

Web 情報検索の深化へ向けて

Towards a New Stage of Web Information Retrieval and Seeking

高久雅生*

Masao Takaku*

* 物質・材料研究機構 科学情報室

Scientific Information Office, National Institute for Materials Science

〒305-0047 茨城県つくば市千現 1-2-1

E-mail: TAKAKU.Masao@nims.go.jp

1 はじめに

1990 年代以降、爆発的な普及をみせたインターネットと WWW (World Wide Web; 以下 Web) は人々の生活に入り込み、ライフスタイルを変えるほどのインパクトを与えつつある。Web 登場以前からの全文テキストに対する検索技術と、Web 特有の特徴を活用した検索技術との融合が進み、Web 上の膨大な情報を探すツールとしてサーチエンジンが構築され、日常的に不可欠のツールとして利用されるに至っている。情報爆発時代とも言われる現在から次の 20 年に向けては、適切な情報やサービスへと導くための情報検索や、意思決定の支援技術がことさら重要になってくると考える。

現在、情報検索研究を進める上ではテストコレクションを用いた研究が大きな役割をはたしている。つまり、新たに開発した検索アルゴリズム等の有効性を示すための効率的な評価手法としては、テストコレクションを用いた評価を採用することが一般的なものとなっている。

たとえば、米国の TREC、日本の NTCIR、欧州の CLEF といった評価ワークショップが継続的に開催され、成果としてのテストコレクションが公開されている。当初はシンプルな文書検索を対象としていたこれらの評価ワークショップも、画像検索、ビデオ検索、メールスパム判定、ブログ検索などと対象メディアを変えたり、対象技術とし

て質問応答技術、要約技術、翻訳技術、分類技術といった多様な研究領域を対象にしたり、文書検索であっても医学分野、法学分野といった分野固有の問題に取り組むものなど、その領域を広げつつある。これらのテストコレクションもしくはそれに類する評価手段を持たない研究手法は、実質的には比較研究の素材となりにくいいため、研究上の評価を得ることすら難しい現状にある。

また、Ingwersen and Jävelin[1] は情報探索における情報検索システムと利用者との接点をめぐって、情報検索研究のこれらの評価技術そのものを新しい方向へと転回する必要があるとの提言を行っている。

本稿では特に、Web 上の情報検索のためのテストコレクション構築にまつわる課題と、それを補完しようとする立場から進めている研究を紹介し、それらの方向性を整理することを目指す。

2 テストコレクション研究とその課題

テストコレクションは検索課題、文書集合、適合文書集合の 3 点セットからなる。このうち、検索課題、適合文書集合の 2 つを作るには通常、人手による作業が必要となり、作成には労力を要する。そのために評価型ワークショップが開催され、研究のための資源を集積して研究の推進力を高めようとする努力が続けられてきた。

Web 文書に対するものとしては、1999 年から TREC Web Track が始まり [2]、NTCIR WEB タスクも 2001 年から開始され [3]、Web 文書に特化した情報検索システムの評価環境は整備されてきた。その中での課題としては、Web 特有の問題として主に、適合判定文書の網羅性の低さ、適合文書の更新頻度の高さの 2 点が問題として取り上げられてきた [4]。1 つ目の課題は、Web 文書は大規模なために、どれだけの労力をかけても人手で判定可能な範囲に限られるため、一定の未判定部分集合が大量に残ることになるため、実用規模のテストコレクション作成が困難となる点である。2 つ目の課題は、Web 文書がいったん判定済となったとしても、あとからその文書が更新されたり消えてしまうことによって、その判定内容の一貫性が保てないという問題である。

これらの点から、現実の Web 空間をそのままテストコレクション化することは難しいとされ、ある時点で Web から収集した文書集合を固定し、その範囲内で評価することが行われてきた。一方で、このような制約のもとに作成されたテストコレクションは、必ずしも現実の Web 空間の反映となっているとは限らず、一定の制限がかかるものとなっている。たとえば、PageRank や HITS アルゴリズムに代表されるリンク解析にもとづく検索手法は、ある程度現実的な規模のリンク密度が得られない場合、その算定値が大きく変化してしまい、現実の Web 空間で汎用的に使用できうる精度を持つのか、分からなくなってしまう [5]。このように動的な Web 空間に対する情報検索研究に取り組むには困難な点も多く、Web 上で現実に行われている運用サービスとの間に壁が生じかねないとの懸念もある。

