情報検索行動における検索の終了

相良佳弘[†]

When will we stop search sessions?

Yoshihiro Sagara[††]

<Abstract>
In the conventional information retrieval research, evaluation of information retrieval system has been performed mainly on the basis of precision and recall of final research result. However, modern retrieval systems are designed to provide results through the interaction with users. The decision by a user to stop his search is therefore influenced by the interactions done up to the point. The survey by questionnaires was conducted to clarify the factors which affect the decision of stopping a search. In the case when users stop search in unsatisfied way, many non-essential reasons are pointed out for giving up further search. The implications of the stop of a search in user behavior study are emphasized.

1.研究の背景

これまで情報検索において、検索者の行動を論じる際には、検索戦略や検索戦術という概念が用いられてきた。検索戦略・検索戦術とは、効率的に適切な文献を検索するためになされる検索の全体的計画や手法、テクニックのことである[(1)]とされる。Bates[(2)(3)]をはじめ、これまで情報要求の分析、検索式の作成および修正など検索の前段階や検索を実行する際の行動については、これまでに様々な分析がなされてきた。しかし、最終的な検索結果を得て検索を終了を決定する際の行動についての研究は、まだ十分になされているとは言えない。

一方で情報検索システムの性能評価を行う際には、得られた検索結果に対して精度・再現率を算出することが広く行われている。しかし、今日の情報検索システムのほとんどは、検索者との相互作用を行い検索結果を出力するシステムとなっている。そのため、どの時点でシステムとの相互作用をやめ、検索を終了するかは検索性能に大きな影響を及ぼすと考えられる。本研究では、情報検索行動における検索セッション終了の意義について考察する。また、検索者がどのような要因に基づいて検索を終了するかは、検索性能の面および実際の検索者の検索行動を探る上で意義があると思われる。そこで、アンケート調査から、ユーザがどのような要因から検索を終了しているのか明らかにした。そのうえで、従

[†] 慶應義塾大学大学院文学研究科図書館・情報学専攻博士課程
[††] Graduate school of Library and Information Science, Keio University.

来の検索戦略に関する議論について、検索の終了の観点から考察を加えることとする。

なお本研究では、ある情報要求に関する検索を開始してから終了にいたるまでを検索セッションと呼ぶことにする。この検索セッションの終了には、マニュアル等を参照するための一時的な中断は含めず、あくまでもその情報要求に関する検索の終了を決断した場合、検索セッションが終了したととらえることとした。

2.情報検索における検索セッション終了の意義

今日の検索システムの多くは、ユーザとの相互作用を行うことを前提として作られている。オンライン検索で広く用いられているブール検索システムでは、ユーザが検索結果を確認した上で検索式の修正を行う適合性フィードバックが広く行われている。適合性フィードバックを行うことによって、検索性能の向上が望めるとともに、検索対象に対する主題知識が不十分な場合でもこの操作を繰り返すうちに、適合する情報を得ることが可能になると言われている。このことは、どの程度ユーザが適合性フィードバックを行ったかによって、得られた検索結果に差異が現れることを示している。つまり、ブール検索システムにおいて、どの時点で検索セッションを終了するかは検索性能に大きな影響を及ぼしているといえる。

一方、インターネット上のサーチエンジン等で広く見られる適合度順出力を行う検索システムには、出力件数をユーザ側で調整が可能であるという特徴がある。そのため、実際に適合度の高い順に検索結果が出力されたと仮定すると、上位数件の出力では精度が高く、多数の結果を出力するに従って再現率が上昇していく。一方で再現率を高めれば高めるほど、精度は下がっていくと考えられる。そのため順位付け出力を行うシステムでは、何件出力したかによって精度・再現率に大きな影響を受けると言える。実際に、Keen が順位付けを行う適合度順検索システムとブール検索システムの精度・再現率を比較した際[(4)]に、適合度順検索システムの精度がブール検索システムに劣る理由として、ユーザが過大な件数を出力している点を指摘している。理論的には優れているとされている適合度順検索システム[(5)]であるが、ユーザが適当な時点で検索セッションの終了を行わない場合、むしろ検索性能の低下を招くおそれがある。

