio
end

@str@
%surface target
$vehicles
%vehicle
and
$deception
%camouflage
$decoys
%decoy
and
%ground clutter
%clutter
end

@srtab@
ir
infrared
infra-red
end

Searcher E

@str@
?00infrared surveillance
%wide area surveillance
?00%camouflage
%flir
and
ground targets
clutter
end

Searcher F

@str@
%ir surveil
%infrared survei
%wide area surv
%wass(wide
?60wass
?00clutter
%ir clutter
%ir radar clutter
%infrared clutter
%infrared radar clutter
%multiband ir
%multiband infrared
%ir detection
%ir detector
%infrared detect
%dual band ir
%dualband ir
%dual band infrared
%dualband infrared
dual band
%dual band flir
%dual band radar
daul band seeker
%dual band seeker
%dual band transmitter
and
targets
ground targets
surface tra
surface targets
radar targets
visual targets
military targets
camouflasge
camouflage
radar camouflage
?60target
?60targets
?60camouflage
?60camouflaged
?60camouflaging
and
color ratio
color ratios
algorithms
?60algorithm
?60algorithms
%weighted difference algorithm
%color radar
%color raster
%color monitor
%color image
%color dis

Searcher D

@swu@
color
and
ratio
ratios
end

@srtab@
color ratio
colors
color-ratio
color-ratios
and
target
targets
end

Searcher E

@swu@
de=infrared surveillance
%wide area surveillance
de=camouflage
%flir
and
ground targets
clutter
end

Searcher F — 2 strategies co

@swu@
%ir surveil
%infrared survei
%wide area surv
%wass(wide
ti=wass
de=clutter
%ir clutter
%ir radar clutter
%infrared clutter
%infrared radar clutter
%multiband ir
%multiband infrared
%ir detection
%ir detector
%color com
%color cont
%color crt
%color cons
%color discrim
end

Searcher 5 — this search had over 13000 hits; no downloading was attempted

@str@
$%infrared surveillance
infrared surface search and surveillance
systems
$%wide area surveillance
$%clutter
$%ground target
$%camouflage
algorithms
$%color ratio
$%multiband infrared
end

Searcher 6

@str@
infrared surveillance
%wide area surveillance
and
?60ground
surface targets
%ground target
and
?60%target
surface targets
%ground target
end

Searcher D

@swu@
color
and
ratio
ratios
end

@srtab@
color ratio
colors
color-ratio
color-ratios
and
target
targets
end

Searcher E

@swu@
de=infrared surveillance
%wide area surveillance
de=camouflage
%flir
and
ground targets
clutter
end

Searcher F — 2 strategies combined in file

@swu@
%ir surveil
%infrared survei
%wide area surv
%wass(wide
ti=wass
de=clutter
%ir clutter
%ir radar clutter
%infrared clutter
%infrared radar clutter
%multiband ir
%multiband infrared
%ir detection
%ir detector
%infrared detect
%dual band ir
%dual band ir
%dualband ir
%dual band infrared
%dualband infrared
dual band
%dual band flir
%dual band radar
%dual band seeker
%dual band transmitter
and
targets
ground targets
surface targets
radar targets
visual targets
military targets
caamouflage
camouflage
radar camouflage
ti=target
ti=targets
ti=camouflage
ti=camouflaged
ti=camouflaging
and
color ratio
color ratios
algorithms
ti=algorithms
%weighted differance algorithm
%color radar
%color raster
%color monitor
%color image
%color dis
%color com
%color crt
%color cons
%colordiscrim
%color discrim
end

@swu@
ti=wassn
not
personnel management
end

Searcher 5 — this searcher had over 10,000 hits; no downloading was attempted

@swu@
sub=$%infrared surveillance
sub=infrared surface search and surveillance systems
sub=$%wide area surveillance
sub=$%clutter
sub=$%ground target
sub=$%camouflage
sub=algorithms
sub=$%color ratio
sub=$%multiband infrared
end

Searcher 6

@swu@
infrared surveillance
%wide area surveillance
end

@srtab@
target
targets
end

11 TR

Searcher A

@str@
$optical detectors
and
%infrared
and
%drug
%narcotic
$drugs
end

Searcher B

@str@
infrared detection
infrared detectors
infrared images
and
drugs
drug interdiction
drug smuggling
infrared detectors
end

@srtab@
drug
drugs
interdiction
end

Searcher F

@str@
infrared detection
%infrared detect
%ir detect
and
drugs
%drug detect
?60drug
?60drugs
narcotics
?60narcotic
?60narcotics
end

Searcher 1

@str@
%ir(infrared
%infrared
and
%detect
$detection
and
$drugs
?00drugs
?00narcotics
?61drugs
?61narcotics
end

Searcher 5

@str@
$%infrared detect
and
$%drug
$%narcoterrorism
$%narcotic
end

11 WU

Searcher A

@swu@
de=$optical detectors
and
sub=%drug
nar=%drug
sub=%narcotic
nar=%narcotic
de=drugs
and
sub=%infrared
nar=%infrared
end

Searcher B

@swu@
infrared detection
infrared detectors
infrared images
and
drugs
drug interdiction
drug smuggling
end

Searcher F

@swu@
infrared detection
%infrared detect
%ir detect
and
drugs
%drug detect
ti=drug
ti=drugs
narcotics
ti=narcotic
ti=narcotics
end

Searcher 1

@swu@
de=%in(infrared
de=%infrared
and
de=$detectors
and
de=%drug
de=$drugs
end

Searcher 5

@swu@
sub=$%infrared detect
and
sub=$%drug
sub=$%narcoterrorism
sub=$%narcotic
end

12 TR

Searcher A

@str@
target acquisition
%target acquisition
and
computerized simulation
mathematical models
iacs=computerized simulation
%visual models
%visual detect
and
$ground vehicles
end

Searcher B

@str@
visual perception
cammouflage
visual detection
and
detection
visual perception
and
$military vehicles
%military vehicles
end

Searcher F

@str@
%military vehicle
%combat vehicle
and
color contrast
%motion detect
%visual acq
camouflage
visual camouflage
?60%detect
?60camouflage
visual model
end

Searcher 4

@str@
%visual acquisition
%visual detection
vision
visual perception
visual surveillance
visual targets
camouflage
coloring
colors
motion
ground speed
and
%visual acquisition
%visual detection
detection
target detection
$optical detection
target acquisition
target discrimination
%motion detection
and
$military vehicles
end

Searcher 5

@str@
$%visual acquisition
$%visual detection
$%visual target acquisition
$%visual target detection
$%camouflage (visual
$%visual camouflage
$%color contrast
$%motion detection
and
$%military ground
$%military vehicle
$%ground vehicle
$%vehicle
$%combat vehicle
end

