DOCUMENT RESUME

ED 376 852 IR 055 328

AUTHOR McJunkin, Monica Cahill
TITLE Precision and Recall in Title Keyword Searches.
PUB DATE Apr 94
NOTE 78p.; Master's Research Paper, Kent State University.
PUB TYPE Dissertations/Theses - Undetermined (040)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Bibliographic Databases; Bibliographic Records; Cataloging; Comparative Analysis; Library Catalogs; *Online Searching; *Performance; *Relevance (Information Retrieval); *Search Strategies
IDENTIFIERS *Library of Congress Subject Headings; *OCLC

ABSTRACT

This study examines precision and recall for title and keyword searches performed in the "FirstSearch" WorldCat database when keywords are used with and without adjacency of terms specified. A random sample of 68 titles in economics were searched in the OCLC (Online Computer Library Center) Online Union Catalog in order to obtain their Library of Congress subject headings. After limiting by year and language, keywords were searched in "FirstSearch" with and without adjacency of the keywords specified. Subject headings of titles retrieved in keyword searches were compared with sample title subject headings to determine the degree of match, or relevancy. Figures for precision (the percentage of retrieved elements which are relevant) and recall (the percentage of relevant items in the database that were retrieved) were compared to determine whether the use of adjacency operators significantly alters the effectiveness of title keyword searches. Precision was improved with little degradation in recall when the keywords were discipline-specific. Other factors affecting overall levels of precision and recall include the number of terms and number of subject headings assigned to the sample titles. It is hoped that the results of this study will help build a framework in which to view keyword search strategies. (Contains 22 references.) (Author)

IR

PRECISION AND RECALL IN TITLE KEYWORD SEARCHES

A Master's Research Paper submitted to the
Kent State School of Library and Information Science
in partial fulfillment of the requirements
for the degree Master of Library Science

by

Monica Cahill McJunkin

April, 1994

PRECISION AND RECALL IN TITLE KEYWORD SEARCHES

ABSTRACT

This study examines precision and recall for title keyword searches performed in the *FirstSearch* WorldCat database when keywords are used with and without the adjacency of terms specified. A random sample of 68 titles in economics were searched in the OCLC Online Union Catalog in order to obtain their Library of Congress subject headings. After limiting by year and language, keywords were searched in *FirstSearch* with and without adjacency of the keywords specified. Subject headings of titles retrieved in keyword searches were compared with the sample title subject headings to determine the degree of match, or relevancy. Figures for precision (the percentage of retrieved elements which are relevant) and recall (the percentage of relevant items in the database that were retrieved) were compared to determine whether the use of adjacency operators significantly alters the effectiveness of title keyword searches. Precision was improved with little degradation in recall when the keywords were discipline-specific. Other factors affecting overall levels of precision and recall include the number of terms and number of subject headings assigned to the sample titles. It is hoped that the results of this study will help build a framework in which to view keyword search strategies.

Master's Research Paper by

Monica Cahill McJunkin

B.S., The Ohio State University, 1987

M.L.S., Kent State University, 1994

Approved by

Adviser Sachera N. Connell **Date** April 27, 1994

TABLE OF CONTENTS

TABLE OF CONTENTS

APPENDIX

LIST OF TABLES

ACKNOWLEDGMENTS

I am grateful to OCLC, Inc. for supporting this study by generously providing access to *FirstSearch*. I would also like to thank Tschera Connell for her input and encouragement, Morgan Jackson for programming advice, and Fred Connell for his guidance on statistics. A special thanks is due to my husband, Kevin, for his constant support throughout the project.

INTRODUCTION

Online catalogs provide many opportunities for creative subject access, including keyword searches. While keyword searches in controlled vocabulary fields allow access to subject headings when entry terms or word order are not known, titles also contain subject-rich terms. These keywords use the authors' own terminology, which is often more current than the Library of Congress Subject Headings (LCSH) (Chan & Hodges 1990), and can be combined or related to each other in order to vary the search. This study investigates the extent to which title keywords convey subject content and compare the relative effectiveness of searching title keywords via two different strategies.

Unlike searches in non-keyword based systems, which must match the beginning of the field, keyword searches involve identifying the requested terms at any position in the field being searched. Multiple terms can be combined in a search using the Boolean operators AND, OR, and NOT. Word stems or truncated terms can be specified, as well as positional operators. These operators can specify the order in which the terms appear, their proximity to each other, or that the terms be adjacent to one another. The options in keyword searching allow the user to broaden or narrow a search as needed.

Peters and Kurth (1991) determined from a study of dial-access transaction logs at the University of Missouri - Kansas City that library patrons were using title keyword searches as a form of uncontrolled vocabulary search. In other studies, users were observed using title terms for subject access both in the catalog and while browsing the shelf (Hancock 1987, Hancock-Beaulieu 1990). These studies make a case for the existing use of subject access

through title keywords, but show no evidence of the success of these searches, or the relative success of different types of keyword searches.

Other studies have found title terms used for subject searching: Larson (1991) has described the decline of subject searching and the concomitant rise in title keyword searching over a six year period, and Ensor (1992) describes several studies which show a rise in keyword searching of all types. Both authors note that keyword use rises with catalog experience. Connell (1991a) observed that experienced users perform title keyword searches as a lead-in to the controlled vocabulary, and Peters and Kurth (1991) recommend this method in addition to using title keyword searches alone.

When users perform title keyword searches as a subject approach to the catalog, how good are the results? More specifically, do items which contain the same terms in their titles cover the same topic, and are certain title keyword search strategies more effective than others for subject searching?

LITERATURE REVIEW

Characteristics of Title Keyword Searches

Title keyword searching has some advantages over controlled vocabulary subject access. Titles terms are more likely to agree with the user's terminology and serve as a complement to the assigned subject headings (Aluri, Kemp & Boll 1991), and have been found by Jamieson to overlap very little with subject cross-references (Yee 1991). Bates (1977) found that subject experts in economics consistently preferred headings that were more precise than the subject headings assigned to works, and that they particularly disliked the subheading "economic conditions" because of the variety of meanings covered by it. In her dissertation, (described by Connell 1991a), Bates also found that users had particular difficulty with subject heading

matching for economics items; economics headings tend to be complex, often including subheadings for time periods and geographic regions.

However, title keywords are only as good as the author makes them. Even after articles, prepositions, and conjunctions are removed from consideration, generic terms like "report" remain, as well as metaphors and cute, catchy phrases; synonyms and spelling variations compound the problem. In general, keyword searching exhibits a lack of tolerance for misspellings and variations of any kind (Akeroyd 1990). Lastly, because the terms are taken out of context, keyword searching can result in what is called a false drop, which occurs when the search terms are used in a different manner in a retrieved record than was intended by the user (Olsgaard & Evans 1981).

Evaluation Methodologies

Many studies have attempted to evaluate the usefulness of title keyword searching. Connell (1991b) used keywords from abstracts in *Book Review Digest* to determine to what extent book descriptions match terms in subject headings or titles. She also looked at fields that are not commonly used, such as the subtitle or other title information, to determine their potential in retrieving items. All words in the descriptions were considered keywords except the following: a, an, and, at, by, for, from, how, in, of, on, the, to, with. Connell compared all keywords from the abstracts with fields in the bibliographic record, and found that for books for which no match was found between the description and the subject headings or LCSH cross-references, 27.8% matched title keywords. Of the books remaining, over a third produced matches in the subtitle field. While some of these last matches were with terms that indicate form of the item, subtitles often provide meaningful keywords when the title proper contains a catchy phrase. This study indicates that titles and subtitles may be useful for subject access;

however, the percentage of matches reflects only those titles for which subject headings and subject cross-references failed to produce a match.

In a study which took the opposite approach, Gerhan (1989) compared the usefulness of terms in titles and subject headings by determining whether they were likely to be used by patrons desiring items on that topic. Catalog cards were examined for terms which had a reasonable probability of being search terms; these terms were subjectively rated according to whether he thought that patrons would use them for subject access. He found that title keywords are effective retrieval terms about 55% of the time, including 10% in which subject headings are absent or extremely lacking, but subject headings were effective about 85% of the time, and so made a better first choice for searching. Gerhan concluded that terms from subject headings and titles are often complementary, and use of both methods may be the most productive.

Cherry (1992) took yet another approach. While Connell started with book descriptions, and Gerhan started with catalog entries in order to determine the likelihood that books would be found based on keywords in the bibliographic record, Cherry examined unsuccessful subject searches (defined as those with zero hits). Actual user subject searches were converted to subject keyword searches, title searches, title keyword searches, and subject cross-reference searches. Title keywords were the most useful, retrieving records in 62% of the cases, as opposed to subject keywords and subject cross-references, which were each successful 33% of the time. Title searches, which must match the beginning of the title, were successful at retrieving records 43% of the time. Although these searches were only performed with requests that had already failed with traditional subject access, this study does indicate that title keyword searching is a useful addition to subject searching, especially since the search terms employed were actual patron search requests.

Aanonson (1987) compared some retrieval sets from subject searches he performed while evaluating keyword searching on six university catalogs. He determined that title keyword searches not only retrieved useful items not found with subject keyword searches, but that they provide useful starting points for getting into the controlled vocabulary. He also found that additional useful records were retrieved when the series title was included in the title keyword search as well.

Evaluation Measures

The previous studies did not evaluate the relevance of retrieved items; books were not examined to determine content, and search terms were accepted as accurate portrayals of desired subjects. Number of records retrieved was the main consideration. However, large retrieval sets can be a disadvantage while searching if the user must browse through many records looking for useful items. Larson (1991) attributes a decline in subject searching over a six year period to increasing database size and the resulting user frustration with large retrieval sets. He notes that keyword-based systems are more likely to cause information overload for the user, and favors ranking of output records according to the number of search terms contained in each record. Yee (1991), on the other hand, suggests that keyword indexing may be improved by locational data to allow searching of keywords combined into "phrases", and Lancaster et al. (1991) include the limiting of keyword searches by date, language, or other factors as a way of improving subject access.

Evaluating retrieved records according to their relevance can be a complex issue. First, one must distinguish between pertinence and relevance. Relevance has been defined as a "relationship between a document and a request", and pertinence as the "relationship between a document and an information need" (Lancaster 1979, 263). In other words, a relevant item is

one that matches the search request, while a pertinent item is one that is judged useful by the user. In the absence of real users with actual information needs, the relevance of an item can be agreed upon by a group of subject experts (Lancaster 1979). Kemp (1974) views relevance as objective and pertinence as subjective, drawing a parallel in psychology with denotation of words (objective) and connotation (subjective). Others disagree, claiming that whenever relevance decisions are made by individuals or groups of individuals, they must be subjective and dependent upon a variety of external factors. In either case, making relevance decisions based on a subjective measure of topicality can be appropriate for initial evaluations of a system's retrieval capabilities (Hersh 1994).

When no users are involved and items are not available for evaluation, other methods must be used to determine relevance. Although finding a matching LC subject heading does not guarantee search success, Bates (1977) claims that a matching score between search terms and subject headings are a good measure of success. The LC subject heading should provide one best heading, controlling for synonyms and related terms, and should match the scope of the item.

Once a method for determining relevance has been determined, records can be weighted according to their usefulness. Unlike known item searches, subject searches need a "measure of degree of success" (Lancaster et al. 1991, 378); some items are more relevant than others. In a study of database coverage for periodical indexes, Sharma weighted items according to the following scale (1982, 36):

fully relevant	1.0
half or moderately relevant	0.5
marginally relevant	0.25
irrelevant	0.0

This type of weighting procedure could apply to any method of deciding relevance.

