

情報検索分野の構造：引用調査による下位領域と発展過程の分析

Structure of Information Retrieval Research: Tracking the
Specialties and Development of Research Using
Co-citation Maps and Citation Diagrams

神 門 典 子, 野 末 道 子, 榛 田 倫 子,
Noriko Kando, Michiko Nozue, Tomoko Harita,
村 上 匡 人, 谷 津 真理子, 上 田 修 一
Tadato Murakami, Mariko Yatsu, Shuichi Ueda

Résumé

This paper attempts to clarify the structure of the field of information retrieval. Co-citation maps and citation diagrams are used to identify specialties and to trace the development of research on information retrieval. In the first-phase co-citation map, following specialties are identified: 1) information retrieval theories, 2) programming and algorithms, 3) fulltext databases and online database retrieval, 4) user modeling in the context of information retrieval, 5) various data structures for information retrieval systems. The developments of research in each specialties are depicted using citation diagrams. To conduct further analysis, the second-phase co-citation maps, i.e. co-citation maps for another aggregate co-citation data, consists of the documents which highly cocited with those are in the each specialties of the first-phase co-citation map. We also identifies the basic literature of each specialties. We discussed the validity of multi-phased co-citation maps and combination with citation diagrams as methods to clarify the structure of disciplines.

I. はじめに

II. 方法

A. 共引用マップと文献間引用

B. 調査手順

神門典子, 野末道子, 榛田倫子, 谷津真理子: 慶應義塾大学大学院文学研究科図書館・情報学専攻, 東京都港区三田 2-15-45.

Noriko Kando; Michiko Nozue; Tomoko Harita; Mariko Yatsu: Graduate School of Library and Information Science, Keio University, 2-15-45, Mita, Minato-ku, Tokyo.

村上匡人: 慶應義塾大学文学部図書館・情報学科, 東京都港区三田 2-15-45.

Tadato Murakami: School of Library and Information Science, Keio University, 2-15-45, Mita, Minato-ku, Tokyo.

上田修一: 慶應義塾大学文学部図書館・情報学科教授, 東京都港区三田 2-15-45.

Shuichi Ueda: Professor, School of Library and Information Science, Keio University, 2-15-45, Mita, Minato-ku, Tokyo.

1992年2月13日受付

III. 情報検索領域の構造

A. 共引用マップ

B. 基礎文献間の引用関係

IV. 検索理論とユーザーモデル領域の分析

A. 検索理論

B. ユーザーモデル

V. おわりに

I. はじめに

情報検索は、図書館・情報学の主要な研究領域の一つであり、情報処理の分野でも研究発表が行われている。図書館・情報学の中では、情報検索は一定数の専門研究者が存在し、論文数が多く、研究成果の積み重ねの努力がなされている数少ない領域である。

周知のように研究領域としての情報検索は 1950 年代に誕生したが、初期の短い間に情報検索を構成する基本的な概念は形成されている。すなわち、再現率や精度などの評価基準が確立し、事後結合索引法、抽出索引法、ブール演算子、転置ファイル方式などが考え出されている。その後、情報検索は二つの道をたどった。

一つの方法はその実用化であり、1960 年代におけるオンライン情報検索システムの開発とデータベースの形成、そして 1970 年代初頭のこれらを組み合わせた商用データベース検索サービスの出現へと続く。これは、基本的に上述の事後結合索引法、抽出索引法、ブール演算子、転置ファイル方式を用いたものである。たとえば、「現在 JICST や DIALOG などの商用のオンライン検索サービスで使用されている各種の近接演算子は AND 演算子の拡張に過ぎず、1950 年代の検索技法とは基本的に差異はない」¹⁾ という評価があるように、現在の商用データベースサービスは索引と検索の面からみれば、かなり古い技術を使っている。

商用データベースサービスで使われている各種の手法は、データベースの量的増大、すなわち大規模化という問題に直面したが、コンピュータ本体の高速化とファイルの大容量化、価格低下によって凌ぐことができていく。数百万件を収録する大規模データベースは、小規模なデータベースとは性質として異なる問題を持っているが、少なくとも現在まではこれは顕在化していない。また、オンライン環境下における対話による探索方式は、検索手法の弱点をカバーしつつ、利用者の探索行動に関

する新しい研究領域を生むことになった。さらに、商用データベースサービスの利用者層の拡大は、情報検索研究の意義を一般に認知させるのに役立っているという側面もある。今日、多数の検索技術に関する教科書が刊行されているが、これらは商用データベースサービスで使われている検索手法の解説であって、後述する現在の検索手法の理論的な研究には言及しないのが常である。

情報検索がたどったもう一つの道は、検索手法の理論的、実験的な研究である。ここでは、データベースの大規模化には背を向けて、もっぱら小規模なデータによって検索実験を行い、新しい検索理論や検索手法の確立を目指してきた。これらの研究は主として *Information Processing and Management* 誌や *Journal of the American Society for Information Science* 誌に発表され、G. Salton²⁾や C. J. Van Rijsbergen³⁾、最近では D. Ellis⁴⁾の教科書が書かれている。検索手法についての文献レビューの中で N. J. Belkin⁵⁾が整理しているように、確率型、ベクトル空間型をはじめとして、多くの検索モデルが開発されてきた。さらに、ファジィ、エキスパート・システムあるいはニューロ・ネットワークなどの流行現象のいち早い取り込みもなされてきた。しかし、D. C. Blair⁶⁾が指摘しているように、検索効率の絶対的な評価基準がないために、多様な検索理論が並列的に存在し、検索理論の研究は拡散しつつある。

このように情報検索の領域は大きく二つに分かれ、実用化したオンラインデータベース検索においては、従来の手法では全文データベースの検索や大規模データベースにおける出力過多 (output overload) に対処できないという問題を抱えている。もう一方の検索手法研究においては、多様な検索理論が存在するために全体像が捉えにくくなりつつある。また、この領域は教科書やレビューが多くはなく、研究に新規参入するのが困難になっている。

どのような研究領域でも同様であるが、情報検索の研

究を行う際にも、研究の現状を明らかにし、必ず参照すべき基礎的な文献や情報検索分野の発展過程を示すことが必要である。

そこで、最近2年間の情報検索に関する文献の引用と共引用とを調査することによって、情報検索分野の下位領域を明らかにするとともに、基礎的な文献を同定し、情報検索に関する主要な研究の流れとその発展過程とを明らかにすることを試みた。

II. 方 法

A. 共引用マップと文献間引用

研究文献をもとに現在の情報検索研究を構成する研究領域や検索理論を示すには、情報検索分野に属する文献を同定し、その中で互いに類似する文献をグループ化することが必要となる。類似する文献をグループ化する手法としては、付与されたディスクリプタの共出現、あるいは標題や抄録中に出現する語によるクラスタ化など様々な方法がある。最近、岸田和明と池谷のぞみ⁷⁾は、共引用を用いる方法が、他の手法よりも妥当な結果を導き出すことを明らかにしている。

共引用とは、二つの文献が同一の文献に引用されている状態を示す。この共引用によって文献間の類似性を測定する手法は広く使われてきた。共引用の強さは二つの文献を同時に引用している文献の数によって測定される。すなわち、AとBという二つの論文を引用している論文の数が多ければ、AとBの類似性は高いと判断する。この数値は「共引用強度」と呼ばれる。共引用強度に基づいて各文献を二次元以上の空間に配置して「共引用によるマップ」を作る方法は、Small⁸⁾によって開発され、共引用マップは文献群あるいは専門領域の構造を表現するために利用されてきた。

一般に単純な引用関係が文献間の類似性を表現するとは言いきれない。ある文献とその文献が引用している文献との関係は様々であって、主題が類似しているという保証はない。しかしながら、共引用強度のように引用関係を累積した場合、一定の操作を加えて、引用関係の持っている主題に関係する要素が引き出されることが期待できる。共引用をはじめとする引用に基づく諸尺度は、広く利用されているにもかかわらず、その妥当性の検証は十分に行われてきたとは言えない。前述の岸田⁷⁾は、共引用と他の尺度とを比較することによって、文献間の類似性を示す尺度としての共引用の相対的な優位を示している。

彼らは、情報学分野の文献を対象として、現在の情報検索システムで使われているディスクリプタあるいは語の共出現に基づく文献間の類似度と、共引用に基づく類似度との比較を行っている。ただし、この場合には別個に「文献間の類似性」を判断する基準が必要である。これについては、文献の主題を構成する研究目的、研究対象、方法などの要素や属性のカテゴリーを「認知的属性」とし、個々の文献に「属性値」を与え、個々の類似性を示す尺度がこれらの認知的属性を一致させる程度によって、妥当性を評価している。

具体的には、情報学分野の主要雑誌4誌の1987年から1989年の掲載論文454件の引用文献を抽出し、ここから一定回数以上引用された53文献を選定した。これらの文献群を対象として、共引用、標題と抄録中の語の共出現、付与されているディスクリプタの共出現を別個に算出し、多次元尺度構成法によるマッピングと階層的クラスタ分析によるクラスタ化を行い、これらを比較した。この結果、共引用マップは、対象文献群を「ビブリオメトリクスの法則」、「情報検索」、「科学計量学」の三つのグループに分離させることに成功した。さらに認知的属性を用いた分析でも、共引用によるクラスタによって示された文献の類似性は、付与されたディスクリプタや文献中に出現する語のように直接に概念を扱っていないにもかかわらず、研究対象などの認知的属性をよく反映していることが明らかとなった。そして、ここで用いられた共引用は、文献の間の類似性を示す尺度として妥当性が高いと結論づけている。

このような成果に基づき、以下の分析では共引用マップによって文献のグループ化を行い、情報検索分野を構成している下位領域に関してより細かな分析を試みたい。また、ここでは情報検索領域の時系列的变化も検討するので、共引用分析に加えて文献間の直接の引用関係も調査する。

B. 調査手順

具体的な調査手順は以下の通りである(第1図参照)。

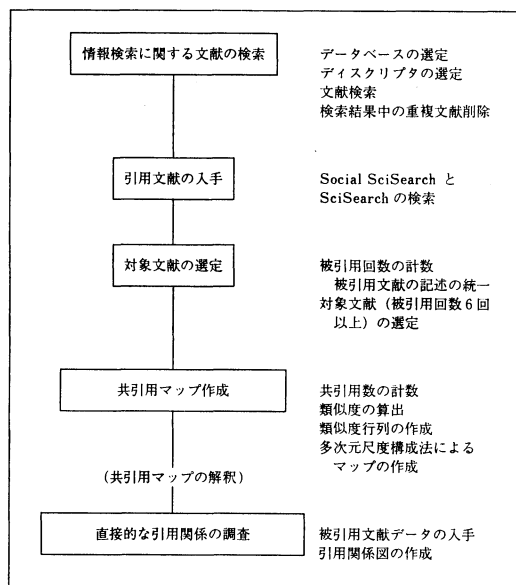
(1) 情報検索に関する文献の検索

最初に1989年と1990年の2年間に公表された情報検索に関する文献を検索した。これらの文献が引用している文献が分析対象となるのであるが、最近の情報検索研究の動向を反映させ、また被引用文献を入手する作業上の制約のために最近の2年間を対象とした。情報検索に関する文献を収録しているデータベースとして、

情報検索分野の構造：引用調査による下位領域と発展過程の分析

INSPEC, ERIC, LISA の 3 種類を選定した。

この最初の段階の検索によって、これ以後の段階の主題範囲が決まることになる。そのため、情報検索領域を



第 1 図 調査手順

示すディスクリプタの選定には注意を払った。各データベースのディスクリプタの種類や関係について、シソーラス (INSPEC と ERIC) と分類表 (LISA) によって検討し、予備検索を繰り返して、第 1 表に示すような検索語群を選んだ。しかし、これらの検索語によって検索される文献群中には、情報検索研究とは関係が薄い CD-ROM やオンラインデータベースの利用方法、ユーザインタフェース等のアプリケーションに関する文献も多く含まれることが想定されたが、次の理由からこれらの文献も含めて検索を行った。

- ① これらのアプリケーションに関する文献のみを適正に除外することは困難である。
- ② 検索式の中での「NOT」の多用は検索もれの危険を高める。
- ③ アプリケーションに関する文献は一般に引用文献数が少なく、共引用による分析結果に大きな影響がないと考えられる。

実際の検索は、DIALOG を用いた。DIALOG を使用したのは、「OneSearch」という複数データベースの検索結果から重複する文献を除去する機能があるためで、これを用いて検索結果から重複文献を取り除いた。これらの 3 つのデータベースにおける検索文献数は第 2 表に

第 1 表 各データベースで用いた検索語

ERIC (CIJE, メジャーディスクリプタのみに限定)	INSPEC	LISA
INFORMATION RETRIEVAL ONLINE SEARCHING SEARCH STRATEGIES SELECTIVE DISSEMINATION OF INFORMATION RELEVANCE (INFORMATION RETRIEVAL) DATABASES BIBLIOGRAPHIC DATABASES FULL TEXT DATABASES NUMERIC DATABASES INDEXING AUTOMATIC INDEXING INFORMATION STORAGE	INFORMATION RETRIEVAL BIBLIOGRAPHIC SYSTEMS INDEXING INFORMATION STORAGE	V? Z? (? はトランケーション) コードに対応するおおよその主題 V: SUBJECT INDEXING Vc: CONTROLLED TERM Z: Zj: COMPUTERIZED SUBJECT INDEXING Zk: COMPUTERIZED INFORMATION RETRIEVAL SEARCHING Zm: ONLINE INFORMATION RETRIEVAL ZmR: DATABASES ZmV: FULLTEXT SEARCH KEYS SEARCHING Zp: AUTOMATIC INDEXING

第2表 ERIC, INSPEC, LISA における検索文献数と重複除去後の文献数

		文献数
		(件)
検索文献数 (3 データベースの合計)		3,339
OneSearch の RD コマンドによる 重複除去後の文献数		2,652
手作業による除去後の文献数		
の除 内去 訳分	重複文献	66
	会議録中のもの	266
	特集のタイトル	4
		2,316