情報検索の過程の中で、「検索の終了」は検索結果に大きな影響を及ぼす行動であると言える。しかし、これまでの検索戦略に関する研究をはじめとした情報検索行動研究では、どのように情報検索を行っていくかについての研究が進められてはいる[(6)]が、どのように情報検索を終了するかに関しては十分な研究はなされてはいない。これまで情報検索行動研究の中では、Fidel がサーチャーが検索を終了する際の条件について触れている。[(7)]サーチャーを、検索式の概念的意味を変更せずにシステムの特性を利用して検索集合を修正していくタイプ(操作派)と、検索式の概念を変更することによって検索集合を修正するタイプ(概念派)とに分類した上で、以下の５つの条件を挙げている。

A)操作派

①主題についておよその見当がついた場合

②十分な数の文献が検索された場合

③これ以上検索集合の質が変化しないだろうと感じた場合

B)概念派

④再現率が満足できるものになった場合

⑤再現率をこれ以上増す必要性のない場合

しかしこれら五つの条件のうち、①②④⑤の条件は「ユーザが検索結果に満足した」時点で検索セッションを終了させていると読み替えることができる。また、③はこれ以上検索集合の質を改善する手だてがなく、検索に行き詰まったケースととらえられる。だが、現実には、ユーザが検索結果に満足しないまま検索を終了させているケースはかなり多いと予想され、これをすべて③の条件のみで説明するには不十分だと思われる。検索結果に不満が残るものの、ユーザに検索を終了させた要因を明らかにすることは、情報検索行動を考える上で重要であると思われる。そこで、情報検索セッションの終了に影響を及ぼした要因について、調査および分析を行った。

3.検索セッションの終了に影響を及ぼす要因

調査は、電子メールによるアンケート調査を中心とし、一部紙によるアンケートも併用して調査を行った。主たる調査対象は、情報知識学会に所属する研究者、図書館・情報学を専攻する大学院生といったエンドユーザであるが、代行検索を行っているサーチャーも一部含まれている。アンケートでの質問内容は、あらかじめ最近行った1件の情報検索セッションを想定した上で、その情報検索セッションの終了を決断する際に影響を及ぼしたと考えられる要因と、その情報検索セッションで得られた検索結果への満足度を質問した。情報検索セッションの終了に影響を及ぼす要因に関しては、あらかじめ影響を及ぼすと考えられる17の要因を設定し、それらの要因がどの程度、情報検索セッションの終了に影響を及ぼしたのか、「A.強く影響した」、「B.やや影響した」、「C.まったく影響しなかった」の三つの選択肢から選択し回答するよう求めた。この17の要因の内容については、次項の表に示す。この17の要因には、情報検索の内容に直接的に関わる要因とともに、要因16の「他の用事」のような検索内容とは関わらない要因も含んでいる。また、あらかじめ設定した17以外の要因が影響を与えたと場合は、その他としてその内容を記述することとした。得られた検索結果のへの満足度については、「A.十分満足」、「B.やや満足」、「C.やや不満」、「D.大いに不満」の四つの選択肢から回答するものとした。

アンケートは全体で27件の回答が得られた。17の要因のそれぞれについて、検索セッションの終了に「A.強く影響している」及び、「B.やや影響している」と回答された割合を表に示している。この結果から、全般的に影響を及ぼしている要因として、「これ以上検索するために必要と思われる主題知識が十分ではないから」や「事前の予想通りの出力件数で、適合度の高い文献が得られた」とユーザが感じたときに、情報検索セッションを終了させていることがわかる。しかし、少数ながら「予定したよりも時間がかかったから」や「他の用事を優先したから」、「体調面の問題」など、情報検索の内容と直接かかわる要因ではない要因が、セッションの終了に影響を及ぼしたという回答もみられた。なお、あらかじめ設定した17の要因のうち、すべての要因が何らかの形で影響を及ぼしており、情報検索セッションの終了は非常に多様な要因から決定されていることが明らかになった。

この情報検索セッションの終了に影響を及ぼす要因と、ユーザの満足度の関係では以下

表　検索セッションの終了に影響を及ぼす要因

	要因	
要因1	予想した通りの出力件数が得られ、かつ適合度の高い文献が得られたから	10.204%
要因2	予想した通りの出力件数が得られたから	7.143%
要因3	検索もれがあると思われるが、適合度の高い文献が得られたから	8.163%
要因4	出力件数は多すぎるが、適合度の高い文献が得られたから	6.633%
要因5	望んでいた情報が検索しているデータベース中に存在しないと思ったから	2.551%
要因6	これ以上検索するために必要と思われる高度な検索機能がシステムにないから	6.122%
要因7	これ以上検索するために必要と思われる高度な検索機能が使いにくいから	5.612%
要因8	これ以上検索するために必要と思われる主題知識が十分でなかったから	11.224%
要因9	新たな検索戦略(検索式の修正など)が思いつかないから	9.184%
要因10	経済的理由(接続料金・出力料金等)	7.653%
要因11	他の人物に検索を行っているところを見られたくなかったから	1.531%
要因12	使用している端末を利用したい人が待っていたから	1.020%
要因13	検索にもう十分な時間をかけたから	9.694%
要因14	開始前に予定したよりも時間がかかってしまったから	4.082%
要因15	眼の疲労	1.020%
要因16	他の用事があり、そちらを優先したから	6.122%
要因17	空腹など体調面に問題が生じたから	2.041%