Once relevance is determined, recall and precision figures put the relevance figures into perspective. Recall is defined as the percentage of relevant documents retrieved, and precision as the percentage of retrieved documents that are relevant (Aluri, Kemp & Boll 1991). Generally, recall and precision are inversely related; improvements in one come at the expense of the other. While precision is easy to calculate because it is based on the ratio of relevant items retrieved to total items retrieved, recall is harder to estimate because it involves the ratio of relevant items retrieved to total relevant items in the database, which is impossible to know. Lancaster (1979) has suggested estimating the total number of relevant items by having several users perform parallel searches (i.e., use different search strategies), then combining the total number of relevant items retrieved to represent the number of relevant items in the database. While this will not disclose indexing failures in the database, it can highlight the usefulness of different search strategies.

Summary

To date, research on title keyword searches has typically focused on comparisons of title keyword searches with subject or subject keyword searches; book descriptions, user searches, or "made up" terms served as the source of keywords. In general, title keyword searching is often characterized by poor precision owing to false drops, and may not improve recall substantially over subject heading searches, especially when time and system costs are taken into account (Hildreth 1983). Truncation and word stemming can increase recall, but searches in large databases often suffer more from a lack of precision. Better precision can be obtained by the use of word proximity or adjacency operators, which combine keywords into meaningful phrases Chan and Hodges 1990). However, it is not known to what extent this degrades recall.

OBJECTIVES AND DEFINITIONS

The objectives of this study are:

To determine the levels of precision and recall obtained with title keyword searching for titles in economics,

To determine the levels of precision and recall obtained with title keyword searching for titles in economics modified by adjacency operators to create keyword "phrases",

To compare the levels of precision and recall obtained via the two methods in order to determine which is the more effective means of subject access.

Unlike previous studies, titles are the source of keywords and provide the searched fields. In effect, works on the same topic are assumed to use the same title terms if they are to be used for subject access. Because of the difficulties with subject access which have been described above, economics was chosen as the subject field for this study. Keywords include all terms except the stop words used by Connell (1991b): a, an, and, at, by, for, from, how, in, of, on, the, to, with. The number of keywords vary from search to search: Keywords were searched singly or in combination with stopwords in the title delineating the search groups. For example, the two keyword groups for the title *Low-income housing in the developing world* are "low-income housing" and "developing world". When more than one term is included, a Boolean AND is implicit in the search.

Searches were performed on *FirstSearch*, using the WorldCat database, which is equivalent in coverage to the OCLC Online Union Catalog. It was chosen to provide the largest possible coverage with the least bias introduced by individual institutional holdings. Title keyword searches cover the title proper field, as well as other title information, uniform titles, added titles, and series titles.

Two different search strategies were used: In the first, keyword(s) were entered, and

searches were performed without regard to word order or proximity. In the second, keyword(s) were entered with adjacency operators which specify the exact phrases to be matched.

Relevance was determined by the degree of LC subject heading match between the source title and the retrieved title. While not ideal, it provides an objective measure which can be used for other studies, and separates the issue of the adequacy of the indexing language from the comparison of keyword search strategies. Sharma's weighting scale has been adapted for this study:

Exact subject match	1.0
Broader or narrower	0.5
Related	0.25
No match	0.00

Broader and narrower matches include headings that omit or include, respectively, subdivisions, in addition to those defined by the LCSH hierarchy. Similarly, related matches include headings with the same main heading but different subdivisions. Because all subject headings from source and retrieved titles were considered, it was possible for an item to receive a relevance score greater than 1.0.

The denominator for calculating recall, the total number of relevant documents in the database, was estimated using the union of the unique relevant records (weighted score) retrieved via the two methods with the number of unique records obtained via an exact phrase subject heading search, using headings from the source titles. Recall, then, is the number of unique relevant records (weighted score) retrieved divided by this denominator. Precision is simpler, and is defined as the number of unique relevant records (weighted score) retrieved divided by the total unique records retrieved.

It is important to note that the scope of this study does not involve comparing title

keyword searching with subject searching. While it is not possible to know how many relevant items may be missed by title keyword searches, the extent to which titles containing the same terminology are on the same subject is an important consideration in title keyword searching. In this respect, it is the relative effectiveness of two title keyword search strategies that is being examined.

METHODOLOGY

Precision and recall of title keyword searches in economics were obtained by analyzing search results from the *FirstSearch* WorldCat database.

Sample

The members of the target population were monograph titles in economics, and the accessible population sampled were the titles in *Economics and Business*, an annotated bibliography that was published from 1984 through 1986. The entries are numbered, which facilitates sampling, and they cover a range of subtopics on economics, such as monetary theory, international economics, and industrial organization, so the vocabulary is varied. Also, because the titles are from a limited number of years, the searches could be limited to these years. The vast majority of titles fall between 1983 to 1985, so only titles in that range are included in the sample.

A random sample of titles was drawn from the bibliography using a table of random numbers. The sample size, n, was chosen to obtain 90% confidence with a margin of error of ten percentage points, using the formula:

$$e = 1.645\frac{s}{\sqrt{n}}$$

where e is the margin of error, and s is the standard deviation. From the results of a pilot study, the standard deviation for precision and recall was estimated to be 0.5. Substituting for e and s, the sample size, n, is 68:

$$0.10 = 1.645\frac{.5}{\sqrt{n}}$$

$$n = 67.65 \equiv 68$$

Procedures

Before title keyword searching began, data for the sample titles were collected. First, each title was searched in the OCLC Online Union Catalog in order to make a list of the LC subject headings assigned to each title. It is important to stress that no subject headings from other authority lists were considered in this study; for this reason, this search could not take place on *FirstSearch*, because the source of subject headings is not displayed with the records. Second, the keyword combinations to be searched for each title were recorded and numbered. An example data form for the sample titles is shown in Appendix A.

Title keyword searches in the *FirstSearch* WorldCat database began after first limiting the searches by language (English) and year of publication (1983-1985). In order to simplify and standardize the searches, keywords were not searched in various forms, such as truncation, word stemming, elimination of plurals, or various spellings. In order to keep the retrieval sets manageable, further limits were imposed: If any subject heading search for a title yielded more than one thousand records, all searches for that title were limited to one year. If any title keyword search yielded more than five hundred records, one hundred records were systematically sampled from the retrieved set. Large retrieved sets which fell into one of the

following categories were not sampled; they were omitted from the study:

Keyword contains the bibliographic format of the item (guide, directory)

Keyword contains the presentation or treatment of data (analysis, survey)

Keyword contains a generic geographic or chronological term (area, nation, era)

Keyword contains a broad geographic or chronological term (United States, 20th century)

When in doubt, the retrieved records were sampled. These limits were necessary because of the significant proportion of overly large retrieval sets: Out of 360 possible searches, 83, or 23% of the retrieval sets contained five hundred or more records. Of these, 65 (18%) were sampled, and 18 (5%) were included in the categories described above and were omitted.

Each title was searched using both strategies. When a keyword stands alone in a title, both strategies were completed in one search. For example, in the title *Agricultural Development in Bangladesh*, "Bangladesh" stands apart from the other keywords, so adjacency operators can not be used. The syntax of the search statements for this title were:

s ti:agricultural development

s ti:bangladesh

s ti:agricultural w development

In the first search, the terms "agricultural" and "development" could appear in the retrieved records in any combination of searched fields, in either order. In the third search, the terms must appear in the same field together as a phrase, in the order specified. The second search retrieves the keyword "bangladesh" for both methods.

Transaction logs from the search sessions were downloaded. The source title, which should appear in the retrieval sets, was removed from consideration. Then data for each search were recorded: the search number and lists of retrieved records (by OCLC record number), in

columns for exact match, broader, narrower, and related. The relevance scores for each retrieved record were determined by comparing subject headings of the sample titles with those of the retrieved titles using the tenth edition of the *Library of Congress Subject Headings*, which most closely corresponds with the time period covered by the study. A data collection form for the retrieved records is shown in Appendix B.

The relevance of each retrieved record was determined by rating each of the retrieved item's subject headings as an exact match (1.0), broader heading (0.5), narrower heading (0.5), related heading (0.25), or no match (0.0).

In order to estimate a denominator for recall, subject headings from the sample titles were searched. Only exact subject heading matches were to be included; however, the WorldCat subject headings can not be searched exactly. Exact phrase searching is available on subject heading fields, but the various segments of the subject headings are indexed separately. Thus, a search for "Government lending -- United States", which is stated as "sh=(government lending and united states)", will also retrieve "Government lending -- Law and legislation -- United States", "Government lending -- United States -- Handbooks, manuals, etc.", as well as a record with the pair of headings "United States -- Small Business Administration" and "Government lending -- Arkansas". Each retrieved set of records was edited to remove the extraneous headings.

Data Analysis

Figures for recall and precision were estimated for each search method by the following method: For each search, the total number of relevant records were calculated. Then recall and precision for each search were estimated for each of the two keyword search strategies using the following formulae:

$$R = \frac{r}{k+s-l_{ks}}$$

$$P = \frac{r}{t}$$

where

R = recall
r = # of relevant records retrieved in this search
k = # of relevant records from title keyword searches for this keyword grouping (w/o adjacency)
s = # of records from exact subject heading searches for this title
l_{ks} = # of records contained in both k and s (overlap)
P = precision
t = total records retrieved in this search

The set of relevant records retrieved when adjacency is specified is always a subset of the set of relevant records retrieved when adjacency is not specified, therefore, only the larger set is necessary for calculating the denominator for recall. Precision and recall were then averaged for each title.

Because every subject heading in a retrieved record is evaluated for relevancy, an individual record may have a relevancy score greater than 1.0; thus precision for a search (and average precision for a title) may also be greater than 1.0. Also, since the denominator for recall includes each relevant record only once, but a retrieved record may have a relevancy score greater than 1.0, it is possible for recall for a search to be greater than 1.0.

DISCUSSION

The titles included in the sample and their LC subject headings are listed in Appendix C, and the number of records retrieved for each keyword search is shown in Appendix D.

Out of 68 titles, 29 required no sampling of retrieved records, and 39 contained retrieval sets which were sampled due to their size; these are referred to as "non-sampled titles" and "sampled

titles", respectively. Actual retrieval set sizes are shown for those which were subsequently sampled.

Subject Heading Searches

The inability to search for exact subject heading matches was unexpected. The searches for three of the sample titles were limited to one year because the retrieval sets for individual subject heading searches were greater than one thousand. For title 47, searching "sh=population", limited to 1984, retrieved 1291 records, only forty of which were found to contain the exact subject heading "Population." Subject headings containing subdivisions pose an additional problem: For title 54, searching "sh=(small business and united states)" retrieved 836 records, and only 235 contained the exact heading "Small business -- United States". Not only were other subdivisions also included in the retrieved records, but "Small business" and "United States" did not have to appear in the same heading in order for a record to be retrieved. While the flexibility allowed by this system has some advantages, ranking of output according to the degree of match to the search statement should be incorporated. If all records containing "Population" were listed before variations including subdivisions, evaluation of records would have been easier, and the search would not have had to be limited to one year.

Unusual Relevancy Scores

Some titles have no precision or recall scores, and others have scores exceeding 1.0. Undefined scores occur when searches retrieve no records (other than the sample title). When there is no set of retrieved records, calculating precision is impossible and calculating recall, although theoretically possible if exact subject heading searches retrieved a nonzero set, is meaningless. There are no undefined precision and recall scores in the sampled titles (because there was always at least a sample of one hundred records retrieved) ; undefined scores occur in

four of the non-sampled titles. For example, title 13 is "Socio-economic accounting". Searching for these keywords either with or without adjacency operators retrieves no records other than the sample title.