示した。なお、Miller⁹⁾ が報告しているように、書誌事項の記述のしかたによって、One Search では重複が除去できない場合もあるので、手作業によって重ねて確認をし、重複を削除した。

(2) 被引用文献の検索

上記で検索した文献が引用している文献は、*Social SciSearch* と *SciSearch* のデータベースから得ることができる。そこで NACSIS-IR において、これら二つのデータベースを用いて論文の書誌事項から検索を行った。(1) で得た検索結果のうち、*Social SciSearch* あるいは *SciSearch* に収録されていた文献数は 833 件であった。また、これらの 833 件の文献が引用している文献数は総計 11,033 件であり、1 文献あたり平均で 13.2 件引用していた。この 11,033 件の文献を被引用文献群とする。

(3) 対象文献群の選定

分析の対象とする文献群は、情報検索に関する文献から高い頻度で引用されている文献とし、それを抽出するために、(2) で得た被引用文献が (1) で得た文献から引用された回数を計数した。

なお、被引用文献の被引用回数を計数する際に大きな技術的な障害となるのは、*Social SciSearch* や *SciSearch* では被引用文献の記述が一定していないという

問題である。(2) で得た被引用文献群では、同一文献が異なる形で記述されている例は極めて多い。これらの被引用文献の同定は、人手によるしかないが、被引用文献数が大量であるため、同定作業を支援するシステムを作った。これを用いて、記述に不統一や誤りがあるレコードを検出し、490 件を他の被引用文献と同一のものと認定してその記述を統一した。最終的に、被引用文献 11,033 件中の異なり文献数は 9,178 件となった。したがって、11,033 件の被引用文献群において、2 回目以降として出現した被引用文献は合わせて 1,855 件であり、そのうち 490 件、すなわち 26.4% は記述の統一作業をしなければ異なる文献として扱われてしまうものであった。

記述を統一した被引用文献ごとに被引用回数を算出した。被引用文献の出現頻度分布を第3表に示した。今回は、対象とする文献群として被引用回数が6回以上である49件を選んだ。これらの書誌事項と内容の概要とを第4表に示した。

(4) 共引用マップの作成

前述の岸田ら⁷⁾の方法に従い、下記の式によって各文献の類似度を求めた。この類似度は、共引用強度をそれぞれの被引用回数によって正規化しているの、共引用強度に対する文献の被引用回数の大きさの影響を除去することができる。

文献AとBの類似度

$$= \frac{A \text{ と } B \text{ が共引用された回数}}{\sqrt{A \text{ の被引用回数} \times B \text{ の被引用回数}}}$$

類似度を各文献の組ごとに算出し、類似度行列を作成し、多次元尺度構成法の中の Kruskal の方法によって共引用マップを作成した。作業には慶應義塾大学三田計算センターの FACOM M760 上の統計パッケージ SPSS-X を用いた。

(5) 直接的な引用による分析

共引用マップに示された各文献間の関係を時系列的にみるために、各文献が引用している文献を調べ、文献相互の関係を時間軸上に配置した引用関係図を作成した。

第3表 被引用回数別にみた被引用文献数

被引用回数	1 回	2 回	3 回	4 回	5 回	6 回	7 回	8 回以上	計	
									異なり	延べ
被引用文献数 (件)	8,135	683	208	70	32	18	11	21	9,178	11,033

情報検索分野の構造：引用調査による下位領域と発展過程の分析

第4表 対象文献（情報検索分野。49件，グループ別）

	コード	書誌事項	内 容	被引用回数
検 索 理 論	BK83	Bookstein, A. Information-retrieval: A sequential learning-process. JASIS. 34(5): 331-342 (1983)	決定理論を用いた確率型検索手法にしレバンスフィードバックを組み込むためのモデルを提示している。*	8
	BK85	Bookstein, A. Probability and fuzzy-set applications to information retrieval. ARIST. 20: 117-151 (1985)	確率型検索とファジ理論を応用した検索についての理論、手法、研究結果および最新動向などをレビューしている。*	9
	LO	Lovins, B. J. Development of a stemming algorithm. Mechanical Translation and Computational Linguistics. 11: 22-31 (1968)	語幹処理のアルゴリズムを開発して、提示している。	6
	MA	Maron, M. E. et al. On relevance, probabilistic indexing and information retrieval. JACM. 7: 216-244 (1960)	レバンスを確率を用いて統計的に表現する事を提唱している初期の文献。	6
	RO76	Robertson, S. E. et al. Relevance weighting of search terms. JASIS. 27(3): 129-146 (1976)	確率型検索の初期の論文。適合文献の情報に基づき、重みづけを行う方法を開発し、その実験結果を提示した。*	14
	RO77	Robertson, S. E. Probability ranking principle in IR. J DOC. 33(4): 294-304 (1977)	確率型検索に対する Cooper が提起した疑義に答える形で書かれた、確率型検索に関するレビュー。	6
	SA71	Salton, G. The SMART retrieval system: experiments in automatic document processing. Prentice-Hall Inc., New Jersey, 1971, 556 p.	自動検索システムである SMART を様々な側面から論じている。特に検索の評価、言語の分析、クラスター化による検索、レバンスフィードバックなどが重点的に扱われている。*	11
	SA75a	Salton, G. et al. A vector-space model for automatic indexing. CACM. 18(11): 613-620 (1975)	ベクトル空間型検索モデルについての主要なアイデアを論じた初期の論文。*	6
	SA75b	Salton, G. et al. Theory of term importance in automatic text analysis. JASIS. 26(1): 33-44 (1975)	識別値を用いた自動索引法を論じている。語の出現頻度に基づくベクトル空間型検索についての初期の論文。*	6
	SA83a	Salton, G. et al. Extended boolean information-retrieval. CACM. 26(11): 1022-1036 (1983)	文献と検索質問に重みづけを行い、従来のブール検索にベクトル空間型検索を導入して拡張する試み。ファジ理論に言及している。*	8
	SA88	Salton G. et al. Paralell text search methods. CACM. 31(2): 202-215 (1988)	並列処理による高速テキスト検索。ブールテキスト検索、ベクトル空間型、SMART などをシリアルに用いた方法と比べて大きな効率の改善はみられなかった。	7
	SP	Sparck-Jones, K. A statistical interpretation of term specificity and its application in retrieval. J DOC. 28(1): 11-21 (1972)	各語が出現する文献の数によって語の重みづけを行う手法を提唱している。Salton のベクトル空間型検索に大きな影響を与えた。*	9
	ST	Stanfill, C. et al. Paralell free-text search on the connection machine system. CACM. 29(12): 1229-1239 (1986)	大規模データベース、特に全文データベースにおいて、並列処理システムにおいて高速テキスト検索を行う手法を提示している。	10
教 科 書	VR77	Van Rijsbergen, C. J. Theoretical basis for use of co-occurrence data in information retrieval. J DOD. 33(2): 106-119 (1977).	語の独立性を仮定しない統計的な重みづけ検索手法を論じている。*	6
	SA68	Salton, G. Automatic information organization and retrieval. McGraw-Hill, New York, 1968, 514 p.	自動的な情報検索に関する初期の教科書。SMART システムの概要や辞書の自動構築、文献の自動的な内容分析、自動抄録作成、等の手順を扱っている。*	7
	SA83b	Salton, G. et al. Introduction to modern information retrieval. McGraw-Hill, New York, 1983, 448 p.	情報検索の理論やシステムに関する包括的な教科書。検索理論の中心はベクトル空間モデルや SMART システムであるが、確率型検索などの手法も解説されている。後半では、ファイル構造などのハードウェアに関するトピックが扱われている。*	54
	VR79	Van Rijsbergen, C. J. Information Retrieval. 2nd ed. Butterworths, London, 1979, 152 p.	情報検索理論の一般的な教科書。特に自動索引、自動分類、検索の評価に重点が置かれている。*	37
プログラミング技法	HA	Harrison, M. C. Implementation of the substring test by hashing. CACM. 14(12): 777-779 (1971).	ハッシングのアルゴリズム。	6
	KN	Knuth, E. D. The art of computer programming. Vol. 3: Sorting and Searching. Addison-Wesley Pub., 1973	プログラミング技法に関する教科書で、ソートと探索を扱っている。	11
全文DBとオンライン検索	BL	Blair, D. C.; Maron, M. E. An evaluation of effectiveness for a full-text document-retrieval system. CACM. 28(3): 289-299 (1985).	大規模な全文データベースの検索効率を評価している。大規模データベースでは自然語による検索は再現率を低下させることを示している。IBM STAIRS を使用。	17
	DI	Directory of Online Databases; 1988. 1988.	オンラインデータベースのディレクトリ。	6
	LA	Lancaster, F. W. Vocabulary control for information retrieval. 2nd ed. Information Resources Press, 1986, 270 p.	語彙統制に関する教科書。主にデータベースで用いられることを前提とした事後統制語彙とシンソーラスを扱っている。自然言語検索と統制語彙の組み合わせ、交換言語、検索シンソーラスにも言及している。	7
	LI	Liddy, E. et al. A study of discourse anaphora in scientific abstracts. JASIS. 38(4): 255-261 (1987)	抄録中の前方照応表現を分類し、それによってテキスト構造を分析している。自然言語検索では、ディスコースレベルの処理が必要であるという前提に立つ。	6
	SA86	Salton, G. Another look at automatic text-retrieval systems. CACM. 29(7): 648-656 (1986)	IBM STAIRS による全文データベースの検索手法について、さらに自動検索システムについても論じている。*	12
	SW	Swanson, D. R. Historical note: Information retrieval and the future of an illusion. JASIS. 39(2): 92-98 (1988)	自動索引作成と情報検索とに関する個人的な展望を述べている。	6

	TE	Tenopir, C. Full text database performance. On-line Review. 9(2): 149-164 (1985)	オンラインの全文データベースで、①全文、②抄録のみ、③統制語、④書誌ユニオンを用いた検索の効率を比較している。	10
ユ ー ザ モ デ ル	BA79a	Bates, M.J. Information search tactics. JASIS. 30(4): 205-214 (1979)	利用者の探索行動を分析し、より進んだ探索をするために用いる探索戦略の要素的単位であるタクティクスを定義して、カテゴリ化をしている。	10
	BA79b	Bates, M.J. Idea tactics. JASIS. 30(5): 280-289 (1979)	心理学的側面から、利用者の探索行動を分析して、その要素的な単位を定義して、カテゴリ化をしている。	7
	BA86	Bates, M.J. Subject access in online catalogs: a design-model. JASIS. 37(6): 357-376 (1986)	OPACでの主題アクセスのための、不確定性・多様性・複雑性を持ったインターフェイスモデルを提示した。	9
	BD	Bellardo, T. An investigation of online searcher traits and their relationship to search outcome. JASIS. 36(4): 241-250 (1985)	サーチャの特性とそれが検索結果に及ぼす影響とを調査している。	7
	BE82	Belkin, N.J.; Brooks, H.M.; Oddy, R.N. ASK for information retrieval. Pt. 1. background and theory. J DOC. 38(2): 61-71 (1982)	知識の変則状態 (ASK) は情報ニーズの中に存在するが、それを表わす試みにもとづく IR システムの理論を述べている。	10
	BO86a	Borgman, C.L. The users mental model of an information retrieval system: an experiment on a prototype online catalog. INT J MMS. 24(1): 47-64 (1986)	オンライン目録の利用者のメンタルモデル。メンタルモデルを提示した利用者教育は特に複雑な検索課題に対して有効であることを示した。	9
	BO86b	Borgman, C.L. Why are online catalogs hard to use: lessons learned from information-retrieval studies. JASIS. 37(6): 387-400 (1986)	OPAC 利用者の検索過程や行動に関する研究のレビュー。利用者の認知的な側面と機械的な側面とにわけて論じている。	7
	FE80	Fenichel, C.H. The process of searching online bibliographic databases: a review of research. Library Research. 2(2): 107-127 (1980)	オンラインは検索システムにおける検索過程および検索戦略についての研究のレビュー。	7
	FE81	Fenichel, C.H. Online searching: measures that discriminate among users with different types of experiences. JASIS. 32(1): 23-32 (1981)	オンライン検索における利用者の行動を調べ、利用者の経験年数別に比較している。	6
	MT	Matthews, J.R. et al. Using online catalogs: a nationwide survey. Neal-Schuman Pub., N.Y., 1983.	OPACの利用に関して、米国で全国規模で行われた調査結果をまとめたもの。	8
デ ィ タ 構 造	SE	Swell, W. et al. Observations of end-user online searching behavior over 11 years. JASIS. 37(4): 234-245 (1986)	米国国立医学図書館のデータベースのエンドユーザ検索のログの解析とフォローアップインタビューによってエンドユーザの行動や認知過程を分析している。	7
	SR88	Saracevic, T.; Kantor, P. A study of information seeking and retrieving. 3. Searchers, searches, and overlap. JASIS. 39(3): 197-216 (1988)	情報探索と検索に含まれる要素の公式的・科学的特点づけを行なう研究の第三部。利用者が質問を発するコンテキストやサーチャの認知過程を分析している。	7
	BE87	Belkin, N.J. et al. Distributed expert-based information systems: an interdisciplinary approach. IPM. 23(5): 395-409 (1987)	エキスパートシステムを用いた情報検索システムの学際的研究。	6
	BU	Bush, V. As we may think. Atlantic Monthly. 176(1): 101-108 (1945)	連想関係によって全ての情報を関連づけ、蓄積し、柔軟に活用することを目指したシステムである Memex を提唱。	7
	CO	Conklin, J. Hypertext: an introduction. Computer. 20(9): 17-41 (1987)	ハイパーテキストに関する基本的な概念や機能を解説した入門的な論文。	15
	CR	Croft, W.B. et al. I3R: a new approach to the design of document-retrieval systems. JASIS. 38(6): 389-404 (1987)	エキスパートシステムを組み込んだ文献検索システム。ユーザーのニーズをより特定化することによって検索効率をあげることをめざしている。	10
	DA	Date, C.J. An introduction to database systems. 3rd ed. Addison-Wesley Pub., 1981	リレーショナル・データベース管理システムに関する教科書。	6
	FI	Fidel, R. Towards expert systems for the selection of search keys. JASIS. 37(1): 37-44 (1986)	エンドユーザーと検索システムとを仲介するエキスパートシステムの必要性を論じ、これは人間の仲介者の行動を分析することによって可能だとしている。	6
	FO	Fox, E.A. Development of the CODER system: a testbed for artificial-intelligence methods in information-retrieval. IPM. 23(4): 341-366 (1987)	様々な種類のテキストを扱う事ができる、人工知能を用いた情報検索システム CODER について論じている。	7
	WE	Weyer, S.A. The design of a dynamic book for information search. International Journal of Man-Machine Studies. 17(1): 87-107 (1982)	柔軟な情報提供を可能にする強力な情報伝達システムとしてのダイナミックブックを論じ、人が情報を探索する方法とダイナミックブックが探索過程とどのように関わらるかを調べている。初歩的なダイナミックブックも提示。	6
独 立	OD	Oddy, R.N. Information-retrieval through man-machine dialog. J DOC. 33(1): 1-14 (1977)	システムがユーザの状況のイメージを形成し、それに基づいて応答するというインタフェースを持ったシステム。	8
	SM	Small, H. Co-citation in scientific literature: new measure of relationship between 2 documents. JASIS. 24(4): 265-269 (1973)	Small が文献間の関係を表す新しい尺度として、共引用分析を提唱した最初の文献。	6
	SR75	Saracevic, T. Relevance: review of a framework for thinking on notion in information-science. JASIS. 26(6): 321-343 (1975)	レバンスを情報学の基本的概念として捉え、科学コミュニケーションの問題として扱っている。知識のコミュニケーションの枠組みで、レバンスに関する様々な見解を整理している理論的な研究。	6