のような傾向が読み取れた。図に、満足度がAないしはBと高かった検索での終了要因と、満足度がCないしはBと低かった検索での終了要因との比較を示している。満足度が高いと評価された検索では、「予想通りの出力件数で、適合度の高い文献が得られた」や「検索もれはあると思われるが、適合度の高い文献が検索された」と感じたときなど、適合文献を入手できたことを確認した上で情報検索セッションを終了させている。一方で、満足度が低い評価の検索では、「新たな検索戦略が思い浮かばない」「主題知識が十分ではない」といった、検索戦略において手詰まりの状況に陥ったと感じたときに終了している。同時に、「検索に十分な時間をかけた」「予定よりも時間をかけた」と感じたことから情報検索セッションを終了させた割合も高い。また、「他の用事がありそちらを優先した」や「体調面の問題」、利用している端末の置かれている環境といった本来検索の内容とは関係のない要因も影響を与えている。全般的に、検索の内容に直接的に関わる要因1から要因9の割合は、満足度が低い回答では比率が低くなり、検索の内容とは直接的に関係のない非本質的な要因である要因10から要因17の比率が高くなっている。これらの結果から、情報検索セッションを終了する際の要因として非本質的要因が影響を及ぼしている事例では、検索結果に対する満足度が低くなる傾向が見られた。また、満足度が低い事例ほど様々な要因が影響を及ぼしている。これは、ユーザが検索内容に関わる要因からだけではなく、非本質的な要因の影響によって検索を終了したために、検索結果に対する満足度が低くなったと思われる。

今回の調査では満足度の低かった事例であっても、この段階で検索を終了せず、検索戦術を変更し検索を継続させていれば、最終的な検索結果に対する満足度も違ったものになった可能性は十分にある。しかし、現実には本来情報検索を行う上で想定していないよう

