

情報検索の適合度判定における非キーワード情報の意義

The Implications of the Non-Keyword Information on the
User's Judgment of Relevance in
Information Retrieval

相 良 佳 弘
Yoshihiro Sagara

Résumé

Most of the present information retrieval systems presuppose subject keywords as an access point as well as the criterion for relevance judgment. In other words, IR systems mainly determine relevant documents by keywords. However, users may look at other document characteristics in addition to the subject when judging relevance of the documents. This means that criteria of judging relevance between the system and end-users may be different. In order to overcome the discrepancy, it is necessary to make clear what kinds of non-keyword information are necessary for user's relevance judgments and how they are identified and used for relevance judgment.

Firstly this paper investigated whether and how each of the 17 characteristics that documents for 3 graduated students and 3 information specialists, usually have was used for the judgment of relevance by IR users. A survey was conducted. The result of the survey showed that users judged relevance using various information from the 17 characteristics. Non-keyword information which contains events, research methods and novelty of the idea, as well as the length of original documents and of abstracts were determined as predominant information for judging relevance. Non-keyword information has the same importance as the information from keywords in retrieving documents.

Secondly, the author discussed the applicability of those kinds of information to enhance retrieval performance. Such kind of information as the length of documents and publication dates will be easily implemented into IR systems since they are parts of bibliographic data. Others such as novelty are more difficult to use, because the value of those kinds of information are dependent on individual user.

相良佳弘：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
Yoshihiro Sagara: Graduate School of Library and Information Science, Keio University, Mita 2-15-45,
Minato-ku, Tokyo
e-mail: sagara@slis.keio.ac.jp
受付日：1999年11月11日 最終改訂稿受付日：2001年1月9日 受理日：2001年2月2日

Finally, this paper suggests the possibility to introduce user models into a IR system since user models can represent such characteristics of each user as novelty. They will be helpful for the system to deal with non-keyword information.

- I. 情報検索における適合度判定
 - A. キーワードの出現頻度に基づく適合度判定
 - B. ユーザによる適合度判定
 - C. 判定に用いる情報の不一致
- II. 適合度判定に用いられる文献特徴
 - A. 文献の特徴
 - B. 文献特徴と非キーワード情報
- III. 適合度判定における文献特徴の利用状況
 - A. 調査の目的と手順
 - B. 調査結果
- IV. 非キーワード情報の意義と影響
 - A. 適合度判定におけるキーワード情報と非キーワード情報
 - B. 非キーワード情報の検索システムへの応用とその課題
- V. 今後の課題

I. 情報検索における適合度判定

A. キーワードの出現頻度に基づく適合度判定

従来、オンライン検索で広く用いられてきたブール検索では、検索式中に含まれるキーワードが存在するか否かの二値的な判定を行っており、検索結果としてユーザに示されるレコードの適合性は、いずれも同じ程度であるとみなされている。したがってユーザは、出力された個々のレコードの適合性を把握するために、全てのレコードを評価しなければならなかった。そのため検索結果が多すぎる場合には、適合性を判定する際のユーザの負担が過大になるという問題が生じる。これは無益点の問題と呼ばれており、ブール検索の大きな問題点として指摘されている¹⁾。

このブール検索の問題点に対して、検索式と蓄積されたレコードとの一致度を多値的に判定しようとする、非ブール型検索手法の実用化が進んでいる。非ブール型検索の多くでは、キーワードの出現頻度から検索結果の順位付け出力を行う適合度順検索が可能である。なお、適合度とは適合性を多値的に判定したものであり、質のみではなく

量的な側面から捉えたものである。

Stibic は、適合度順検索の利点について、以下のようにまとめている²⁾。

- ①広い意味範囲を持つ語や曖昧な語を用いて検索することができる
- ②非常に限定的な意味を持つ語を用いて検索することができる

これまでのブール検索では、広い意味範囲を持つキーワードや曖昧なキーワードを用いた検索は、出力件数が多くなりすぎユーザの負担が大きくなるとされてきた。しかし、適合度順検索ではキーワードの出現頻度により順位づけて出力を行うため、そのキーワードが示す主題を一層中心的に扱った文献についてのレコードが上位に出力される。したがって、①のような検索においても十分情報要求を満たすことができる²⁾。

また、非常に限定的な意味を持つキーワードを多数用いる検索も避けるべきであるとされてきた。ブール検索では検索結果が得られなかったり、検索もれが生じる可能性が高くなるためであ

る。しかし、適合度順検索では、検索式中の全てのキーワードを含まなくても検索結果が得られるため、ある程度の再現率を期待することができる²⁾。このように、検索結果を順位として出力する検索システムは、Salton による SMART システム³⁾を先駆けとして、今日まで様々な研究が行われており、多くのサーチエンジンなどで用いられるようになってきている。これは、検索結果が膨大な場合には、適合度、有用度、満足度が高いと考えられるレコードから順に出力する方が、情報検索システムの効果をあげると考えられるからである。

Robertson は、この適合度順に出力することの重要性を指摘し、適合度や有用度の確率が高いものから順に検索結果を出力するというアプローチを示した⁴⁾。この Robertson の考えは確率的順位付け原理と呼ばれ、ある情報要求に対して特定のキーワードを含む文献が適合する確率に基づいて順位付けを行うものである。今日広く用いられているキーワードの出現頻度から算出するアルゴリズムとは異なっているが、ユーザが望む順に結果を出力することが、検索システムにとって重要な意義を持つことを示したものである。

これまでに、適合度順検索システムの順位付けの手法は、統計的手法や文献間の関連性に基づいた活性化伝播法、クラスター法など様々な手法が研究されてきている。その中でも近年実用化されつつあるシステムの多くは、レコード中のキーワードの出現頻度に基づいて、適合度順に順位付けする手法を用いたものである。この手法は、キーワードの出現頻度からその文献の主題と検索式の類似性を算出し、順位付けを行うものであり、しばしばキーワードの出現位置や近接度、逆文献頻度⁵⁾などで補正される。キーワードの出現頻度に基づいた手法は、検索式と蓄積レコードの類似度から適合度順出力を行おうとする手法である。この順位付け手法は、WWW のサーチエンジン⁶⁾⁷⁾などで広く用いられている手法であるが、商用データベースにおいても DIALOG の TARGET などで利用されている。

例えば、TARGET の順位付けアルゴリズムは、

以下の四つの手法を用いていると公式に発表されている⁸⁾。

- ①入力されたキーワードのレコード内での出現頻度
- ②入力されたキーワードのレコード内での出現位置
- ③入力されたキーワード間の近接度
- ④入力されたキーワードのデータベース内での出現頻度

アルゴリズムの詳細は公開されていないが、Tenopir 等⁹⁾ および Keen¹⁰⁾ は、TARGET のアルゴリズムが一般的なキーワードの出現頻度に基づく手法であると指摘している。

しかし、キーワードの出現頻度に基づく手法は広く用いられているものであるが、適合度の順に検索結果を出力できているかという点と不十分な面もある。第 1 表は 25 件の情報要求を対象として、DIALOG の TARGET システムによる検索結果の順位付けと、ユーザによる適合度順とを比較した実験の結果を示したものである¹¹⁾。ここでは順位付けの一致度を比較するために順位相関係数を用いている。この係数が +1 であればシステムによる順位付けとユーザによる順位付けは完全に一致したことになり、-1 であれば逆順になったことを示している。多くの事例で正の値は得られているものの 0 に近い値となっており、ユーザが望んだ順位付けとは相関性の無い順位で出力がなされていると言える。つまり、キーワードの出現頻度情報を用いた適合度順検索では、ブール検索の問題点の一部が改善されたものの、十分な順位付けで検索結果を提供できているとは必ずしも言えないのである。

理想的な情報検索とは、ユーザが実際に利用したいと考える情報を含むレコードのみを的確に出力することである。しかし、実際には検索もれやノイズを避けることは困難であり、検索システムからの出力と、ユーザが望む文献とが一致することはまれである。とくに非ブール型の検索システムでは、システムによって判定された順位付けと

第1表 適合度順検索システムとユーザとの順位付けの一致度(相関係数)¹¹⁾

情報要求	相関係数
1	-0.09
2	0.02
3	0.29
4	0.30
5	0.26
6	0.44
7	-0.03
8	0.32
9	0.06
10	0.34
11	0.27
12	0.00
13	0.29
14	-0.16
15	0.00
16	0.30
17	0.61
18	0.58
19	-0.22
20	0.06
21	0.29
22	0.09
23	0.00
24	0.36
25	-0.16
平均	0.17

注) 相良¹¹⁾の実験データを表にまとめなおした

ユーザにとって望ましい順位付けとが、より近づくことが求められている。つまり検索システムには、よりユーザの適合度判定に近い検索結果を出力することが求められていると言えよう。

B. ユーザによる適合度判定

理想的な情報検索を考えるうえで、まずユーザにとって望ましい順位付けとはどのような順位付けかを明らかにするために、ユーザにとっての適合度はどのように判定されているのか検討を行う。これまでに、T.K. Park¹²⁾、Barry¹³⁾、H. Park¹⁴⁾は、ユーザの行う適合性判定について様々な研究を行っている。

T.K. Parkは、ユーザが検索結果を評価する過程でインタビューを行い、適合性の評価に影響を与える要因を収集した。そして、利用者が文献のタイトル、著者名、雑誌名などの各書誌要素の価値をどのように解釈しているかについての分析を行った。このとき分析した書誌要素は、「タイトル」、「タイトルの形式」、「著者名」、「雑誌名および文献の形態」、「抄録」の五つである。しかし、六番目に「異なる書誌要素間の相互関連」を挙げ、これも適合性の評価に影響を及ぼしていると指摘した¹²⁾。

- ①「タイトル」……文献の主題を示す要素であり、タイトル中のキーワードはその文献の扱っている分野を知らせる働きがある
- ②「タイトルの形式」……文献の信頼性を判断する
- ③「著者名」……著名な著者は重要な適合性の判断基準となる
- ④「雑誌名および文献の形態」……その文献の信頼性や入手が容易か否かを示す
- ⑤「抄録」……個々のユーザによって認識が変化する要素であり、他の要素で十分内容が把握できている場合は補完的な存在だが、そうではない場合、判断の中心的な基準として用いられる
- ⑥「複数の要素間の相互関連」……単独の基準を用いて判断するのではなく、複数の基準を組み合わせ用いられる