High precision and recall scores occurred for both sampled and non-sampled titles. As described in the data analysis section, retrieved records may receive a relevancy score greater than 1.0. This usually occurred when there were few exact subject heading matches, and title keyword searches retrieved small sets of records consisting mostly of other editions of the same work. These cases have been included in the overall data calculations, even though they are artificially large. The alternative would be to evaluate all retrieved records in order to eliminate those which are considered duplicates of the sample, or of each other, to determine that the retrieval sets contain only unique records. With so many records, however, and none of the items in hand, this alternative is not feasible. It was assumed that these duplicates are evenly distributed throughout the retrieval sets, and would not affect the comparison between the two search strategies.

Non-sampled Titles

Precision and recall for the non-sampled titles via both strategies are shown in Appendix E. For convenience, title keyword searches performed without adjacency specified are referred to as keyword searches, and title keyword searches performed with adjacency specified are referred to as phrase searches. Table 1 contains summary data for the non-sampled titles. Confidence intervals were generated using the z-statistic at a level of significance of 0.10. The mean precision scores for keyword and phrase searching are 44% and 53%, respectively. The confidence intervals overlap quite a bit, yet it is clear that higher precision was obtained from phrase searching. The mean difference between the scores is 7.8%.

The mean recall scores for keyword and phrase searching are 17.5% and 15%, respectively. The loss in recall obtained with phrase searching is much less than the gain in precision, and the confidence intervals almost totally overlap. Recall scores for keyword searches were, on average, 3% higher than recall scores for phrase searches.

The values of the keyword and phrase scores relative to each other are what would be expected; phrase searching results in higher precision with only a slight loss in recall. In other words, the number of false drops eliminated exceeded the relevant records which were missed. It is significant to note that because keywords occurring singly were searched singly for both strategies, the difference in precision is not as large as it might be if only multiple-word keyword phrases were included in the study. They were included to obtain a more realistic sense of how the strategies would perform against each other in natural settings, in which keywords would often be searched singly despite user strategy preferences or system defaults. The precision and recall obtained when single keywords are excluded is explored later.

Table 1.-- Non-Sampled Titles

Type of Score	Confidence Interval
Keyword Precision	$.3016 \leq .4402 \leq .5787$
Phrase Precision	$.3580 \leq .5335 \leq .7090$
Difference Between Strategies (Keyword Precision - Phrase Precision)	$-.1488 \leq -.0783 \leq -.0078$
Keyword Recall	$.1134 \leq .1745 \leq .2357$
Phrase Recall	$.0951 \leq 15023 \leq .2054$
Difference Between Strategies (Keyword Recall - Phrase Recall)	$.0028 \leq .0309 \leq .0590$

Sampled Titles

Table 2 contains summary data for precision and recall for the sampled titles. Confidence intervals were again generated using the z-statistic. The mean precision scores for keyword and phrase searching are almost identical, 23.5% and 25.4%, respectively, with a mean difference of less than 2%. The confidence intervals overlap almost completely. Mean recall figures are also similar to each other, 21.5% and 20% for keyword and phrase searching, respectively. The lack of difference between strategies may be due to the preponderance of single keyword searches in the sampled titles, for which keyword and phrase searching are identical.

Comparing the data in Table 2 with the data for non-sampled titles in Table 1, it is apparent that some factor is causing a significant difference in the relevance scores. Precision for sampled titles is much lower than the precision for non-sampled titles. The keywords in the sampled titles are more likely to be general, non-discipline specific terms (hence the need to sample from large retrieval sets). These terms are used in a variety of ways, resulting in a lot of false drops. Also, the sampled titles tend to contain more keywords, and more keyword

Table 2.-- Sampled Titles

Type of Score	Confidence Interval
Keyword Precision	$.1376 \leq .2352 \leq .3328$
Phrase Precision	$.1520 \leq .2541 \leq .3562$
Difference Between Strategies (Keyword Precision - Phrase Precision)	$-.0334 \leq -.0006 \leq -.0045$
Keyword Recall	$.1353 \leq .2153 \leq .2954$
Phrase Recall	$.1214 \leq .2020 \leq .2827$
Difference Between Strategies (Keyword Recall - Phrase Recall)	$-.0014 \leq .0133 \leq .2796$

groupings, than the non-sampled titles. This reduces the probability that any one keyword (or keyword grouping) adequately describes the content of the item, and lessens the probability that retrieved records will have matching subject headings. The precision and recall scores for the sampled titles are in Appendix F.

There is no significant difference in precision and recall between keyword and phrase search strategies. Again, this may be because these titles contain more keywords which were searched singly; this will be examined later in the paper. An analysis of the titles that contained three or more single keywords to be searched shows that 13 out of a total of 15 are sampled titles. Precision and recall data for these titles are shown in Appendix G. Confidence intervals for all following tables were generated using the t-statistic for a two-tailed test at the 0.10 level of significance. As summary data for the 13 sampled titles in Table 3 is shown, precision for both keyword and phrase searching is low, only 18% and 20%, respectively, which indicates that searching single keywords, which are less specific in meaning than multiple word phrases, lowers precision.

Recall is very similar to the recall obtained for all of the sampled titles: 21% and 21.5%

Table 3.-- Sampled Titles with Three or More Single Keywords

Type of Score	Confidence Interval
Keyword Precision	$.0438 \leq .1829 \leq .3220$
Phrase Precision	$.0401 \leq .2017 \leq .3633$
Difference Between Strategies (Keyword Precision - Phrase Precision)	$-.0447 \leq -.0188 \leq .0071$
Keyword Recall	$.1220 \leq .2114 \leq .3009$
Phrase Recall	$.1235 \leq .2154 \leq .3072$
Difference Between Strategies (Keyword Recall - Phrase Recall)	$-.0165 \leq -.0039 \leq .0086$

(Table 2) versus 21.5% and 20% (Table 3). This is slightly higher than the recall obtained for the non-sampled titles (See Table 1). Although the general terms found in the sampled titles result in larger retrieval sets containing many false drops, they pick up more of the relevant records.

In a similar analysis, titles which contained four or more keyword groups were combined. This set does not quite overlap completely with the titles containing three or more single keywords, but it is also composed almost entirely of sampled titles (15 out of 17). (See Appendix H for precision and recall for individual titles.) Table 4 contains summary data for the 15 sampled titles in this category. Precision, which is 17% and 18.6% for keyword and phrase searching, respectively, is slightly, but not significantly lower than the precision found with all sampled titles in Table 2 or that found for the sampled titles with three or more single keywords shown in Table 3. Recall is significantly lower, at 14.5% and 14.6%. Scores for both precision and recall may be lower than for the entire group because as the number of keyword groups increases, it is less likely that any one group approximates the content adequately. Fewer relevant records are retrieved, and thus recall suffers as well as precision.

Table 4. -- Sampled Titles with Four or More Keyword Groups

Type of Score	Confidence Interval
Keyword Precision	$.0518 \leq .1692 \leq .2866$
Phrase Precision	$.0489 \leq .1863 \leq .3237$
Difference Between Strategies (Keyword Precision - Phrase Precision)	$-.0400 \leq -.0171 \leq .0058$
Keyword Recall	$.0634 \leq .1454 \leq .2274$
Phrase Recall	$.0612 \leq .1462 \leq .2312$
Difference Between Strategies (Keyword Recall - Phrase Recall)	$-.0123 \leq -.0008 \leq .0108$

Relevance Scores in Relation to Number of Subject Headings

Since relevance is evaluated based on subject heading matches, the number of subject headings assigned to the sample titles was analyzed to see if this affected precision and recall. Appendix I shows the precision and recall data for the titles which have three or more subject headings assigned to them; summary data is shown in Table 5.

For both non-sampled and sampled titles, there is little difference in precision and recall due to strategy. For non-sampled titles, precision and recall both dropped significantly from the scores for all the non-sampled titles. (See Table 1.) This may indicate that titles with three or

Table 5. -- Titles with Three or More Subject Headings

Type of Score -- Non-sampled Titles	Confidence Interval
Keyword Precision	$.1285 \leq .2425 \leq .3565$
Phrase Precision	$.1207 \leq .2567 \leq .3927$
Difference Between Strategies (Keyword Precision - Phrase Precision)	$-.0674 \leq -.0142 \leq .0390$
Keyword Recall	$.0343 \leq .0920 \leq .1496$
Phrase Recall	$.0240 \leq .0769 \leq .1297$
Difference Between Strategies (Keyword Recall - Phrase Recall)	$.0039 \leq .0151 \leq .0262$
Type of Score -- Sampled Titles	
Keyword Precision	$.1259 \leq .3306 \leq .5353$
Phrase Precision	$.1398 \leq .3482 \leq .5566$
Difference Between Strategies (Keyword Precision - Phrase Precision)	$-0.0441 \leq -.0176 \leq .0089$
Keyword Recall	$.0629 \leq .2343 \leq .4058$
Phrase Recall	$.0321 \leq .2042 \leq .3762$
Difference Between Strategies (Keyword Recall - Phrase Recall)	$-0.0015 \leq .0302 \leq .0619$

more subject headings have complex or varied topics that can not be described with only one or two subject headings. Non-sampled titles tend to have fewer, more specific keyword groupings (2.25 per title versus 4 for sampled titles), and may have specific terms which match none of the subject headings. For sampled titles, recall is not significantly different from the recall obtain for all the sampled titles. Precision, however, is improved (33% and 35% versus 23.5% and 25.4% - See Table 2). This is likely because sampled titles tend to contain more keyword groupings, which have a greater chance of match against several subject headings.

Appendix J shows the precision and recall scores for titles which have only one subject heading. Confidence intervals are displayed in Table 6. For the non-sampled titles, precision for both strategies is similar to precision for all non-sampled titles (46% and 50% versus 44% and 53% in Table 1). Recall, however, is greatly improved. Since non-sampled titles, on average, have fewer keyword groupings than the sampled titles, when only one subject heading is assigned, the topic of the work is covered by one phrase. So, recall may be improved because the few keyword groupings are more likely to match the single subject heading.

For the sampled titles, precision and recall are both lower than for sampled titles as a whole. This is probably because the large number of terms or phrases do not match well individually to a single subject heading.

Single Keyword or Keyword Group

It was found that eleven of the non-sampled titles contain only a single keyword or keyword grouping. These are shown in Appendix K. Searches for three of them retrieved zero records, so these have undefined precision and recall scores. None of the sampled titles fall into this category. Confidence intervals are shown in Table 7.