12 WU

Searcher A

@swu@
target acquisition
%target acq
and
computerized simulation
iacs=computerized simulation
algorithms
%visual model
%visual detect
$mathematical models
and
$ground vehicles
end

Searcher B

@swu@
visual perception
camouflage
and
detection
and
$military vehicles
%military vehicle
end

Searcher F

@swu@
%military vehicle
%combat vehicle
and
color contrast
motion detection
%motion detect
%visual acq
camouflage
visual camouflage
ti=%detect
ti=camouflage
visual model
end

Searcher 4

@swuwps@
%visual acquisition
%visual detection
vision
visual perception
visual surveillance
visual targets
camouflage
ccloring
colors
motion
ground speed
and
%visual acquisition
%visual detection
detection
target detection
$optical detection
target acquisition
target discrimination
%motion detection
and
$military vehicles
end

Searcher 5

@swu@
sub=$%visual acquisition
sub=$%visual detection
sub=$%visual target acquisition
sub=$%visual target detection
sub=$%camouflage (visual
sub=$%visual camouflage
sub=$%color contrast
sub=$%motion detection
and
sub=$%military ground
sub=$%military vehicle
sub=$%ground vehicle
sub=$%vehicle
sub=$%combat vehicle
end

13 TR

Searcher A

@str@
night vision
night vision devices
and
thermal images
%thermal imag
iacs=%thermal imag
%flir
forward looking
%forward looking
end

Searcher B

@str@
night vision
night sights
night vision devices
and
thermal images
forward looking infrared systems
flir
thermal imagery
electrooptical photography
end

Searcher C

@str@
%night vision
and
%flir
forward looking infrared radar
%forward looking infrared (flir
infrared images
%thermal imager
thermography
electrooptics
electrooptical photography
end

Searcher 3

@str@
night vision
night vision devices
%day/night
%night sight
and
thermal images
thermal imagery
electron optics
?60electrooptical
?00electrooptical
forward looking infrared systems
?00flir
?60flir
end

Searcher 4

@str@
%night sight%
%night seeing
night vision
and
%thermal imag
%electrooptic
%electrooptical
%electro-optic
%flir%
%flir(forward
%forward looking i
and
$sights
%sighting
%night vision devices
%flir%
%flir(forward
%forward looking i
end

13 WU

Searcher A

@swu@
de=night vision
de=night vision devices
and
%flir
de=forward looking
%forward looking
end

Searcher B

@swu@
night vision
night sights
night vision devices
thermal images
forward looking infrared systems
flir
thermal imagery
electrooptical photography
end

@srtab@
thermal imagery
forward looking infrared
electrooptical
flir
sighting devices
and
night vision
end

Searcher C

%flir
forward looking infrared radar
%forward looking infrared (flir
infrared images
%thermal imager
thermography
electrooptical photography
and
night vision
end

Searcher 3

@swuwps@
night vision
night vision devices
%day/night
%night sight
and
thermal images
thermal imagery
electron optics
ti=electrooptical
sub=electrooptical
forward looking infrared systems
sub=flir
ti=flir
end

Searcher 4

@swuwps@
%night sight%
%night seeing
night vision
and
%thermal imag
%electrooptic
%electrooptical
%electro-optic
%flir%
%flir(forward
%forward looking i
and
$sights
%sighting
%night vision devices
%flir%
%flir(forward
%forward looking i
end

Searcher A

@str@
$detection
iacs=%detect
and
$mines(ordnance)
and
%imag
end

Searcher B

@str@
beach reconnaissance
minefield detection
beaches
minefields
and
target detection
imaging detection
imaging devices
infrared images
image processing
target recognition
beach reconnaissance
end

Searcher D — 4 strategies combined in user file

@str@
?60%beach
?60shore
?60shores
?60littoral
?60surf
and
?60reconnaissance
end

@str@
$coastal regions
and
$mines(ordnance)
$mine warfare
and
$detection
end

@str@
%coastal
coast
coasts
%coastline
%littoral
?00%surf
and
$mines(ordnance)
$mine warfare
and
$detection
end

@str@
$shores
and
$mines(ordnance)
%minefield
$mine warfare
end

@srtab@
detect
detects
detected
detection
detecting
end

Searcher E

@str@
%minefield detection
%minefield breach
end

@qsrtab@
beach
reconnaissance
imaging
end

Searcher 3

@str@
*mine detection
?60minefield
?00minefield
and
*mine detection
?60%detect
?00%detect
end

@srtab@
beach
shore
beaches
shores
end

Searcher 6

@str@
mine detection
mine detectors
minefields
and
$images
target recognition
target detection
and
shores
beaches
beach heads
end

14 WU

Searcher A

@swu@
de=detection
%detect
and
de=$mines(ordnance)
and
%imag
end

Searcher B

@swu@
beach reconnaissance
minefield detection
beaches
and
target detection
imaging detection
imaging devices
infrared images
image processing
target recognition
beach reconnaissance
end

Searcher D — 5 strategies combined in file

@swu@
beach
and
reconnaissance
end

@srtab@
beach reconnaissance
end

@swu@
%beach reconnaiss
end

@swu@
littoral
surf
and
mine
mines
minefield
minefields
end

@srtab@
detect
detection
detects
detect
detecting
end

@swu@
mine
mines
ordnance
and
shore
shores
beach
beaches
end

@srtab@
detect
detects
detecting
detection
detected
detector
detectors
end

@swu@
$shores
beach
beaches
%littoral
surf
surfing
and
$mines(ordnance)
and
$detection
$tracking
%automatic target track
%atr(automatic
end

Searcher E

@swu@
%minefield detection
%minefield breach
end

@qsrtab@
beach
reconnaissance
imaging
end

Searcher F

@swu@
?00beach
ti=beach
ti=beaches
ti=beachhead
ti=beachheads
ti=beachheads
%beach recon
surf
surf zone
%surf zones
%littoral
%littoral zon
ti=littoral
and
mines(ordnance)
mines
%mine detect
%minefield
%minefields
%mine imag
%navakxxx
%naval mine
%underwater mine
automatic target recognition
%atr(automatic target r
end

Searcher 3

@swuwps@
*mine detection
?60minefield
?00minefield
and
*mine detection
?60%detect
?00%detect
end

@srtab@
beach
beaches
shore
shores
shoreline
shorelines
end

Searcher 6

@swuups@
mine detection
mine detectors
minefields
and
$images
target recognition
target detection
and
shores
beaches
beach heads
end