また、検索課題の代表性も検討を要する課題のひとつである。たとえば、TREC Web Track においては、適合判定の労を減らすため、あらかじめ正解集合を決め、そこから検索課題を作成するというアプローチ

や、一定の検索クエリ集合のサンプルを元にした課題を作成するというアプローチを取っていた。一方、筆者も運営に加わった NTCIR WEB タスクでは、課題作成者に自由にトピックを想起してもらう方式で多数のトピック候補を作成した上で、その中から適切な課題を選択して検索課題を作成した [6]。このようにして作られた両者の検索課題は一見似たものとなっているものの、ある意味では全く異なる性質を持っているもとらえられる。つまり、TREC Web 方式では所与のサイト・ページに到達するためのトピックが与えられており、NTCIR WEB 方式では Web サイト・ページを所与のものとするのではなく、実世界の事物に密着したモノを対象としている。これらはタスク設計の違いであり、どちらが好ましいというわけでもないが、テストコレクションを利用する者は、そのタスク設計の背後にある意図と、自らが構築しようとするシステムの要件とを比較し、適切なものであるか、判断しなければならない。テストコレクションを用いた評価は、精度・再現率を加味した評価指標により、明示的で定量的な評価がなされるように見える一方で、上述したように、タスクとしての利用者モデル、テストコレクション設計上の意図、判定者の主観評価が加わっており、本質的に定性的な性質をもつことに注意が必要である。

一方で、観点を変えて Web 利用者の側から見た場合には、また別の難しさもある。これは主に、Web サービスのインタラクティブ性と、より広い観点からの文脈依存性との 2 点による。

Web 上で展開されているサービスは、たとえば旅行案内であれば、宿泊施設や列車の予約、地図上での動的な移動表示、ビデオや音声を用いたマルチメディア観光案内コンテンツなどといった具合に、利用者との動的なインタラクションによるサービス利用を伴うものも多く、従来からの情報検

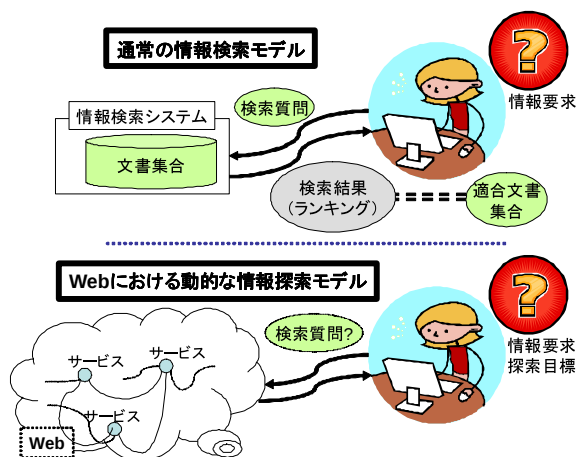


図1 従来型および Web 型の情報検索システムと利用者モデル

索研究で対象となってきた「文書」の枠を超えつつある。近年では Ajax と呼ばれるような明示的なページ遷移を行わずにインタラクティブ性を付加したサービスも多く、もともとの Web 文書が持つハイパーテキスト性ともあいまって、どこまでが一文書かといった文書単位を概念を確定することすら困難なものとなっている。このような環境において利用者は、Web サイトが提供する検索や予約・購入といったサービス機能を頻繁に使い、さらには、単一のサイトに限らずに複数のサイトにまたがってそれぞれのサービスから得た情報を関連付け、情報収集やブラウジングをしながら、学習や調査活動を行っている。図1に示すように、既存の情報検索のモデルがある程度固定化された文書集合、明示的な検索質問、ランキングによる検索結果と主題適合性に基づく適合判定結果から構成されていたのとは対照的に、Web におけるサービス利用を前提とした環境では、これらが多くの点で変化しており、既存モデルの再考がせまられている。つまり、既存の情報検索システムが前提としていた文書集合はより動的な環境におきかわっており、個々の情報要求も単一の検索セッションを超え、単一システムの枠も超えたような広い文脈の中での情報

要求や探索目標が重要となっており、情報検索研究においてもこれらを反映した、より Web 型の利用者の行動モデルを前提としたものへと変更する必要がある。

3 今後の課題

前節で述べたように、現在の Web 検索研究においては、動的な Web 上の利用者インタラクションや検索課題の代表性の課題が大きな問題として残されている。これらに取り組むにあたって、ひとつの鍵となるのは「人」を中心に置いた研究への取り組みである。

