

オルトメトリクスは論文評価を変えるか

——ソーシャルメディアで算出する新たな指標

佐藤 翔¹・吉田光男²

¹同志社大学免許資格課程センター・²豊橋技術科学大学大学院工学研究科

従来のピア・レビューや被引用数の問題点や限界が指摘されるなか、TwitterやFacebookなどのソーシャルメディアで言及された指標をもとに評価する「オルトメトリクス」が登場し、注目されている。新たな指標から何がわかるのか、これまでの指標にとって代わるのか、その可能性を探ってみよう。

従来の研究評価手法とその問題

研究者は日夜、研究評価にさらされている。論文の査読においては、研究成果そのものが同じ分野の研究者によって評価される。採用や昇進、年次評価にあたっては、個人の単位でこれまでにだした研究成果(主として雑誌論文)の数と質が問われる。“Publish or Perish”(発表せよ、さもなくば滅びよ)とは使い古された格言だが、最近では“Publish and Perish”といい換えられることも多い。すなわち、現代の研究者には多くの論文を発表するだけではなく、高く評価されるような論文を発表することが求められるようになったのである。

従来、研究成果の質の評価は、査読のようにほかの研究者が中身を読んで判断する方法(ピア・レビュー)と、なんらかの計量的指標に基づいて判断する方法で行われてきた。後者では、多くの分野で論文が引用された回数(被引用数)に基づく指標が用いられている。たとえば、学術雑誌のインパクトファクターも被引用数に基づく指標の一つである^{*1)}。

さとう・しょう ● 同志社大学免許資格課程センター助教, 2013年筑波大学大学院図書館情報メディア研究科博士後期課程修了, <専門分野>計量書誌学, 学術情報流通, 情報利用行動, <趣味>読書, 旅行
よしだ・みつお ● 豊橋技術科学大学大学院工学研究科助教, 2014年筑波大学大学院システム情報工学研究科博士後期課程修了, <専門分野>ウェブ工学, 自然言語処理, 計算社会科学

しかしながら、これまでの研究評価手法には問題があることも指摘されている。ピア・レビューは個々の研究の評価に人的・時間的コストがかかるうえ、評価者によって結果が左右されたり、国籍や性別など、被評価者の属性によってもバイアスがかかる。一方、被引用数を用いる方法は、ピア・レビューに比べればコストを削減できるが、ある論文が出版されてから引用されはじめるまでには年単位で時間がかかるため、出版直後の論文は評価できない。また、質とは別に、分野やテーマによって論文の引用のされやすさが異なる^{*2)}こともあげられる。

そもそも被引用数も、算出に使われるのはもっぱら査読誌に掲載された論文からの引用であり、その論文を書いているのは研究者である。ピア・レビューも被引用数も、研究の質を「ほかの研究者からの評価」、つまり研究者コミュニティ内での価値観のみから判断していることには変わりはない。社会的に注目度の高い問題を解決する、多くの人命を救うなど、ある研究が社会に与えた影響も研究評価の観点になりうるが、このような観点は従来の研究評価のしくみ、とりわけ被引用数に基づく指標のなかでは見落とされがちであった。

そこで、被引用数に基づく指標の欠点を補うものとして近年注目を集めているのがTwitterやFacebookといったソーシャルメディアで言及された回数に基づいて算出される新たな指標、オルトメトリクス(altmetrics)である。本稿ではオルトメトリクスの概要と日本の化学論文を対象としたオルトメトリクス分析の実例を紹介していくとともに、この新指標をどう捉えるべきかを論じていく。

^{*1)} 被引用数に基づく指標の代表格はインパクトファクターであるが、これはあくまで雑誌の評価指標であり、掲載論文や研究者の評価に用いるのは明らかな誤用である。詳細は文献1にも記述があるので参照されたい。

^{*2)} 研究者人口が多く論文数の多い分野・テーマほどよく引用される。

オルトメトリクスとは：概要と現状

オルトメトリクス (altmetrics) は alternative (代替的な) と metrics (指標) を組み合わせた造語である。論文をインターネットで閲覧することが当たり前になるにつれ、それらの論文に向けられたリンクの数など、被引用数以外の要素を用いた指標をつくる試みも行われはじめた。これらの指標を“オルトメトリクス”と名づけ、その普及を後押ししたのは、情報学を専攻する大学院生であった Priem らが 2010 年に発表した文書, “Altmetrics: A manifesto” である²⁾。この文書で Priem らは、査読や被引用数に基づく評価について、とくにそれらにかかる時間の面での限界を指摘している。そのうえで、研究者が Twitter などのソーシャルメディアを活用しはじめていることから、それらの言及数などを分析することで、新たな指標を構築することを提案している。