* JICST データベースの抄録を加工, ** 文献 (7) の報告書 p. 65-71 より抜粋・加工, ARIST: Annual Review of Information Science and Technology, INT J MMS: International Journal of Man-Machine Studies, IPM: Information Processing and Management, JASIS: Journal of the American Society for Information Science, J DOC: Journal of Documentation

III. 情報検索領域の構造

A. 共引用マップ

第2図に示したごとく、情報検索分野の共引用マップでは、①検索理論、②プログラミング技法、③全文データベースおよびオンライン検索、④ユーザモデル、⑤データ構造の五つのグループが認められた。

これらの類似性を持った文献からなるグループは、共引用マップにおける各文献の位置にもとづいて、文献の内容を考慮して認定した。なお、文献の内容は、主として文献の抄録によって把握し、必要に応じて本文を参照した。

①「検索理論」は、確率型検索、ベクトル空間型検索、ファジィ理論を応用した検索など様々な検索理論に関する文献からなるグループである。また、この中に検索理論の包括的な教科書の小グループがある。

②「プログラミング技法」は、アルゴリズムやハッシングなど情報検索のコンピュータ処理の技術面に関わる文献からなるグループである。

③「全文データベースおよびオンライン検索」は、全文データベースの検索効率を扱った文献やオンライン検索システムに関する文献からなるグループである。これは、情報検索の実用化の流れの中に位置づけられるグループといえる。この中には Lancaster の語彙統制に関する

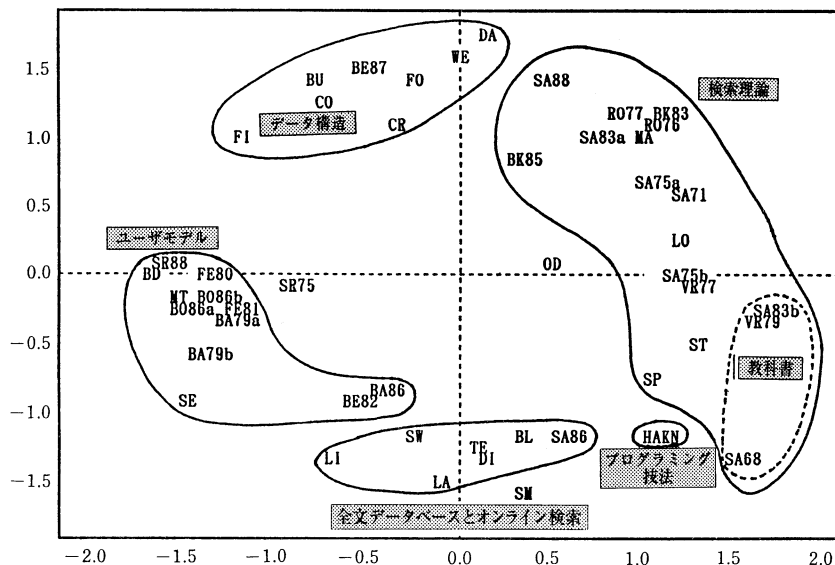
教科書 (LA, 1986) があるが、現在実用化されているオンライン検索システムにおいては、自然語検索か統制語彙検索かという観点が生じることから、このグループに入っていると考えられる。

④「ユーザモデル」は、オンライン検索やオンライン目録 (以降 OPAC) の利用者のふるまいやその認知構造、あるいは検索結果に影響を及ぼすような検索者の特性などを扱った文献からなっているグループである。

⑤「データ構造」は、Date (DA, 1986) のリレーショナルデータベース管理システム (以降 RDBMS)、ハイパーテキスト、知識ベースを用いた知的情報検索システムなどに関する文献からなっている。これらは、いずれも情報検索システムにおける種々のデータ構造を扱っていると考えられることから「データ構造」グループと名づけた。

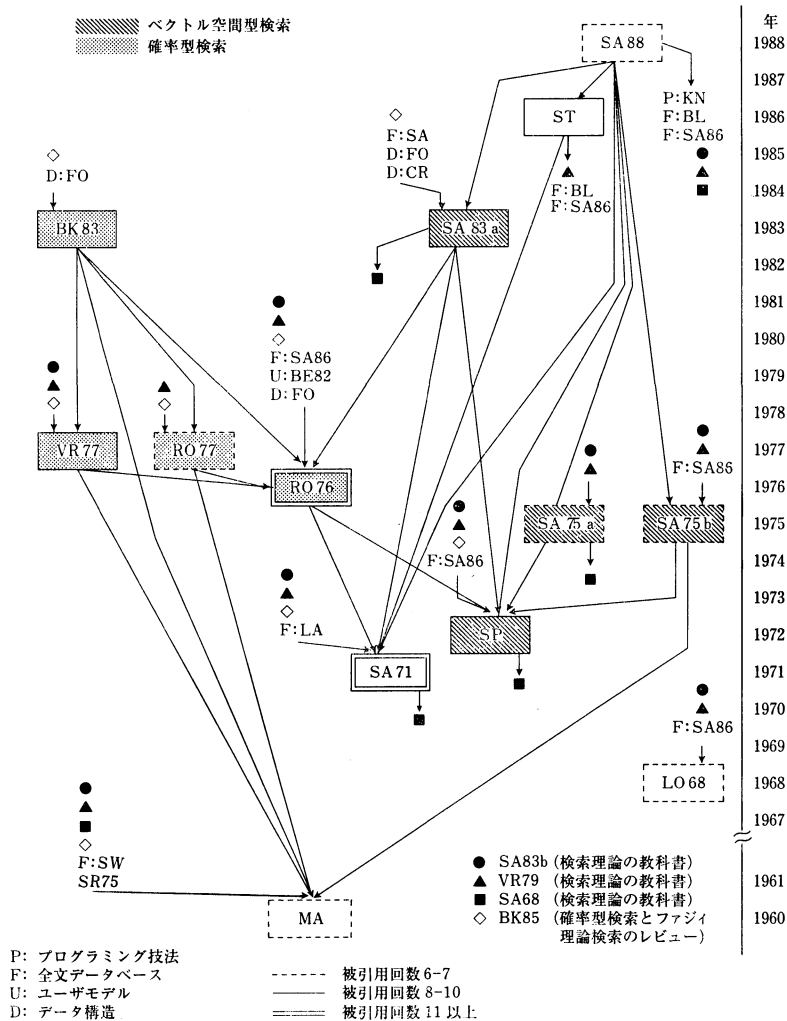
なお、以上に含まれない独立した文献として Oddy (OD, 1977), Saracevic (SR75, 1975), Small (SM, 1973) があつた。

以上、最近2年間の情報検索に関する文献から高頻度で引用された文献から作成した共引用マップを、内容を考慮して分析した結果、検索理論、プログラミング技法、全文データベースおよびオンライン検索、ユーザモデル、データ構造という五つのグループが認定できた。これらの五つが情報検索に関する主要な研究領域と考えられ



第2図 情報検索領域の共引用マップ (S-STRESS : 0.269)

文献を示す略号については、第4表を参照。



第3図 引用関係図 (検索理論)
 文献を示す略号については、第4表を参照。

る。

B. 文献間の直接引用関係

次に、研究の流れや発展過程を調べるために、各グループ内の文献間の直接的な引用関係を調べた。

(1) 検索理論

第3図は、検索理論グループの引用関係図である。この図では、下部から上部へ年代順に文献を配置し、このグループ中の個々の文献は四角で囲んで示した。矢印は、引用関係を示しており、ある文献からその文献が引用し

ている文献へと向かっている。また、このグループ以外の文献は、四角で囲まずに、発表年代に関わらず、引用関係のある文献の近くに記載してある。

さて、この検索理論グループの文献の引用関係図では Robertson (RO76, 1976) と Sparck-Jones (SP, 1972) に多くの矢印が集まっており、これらの文献はこのグループ内の文献から多く引用されている中心的な文献と考えられる。

Sparck-Jones (SP, 1972) は、語の出現頻度を用いて索引語への重み付けを行った研究で、ベクトル空間型検

索に影響を与えたとされている文献である。これは、斜線で示したベクトル空間型検索の文献から引用されている。また、Salton (SA83a, 1983) は、文献と検索質間に重み付けを行い、従来のブール演算を用いた検索にベクトル空間型の検索を導入した研究であるが、ファジィ理論にも言及しており、ベクトル空間型とファジィ理論検索の両者に関連している。

さらに、図中右上方の Stanfill (ST, 1986) と Salton (SA88, 1988) は、並列処理検索に関する文献である。これらは、Blair (BL, 1985) と Salton (SA86, 1986) の全文データベースに関する文献を引用しており、引用関係からみても、この検索理論のグループの他の文献とは別の新しいグループが形成されつつあることが分かる。

一方、Robertson (RO76, 1976) は、確率型検索を扱った初期の論文である。網かけで示した確率型検索を扱った文献から引用されている。Robertson (RO76, 1976) から出発しているこれらの確率型検索の文献の全ては、確率型検索とファジィ理論検索に関するレビュー論文である Bookstein (BK85, 1985) に引用されている。なお、この Bookstein (BK85, 1985) もこの検索理論グループ内の文献であるが、引用している文献が多いので便宜的に図中では ◇ という記号で示してある。さらに、これらの確率型検索に関する文献のほとんどは Maron (MA, 1960) を引用している。この Maron (MA, 1960) は、適合性を確率で表わすことを提唱した文献であるが、このように多くの確率型検索の文献から引用されていることから、確率型検索の基盤の一角をなす文献と判断される。

以上のように、検索理論グループでは、ベクトル空間型検索と確率型検索との大きな二つの流れがあることが分かる。中でも、Sparck-Jones (SP, 1972) と Robertson (RO76, 1976) は二つの流れの出発点になっていると同時に、最近2年間の情報検索に関する文献からの被引用回数も多く、中心的な文献であると考えられる。さらに、ファジィ理論を応用した検索や並列処理に関する文献も見られたが、ここでは流れとして捉えるのに十分な文献数は得られなかった。

さらに、この検索理論グループには、検索理論の包括的な三種の教科書 (Salton (SA83b, 1983b, 図中 ●), Van Rijsbergen (VR79, 1979, 図中 ▲), Salton (SA68, 1968, 図中 ■) が含まれている。これらは、いずれも、その教科書が出版される以前に出版された、この検

索理論グループ内のほとんど全ての文献を引用していた。また、グループ外からは、個々の研究文献よりもこれらの教科書が多く引用されている。

(2) 全文データベースおよびオンライン検索

次に、全文データベースやオンライン検索に関するグループでは、グループ内での直接的な引用関係はほとんどなく、Blair (BL, 1985) が Salton (SA86, 1986) と Swanson (SW, 1988) から引用されているのみであった。グループ外では検索理論グループの文献を多く引用していた。そのため、この全文データベースやオンライン検索に関するグループの文献は、理論的には検索理論グループの文献に依存していると考えられる。特に Salton (SA86, 1986) は IBM STAIRS による全文データベースの検索手法を中心的に扱っているが、自動検索システムにも言及していることから、検索理論グループの文献を多く引用していた。