ここ数年、筆者は利用者実験を取り入れたシステム評価と、既存評価研究との接点を探る共同研究に参加している [7][8][9][10]。たとえば、既存の NTCIR テストコレクションを使い、利用者がシステムの出力結果にどのように反応するかを観察する実験を実施し、検索遂行中の経過を分析した。その分析結果によれば、既存テストコレクションによる評価において高い性能を示したシステムであっても、低い性能であったシステムでも、利用者による遂行時間平均には有意な差は見られないとの結果を得た [7]。これと同様の結果が TREC テストコレクションにおける既往研究 [11][12] でも指摘されており、より利用者に近いサービスを開発する場合には、システムの違いを反映できるレベルでの評価軸が必要とされているとも言える。このような結果が何に起因するのか、どのようにその弱点を克服すべきか、不明な点も多く、研究手法も成熟していないため、研究の蓄積が必要となっている。

最後に、筆者も参加した、2007 年のオランダ・アムステルダムで開催された SIGIR 併催ワークショップ WISI (Workshop on Information Seeking and Interaction) の全体討議の中で出てきた議論を紹介して、まとめたい。WISI における全体討議において、情報検索研究は今後どのような方向性を目

指すべきかの問いに，ある参加者は「Improving systems vs Educating users」という大枠を提示し，両者にバランス良く取り組むことが重要だとしながら，「賢いシステム」よりも「賢い利用者」をどのように実現するかが最終的には重要になるのではないかと議論を行った．ともすれば技術的な側面に偏りがちになる情報検索研究にあたっては，このような観点を重視しつつ研究の知見を蓄積してゆくことが重要であろう．

参考文献

- [1] Ingwersen, P.; Jävelin, K. : *The Turn: Integration of Information Seeking and Retrieval in Context*. Springer, 448p., 2005.
- [2] Hawking, D.; Craswell, N.; Thistlewaite, P.; Harman, D. : “Results and challenges in Web search evaluation”, *WWW '99: Proceedings of the eighth international conference on World Wide Web*, pp. 1321–1330, 1999.
- [3] 江口浩二; 大山敬三; 石田栄美; 神門典子; 栗山和子 : 「NTCIR-3 WEB: Web 検索のための評価ワークショップ」, *NII Journal*, No. 6, pp. 31–56, 2003.
- [4] Buckley, C.; Voorhees, E. M. : “Retrieval evaluation with incomplete information”, *SIGIR '04: Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 25–32, 2004.
- [5] Gurrin, C.; Smeaton, A. F. : “Replicating Web Structure in Small-Scale Test Collections”, *Information Retrieval*, Vol. 7, No. 3-4, pp. 239–263, 2004.
- [6] Oyama, K.; Takaku, M.; Ishikawa, H.; Aizawa, A.; Yamana, H. : “Overview of the NTCIR-5 WEB navigational retrieval subtask 2 (Navi-2)”, *Proceedings of NTCIR-5 Workshop Meeting*, pp. 242–222, 2005.
- [7] Takaku, M.; Egusa, Y.; Saito, H.; Terai, H. : “Comparing System Evaluation with User Experiments for Japanese Web Navigational Retrieval”, *Proceedings of the First Workshop on Web Information-Seeking and Interaction (WISI 2007) (SIGIR 2007 Workshop)*, pp. 49–52, 2007.
- [8] 高久雅生; 寺井仁; 江草由佳; 齋藤ひとみ; 三輪眞木子; 神門典子 : 「Web 情報探索における視線データの予備的分析」, 情報知識学会誌, Vol. 18, No. 2, pp. 181–188, 2008.
- [9] Terai, H.; Saito, H.; Takaku, M.; Egusa, Y.; Miwa, M.; Kando, N. : “Differences between Informational and Transactional Tasks in Information Seeking on the Web”, *Proceedings of the Second Symposium on Information Interaction in Context (IiX 2008)*, pp. 152–159, 2008.
- [10] 江草由佳 : 「CRES プロジェクトの紹介」, 情報知識学会誌, Vol. 18, 情報知識学会創立 20 周年記念号, 2008.
- [11] Hersh, W. R.; Turpin, A.; Price, S.; Kraemer, D.; Olson, D.; Chan, B.; Sacherek, L. : “Challenging conventional assumptions of automated information retrieval with real users: Boolean searching and batch retrieval evaluations.”, *Inf. Process. Manage.*, Vol. 37, No. 3, pp. 383–402, 2001.
- [12] Turpin, A.; Hersh, W. : “Why batch and user evaluations do not give the same results”, *SIGIR '01: Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 225–231, 2001.