

研究フローにおける 抄録・引用データベース「Scopus」の 利用価値

清水毅志

●Scopus誕生の経緯

エルゼビア社は1580年にオランダ国ライデンにて創立された長い歴史を持つ学術出版社である。400年以上に渡り学術雑誌や専門書を出版してきており、ガリレオ・ガリレイの著作であるDiscorsi e Dimostrazioni Matematiche（邦題：新科学対話）も出版物の一つである。1997年には、サイエンス・ダイレクトという名称で発行ジャーナルの全電子化を進め、現在の電子ジャーナルの先駆けとなった。

学術論文コンテンツを膨大に有するエルゼビア社は、他の出版社の書誌情報も取り入れて新しい抄録・引用データベースであるScopus（スコパス）を2004年に発表した。スコパスという名称は*Phylloscopus Collybita*という渡り鳥の名前に由来する。この小さな鳥は長距離を正確に飛行する優れたナビゲーション能力を持つことで知られており、膨大な学術情報のなかから必要な文献に正確に辿り着くデータベースを作りたいという願いを込めたものである。当時の方針からもわかる通り、あくまでも利用の目的は研究そのものを対象としており、利用者の中心は研究者やレファレンスライブラリアンであった。

この研究業務目的に研究分析（評価）目的が加わってきたのは、Scopusのユニークな機能である「著者名名寄せ」と「研究機関名名寄せ」が実用化されたことが起点となっている。時を同じくして、大学への運営交付金や研究資金提供が世界的に厳しい経済状況になり、政府や研究資金提供団体は研究の成果と優位性を示す具体的な数値指標を求めるようになった。そのため、Scopusは単なる文献検索用途だけではなく、研究の一般的サイクル、つまり研究フローの全過程において必須なツールとなっている。

●Scopus作成における基本ポリシー

Scopusは発表当初から、次に掲げる基本ポリシーを基に開発を行ってきた。

(1) ユーザー中心デザイン

利用者が初めて画面操作する場合であっても、直感的に入力箇所や操作手順がわかるようなシンプルなデザインにすること。

(2) クリック数の最小化

論文フルテキストの表示やダウンロードまでが最少のクリック数で完結するようなサイト構成にすること。

(3) 世界最大の収録コンテンツ

影響力の高いコアジャーナルに限定するのではなく、ニッチな研究領域や欧米以外の地域での発行ジャーナルも精査して広く収集すること。

(4) ワンストップサービス

研究分野、文献タイプ、出版年等による細分コレクション化を避け、どの研究分野やポジションの利用者であっても、ワンストップサイトで包括的に操作できること。論文情報のみに限らず、研究に必要な他の情報（特許等）も同時に検索できること。

(5) 情報スピード

雑誌発行からScopus収録までの時間を最短化し、収録文献や被引用数のアップデートを毎日実施すること。

Scopusはこれらの基本ポリシーを遵守しながら、収録文献やユニークな機能を追加してきた。発表当時は収録文献数は約2500万レコード程度であったが、2017年2月現在では、6700万レコードまで拡大している。

●研究者の立場からみたScopusの活用

ここでは研究フローにおいて、研究者からの視点と、研究分析者からの視点、それぞれの立場から利用度の高い機能の一部を紹介したい。

(1) 研究関連文献の調査を目的として

Scopusにアクセスすると最初に「文献検索」画面が表示される。検索画面は非常にシンプルで、検索語を入力し検索ボタンをクリックするのみである。検索に際して予め検索範囲（検索場所や出版年など）が決まっている場合にはこの時点で入力することもできるが、検索結果の絞り込み機能が充実しているため、筆者としては最初の検索は広めに検索することを推薦する。

検索結果は「文献検索結果」画面として表示される（図1参照）。通常は検索結果が多いため「並び替え」と「絞り込み」作業が必要となるが、Scopusではこれらの作業をクリックのみでサポートする仕掛けと機能が多数用意されている。特に「絞り込み」機能においては内訳表示として、出版年、著者名、研究分野、出版物タイプ、文献タイプ、キーワード、著者所属機関、機関所属国、本文言語、各項目それぞれの上位160位までを表示する。その内訳を確認して絞り込むことがクリックのみで行うことができる。特に上位キーワードでは予想外のキーワードが含まれていることがあり、次に続く検索語のヒントになる場合も多い。

(2) 論文フルテキストのダウンロードを目的として

抄録を確認して関連度が高いと判明した文献は、研究者としては当然フルテキストまでを読む要求が出てくる。これまでは対象文献を一つずつクリックしながらダウンロードしていたが、Scopusでは必要な文献を一括して自動ダウンロードする機能が用意されている。通常フルテキストはその機関で購読契約している雑誌しかダウンロードできないが、近年はオープンアクセスジャーナル（著者が投稿料を支払うビジネスモデルで、読者はフルテキストを自由に閲覧できる）やエンバーゴジャーナル（発行から一定期間が過ぎたものを公開する）など、フルテキストを無料で閲覧できるジャーナルが増えている。このため自動ダウンロード機能を使うことにより、予想していた以上に多くのフルテキストがダウンロードできたケースが頻繁に見受けられる。

(3) 文献管理ツールによる管理を目的として

研究に関連度の高い文献は論文執筆の際、特に参考文献として必要となることが多い。Scopusは気になった論文の書誌情報を簡単に文献管理ツールへエクスポートすることができる。これらの文献リストにより参考文献記載時の負担を軽減できる。無