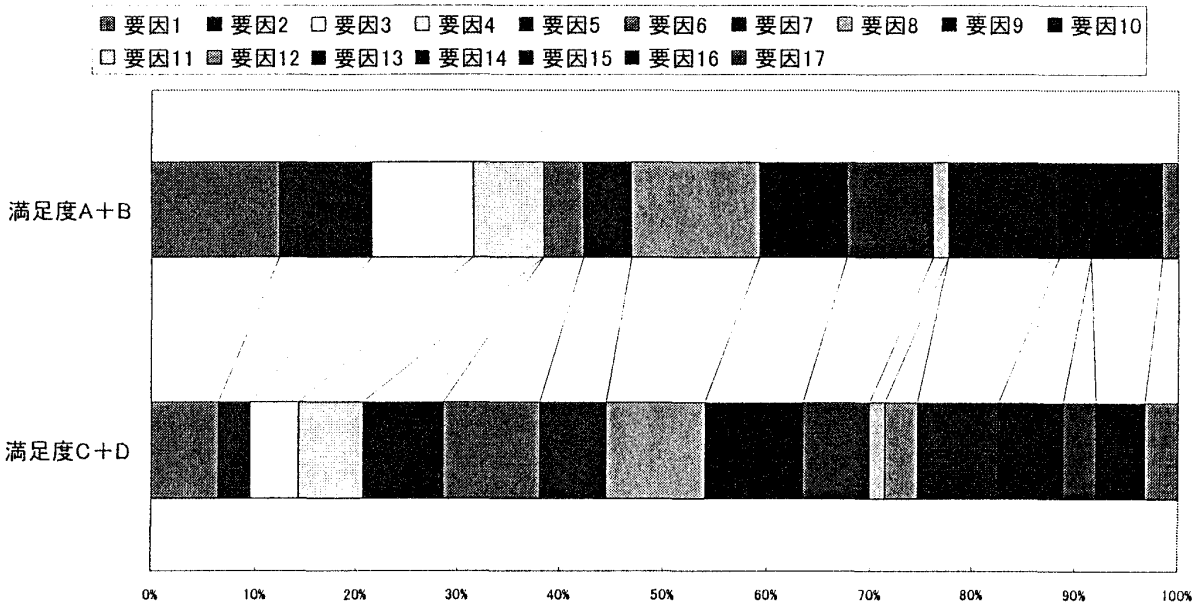

図 満足度からみた検索終了要因

な、「他の用事」や「体調面の問題」などの要因も影響を及ぼし、情報検索セッションを終了させているのである。

4.考察

情報検索セッションの終了要因に関する調査では、終了を決定する際の要因が多様であることと、これら要因と満足度の間に関係があることが明らかになった。これらの要因の中には、「体調面の問題」や「他の用事」、利用している端末の置かれている環境など、本来の情報検索行動とは本質的な関わりのない要因も存在する。これは、情報検索システムとユーザとの相互作用のみに対象を限定してきた従来の情報検索行動研究の限界を示唆していると考えられる。従来の情報探索行動研究では、検索の内容に本質的に関わりのある検索式、コマンド、システムからの出力といった対象を中心に、それらとユーザとの関わりを取り上げてきた。しかし、実際に情報検索セッションの終了に影響を及ぼす要因は非常に多様であり、情報検索行動を考える上で情報検索とは本質的に関わらないような要因も影響を及ぼしている。これまでの研究では対象とされていなかった情報検索行動における非本質的な要因が、実際の検索行動には存在することが今回の検索セッションの終了要因に関する調査から明らかになった。これらの多様な非本質的要因は、検索セッションの終了にとどまらず、キーワードの選定や検索式の構築といった検索行動全般に影響を及ぼしている可能性も考えられる。特にキーワードの選定等に、端末のおかれている環境が影響を及ぼすことは、十分に考えられる。今後、自身の要求に基づいて行われるエンドユーザの検索行動を考える上では、業務で検索を行うサーチャー以上に様々な要因が検索行動に影響を及ぼすことを考慮しなければならないだろう。

今後は、これらの点を踏まえた上で、エンドユーザがどのように情報を探し、アクセスし、入手するかについて研究を行っていく必要があるだろう。一般的に従来のブール検索では、まず情報要求から構築した検索式を入力し、何度か検索式を修正した後に適当だと思われる検索集合を出力するという流れで検索が行われてきた。多くの場合、この最後に

出力した検索集合が、その情報検索セッションにおいてユーザが望んでいた情報だと言える。しかし、インターネットのサーチエンジン等で用いられている適合度順検索システムは、検索集合を作成するのではなく、検索対象の検索式に対する順位付けを行うシステムである。そのため、ユーザが情報検索セッションの終了を決断するまで、検索対象が順に出力され続けることとなる。つまり、適合度順検索ではブール検索以上に、望んでいた情報を入手して検索を終了するのではなく、望まなかった情報が出力されることによって検索セッションを終了させる事例が増加すると思われる。より適切な時点で検索セッションを終了させるためには、エンドユーザがどのような要因から情報検索セッションを終了させたかを明らかにすることが重要である。特に、検索セッションを終了に影響を及ぼす要因と満足度の関係から、ユーザにとって最適な検索セッションの終了とはどのようなものかを明らかにしていくことは、今後の課題である。

---

(1) 下沢ゆりあ，倉田敬子．“オンライン検索における検索戦略と検索戦術”．Library and Information Science. No. 30, p147-171(1992)

(2) Bates, M. J. "Information search tactics". Journal of the American Society for Information Science. Vol. 30, No. 4, p205-213(1979)

(3) Bates, M. J. "Idea tactics". Journal of the American Society for Information Science. Vol. 31, No. 5, p280-289(1979)

(4) Keen,E.M. “Performance comparison of Boolean and ranked output retrieval". Proceedings of the 14th information retrieval colloquium,p89-101(1992)

(5) 安形輝，“情報検索評価における戦略性の導入：順位付け出力の非順位付け出力に対する優位性”．三田図書館・情報学会1997年度研究大会．三田図書館・情報学会編．東京，1997-10，三田図書館・情報学会．東京，1997，p57-60.

(6) 野添篤毅，池田順子．“オンライン検索過程における検索戦略と戦術”．図書館情報大学研究報告．Vol. 7, No. 1, p27-46(1988)

(7) Fidel, R. Online searching styles: a case-study-based model of searching behavior. Journal of the American society for information science. Vol35,No.4, p.211-221(1984)