Table 6. -- Titles with One Subject Heading

Type of Score -- Non-sampled Titles	Confidence Interval
Keyword Precision	$.2804 \leq .4560 \leq .6396$
Phrase Precision	$.2921 \leq .5023 \leq .7124$
Difference Between Strategies (Keyword Precision - Phrase Precision)	$-.0951 \leq -.0423 \leq .0106$
Keyword Recall	$.1377 \leq .2764 \leq .4152$
Phrase Recall	$.1226 \leq .2521 \leq .3816$
Difference Between Strategies (Keyword Recall - Phrase Recall)	$-.0190 \leq .02433 \leq .0677$
Type of Score -- Sampled Titles	
Keyword Precision	$.0505 \leq .0918 \leq .1331$
Phrase Precision	$.0460 \leq .1144 \leq .1827$
Difference Between Strategies (Keyword Precision - Phrase Precision)	$-0.0540 \leq -.0226 \leq .0088$
Keyword Recall	$.0695 \leq .1504 \leq .2314$
Phrase Recall	$.0589 \leq .1454 \leq .2319$
Difference Between Strategies (Keyword Recall - Phrase Recall)	$-0.0143 \leq .0005 \leq .0244$

Table 7. -- Titles with a Single Keyword or Keyword Group

Type of Score Non-Sampled Titles	Confidence Interval
Keyword Precision	$.3561 \leq .7500 \leq 1.1438$
Phrase Precision	$.4657 \leq .9402 \leq 1.415$
Difference Between Strategies (Keyword Precision - Phrase Precision)	$-.4189 \leq -.1903 \leq .0383$
Keyword Recall	$.0372 \leq .1842 \leq .3311$
Phrase Recall	$.0120 \leq .1530 \leq .2861$
Difference Between Strategies (Keyword Recall - Phrase Recall)	$-.0112 \leq .0311 \leq .0735$

The figures for precision are much higher than those of the entire non-sampled group, and there is a significant difference between precision for keyword and phrase searching (75% and 94%). However, these figures are artificially inflated by the occurrence of small retrieval sets that include records which are duplicates or close matches to the sample title and thus have unusually large relevance scores.

Keywords that Match Subject Headings

Titles containing keywords which matched topical or geographic terms in the assigned subject headings were also analyzed. Precision and recall scores for these titles are shown in Appendix L; confidence intervals are displayed in Table 8. There is no significance difference between the two strategies for precision or recall. The matches were thought to indicate standardized terminology. However, the matching terms also appear to be general, non-discipline specific, which caused precision to decrease for both non-sampled and sampled titles. Recall for non-sampled titles, which is only 17.4% and 14.8% for keyword and phrase searching respectively, is slightly higher at 19% and 18.7%, showing the improvement obtained by the by the standardization of terms. Recall drops slightly for sampled titles from 21.5% and

Table 8. -- Titles with Keywords which Match Subject Headings

Type of Score -- Non-sampled Titles	Confidence Interval
Keyword Precision	$.2781 \leq .4228 \leq .5675$
Phrase Precision	$.2638 \leq .4207 \leq .5776$
Difference Between Strategies (Keyword Precision - Phrase Precision)	$-.0179 \leq -.0021 \leq .0221$
Keyword Recall	$.0262 \leq .1919 \leq .3576$
Phrase Recall	$.0196 \leq .1875 \leq .3554$
Difference Between Strategies (Keyword Recall - Phrase Recall)	$-.00036 \leq .0043 \leq .0122$

Table 8.-- continued	Confidence Interval
Type of Score -- Sampled Titles	
Keyword Precision	$.0662 \leq .1653 \leq .2645$
Phrase Precision	$.0755 \leq .1790 \leq .2825$
Difference Between Strategies (Keyword Precision - Phrase Precision)	$-0.0348 \leq -.0137 \leq .0073$
Keyword Recall	$.0937 \leq .1832 \leq .2728$
Phrase Recall	$.0762 \leq .1626 \leq .2490$
Difference Between Strategies (Keyword Recall - Phrase Recall)	$-0.0346 \leq .0207 \leq .0759$

20%, to 18% and 16%. More general terms tend to increase recall; in this case the decrease is not significant and is most likely due to the small sample size.

Single Keywords Excluded

Lastly, keywords which were searched singly were removed from consideration in order to determine their effect on precision and recall. Only seven titles remained in the "sampled" titles category. This demonstrates that single keywords tended to be general terms which resulted in large retrieval sets, and thus required sampling. Precision and recall for each title are shown in Appendix M (non-sampled titles) and Appendix N (Sampled titles).

Table 9 shows the data for non-sampled titles. Confidence intervals were generated using the z-statistic. Precision levels for keyword and phrase searching are 49.5% and 59%, respectively, with a mean difference of 7%. These are similar to, but slightly higher than the levels for non-sampled titles including single keywords, shown in Table 1. and much higher than the levels obtained for sampled titles including single keywords, shown in Table 2. As expected, the removal of the single keywords, which are more general in meaning, results in higher precision with a significant difference between keyword and phrase searching strategies. Recall levels are 22.5% and 19%, with a mean difference of 5%.

Table 9.-- Non-sampled Titles Excluding Single Keywords

Type of Score	Confidence Interval
Keyword Precision	$.3513 \leq .4956 \leq .6399$
Phrase Precision	$.4277 \leq .5928 \leq .7578$
Difference Between Strategies (Keyword Precision - Phrase Precision)	$-.1137 \leq -.0692 \leq -.0247$
Keyword Recall	$.1211 \leq .2255 \leq .3299$
Phrase Recall	$.0837 \leq .1912 \leq .2987$
Difference Between Strategies (Keyword Recall - Phrase Recall)	$.0209 \leq .0497 \leq .0786$

Table 10 contains the summary data for sampled titles; confidence intervals were generated using the t-statistic with 6 degrees of freedom. Precision is low, only 20.3% and 21.6% for keyword and phrase searching, respectively. Even though the single keywords have been excluded, these titles still contain general terms which required sampling of retrieval sets; the use of non-specific terms results in lower precision with little difference between search strategies. Recall, at 21% and 22% is not significantly different from the recall obtained with non-sampled titles (See Table 9).

Table 10.-- Sampled Titles Excluding Single Keywords

Type of Score	Confidence Interval
Keyword Precision	$.0079 \leq .2039 \leq .3998$
Phrase Precision	$.0098 \leq .2167 \leq .4236$
Difference Between Strategies (Keyword Precision - Phrase Precision)	$-.0409 \leq -.0128 \leq .0152$
Keyword Recall	$.1010 \leq .2083 \leq .3156$
Phrase Recall	$.1051 \leq .2172 \leq .3293$
Difference Between Strategies (Keyword Recall - Phrase Recall)	$-.0266 \leq -.0089 \leq .0088$

SUMMARY AND CONCLUSIONS

This study has examined precision and recall obtained from title keyword searches performed with and without adjacency operators. When keywords are limited in meaning, precision is significantly improved by the use of adjacency operators and recall declines to a lesser extent. Because of the design of this study, other factors were larger influences, such as the level of specificity of the terms, the length of the sample title, the number of subject headings assigned to the sample title, and the extent to which titles contained standardized terminology.

Overall, precision and recall were quite low; many exact subject heading matches were missed by title keyword searches. Precision can be improved by choosing search terms carefully; discipline-specific, subject-rich terms are best. Care should also be taken when using title keyword searches as a lead-in to the controlled vocabulary: The user should be aware of the standard terminology in the field and the level of specificity needed. As with any search, one who is not familiar with a subject's terminology may not end up with the one best heading. For example, a keyword search for "macroeconomics" would pull up records with the subject heading "Macroeconomics". However, the user may really have something like "Supply-side economics" in mind, but does not know how to phrase it for a search. One who is not a subject expert should consult the LCSH or online cross-references in order to find the correct terminology. On the other hand, one should also be knowledgeable about the online system in order to use it effectively. A user who does not know that *FirstSearch* may retrieve terms from several fields in the same record may be confused by the results: A search for "industrial structure" may retrieve a record with "pricing structure" in the title, and "Industrial commission" in the series title. Here, certainly, knowledge about the system's search logic and the availability of adjacency operators is helpful. Although the results of this study seem to support the use of

adjacency operators to improve searching effectivness, a user for whom absolute recall is more important may wish to use a broader search strategy.

Title keyword searching, with or without adjacency operators, is available in many online catalogs, and is sure to be added to more in the future. Evidence suggests that library patrons are using title keyword searching as a means of subject access, but we have few measures of its effectiveness, and as database sizes increase, precision will be an ever-growing problem. Whenever title fields are searched, alone or in combination with other content-bearing fields such as subject headings or notes, precision requires that title terms be indicative of the content of the item. More studies are needed to clarify the extent to which adjacency operators affect precision and recall. Future research could repeat this study with a larger sample size, using only discipline-specific and/or multiple-word keyword phrases in order to magnify the relationship between adjacency operators and precision and recall. Other disciplines could be examined, or the focus could be on journal article titles. Future research could take another direction and repeat this study using truncation of terms, or proximity operators in place of adjacency. If studies support certain strategies as being more helpful than others, this could have implications in several areas. First, more systems can be designed to support these strategies. Second, users can be instructed on the relative merits of different strategies, either formally or through help screens. Lastly, retrieval systems could be designed to default to certain strategies under some conditions, or to rank the output based on adjacency or proximity, in order to increase search success without increasing user effort. It is hoped that the results of this study will help build a framework in which to view keyword search strategies.

REFERENCES

Aanonson, John. 1987. A comparison of keyword subject searching on six British university OPACs. Online Review 11(5): 303-313.

Akeroyd, John. 1990. Information seeking in online catalogs. Journal of Documentation 46(1): 33-52.

Aluri, Rao, D. Alasdair Kemp, and John T. Boll. 1991. Subject analysis in online catalogs. Englewood, Colo.: Libraries Unlimited.

Bates, Marcia J. 1977. Factors affecting subject catalog search success. Journal of the American Society for Information Science (May): 161-169.

Chan, Lois Mai, and Theodora Hodges. 1990. Subject cataloguing and classification: the late 1980s and beyond. In Technical Services Today and Tomorrow, comp. Michael Gorman, 74-85. Englewood, Colo.: Libraries Unlimited.

Cherry, Joan M. 1992. Improving subject access in OPACs: an exploratory study of conversion of users' queries. Journal of Academic Librarianship 18(2): 95-99.

Connell, Tschera Harkness. 1991a. Librarian subject searching in online catalogs: an exploratory study of knowledge used. Ph.D. diss., University of Illinois at Urbana-Champaign.

Connell, Tschera Harkness. 1991b. Techniques to improve subject retrieval in online catalogs: flexible access to elements in the bibliographic record. Information Technology and Libraries (June): 87-98.

Ensor, Pat. 1992. User characteristics of keyword searching in an OPAC. College & Research Libraries (January): 72-80.

Gerhan, David R. 1989. LCSH *in vivo*: subject searching performance and strategy in the OPAC era. Journal of Academic Librarianship 15(2): 83-89.

Hancock, Michelene. 1987. Subject searching behaviour at the library catalogue and at the shelves: implications for online interactive catalogues. Journal of Documentation 43(4): 303-321.

Hancock-Beaulieu, Michelene. 1990. Evaluating the impact of an online library catalogue on subject searching behaviour at the catalogue and at the shelves. Journal of Documentation 46(4): 318-338.

Hersh, William. 1994. Relevance and retrieval evaluation: perspectives from medicine. Journal of the American Society for Information Science 45(3): 201-206.

Hildreth, Charles R. 1983. To Boolean or not to Boolean? Information Technology and Libraries (September): 235-237.

Kemp, D.A. 1974. Relevance, pertinence and information system development. Information Storage and Retrieval 10: 37-47.

Lancaster, F.Wilfred. 1979. Information retrieval systems: characteristics, testing and evaluation. 2d ed. New York: Wiley & Sons.

Lancaster, F.Wilfred, et al. 1991. Identifying barriers to effective subject access in library catalogs. Library Resources and Technical Services 35(4): 377-391.

Larson, Ray R. 1991. The decline of subject searching: long-term trends and patterns of index use in an online catalog. Journal of the American Society for Information Science 42(3): 197-215.