T.K. Parkは、判断基準としてこれら六つの要素がどのように使用されていたかを分析した上で、適合性評価に影響を与える要因を、以下の三

つのカテゴリに分類している。

- 1) 内的文脈 (Internal Context)……過去の経験や専門知識の程度
- 2) 外的文脈 (External Context)……検索の目的や研究の段階、最終的な目的といった状況的な要因
- 3) 問題文脈 (Problem Context)……その問題の捉え方

これら三つのカテゴリは、個々のユーザの問題意識、経験、知識などから生じるものである。ある文献がユーザにとって適合か否かの判定は、文献が持つ要素を判断し決定されるが、内的文脈、外的文脈、問題文脈といった三つのカテゴリで示されるユーザやユーザのおかれた状況の影響を受ける。適合性の判断は、ユーザの質問と文献との間の客観的な対応関係によってのみ行われるものではなく、非常に複雑な行為であると指摘している¹²⁾。

Barry は、適合性の判定を行う基準として 23 のカテゴリを導き出している。この 23 のカテゴリのうち、半数以上は文献内容の新奇性や正確さ、明瞭さなど主題と関連のないカテゴリであった。このことから、ユーザは主題だけではなく、情報要求を持つに至った状況など様々な観点から評価を行っていると言われている。しかし、これらの基準は個々のユーザごとに全く異なるのではなく、ある程度共通していると述べている¹³⁾。

さらに H. Park は、ユーザが行う適合性判定を観察した結果から、適合性判定には「応用性」、「有用性」、「類似性」など 33 の次元が影響すると指摘している。この中で「応用性」や「有用性」は、得られた文献がユーザの問題に应用できるか否かや、役に立つか否かといった次元である。一方「類似性」は、主題志向の適合性判定に見られるような、情報要求と文献とが共通の記述を持っている場合を示すものである。H. Park は、多くの検索システムが、「類似性」しか考慮していないと指摘し、今日の情報検索システムはユーザが行う適合性判定の一部分しか反映していないと述べ

ている。さらに、文献と情報要求の「類似性」以外にも、「応用性」や「有用性」をはじめとした多くの次元から、ユーザは文献に対する評価を行っていると言われている¹⁴⁾。

T. K. Park や Barry が指摘したように、ユーザが適合性判定をする際には、単なる主題の一致度だけではなく、文献について様々な観点から評価を行っていると考えられる。これは、H. Park の言う「応用性」や「有用性」などの次元が、ユーザによる適合性判定には重要であることを示している。つまり、「類似性」が高く主題が一致するからといって、適合性判定における他の次元である「応用性」や「有用性」も高いとは言えないのである。

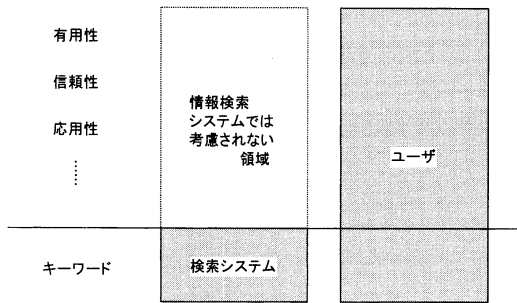
C. 判定に用いる情報の不一致

主題検索は、ユーザが持つ情報要求の主題概念と、文献に含まれる主題概念とを照合することによって検索結果を得ることを理想としている。しかし、実際には主題概念を直接照合することは不可能と言える。そこで、情報要求の主題概念を示すキーワードと、文献中に含まれるキーワードとを照合することによって検索結果を出力することが行われている¹⁵⁾。ほとんどの非ブール型検索システムも、主題概念を示すキーワードの出現頻度に基づいた検索を行っており、キーワードの有無や出現頻度といった情報を用いた検索である。

このように今日の検索システムでは、検索式と文献がキーワードで表現され、このキーワード間の関係によって、適合文献であるか否かの判定を行っている。ここで判定に用いられている文献の主題概念と検索式が示す主題概念との一致度は、主題志向の適合性と呼ばれている。これまで広く用いられてきたキーワードに基づいた検索システムは、主題志向の適合度判定に基づいて検索を行っていると言える。

一方、ユーザが文献の適合度を判定する際には、キーワードで表現される文献の主題だけではなく、文献に含まれる情報の信頼性や新奇性、有用性といった様々な観点から適合度の判定を行っている。システムが出力する文献とユーザが望む

情報検索の適合度判定における非キーワード情報の意義



第1図 システムとユーザの適合度判定の不一致

- ②意識された要求
- ③具体化された要求
- ④調整された要求

実際に検索システムに入力される段階の情報要求は、段階④の調整された要求すなわち検索式である。これは必ずしもユーザが本来持っていた情報要求と同じではない。なぜならば、検索式はキーワードの類似性のみで判定するというシステムの制限にあわせて調整した情報要求であって、ユーザの情報要求を的確に反映しているとは限らないからである。ユーザ本来の情報要求は、段階②の意識された要求であり、それを表現した段階③の具体化された要求であろう。これらの情報要求は、検索式のような妥協の産物ではない。

文献とは、それぞれ異なる判定によって決定されたものと言える。システムによる判定は、キーワードが示す主題の類似性によってなされているのに対して、ユーザが望んだ文献は主題の類似性以外に、情報の信頼性や応用性といったキーワード以外から得られる特徴も用いた総合的な判定である。そのため、システムは文献が持つ様々な文献の特徴のうちキーワードのみを利用しているのに対して、ユーザが行う適合度判定ではより多くの文献の特徴が用いられていると予想される。

情報検索システムが、調整された情報要求である検索式に基づいて結果を出力しているのに対し、検索結果の適合度を判定するユーザは、キーワードの類似性のみではなく総合的な適合度判定を行っており、そこでの情報要求はキーワード以外の情報も多く含んでいると思われる。つまり、ユーザは自らの意識の中で表現されている情報要求や、具体化された情報要求との比較において、検索結果の評価を行っているのである。

確かに文献の特徴のうち、キーワードで表現される主題は、適合度を判定する上で非常に大きな影響を及ぼしている。そのため、キーワードを用いた主題検索が情報検索の中心的な存在として研究対象となってきた。しかしユーザは、キーワードで表現しうる特徴だけではなく、信頼性や有用性までを含めて文献を捉えている。つまりシステムとユーザの間には、第1図のように適合度の判定に用いる情報に不一致がある。

同様なことは、Borlund等¹⁷⁾も指摘している。Borlund等は、情報検索システムの評価手法に関する研究の中で、状況の適合性と主題志向の適合性の妥当性について述べている。情報要求はユーザの認知的状況に応じて様々に変化する。したがって状況的適合性は、文献のような検索対象とこの様々に変化する情報要求との関係に基づいて判定される。つまり状況的適合性は、ユーザの認知的状況に応じて同じ検索対象でも異なった判定がなされることを示している。一方、主題志向の適合性は、ユーザの認知的状況の変化に関わらず、情報要求を記述した検索式と検索対象との間の関連性を示しているにすぎないと指摘している。TREC (Text REtrival Conference) 等で用いられているような、情報要求の認知的背景を一切考慮せずに作成された検索式を用いた適合性判定が、主題志向の適合性である。Borlund等は、こ

また、適合度は文献と情報要求との関係を示すものである。そのため、適合度判定の不一致は、判定に用いられる情報に不一致があることを示していると言える一方で、判定の基準となる情報要求にも不一致があることを示しているとも見ることができよう。この適合度判定の基準となる情報要求の不一致を、Taylorの情報要求形成の四段階を用いて検証することが可能である。

Taylorの情報要求形成の四段階は、以下のように示される¹⁶⁾。

- ①直感的要求

の主題志向の適合性と状況的適合性の間には、関連性は全く認められないと指摘した上で、主題志向の適合性評価では不十分だと述べている。情報要求としてシステムが用いている検索式と、検索結果の評価を行う際の基準となる本来の情報要求とは、同じものではない。そのため、たとえキーワードが一致した検索結果であったとしても、ユーザが求める文献とは異なったものが得られることがしばしば見られる。

この問題は、キーワードのみを用いて検索を行う検索システムでは、常に考慮しなければならない。そのため、適合性フィードバックや、書誌情報などを手がかりとして情報要求に合致するレコードのみに検索結果を絞り込むことが望ましい。しかし、根本的な解決は未だなされていない。

このような問題を解決し、検索システムがユーザの望む検索結果を適切に出力するためには、文献が持つ多様な特徴をユーザがどのように捉え、判定に利用しているかを明らかにする必要がある。本論文では、文献の特徴のうち検索時にシステムが利用するものと、ユーザが適合度判定を行う際に用いるものとを比較することによって、両者の違いを示す。さらに、適合度を判定する際に、個々の特徴がどのように用いられているかも明らかにし、検索システムへ応用する場合の課題を検討する。

II. 適合度判定に用いられる文献特徴

A. 文献の特徴

適合度判定に使用することができる文献特徴には、様々なものがある。一般に書誌情報と呼ばれ、文献タイトル、著者名、出版年などがその例

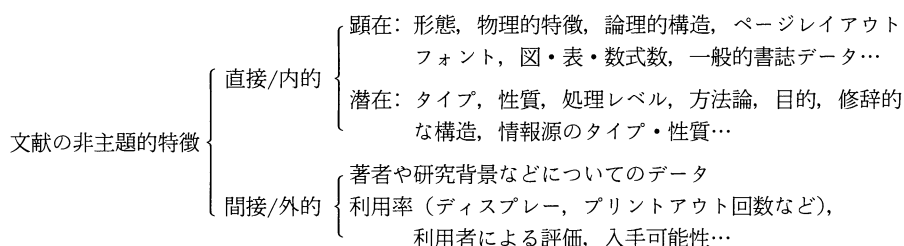
であるし、文献でどのような事柄が表現されているか、つまり主題も、文献の特徴であると言える。このように文献の特徴には様々な種類が存在する。前章で述べた T.K. Park の分析した書誌要素¹²⁾ は、文献が持つ多様な特徴の一面を示している。