この提案以後、オルトメトリクスに関する研究やサービス開発はおおいに進展した。その提供するサービスとして、提唱者であった Priem らが立ち上げた ImpactStory³⁾ や、その名も Altmetric 社⁴⁾ の提供する可視化サービスなどができている。Altmetric 社のサービスを見たことがある読者もいるだろう。Tweet 数や Facebook 言及数などを色分けし、ドーナツ型に表現したグラフが特徴的で、Nature Publishing Group や Wiley など、著名な出版者も導入している。

しかし、これらの会社のサービスなどでは DOI のついた論文しか集計対象にならず、DOI がついていない日本の論文誌などまではフォローが難しい。そういった論文も含め、CiNii をはじめとする主要な日本の論文データベースに対応したオルトメトリクス提供サービスとして、筆者(吉田)らが運営する Ceek.jp Altmetrics⁵⁾ がある。はてなブックマークなど、日本特有のソーシャルメディアにも対応しており、日本語の Web サイトで話題になった論文を知ることができる。

同様に、オルトメトリクスの集計対象に何を含むかは提供元によって異なる。オルトメトリクスといいつつも、多くのサービスでは従来の被引用数も集計・提供しているし、ダウンロード数などの閲覧統計を含む場合もある。また、CiteULike⁶⁾ や Mendeley⁷⁾ のように、文献管理ツールにおける登録数を盛り込む場合も多い。そのなかで、どのサービスでも共通して集計の対象としているのは Twitter や Facebook などの研究者以外も用いるソーシャルメディアであり、オルトメトリクスの研究論文でもこれらを扱ったものがとくに多い。

オルトメトリクスのメリット・デメリット

前述のようにオルトメトリクスは、ピア・レビューや被引用数に基づく指標の欠点を補うものとして登場した。オルトメトリクスに関する研究をレビューした Bornmann は、そのメリットとデメリット(限界)としてそれぞれ次の 4 点をあげている⁸⁾。

■メリット

- ①広域性：被引用数では測れない、広く社会に与えた影響を測ることができる可能性がある。
- ②多様性：研究データなど、論文以外の対象も評価しうる。
- ③即時性：論文公開当日～数週間で集計できる。
- ④公開性：Tweet など Web で公開された情報に由来するため、誰もが同一情報源に基づいて研究評価を行う。

■デメリット

- ①商業主義：集計元となる Twitter や Facebook など、企業のサービス方針が評価に影響しかねない。
- ②データの質：集計基準が統一されておらず、そもそもソーシャルメディア利用者層自体に偏りがありうる。
- ③オルトメトリクスが研究評価に適用可能というエビデンスが欠如。
- ④自身の論文への言及数を水増しすることなどが容易。

このように、オルトメトリクスの利用にはメリットもあるものの、問題点も多いといえるだろう。

日本の化学論文のオルトメトリクス ～Altmetric 社のスコア～

ここからは、オルトメトリクスによる評価の具体例として、本誌読者に馴染み深いであろう化学分野を例に、被引用数とオルトメトリクスとの比較や、オルトメトリクス上位論文がどんなものであるかを見ていきたい。対象とするのはオルトメトリクスが提案された 2010 年から 2014 年までの 5 年間に発表され、Web of Science で“Chemistry”に分類された原著論文で、かつ著者のなかに日本の研究機関に属する者を含む論文、45,324 報である^{*3}。これらの論文について、Altmetric 社の Web API を用いて取得したスコア(個々

^{*3} 日本の研究機関に限定したのは、データ取得の都合上、ある程度対象論文数を絞り込む必要があったためである。また、正確には条件に合致する論文は Web of Science 上で 49,823 報ヒットしたが、そのうち被引用数データの取得に成功したのが 45,324 報であった。

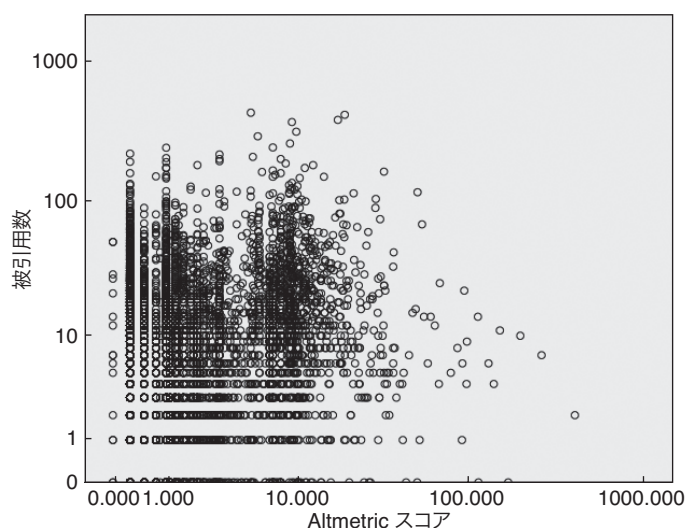


図1 Altmetricスコアと被引用数の関係 (N = 7759)