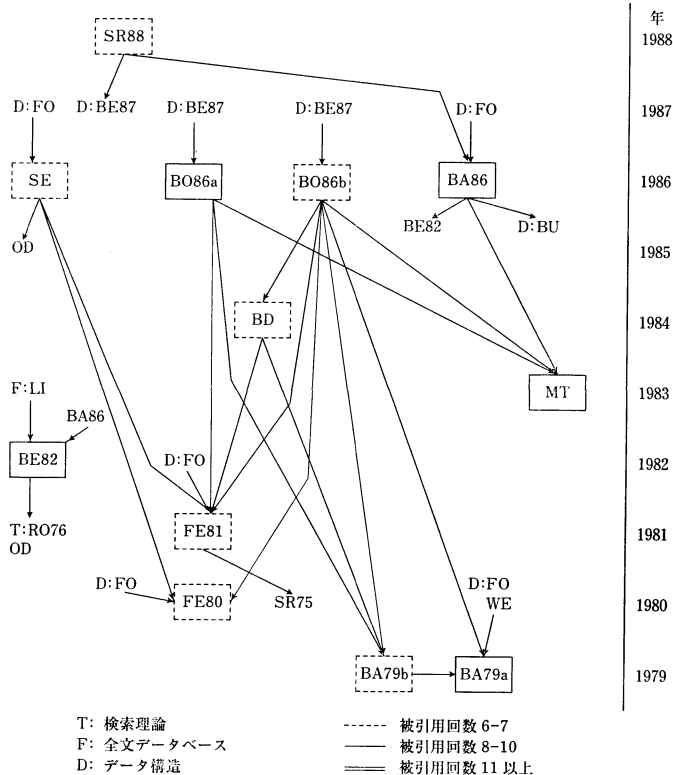
その中で、Blair (BL, 1985) は、このグループ内のみならず、検索理論やデータ構造のグループからも引用されており、最近2年間の情報検索に関する文献からの被引用回数も多いことから、全文データベースの研究に関して他に影響を与えた基礎的な文献であると考えられる。

(3) ユーザモデル

第4図のユーザモデルのグループの引用関係では、Fenichel (FE81, 1981) と Bates (BA79b, 1979) に多くの矢印が集まっており、このグループ内の文献に影響を与えた文献であると考えられる。Fenichel (FE81, 1981) は、オンライン検索利用者の行動を経験年数の差によって比較した研究であり、Bates (BA79b) は、情報探索行動を心理学的な側面から分析・整理することを試みた研究である。これらを引用している文献は、オンライン検索や OPAC の検索過程やそれに影響を与える因子に関する研究であった。

また、Matthews (MT, 1983) は、OPAC の利用状況に関する調査であり、ユーザモデルとは直接には関係しないが、OPAC の検索者のユーザモデルを扱っている Borgman (BO86a, 1986a) や Borgman (BO86b, 1986b), Bates (BA86, 1986) から引用されている。

Borgman (BO86b, 1986) は、OPAC 利用者の行動や検索戦略に関するレビューであり、このグループ内のほとんどすべての文献を引用している。一方、Belkin (BE82, 1982) は、この Borgman (BO86b, 1986) から引用されず、Bates (BA86, 1986) やグループ外の文献から



第4図 引用関係図 (ユーザモデル)
文献を示す略号については、第4表を参照。

引用されているが、グループ内の他の文献との引用関係が少なく、独得な位置づけにあると考えられる。

以上のごとく、このグループでは、矢印が集まっている文献は見られたが、これらの文献を引用している文献は互いに重なりあい、顕著なサブグループは見いだせなかった。

全体として、グループ外の文献を引用することはほとんどなく、比較的グループ内でまともになっている。データ構造グループ中のエキスパートシステムの文献から、多く引用されている。これは、エキスパートシステムではシステム設計においてユーザモデルを考慮する必要があるためと考えられる。

(4) データ構造

この「データ構造」グループでは、Bush (BU, 1945)を除いて比較的新しい文献が多いため、最近2年間の情報検索に関する文献からの共引用関係にもとづいて作成した共引用マップでは、サブグループを認定することが

できなかった。また、文献の出版年がほぼ同じであったため、文献間の直接的な引用関係も多くなく、引用が集中した文献も見られなかった。しかし、直接引用関係の分析からそれぞれ異なる背景を持った「エキスパートシステム」と「ハイパーテキスト」という流れが生じていることが示唆され、①ハイパーテキスト、②エキスパートシステム、③RDBMSの三つのサブグループが認められた。

- ① 「ハイパーテキスト」は、Bush の MEMEX (BU, 1945) とそれを引用している文献 Weyer (WE, 1982) と Conklin (CO, 1987) からなるグループである。
- ② 「エキスパートシステム」は、ユーザモデルの文献を多く引用している文献やそれと結びついている文献からなるグループで、具体的には、Belkin (BE87, 1987), Croft (CR, 1987), Fidel (FI, 1986), Fox (FO, 1987) からなっている。

情報検索分野の構造：引用調査による下位領域と発展過程の分析

第5表 拡張検索理論グループの文献

	コード	書誌事項	内 容
統計的手法に基づいた重み付け検索モデル	BK85	Bookstein, A. Probability and fuzzy-set applications to information retrieval. ARIST. 20: 117-178 (1985)	確率型検索とファジ理論を応用した検索についての理論, 手法, 研究結果および最新動向などをレビューしている。*
	SA75b	Salton, G. et al. A theory of term importance in automatic text analysis. JASIS. 26(1): 33-50 (1975)	識別値を用いた自動索引法を論じている。語と出現頻度に基づくベクトル空間型検索モデルについての初期の論文。*
	SA71	Salton, G. ed. The SMART Retrieval System. Englewood Cliffs, N. J., Prentice-Hall, 1971, 556 p.	自動検索システムである SMART システムを様々な側面から論じている。特に検索の評価, 言語の分析, クラスター化による検索, レlevance・フィードバックなどが重点的に扱われている。*
	SA75a	Salton, G. et al. A vector space model for automatic indexing. CACM. 18(11): 613-620 (1975)	ベクトル空間型検索モデルについての主要なアイデアを論じた初期の論文。*
	SA83b	Salton, G. et al. Introduction to Modern Information Retrieval. New York, McGraw-Hill, 1983, 448 p.	情報検索の理論やシステムに関する包括的な教科書。検索理論の中心はベクトル空間モデルや SMART システムであるが, 確率型検索などの手法も解説している。また, 後半ではファイル構造などのハードウェアに関するトピックが扱われている。*
	SA86	Salton, G. Another look at automatic text-retrieval system. CACM. 29(7): 648-656 (1986)	IBM STAIRS による全文データベースの検索手法について。さらに自動検索システムについても論じている。*
	SP	Spark-Jones, K. A statistical interpretation of term specificity and its application in retrieval. J DOC. 28(1): 11-21 (1972)	各語が出現する文献の数によって語の重み付けを行う手法を提唱している。Salton のベクトル空間型検索に大きな影響を与えた。*
	VR79	Van Rijsbergen, C. J. Information Retrieval. 2nd. London, Butterworth, 1979. 152 p.	情報検索理論の一般的な教科書。特に自動索引, 自動分類, 検索の評価に重点がおかれている。*
	YU76	Yu, C. T.; Salton, G. Precision weighting: an effective automatic indexing method. JACM. 23(1): 76-88 (1976)	適合文献と非適合文献の情報を用いて利用者の検索質問中の語を重み付ける手法を開発し, その効果を実験している。一種のベクトル空間型検索モデルである。*
ファジ理論検索	BK80	Bookstein, A. Fuzzy requests: an approach to weighted Boolean searches. JASIS. 31(4): 240-247 (1980)	従来のブール検索のシステム特性を検討した上で, ブール検索を拡張するために, ファジ理論を導入している。*
	BUE	Buell, D. A.; Kraft, D. H. A model for a weighted retrieval system. JASIS. 32(3): 211-16 (1981)	質問に対し, 閾値を設定し, 語の重み付けを行ったシステムにおいてファジ理論を導入して検索を行っている。*
	SA83a	Salton, G. et al. Extended boolean information retrieval. CACM. 26(12): 1022-1036 (1983)	文献と検索質問に重みづけを行ない, 従来のブール検索にベクトル空間型の検索を導入して拡張する試み。ファジ理論に言及している。*
	WA	Waller, W. G.; Kraft, D. H. A mathematical model of a weighted Boolean retrieval system. IPM. 15(5): 235-45 (1979)	語の重み付けを行った検索システムに対し, ファジ理論を応用した論理演算を導入し, これを拡張したモデルを提示している。*
確率型検索とレlevance フィードバック手法	BK83	Bookstein, A. Information retrieval: a sequential learning process. JASIS. 34(5): 331-242 (1983).	決定理論を用いた確率型検索の手法に, レlevance・フィードバックを組み込むためのモデルを提示している。*
	BK83a	Bookstein, A. Outline of a general probabilistic retrieval model. J DOC. 39(2): 63-72 (1983)	より一般的な関連性尺度を導入したモデルによって, 理論の拡張を行った。このモデルは, 既存の理論を包含しており, これまでのモデルにおける回答のランキングの原理を, より満足された形で説明することができる。*
	CR79	Croft, W. B.; Harper, D. J. Using probabilistic models of document retrieval without relevance information, J DOC. 35(4): 285-95 (1979)	適合文献についての情報を用いない確率的な検索手法を, 確率モデルに基づいて提案している。*
	CR81	Croft, W. B. Document representation in probabilistic models of information retrieval. JASIS. 32(6): 451-457 (1981)	有効な検索戦略のもとになる確率モデルを検討。このモデルの限界は, 文献は2つの索引語で示されるとする仮説にある。索引語の重要性としての重みづけを伴う文献表式法について, 検索モデルは独立した仮説または依存した仮説を用いて, 拡大しうるかを論じた。変形モデルに基づく検索戦略を他の研究と比較した。*
	FS	Fishburn, P. C. Utility theory for decision making. New York, Wiley, 1970.	意思決定のための効用理論
	LS	Losee, R. M. Parameter estimation for probabilistic document-retrieval models. JASIS. 39(1): 8-16 (1988)	適合文献と不適合文献の情報をを用いてフィードバックを行ない, 逐次学習モデルを用いた確率型検索を行っている。*
	MA	Maron, M. E.; Kuhns, J. L. On relevance, probabilistic indexing and information retrieval. JACM. 7: 216-244 (1950)	レlevanceを確率を用いて表現することを提唱している初期の文献。*
	RO76	Robertson, S. E.; Spark-Jones, K. Relevance weighting of search terms. JASIS. 27(3): 129-146 (1976)	適合文献の情報に基づき, 重み付けを行う手法を開発し, その実験結果を提示している。確率型検索についての初期の論文。*

確率型検索とレバンスフィードバック手法	RO77	Robertson, S.E. Probability ranking principle in IR. J DOC. 33(4): 294-304 (1977)	確率型検索に対する Cooper が提起した疑義に答える形で書かれた確率型検索に関するレビュー。
	RO81	Robertson, S.E.; Van Rijsbergen, C.J.; Porter, M.F. "4, Probabilistic models of indexing and searching", In: Oddy, R.N.; Robertson, S.E.; Van Rijsbergen, C.J.; Williams, P.W. (eds.) Information retrieval research. London, Butterworths, 1981, p. 35-56.	確率型検索に対する理論と概念を説明し、語の出現頻度に基づく自動索引についての初期理論と実験についてまとめている。
	RO82	Robertson, S.E.; Maron, M.E.; Cooper, W.S. Probability of relevance: a unification of two competing models for document retrieval. Information Technology: Research and Development. 1(1): 1-21 (1982)	Maron (1960) と Robertson (1976) が提案している 2 つの基本的な確率型検索の手法を分析し、これを統合した手法を提案している。
	RO86	Robertson, S.E. On relevance weight estimation and query expansion. J DOC. 42(3): 182-188 (1986)	フィードバック検索で、使用する重みづけ関数の改良を議論。Robertson, Sparck-Jones 関連重みづけの式に、Bayes の理論に基づく補正項を加えた。*
	SN	Smeaton, A.F.; Van Rijsbergen, C.J. The retrieval effects of query expansion on a feedback document retrieval system. Comput J. 26(3): 239-246 (1983)	フィードバック文獻検索システムにおける問合せ拡張の検索効果: 種々のタイプの質問拡張の方策を示し、その検索効果を分析した。*
	YU83	Yu, C.T.; Buckley, C.; Lam, K.; Salton, G. A generalized term dependence model in information retrieval. Information Technology: Research and Development. 2(4): 129-54 (1983)	語の独立性をもとにして、ファイルのクラスター化を行うためのアルゴリズムを提示している。
並列処理検索	CR88	Croft, W.B.; Savino, P. Implementing ranking strategies using text signatures. ACM Trans Off Inf Syst. 6(1): 42-62 (1988)	シグネチャファイルを用いた検索に対し、確率手法を用いた文獻への順位付けを行うことを提案。またこの際に、用いるシステムとして並列処理システムについて論じている。
	FA	Faloutsos, C.; Christodoulakis, S. Description and performance analysis of signature file methods for office filing. ACM Trans Off Inf Syst. 5(3): 237-257 (1987)	オフィス環境下における効果的な全文データ文書処理の手法として、シグネチャファイルを用いてテキストをコード化記述を行い、検索実験により評価を行っている。並列処理システムについても論じている。
	SA88	Salton, G.; Buckley, C. Parallel text search methods. CACM. 31(2): 202-215 (1988)	並列処理によるテキスト検索。ブールテキスト検索、ベクトル型、SMART などをシリアルに用いた方法と比べて大きな効果の改善はみられなかった。
	ST	Stanfill, C.; Kahle, B. Parallel free-text search on the connection machine system. CACM. 29(12): 1229-1239 (1986)	大規模データベース、特に全文データベースにおいて、並列処理システムにおいて検索を行う手法を提示している。
	SO	Stone, H.S. Parallel querying of large databases: A case study. Computer. 20(10): 11-21 (1987)	大規模データベースの並列問合せケーススタディ: パラレリズム自身は高速度を保障せず、アルゴリズムに依存する事をケース・スタディで示す。有効なアルゴリズムの下で走る 1 プロセッサと、パラレルだが非効率的に走る 64 K のプロセッサ群とを比較する。*
	WZ	Waltz, D.L. Applications of the connection machine. Computer. 20(1): 85-97 (1987).	コネクション・マシンは多量並列システムであり、容易にプログラムでき、種々の応用に適用可能である上に、さらに多くのプロセッサを付加することによって大きな問題に規模を合わせることができる。ドキュメント回復、メモリの推論、自然言語処理などの応用について述べる。*
用語の処理	LO	Lovins, B.J. Development of a stemming algorithm. Mechanical Translation and Computational Linguistics. 11: 22-31 (1968)	語幹処理のアルゴリズムを開発して、提示している。
	VR77	Van Rijsbergen, C.J. A theoretical basis for the use of co-occurrence data in information retrieval. J DOC. 33(2): 106-119 (1977)	語の独立性を仮定しない統計的な重みづけ検索手法を論じている。*
	SA76	Salton, G.; Wong, A. On the role of words and phrases in automatic text analysis. Computers and the Humanities. 10(2): 69-87 (1976)	索引作業における人手による索引と、自動索引を比較検討し、索引語の選定について論じている。