Searcher C

@str@
$coatings
and
aluminum
%ivd(ion
?60ivd
%ion vapor deposi
?60vapor
and
fastenings
end

Searcher D — 3 strategies combined in user file

@str@
?60ivd
end

@srtab@
aluminum
al
and
coating
coatings
coated
coats
end

@str@
?60%fasten
and
?60%coat
end

@srtab@
al
aluminum
and
ivd
ion vapor
end

@str@
$fastenings
and
$coatings
end

@srtab@
al
aluminum
and
ivd
ion vapor
end

Searcher E

@str@
fasteners--c
fasteners--m
fasteners--nf
asteners--p
fasteners--t
fasteners/corrosion
?00fastenings
and
coating
coatings
end

Searcher 4

@str@
$fastening
%fastener
?60%fastener
and
$coatings
%ivd(ion
%ion vapor deposition
$corrosion inhibition
%corrosion protection
end

Searcher 5

@str@
$%metal coating
$%aluminum coating
$%fastener
$%coaating
$%coating
and
$%ivd(ion
$%ion vapor deposit
$%corrosion protection
end

Searcher C

```
@swuwps@
$coatings
and
aluminum
%ivd(ion
?60ivd
%ion vapor deposi
?60vapor
and
fastenings
end
```

Searcher D — 2 strategies combined in user file

```
@swu@
ivd
and
%coating
and
aluminum
end

@swu@
vapor
and
ion
and
%deposit
and
aluminum
al
and
%coat
end

@uftab@
fastener
fasteners
fastenings
end
```

Searcher E

```
@swu@
%fasteners
de=fasteners
sub=fasteners
fasteners, seals, clamps
nar=fasteners
and
coatings
end
```

Searcher 4

```
@swu@
$fastening
%fastener
ti=%fastener
and
$coatings
%ivd(ion
%ion vapor deposition
$corrosion inhibition
%corrosion protection
end
```

Searcher 5

```
@swu@
sub=$%metal coating
sub=$%aluminum coating
sub=$%fastener
sub=$%coating
and
sub=$%ivd(ion
sub=$%ion vapor deposit
sub=$%corrosion protection
end
```

Searcher A

@str@
$grenades
iacs=%grenade
and
assembly
%assembl
%fabricat
%manufactur
disassembly
end

Searcher C

@str@
grenades
and
assembly
end

Searcher D

@str@
?60%grenade
end

@srtab@
m-77
m77
m/77
end

Searcher E

@str@
m77 shaped charges
end

Searcher 1 — 2 strategies combined in user file

@str@
%grenade
and
%assemb
end

@str@
%grenade
and
%safety pin
%safety
and
%pin
%device
%fastening
%fastener
end

Searcher 5

@str@
$%grenade systems
$%grenades
$%grenades xm-77
and
$%assembly
$%safety pins
end

16 WU

Searcher A

@swu@
%grenade
de=$grenade
and
de=assembly
de=disassembly
%assembl
%fabricat
%manufact
de=$fabrication
end

Searcher C

@swu@
grenades
and
assembly
end

Searcher E

@swu@
m77
m-77
m/77
end

Searcher 1 — 2 strategies combined in user file

@swu@
%grenade
de=%grenade
kw=%grenade
and
%assembl
de=%assembl
kw=%assembl
end

@swu@
%grenade
kw=%grenade
de=%grenade
and
%safety pin
kw=%safety ping
kw=%safety pin
de=%safety pin
%safety
and
%pin
%device
%fastening
%fastener
end

Searcher 5 — No hits; searcher did not try again

@swu@
sub=$%grenade systems
sub=$%grenades
sub=$%grenades xm-77
and
sub=$%assembly
sub=$%safety pins
end

17 TR

Searcher A

@str@
%neural net
iacs=%neural net
and
manufacturing
iacs=%manufact
end

Searcher E

@str@
%neural network
end

Searcher F — 2 strategies combined in user file

@str@
%neural net
%expert suys
%expert sys
and
manufacturing
industrial plants
fabrication
$fabrication
$molding techniques
material forming
materials forming
composite fabrication
composites fabrication
iacs=manufacturing
iacs=fabrication
iacs=molding techniques
iacs=material forming
iacs=materials forming
iacs=composites
iacs=composite materials
$composite materials
end

@str@
?60neural
and
?60net
?60nets
?60network
?60networks
?60networking
?60networked
and
industrial plants
fabrication
?60fabricate
?60fabricated
?60fabricating
?60fabrication
manufacturing
?60manufacture
?60manufacturing
?60manufacturing
?60manufactured
$molding techniques
material forming
composite fabrication
composites fabricatop^H^Hion
composites fabrication
end

Searcher 2

This searcher found it inconvenient to download the search strategies. While the investigator attempted to record each strategy as it was executed, the strategy for this query appears to have been missed.

Searcher B

@str@
work measurement
job analysis
workload assessment
%time in motion stud
time studies
and
workload
work measurement
workload assessment
process improvement
end

Searcher E

@str@
?00%work measurement
%work measurement
end

Searcher 2 — The first search below had over 6100 hits and the second nearly 1200; no downloading was attempted.

@str@
$work measurement
%job analysis
%work load
%job shop sched
%systems engineering
end

@str@
$work measurement
%job analysis

Searcher 4

@str@
%process and product improvement
*work
*work measurement
%process improvement
process
processes
?60process
?60processes
and
*measurement
*work measurement
%process improvement
?60improvement
?60enhancement
end

@srtab@
work measurement
work enhancement
process improvement
process enhancement
process measurement
work measurements
work enhancements
process improvements
process enhancements
process measurements
end

18 WU

Searcher B

@swu@
work measurement
workload assessment
job analysis
time studies
process improvement
end

@srtab@
work measurement
process improvement
end

Searcher E

@swu@
de=work measurement
%work measurement
end

Searcher 2 — The first search below had over 4800 hits and the second over 1000; no downloading was attempted.