のソーシャルメディアなどでの言及を重みづけし、合算した値)と被引用数の関係を分析していこう。

45,324 報の論文のうち、なんらかの Altmetric スコアが付与されていた論文は約 17% の 7759 報にとどまり、それらについて、Altmetric スコアと Web of Science における被引用数の関係を散布図に示したものが図 1 である。また、表 1 は Altmetric スコアが付与されていた論文の割合と、スコアの平均値 (母数は論文 7759 報に限定)、被引用数が 1 以上の論文の割合と平均値を出版年別に示したものである。

表 1 より、被引用数が 1 以上の論文の割合が 89% であるのに対し、Altmetric スコアが付与された、つまりなんらかのソーシャルメディアなどで言及された論文の割合は前述のとおり 17% にとどまっていた。オルトメトリクスは被引用数による評価の欠点を補うことが期待されていると先に述べたが、現状、日本の化学分野ではそもそもオルトメトリクスでは評価できない論文が大半で、実用には耐えがたいという

点をまず頭に入れておくべきだろう。もっとも、Altmetric スコアの付与される論文の割合は一貫して増加傾向にあり、近い将来には多くの論文がなんらかの Altmetric スコアをもつようになる可能性もある。

図 1 を詳細に解析すると、Altmetric スコアと被引用数とのあいだには有意水準 1% で正の相関があるが、その相関係数はスピアマンの ρ^{*4} が 0.164 と、ごく小さい。これは Altmetric スコアが付与されていない論文を除いた値であって、スコアのない論文を加えれば相関はさらに弱まる。この相関の弱さの一因は、表 1 から明らかなように、出版年との関係にある。被引用数は出版年が古い論文ほど増える傾向にあるが、Altmetric スコアにそのような傾向は見られない。論文の引用は比較的長期にわたって行われ続けるため、出版から年数が経つほど増えていくのに対し、ソーシャルメディアでの言及は過去の論文に対してはあまりなされないため、出版年から時間が経過した論文のスコアがあまり上がらないものと考えられる。そのため、出版年を統制すれば被引用数と Altmetric スコアの相関は高まるが、その場合でも相関係数が 0.3 を上回ることではなく、あくまで弱い相関関係をもつにとどまった。

Altmetric スコアと被引用数とのあいだにあまり相関がないということは、オルトメトリクスを被引用数よりも即時的に研究評価ができるものとして、被引用数による評価の代替と見る場合には問題が生じる。一方で、社会的な影響度のような、被引用数では測れない側面を評価するものとする場合には、相関は弱いほうが好ましい。それは、これまでの指標ではあまり評価されなかった論文を高く評価できている可能性があるからである。実際に Altmetric スコアの高い論文の

*4 対象とするデータの分布について、仮定を置かずに相関関係を評価する指標のこと。一般に相関の評価に用いるピアソンの r はデータが正規分布することを仮定するが、被引用数はデータの偏りが大きいため、スピアマンの ρ を用いることが多い。

表 1 日本の化学分野論文の Altmetric スコアと被引用数の概要

出版年	Altmetric スコア (2015 年 9 月現在)				被引用数 (2015 年 10 月現在)		
	全論文数	付与論文数	割合	平均スコア ^{†1}	1 以上論文数	割合	平均被引用数 ^{†2}
2010	8,118	218	2.7%	4.04	7,852	96.7%	17.2
2011	8,756	710	8.1%	3.28	8,383	95.7%	14.6
2012	9,386	1,941	20.7%	2.58	8,746	93.2%	10.3
2013	9,556	2,246	23.5%	3.24	8,403	87.9%	6.5
2014	9,508	2,644	27.8%	3.08	6,940	73.0%	3.1
合計	45,324	7,759	17.1%	3.05	40,324	89.0%	10.1

†1 平均スコア算出の母数は Altmetric スコア付与論文のみ。小数点以下第 3 位で四捨五入。

†2 平均被引用数算出の母数は被引用数 0 論文を含む。小数点以下第 2 位で四捨五入。

表2 Altmetric スコア上位5位論文の詳細(2015年9月現在)