Olsgaard, John N., and John Edward Evans. 1981. Improving keyword indexing. Journal of the American Society for Information Science (January): 71-72.

Peters, Thomas A., and Martin Kurth. 1991. Controlled and uncontrolled vocabulary subject searching in an academic library online catalog. Information Technology and Libraries (September): 201-211.

Sharma, V.S. 1982. Information resources for recreation, leisure and sport. Database (August): 32-50.

Yee, Martha M. 1991. System design and cataloging meet the user: user interfaces to online public access catalogs. Journal for the American Society for Information Science 42(2): 78-98.

APPENDIX A -- SOURCE TITLE DATA FORM

Title #___

Title:_____

LCSH:
# of records
retrieved subject heading

Search statements:
Type Statement

Search statement type codes:
k -- keyword search, no adjacency specified
p -- keyword "phrase" search with adjacency specified

APPENDIX B -- RETRIEVED RECORD DATA FORM

Title # ____ Search #_____ Search Type_____

Exact Match_ Broader Narrower Related

APPENDIX C - SAMPLE TITLES AND SUBJECT HEADINGS

Title #	Title	Subject Headings
1	Macroeconomics: Keynesian, monetarist, and Marxist views	Macroeconomics.
2	The Norwegian economy, 1920-1980	Norway -- Economic conditions -- 1918-
3	Silicon Valley fever : growth of high-technology culture	Microelectronics industry -- California -- Santa Clara County -- History. Semiconductor industry -- California -- Santa Clara County -- History. High technology industries -- California -- Santa Clara County -- History.
4	The mechanics of Baltimore : workers and politics in the age of revolution, 1763-1812	Skilled labor -- Maryland -- Baltimore -- Political activity. Working class -- Maryland -- Baltimore -- Political activity. Baltimore (Md.) -- Politics and government.
5	Crisis in Soviet agriculture	Agriculture and state -- Soviet Union. Agriculture -- Economic aspects -- Soviet Union. Agriculture -- Soviet Union.
6	The West German economy	Germany (West) -- Economic conditions.
7	Agricultural computer guide & directory : here's how to decide if a computer is in your future	Farm management -- Data processing -- Handbooks, manuals, etc. Agriculture -- Data processing -- Handbooks, manuals, etc. Microcomputers -- Handbooks, manuals, etc. Computer industry -- United States -- Directories.
8	Beyond monetarism : finding the road to stable money	Monetary policy -- United States. Federal Reserve banks. Chicago school of economics.

APPENDIX C - SAMPLE TITLES AND SUBJECT HEADINGS

Title #	Title	Subject Headings
9	The theory of international trade	International trade -- Mathematical models. International economic relations -- Mathematical models.
10	World economy : changes and challenges	Economic history -- 1971-1990.
11	The political economy of China's changing relations with Southeast Asia	China -- Foreign economic relations -- Asia, Southeastern. Asia, Southeastern -- Foreign economic relations -- China.
12	Business and its public	Public relations. Industry -- Social aspects. Corporate image.
13	Socio-economic accounting	Social accounting.
14	Accounting for pensions : results of applying the FASB's preliminary views	Pension trusts -- United States -- Accounting.
15	Tax aspects of real estate transactions	Real property and taxation -- United States.
16	Industrial structure, pricing, and inflation	Inflation (Finance) -- Great Britain. Industrial concentration -- Great Britain. Great Britain -- Economic policy -- 1945- Inflation (Finance) -- United States. Industrial concentration -- United States. United States -- Economic policy -- 1981-
17	The economic analysis of technological change	Technological innovations -- Economic aspects. Diffusion of innovations.

APPENDIX C - SAMPLE TITLES AND SUBJECT HEADINGS

Title #	Title	Subject Headings
18	The energy crisis ten years later	Energy policy -- Congresses. Energy industries -- Congresses. Petroleum industry and trade -- Congresses.
19	Mass unemployment : plant closings and community mental health	Unemployment -- United States -- Psychological aspects. Plant shutdowns -- United States -- Psychological aspects. Community mental health services -- United States.
20	Japan's reshaping of American labor law	Labor laws and legislation -- United States. Labor laws and legislation -- Japan.
21	The birth of solidarity : the Gdansk negotiations, 1980	Labor policy -- Poland -- History -- Sources. Collective bargaining -- Poland -- History -- Sources.
22	The 100 best companies to work for in America	Personnel management -- United States -- Handbooks, manuals, etc. Job satisfaction -- United States -- Handbooks, manuals, etc. Corporations -- United States -- Handbooks, manuals, etc.
23	Comparative international budgeting and finance	Budget -- Case studies. Finance, Public -- Case studies.
24	Macroeconomic theory, a survey	Macroeconomics.
25	Public enterprise economics	Government business enterprises. Microeconomics.

APPENDIX C - SAMPLE TITLES AND SUBJECT HEADINGS

Title #	Title	Subject Headings
26	Keynes, the instability of capitalism	Keynesian economics. Capitalism. Keynes, John Maynard, -- 1883-1946.
27	American enterprise in foreign markets : studies of Singer and International Harvester in imperial Russia	Corporations, American -- Soviet Union -- History -- Case studies. International business enterprises -- Soviet Union -- History -- Case studies. International business enterprises -- United States -- History -- Case studies.
28	The rise of the corporate economy	Corporations -- Great Britain -- History. Big business -- Great Britain -- History. Industrial concentration -- Great Britain -- History.
29	Inequality and poverty in Malaysia : measurement and decomposition	Income distribution -- Malaysia. Poor -- Malaysia.
30	Cuba, dilemmas of a revolution	Cuba -- Politics and government -- 1959-
31*	Industrial structure and policy in less developed countries	Industry and state -- Developing countries. Developing countries -- Industries. Economic development. International business enterprises.
32	To promote prosperity : U.S. domestic policy in the mid-1980s	United States -- Economic policy -- 1981- United States -- Social policy -- 1980- United States -- Politics and government -- 1981-
33	Hidden spending : the politics of federal credit programs	Government lending -- United States. Loans -- United States -- Government guaranty.

* - Searches for this title were limited to 1985.

APPENDIX C - SAMPLE TITLES AND SUBJECT HEADINGS

Title #	Title	Subject Headings
34	Multinational excursions.	International business enterprises.
35	The international economy since 1945	International economic relations. Economic history -- 1945-
36	Banking deregulation and the new competition in financial services	Banks and banking -- United States -- Deregulation. Financial services industry -- United States -- Deregulation. Banking law -- United States.
37	The exclusive economic zone : a Latin American perspective	Economic zones (Maritime law) -- Latin America. Territorial waters -- Latin America.
38	Shopping centre development	Shopping centers.
39	Youth, expectations and transitions	Youth -- Austrlia -- Victoria. Young adults -- Austrlia -- Victoria. Education, secondary -- Australia -- Victoria. Youth -- Employment -- Australia -- Victoria.
40	American jobs and the changing industrial base	Labor supply -- United States -- Effect of technological innovations on. Occupations -- United States -- Forecasting.
41	The plant closure policy dilemma : labor, law and bargaining	Plant shutdowns -- Government policy -- United States. Plant shutdowns -- Law and legislation -- United States. Collective bargaining -- United States. Labor laws and legislation -- United States. Labor policy -- United States.

APPENDIX C - SAMPLE TITLES AND SUBJECT HEADINGS

Title #	Title	Subject Headings
42	The elements of industrial relations	Industrial relations.
43	Multinational enterprises and the OECD industrial relations guidelines	Collective bargaining -- International business enterprises. Organisation for Economic Co-operation and Development. Industrial relations.
44	The mediators.	Mediation and conciliation, Industrial -- United States.
45	Negotiating a labor contract : a management handbook	Collective bargaining -- Handbooks, manuals, etc. Collective labor agreements -- Handbooks, manuals, etc.
46	Comparative industrial relations : a trans-Atlantic dialogue	Industrial relations -- Great Britain -- Congresses. Industrial relations -- United States -- Congresses.
47*	Multidisciplinary perspectives on population and conflict	Population. Interpersonal conflict. International relations.
48	Forecasting use of health services : a provider's guide	Medical care -- Utilization -- Forecasting. Health facilities -- Utilization -- Forecasting. Medical care -- Utilization -- Mathematical models. Health facilities -- Utilization -- Mathematical models.
49	Affordable housing : new policies for the housing and mortgage markets	Housing policy -- United States. Housing -- United States -- Finance. Mortgages -- United States.

* - Searches for this title were limited to 1984.

APPENDIX C - SAMPLE TITLES AND SUBJECT HEADINGS

Title #	Title	Subject Headings
50	Policy studies on capital formation : a selected bibliography	Saving and investment -- Bibliography.
51	Statistics sources : a subject guide to data on industrial, business, social, educational, financial, and other topics for the United States and internationally	Statistics -- Bibliography. Statistical services.
52	Trade names dictionary : a guide to approximately 194,000 consumer-oriented trade names, brand names, product names, coined names, model names, and design names, and names and addresses of their manufacturers, importers, marketers, or distributors	Trademarks -- United States. Business names -- United States.
53*	Economics : what went wrong, and why, and some things to do about it	Economics.
54	Innovation and entrepreneurship : practice and principles	Small business -- United States. New business enterprises -- United States. Entrepreneurship.
55	Neoclassical political economy : the analysis of rent-seeking and DUP activities	Neoclassical school of economics -- Congresses. Rent -- Congresses. Monopolies -- Congresses. Competition -- Congresses.
56	Macroeconomic conflict and social institutions	Income distribution. Social institutions. Institutional economics. Macroeconomics.

* - Searches for this title were limited to 1983.

APPENDIX C - SAMPLE TITLES AND SUBJECT HEADINGS

Title #	Title	Subject Headings
57	The rules of the game : the logical structure of economic theories	Economics -- Methodology.
58	Marx, an introduction	Marx, Karl, -- 1818-1883.
59	Aspects of efficiency in a socialist developing country, Iraq.	Industrial productivity -- Iraq -- Mathematical models. Efficiency, Industrial -- Iraq -- Mathematicals models.
60	Planning the Mexican economy : alternative development strategies.	Mexico -- Economic conditions -- 1970-1982. Mexico -- Economic conditions -- 1970- -- Mathematical models. Agriculture -- Economic aspects -- Mexico. Mexico -- Economic conditions -- 1982-
61	Making America work again	United States -- Economic policy -- 1981-
62	Business and technological dynamics in newly industrializing Asia	Asia -- Industries -- Case studies. Technological innovations -- Asia -- Case studies.
63	Rhythms in politics and economics	Political stability -- Economic aspects. Business cycles -- Political aspects.
64	Mathematical models in agriculture : a quantitative approach to problems in agriculture and related sciences	Agriculture -- Mathematical models.
65	Market demand : an analysis of large economies with non-convex preferences	Demand (Economic theory) -- Mathematical models. Smoothing (Statistics).

APPENDIX C - SAMPLE TITLES AND SUBJECT HEADINGS

Title #	Title	Subject Headings
66	The threat of Japanese multinationals : how the West ca·. respond	Corporations, Japanese. International business enterprises.
67	Developing assistance policies and the performance of aid agencies : studies in the performance of DAC, OPEC, the regional development banks, and the World Bank Group.	Economic assistance -- Evaluation.
68	Managing in turbulent times.	Industrial management.

