情報検索によって新しい仮説は発見できるか

Don R. Swanson の新理論紹介

Can a New Scientific Hypothesis be Drawn from a Systematic Literature Search? : Study on "Undiscovered Public Knowledge" by Don R. Swanson

菊池 しづ子*

学習院女子短期大学

I.

現在行われている文献情報検索は、利用者が今関心を持っているのと同じ特定のテーマについて書かれた過去の文献を得るためのものである。従って、そのテーマに関する先行した研究がわかり、未だ解明されていない点はわかるにしても、これから先の研究に思わぬヒントを与えてくれるとか、全く新しい仮説を思いつかせてくれるとは限らない。

異なるテーマの文献間の関連を知るのには、引用文献索引がある。書誌結合 (bibliographic coupling) や組引用 (co-citation) は、直接の関連はない文献間の関連や結びつきを見いだすのに使われてきた。ただしこれらは、図書館情報学ではむしろ過去の各分野におけるコア文献を確定したり、あるテーマが時代とともにどのように発展してきたかといった歴史的分析に利用されてきたようである。また、書誌結合は過去において何のつながりもないが将来はつながりを持つであろう2文献の関係を見つけることには殆ど使えない。組引用の方が本来関係のなかった文献間に新たな結合をもたらすことには役にたつようである。それにしても引用索引を使っていままで気づかれなかった関連を発見し、今後の研究課題にしようという発想を体系だてたものはこれまでなかった。

多くの研究者に経験のあることと思われるが、研究

上の新しいアイデアの類はむしろインフォーマル・コミュニケーションから、それも違った分野の人との接触から得られることが多い。つまり、同じ研究の上を辿っているだけでは発展はないのである。ところが残念なことに、他の分野のこととなると文献から得ることは極めて難しい。ただでさえ今日では文献の量が多すぎるのに、一体何処に研究のヒントを与えてくれるものがあるのかなど、なかなかわかるものではない。勿論第一線の研究者は常日頃新しい文献を出来るだけ広範に目を通すようにしているではあろうが、結局、いろいろな情報の中から関連のありそうなものをかき分けて自分の研究に結びつけるある種のカンが、有能な研究者には必要だということになる。

そういうわけで、ひょっとしたらガンの決定的な治療薬さえも既に発見されているのかもしれない。ただそれがガン研究者の知らないところでのことなのかもしれないのである。有効な結び付きに気づかずにいるための研究の無駄がどのくらいあるのかは調べようがないが、そういった無駄を出来るだけなくすための研究方法が必要であらう。

シカゴ大学 Library School の Don R. Swanson は従来の情報検索の限界をまとめた上で、²⁾³⁾新しい検索法を考え出した。これから紹介する Swanson の文献検索法は、カンや幸運がなくても根気よく綿密に検討すれば、文献の操作だけで従来関連づけられていなかった文献どうしの関連を発見できるというものである。

* Shizuko KIKUCHI : Gakushuin Women's College. 3-20-1 Toyama, Shinjuku-ku, Tokyo 162, Japan.

II.

Swanson は、上記のテーマに関して極めてわかりやすい順序で次々に論文を発表している。従って、書かれた順に（刊行された順とは少しずれている）読んでいくことでよくわかるようになっている。

2)5) は、この一連の研究の背景を説明し、その原理を説明してからレイノー病と魚油の関係の例について簡単に紹介している。3)4) は専らこのレイノー病と魚油の例について述べているのだが、まだかなり原理的な説明が中心で、具体的な説明に欠けている。というのは、この2論文ではふたつの文献群が互いに無関係で書誌的に独立していることを説明しているが、その両者の関係をどうやって見出したのかの説明がないからである。しかし7) では、このふたつの文献群の間の関係を見つけるためのシステムティックな検索方法について順序よく詳しく述べている。ここに至って、初めレイノー病から検索を始めた時点では対象となる関連文献群（この場合魚油）は未知であり、両者の関係を見出すことこそがこの研究のコアだということが明らかになるのである。

第2の例は偏頭痛とマグネシウムに関する文献群の関連を見出すというもので6)と8)である。やり方は全く同じで、この研究の妥当性を検証するために書かれている。どちらの場合も医学的な正当性を問うためにまず医学雑誌に発表している。それゆえ、3)6)では、文献の医学的な内容に力点がおかれ、JASISに発表した4)7)8)では検索の面に力点がおかれている。さらにSwansonはこれらの研究の「まとめ」にあたる論文も用意している。⁹⁾