@str@
$work measurement
%job analysis
%work load
%job shop sched
%systems engineering
end

@str@
$work measurement
%job analysis

Searcher 4

@swu@
%process and product improvement
*work
*work measurement
%process improvement
process
processes
ti=process
ti=processes
and
*measurement
*work measurement
%process improvement
ti=improvement
ti=enhancement
end

@srtab@
work measurement
work enhancement
process improvement
process enhancement
process measurement
work measurements
work enhancements
process improvements
process enhancements
process meaurements
process measurements
end

19 TR

Searcher A

@str@
$radar receivers
and
$warning systems
and
$rotary wing aircraft
fixed wing aircraft
end

Searcher B

@str@
$radar
and
$receivers
and
$aircraft
end

Searcher C

%radar warning rec
radar receivers
and
$warning systems
%radar warning rec
and
$rotary wing aircraft
$helicopters
fixed wing aircraft
$jet aircraft
$military aircraft
tank aircraft
commercial aircraft
tanker aircraft
training aircraft
end

Searcher 1

@str@
%warning
and
%receiver
and
%radar
$radar
and
%helicopter
%fixed wing aircraft
end

Searcher 5

@str@
radar receiver
%radar warning receiver
%rwr%
%rwr(radar
and
%warning receiver
%radar warning receiver
%rwr%
%rwr(radar
and
$aircraft
end

Searcher 5

@swuups@
radar receiver
%radar warning receiver
%rwr%
%rwr(radar
and
%warning receiver
%radar warning receiver
%rwr%
%rwr(radar
and
$aircraft
end

Appendix 4

Sample Search

Query 1 executed by searcher C in TR Bibliographic Database was selected as an example.

Query: Flight control and instrumentation: including displays and related topics

Search strategy:

@str@
$*flight control systems
%flight display
flight instruments
and
instrumentation
$flight instruments
end

This strategy retrieved 319 hits, of which 70 were included in the sample which was judged for relevance. Relevance judgments for all of the 100 citations in the sample for this query were as follows:

Judged relevant or partially relevant by both judges	62
Judged relevant or partially relevant by judge 1, not relevant by judge 2	27
Judged relevant or partially relevant by judge 2, not relevant by judge 1	1
Judged not relevant by both judges	5
Relevance considered non-determinable by at least one judge	5

The 95 citations for which both judges provided judgments were used in the analysis. Of these 95 sample citations, 62 (70 percent) were judged relevant using the criterion of concurrence of judges, and 90 (95 percent) were judged relevant by either judge.

Of searcher C's 70 hits that were included in the sample, 42 were judged relevant by concurrence of judges, and 45 by either judge. This searcher's precision and recall ratios were calculated as follows:

Concurrence of judges:

Precision = 42 relevant retrieved ÷ 70 total retrieved = .60
Recall = 42 relevant retrieved ÷ 62 relevant in sample = .68

Either judge:

Precision = 45 relevant retrieved ÷ 70 total retrieved = .64
Recall = 45 relevant retrieved ÷ 90 relevant in sample = .50

Appendix 5

Concurrence between Relevance Judges

Search TR01

	1: Rel.	2: Not rel.	2: Total
2: Rel.	62	27	89
%	69.7	30.3	93.7
2: Not rel.	1	5	6
%	16.7	83.3	6.3
1: Total	63	32	95
%	66.3	33.7	100.0

Search TR02

	1: Rel.	2: Not rel.	2: Total
2: Rel.	11	5	16
%	68.8	31.3	88.9
2: Not rel.	0	2	2
%	.0	100.0	11.1
1: Total	11	7	18
%	61.1	38.9	100.0

Search TR03

	1: Rel.	2: Not rel.	2: Total
2: Rel.	28	34	62
%	45.2	54.8	66.7
2: Not rel.	11	20	31
%	35.5	64.5	33.3
1: Total	39	54	93
%	41.9	58.1	100.0

Search TR04

	1: Rel.	2: Not rel.	2: Total
2: Rel.	67	23	90
%	74.4	25.6	94.7
2: Not rel.	0	5	5
%	.0	100.0	5.3
1: Total	67	28	95
%	70.5	29.5	100.0

Search TR05

	1: Rel.	2: Not rel.	2: Total
2: Rel.	72	0	72
%	100.0	.0	74.2
2: Not rel.	0	25	25
%	.0	100.0	25.8
1: Total	72	25	97
%	74.2	25.8	100.0

Search TR06

	1: Rel.	2: Not rel.	2: Total
2: Rel.	91	0	91
%	100.0	.0	95.8
2: Not rel.	0	4	4
%	.0	100.0	4.2
1: Total	91	4	95
%	95.8	4.2	100.0

Search TR07

	1: Rel.	2: Not rel.	2: Total
2: Rel.	93	0	93
%	100.0	.0	98.9
2: Not rel.	0	1	1
%	.0	100.0	1.1
1: Total	93	1	94
%	98.9	1.1	100.0

Search TR08

	1: Rel.	2: Not rel.	2: Total
2: Rel.	87	0	87
%	100.0	.0	96.7
2: Not rel.	0	3	3
%	.0	100.0	3.3
1: Total	87	3	90
%	96.7	3.3	100.0

Search TR09

	1: Rel.	2: Not rel.	2: Total
2: Rel.	33	2	35
%	94.3	5.7	41.2
2: Not rel.	1	49	50
%	2.0	98.0	58.8
1: Total	34	51	85
%	40.0	60.0	100.0

Search TR10

	1: Rel.	2: Not rel.	2: Total
2: Rel.	32	11	43
%	74.4	25.6	55.8
2: Not rel.	7	27	34
%	20.6	79.4	44.2
1: Total	39	38	77
%	50.6	49.4	100.0

Search TR11

	1: Rel.	2: Not rel.	2: Total
2: Rel.	2	1	3
%	66.7	33.3	50.0
2: Not rel.	0	3	3
%	.0	100.0	50.0
1: Total	2	4	6
%	33.3	66.7	100.0

Search TR12

	1: Rel.	2: Not rel.	2: Total
2: Rel.	24	10	34
%	70.6	29.4	41.0
2: Not rel.	17	32	49
%	34.7	65.3	59.0
1: Total	41	42	83
%	49.4	50.6	100.0

Search TR13

	1: Rel.	2: Not rel.	2: Total
2: Rel.	35	6	41
%	85.4	14.6	52.6
2: Not rel.	22	15	37
%	59.5	40.5	47.4
1: Total	57	21	78
%	73.1	26.9	100.0

Search TR14

	1: Rel.	2: Not rel.	2: Total
2: Rel.	31	16	47
%	66.0	34.0	61.0
2: Not rel.	5	25	30
%	16.7	83.3	39.0
1: Total	36	41	77
%	46.8	53.2	100.0

Search TR15

	1: Rel.	2: Not rel.	2: Total
2: Rel.	7	9	16
%	43.8	56.3	64.0
2: Not rel.	0	9	9
%	.0	100.0	36.0
1: Total	7	18	25
%	28.0	72.0	100.0

Search TR16

	1: Rel.	2: Not rel.	2: Total
2: Rel.	12	9	21
%	57.1	42.9	87.5
2: Not rel.	1	2	3
%	33.3	66.7	12.5
1: Total	13	11	24
%	54.2	45.8	100.0