* JICST データベースの抄録を加工

** 文献 (7) の報告書 p.65-71 より抜粋・加工

ARIST: Annual Review of Information Science and Technology

CACM: Communications of the ACM

INT J MMS: International Journal of Man-Machine Studies

IPM: Information Processing and Management

JACM: Journal of the Association for Computing Machinery

JASIS: Journal of the American Society for Information Science

J DOC: Journal of Documentation

情報検索分野の構造：引用調査による下位領域と発展過程の分析

第6表 拡張ユーザモデルグループの文献

	コード	書誌事項	内 容
探 索 行 動 の 分 析	BA79a	Bates, M. J. Information search tactics. JASIS. 30(4): 205-214 (1979)	利用者の探索行動を分析し、より進んだ探索をするために用いる探索戦略の要素の単位であるタクティクスを定義して、カテゴリ化をしている。
	BA79b	Bates, M. J. Idea tactics. JASIS. 30(5): 280-289 (1979)	心理学的側面から、利用者の探索行動を分析して、その要素的な単位を定義して、カテゴリ化をしている。
	BA86	Bates, M. J. Subject access in online catalogs: a design-model. JASIS. 37(6): 357-376 (1986)	OPACでの主題アクセスのための、不確定性・多様性・複雑性を持ったインターフェイスモデルを提示している。
	FI85	Fidel, R. Moves in online searching. Online Review. 9(1): 61-74 (1985)	検索集合が不適切な場合の質問式改良法を論ずる。約90回の検索結果の分析と考察に基づき、検索語の意味に立ち入った改良法と立ち入らない改良法に分けて論じ、18種の改良法を一覧表にし、改良法と検索戦略との関係を論ずる。*
	MK	Markey, K. Subject Searching in Library Catalogs: Before and After the Introduction of Online Catalogs. OCLC, Inc., 1984.	図書館目録の主題からの探索について、OPAC導入前後を比較しながら論じている。
	MT	Matthews, J. R. et al. Using online catalogs: a nationwide survey. Neal-Schuman Pub., N.Y., 1983.	OPACの利用に関して、米国で全国規模で行われた調査結果をまとめたもの。
	NO	Noerr, P. L.; Bivins Noerr, K. T. Browse and navigate: an advance in database access methods. IPM. 21(3): 205-213 (1985)	新しいマルチリンク DBMS (TINMAN) を開発。ブラウジングとリンクをたどるナビゲーションによって DB にアクセスする。
	WI	Williams, M. D. What makes RABBIT run? INT J MMS. 21(4):333-352 (1984)	新しいインターフェイスの設計が完了したときに生じる問題点を挙げ、DB 検索インターフェイスの RABBIT を例に、1) 現在の行動の記述、2) 再定式化による検索の特性の定義、3) 新しい特徴を明確にする試みの特徴づけを行った。*
検 索 へ の 影 響 因 子	BA81	Bates, M. J. Search Techniques. ARIST. 16: 139-169 (1981)	情報探索の技法を、心理学的な側面から論じたレビュー。
	BD	Bellardo, T. An investigation of online searcher traits and their relationship to search outcome. JASIS. 36(4): 241-250 (1985)	サーチャの特性とそれが検索結果に及ぼす影響とを調査している。
	BO86b	Borgman, C. L. Why are online catalogs hard to use: lessons learned from information-retrieval studies. JASIS. 37(6): 387-400 (1986)	OPAC 利用者の検索過程や行動に関する研究のレビュー。利用者の認知的な側面と機械的な側面とにおいて論じている。
	EL	Elkerton, J.; Williges, R. C. Information retrieval strategies in a file-search environment. Human Factors. 26(2): 171-184 (1984)	探索方法、ユーザの経験の有無、ファイル型、ファイル長、目標型、ウィンド・サイズを変数として、利用者が用いる探索戦略を調べている。*
	FE80	Fenichel, C. H. The process of searching online bibliographic databases: a review of research. Library Research. 2(2): 107-127 (1980)	オンライン検索システムにおける検索過程および検索戦略についての研究のレビュー。
	FE81	Fenichel, C. H. Online searching: measures that discriminate among users with different types of experiences. JASIS. 32(1): 23-32 (1981)	オンライン検索における利用者の行動を調べ、利用者の経験年数別に比較している。
	FI83	Fidel, R.; Soergel, D. Factors affecting online bibliographic retrieval: a conceptual framework for research. JASIS. 34(3): 163-180 (1983)	オンライン文献検索に影響を与える因子として、状況、利用者、質問要求、データベース、検索システム、検索者、検索プロセス、検索結果について考察した。*
メンタルモデル	BO86a	Borgman, C. L. The users mental model of an information retrieval system: an experiment on a prototype online catalog. INT J MMS. 24(1): 47-64 (1986)	オンライン目録の利用者のメンタルモデル。メンタルモデルを提示した利用者教育は特に複雑な検索課題に対して有効であることを示した。
	FI84	Fidel, R. Online searching styles: a case study based model of searching behavior. JASIS. 35(4): 211-221 (1984)	通常の検索を体系的に観察し、事例研究として10の検索を分析して、操作主義者と概念論者という2種メンタルモデルに基づいて、個々の検索を説明づけている。
	GE	Gentner, D.; Stevens, A. L. ed. Mental Models. Lawrence Erlbaum Assoc. Inc., 1983.	心理学分野のメンタルモデルに関する包括的な単行書。
情報探索の包括的モデル	BE82	Belkin, N. J.; Brooks, H. M.; Oddy, R. N. ASK for information retrieval. Pt. 1. background and theory. J DOC. 38(2): 61-71 (1982)	知識の変則状態 (ASK) は情報ニーズの中に存在するが、それを表わす試みにもとづくシステムの理論を述べている。*
	BE82a	Belkin, N. J.; Brooks, H. M.; Oddy, R. N. ASK for information retrieval. Pt. 2. Results of a design study. J DOC. 38(3): 145-164 (1982)	ASK を表す試みに基づいて、利用者を対象とし、IR についての見解、抄録の評定、テキストに表われる語句の連想関係などをインタビューと質問紙によって調べた結果。*
	BO85	Borgman, C. L. Individual differences in the use of technology: work in progress. Proceedings of the ASIS. 22: 243-249 (1985)	情報技術の利用の要因と考えられる個人特性の測定を、学術的な測定と比較した。問題提起をし、関連研究、研究方法等について述べている。結果は後で報告する予定。*
	OD	Oddy, R. N. Information-retrieval through man-machine dialog. J DOC. 33(1): 1-14 (1977)	システムがユーザの状況のイメージを形成し、それに基づいて応答するというインタフェースを持ったシステム。

情報探索の包括的モデル	SR88a	Saracevic, T.; Kantor, P.; Chamis, A. Y.; Trivison, D. J. A study of information seeking and retrieving. I. Background and methodology. JASIS. 39(3): 161-176 (1988)	情報探索と検索に含まれる要素の公式的・科学的特徴づけを行なう目的で、情報検索における利用者、質問、検索者、検索、検索項目についての一連の研究の第Ⅰ部。背景、モデル、尺度、方法、手順、統計的分析手法を論じる。*
	SR88b	Saracevic, T.; Kantor, P. A study of information seeking and retrieving. II. Users, questions, and effectiveness. JASIS. 39(3): 177-196 (1988)	上記の第Ⅱ部。情報検索における利用者、質問、有効な尺度として適合率と再現率、有用性、検索された文献、出力サイズについて検討した。40の検索質問を9人の検索者が、延360回の検索を行なった結果を比較分析した。
	SR88	Saracevic, T.; Kantor, P. A study of information seeking and retrieving. III. Searchers, searches, and overlap. JASIS. 39(3): 197-216 (1988)	上記の第Ⅲ部。利用者が質問を発するコンテキストやサーチの認知過程を分析している。

* JICST データベースの抄録を加工

ARIST: Annual Review of Information Science and Technology

CACM: Communications of the ACM

INT J MMS: International Journal of Man-Machine Studies

IPM: Information Processing and Management

JASIS: Journal of the American Society for Information Science

J DOC: Journal of Documentation

- ③ 上記以外のものとして、RDBMS の教科書である Date (DA, 1986) があり、これを「RDBMS」サブグループとした。

また、これらの三つのサブグループ間の関係を見るために、このデータ構造グループの文献と最近2年間の情報検索に関する文献との引用関係を調べた。その結果、「RDBMS」サブグループの文献を引用している最近2年間の文献は、RDBMS あるいはエキスパートシステムに関するものであり、「エキスパートシステム」サブグループを引用している最近2年間の文献は、エキスパートシステムあるいはハイパーテキストに関するものであり、「ハイパーテキスト」サブグループを引用している最近2年間の文献は、ハイパーテキストに関するものであるという関係が見られた。

さらに、この三つのサブグループ間の共引用関係を見ると、「RDBMS」と「エキスパートシステム」サブグループを共引用している最近2年間の文献は、エキスパートシステムに関するものであり、「エキスパートシステム」と「ハイパーテキスト」とを共引用している最近2年間の文献は、ハイパーテキストに関するものであるという関係が見られた。このことから、情報検索領域でのデータ構造に関する関心は、RDBMS から、エキスパートシステム、さらにハイパーテキストへと広がっていることも推察できる。

以上のような情報検索領域のグループ化を、1990年に出版された Ellis の情報検索の教科書³⁾の内容と比較した。まず、Ellis の教科書では1章の緒言に続き、2章は「統計的検索と確率型検索」であるが、これは今回の結果では検索理論グループに相当する。また、3章の「認知的なユーザのモデル化」が今回のユーザモデルグループ、4章は、ユーザインタフェースにエキスパートシステムを応用したものを扱っているが、これは、今回のデータ構造グループの中のエキスパートシステムにあたり最後の5章の「連想、関係、ハイパーテキスト」は、今回のデータ構造グループの中のハイパーテキストに相当する内容であると考えられる。このように、今回の分析結果は、この Ellis の情報検索の教科書の章立てとほぼ一致している。つまり、今回の共引用分析による情報検索分野の文献のグループ化と最近の教科書に示された情報検索の枠組みとがほぼ一致したといえる。

IV. 検索理論とユーザーモデル領域の分析

次に、この五つのグループの中で主要な位置を占める

検索理論とユーザモデルのグループを取り上げ、上記の情報検索領域全体の分析と同様の手法でその発展過程や研究の流れをさらに詳細に分析した。

まず、最近2年間の情報検索に関する文献から引用されている11,033件の文献からなる被引用文献群の中から、検索理論グループまたはユーザモデルグループに属する文献と共引用強度が3以上の文献を全て抽出し、これをそれぞれ、「拡張検索理論グループ」(第5表参照)、「拡張ユーザモデルグループ」(第6表参照)とした。この拡張検索理論グループと拡張ユーザモデルグループに含まれる文献を対象として、II B. に示した方法と同様に、それぞれの共引用マップを作成した。さらに、これらの文献間の直接的な引用関係を調べた。

A. 検索理論

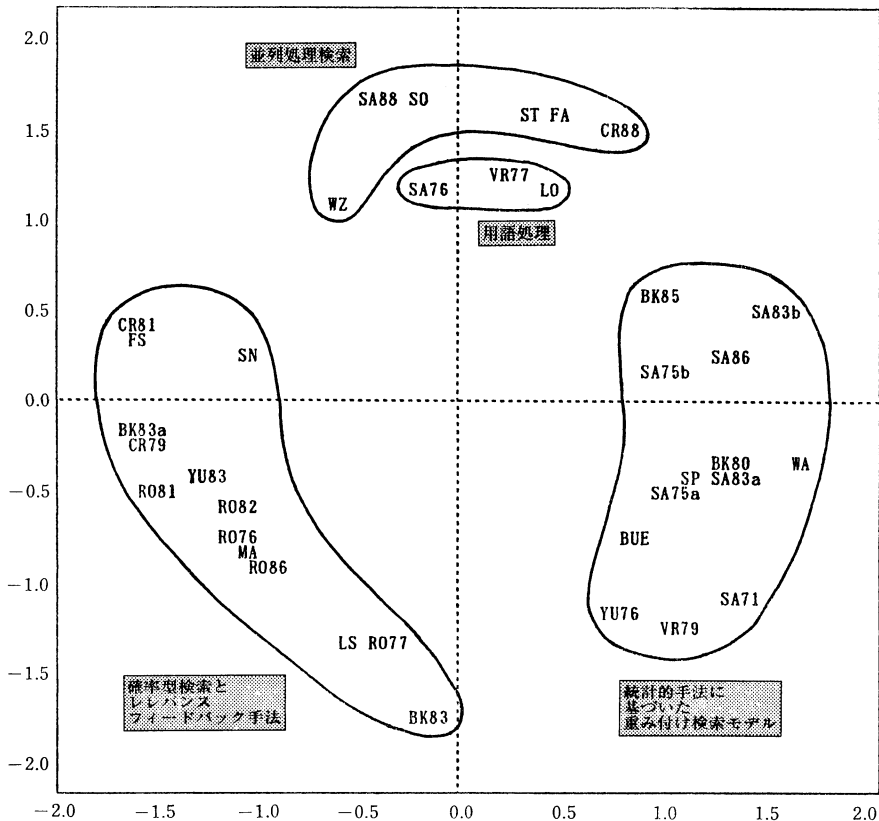
拡張検索理論グループの共引用マップを第5図に示す。その結果、①統計的手法に基づいた語への重み付けを行った検索モデル、②確率型検索およびこれに関連するレバンスフィードバック手法、③並列処理検索、④用語の処理、の四つのグループが認められた。