Swansonの言わんとするところは簡単である。もし、「AはBの原因である」と「BはCの原因である」という2つの確立した説があったとすれば、「AはCの原因である」という仮説が想定される。もし誰もこの三段論法による仮説に気づいていなかったらこの仮説は前の二つの仮説が公的知識であるにもかかわらず「未発見の公的知識」だということになる。実際の検索では、「AはBの原因である」と論じている文献群と「BはCの原因である」と論じている文献群との間のつながりを見つけ出すことにより導き出される。

Swansonは、論理的にはつながっているにもかかわらず書誌的にはつながっていない2つの文献群を見出し、確定するプロセスを明らかにしたのである。このプロセスは、大きく分けて2つの部分からなっている。

前半：論理的につながりのある文献群の対を見つける
為の探求のプロセス
後半：そのなかからすでに気づいた人がいて検証されている対を全て除くこと

そして実際の探索方法は4つの段階に分かれているといえる。ここではレイノー病と魚油の例を示そう。この探索全体の目的はまだ誰も気づいていないこの病気の治療法の可能性を捜し出すことである。

前 半

第1段階：まずレイノー病からスタートする。Medlineを使ってレイノー病がタイトルとしてついている文献を見つけ、試行錯誤によってこの病気の治療法として可能性のあるような副タイトルのついた文献を探す。このケースではそれは血液に関わりのある語——病気の状態によっておこる血液の変化を示すのに使われている——ではないかと思われた。これらの血液関係の語がレイノー病の治療の可能性の手がかりを示唆しているのではないか。

第2段階：次にこの手がかり（血液に関する語）を検索語として文献の検索を行う。そうすると、レイノー病にみられるような血液症状を魚油によって改善することが出来ると指摘している文献があった。もちろん、こういった文献はレイノー病について触れているわけではない。こうして血液症状を手がかりにして出発点（レイノー病）と未知のターゲット（治療法すなわち魚油）とが結び付いた。

後 半

第3段階：次にこの仮説（レイノー病と魚油の関係）がすでに知られているものか否かを検討する。オンライン検索でレイノー病と魚油とを同時に含む全ての文

献を調べるのは簡単である。ひとつもなかった。もちろん、このようなプロセスの過程では他にも多くの仮説が候補に上がり、そしてすでに検証済みだと判明した時点で、そうした仮説は棄却された。

第4段階：もし誰かがすでにレイノー病と魚油との関係について考察しているのなら、それぞれの文献群の間で引用がなされているはずである。Scisearchを使ってこの2つの文献群間には co-citation はないことがわかった。こうして「レイノー病は魚油によって多少なりとも治癒されるのではないか」という新しい仮説を立てることが出来た。

最後の仕事：最後にすべきことは、仮説としてたてたこの両者の関係が医学的に筋がとおるか否かを、文献を読んで慎重に調べることである。

手がかり(B), 更に(B)からターゲット(A)を見出すプロセスがこの探索のコアである。Swanson は言う、

「Cすなわち既知の病気から未知のターゲットAへ到達するためには、そのあいだに介在する未知の結合役Bを探さねばならない。Bとはその病気のメカニズム、原因、或は悪化させる要因などに関連するもののはずである。Bを確定するために様々な推測が試行錯

誤で検討される。」

「・・・この第2の探索では、Cについて触れていないBに関する文献を広く検討するのである。ここでもオンラインサーチの結果をブラウジングすることによって、AすなわちBという結果をもたらす可能性のあるなんらかの要因或は物質との関連を探るのである。」

ちなみに、第2の例では偏頭痛の治療にマグネシウムが有効ではないかという仮説をいくつかのキーワードを介在して導き出している。レイノー病の例についてこのプロセスをSwansonが図にしたものを下に転記する。(検索語がわかるように英語のままとした。)

こうしてみると、これは一種の研究方法そのものだといえよう。多分研究者は無意識に、或はもっとシステムティックでないやり方にしろ、これに近い方法で新しい仮説を見つけていることもあるだろう。例えば偏頭痛を扱っている研究者は偏頭痛のメカニズムや特徴を分析して各要素に対する治療法を見つけようとするだろう。しかし新しい仮説に役立つような関連には気づかずに過ぎてしまうことも多いだろう。

新しい知識は実験室と同様に図書室からも(つまり文献を分析したり探索したりするプロセスを通じて)得られるのだという考えを科学者たちはもっと真剣に取り合うべきだとSwansonは主張しているのである。

III.