Search TR17

	1: Rel.	2: Not rel.	2: Total
2: Rel.	33	56	89
%	37.1	62.9	97.8
2: Not rel.	0	2	2
%	.0	100.0	2.2
1: Total	33	58	91
%	36.3	63.7	100.0

Search TR18

	1: Rel.	2: Not rel.	2: Total
2: Rel.	57	14	71
%	80.3	19.7	85.5
2: Not rel.	8	4	12
%	66.7	33.3	14.5
1: Total	65	18	83
%	78.3	21.7	100.0

Search TR19

	1: Rel.	2: Not rel.	2: Total
2: Rel.	31	28	59
%	52.5	47.5	78.7
2: Not rel.	1	15	16
%	6.3	93.8	21.3
1: Total	32	43	75
%	42.7	57.3	100.0

TR Pooled

	1: Rel.	2: Not rel.	2: Total
2: Rel.	808	74	882
%	91.6	8.4	63.9
2: Not rel.	251	248	499
%	50.3	49.7	36.1
1: Total	1059	322	1381
%	76.7	23.3	100.0

Search WU01

	1: Rel.	2: Not rel.	2: Total
2: Rel.	64	23	87
%	73.6	26.4	88.8
2: Not rel.	4	7	11
%	36.4	63.6	11.2
1: Total	68	30	98
%	69.4	30.6	100.0

Search WU02

	1: Rel.	2: Not rel.	2: Total
2: Rel.	57	26	83
%	68.7	31.3	83.0
2: Not rel.	2	15	17
%	11.8	88.2	17.0
1: Total	59	41	100
%	59.0	41.0	100.0

Search WU03

	1: Rel.	2: Not rel.	2: Total
2: Rel.	28	27	55
%	50.9	49.1	55.6
2: Not rel.	3	41	44
%	6.8	93.2	44.4
1: Total	31	68	99
%	31.3	68.7	100.0

Search WU04

	1: Rel.	2: Not rel.	2: Total
2: Rel.	80	18	98
%	81.6	18.4	99.0
2: Not rel.	0	1	1
%	.0	100.0	1.0
1: Total	80	19	99
%	80.8	19.2	100.0

Search WU05

	1: Rel.	2: Not rel.	2: Total
2: Rel.	78	1	79
%	98.7	1.3	82.3
2: Not rel.	3	14	17
%	17.6	82.4	17.7
1: Total	81	15	96
%	84.4	15.6	100.0

Search WU06

	1: Rel.	2: Not rel.	2: Total
2: Rel.	61	0	61
%	100.0	.0	64.2
2: Not rel.	0	34	34
%	.0	100.0	35.8
1: Total	61	34	95
%	64.2	35.8	100.0

Search WU07

	1: Rel.	2: Not rel.	2: Total
2: Rel.	82	1	83
%	98.8	1.2	83.8
2: Not rel.	1	15	16
%	6.3	93.8	16.2
1: Total	83	16	99
%	83.8	16.2	100.0

Search WU08

	1: Rel.	2: Not rel.	2: Total
2: Rel.	81	1	82
%	98.8	1.2	84.5
2: Not rel.	0	15	15
%	.0	100.0	15.5
1: Total	81	16	97
%	83.5	16.5	100.0

Search WU09

	1: Rel.	2: Not rel.	2: Total
2: Rel.	28	0	28
%	100.0	.0	29.2
2: Not rel.	2	66	68
%	2.9	97.1	70.8
1: Total	30	66	96
%	31.3	68.8	100.0

Search WU10

	1: Rel.	2: Not rel.	2: Total
2: Rel.	16	21	37
%	43.2	56.8	43.0
2: Not rel.	14	35	49
%	28.6	71.4	57.0
1: Total	30	56	86
%	34.9	65.1	100.0

Search WU11

	1: Rel.	2: Not rel.	2: Total
2: Rel.	2	1	3
%	66.7	33.3	50.0
2: Not rel.	1	2	3
%	33.3	66.7	50.0
1: Total	3	3	6
%	50.0	50.0	100.0

Search WU12

	1: Rel.	2: Not rel.	2: Total
2: Rel.	17	12	29
%	58.6	41.4	33.3
2: Not rel.	12	46	58
%	20.7	79.3	66.7
1: Total	29	58	87
%	33.3	66.7	100.0

Search WU13

	1: Rel.	2: Not rel.	2: Total
2: Rel.	50	7	57
%	87.7	12.3	67.9
2: Not rel.	8	19	27
%	29.6	70.4	32.1
1: Total	58	26	84
%	69.0	31.0	100.0

Search WU14

	1: Rel.	2: Not rel.	2: Total
2: Rel.	14	31	45
%	31.1	68.9	64.3
2: Not rel.	1	24	25
%	4.0	96.0	35.7
1: Total	15	55	70
%	21.4	78.6	100.0

Search WU15

	1: Rel.	2: Not rel.	2: Total
2: Rel.	9	34	43
%	20.9	79.1	81.1
2: Not rel.	0	10	10
%	.0	100.0	18.9
1: Total	9	44	53
%	17.0	83.0	100.0

Search WU16

	1: Rel.	2: Not rel.	2: Total
2: Rel.	14	59	73
%	19.2	80.8	84.9
2: Not rel.	1	12	13
%	7.7	92.3	15.1
1: Total	15	71	86
%	17.4	82.6	100.0

Search WU17

	1: Rel.	2: Not rel.	2: Total
2: Rel.	26	60	86
%	30.2	69.8	90.5
2: Not rel.	0	9	9
%	.0	100.0	9.5
1: Total	26	69	95
%	27.4	72.6	100.0

Search WU18

	1: Rel.	2: Not rel.	2: Total
2: Rel.	67	10	77
%	87.0	13.0	80.2
2: Not rel.	8	11	19
%	42.1	57.9	19.8
1: Total	75	21	96
%	78.1	21.9	100.0

Search WU19

	1: Rel.	2: Not rel.	2: Total
2: Rel.	21	24	45
%	46.7	53.3	90.0
2: Not rel.	0	5	5
%	.0	100.0	10.0
1: Total	21	29	50
%	42.0	58.0	100.0

WU Pooled

	1: Rel.	2: Not rel.	2: Total
2: Rel.	795	356	1151
%	69.1	30.9	72.3
2: Not rel.	60	381	441
%	13.6	86.4	27.7
1: Total	855	737	1592
%	53.7	46.3	100.0