さらに、鄒等¹⁸⁾ は文献にはキーワードのような主題的な特徴だけではなく、形態、物理的特徴、論理的構成といった顕在的な特徴と、文献のタイプ、性質、文献で用いられた方法論のような潜在的な特徴とからなる非主題的な特徴があるとしている。第2図は、鄒等の研究における非主題的な特徴である。これは、どのような非主題的な特徴を文献が持っているかを理論的にまとめたものである。

文献には主題を示す特徴に加え、鄒等が示す多様な非主題的な特徴が含まれており、文献は多くの特徴を持っていると言える。しかし、検索時のユーザが、これら文献の特徴の全てを利用できるとは限らない。例えば、実際のページレイアウトやフォントといった特徴は、その文献を実際に入手して初めて知りうる特徴である。書誌情報データベースを利用した検索において、これらの特徴を適合度判定に用いることは困難である。鄒等の示した特徴の中には、検索結果を出力した時点で利用できるものと、実際に文献を入手した後に利用可能になるものと両方が含まれている。この点を踏まえ、検索されたレコードの適合度判定を行う時点で利用できるかという観点から、文献の特徴を以下のようにまとめることができる。

①直接利用可能な特徴

主題を示すキーワード、文献の形態、文献



第2図 鄒等による文献の非主題的特徴¹⁸⁾

の長さ、出版年月¹⁹⁾、情報源についてのデータや著者についてのデータなどの書誌情報

②抄録、書誌情報等から類推しなければ利用できない特徴

著者の専門分野、情報源の入手可能性、文献で扱っている方法論、目的、論旨の明確さ、図・表の有無、事例の有無、新奇性など文献の性質

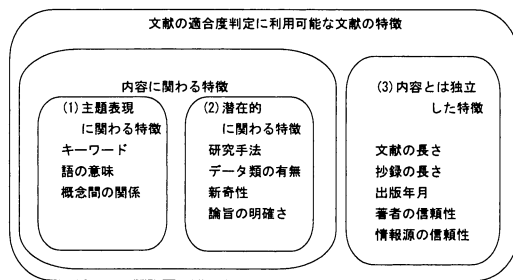
③利用できない特徴(原文献入手以前には利用が困難な特徴)

原文献が持つ物理的特徴、ページレイアウト、フォントなど

書誌データベースを検索する場合には、ユーザは検索されたレコードに記録されていない文献特徴を知ることができない。そのため、検索結果を出力した時点で利用可能な文献特徴は限られる。しかし、直接利用できない特徴の中には、抄録等から類推可能なものもある。実際、方法論や目的は抄録の中で記述されることがあり、またユーザは抄録から文献内容の新奇性や論旨の明確さについて類推を行っていると考えられる。そのため、多くの非主題的特徴は、検索結果を出力した時点で利用可能であると言える。しかし、原文献を入手しなければ判断できない情報も存在する。本論文では検索結果を出力した時点での適合度判定を取り上げるため、この時点で利用できない文献特徴については、対象としないことにする。

B. 文献特徴と非キーワード情報

文献が持つ多様な特徴を、適合度判定の観点から分類したものが第3図である。文献の特徴は、内容に関わる特徴と、内容から独立して存在する特徴とに分けることができる。さらに、内容に関わる特徴は、主題表現に関わるものと、明確に表現されることの少ない潜在的なものに分類できる。いわゆる主題は、文献の内容に関わる特徴の中でも主題表現に関わる特徴によって示される部分であり、ここでは文献の内容の方がより広い概念とする。



第3図 適合度判定における文献の特徴

文献の内容と潜在的に関わる特徴(2)、および内容から独立した特徴(3)は、鄒等の指摘する非主題的特徴と本質的に重なり合う部分がある。本論文では検索結果を出力した時点で利用可能な文献特徴を、文献の内容と関わるか否かという観点から、第2表のように17の特徴に再度整理している。この17の特徴は、T.K. Parkの分析した書誌要素¹²⁾や鄒等¹⁸⁾の示した文献の特徴を参考にしてまとめたものである。第2表にあげた17の特徴以外にも、検索結果の評価を行う際に用いられる特徴がある可能性は避けられないが、ほとんどの検索事例はこの17の特徴によって適合度判定が行われていると思われる。

主題表現に関わる特徴は、いわゆる文献の主題を表現するために用いられてきた特徴である。17の特徴のうち主題表現と関わる特徴として、まずキーワードがあげられる。キーワードは、蓄積レコードや検索式に含まれる主題を表す語であり、適合度の判定に大きな影響を及ぼす文献特徴と言える。ユーザにとって適合度を高いと判定するために必要なキーワードが、質的および量的にすべて含まれていれば、ユーザはこれらのキーワードからそのレコードの適合度を判定すると考えられる。一方、ユーザにとって必要なキーワードが欠けているために適合度が低いと判断された場合も、ユーザはキーワードから適合度判定を行ったと考えられる。

しかし文献の主題は、キーワードだけで示されるものではない。今日、多くの検索システムにおいては、個々のキーワードがどのような意味で用いられているかの判別は行わず、表面的な文字列

第2表 適合度判定に用いられる文献の17の特徴

主題表現に関わる情報 (キーワード情報)	
1	タイトル中に必要な概念 (キーワード) 群がすべて含まれているか
2	ディスクリプタに必要な概念 (キーワード) 群がすべて含まれているか
3	抄録中に必要な概念 (キーワード) 群がすべて含まれているか
主題表現に関わる情報 (非キーワード情報)	
4	文献中で用いられているキーワードが望ましい定義で用いられているか
5	概念 (キーワード) 間の関係が情報要求と一致しているか
内容と潜在的に関わる非キーワード情報	
6	文献で扱われる事象や時期、地域などが情報要求と一致しているか
7	文献が扱っている研究手法が情報要求にとって望ましい手法であるか
8	文献に必要なデータ・事例が含まれているか
9	内容に新奇性があるか
10	抄録から文献の内容 (論旨) が明確であると感じられるか
内容とは独立している非キーワード情報	
11	文献の長さが適当であるか
12	抄録の長さが適当であるか
13	出版日は情報要求から見て適当であるか
14	情報源 (掲載誌, 掲載紙) は信頼できるか
15	文献の形態 (雑誌論文, 会議録, コラムなど) が情報要求にあっているか
16	著者及び著者の所属機関に信頼がおけるか
17	著者の専門分野は情報要求にあっているか

の一致によって検索を行っている。例えば distributor というキーワードは、流通業者や販売代理店を示す語であるが、配電器という意味もある。文字列のみの一致では、レコード中に含まれている distributor が、流通業者なのか配電器なのかは考慮されることなく検索されてしまう。文献の内容からみれば、個々のキーワードの意味、キーワード間の関係や役割は、文献の主題を表現

する上で重要であり、主題表現に関わる特徴として分類した。しかし、これらの特徴は、現在の検索システムの多くでは導入されていない。

次に、文献の内容と潜在的に関わる特徴である。ここでは、明確にキーワードでは表現されていないが、文献の内容に密接に関わる特徴が含まれる。これらの特徴は、キーワードとして明示的には現れないが、ユーザの類推によって読みとることが可能な潜在的な特徴といえる。例えば、文献内で用いられている研究手法やアプローチは、必ずしもタイトル中や抄録で触れられるとは限らないが、文献の内容を決める上で重要な特徴である。同じ主張がなされている論文であったとしても、どのようなアプローチでその主張を展開しているかによって、文献の内容は大きく異なると言える。同様に、その文献中にどのようなデータや事例が含まれているかは、キーワードとしてタイトル中や抄録中で触れられていなくても、重要な特徴である。主題は共通であったとしても、実験や調査等を行っている文献であれば、具体的なデータが示されるか否かによって、内容は大きな影響を受けると考えられる。このように、その文献で用いられている研究手法や扱っている地域や事象、データや事例の有無などは、キーワードで表現されていなくともタイトルや抄録から類推することが可能な特徴である。また、文献内容の目新しさを示す新奇性は、キーワードとして表現されることは無く、ユーザが内容全体に基づいて類推することも多いと思われる。なお、ここでは文献の内容についての新奇性であるので、既読文献か否かは含めないものとする。これらの類推によって得られる特徴を、内容と潜在的に関わる特徴とする。

文献には、内容と密接に関わる特徴の他に、内容から独立して存在しうる特徴もある。これらの特徴には、文献の長さや抄録の長さ、論文なのか会議録なのかといった文献の形態、出版年月のようないわゆる書誌情報で示されるデータの的なものと、著者が著名で信頼できるといった著者の特性のようなユーザの経験等から判断されるものがある。また、著者に対する評価を著者の所属機関

を用いて行うことも考えられる。これらの特徴は、文献に記述される内容に関わらずユーザが評価を行うことが可能な特徴と言える。

この 17 の文献特徴には、キーワードで表現される特徴と、キーワードで直接表現されない特徴とが混在している。第 2 表に示した特徴のうち、こうしたキーワード以外の特徴から得られる情報を、非キーワード情報と呼び、キーワードから得られる情報をキーワード情報とする。非キーワード情報は、「内容と潜在的に関わる特徴」、「内容とは独立している特徴」から得られる情報が中心となっており、それぞれ、内容と潜在的に関わる非キーワード情報、内容とは独立している非キーワード情報と呼ぶ。

一方、「主題表現に関わる特徴」のうち、個々の語の意味の判別およびキーワード間の関係や役割が、キーワードと同様に主題を表現する際に使用できるか否かは、システム次第である。多くの商用システムにおいても、技術的、経済的な面での制約が解決されれば、利用可能になることが十分に予想される。しかし、現状の商用データベースシステムでは、この特徴を検索式での主題表現に利用できないため、今回、この二つの文献特徴から得られる情報は、主題表現に関わる非キーワード情報として分類することとした。

次章では上述した文献特徴が、ユーザによる適合度判定にどのように用いられるかを調査から明らかにする。その上で、これら文献特徴から得られる非キーワード情報やキーワード情報の利用のされ方について検討を行う。

III. 適合度判定における文献特徴の利用状況

A. 調査の目的と手順

ここでは、前章で述べた検索結果の出力時点で利用可能な文献特徴のそれぞれが、どの程度適合度判定に影響し、どのような判定基準として用いられているのかを明らかにする。その結果、適合度判定に用いられる文献特徴を、検索システムにどのように導入すべきかを考えることが可能になる。