APPENDIX D - NUMBER OF RECORDS RETRIEVED

Title #	Keyword(s)	Records Retrieved (Keyword)	Records Retrieved (Phrase)
1	macroeconomics	197	197
	keynesian	45	45
	monetarist	24	24
	marxist views	2	2
2	norwegian economy	4	3
	1920-1980	18	18
3	silicon valley fever	4	4
	growth	6516	6516
	high-technology culture	4	4
4	mechanics	1443	1443
	baltimore	497	497
	workers	2767	2767
	politics	4381	4381
	age (omitted)	4254	4254
	revolution	299	299
	1763-1812	2	2
5	crisis	2422	2422
	soviet agriculture	22	10
6	west german economy	4	4
7	agricultural computer guide	10	1
	directory (omitted)	11585	11585
	here's	237	237
	decide if	2	1
	computer is*	13882	13882
	your future	146	101

* "is" is treated as a stopword by *FirstSearch*

APPENDIX D - NUMBER OF RECORDS RETRIEVED

Title #	Keyword(s)	Records Retrieved (Keyword)	Records Retrieved (Phrase)
8	beyond monetarist	1	1
	finding	914	914
	road	3743	3743
	stable money	13	12
9	theory (omitted)	8168	8168
	international trade	1206	737
10	world economy	310	207
	changes	3667	3667
	challenges	689	689
11	political economy	767	741
	china's changing relations	2	2
	southeast asia	402	387
12	business	14743	14743
	its public	213	63
13	socio-economic accounting	1	1
14	accounting	3121	3121
	pensions	283	283
	results (omitted)	3020	3020
	applying	353	353
	FASB's preliminary views	2	2
	federal accounting standards bureau preliminary views	0	0
15	atlas' tax aspects	1	1
	real estate transactions	103	89

APPENDIX D - NUMBER OF RECORDS RETRIEVED

Title #	Keyword(s)	Records Retrieved (Keyword)	Records Retrieved (Phrase)
16	industrial structure	55	26
	pricing	906	906
	inflation	637	637
17	economic analysis	1653	846
	technological change	252	238
18	energy crisis ten years after	9	2
19	mass unemployment	6	6
	plant closings	38	38
	community mental health	204	155
20	japan's reshaping	1	1
	american labor law	15	3
21	birth	1264	1264
	solidarity	140	140
	gdansk negotiations	3	3
	1980 (omitted)	5371	5371
22	100 best companies	4	4
	one hundred best companies	3	3
	work	8539	8539
	America (omitted)	9251	9251
23	comparative international budgeting	1	1
	finance	4008	4008
24	macroeconomic theory	19	14
	survey (omitted)	21751	21751

APPENDIX D - NUMBER OF RECORDS RETRIEVED

Title #	Keyword(s)	Records Retrieved (Keyword)	Records Retrieved (Phrase)
25	public enterprise economics	8	3
26	keynes	96	96
	instability	271	271
	capitalism	393	393
27	american enterprise	50	35
	foreign markets	54	9
	studies (omitted)	28724	28724
	singer	139	139
	international harvester	14	14
	imperial russia	17	16
28	rise	863	863
	corporate economy	17	2
29	inequality	301	301
	poverty	1234	1234
	malaysia	721	721
	measurement (omitted)	2781	2781
	decomposition	322	322
30	cuba	253	253
	dilemmas	286	286
	revolution	2020	2020
31*	industrial structure	20	9
	policy	5524	5524
	less developed countries	27	26

* - Searches for this title were limited to 1985.

APPENDIX D - NUMBER OF RECORDS RETRIEVED

Title #	Keyword(s)	Records Retrieved (Keyword)	Records Retrieved (Phrase)
32	promote prosperity	1	1
	u.s. domestic policy	20	1
	united states domestic policy	8	0
	mid-1980s	17	17
33	hidden spending	1	1
	politics	4380	4380
	federal credit programs	10	9
34	multinational excursions	1	1
35	international economy since 1945	2	2
36	banking deregulation	43	14
	new competition	59	9
	financial services	409	259
37	exclusive economic zone	42	42
	latin american perspective	12	6
38	shopping center development	13	6
39	youth	2801	2801
	expectations	895	895
	transitions	283	283
40	american jobs	7	1
	changing industrial base	3	1
41	plant closure policy dilemma	1	1
	labor	5025	5025
	law	16883	16883
	bargaining	650	650

APPENDIX D - NUMBER OF RECORDS RETRIEVED

Title #	Keyword(s)	Records Retrieved (Keyword)	Records Retrieved (Phrase)
42	elements	1452	1452
	industrial relations	685	614
43	multinational enterprises	68	64
	OECD industrial relations guidelines	1	1
	organisation for economic co-operation and development industrial relations guidelines	0	0
44	mediators	90	90
45	negotiating	230	230
	labor contract	21	13
	management handbook	401	144
46	comparative industrial relations	21	4
	trans-atlanticdialogue	1	1
47*	multidisciplinary perspectives	3	3
	population	1983	1983
	conflict	684	684
48	forecasting use	41	1
	health services	1751	862
	provider's guide	31	6
49	affordable housing	99	84
	new policies	109	12
	housing	6430	6430
	mortgage markets	37	33
	twentieth century fund report	6	1

* - Searches for this title were limited to 1984.

APPENDIX D - NUMBER OF RECORDS RETRIEVED

Title #	Keyword(s)	Records Retrieved (Keyword)	Records Retrieved (Phrase)
50	policy studies	1190	488
	capital formation	79	70
	selected bibliography (omitted)	1238	737
51	statistics sources	36	6
	subject guide	110	53
	data (omitted)	14261	14261
	industrial	6944	6944
	business	14742	14742
	social	16882	16882
	educational	5601	5601
	financial	7704	7704
	other topics	35	18
	united states (omitted)	23970	23944
	internationally	18	18
52	trade names dictionary	7	7
	guide (omitted)	41593	41593
	approximately 194,000 consumer-oriented trade names	1	1
	brand names	21	21
	product names	9	6
	coined names	6	6
	model names	6	6
	design names	7	6
	names	1132	1132
	addresses	318	318
	their manufacturers	18	17
	importers	50	50
	marketers	45	45
	distributors	97	97

APPENDIX D - NUMBER OF RECORDS RETRIEVED

Title #	Keyword(s)	Records Retrieved (Keyword)	Records Retrieved (Phrase)
53*	economics	2607	2607
	what went wrong	4	4
	why	511	511
	some things	8	7
	do about it	63	28
54	innovation	1026	1026
	enterpreneurship	229	229
	practice (omitted)	9398	9398
	principles (omitted)	3296	3296
55	neoclassical political economy	1	1
	analysis (omitted)	31313	31313
	rent-seeking	8	8
	DUP activities	4	1
	directly-unproductive profit-seeking activities	2	1
56	macroeconomic conflict	2	2
	social institutions	50	13
57	rules	3932	3932
	game	2097	2097
	logical structure	10	5
	economic theories	29	8
58	marx	210	210
	introduction (omitted)	7173	7173
59	aspects	4548	4548
	efficiency	1459	1459
	socialist developing country	2	1
	iraq	114	114

* - Searches for this title were limited to 1983.

APPENDIX D - NUMBER OF RECORDS RETRIEVED

Title #	Keyword(s)	Records Retrieved (Keyword)	Records Retrieved (Phrase)
60	planning	13451	13451
	mexican economy	26	14
	alternative development strategies	20	6
61	making america work again	1	1
62	business	14742	14742
	technological dynamics	5	1
	newly industrializing asia	2	1
63	rhythms	169	169
	politics	4382	4382
	economics	7877	7877
64	mathematical models	146	111
	agriculture	5850	5850
	quantitative approach	24	15
	problems	7530	7530
	related sciences	62	15
65	market demand	91	36
	analysis (omitted)	31313	31313
	large economies	4	2
	non-convex preferences	1	1
66	threat	426	426
	japanese multinationals	12	7
	west can respond	2	1

APPENDIX D - NUMBER OF RECORDS RETRIEVED

Title #	Keyword(s)	Records Retrieved (Keyword)	Records Retrieved (Phrase)
67	development assistance policies	10	2
	performance	9657	9657
	aid agencies	19	9
	studies (omitted)	28778	28778
	DAC	11	11
	development assistance committee	29	8
	OPEC	106	106
	organization of petroleum exporting countries	6	6
	regional development banks	2	2
	world bank group	6	2
68	managing	1945	1945
	turbulent times	10	9

APPENDIX E - NON-SAMPLED TITLES

	Precision			Recall		
	Keyword	Phrase	Difference	Keyword	Phrase	Difference
1	0.2435	0.2435	0	0.1283	0.1283	0
2	0.5	0.5	0	0.5	0.5	0
6	0.6667	0.6667	0	0.1317	0.1317	0
13						
15	0.2181	0.2528	-0.0347	0.2018	0.2018	0
18	0.2813	0.5	-0.2187	0.0471	0.0105	0.0366
19	0.4653	0.4725	-0.0072	0.225	0.1979	0.0271
20	0.75	0.875	-0.125	0.0395	0.0066	0.0329
24	0.8056	0.8846	-0.079	0.0503	0.0399	0.0104
25	1	2	-1	0.0417	0.0238	0.0179
26	0.3545	0.3545	0	0.1986	0.1986	0
27	0.0251	0.0288	-0.0037	0.2	0.1555	0.0445
32	0.1456	0.0469	0.0987	0.0027	0.0011	0.0016
34						
35	2	2	0	0.0031	0.0031	0
36	0.2271	0.1827	0.0444	0.1044	0.0805	0.0239
37	0.1203	0.1476	-0.0273	0.5163	0.5163	0
38	0.6923	0.9167	-0.2244	0.4737	0.2895	0.1842
40	0.0625			0.0095		
43	0.153	0.1627	-0.0097	0.0443	0.0443	0
44	0.0393	0.0393	0	0.1429	0.1429	0
45	0.105	0.124	-0.019	0.3196	0.3137	0.0059
46	0.6	1	-0.4	0.5455	0.1364	0.4091
55	0	0	0	0	0	0
56	0.5306	0.5625	-0.0319	0.0055	0.0034	0.0021
58	0.5144	0.5144	0	0.5827	0.5827	0
61						
65	0.0028	0	0.0028	0.0062	0	0.0062
66	0.9414	0.8631	0.0783	0.0175	0.0113	0.0062
Mean	0.440169	0.533532	-0.07826	0.174535	0.148792	0.032344
Std. dev	0.429494	0.533525	0.214317	0.189595	0.168771	0.085181
Error	0.13856	0.17553	0.07051	0.061165	0.055526	0.028024
M + Err	0.578729	0.709062	-0.00775	0.2357	0.204318	0.060368
M - Err	0.30161	0.358002	-0.14877	0.113369	0.093266	0.00432

Error estimated with z-statistic:

Keyword Precision Error = 1.645 * Std dev / sqrt(26)
Phrase Precision Error = 1.645 * Std dev / sqrt(25)
Keyword Recall Error = 1.645 * Std dev/ sqrt(26)
Phrase Recall Error = 1.645 * Std dev/ sqrt(25)