以下に、各グループ毎に、内容と文献間の直接の引用関係から検討した結果を述べる。

(1) 統計的手法に基づいた重み付け検索モデル

これは、検索語への重み付けの手法とそれを用いたベクトル空間型検索、フェジィ理論検索および拡張論理型検索を中心とした文献からなるグループである。なお、(2)の確率型検索も語への重み付けを行うが、この(1)のグループでは索引作業の段階で重み付けを行なうのに対し、確率型検索では基本的には検索式を修正するフィードバックの際に重み付けを行うという点で異なっている。

共引用マップ上で、この重み付け検索モデルグループのほぼ中央に文献が集まっている部分が認められる。この中で Spark-Jones (SP, 1972) は語の出現頻度による重み付けに関する初期の論文であり、Salton (SA75a, 1975) はベクトル空間型検索の主要なアイデアを論じた初期の論文である。いずれもこの重み付け検索モデル研究の初期に公刊された文献であり、このグループ内のほとんどの文献と関連するものである。一方、同じく中央部に集まっている Bookstein (BK80, 1980), Waller (WA, 1979), Salton (SA83a, 1983), Buell (BUE, 1981) は、互いに強い共引用強度で結び付いている一群であり、いずれもフェジィ集合理論を導入して論理型検索を



第5図 拡張検索理論グループの共引用マップ (S-STRESS : 0.294)

文献を示す略号については、第5表を参照。

拡張した検索手法に関する文献からなっている。すなわち、ファジィ理論検索、もしくは拡張論理型検索の文献からなる一群である。

Salton (SA83b, 1983) と Van Rijsbergen (VR79, 1979) は、検索理論全般を扱う教科書であり、この重み付け検索モデルのみでなく、拡張検索理論グループ全体の文献の多くと共引用関係がある。

また、直接引用関係を第6A図に示した。この重み付け検索モデルグループに含まれる文献のほとんどは、Spark-Jones (SP, 1972) を引用している。これは、語の重み付け手法に関する文献であり、このグループで最も基礎的な文献であると考えられる。

また、ファジィ理論検索に関する文献は、直接的な引用関係においても、結びつきが明確になっている。

確率型検索とファジィ理論検索に関するレビュー論文である Bookstein (BK 85, 1985) は、(2) の確率型検索

グループの文献も多く引用している。

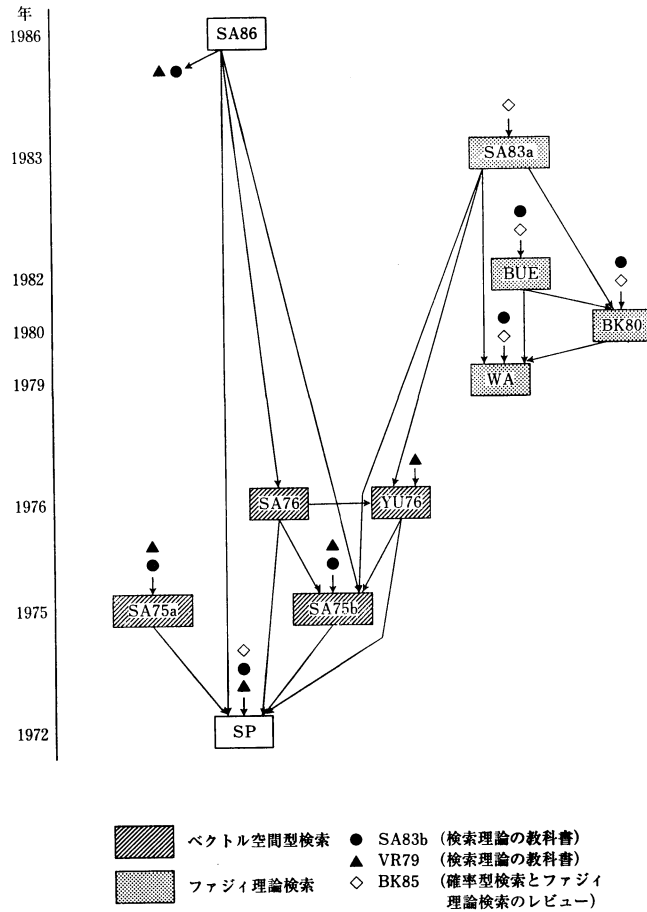
(2) 確率型検索とレバンスフィードバック手法

このグループは、確率型モデルを用いた検索実験を行った文献が中心となっている。また、確率型検索において用いられるフィードバック手法についての論文も含まれている。Robertson (RO 76, 1976) は、マップ上でこのグループのほぼ中央に位置しているが、これは確率型検索の概念を提案している初期の論文で、このグループのほぼ全ての文献と共引用関係があり、このグループの中心的な文献であると考えられる。

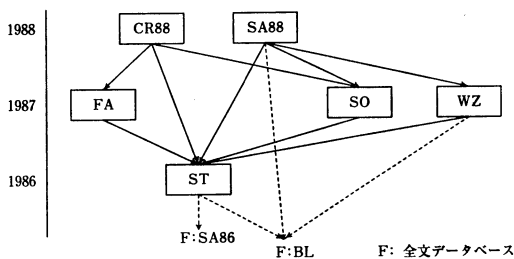
また、直接の引用関係においても、この Robertson (RO76, 1976) は、このグループ内のほとんど全ての文献から引用されていた。

以上の結果から同じ適合度順出力を行う検索モデルであっても、ベクトル空間型、ファジィ理論検索（もしくは拡張論理型）、確率型は、それぞれ中心となる文献が異

情報検索分野の構造：引用調査による下位領域と発展過程の分析



第 6A 図 引用関係図 (統計的手法に基づいた重み付け検索モデル)
文献を示す略号については、第 5 表を参照。



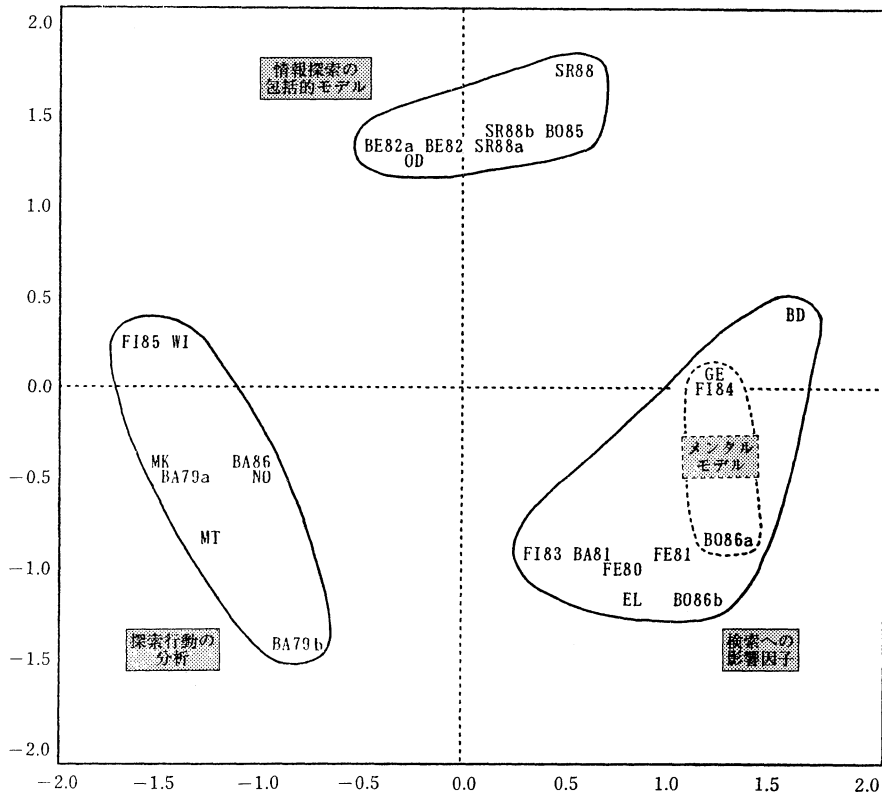
第 6B 図 引用関係図 (並列処理検索)
文献を示す略号については、第 5 表を参照。

っており、(1) の重み付け検索モデルグループと (2) の確率型検索のグループにおいては別個の研究動向があることが明確になった。

(3) 並列処理検索

これは、並列処理アルゴリズムを用いた高速テキスト検索に関する文献からなるグループである。1986年以降に公開された比較的新しい論文のみからなっている。また、論文掲載誌は *Communications of the ACM* や *Computer* などの情報処理分野の雑誌であり、情報処理領域での並列処理アルゴリズム研究との関連において出現したものであると考えられる。

直接の引用関係を第 6B 図に示した。Stanfill (ST, 1986) は、このグループ内の全ての文献から引用され、このグループ内の文献に影響を与えた基礎的な文献であると考えられる。この Stanfill (ST, 1986) は、Blair (BL, 1985) や Salton (SA86, 1986) の全文データベー



第7図 拡張ユーザモデルグループの共引用マップ (S-STRESS : 0.309)

文献を示す略号については、第6表を参照。

スの研究を引用している。一方, Salton (SA88, 1988) は従来の検索手法との比較を行っている。Croft (CR88, 1988) は、並列処理システムにおいて文献に確率的な順位づけを行った研究である。このように、並列処理検索の研究は、次第に従来の情報検索の手法と関連を持ちながら発展していると考えられる。

(4) 用語の処理

このグループは、自動索引による索引語の抽出法や語幹処理のアルゴリズムなど、自動化した検索システムで使用する用語の処理を扱った文献からなるグループである。このような語句の処理は、いずれの検索手法を実現する際にも直面する問題であるが、検索手法の一つとして位置づけられるものではない。

共引用マップ上では、並列処理検索のグループと近接しているが、これらは直接引用関係はなく、内容面からも別のグループと見なせるものである。

B. ユーザモデル

第7図に示すごとく、拡張ユーザモデルグループの共引用マップでは、①探索行動の分析、②検索への影響因子、③情報探索の包括的モデルの三つのグループが認められた。

以下に、各グループ毎に、内容と文献間の直接の引用関係から検討した結果を述べる。

(1) 探索行動の分析

これは、情報探索行動を分析し、それを構成する要素的な単位を整理して定義し、それらを用いて探索行動のモデル化を試みた研究と、利用者の探索行動を考慮したユーザインタフェースの研究からなるグループである。

Bates (BA79a, 1979) や Fidel (FI85, 1985) は、探索式の改善や検索結果の検査など、外部から観察可能な、探索戦略の一部となる行動のモデル化をしている。これらは、ユーザインタフェースに関する文献の近くに位置づけられている。このような探索戦略の一部として探索

者がとる行動は、検索システムのユーザインタフェースで実現されることが望ましい機能であり、ユーザインタフェース研究の基礎として密接な関連があると考えられることから、妥当な位置づけであると考えられる。

それに対し、Bates (BA79b, 1979) は、心理学的側面から探索者が考えることに着目して探索行動をモデル化した研究である。これは、探索者がある検索戦略をもちいた理由に関連するものであり、検索へ影響を及ぼした因子と関連があると考えられ、マップ上で「検索への影響因子」グループにやや近い位置にあることは妥当であると考えられる。

また、Matthews (MT, 1983) は OPAC に関する調査であり、ユーザモデルとは直接的な関係はないが、最近の OPAC のユーザインタフェースを扱った文献から、この「探索行動の分析」グループの他の文献と共引用されることが多い。したがって、このグループは、「OPAC」という対象となる検索システムの種類という要素も絡んで形成されており、複雑な構造になっている。

直接的な引用関係においても、第 8 図のごとく、探索戦略に関わる探索行動を扱った Bates (BA79a, 1979) は、検索式の改善法をモデル化した Fidel (FI85, 1985) から引用され、OPAC での主題探索法を扱った Markey (MK, 1984) を通して、Bates (BA86, 1986) などのユーザインタフェースの研究へ通じている。それに対し、心理学的な側面を扱った Bates (BA79b, 1979) は「検索への影響因子」グループから多く引用され、「検索への影響因子」研究に多くの影響を与えたと考えられる。

OPAC に関する Matthews (MT, 1983) は、Bates (BA86, 1986) や「検索への影響因子」グループの Borgman (BO86a, 1986)、Borgman (BO86b, 1986) といった OPAC 検索でのユーザモデルを扱った文献から引用されている。ユーザインタフェースの文献の中で、Williams (WI, 1984) は、このグループの文献と共引用関係はあったが、直接の引用関係はなく、直接影響を与えた関係ではない。

以上のように、この「探索行動の分析」グループでは、探索戦略およびユーザインタフェース研究と、「検索への影響因子」の研究へ向かう 2 つの流れが認められ、Bates (BA79a, 1979) と Bates (BA79b, 1979) は、それぞれその初期の基礎となっている文献である。

(2) 検索への影響因子

これは、オンライン検索や OPAC の検索の検索過程における探索者の行動や検索戦略、検索結果に影響を及

ぼす因子に関する文献からなるグループである。因子としては、経験年数の差、探索者の属性など種々ものが扱われている。この中には、検索に影響を及ぼす因子として、探索者のメンタルモデルを提示している文献からなる「メンタルモデル」のサブグループが認められた。