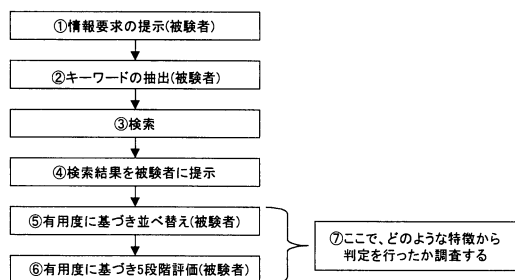
本調査は、被験者による適合度判定と、従来のキーワードに基づいた適合度順検索システムによる判定とを、比較することによって行った。調査に用いた検索システムは、DIALOG から提供されている TARGET を用いた。DIALOG の提供するデータベースは多くのユーザが使用しており、TARGET を用いることによって、特殊な実験用のデータベースではなく、被験者が普段行っている適合度判定に近い環境で調査を行うことができる。そのため、調査に用いるシステムとして TARGET は妥当であると思われる。

今回の調査では、ユーザが自らの情報要求に基づいて行った判定とシステムによる判定との差違を明らかにすることを目的としている。そこで、被験者には以下の 3 点を求めた。

- ①明確な情報要求を持っている
- ②キーワードを用いて情報要求を表現できる
- ③検索結果を自らの情報要求と照らし合わせて評価できる

以上の点を満たす被験者として、大学院修士課程在籍以上の研究者および検索に習熟した情報専門家 7 名を対象とした。内訳は、慶應義塾大学大学院修士課程在籍者が 3 名、同博士課程在籍者が 1 名および社団法人情報科学技術協会分類/シソーラス/Indexing 部会に所属している情報専門家が 3 名である。被験者には、それぞれ自らの情報要求を提供するよう求め、被験者一人につき 1 件ずつ 7 件の情報要求が得られた²⁰⁾。全ての被験者が、情報学分野の知識を持っていたため、図書館・情報学分野の書誌情報データベースである LISA を用いることとした。また、被験者には検索システムの特徴などをあらかじめ説明し、調査を行った。

具体的な調査の手順は第 4 図に示す通りである。まず被験者は、インタビューを通じて、自らの関心に基づいた情報要求を示す。この情報要求は、被験者のおかれている問題状況などを含んでいる。なお、適合する文献がデータベース中に全く存在しない情報要求は、適合度の判定という調



第4図 調査の流れ

査の目的に適さないため、調査から除外した。

続いて、情報要求から検索語として入力するキーワードを選択するよう求めた。まず、自然文で示された情報要求から、重要な概念を示すキーワードを選択し、必要に応じて同義語や関連語を加えるよう求め、これらの語をキーワードとして検索に用いた。この時、インタビューを行い、被験者間でキーワードの選択に差異が生じないように配慮した。また、検索されたレコードに特定のキーワードが必ず含まれることを求める情報要求の場合、指定されたキーワードが必ず含まれるレコードのみを出力する検索機能が TARGET にはある。検索結果として得られるレコードに、ある概念が必ず含まれなければならないと被験者が考える場合、その概念を示すキーワードにこの指定を行うこととした。この指定を行わなかった場合、そのキーワードが検索結果に含まれていないことがありうる。さらに、TARGET には複数のキーワードが同義語の関係にあることを指定する機能もある。同義のキーワードのいずれかが含まれれば十分であると被験者が考える場合、この機能を利用する。

選択されたキーワードによる検索結果のうち、本調査では適合度順に出力された上位 30 件を調査対象レコードとした。この 30 件について、被験者が各自の有用度に基づいて適合度を評価し、1 位から 30 位まで全てのレコードを順位づけた。その際、個々のレコードは全て無作為に並び替えてから被験者に提示した。これはあらかじめシステムの評価である出力順位を知ることによって、被験者の評価が影響を受ける可能性を排除するた

第3表 五段階評価の基準

5	完全に情報要求に一致する
4	周辺的な内容も扱っているが、十分に情報要求を満たす
3	周辺的な内容が中心であるが、十分参考になる
2	周辺的な内容であり、あまり参考にならない
1	情報要求と全く一致しない

めである。

被験者は、有用度に基づく順位の評価とともに、第3表に示された5段階での適合度評価も同時に行った。これは、順位のような個々のレコード間の相対的な評価だけではなく、それぞれのレコードがどの程度有用なのかの絶対的な判定についても明らかにするためである。また、被験者が適合度を評価する際に、出力されたレコードが既知であるか否かに関わらず、被験者にとってそれが有用か否かのみに基づいて判定するように求めた。

そして、ユーザがどのような特徴に着目して 30 件のレコードの順位および 5 段階評価を行ったかを明らかにするために、アンケートおよびインタビューを行った(質問紙は付録に示した)。インタビューはアンケートの回答を確認する目的で、すべての被験者を対象に行った。質問項目はアンケートと同じ内容であり、質問内容は以下の五つである。

- ①5 段階評価の 5 と判定したレコードは、どのような文献特徴を用いて判定したのか
- ②5 段階評価で 4 から 1 とされたレコードは、4 段階のそれぞれについて、どのような文献特徴から上位のレコードよりも劣っていたと判定したのか
- ③5 段階評価が同じであったレコード間の順位を判定する際に、どのような文献特徴を用いたか
- ④被験者が有用度を 1 位から 10 位と判定した 10 件のレコードを、それぞれのレコードごとにどのような文献特徴を用いて判定

したのか

- ⑤TARGET では上位 10 位以内で出力しているが、ユーザは 21 位から 30 位と低く評価したレコードは、どのような文献特徴を用いて判定したのか

質問①および質問②では、検索されたレコードが、第 3 表に示した 5 段階評価のどの段階に該当するかを判定する際に、前章の第 2 表の文献特徴をどのように用いたか回答するように求めた。質問は、5 段階評価のそれぞれについて行われるため、回答項目数は 5 問となる。また、質問③④⑤は、第 2 表の文献特徴を個々のレコードの順位を判定する際にどのように用いたかを質問している。この中で、質問③は順位を評価する際に全体を通じて用いたもの、質問④は適合度が高いと判定したレコードの場合、質問⑤は適合度が低いと判定したレコードの場合についての質問である。質問④は、10 件のレコードのそれぞれについて回答するように求めた。質問③および質問⑤はそれぞれ 1 問の回答項目となるため、全体での回答項目数は情報要求 1 件につき 17 問となる。今回の調査は、情報検索時の適合度判定全体の傾向を明らかにすることを目的としているため、質問①から質問⑤の質問項目を総合して分析を行った。

17 の特徴のそれぞれは、その特徴を持つレコードの適合度が高いと判定する場合と、逆に低いと判定する場合とでは、用いられ方が異なると考えられる。例えば、適合度判定をする上でユーザにとって必要な概念（キーワード）がすべて含まれていた場合、ユーザはこれらの概念（キーワード）に基づいて適合度が高いと判定すると思われる。一方、必要な概念（キーワード）が欠けていて適合度が低いと判定される場合には、その概念（キーワード）が含まれていないという情報が判定に用いられたと考えられる。

また文献の長さやレコードに含まれる抄録の長さのような文献特徴は、文献が長すぎる場合の評価と短すぎる場合の評価とでは、意味合いが異なると思われる。文献の長さを基準として適合度が低いと判定した場合には、文献が長いと低く評

価した場合と、短いと低く評価した場合との二通りが考えられる。そのため、適合度が高いと判定する場合の選択肢は 17 であるが、低いと判定する場合の選択肢は 19 となり、全体では 36 の選択肢を用いて調査を行った。第 4 表は、17 の文献特徴が、適合度を高く判定する根拠となった場合の選択肢、第 5 表は低く判定する根拠となった場合の選択肢を示したものである。調査では、質問に応じて第 4 表および第 5 表、またはその両方の選択肢から選択して回答するよう被験者に求めた。従来のユーザによる適合度判定研究では、適合度判定に用いられる情報が何であるかを明らかにしようとするものは多かったが、その情報が適合度が高いと判定する根拠になるのか、あるいは低いと判定する根拠になっているのかには触れられていなかった。そこで、本調査では個々の文献特徴がどのように適合度判定に用いられているかを明らかにするために、17 の特徴（36 の選択肢）を用いることにした。また、これら 17 の特徴（36 の選択肢）に該当しない文献特徴を用いてレコードを評価した場合、どのような理由でそのレコードに対する順位付けや、適合度の評価を行ったかを被験者が自由に記入できるようにした。

B. 調査結果

1. 全体的傾向

1997 年 4 月から 12 月にかけて、7 種類の情報要求のそれぞれについて、書誌レコードを 30 件出力し調査を行った。対象とした書誌レコードの数は、合計で 210 件である。1 件の情報要求に対して回答項目は 17 問であるので、7 種類の情報要求を用いた調査全体での回答項目数は 119 問となる。なお、5 段階評価の判定に関わる質問で、第 3 表に示した基準と照らし合わせたところ、ある段階にあたるレコードが存在しないとの回答が 2 件あったため、得られた総回答項目数は 117 問となった。今回の調査では複数回答が許されており、これらの質問に対して 359 の文献特徴が回答された。17 問の質問項目ごとの回答件数は、5 段階評価が同じレコード間の順位を判定する際に用いた文献特徴が平均 5.1 件とやや多い回答件数が

第4表 適合度が高いと判定する根拠となる文献特徴

(A)	タイトル中に必要な概念（キーワード）群がすべて含まれる
(B)	ディスクリプタ中に必要な概念（キーワード）群がすべて含まれる
(C)	抄録中に必要な概念（キーワード）群がすべて含まれる
(D)	文献中で用いられているキーワードが望ましい定義で用いられている
(E)	概念（キーワード）間の関係が情報要求と一致する
(F)	文献の対象（事象，地域，時期など）が情報要求と一致する
(G)	文献が扱っている研究手法が情報要求から望ましい手法である
(H)	文献に必要なデータ・事例が含まれている
(I)	内容に新奇性がある
(J)	文献（抄録）の内容（論旨）が明確である
(K)	文献の長さが適当である
(L)	抄録の長さが適当である
(M)	出版日が新しい
(N)	情報源（掲載誌など）に信頼がおける
(O)	文献の形態（雑誌論文・会議録など）に信頼がおける
(P)	著者および著者の所属機関（団体著者を含む）に信頼がおける
(Q)	著者の専門分野が情報要求と一致する