Appendix 6

Precision and Recall

by Query and Searcher

TR Bibliographic Database
Precision: Relevance Based on Concurrence of Judges

Query	A	B	C	D	E	F	1	2	3	4	5	6	p	L-BOUND	U-BOUND
1	.82		.60		.68				1.00		.79		.69	.62	.77
2	.50		1.00				.60				.57		.63	.41	.85
3			.60			.05		.15		.18			.28	.20	.36
4					.74	.66			.73	.76			.71	.65	.77
5	.83	.68	1.00		.83				1.00		.80		.73	.65	.82
6	.25					.95	.93					.93	.93	.89	.97
7	.94				1.00						1.00	.90	.93	.89	.97
8		.88	1.00	1.00						1.00	.86		.88	.82	.94
9	.16				.22			.64		.33			.32	.24	.40
10				.00	.44	.29						.75	.33	.24	.42
11	.33	.20				.33	.40				.25		.30	.10	.50
12	.33	.30				.20				.50			.31	.23	.39
13	.47	.43	.32						.41	.33			.39	.32	.46
14	.54	.57		.21	1.00				.00			.00	.37	.28	.46
15			.40	.20	.33					.50	.67		.35	.20	.51
16	.48		.60		.00		.58				.60		.54	.41	.67
17	1.00				.19	.45		.47		1.00			.43	.35	.51
18		.65			.74					.33			.64	.57	.72
19	.63	.31	.52				.00			.50			.41	.34	.48
p	.63	.53	.52	.22	.55	.52	.75	.40	.61	.49	.78	.90			
L-BOUND	.57	.48	.46	.13	.49	.47	.63	.30	.50	.43	.71	.84			
U-BOUND	.68	.58	.59	.31	.61	.57	.87	.50	.72	.56	.85	.96			

NOTES:
p DENOTES THE OVERALL ROW (COLUMN) PERCENTAGE
L-BOUND, U-BOUND DENOTE THE 95% LOWER AND UPPER BOUND CONFIDENCE LIMITS
*) DENOTES THAT BOTH JUDGES FOUND THE DOCUMENT RELEVANT OR PARTIALLY RELEVANT

TR Bibliographic Database
Precision: Relevance Based on Either Judge

Query	A	B	C	D	E	F	1	2	3	4	5	6	p	L-BOUND	U-BOUND
1	.82		.64		.82				1.00		.87		.75	.68	.82
2	.75		1.00				.60				.71		.74	.54	.93
3			.69			.35		.24		.23			.40	.31	.49
4					.95	.75			.96	.96			.87	.83	.91
5	.83	.68	1.00		.83				1.00		.80		.73	.65	.82
6	.25					.95	.93					.93	.93	.89	.97
7	.94				1.00						1.00	.90	.93	.89	.97
8		.88	1.00	1.00						1.00	.86		.88	.82	.94
9	.18				.22			.72		.33			.34	.26	.42
10				.00	.50	.33						.75	.37	.28	.46
11	.33	.20				.33	.40				.25		.30	.10	.50
12	.33	.40				.33				.59			.41	.33	.50
13	.58	.52	.41						.52	.33			.48	.41	.55
14	.65	.57		.34	1.00				.33			1.00	.48	.39	.57
15			.80	.33	.50					.75	.67		.54	.38	.70
16	.65		.80		1.00		.75				.80		.73	.61	.85
17	1.00				.50	.65		.74		1.00			.66	.58	.74
18		.73			.80					.43			.72	.65	.79
19	.75	.42	.65				1.00			.73			.55	.48	.62
p	.68	.60	.61	.36	.69	.61	.81	.55	.78	.60	.83	.91			
L-BOUND	.63	.55	.55	.25	.63	.56	.70	.45	.69	.53	.76	.86			
U-BOUND	.73	.64	.67	.47	.75	.66	.92	.65	.87	.66	.89	.96			

NOTES:
p DENOTES THE OVERALL ROW (COLUMN) PERCENTAGE
L-BOUND, U-BOUND DENOTE THE 95% LOWER AND UPPER BOUND CONFIDENCE LIMITS
*) EITHER DENOTES THAT ONE JUDGE, THE OTHER JUDGE, OR BOTH JUDGES FOUND THE DOCUMENT RELEVANT OR PARTIALLY RELEVANT

WUIS

Precision: Relevance Based on Concurrence of Judges

Query	A	B	C	D	E	F	1	2	3	4	5	6	p	L-BOUND	U-BOUND
1	.92		.68		.46						.65		.66	.58	.73
2	.33		.80	.88			.59				.65		.62	.53	.71
3			.69			.67		.25		.31			.32	.24	.40
4					.78	.79			.74	.80			.78	.73	.83
5	.88	.60	.77		.75				1.00		1.00		.81	.74	.88
6	1.00	.87				.55	.91					.92	.80	.75	.85
7	1.00	.78		1.00	1.00						1.00	1.00	.91	.87	.95
8		.94	.67	1.00						.67	.79		.82	.75	.89
9	.50		.48		.18			.36		.27			.34	.26	.41
10					.19	.17						.16	.17	.10	.25
11	.00	.40				.33	.33				.33		.31	.08	.54
12	.22	.47				.13				.25	.50		.23	.16	.31
13	.48	.55	.50						.56	.43			.51	.44	.58
14	.19	.20		.13	1.00	.17			.19				.20	.12	.27
15			.42	1.00	.29					.21	.15		.25	.16	.33
16	.15		.00		1.00		.15						.15	.09	.21
17	.65				.44	.62		.27		.80			.42	.34	.50
18		.64			.69					.50			.67	.58	.75
19	.26	.45	.24				.25			.29			.29	.22	.37
p	.47	.72	.55	.41	.60	.45	.59	.28	.55	.41	.67	.82			
L-BOUND	.42	.67	.49	.25	.55	.39	.51	.22	.47	.35	.61	.75			
U-BOUND	.52	.77	.61	.56	.66	.50	.66	.35	.64	.46	.74	.89			