Swanson の考えでは、これは情報検索の新しい手法であり、同時に科学者に対しては一種の研究手法の提案であった。また、科学者に対して文献の重要性を示したかったとも思われる。医学関係の雑誌に載せているのはその表れであろう。Swanson は医学的知識があまりなくても文献検索のみで新しい仮説を見出せるということを証明したといっている。もしこういったことが進めば研究活動において実験室と図書室の分業、あるいは研究者とライブラリアンとの協力がもっと有効になるのではないだろうか。図書館員の問題にもっと引き付けていえば、研究活動の一分野を担うレファレンス・ライブラリアンの存在意義はこのようなところから生まれてくるのではなからうか。

Swanson はかなりこの研究に力を注いだようで、また自信も持っていたようである。惜しむらくは、JASIS におけるタイム・ラグがひどく、プレ・プリントが出てから雑誌上で刊行されるまでに2年もの時間が経過してしまったことである。

この論文群に対しての図書館情報学関係の人々の反応は概ね高い評価であるように思われる。少なくとも従来の情報検索の研究とは違った独創性を持っているということは誰もが認めている。また、3段論法のみならず4段論法でもいいのではないかといった意見も提出されている。¹⁾

私の思うこの論文群の独創性とは、情報検索を情報検索だけの論議にとどめず、科学研究の一部として為し得ることを考えたということである。もうひとつは、組引用の使い方である。従来は組引用にしろ、書誌結合にしろ、それがあるかどうかという話だったが、Swanson の場合、それがなくことによって新しい仮説を見つけるという点である。

問題はこれを実際の科学研究の場で利用できるのか、どの程度有効性があるのかということである。それがはっきりしないと、この方法の価値は定まらない。最近 Swanson は、レイノー病に魚油がきくという仮説の正しさは実験によって医学的に証明されたと言っている。¹⁰⁾ 実は私は、魚油の効用は周知のことであって、わざわざ時間をかけ検索費用をかけてレイノー病と結

びつけるほどのことがあろうかと思っていたのである。¹¹⁾

そこで必要なことはもっと研究者にこれを知らせ、その有効性を検証してもらうことだと思う。レイノー病や偏頭痛では問題が小さいかも知れないが、もしこの方法が役に立つのならもっと重大な問題の解決にも役立つかも知れない。私がこのような紹介記事を書いたのはそのような反応が知りたかったからなのである。残念ながら医学研究者は忙しくて図書館情報学関係の論文などは読まないであろうし、*Perspectives in Biology and Medicine* という雑誌は日本ではあまり読まれていないらしい。それゆえ、この記事を読まれた方はもし機会があったら是非医学研究者の方々に意見を伺っていただきたく、お願いしたい。

引用・参考文献

Swanson, Don R. の文献

- 1) Information retrieval as a trial-and-error process. *Library Quarterly* 1977 ; 47 (2) : 128-48. これは直接には今回のテーマに触れているのではないが背景となる Swanson の考え方がわかる。
- 2) Undiscovered public knowledge. *Library Quarterly* 1986 ; 56 (2) : 103-18.
- 3) Fish oil, Raynaud's syndrome, and undiscovered public knowledge. *Perspectives in Biology and Medicine* 1986 ; 30 (1) : 7-18.
- 4) Two medical literatures that are logically but not bibliographically connected. *JASIS* 1987 ; 38 (4) : 228-33.
- 5) Historical note : information retrieval and the future of an illusion. *JASIS* 1988 ; 39 (2) : 92-8.
- 6) Migraine and magnesium : eleven neglected connections. *Perspectives in Biology and Medicine* 1988 ; 31 (4) : 526-58.
- 7) Online search for logically-related non-interactive medical literatures : a systema-

tic trial-and-error strategy. JASIS 1989 ; 39 (5) : 356-8.

- 8) A second example of mutually-isolated medical literatures related by implicit, unnoticed connections. JASIS 1989 ; 39 (6) : 432-5.
- 9) Logically-related, mutually isolated medical literatures as sources of new knowl-

edge. 未刊行草稿.

その 他

- 10) Letters to the Editor. JASIS 1989 ; 40 (3) : 152.
- 11) 魚油とコレステロールや血圧との関係については家庭用の医学書にも書いてある。