注) 付録の別紙・判定基準の A 群に相当する

得られたが、その他の質問項目は平均約 2.4 件から約 3.9 件の間に収まっており大きなばらつきはなかった。また、被験者間での回答数のばらつきも見られなかった。359 の回答のうち、適合度が高いと判定する根拠として用いられた文献特徴は 172 件、適合度が低いと判定する根拠として用いられた文献特徴は 187 件であった。第 5 図は、調査の全体を通じて 17 の文献特徴が、どの程度用いられていたかを集計したものである。

まず調査の結果から、T. K. Park¹²⁾, Barry¹³⁾ の研究でも指摘された通り、文献の適合度判定に影

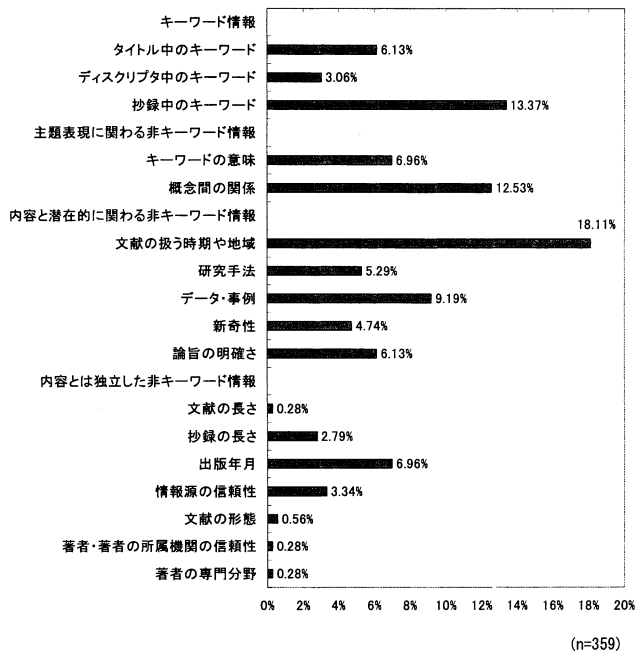
第5表 適合度が低いと判定する根拠となる文献特徴

(ア)	タイトル中に必要な概念（キーワード）群が欠けている
(イ)	ディスクリプタ中に必要な概念（キーワード）群が欠けている
(ウ)	抄録中に必要な概念（キーワード）群が欠けている
(エ)	文献中で用いられているキーワードの定義が望ましい定義ではない
(オ)	概念（キーワード）間の関係が情報要求と異なる
(カ)	文献の対象（事象，地域，時期など）が情報要求と異なる
(キ)	文献が扱っている研究手法が情報要求から望ましい手法ではない
(ク)	文献に必要なデータ・事例が含まれていない
(ケ)	内容に新奇性が認められない
(コ)	文献（抄録）の内容（論旨）がはっきりしない
(サ)	文献の長さが長すぎる
(シ)	文献の長さが短すぎる
(ス)	抄録の長さが長すぎる
(セ)	抄録の長さが短すぎる
(ソ)	出版日が古い
(タ)	情報源（掲載誌）に信頼がおけない
(チ)	文献の形態（雑誌論文，会議録など）が信頼のおける形態ではない
(ツ)	著者および著者の所属機関（団体著者を含む）に信頼がおけない
(テ)	著者の専門分野が情報要求とは異なる

注) 付録の別紙・判定基準の B 群に相当する

響を及ぼす文献の特徴は多様であることが裏付けられた。今回、調査に用いた 17 の文献特徴は、使用頻度に差はあるものの全てが何らかの形で適合度判定に用いられていた。また、あらかじめ用意した選択肢以外の特徴から適合度を判定した場合には、その文献特徴を記述するよう求めた。今回の調査では 17 の文献特徴以外に、「なんとなく」といった明確に表現できない回答が 4 件得られ

情報検索の適合度判定における非キーワード情報の意義



第5図 適合度判定に影響を及ぼした文献特徴

た。このような適合度判定の根拠をどのように扱うかについては、今後の検討が必要である。しかし、本論文では文献特徴の適合度判定に与える影響を考察することが目的であるので、不明確なこの種の回答は調査結果の対象から外すこととし、359件の回答には含めていない。今回の調査では、これ以外に17の文献特徴に含まれない回答は得られなかった。

調査の全体を通じて、適合度判定に用いられた文献特徴の中で、タイトル、ディスクリプタ、抄録中のキーワードといったキーワード情報は、約23%にとどまった。このキーワード情報に、キーワードでは表現されない個々の語の意味や、概念間の関係や役割といった主題表現に関わる非キーワード情報を加えた主題表現に関わる特徴全体は、約42%であった。それに対して、第2表で内容と潜在的に関わる非キーワード情報として分類される文献特徴は、合計で約43%であった。主題表現に関わる文献特徴は、適合度判定において約4割を占め、重要であると言えるが、一方で主題のみによって適合度判定が行われているわけ

はないことも明らかになった。さらに、文献の内容とは独立している非キーワード情報も約14%の割合を示しており、鄒等が指摘した非主題的の特徴も無視し得ない影響を及ぼしていると言える。このことから、ユーザは検索時に用いたキーワード以外の特徴も用いて、全体的に適合度判定を行っていることが読みとれる。

全体を通じて利用度が高かった文献特徴として、文献で扱われている事象や時期、地域などが情報要求に合致しているかがあげられ、全体で18.11%の回答が得られた。これは、文献で扱われている時期、地域、事象についての情報が、タイトルや抄録の中で明確に示されていなくても、抄録などからの類推によって間接的に得られ、その情報が適合度を判定する根拠となった事例が多かったことを示している。文献で扱われている時期や地域は、適合度を高いと判定する場合と低いと判定する場合との両方で多く用いられていた文献特徴である。

次に、抄録中の概念（キーワード）の利用度が高く、続いて概念間の関係が情報要求と合致して

いることも重要視されており、13.37% および12.53%の回答が得られた。一方、タイトルやディスクリプタはそれほど利用されることは無かった。概念間の関係が情報要求に合致しているかどうかが高い使用頻度を示したことは、概念間の関係や役割を用いることの有効性を示唆していると考えられる。

研究手法やデータ・事例の有無、論旨の明確さや文献の新奇性といった内容と潜在的に関わる情報は、いずれも利用度が高かった。適合度を判定する上で、文献中にどのようなデータや事例が含まれているかの情報は、タイトル中のキーワード以上に利用されていたことになる。

また、内容とは独立している非キーワード情報では、出版年月の利用度が高かった。次いで、情報源の信頼性や抄録の長さといった文献特徴が多く用いられていた。

2. 適合度を高く判定する根拠となった文献特徴

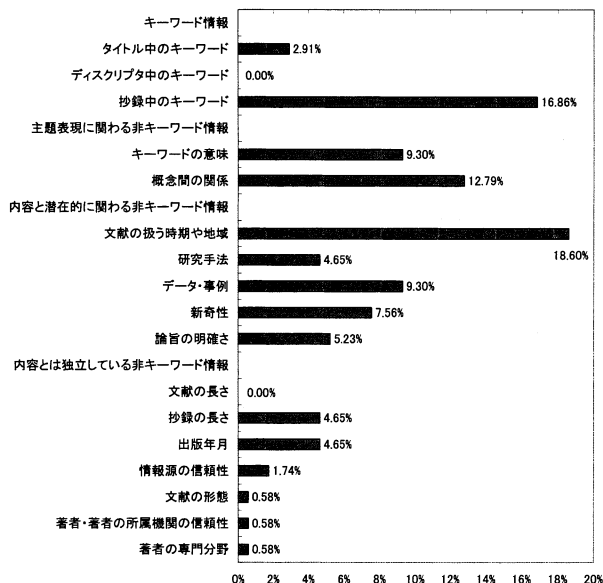
今回の調査では、それぞれの文献特徴を、適合度を高く判定する根拠として用いたのか、低く判定する根拠として用いたのかを区別して質問して

いる。適合度を高く判定する根拠として用いられた文献特徴を集計したものが、第6図である。

適合度が高いと判定する根拠として、文献の扱う時期や地域、抄録中のキーワードが用いられることが多かった。しかし、文献の扱う時期や地域は、後述する第7図で示すように、適合度が低いと判定する場合にも多く用いられており、全体を通じて多く用いられる文献特徴であったと言える。適合度を高く判定する根拠として特徴的に用いられる傾向が強かった文献特徴は、抄録中のキーワード、キーワードの意味、新奇性、抄録の長さの四つである。抄録中のキーワードは、全体の中でも利用度が高く、ユーザの多くは抄録を適合度を高いと判定する際に判定材料としていた。

キーワード情報では、抄録のみが多く用いられていた。タイトル中やディスクリプタ中のキーワードは、適合度が高いと判定する根拠としてはほとんど用いられていなかった。主題表現に関わる文献特徴の中で、適合度が高いと判定する場合に用いられる傾向が強かった特徴は、個々のキーワードの意味である。

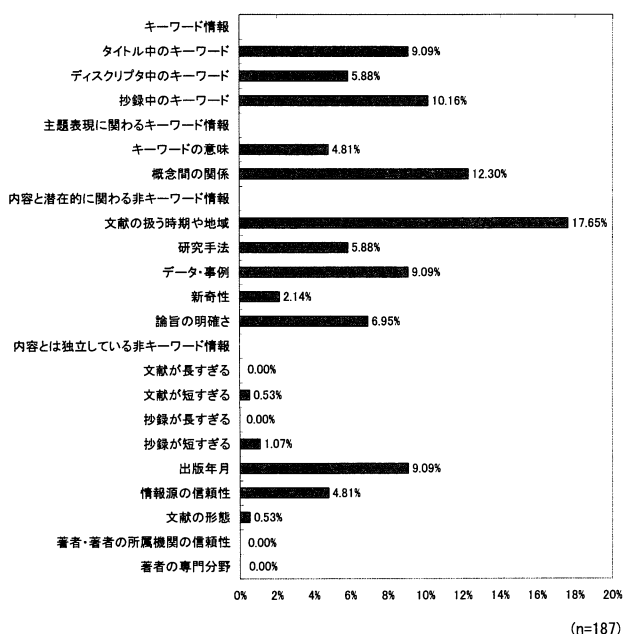
内容と潜在的に関わる文献特徴の中では、新奇



(n=172)