APPENDIX F - SAMPLED TITLES

	Precision			Recall		
	Keyword	Phrase	Difference	Keyword	Phrase	Difference
3	1.444	1.444	0	1.7333	1.7333	0
4	0.016	0.016	0	0.1578	0.1578	0
5	0.4	0.5568	-0.1568	0.3149	0.1896	0.1253
7	0.0498	0.0158	0.034	0.0114	0.012	-0.0006
8	0.2569	0.25	0.0069	0.0089	0.0087	0.0002
9	0.035	0.0425	-0.0075	0.0651	0.0791	-0.014
10	0.0237	0.0246	-0.0009	0.2333	0.1639	0.0694
11	0.7062	0.7084	-0.0022	0.3825	0.3843	-0.0018
12	0.0103	0.0038	0.0065	0.007	0.0015	0.0055
14	0.1405	0.1405	0	0.1607	0.1607	0
16	0.0845	0.0842	0.0003	0.0122	0.0114	0.0008
17	0.0867	0.086	0.0007	0.157	0.1471	0.0099
21	0.3417	0.3417	0	0.7333	0.7333	0
22	1.6733	1.6733	0	0.2456	0.2456	0
23	0.0175	0.0175	0	0.6364	0.6364	0
28	0.056	0.0013	0.0547	0.3333	0.0417	0.2916
29	0.0429	0.0429	0	0.2415	0.2415	0
30	0.0977	0.0977	0	0.2861	0.2861	0
31	0.2061	0.2094	-0.0033	0.019	0.0188	0.0002
33	0.3611	0.3438	0.0173	0.0268	0.0227	0.0041
39	0.0895	0.0895	0	0.3336	0.3336	0
41	0.1067	0.1067	0	0.0853	0.0853	0
42	0.225	0.38	-0.155	0.0946	0.1124	-0.0178
47	0.125	0.125	0	0.0554	0.0554	0
48	0.0283	0.0625	-0.0342	0.2564	0.2308	0.0256
49	0.2789	0.292	-0.0131	0.05	0.0467	0.0033
50	0.0586	0.0587	-0.0001	0.3937	0.4362	-0.0425
51	0.0894	0.1457	-0.0563	0.0978	0.0664	0.0314
52	1.0213	1.2066	-0.1853	0.1132	0.1121	0.0011
53	0.048	0.048	0	0.0188	0.0188	0
54	0.4075	0.4075	0	0.1477	0.1477	0
57	0.0089	0.0179	-0.009	0.0086	0.0043	0.0043
59	0.0061	0.0081	-0.002	0.25	0.3333	-0.0833
60	0.3871	0.5816	-0.1945	0.2118	0.151	0.0608
62	0	0	0			
63	0	0	0	0	0	0
64	0.0186	0.0192	-0.0006	0.1418	0.1418	0
67	0.1121	0.1433	-0.0312	0.1021	0.0699	0.0322
68	0.1118	0.1188	-0.007	0.0561	0.0561	0
Mean	0.235197	0.254136	-0.01894	0.215342	0.202034	0.013308
Std dev.	0.370665	0.387611	0.054777	0.299868	0.302174	0.054923
Error	0.097637	0.102101	0.014429	0.080021	0.080636	0.014657
M + Err	0.332835	0.356237	-0.00451	0.295363	0.282671	0.027964
M - Err,	0.13756	0.152035	-0.03337	0.135321	0.121398	-0.00135

Error estimated with z-statistic:

Precision Error = 1.645 * Std dev / sqrt(39)
Recall Error = 1.645 * Std dev / sqrt(38)

APPENDIX G - TITLES WITH THREE OR MORE SINGLE KEYWORDS

NON-SAMPLED TITLES

	Precision			Recall		
	Keyword	Phrase	Difference	Keyword	Phrase	Difference
1	0.2435	0.2435	0	0.1283	0.1283	0
26	0.3545	0.3545	0	0.1986	0.1986	0

SAMPLED TITLES

	Precision			Recall		
	Keyword	Phrase	Difference	Keyword	Phrase	Difference
4	0.016	0.016	0	0.1578	0.1578	0
14	0.1405	0.1405	0	0.1607	0.1607	0
21	0.3417	0.3417	0	0.7333	0.7333	0
29	0.0429	0.0429	0	0.2415	0.2415	0
30	0.0977	0.0977	0	0.2861	0.2861	0
39	0.0895	0.0895	0	0.3336	0.3336	0
41	0.1067	0.1067	0	0.0853	0.0853	0
51	0.0894	0.1457	-0.0563	0.0978	0.0664	0.0314
52	1.0213	1.2066	-0.1853	0.1132	0.1121	0.0011
54	0.4075	0.4075	0	0.1477	0.1477	0
59	0.0061	0.0081	-0.002	0.25	0.3333	-0.0833
63	0	0	0	0	0	0
64	0.0186	0.0192	-0.0006	0.1418	0.1418	0
Mean	0.182915	0.2017	-0.01878	0.211446	0.215354	-0.00391
Std. dev	0.270337	0.314063	0.050324	0.17382	0.178515	0.024381
Error	0.139074	0.161569	0.025889	0.089421	0.091836	0.012543
M + Err	0.32199	0.363269	0.007105	0.300867	0.30719	0.008635
M - Err	0.043841	0.040131	-0.04467	0.122025	0.123517	-0.01645

Error estimated with T-statistic for 12 degrees of freedom:

Error =1.782 * Std dev / sqrt(12)

APPENDIX H - TITLES WITH FOUR OR MORE KEYWORD GROUPS

NON-SAMPLED TITLES

	Precision			Recall		
	Keyword	Phrase	Difference	Keyword	Phrase	Difference
1	0.2435	0.2435	0	0.1283	0.1283	0
27	0.0251	0.0288	-0.0037	0.2	0.1555	0.0445

SAMPLED TITLES

	Precision			Recall		
	Keyword	Phrase	Difference	Keyword	Phrase	Difference
4	0.016	0.016	0	0.1578	0.1578	0
7	0.0498	0.0158	0.034	0.0114	0.012	-0.0006
8	0.2569	0.25	0.0069	0.0089	0.0087	0.0002
14	0.1405	0.1405	0	0.1607	0.1607	0
21	0.3417	0.3417	0	0.7333	0.7333	0
29	0.0429	0.0429	0	0.2415	0.2415	0
41	0.1067	0.1067	0	0.0853	0.0853	0
49	0.2789	0.292	-0.0131	0.05	0.0467	0.0033
51	0.0894	0.1457	-0.0563	0.0978	0.0664	0.0314
52	1.0213	1.2066	-0.1853	0.1132	0.1121	0.0011
53	0.048	0.048	0	0.0188	0.0188	0
57	0.0089	0.0179	-0.009	0.0086	0.0043	0.0043
59	0.0061	0.0081	-0.002	0.25	0.3333	-0.0833
67	0.1121	0.1433	-0.0312	0.1021	0.0699	0.0322
64	0.0186	0.0192	-0.0006	0.1418	0.1418	0
Mean	0.169187	0.186293	-0.01711	0.145413	0.146173	-0.00076
Std. dev	0.249434	0.291904	0.048683	0.174246	0.180625	0.024473
Error	0.117395	0.137384	0.022912	0.082008	0.085011	0.011518
M + Err	0.286582	0.323677	0.005806	0.227422	0.231184	0.010758
M - Err	0.051791	0.048909	-0.04002	0.063405	0.061163	-0.01228

Error estimated with t-statistic for 14 degrees of freedom:

Precision Error =1.761 * Std dev / sqrt(14)
Recall Error =1.761 * Std dev / sqrt(14)

APPENDIX I - TITLES WITH THREE OR MORE SUBJECT HEADINGS

NON-SAMPLED TITLES

	Precision			Recall		
	Keyword	Phrase	Difference	Keyword	Phrase	Difference
18	0.2813	0.5	-0.2187	0.0471	0.0105	0.0366
19	0.4653	0.4725	-0.0072	0.225	0.1979	0.0271
26	0.3545	0.3545	0	0.1986	0.1986	0
27	0.0251	0.0288	-0.0037	0.2	0.1555	0.0445
32	0.1456	0.0469	0.0987	0.0027	0.0011	0.0016
36	0.2271	0.1827	0.0444	0.1044	0.0805	0.0239
43	0.153	0.1627	-0.0097	0.0443	0.0443	0
55	0	0	0	0	0	0
56	0.5306	0.5625	-0.0319	0.0055	0.0034	0.0021
Mean	0.2425	0.256733	-0.01423	0.091956	0.076867	0.015089
Std. dev	0.173332	0.206829	0.080895	0.087612	0.080399	0.016942
Error	0.113985	0.136013	0.053198	0.057614	0.052871	0.011141
M + Err	0.356485	0.392746	0.038964	0.14957	0.129738	0.02623
M - Err	0.128515	0.120721	-0.06743	0.034341	0.023996	0.003948

Error estimated with t-statistic for 8 degrees of freedom:

Precision Error = 1.860 * Std dev / sqrt(8)
Recall Error = 1.860 * Std dev / sqrt(8)

SAMPLED TITLES

	Precision			Recall		
	Keyword	Phrase	Difference	Keyword	Phrase	Difference
3	1.444	1.444	0	1.7333	1.7333	0
4	0.016	0.016	0	0.1578	0.1578	0
5	0.4	0.5568	-0.1568	0.3149	0.1896	0.1253
7	0.0498	0.0158	0.034	0.0114	0.012	-0.0006
8	0.2569	0.25	0.0069	0.0089	0.0087	0.0002
12	0.0103	0.0038	0.0065	0.007	0.0015	0.0055
16	0.0845	0.0842	0.0003	0.0122	0.0114	0.0008
22	1.6733	1.6733	0	0.2456	0.2456	0
28	0.056	0.0013	0.0547	0.3333	0.0417	0.2916
31	0.2061	0.2094	-0.0033	0.019	0.0188	0.0002
39	0.0895	0.0895	0	0.3336	0.3336	0
41	0.1067	0.1067	0	0.0853	0.0853	0
47	0.125	0.125	0	0.0554	0.0554	0
48	0.0283	0.0625	-0.0342	0.2564	0.2308	0.0256
49	0.2789	0.292	-0.0131	0.05	0.0467	0.0033
54	0.4075	0.4075	0	0.1477	0.1477	0
60	0.3871	0.5816	-0.1945	0.2118	0.151	0.0608
Mean	0.330582	0.3482	-0.01762	0.234329	0.204171	0.030159
Std Dev.	0.468937	0.47752	0.060745	0.392851	0.394147	0.072624
Error	0.204691	0.208437	0.026515	0.171479	0.172045	0.0317
M + Err	0.535273	0.556637	0.008898	0.405809	0.376216	0.061859
M - Err	0.125891	0.139763	-0.04415	0.06285	0.032125	-0.00154

Error estimated with t-statistic for 16 degrees of freedom:

Precision Error = 1.746 * Std dev / sqrt(16)
Recall Error = 1.746 * Std dev / sqrt(16)

APPENDIX J - TITLES WITH A SINGLE SUBJECT HEADING

NON-SAMPLED TITLES

	Precision			Recall		
	Keyword	Phrase	Difference	Keyword	Phrase	Difference
1	0.2435	0.2435	0	0.1283	0.1283	0
2	0.5	0.5	0	0.5	0.5	0
6	0.6667	0.6667	0	0.1317	0.1317	0
13						
15	0.2181	0.2528	-0.0347	0.2018	0.2018	0
24	0.8056	0.8846	-0.079	0.0503	0.0399	0.0104
34						
38	0.6923	0.9167	-0.2244	0.4737	0.2895	0.1842
44	0.0393	0.0393	0	0.1429	0.1429	0
58	0.5144	0.5144	0	0.5827	0.5827	0
61						
Mean	0.459988	0.50225	-0.04226	0.276425	0.2521	0.024325
Std. dev	0.250765	0.293397	0.073749	0.193689	0.180782	0.060523
Error	0.179608	0.210144	0.052822	0.138729	0.129483	0.043349
M + Err	0.639596	0.712394	0.01056	0.415154	0.381583	0.067674
M - Err	0.280379	0.292106	-0.09508	0.137696	0.122617	-0.01902