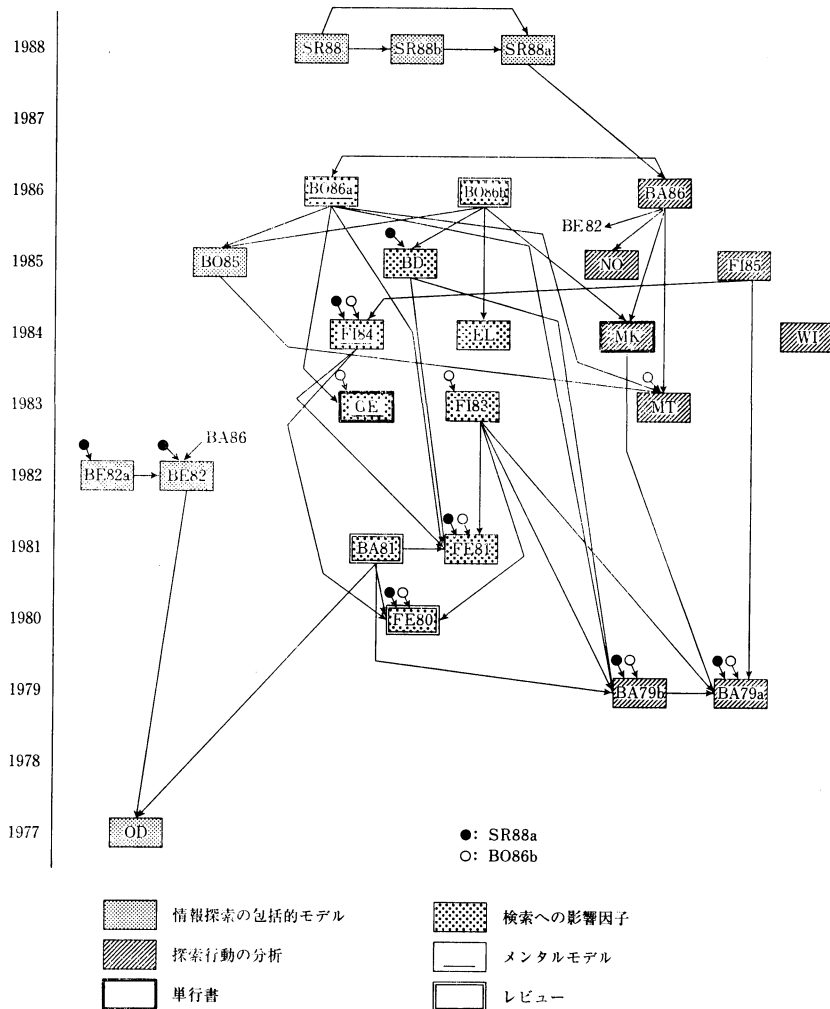
直接引用関係を見ると、第 8 図のごとく、「探索行動の分析」グループの Bates (BA79b, 1979) を多く引用している。Fenichel (FE81, 1981) は多くの矢印が集まっており、このグループの基礎的な文献の一つであると考えられる。この Fenichel ような経験年数による差など外的な要因を扱ったものから、次第に、Fidel (FI84, 1984) など検索に影響する因子としてメンタルモデルを提示した研究など、内的な要因にも関心が広がってきている。また、このグループ内の文献のほとんどは、Borgman (BO86b, 1986, 図中一部 ○) のレビューに引用されている。

(3) 情報探索の包括的モデル

(1) と (2) は、検索質問が提示された後の具体的な検索の過程を対象としているのに対し、このグループは、検索質問が提示される前の状況をも含めたより広いコンテキストで情報探索行動を捉えようとした文献から成っている。

Belkin (BE82, 1982) と Belkin (BE82a, 1982) は、利用者の「問題状況」に着目して、利用者が情報を必要とする状況を扱ったものである。また、Saracevic による一連の情報探索行動に関する研究 (SR88; SR88a; SR88b, 1988) も、利用者の「問題」を考慮する観点に立ち、検索質問が発せられる状況など顕在化していない利用者の情報要求に関連することから、検索質問の構造と分類、探索者の認知過程と判断、同一の検索質問に対する異なる探索法まで、情報探索行動に関わることがらを広く扱っている。Oddy (OD, 1977) は、利用者の状況をシステムが考慮するインタフェースを持ったシステムに関する研究である。Borgman (BO85, 1985) は、情報技術に対する利用者の態度に関する研究であるが、情報検索そのものではなく、情報検索をとりまくものを扱っていることからこのグループに入っていると考えられる。

直接引用関係を見ると、Belkin (BE82, 1982) と Belkin (BE82a, 1982) は、Oddy (OD, 1977) を引用しているが、拡張ユーザモデルグループの他の文献とほとんど引用関係がない。情報探索行動の諸側面を広く扱っている一連の Saracevic の研究 (SR88; SR88a; SR88b, 1988, SR88a は図中一部 ●) は、この拡張ユーザモデル



第8図 引用関係図 (拡張ユーザモデルグループ)
文献を示す略号については、第6表を参照。

グループの文献を広く引用しており、このグループ全体の研究成果の上に成り立っている研究であると考えられる。

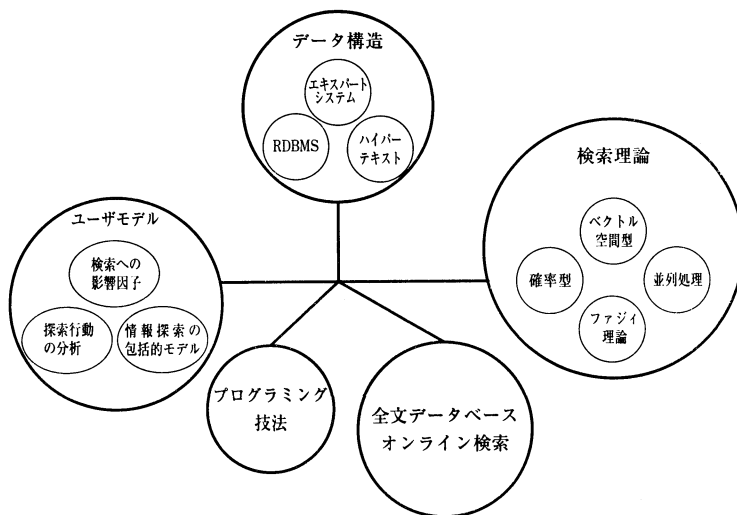
さらに、この拡張ユーザモデルグループ全体をみると、「探索行動の分析」グループから始まって、探索行動とそれを考慮したユーザモデル研究と「検索への影響因子」グループへと発展していく大きな二つの流れがある。この「検索への影響因子」グループの中でも、外的な因子を扱ったものからメンタルモデルなど内的な因子を扱ったものへと対象が広がっている。また、この流れとは別

に、個人の問題状況に焦点を当てた Belkin (BE82; BE82a, 1982) の研究から発展した流れがある。そして、この両者の流れは、情報探索行動の諸側面を広く扱った Saracevic の一連の研究 (SR88; SR88a; SR88b, 1988, SR88a は図中一部 ●) へ通じ、この Saracevic の研究がこのグループ全体の研究成果の上に成り立っていることがわかる。

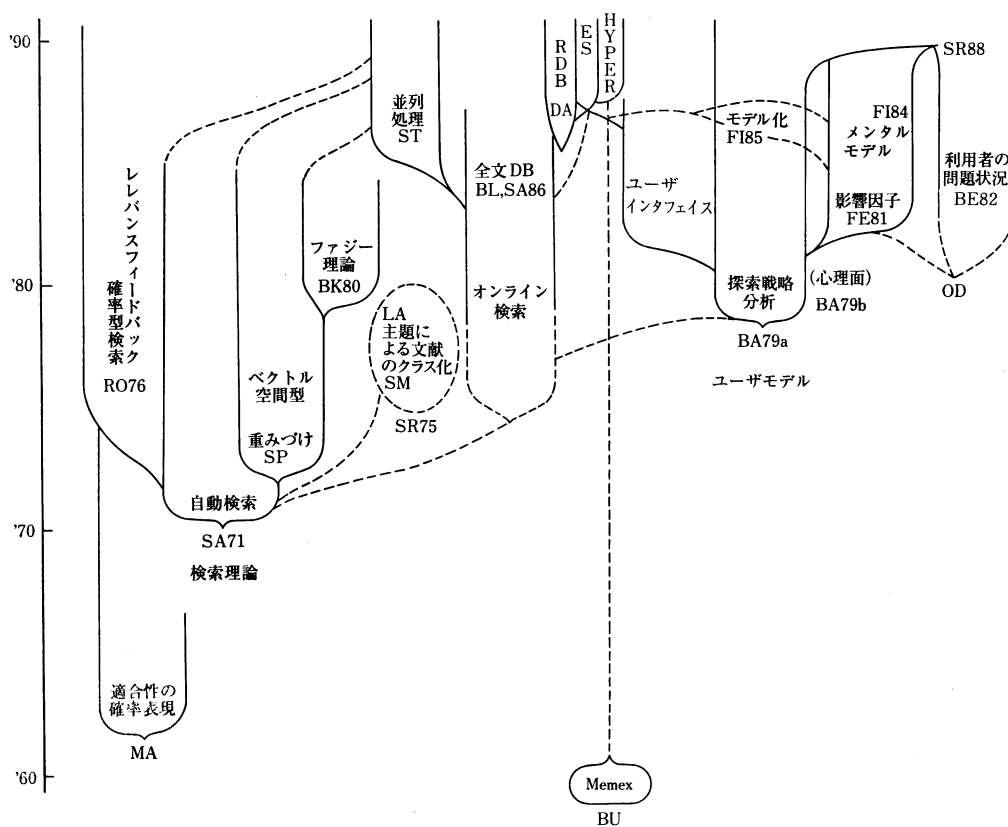
V. おわりに

以上のごとく、情報検索領域の主要な研究領域とその

情報検索分野の構造：引用調査による下位領域と発展過程の分析



第9図 情報検索領域の構造



第10図 情報検索領域の発展過程

文献を示す略号については、第5表・第6表を参照。

発展過程を分析した。その結果を総合的に判断すると、情報検索領域は、大きく第9図のように構成されている。また、発展の経緯を見ると第10図のようになる。

また、ここで用いた共引用マップと直接引用を用いる方法は、一つの研究領域の構造や発展を認識する手法として、大きな可能性があると考えられる。

まず、前述のごとく、共引用は文献間の類似性を示す尺度として妥当性がすでに示されている⁷⁾。そして、このような共引用を用いたマッピングによって類似性が認められた文献群内で、文献間の直接的な引用関係を見ることによって、共引用分析では明らかにならなかった研究領域の経時的な変化を見ることが出来る。また、主題の類似性が認められた文献間での直接的な引用関係は、その共通する主題に関わる部分での影響関係が表れることが期待される。したがって、共引用マッピングによって認定された下位領域を構成する文献群内の直接的な引用関係を分析することによって、各下位領域の中心的な文献や大きな影響を与えた文献を明らかにし、研究の発展過程を示すことが可能であると考えられる。

経時的な変化を示す方法としては、複数時点で作成した共引用マップの比較¹⁰⁾や引用文脈分析¹¹⁻¹²⁾がある。これらの方法はそれぞれ利点があるが、今回用いた共引用マップと直接引用関係の分析という方法は、これらに比べて作業的な負担が少ないという利点もある。ただし、直接引用図において、文献の配置のしかたに決定的な方法がないことは一つの弱点といえるであろう。

また、今回は、共引用マッピングを二段階で行なった。すなわち、一回目の共引用マッピングは、被引用回数に下限を設定して対象文献を選定した。そのマッピングによって示された文献のグループ毎に再度共引用マッピングを行なった。その際、分析に十分な文献数を得るために、各グループ内の文献と類似性が強いと考えられる文献を、一定の共引用強度を下限として抽出し、新たな対象文献群とした。

その結果、一回目も二回目の分析でも、各領域毎に中心となる文献は、検索理論では Sparck-Jones (SP, 1972) と Robertson (RO76, 1976)、ユーザモデルでは (BA79b, 1979) と Fenichel (FE81, 1981) というように同一のものであり、基本的な文献間の構造はほとんど変化がなかった。

また、二回目の分析では、特定の下位領域内の類似性が認められた文献だけを対象として、文献数を増加させて分析を行なったことによって以下のような効果がみら

れた。すなわち、

- ① 一回目の分析では、直接引用でしか確認できなかった下位領域が、共引用マップで確認できる。
- ② 一回目の分析では、直接引用でも明確に確認できなかった微かな動きが、共引用マップで確認できる。
- ③ 一回目の分析では、直接引用でも分けることができなかった下位領域が共引用マップで確認できる。

拡張検索理論グループに対する二回目の共引用マップ上で、確率型検索とベクトル空間型検索が分離されたことは①の例であり、並列処理検索やファジィ理論検索が認められたことは②の例である。また、ユーザモデル領域は③の例である。この様に、分析対象文献を拡張して共引用分析を行なったことによって、さらに詳しい分析が可能になった。

しかし、繰り返し対象文献を拡張するには、対象文献を選定する共引用強度の下限をさらに引き下げることになる。共引用強度が低下すると文献間の類似性を表す力が弱くなるという報告もあり¹³⁾、拡張された文献群内での類似性を保つには、自ずと拡張できる回数には限界があり、何段階も繰り返し拡張することはできない。

この複数段階で共引用マップを作成する手法と直接引用関係の分析を組み合わせる方法は、研究領域の構造を分析する有望な方法の一つであると考えられるが、より精緻化するために、さらに詳細な検討を重ねることが必要である。

この調査にあたってご助言を頂いた慶應義塾大学文学部細野公男、田村俊作、図書館情報大学の岸田和明、慶應義塾大学大学院池谷のぞみの各氏に感謝申し上げます。

- 1) 細野公男. 情報検索理論・技法の問題点とその解決の方向. 情報処理学会研究報告. Vol. 91, No. 98 (情報学基礎 24 (91-FI-24)), p. 1-7 (1991)
- 2) Salton, G.; McGill, M. J. Introduction to Modern Information Retrieval. New York, McGraw-Hill, 1983, 448 p.
- 3) van Rijsbergen, C. J. Information Retrieval. 2nd ed. London, Butterworth, 1979, 152 p.
- 4) Ellis, David. New Horizons in Information Retrieval. London, Library Assn., 1990, 138 p.
- 5) Belkin, N. J.; Croft, W. B. Retrieval Techniques, Annual Review of Information Science and Technology, Vol. 22, p. 109-145 (1987)
- 6) Blair, D. C. Language and Representation in Information Retrieval. Amsterdam, Elsevier Science Pub., 1990, 335 p.