第6図 適合度が高いと判定する根拠となった文献特徴

情報検索の適合度判定における非キーワード情報の意義



第7図 適合度が低いと判定する根拠となった文献特徴

性が適合度を高く判定する根拠として用いられることが多かった。ユーザは、文献内容に新たな知見が含まれていることが推測できれば、適合度が高いと判定する根拠としていたと思われる。また、内容とは独立している文献特徴では、抄録の長さが適切であることが、多く用いられていた。このことは、被験者の多くが、抄録中のキーワードを用いていたことから、ある程度の長さを持つレコードを高く評価したためと考えられる。

3. 適合度を低く判定する要因となった文献特徴

適合度を低く判定する根拠として用いられた特徴を、第7図に示す。文献で扱われている地域や時期は、適合度を高いと判定する根拠としても多く用いられていたが、低いと判定する根拠としても多く用いられていた。文献で扱われている地域や時期が情報要求と合致していれば、その文献の適合度を高く評価する要因となることは予想されたが、逆に合致しないため低く評価するという事例も多数見られた。望んでいた主題について触れていたものの、抄録を読んだ結果、要求とは違う内容であると類推されたケースである。情報要求

に照らし合わせて特殊な事例として低く評価された具体的な例は、「図書館サービスの設計について知りたい」という情報要求に対して、タスマニアの事例や、病院図書館を対象とした文献、サービス対象を障害者に限定した文献等である。

適合度が高いと判定する場合よりも、適合度が低いと判定する場合に多く用いられた文献特徴として、タイトル中のキーワード、ディスクリプタ中のキーワード、出版年月、情報源の信頼性の四つがあげられる。タイトル中のキーワードは、適合度が高いと判定するよりも低いと判定する根拠として用いられることが明らかに多かった。これは、現状のキーワード検索では、あるキーワードが抄録中に一語でも用いられていれば、そのレコードが検索されてしまう。そのため、レコード中にキーワードが含まれていたものの、必要な概念について中心的に記述されていないレコードも検索されてしまう。このようなレコードを、被験者がタイトルを用いて除去したため、タイトル中のキーワードが適合度が低いと判定する根拠として用いられたと考えられる。ディスクリプタ中のキーワードも、本来望んでいたレコードとは異なる

るディスクリプタが付与されていた場合に、そのレコードを適合していないと判定する根拠として用いられていた。タイトルおよびディスクリプタは、それが情報要求に合致しているからといって適合度が高いと評価する根拠にはならないが、情報要求に合致しない場合には適合度が低いと評価する根拠にはなると言える。タイトルやディスクリプタフィールドは文献の内容を端的に表現している部分であるため、これらのフィールドが明らかに情報要求とは違う場合には、適合していないと判断する大きな要因になると考えられる。しかし、ディスクリプタが全体的にあまり用いられなかったことに関しては、LISA というデータベースの性格も影響を及ぼしている可能性がある²¹⁾。

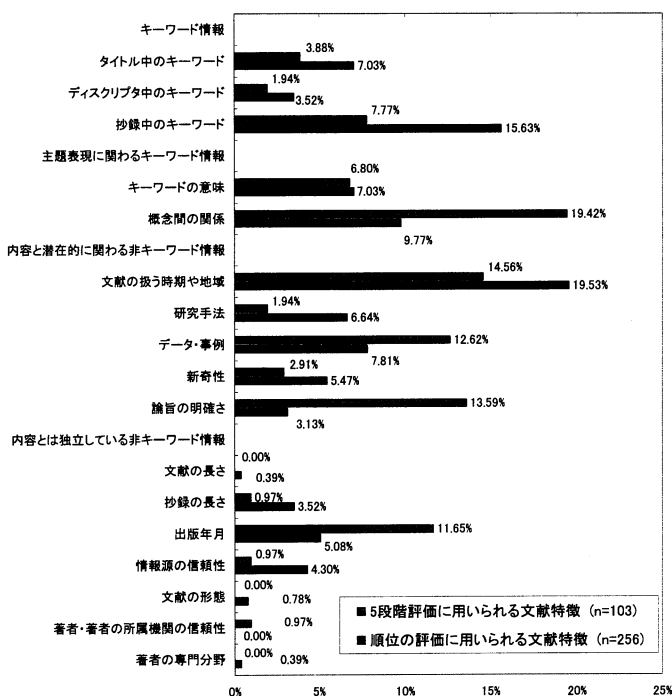
また出版年月、情報源の信頼性の二つの文献特徴は、適合度が高いと判定する根拠として全く用いられないわけではないが、主に適合度が低いとする根拠として用いられることが多かった。これらの文献特徴は、いわゆるフィルターとして情報要求からはずれている可能性の高いレコードや内

容的に不十分なレコードを、除去するための判断基準として利用されていると考えられる。

4. その他の結果

本調査では、適合度を5段階で評価した場合と、レコード間の相対的な順位を評価した場合とのそれぞれについて、根拠として用いた文献特徴の差異の調査も行った。5段階評価は、文献に適合度に応じていくつかのグループに分ける適合度判定であり、順位の評価は全ての文献に差をつける詳細な適合度判定であると言える。5段階評価については質問①②から、順位の評価については質問③④⑤で得られた回答から、それぞれの適合度判定で用いられた文献特徴の比率を分析した。第8図は、5段階評価および順位の評価に用いられた文献特徴の比率を示している。

5段階評価において用いられることが多かったのは、概念間の関係や役割(19.42%)、文献で扱われている時期や地域などの適切さ(14.56%)、論旨の明確さ(13.59%)、データや事例の有無



第8図 5段階評価と順位の評価で用いられた文献特徴

情報検索の適合度判定における非キーワード情報の意義

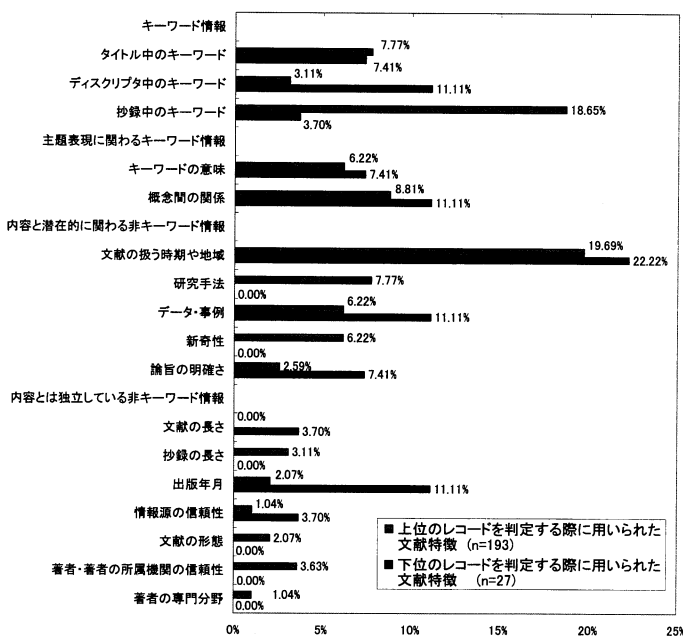
(12.62%), 出版年月(11.65%)である。それに対して、順位を評価する際に用いられることが多かったのは、文献で扱われている時期や地域などの適切さ(19.53%), 抄録中のキーワード(15.63%), 概念間の関係や役割(9.77%)などがあげられる。文献で扱われている時期や地域などの適切さや概念間の関係や役割は、全体を通じても多く使用されており、全般的に重要視されていたと思われる。

5段階評価と順位の評価とで、適合度判定の際に用いられる比率に大きな差がでた文献特徴もあった。5段階評価での適合度判定により多く用いられていた文献特徴は、論旨の明確さ、概念間の関係と出版年月である。論旨の明確さは、抄録などから類推して得られる情報であり、明確に判定できる文献特徴ではない。そのため、個々のレコードの差異を表す順位の評価よりも、5段階評価で用いられることが多かったと思われる。出版年月も、相対的な順位の評価よりも5段階評価で用いられることが多かった。このことは、出版年月が情報要求にあわない場合、適合度が相対的に

ではなく、絶対的に低いと評価されるためと思われる。

順位を判定する際に特徴的に用いられていた文献特徴は、抄録中のキーワード、文献で扱われている時期や地域、研究手法である。順位の評価は、5段階評価以上に個々のレコードの有用度について細かく判定する必要がある。今回の調査では、このような詳細な適合度判定を可能にするような文献特徴は、抄録から得られることが多いことを示している。しかし、抄録から得られる情報でも、論旨の明確さのような具体性に欠けるとと思われる文献特徴は、順位の評価にはあまり用いられていなかった。

また、質問④と質問⑤から、ユーザによる適合度が高いレコードで多く用いられた文献特徴と、逆に適合度が低いレコードで用いられた文献特徴について明らかにした。ここで、適合度が高いレコードとは、ユーザの適合度判定で上位とされた10件のレコードである。一方、適合度が低いと判定されたレコードは、TARGETでは10位以内と高く判定されたがユーザが適合度を低く判定し



第9図 上位のレコードと下位のレコードを判定する際に用いられた文献特徴

たレコードを対象としている。これらのレコードは、レコード中にある程度のキーワードが含まれており、かつ被験者がキーワード情報および非キーワード情報の両面から判断し、適合度が低いと判定したレコードである。本研究では、キーワード情報と非キーワード情報が適合度判定に及ぼす影響を明らかにすることを目的としている。そのため、TARGETによる順位が低く、かつユーザによる適合度も低く判定されたレコードは、キーワードがほとんど含まれずキーワード情報が得られないと予想されたため、調査の目的にはあわないと考え、分析の対象とはしなかった。第9図は、ユーザによる適合度が高いレコードで多く用いられた文献特徴と、逆に適合度が低いレコードで用いられた文献特徴の比率を示している。

適合度が高いレコードと低いレコードについても、それぞれでの用いられ方に大きな差異がある文献特徴があった。適合度が高いレコードは、抄録中のキーワードや研究手法から適合度が判定されていることが多い。一方、適合度が低いレコードでは、出版年月、ディスクリプタ中のキーワード、データや事例の有無が根拠として多く用いられている傾向が見られた。出版年月とディスクリプタ中のキーワードは、適合度を低く判定する根拠としても多く用いられており、実際に適合度が低いと判定されたレコードで多く用いられていることから、非適合文献を除去する際に用いられていたと考えられる。