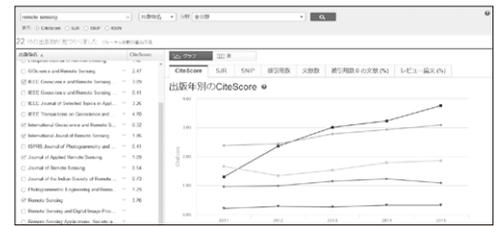
料の文献管理ツールであるMendeleyは、指定されたメンバー内で文献の共有やコメント追加も可能であるため世界中でユーザー数が増加している。

図1 Scopusの「文献検索結果」画面



(出所) www.scopus.com/search/form.uri

図2 「ジャーナル比較」の画面



(出所) www.scopus.com/source/eval.uri

(4) 論文投稿時の候補ジャーナルの検討を目的として

研究者は自身の研究内容や成果の重要性に応じて、論文を投稿するジャーナルを予め決めていることが多い。もちろん特定の著名ジャーナルに投稿することを目的にしている場合もある。しかしながら出版社で業務している者として、ジャーナルは生き物でありその評価は常に変動していることを実感している。特に近年オープンアクセスを中心として新しいジャーナルが多数発刊されているため、専門分野に詳しい研究者といえどもすべての関連ジャーナルの動向を知ることは困難である。Scopusではジャーナル比較機能として研究内容の近いジャーナルを視覚的に比較検討することが可能である（図2参照）。3つのジャーナル指標（CiteScore、SJR、SNIP）はもちろん、論文数や被引用数、レビュー論文比率、またそのジャーナルで1回も引用されなかった論文の割合など、かなり赤裸々な指標をもって複数ジャーナルの年次変化を比較することができる。ある研究分野において数年前までトップジャーナルとして認識されていたものが、実は今では他のジャーナルの後塵を拝している場合もよく見受けられるため、先入観を取り払って再度投稿ジャーナル候補の検討を行うことをお勧めする。

●研究分析や研究評価者の立場からみたScopusの活用

研究フローにおける研究評価作業は、論文のインパクトを客観的に分析し、研究の次ステージへの展開を

図るために必要不可欠なものである。優れた研究論文は研究者個人のみならず、その所属機関や研究資金提供団体にも高い評価をもたらし、研究フローの好循環を導くことができるからである。論文のインパクトを数値的に表すものとして被引用数が有名であるが、Scopusではビッグデータをベースにした多面的な数値指標を用意している。ちなみに研究者個人の評価指標としてインパクトファクターが利用されることがあるが、インパクトファクターはあくまでもジャーナルの評価指標であり、研究者個人や研究機関への評価に用いることは明らかに誤用といえるだろう。

(1) 特定論文の評価分析を目的として

各論文の「評価指標の詳細」ページでは下記のような論文評価指標が表示される（図3参照）。

図3 「評価指標の詳細」画面



(出所) www.scopus.com/record/pubmetrics.uri

- ① 被引用数に関するもの：被引用数、FWCI（論文あたりの平均引用数を分野と世界平均値で正規化したもの）、被引用ベンチマーク（同出版年、同分野、同論文種類のなかで被引用数のランキングをパーセンタイルで表したもの）。
- ② 学術活動に関するもの：文献管理サービス（Mendeley とCiteULike）での文献保存数。
- ③ 学術的コメントに関するもの：ブログ、レビュー、Wikipediaなどで言及された数。
- ④ マスメディアに関するもの：新聞や雑誌などのマスメディア媒体（欧米）での紹介数。
- ⑤ 社会的活動に関するもの：Twitter、Facebook、Google+などの一般的なソーシャルメディアでの言及数。

これまでのデータベースでは被引用数をベースにした指標しか算出できなかった。しかしながら被引用数は論文発表後から最大値になるまでに2~5年間の時間を要するため、新しい論文では経過時間が短く即時的な評価には適していなかった。これに対し文献管理サービスへの保管ダウンロード数は論文発表直後に数字として表れる。またソーシャルネットワークやマス

メディアでの報告も短期間に増加することでは同様であるが、こちらは研究者による学術界での評価ではなく、研究者以外も含め広く一般人を対象としている点が大きく異なる。このような評価指標を学術論文に利用することには賛否両論があるのだが、研究イノベーションや社会還元という側面では今後広く認知されていくものと思われる。

(2) 特定研究者の評価分析を目的として

Scopusに収録された文献のすべての著者には自動的に著者番号（Author ID）が作成され、同一著者の文献であると判断されたものはその著者番号にグループ化される、いわゆる著者名寄せ作業が行われている。この作業は独自の名寄せアルゴリズム（研究分野、所属機関、住所、共著者、出版年、メールアドレス等から同一性を判断）によって機械的に行われた後、著者本人からの修正依頼に基づきマニュアルでもブラッシュアップされていく。同一人物である確率が低い場合にはあえてグループ化しないため、1人の研究者が複数の著者番号を持つこともある。KAKENデータを用いた日本人研究者の名寄せ正確性（precision）は98%以上、再現性（recall）は99%以上であると報告されている（参考文献①）。このような著者名寄せ機能を最大限に活用して研究者個人の業績プロフィールがまとめられており、研究分野、研究業績（文献リスト）、総被引用数、共著者やh-indexまで、多様な業績指標を俯瞰することができる（図4参照）。

図4 「著者詳細」の著者プロフィール画面



(出所) www.scopus.com/authid/detail.uri

(3) 特定研究機関の評価分析を目的として

著者名寄せ機能と同様に世界の研究機関名でも名寄せが行われ、研究機関での発表論文数、共同研究機関、発表数の多いジャーナルや研究分野などが機関プロフィールとして表示される（図5参照）。研究推進や研究企画に関する検討では必須なデータだと考えられるが、研究機関全体の発表論文数など基本的な数字を把握している研究機関は思った以上に少ない。