NOTES:

p DENOTES THE OVERALL ROW (COLUMN) PERCENTAGE

L-BOUND, U-BOUND DENOTE THE 95% LOWER AND UPPER BOUND CONFIDENCE LIMITS

*) DENOTES THAT BOTH JUDGES FOUND THE DOCUMENT RELEVANT OR PARTIALLY RELEVANT

WUIS
Precision: Relevance Based on Either Judge

Query	A	B	C	D	E	F	1	2	3	4	5	6	p	L-BOUND	U-BOUND
1	.92		.69		.46						.72		.68	.61	.76
2	.33		.80	1.00			.63				.65		.66	.57	.74
3			.69			1.00		.29		.36			.37	.29	.45
4					.97	.93			.97	.97			.95	.93	.98
5	.88	.60	.85		.77				1.00		1.00		.82	.76	.89
6	1.00	.87				.55	.91					.92	.80	.75	.85
7	1.00	.78		1.00	1.00						1.00	1.00	.91	.87	.95
8		.94	.67	1.00						.67	.81		.83	.76	.89
9	.50		.52		.18			.40		.27			.36	.28	.43
10					.19	.25						.16	.22	.14	.30
11	.50	.60				.33	.33				.33		.44	.20	.68
12	.22	.65				.23				.34	.50		.33	.25	.42
13	.48	.55	.50						.56	.43			.51	.44	.58
14	.44	.20		.17	1.00	.17			.23				.29	.21	.37
15			.58	1.00	.42					.28	.19		.33	.24	.42
16	.28		.50		1.00		.37						.32	.23	.40
17	.70				.63	.69		.40		.80			.54	.46	.62
18		.73			.74					.60			.73	.65	.80
19	.42	.60	.53				.38			.39			.45	.36	.53
p	.56	.75	.61	.46	.68	.52	.65	.36	.63	.48	.70	.82			
L-BOUND	.51	.70	.55	.30	.63	.47	.57	.29	.54	.43	.63	.75			
U-BOUND	.61	.80	.67	.62	.73	.58	.72	.43	.71	.53	.77	.89			

NOTES:
p DENOTES THE OVERALL ROW (COLUMN) PERCENTAGE
L-BOUND, U-BOUND DENOTE THE 95% LOWER AND UPPER BOUND CONFIDENCE LIMITS
*) EITHER DENOTES THAT ONE JUDGE, THE OTHER JUDGE, OR BOTH JUDGES FOUND THE DOCUMENT RELEVANT OR PARTIALLY RELEVANT

TR Bibliographic Database
Recall: Relevance Based on Concurrence of Judges

Query	A	B	C	D	E	F	1	2	3	4	5	6	p	L-BOUND	U-BOUND
1	.29		.68		.24				.02		.48		.34	.00	.76
2	.18		.27				.27				.36		.27	.00	.71
3			.75			.04		.18		.14			.28	.00	.72
4					.48	.96			.49	.55			.62	.14	1.00
5	.07	.74	.08		.07				.06		.06		.18	.00	.48
6	.01					.90	.30					.46	.42	.00	.90
7	.86				.08						.01	.59	.38	.00	.86
8		.32	.05	.02						.05	.66		.22	.00	.58
9	.18				.18			.48		.39			.31	.00	.76
10				.00	.25	.72						.09	.27	.00	.71
11	.50	.50				.50	1.00				.50		.60	.17	1.00
12	.13	.25				.50				.71			.40	.00	.88
13	.57	.54	.63						.34	.17			.45	.02	.89
14	.45	.52		.39	.03				.00			.00	.23	.00	.67
15			.29	.43	.29					.57	.29		.37	.00	.79
16	.92		.50		.00		.58				.50		.50	.06	.94
17	.30				.27	.55		.48		.21			.36	.00	.79
18		.93			.86					.12			.64	.09	1.00
19	.16	1.00	.87				.00			.42			.49	.05	.93
p	.22	.26	.16	.02	.17	.25	.05	.05	.06	.14	.13	.12			
L-BOUND	.19	.23	.14	.01	.14	.22	.03	.03	.05	.11	.11	.10			
U-BOUND	.25	.29	.19	.03	.19	.28	.06	.06	.08	.16	.15	.15			

NOTES:
p DENOTES THE OVERALL ROW (COLUMN) PERCENTAGE
L-BOUND, U-BOUND DENOTE THE 95% LOWER AND UPPER BOUND CONFIDENCE LIMITS
*) DENOTES THAT BOTH JUDGES FOUND THE DOCUMENT RELEVANT OR PARTIALLY RELEVANT

TR Bibliographic Database
Recall: Relevance Based on Either Judge

Query	A	B	C	D	E	F	1	2	3	4	5	6	p	L-BOUND	U-BOUND
1	.20		.50		.20				.01		.37		.26	.00	.64
2	.19		.19				.19				.31		.22	.00	.62
3			.33			.10		.11		.07			.15	.00	.50
4					.46	.81			.48	.52			.57	.08	1.00
5	.07	.74	.08		.07				.06		.06		.18	.00	.48
6	.01					.90	.30					.46	.42	.00	.90
7	.86				.08						.01	.59	.38	.00	.86
8		.32	.05	.02						.05	.66		.22	.00	.58
9	.19				.17			.50		.36			.31	.00	.76
10				.00	.18	.52						.06	.19	.00	.57
11	.33	.33				.33	.67				.33		.40	.00	.83
12	.06	.16				.39				.39			.25	.00	.67
13	.40	.37	.44						.24	.10			.31	.00	.71
14	.33	.31		.38	.02				.02			.02	.18	.00	.49
15			.25	.31	.19					.38	.13		.25	.00	.63
16	.68		.36		.05		.41				.36		.37	.00	.79
17	.11				.27	.29		.28		.08			.21	.00	.56
18		.76			.67					.11			.51	.00	1.00
19	.10	.70	.57				.02			.32			.34	.00	.75
p	.17	.20	.14	.02	.15	.21	.04	.05	.06	.12	.10	.09			
L-BOUND	.15	.18	.12	.01	.13	.18	.03	.03	.04	.10	.08	.07			
U-BOUND	.19	.23	.16	.03	.17	.23	.05	.06	.07	.14	.12	.10			

NOTES:
p DENOTES THE OVERALL ROW (COLUMN) PERCENTAGE
L-BOUND, U-BOUND DENOTE THE 95% LOWER AND UPPER BOUND CONFIDENCE LIMITS
*) EITHER DENOTES THAT ONE JUDGE, THE OTHER JUDGE, OR BOTH JUDGES FOUND THE DOCUMENT RELEVANT OR PARTIALLY RELEVANT

WUIS

Recall: Relevance Based on Concurrence of Judges

Query	A	B	C	D	E	F	1	2	3	4	5	6	p	L-BOUND	U-BOUND
1	.19		.72		.17						.44		.38	.00	.85
2	.02		.07	.12			.86				.19		.25	.00	.63
3			.32			.07		.68		.50			.39	.00	.87
4					.56	.81			.36	.64			.59	.11	1.00
5	.27	.08	.13		.55				.09		.18		.22	.00	.54
6	.15	.97				.67	.69					.79	.65	.24	1.00
7	.60	.77		.04	.11						.11	.51	.36	.00	.74
8		.36	.10	.01						.02	.72		.24	.00	.62
9	.11		.54		.11			.57		.50			.36	.00	.79
10					.25	.69						.19	.38	.00	.92
11	.00	1.00				.50	.50				.50		.50	.06	.94
12	.12	.47				.35				.65	.06		.33	.00	.74
13	.30	.24	.60						.64	.32			.42	.00	.85
14	.50	.07		.21	.21	.21			.43				.27	.00	.63
15			.56	.11	.78					.89	.44		.56	.12	.99
16	.93		.00		.07		.29						.32	.00	.77
17	.50				.46	.31		.77		.31			.47	.03	.91
18		.21			.91					.07			.40	.00	.95
19	.52	.43	.19				.10			.67			.38	.00	.81
p	.20	.26	.16	.02	.25	.17	.12	.07	.09	.18	.16	.12			
L-BOUND	.17	.23	.14	.01	.22	.15	.10	.05	.07	.15	.13	.09			
U-BOUND	.22	.29	.19	.03	.28	.20	.15	.09	.11	.21	.18	.14			