IV. 非キーワード情報の意義と影響

A. 適合度判定におけるキーワード情報と非キーワード情報

ユーザによる適合度判定の際に、キーワード情報と非キーワード情報がどのように用いられていたかは、調査の結果から以下のようにまとめることができる。

- 1) キーワードでは表現できない非キーワード情報は、キーワード情報と共に適合度判定に大きな影響を与える

- 2) 適合度判定に用いられる情報には、検索結果を評価する時点でフィルターとして使われるものと、使われないものがある

文献内容と潜在的に関わる非キーワード情報や、内容とは独立している非キーワード情報は、適合度判定においてキーワード情報と同様に重要な役割を果たしていることが明らかである。これは、H. Park¹⁴⁾ や Janes²²⁾ の指摘と同様に、ユーザにとって役に立つか否かという観点から判断を行った場合、単純に主題が一致しても必ずしもその文献が有用とはならないことを示す結果と言える。ユーザが情報検索システムを利用する目的は、主題が合致する文献を入手することではなく、ユーザがおかれた問題状況に対処できる情報を入手するためと考えられる。そのため、ユーザの適合度判定は単なる主題の類似性だけではなく、文献の様々な特徴から適合度判定を行っている。つまり、文献の適合度判定は主題に関わるキーワード情報だけではなく、非キーワード情報も用いて行われているのである。

次に、キーワード情報と非キーワード情報がどのように利用されているかについて検証する。鄒等の研究では、文献の主題とは直接的に関わらない非主題的特徴が、検索時にノイズを除去するフィルターとして利用可能であることが指摘されている¹⁸⁾。従来の情報検索システムにおいても、出版年月や情報源などを利用した検索結果の絞り込みが行われてきた。そのため、書誌情報を適合度判定に利用する場合、多くは非適合文献を取り除くフィルターとして利用することを想定してきた。しかし今回の調査では、非キーワード情報の中でも、フィルターとして利用される情報と利用されない情報とがあることが明らかになった。

従来、フィルターとして利用されると考えられていた情報源の信頼性、文献の長さ、出版年月は、今回の調査でも適合度が高い根拠というよりは低い場合の根拠として用いられることが多く、フィルターとして利用されていたと言えよう。これらの文献特徴をフィルターとして利用することの妥当性は、今回の調査結果から裏付けられたと言え

よう。

また、今回の調査では主題表現であるタイトルやディスクリプタ中のキーワードが、文献の適合度を低くする要因として用いられていた事例が多かったことを指摘しておく必要がある。Fidel²³⁾は、検索者のタイプを出版年や検索対象フィールドを限定するなどの手法で件数を絞り込む操作派と、概念を組み合わせることでよりの確な検索結果を得ようとする概念派とに分類した。被験者は、タイトル、ディスクリプタ、抄録の全てを対象としたキーワード検索の結果が、実際に情報要求に合致するか否かを判定する際に、タイトルフィールドやディスクリプタフィールドに何が記述されているかを確認したと考えられる。つまりフィールドを限定して絞り込むといった、Fidelの操作派的な行動をとったと思われる。適合度順検索ではキーワードの出現頻度に基づいて検索が行われるので、抄録におけるキーワードが検索結果に与える影響は大きく、抄録中に頻出していけば上位で出力されてしまう。そこでユーザは、タイトルやディスクリプタ中の概念の適切さを用いて、キーワードが頻出しているが適合していない文献を除去していると思われる。とくにディスクリプタについては、ほとんどが適合度を低く判定する根拠として用いられていた。これまでフィルターとして利用される文献特徴として前述の出版年月などが考えられていたが、タイトルやディスクリプタ中のキーワードも、非適合文献を除去する根拠として用いられうるといえよう。

T. K. Park¹²⁾は、「タイトル中のキーワード」にはその文献が扱っている分野を示す機能があると指摘したが、今回の調査でも同様に、タイトルはユーザが必要とする内容の文献であるかどうかを判別するフィルターとして用いられていた傾向が確認できた。タイトルやディスクリプタは、文献が扱う主題や分野を確認する情報であると言える。今後、適合度判定におけるキーワード情報の役割について、より詳細に検討していく必要があると思われる。

一方、文献内容と潜在的に関わる非キーワード情報には、文献の適合度を高く評価する要因とな

るものが含まれていた。これらの文献特徴は、適合度判定において重要な情報であると考えられる。特に文献の新奇性や文献の扱う時期や地域の適切さは、適合度が高いと判定される根拠として用いられる傾向が強かった。従来、文献の主題は、まずタイトルやディスクリプタから判断し、その後抄録を用いてその判断の適否を判定すると考えられてきた。今回の調査では、被験者は、内容と関わる非キーワード情報をはじめとして多くの情報を抄録から得て、それを元にユーザは適合度を判断していた。適合度判定において、抄録は重要な役割を果たしていたと言える。ユーザによる適合度判定の中心的役割を果たしているのは、抄録から得られるキーワード情報や非キーワード情報であった。また、出版年月や著者の信頼性などその他の情報は、抄録から得られた情報を補完していると考えられる。

従来の検索システムは、ユーザが適合度判定に用いている情報の一部に過ぎないキーワードに基づいて、検索を行ってきた。これが、システムとユーザとの間にギャップが発生する要因の一つである。ユーザが適合度を判定する際には、キーワード情報のみではなく非キーワード情報も重要な役割を果たしている。検索システムがよりユーザの望んでいる文献を提供するためには、ユーザによる適合度判定の中で非キーワード情報に基づいて行われている部分を、何らかの形で検索システムに反映させることが必要であろう。

B. 非キーワード情報の検索システムへの応用とその課題

ユーザが、適合度判定において非キーワード情報も用いていることを考えれば、システムがユーザの望む文献を的確に検索するためには、こうした非キーワード情報を検索に応用する意義は高い。今回の調査で明らかになったように、ユーザによる適合度判定に非キーワード情報が大きな影響を及ぼしていることは、ユーザが蓄積レコード中に明確に示されたキーワード以外の情報から、適合度判定を行っていることを意味している。このことは同時に、検索式には記述できなかった情

報を適合度判定に用いているとも言える。

今回の調査では、文献の扱う時期や地域が適切か否かや、研究手法、データ・事例の有無など、キーワードでは示されていない文献の内容を、ユーザが適合度判定に用いた事例が数多く見られた。ユーザは、自分のおかれた状況に対して、あるデータや事例の有無が必要であるならば、抄録等からそれらの有無を類推していたと思われる。このような情報が蓄積レコード内で明確に示され、情報要求にも含まれていれば、非キーワード情報を用いた検索を行うことが可能になる。そのためには、蓄積レコードと検索式との双方を、従来のキーワード情報のみではなく非キーワード情報も含めて、より詳細に記述する必要がある。

例えば、非キーワード情報について適切な索引語を付与することにより、蓄積レコードをより詳細にすることが可能となる。また、全文を収録することによっても、蓄積レコードの詳細な記述は可能であろう。今回の調査結果から、潜在的に内容と関わる非キーワード情報のうち、特に文献で扱われている時期や地域などの適切さ、文献にどのような事例・データが含まれるか、文献で扱っている研究手法の三つが、研究論文を中心的な収録対象とするデータベースで検索キーとして使用されることは、十分に予測できる。このように必要と思われる情報について蓄積レコード中に記述することは、いくつかのデータベースで試みられている。例えば、MEDLINEでは、検索結果を、ヒトが主題になっているレコードに限定する機能がある²⁴⁾。これは、通常のタイトルや抄録中では明確に表現されていないこともある文献の内容について、詳細に記述した例である。しかし、人手による索引語の付与が必要になり費用がかかるなどの理由で一部のデータベースでしか導入されていない。また、データ・事例の有無は、図表のキャプション等をデータベース内に蓄積することより、ある程度把握することができると思われる。実際、文献に含まれる図表のリスト、キャプション等のリストをユーザに示し、検索の際に活用しようとする試みがなされている²⁵⁾。これらは、キーワードで明確に表現されるとは限らない

情報を提供することで、適合度判定をより行いやすくしようというものである。

また、検索式を詳細に記述することも求められる。ユーザは、非キーワード情報を適合度判定に用いており、文献を取捨選択する基準としていると言えよう。そのため、従来のキーワードで示される主題のみではなく、これら非キーワード情報も含めて詳細な記述が行えるならば、より情報要求に合致した検索が行えると思われる。

ある非キーワード情報が適合度を高く判定する根拠として用いられるのか、あるいは低く判定する根拠として用いられるのかといった非キーワード情報の用いられ方は、個々のユーザや検索の目的などに依って異なると考えられる。例えば、新奇性は当該分野に関してどれだけ知識を持っているかによってその判断は異なるであろうし、情報源の信頼性、著者の信頼性も主題知識や過去の経験、つまりユーザの特徴に依存するものと考えられる。同様に、論旨の明確さも個々のユーザの好みによって影響を受けると考えられる。このように、適合度判定に中心的に用いられていた非キーワード情報の多くは、ユーザごとに用い方が異なると予想される。また、同一のユーザであったとしても、検索の目的が異なれば同じ非キーワード情報に対しても異なった判断がなされると思われる。このように、非キーワード情報の用いられ方が個々のユーザや情報要求によって異なることは、ユーザの特徴や、キーワードで示されない情報要求の特徴が現れているためと言えよう。このような、非キーワード情報の用いられ方のユーザ間や情報要求間での異なりは、従来の検索式では記述できなかったのである。

このような非キーワード情報を検索式の一部として検索に利用するためには、何らかの形でこの情報をシステムに入力する必要がある。しかし、非キーワード情報までを含めて詳細に検索式を記述することは、ユーザにとって大きな負担となる。特にエンドユーザを想定した検索システムでは、ユーザの負担が大きくなることが求められている。そのため、検索ごとに詳細な検索式をユーザに記述させることは、現実的ではないと思