情報検索分野の構造：引用調査による下位領域と発展過程の分析

- 7) 岸田和明, 池谷のぞみ. 文献間の類似性を測定する尺度としての共引用の特徴について. 1991年度日本図書館学会春季研究集会発表資料, 1991, 14 p.
詳細は以下の資料による:
慶應義塾大学文学部図書館・情報学科. 文献間の類似性を測定する尺度としての共引用の妥当性についての評価:「情報学」関連文献を事例として. 平成2年度慶應義塾大学学事振興基金による研究(共同研究)「主題分野のマッピングの手法に関する研究」報告. 東京, 慶應義塾大学文学部図書館・情報学科, 1991, 77 p.
- 8) Small, Henry. Co-citation in the scientific literature: a new measure of the relationship between two documents. *Journal of the American Society for Information Science*. Vol. 24, No. 4, p. 265-269 (1973)
- 9) Miller, Carmen. Detecting duplicates; a searcher's dream come true. Online. Vol. 14, No. 4, p. 27-34 (1990)
- 10) McCaine, Katherine W. Longitudinal author cocitation mapping; the changing structure of macroeconomics. *Scientometrics*. Vol. 6, No. 1, p. 33-53 (1984)
- 11) Small, Henry. Citation context analysis and the structure of paradigms. *Journal of Documentation*. Vol. 36, No. 3, p. 186-196 (1980)
- 12) Small, Henry; Greenlee, E. Citation context analysis of a co-citation cluster; recombinant DNA. *Scientometrics*. Vol. 2, No. 4, p. 277-301 (1980)
- 13) Shaw, W.M., Jr. Critical thresholds in co-citation graphs. *Journal of the American Society for Information Science*. Vol. 36, No. 1, p. 38-43 (1985)

付録 文献コードと書誌事項

- BA79a Bates, M. J. Information search tactics. JASIS. 30(4): 205-214 (1979).
- BA79b Bates, M. J. Idea tactics. JASIS. 30(5): 280-289 (1979).
- BA81 Bates, M. J. Search Techniques. ARIST. 16: 139-169 (1981).
- BA86 Bates, M. J. Subject access in online catalogs: a design-model. JASIS. 37(6): 357-376 (1986).
- BE82 Belkin, N. J.; Oddy, R. N.; Brooks, H. M. ASK for information-retrieval. 1. background and theory. J DOC. 38(2): 61-71 (1982).
- BE82a Belkin, N. J.; Brooks, H. M.; Oddy, R. N. ASK for information retrieval. Part II. Results of a design study. J DOC. 38(3): 145-164 (1982).
- BE87 Belkin, N. J.; Borgman, C. L.; Brooks, H. M.; Bylander, T.; Croft, W. B.; Daniels, P.; Deerwester, S.; Fox, E. A.; Ingwersen, P.; Rada, R.; Sparck-Jones, K.; Thompson, R.; Walker, D. Distributed expert-based information-systems: an interdisciplinary approach. IPM. 23(5): 395-409 (1987).
- BD Bellardo, T. An investigation of online searcher traits and their relationship to search outcome. JASIS. 36(4): 241-250 (1985).
- BL Blair, D. C.; Maron, M. E. An evaluation of effectiveness for a full-text document-retrieval system. CACM. 28(3): 289-299 (1985).
- BK80 Bookstein, A. Fuzzy requests: an approach to weighted Boolean searches. JASIS. 31(4): 240-7 (1980).
- BK83 Bookstein, A. Information retrieval: a sequential learning process. JASIS. 34(5): 331-342 (1983).
- BK83a Bookstein, A. Outline of a general probabilistic retrieval model. J DOC. 39(2): 63-72 (1983).
- BK85 Bookstein, A. Probability and fuzzy-set applications to information retrieval. ARIST. 20: 117-178 (1985).
- BO85 Borgman, C. L. Individual differences in the use of technology: work in progress. Proceedings of the American Society for Information Science. 22: 243-249 (1985).
- BO86a Borgman, C. L. The users mental model of an information-retrieval system: an experiment on a prototype online catalog. INT J MMS. 24(1): 47-64 (1986).
- BO86b Borgman, C. L. Why are online catalogs hard to use: lessons learned from information-retrieval studies. JASIS. 37(6): 387-400 (1986).
- BUE Buell, D. A.; Kraft, D. H. A model for a weighted retrieval system. JASIS. 32(3): 211-16 (1981).
- BU Bush, V. As we may think. Atlantic Monthly. 176(1): 101-108 (1945).
- CO Conklin, J. Hypertext: an introduction. Computer. 20(9): 17-41 (1987).
- CR79 Croft, W. B.; Harper, D. J. Using probabilistic models of document retrieval without relevance information. J DOC. 35(4): 285-95 (1979).
- CR81 Croft, W. B. Document representation in probabilistic models of information retrieval. JASIS. 32(6): 451-457 (1981).
- CR Croft, W. B.; Thompson, R. H. I3R: a new approach to the design of document-retrieval systems. JASIS. 38(6): 389-404 (1987).
- CR88 Croft, W. B.; Savino, P. Implementing ranking strategies using text signatures. ACM transactions on Office Information Systems. 6(1): 42-62 (1988).
- DA Date, C. J. An Introduction to Database Systems. 3rd ed. Addison-Wesley Publishing Company, 1981.
- EL Elkerton, J.; Williges, R. C. Information retrieval strategies in a file: search environment. Human Factors. 26(2): 171-184 (1984).
- FA Faloutsos, C.; Christodoulakis, S. Description and performance analysis of signature file methods for office filing. ACM Transactions on Office information Systems. 5(3): 237-257 (1987).
- FE80 Fenichel, C. H. The process of searching online bibliographic databases: a review of research. Library Research. 2(2): 107-127 (1980).
- FE81 Fenichel, C. H. Online searching: measures that discriminate among users with different types of experiences. JASIS. 32(1): 23-32 (1981).

- FI83 Fidel, R.; Soergel, D. Factors affecting online bibliographic retrieval: a conceptual framework for research. JASIS. 34(3): 163-180 (1983).
- FI84 Fidel, R. Online searching styles: A case study based model of searching behavior. JASIS. 35(4): 211-221 (1984).
- FI85 Fidel, R. Moves in online searching. Online Riview. 9(1): 61-74 (1985).
- FI Fidel, R. Towards expert systems for the selection of search keys. JASIS. 37(1): 37-44 (1986).
- FS Fishburn, P.C. Utility theory for decision making. New York: Wiley, 1970.
- FO Fox, E.A. Development of the CODER system: a testbed for artificial-intelligence methods in information-retrieval. IPM. 23(4): 341-366 (1987).
- GE Gentner, D.; Stevens, A.L. ed. Mental Models. Lawrence Erlbaum Associates, Inc. Publishers, 1983.
- HA Harrison, M.C. Implementation of the substring test by hashing. CACM. 14(12): 777-779 (1971).
- KN Knuth, E.D. The Art of Computer Programming. Vol. 3: Sorting and Searching. Addison-Wesley Publishing Company, 1973.
- LA Lancaster, F.W. Vocabulary Control for Information Retrieval. 2nd ed. Information Resources Press, 1986.
- LI Liddy, E.; Bonzi, S.; Katzer, J.; Oddy, E. A study of discourse anaphora in scientific abstracts. JASIS. 38(4): 255-261 (1987).
- LS Losee, R.M. Parameter estimation for probabilistic document-retrieval models. JASIS. 39(1): 8-16 (1988).
- LO Lovins, B.J. Development of a stemming algorithm. Mechanical Translation and Computational Linguistics. 11: 22-31 (1968).
- MK Markey, K. Subject Searching in Library Catalogs: before and after the Introduction of Online Catalogs. OCLC Inc., 1984.
- MA Maron, M.E.; Kuhns, J.L. On relevance, probabilistic indexing and information retrieval. JACM. 7(): 216-244 (1960).
- MT Matthews, J.R.; Lawrence, G.S.; Ferguson, D.K. Using Online Catalogs: A Nationwide Survey. Neal-Schuman Publishers, 1983.
- NO Noerr, P.L.; Bivins Noerr, K.T. Browse and navigate: An advance in database access methods. IPM. 21(3): 205-213 (1985).
- OD Oddy, R.N. Information-retrieval through man-machine dialog. J DOC. 33(1): 1-14 (1977).
- RO76 Robertson, S.E.; Jones, K.S. Relevance weighting of search terms. JASIS. 27(3): 129-146 (1976).
- RO77 Robertson, S.E. Probability ranking principle in IR. J DOC. 33(4): 294-304 (1977).
- RO81 Robertson, S.E. "4. Probablistic models of indexing and searching." In: Oddy, R.N.; Robertson, S.E.; Van Rijsbergen, C.J.; Williams, P.W. (eds.) Information Retrieval Research. Butterworths, 1981, p. 35-56.
- RO82 Robertson, S.E.; Maron, M.E.; Cooper W.S. Probability of relevance: a unification of two competing models for document retrieval. Information Technology: Research and Development. 1(1): 1-21 (1982).
- RO86 Robertson, S.E. On relevance weight estimation and query expansion. J DOC. 42(3): 182-188 (1986).
- SA68 Salton, G. Automatic Information Organization and Retrieval. McGraw-Hill, 1968.
- SA71 Salton, G. ed. The SMART Retrieval System. Englewood Cliffs, N.J., PrSentic-Hall, 1971.
- SA75a Salton, G. et al. A vector space model for automatic indexing. CACM. 18(11): 613-620 (1975).
- SA75b Salton, G. et al. A theory of term importance in automatic text analysis. JASIS. 26(1): 33-50 (1975).
- SA76 Salton, G.; Wong, A. On the role of words and phrases in automatic text analysis. Computers and the Humanities. 10(2): 69-87 (1976).
- SA83a Salton, G. et al. Extended boolean information retrieval. CACM. 26(12): 1022-1036 (1983).
- SA83b Salton, G. Introduction to Modern Information Retrieval. New York. McGraw-Hill, 1983.
- SA86 Salton, G. Another look at automatic text-retrieval system. CACM. 29(7): 648-656 (1986).

- SA88 Salton, G.; Buckley, C. Parallel text search methods. CACM. 31(2): 202-215 (1988).
- SR75 Saracevic, T. Relevance: review of a framework for thinking on for thinking on notion in information-science. JASIS. 26(6): 321-343 (1975).
- SR88a Saracevic, T.; Kantor, P.; Chamis, A. Y.; Trivison, D. J. A study of information seeking and retrieving. I. Background and methodology. JASIS. 39(3): 161-176 (1988).
- SR88b Saracevic, T.; Kantor, P. A study of information seeking and retrieving. II. Users, questions, and effectiveness. JASIS. 39(3): 177-196 (1988).
- SR88 Saracevic, T. A study of information seeking and retrieving. 3. Searchers, searches, and overlap. JASIS. 39(3): 197-216 (1988).
- SE Sewell, W.; Teitelbaum, S. Observations of end-user online searching behavior over 11 years. JASIS. 37(4): 234-245 (1986).
- SN Smeaton, A. F.; Van Rijsbergen, C. J. The retrieval effects of query expansion on a feedback document retrieval system. Computer Journal. 26(3): 239-246 (1983).
- SM Small, H. Co-citation in scientific literature: new measure of relationship between 2 documents. JASIS. 24(4): 265-269 (1973).
- SP Sparck-Jones, K. A statistical interpretation of term specificity and its application. J DOC. 28(1): 11-21 (1972).
- SO Stone, H. S. Parallel querying of large databases: A case study. Computer. 20(10): 11-21 (1987).
- ST Stanfill, C.; Kahle, B. Parallel free-text search on the connection machine system. CACM. 29(12): 1229-1239 (1986).
- SW: Swanson, D. R. Historical note: Information-retrieval and the future of an illusion. JASIS. 39(2): 92-98 (1988).
- TE Tenopir, C. Full text database performance. Online Review. 9(2): 149-164 (1985).
- VR77 Van Rijsbergen, C. J. A theoretical basis for the use of co-occurrence data in information retrieval. J DOC. 33(2): 106-119 (1977).
- VR79 Van Rijsbergen, C. J. Information Retrieval. 2nd. ed. London, Butterworth, 1979, 152 p.
- WA Waller, W. G.; Kraft, D. H. A mathematical model of a weighted Boolean retrieval system. IPM. 15(5): 235-45 (1979).
- WZ: Waltz, D. L. Applications of the connection machine. Computer. 20(1): 85-97 (1987).
- WE Weyer, S. A. The design of a dynamic book for information search. INT J MMS. 17(1): 87-107 (1982).
- WI Williams, M. D. What makes RABBIT run? INT J MMS. 21(4): 333-352 (1984).
- YU76 Yu, C. T.; Salton, G. Precision weighting: an effective automatic indexing method. JACM. 23(1): 76-88 (1976).
- YU83 Yu, C. T.; Buckley, C.; Lam, K.; Salton, G. A generalized term dependence model in information retrieval. Information Technology: Research and Development. 2(4): 129-54 (1983).
- DI Directory of Online Databases 1988, 1888.

ARIST: Annual Review of Information Science and Technology.

CACM: Communications of the ACM.

INT J MMS: International Journal of Man-Machine Studies.

IPM: Information Processing and Management.

JACM: Journal of the Association for Computing Machinery.

JASIS: Journal of the American Society for Information Science.

J DOC: Journal of Documentation.