Error estimated with t-statistic for 7 degrees of freedom:

Precision Error = 1.895 * Std dev/sqrt(7)
Recall Error = 1.895 * Std dev / sqrt(7)

SAMPLED TITLES

	Precision			Recall		
	Keyword	Phrase	Difference	Keyword	Phrase	Difference
10	0.0237	0.0246	-0.0009	0.2333	0.1639	0.0694
14	0.1405	0.1405	0	0.1607	0.1607	0
30	0.0977	0.0977	0	0.2861	0.2861	0
42	0.225	0.38	-0.155	0.0946	0.1124	-0.0178
50	0.0586	0.0587	-0.0001	0.3937	0.4362	-0.0425
53	0.048	0.048	0	0.0188	0.0188	0
57	0.0089	0.0179	-0.009	0.0086	0.0043	0.0043
67	0.1121	0.1433	-0.0312	0.1021	0.0699	0.0322
68	0.1118	0.1188	-0.007	0.0561	0.0561	0
Mean	0.091811	0.114389	-0.02258	0.150444	0.145378	0.005067
Std Dev.	0.062797	0.103948	0.047774	0.123087	0.131503	0.029371
Error	0.041296	0.068357	0.031417	0.080943	0.086478	0.019314
M + Err	0.133107	0.182746	0.008839	0.231387	0.231856	0.024381
M - Err	0.050515	0.046032	-0.05399	0.069502	0.0589	-0.01425

Error estimated with t-statistic for 8 degrees of freedom:

Precision Error = 1.860 * Std dev / sqrt(8)
Recall Error = 1.860 * Std dev / sqrt(8)

APPENDIX K - TITLES WITH A SINGLE KEYWORD OR KEYWORD GROUP

NON-SAMPLED TITLES

	Precision			Recall		
	Keyword	Phrase	Difference	Keyword	Phrase	Difference
6	0.6667	0.6667	0	0.1317	0.1317	0
13						
18	0.2813	0.5	-0.2187	0.0471	0.0105	0.0366
24	0.8056	0.8846	-0.079	0.0503	0.0399	0.0104
25	1	2	-1	0.0417	0.0238	0.0179
34						
35	2	2	0	0.0031	0.0031	0
38	0.6923	0.9167	-0.2244	0.4737	0.2895	0.1842
44	0.0393	0.0393	0	0.1429	0.1429	0
58	0.5144	0.5144	0	0.5827	0.5827	0
61						
Mean	0.74995	0.940213	-0.19026	0.18415	0.153013	0.031138
Std. dev	0.549846	0.662524	0.319145	0.205214	0.185772	0.059099
Error	0.393823	0.474528	0.228585	0.146983	0.133058	0.042329
M + Err	1.143773	1.41474	0.038323	0.331133	0.28607	0.073467
M - Err	0.356127	0.465685	-0.41885	0.037167	0.019955	-0.01119

Error estimated with t-statistic for 7 degrees of freedom:

Precision Error =1.895 * Std dev / sqrt(7)
Recall Error =1.895 * Std dev / sqrt(7)

APPENDIX L - TITLES WITH KEYWORDS THAT MATCH SUBJECT HEADINGS

NON-SAMPLED TITLES

	Precision			Recall		
	Keyword	Phrase	Difference	Keyword	Phrase	Difference
1	0.2435	0.2435	0	0.1283	0.1283	0
6	0.6667	0.6667	0	0.1317	0.1317	0
26	0.3545	0.3545	0	0.1986	0.1986	0
36	0.2271	0.1827	0.0444	0.1044	0.0805	0.0239
56	0.5306	0.5625	-0.0319	0.0055	0.0034	0.0021
58	0.5144	0.5144	0	0.5827	0.5827	0
Mean	0.4228	0.420717	0.002083	0.191867	0.187533	0.004333
Std. dev	0.160525	0.174083	0.022222	0.183866	0.186309	0.008784
Error	0.144655	0.156872	0.020025	0.165688	0.16789	0.007916
M + Err	0.567455	0.577589	0.022108	0.357555	0.355423	0.012249
M - Err	0.278145	0.263844	-0.01794	0.026179	0.019644	-0.00358

Error estimated with z-statistic:

Precision Error = 2.015 * Std dev / sqrt(5)
Recall Error = 2.015 * Std dev / sqrt(5)

SAMPLED TITLES

	Precision			Recall		
	Keyword	Phrase	Difference	Keyword	Phrase	Difference
4	0.016	0.016	0	0.1578	0.1578	0
9	0.035	0.0425	-0.0075	0.0651	0.1639	-0.0988
11	0.7062	0.7084	-0.0022	0.3825	0.0015	0.381
16	0.0845	0.0842	0.0003	0.0122	0.0114	0.0008
23	0.0175	0.0175	0	0.6364	0.6364	0
30	0.0977	0.0977	0	0.2861	0.2861	0
39	0.0895	0.0895	0	0.3336	0.3336	0
42	0.225	0.38	-0.155	0.0946	0.1124	-0.0178
47	0.125	0.125	0	0.0554	0.0554	0
49	0.2789	0.292	-0.0131	0.05	0.0467	0.0033
53	0.048	0.048	0	0.0188	0.0188	0
54	0.4075	0.4075	0	0.1477	0.1477	0
64	0.0186	0.0192	-0.0006	0.1418	0.1418	0
Mean	0.165338	0.179038	-0.0137	0.183231	0.162577	0.020654
Std Dev.	0.192666	0.201179	0.040967	0.174109	0.167928	0.107312
Error	0.099111	0.10349	0.021074	0.089565	0.086385	0.055203
M + Err	0.26445	0.282529	0.007374	0.272796	0.248962	0.075857
M - Err	0.066227	0.075548	-0.03477	0.093666	0.076192	-0.03455

Error estimated with t-statistic for 12 degrees of freedom:

Precision Error = 1.782 * Std dev / sqrt(12)
Recall Error = 1.782 * Std dev / sqrt(12)

APPENDIX M - NON-SAMPLED TITLES EXCLUDING SINGLE KEYWORDS

	Precision			Recall		
	Keyword	Phrase	Difference	Keyword	Phrase	Difference
1	0	0	0	0	0	0
2	1	1	0	1	1	0
3	2.1667	2.1667	0	2.6	2.6	0
4						
5	0.7976	1.1111	-0.3135	0.6204	0.3704	0.25
6	0.6667	0.6667	0	0.1317	0.1317	0
8	0.7708	0.75	0.0208	0.0268	0.0261	0.0007
10	0.0661	0.0688	-0.0027	0.6833	0.475	0.2083
12	0.013	0	0.013	0.0109	0	0.0109
13						
14	0.5	0.5	0	0.0179	0.0179	0
15	0.2181	0.2528	-0.0347	0.2018	0.2018	0
16	0.0509	0.05	0.0009	0.0044	0.002	0.0024
18	0.2813	0.5	-0.2187	0.0471	0.0105	0.0366
19	0.4653	0.4725	-0.0072	0.225	0.1979	0.0271
20	0.75	0.875	-0.125	0.0395	0.0066	0.0329
21	1	1	0	0.8	0.8	0
22	2.5	2.5	0	0.3158	0.3158	0
23						
24	0.8056	0.8846	-0.079	0.0503	0.0399	0.0104
25	1	2	-1	0.0417	0.0238	0.0179
26						
27	0.0314	0.036	-0.0046	0.25	0.1944	0.0556
28	0.1094	0	0.1094	0.5833	0	0.5833
29						
30						
31	0.2979	0.3029	-0.005	0.0251	0.0247	0.0004
32	0.195			0.0036		
33	0.7222	0.6875	0.0347	0.0535	0.0453	0.0082
34						
35	2	2	0	0.0031	0.0031	0
36	0.2271	0.1827	0.0444	0.1044	0.0805	0.0239
37	0.1203	0.1476	-0.0273	0.5163	0.5163	0
38	0.6923	0.9167	-0.2244	0.4737	0.2895	0.1842
39						
40	0.0625			0.0095		
41						
43	0.153	0.1627	-0.0097	0.0443	0.0443	0
44						
45	0.0756	0.1042	-0.0286	0.0383	0.0294	0.0089
46	0.6	1	-0.4	0.5455	0.1364	0.4091
47	0	0	0	0	0	0
49	0.3161	0.3335	-0.0174	0.0525	0.048	0.0045
51	0.1813	0.3502	-0.1689	0.2289	0.1344	0.0945
52	1.6562	1.9738	-0.3176	0.1241	0.1223	0.0018
53	0	0	0	0	0	0
54						
55	0	0	0	0	0	0
56	0.5306	0.5625	-0.0319	0.0055	0.0034	0.0021
57	0.0179	0.0357	-0.0178	0.0173	0.0086	0.0087
58						
59	0			0		
60	0.5795	0.8712	0.2917	0.3149	0.2238	0.0911
61						
62	0	0	0			
63						
64	0.0034	0.0045	-0.0011	0.0364	0.0364	0
65	0.0028	0	0.0028	0.0067	0	0.0067
66	1.4091	1.2917	0.1174	0.0187	0.0093	0.0094
67	0.1477	0.1947	-0.047	0.0914	0.043	0.0484
68	0.1111	0.125	-0.0139	0.0092	0.0092	0
Mean	0.495628	0.592757	-0.06919	0.225485	0.191202	0.049709
Std. dev	0.601393	0.665647	0.179566	0.43037	0.428591	0.114981
Error	0.144303	0.163076	0.044531	0.104383	0.107516	0.028844
M + Err	0.639931	0.757833	-0.02466	0.329868	0.298719	0.078553
M - Err	0.351325	0.427681	-0.11372	0.121102	0.083686	0.020865

Error estimated with z statistic:

Keyword Precision Error = 1.645 * Std dev / sqrt(47)
Phrase Precision Error = 1.645 * Std dev / sqrt(44)

Keyword Recall Error = 1.645 * Std dev / sqrt(46)
Phrase Recall Error = 1.645 * Std dev / sqrt(43)

APPENDIX N - SAMPLED TITLES EXCLUDING SINGLE KEYWORDS

	Precision			Recall		
	Keyword	Phrase	Difference	Keyword	Phrase	Difference
7	0.0622	0.0238	0.0384	0.0143	0.0181	-0.0038
9	0.035	0.0425	-0.0075	0.0651	0.0791	-0.014
11	0.7062	0.7084	-0.0022	0.3825	0.3843	-0.0018
17	0.0867	0.086	0.0007	0.157	0.1471	0.0099
42	0.45	0.535	-0.085	0.1891	0.2248	-0.0357
48	0.0283	0.0625	-0.0342	0.2564	0.2308	0.0256
50	0.0586	0.0587	-0.0001	0.3937	0.4362	-0.0425
Mean	0.203857	0.2167	-0.01284	0.2083	0.2172	-0.0089
Std Dev.	0.247028	0.260895	0.035408	0.135317	0.141296	0.022339
Error	0.195949	0.206949	0.028087	0.107337	0.11208	0.01772
M + Err	0.399806	0.423649	0.015244	0.315637	0.32928	0.00882
M - Err	0.007908	0.009751	-0.04093	0.100963	0.10512	-0.02662

Error estimated with t-statistic for 6 degrees of freedom:

Precision Error = 1.943 * Std dev / sqrt(6)
Recall Error = 1.943 * Std dev / sqrt(6)