われる。

これらの非キーワード情報の中には、同じユーザであれば一定期間判断の基準が変わらず、ユーザに固有の特徴が見られると予想されるものもある。例えば、ある論文をまとめるという状況におかれたユーザであれば、文献の扱う時期や地域の適切さや、文献中で用いられている研究手法についての判断は、ある程度の期間は変化しないことが予想される。このような非キーワード情報の用いられ方は、個々のユーザに固有のものであるから、ユーザモデルとして検索システム内にあらかじめ構築しておくことが考えられる。このユーザモデルに応じて検索結果を重み付けすることにより、ユーザの負担を過大にすることなく検索システムに非キーワード情報を取り込むことができると思われる。

なお今回の調査では、抄録は長いものほど適合度を高く評価され、出版年月は古いものの適合度を低く判定する場合が多かった。このように、個々のユーザ間で用いられ方に差異が無く、ある一定の範囲内に収まる情報は、前述のユーザモデルを用いるまでもなく、重み付けに反映させやすいと考えられる。特に、ユーザによる適合度判定で出版年月は多く用いられており、重要であると思われる。実際に、新しい情報が望まれると考えられている WWW 検索においては、その情報が更新された日付、つまり更新日時といった情報によって重み付けをすることが既に試みられている²⁶⁾。これは、非キーワード情報を検索に応用した例と言える。このような出版年月や文献の長さといった非キーワード情報は、多くのデータベースで蓄積レコード中に含まれており、検索システムに応用しやすい非キーワード情報と言える。

検索式と蓄積レコードにおける詳細な記述とそれらを処理できる検索システムは、ユーザが望む結果を出力するという情報検索の目的を達成するためには重要である。非キーワード情報の中には、容易に応用が可能なものも含まれている。しかし、非キーワード情報の用いられ方がユーザごとに異なる個別性の問題を多くの非キーワード情報は抱えている。この個別性の問題を解決し

つ、検索時のユーザの負担を増加させないためには、ユーザモデルを用いた検索システムであることが求められるであろう。

V. 今後の課題

本研究では、ユーザとシステムとの間に適合文献を判定する過程に大きなギャップがあることを指摘するとともに、ユーザがどのように文献の適合度を判定しているかを明らかにした。その上で、現在のキーワードを中心とした検索システムの性能は十分とは言えず、キーワードでは表現されない非キーワード情報を導入することが重要であると指摘した。

しかし、被験者の数や課題数に関して十分とは言えない面もあり、17の文献特徴がユーザの適合度判定に影響を及ぼしていることは明らかであるが、その詳細については、今回の調査では対象としなかった既読文献の扱いも含めて、より大規模な調査を行う必要がある。また、適合度判定の根拠として「なんとなく」といった記述できない部分があることも確認されており、ユーザの情報利用についての質的な研究も重要であると思われる。今後は、ユーザと検索システムとの相互作用について、量的および質的な側面からの研究が十分になされる必要がある。

非キーワード情報を検索システムに応用するためには、ユーザごとに異なると思われる非キーワード情報の用いられ方を反映させ、さらに検索時のユーザの負担を軽減するために、ユーザモデルを用いた検索システムが必要となる。ユーザがどのような点を重視して適合度判定を行うかや、ユーザが情報要求を持っている背景等をモデル化することによって、ユーザの個性を検索に反映させようという試みは過去にもいくつか見られる。例えば Rich²⁷⁾ の GRUNDY におけるユーザの定型を利用し検索を行おうというアプローチは、ユーザモデルを検索に利用しようとする試みの先駆けと言える。GRUNDY は、小説を対象とした検索システムであるが、いわゆる主題だけではなくユーザの好みを検索に取り込もうという試みであった。ユーザの好みは、キーワードで明確

に表現できるものではなく、このような情報を検索に応用するためにユーザモデルを用いていた。しかし、GRUNDYで扱っている定型だけではユーザが情報要求に至る背景やその問題の状況を十分に表すことはできず、限界があったと指摘されており¹⁾、必ずしも十分なものではなかった。このことは、ユーザがどのような文献を望んでいるかをより詳細にモデル化することが必要であることを示している。

また、近年注目を集めているのが認知的アプローチである²⁸⁾。認知的アプローチでは、ユーザの行動を、表面的な行動ではなく、心理的あるいは認知的な内面までを考慮する必要があると指摘している。これは、情報要求や適合度判定が、ユーザの認知状態や知識状態の影響を受けるためである。本研究で取り上げた非キーワード情報の多くは、個々のユーザの認知状態や知識状態の影響を受けていると考えられる。ユーザの知識状態や情報要求の背景を情報検索に反映させようという試みは、キーワードで表現不可能な非キーワード情報をシステムが利用できるようにするためには重要である。

検索システムがユーザの情報要求をより把握し、ユーザが望む結果を出力するためには、非キーワード情報を検索システムに応用することが重要である。非キーワード情報の中には、出版年月のように既に実用化が試みられているものもある。しかし、本研究でも明らかのように、中心的に用いられていた非キーワード情報の多くは、ユーザごとに個別性を持つと思われる情報である。この個別性を検索システムに反映させるためには、ユーザの認知状態まで含めたユーザモデルを用いることが必要であると思われる。今後、ユーザがどのような基準によって情報の取捨選択を行っているかをより緻密に検証していくとともに、このような非キーワード情報を取り込んだ検索システムの有効性を検証していくことが重要であろう。

注・引用文献

1) Ellis, D. 情報検索論：認知的アプローチへの展

- 望. 細野公男監訳. 東京, 丸善, 1994. 180 p.
- 2) Stibic, V. "Influence of unlimited ranking on practical online search strategy". *Online Review*. Vol. 4, No. 3, p. 273-279 (1980)
- 3) Salton, G. ed. SMART 情報検索システム：データベースへのアプローチ. 神保健二訳. 東京, 企画センター, 1974. 647 p.
- 4) Robertson, S. E. "The probability ranking principle in IR". *Journal of Documentation*. Vol. 33, No. 4, p. 294-304 (1977)
- 5) キーワードの逆文献頻度とは、データベース全体でのキーワードの出現頻度である。キーワードの出現頻度を用いた適合度順検索システムでは、キーワードのレコード内での出現頻度から、どのような概念が多く含まれているかを判定している。この時、そのキーワードのデータベース全体を通じての出現頻度、つまり逆文献頻度によってそのキーワードの重要度を修正することが行われる。
- 6) <http://www.goo.ne.jp/help/faq2.html> (access: 1999-8-25)
- 7) http://www.infoseek.co.jp/Help?pg=help_s_kekka.html (access: 1999-8-25)
- 8) Dialog. "TARGET: a new way to search on DIALOG". *Dialog Chronolog*. November/December, p. 319-321 (1993)
- 9) Tenopir, C., Chan, P. "TARGET & FREE-STYLE". *Online*. Vol. 13, No. 2, p. 31-47 (1994)
- 10) Keen, E. M. "How does Dialog's target work?" *Online and CDROM Review*. Vol. 18, No. 5, p. 285-288 (1994)
- 11) 相良佳弘. "語の出現頻度情報に基づく適合度順検索の問題点". 情報知識学会第6回(1998年度)研究報告会講演論文集. p. 23-28 (1998)
- 12) Park, T. K. "The nature of relevance in information retrieval: an empirical study". *Library Quarterly*. Vol. 63, No. 3, p. 318-351 (1993)
- 13) Barry, C. L. "User-defined relevance criteria: an exploratory study". *Journal of the American Society for Information Science*. Vol. 45, No. 3, p. 149-159 (1994)
- 14) Park, H. "Relevance of science information: origins and dimensions of relevance and their implications to information retrieval". *Information and Processing Management*. Vol. 33, No. 3, p. 339-352 (1997)
- 15) 細野公男. "索引作業の今日の特徴と課題". *情報の科学と技術*. Vol. 46, No. 11, p. 606-612 (1996)
- 16) Taylor, R. S. "Question negotiation and information seeking in libraries". *College and Research Libraries*. Vol. 29, No. 3, p. 178-194 (1968)
- 17) Borlund, P.; Ingwersen, P. "The development

- of a method for the evaluation of interactive information retrieval systems". *Journal of Documentation*. Vol. 53, No. 3, p. 225-250 (1997)
- 18) 鄒永利, 相良佳弘. “文献の非主題的特徴と情報検索におけるその意義”. 情報知識学会第7回(1999年度)研究報告会講演論文集. p. 3-8 (1999)
- 19) 一般に図書の本誌データには, 出版された時期(出版日)として出版年が含まれることが多い. 一方, 学術論文等を収録対象としたデータベースの中には, 出版年月が含まれるものもある. そのため本論文では, 本誌データベースに関する記述の場合, 出版日に関して出版年月と表記することにした.
- 20) 今回の調査対象とした検索事例は, 第1表に示された事例の情報要求1から3および情報要求7から10の合計7件である.
- 21) LISAは, 「DE-」フィールドすなわちディスクリプタフィールドを持っているが, これはMEDLINEに対するMeSHのような厳密な統制語彙に基づいて付与されているわけではない.
- 22) Janes, J. W. “Other people judgment: a comparison of user's and others' judgments of document relevance, topicality, and utility”. *Journal of the American Society for Information Science*. Vol. 45, No. 3, p. 160-171 (1994)
- 23) Fidel, R. “Online searching styles: a case-study-based model of searching behavior”. *Journal of the American Society for Information Science*. Vol. 35, No. 4, p. 211-221 (1984)
- 24) ヒトが主題のものに限定する機能の詳細については, ジーサーチ社が発行している『Dialogブルーシート日本語版』に記述されている. なお, MEDLINEの英文版ブルーシートは, 以下のURLで最新のものを参照する事ができる. <http://library.dialog.com/bluesheets/html/b0154.html> (access: 2000-7-22)
- 25) Bishop, A. P. “Document structure and digital libraries: how researchers mobilize information in journal articles”. *Information Processing and Management*. Vol. 33, No. 3, p. 255-279 (1999)
- 26) 福島俊一; 松田勝志; 高野元. “Webページの重要度ファクタに関する一考察”. 情報知識学会第7回(1999年度)研究報告会講演論文集. p. 77-80 (1999)
- 27) Rich, E. A. “User modelling via stereotype”. *Cognitive Science*. Vol. 3, p. 329-54 (1979)
- 28) Ingwersen, P. 情報検索研究: 認知的アプローチ. 藤原鎮男監訳. 東京, トップラン, 1995, 378 p.