NOTES:

p DENOTES THE OVERALL ROW (COLUMN) PERCENTAGE

L-BOUND, U-BOUND DENOTE THE 95% LOWER AND UPPER BOUND CONFIDENCE LIMITS

*) DENOTES THAT BOTH JUDGES FOUND THE DOCUMENT RELEVANT OR PARTIALLY RELEVANT

WUIS
Recall: Relevance Based on Either Judge

Query	A	B	C	D	E	F	1	2	3	4	5	6	p	L-BOUND	U-BOUND
1	.13		.52		.12						.34		.28	.00	.72
2	.01		.05	.09			.61				.13		.18	.00	.51
3			.16			.05		.38		.28			.22	.00	.62
4					.57	.78			.39	.63			.59	.11	1.00
5	.26	.07	.13		.54				.09		.17		.21	.00	.53
6	.15	.97				.67	.69					.79	.65	.24	1.00
7	.58	.75		.04	.11						.11	.50	.35	.00	.73
8		.35	.10	.01						.02	.72		.24	.00	.62
9	.10		.53		.10			.60		.47			.36	.00	.78
10					.08	.31						.06	.15	.00	.55
11	.25	.75				.25	.25				.25		.35	.00	.77
12	.05	.27				.27				.37	.02		.20	.00	.54
13	.23	.18	.46						.49	.25			.32	.00	.73
14	.35	.02		.09	.07	.07			.15				.12	.00	.39
15			.16	.02	.23					.26	.12		.16	.00	.48
16	.34		.04		.01		.14						.13	.00	.46
17	.16				.20	.10		.35		.09			.18	.00	.52
18		.19			.76					.07			.34	.00	.88
19	.40	.27	.20				.07			.42			.27	.00	.66
p	.15	.18	.12	.01	.18	.13	.09	.06	.07	.14	.11	.08			
L-BOUND	.13	.15	.10	.01	.16	.11	.07	.04	.06	.12	.09	.06			
U-BOUND	.17	.20	.14	.02	.21	.15	.11	.07	.08	.16	.13	.09			

NOTES:
p DENOTES THE OVERALL ROW (COLUMN) PERCENTAGE
L-BOUND, U-BOUND DENOTE THE 95% LOWER AND UPPER BOUND CONFIDENCE LIMITS
*) EITHER DENOTES THAT ONE JUDGE, THE OTHER JUDGE, OR BOTH JUDGES FOUND THE DOCUMENT RELEVANT OR PARTIALLY RELEVANT

Appendix 7

Hits in Sample

TR Bibliographic Database
Hits in Sample
Total Hits / Relevant Hits: Concurrence of Judges / Relevant Hits: Either Judge

Query	Searcher											
	A	B	C	D	E	F	1	2	3	4	5	6
1	22/18/18		70/42/45		22/15/18				1/1/1		38/30/33	
2	4/2/3		3/3/3				5/3/3				7/4/5	
3			35/21/24			20/1/7		34/5/8		22/4/5		
4					43/32/41	97/64/73			45/33/43	49/37/47		
5	6/5/5	78/53/53	6/6/6		6/5/5				4/4/4		5/4/4	
6	4/1/1					86/82/82	29/27/27					45/42/42
7	85/80/80				7/7/7						1/1/1	61/55/55
8		32/28/28	4/4/4	2/2/2						4/4/4	66/57/57	
9	38/6/7				27/6/6			25/16/18		39/13/13		
10				1/0/0	18/8/9	79/23/26						4/3/3
11	3/1/1	5/1/1				3/1/1	5/2/2				4/1/1	
12	9/3/3	20/6/8				60/12/20				34/17/20		
13	43/20/25	44/19/23	69/22/28						29/12/15	18/6/6		
14	26/14/17	28/16/16		58/12/20	1/1/1				3/0/1			1/0/1
15			5/2/4	15/3/5	6/2/3					8/4/6	3/2/2	
16	23/11/15		10/6/8		1/0/1		12/7/9				10/6/8	
17	10/10/10				48/9/24	40/18/26		34/16/25		7/7/7		
18		82/53/60			66/49/53					21/7/9		
19	8/5/6	99/31/42	52/27/34				1/0/1			26/13/19		

WUIS
Hits in Sample
Total Hits / Relevant Hits: Concurrence of Judges / Relevant Hits: Either Judge

Query	Searcher											
	A	B	C	D	E	F	1	2	3	4	5	6
1	13/12/12		68/46/47		24/11/11						43/28/31	
2	3/1/1		5/4/4	8/7/8			83/49/52				17/11/11	
3			13/9/9			3/2/3		75/19/22		45/14/16		
4					58/45/56	82/65/76			39/29/38	64/51/62		
5	24/21/21	10/6/6	13/10/11		57/43/44				7/7/7		14/14/14	
6	9/9/9	68/59/59				75/41/41	46/42/42					52/48/48
7	49/49/49	81/63/63		3/3/3	9/9/9						9/9/9	42/42/42
8		31/29/29	12/8/8	1/1/1						3/2/2	73/58/59	
9	6/3/3		31/15/16		17/3/3			45/16/18		52/14/14		
10					21/4/4	64/11/16						19/3/3
11	2/0/1	5/2/3				3/1/1	3/1/1				3/1/1	
12	9/2/2	17/8/11				48/6/11				44/11/15	2/1/1	
13	31/15/15	22/12/12	60/30/30						57/32/32	37/16/16		
14	36/7/16	5/1/1		24/3/4	3/3/3	18/3/3			31/6/7			
15			12/5/7	1/1/1	24/7/10					39/8/11	26/4/5	
16	89/13/25		6/0/3		1/1/1		27/4/10					
17	20/13/14				27/12/17	13/8/9		75/20/30		10/8/8		
18		22/14/16			88/61/65					10/5/6		
19	43/11/18	20/9/12	17/4/9				8/2/3			49/14/19		