	DOI	Altmetric スコア	被引用数	言及の多い ソーシャルメディア	詳 細
1	10.1021/ac501998d	405.1	2	Twitter (794)	つくば市で採取した微粒子からウランが検出されたとする論文。福島第一原発の問題と関連して Twitter など話題に
2	10.1021/ol401012u	262.3	7	Twitter (362)	グラフィカルアブストラクトの化学結合の図でいわゆる「萌え」イラストを使用
3	10.1021/jf401819s	198.6	10	Mendeley (56) や ニュースメディアなど	世界一高いコーヒー豆ともいわれるコピ・ルアク(ジャコウネコのフンから採取した未消化の豆)について、ほかの豆との違いを化学的に分析
4	10.1255/jnirs.1124	169.7	0	Twitter (313)	論文掲載誌のツイートが300回以上 Retweet されている
5	10.1002/adma.201400420	152.2	11	Mendeley (58) や ニュースメディアなど	人体のなかで形態を変化させられる電子デバイスの開発として、おもに海外メディアで話題に

上位について、DOI、Altmetric スコア、被引用数、どのソーシャルメディアでの言及数が多かったか、またその理由をまとめたものが表2である。最も Altmetric スコアが高かったのは福島第一原発からの放射性物質の拡散状況を扱った論文で、Twitter において 800 回近い言及を集めている。一方で被引用数は出版から約1年が経過した時点で2にとどまる。研究としての注目度以上に社会的な関心の高いトピックを扱う、このような論文を高く評価できることは、オルトメトリクスのメリットを端的に示している。

しかし、話はそう単純ではない。第2位の論文も Twitter での言及が多いが、その理由はグラフィカルアブストラクト中にいわゆる「萌え」イラストを用いているためである。話題になったのは確かだが、それを研究評価に用いるべきと考える人は少ないと思われる。第3位、第5位はともにニュースメディアと Mendeley での言及が多い論文だが、5位の論文が、開発された技術の有用性で注目されているのに対し、3位の論文は「世界一高いコーヒー豆」「動物のフンから採取する豆」という話題性を理由に注目されている。最も問題なのは第4位の論文で、これは論文掲載誌による広報 Tweet が多数 Retweet (他者の投稿を自分の関係者に共有する機能) されただけで、人びとが何に注目したのかは不明である。上位5位までを見ただけでも、オルトメトリクスの利点と問題点がともに一覧できる状況にあることがわかる。

日本のソーシャルメディアではどうか？ ～Ceek.jp Altmetrics～

続いて、日本のソーシャルメディアにおける化学論文への言及を見るべく、日本の論文を対象とする Ceek.jp

Altmetrics における言及状況も調査した。対象は日本の学会が発行する ISSN をもち、CiNii Articles に採録され、国立国会図書館において「化学」に分類されている雑誌に掲載された論文とした。

Ceek.jp Altmetrics では Altmetric 社で実施しているような重みづけは行わず、各ソーシャルメディアなどでの言及数の合計をスコアとして用いている。表3は日本の化学論文について、Ceek.jp Altmetrics でのスコア上位5位を見たものである。分析対象として条件を満たした雑誌は全部で39誌存在したが、上位5位はすべて日本化学会の機関誌である『化学と教育』に掲載された論文で、研究成果というよりは解説・論説記事などである。すなわち、日本語 Web 空間で注目を集めるのは、原著論文よりも馴染みやすく、専門家以



オルトメトリクスを語るうえで欠かせないデジタルツール

表 3 Ceek.jp Altmetrics スコア上位 5 位論文の詳細 (2015 年 10 月現在)

	URL	タイトル	出版年	言及数	言及の多いソーシャルメディア
1	http://ci.nii.ac.jp/naid/110009625158	「昇華」の逆は「凝華」	2013	68	Twitter (62), はてなブックマーク (6)
2	http://ci.nii.ac.jp/naid/110008906710	紙類に含まれる放射能 (レーダー)	2005	39	Twitter (39)
3	http://ci.nii.ac.jp/naid/110001827316	折り紙で作る分子模型	1987	33	はてなブックマーク (26), Twitter (6)
4	http://ci.nii.ac.jp/naid/110009625159	高校化学における有機電子論	2013	15	Twitter (15)
5	http://ci.nii.ac.jp/naid/110001833856	だれが原子や分子をみたか	1993	12	はてなブックマーク (6), Twitter (6)

外も理解できる記事といえそうである。ただし分野によって傾向は異なり、情報工学分野などでは専門的な話題を扱った論文のほうがより多くの言及を集めている。

Ceek.jp Altmetrics で特徴的なのは、言及数上位論文の出版年と、言及元ソーシャルメディアの傾向だろう。前者については、1987 年や 1993 年など、ソーシャルメディアが存在しなかった時代の論文が多く言及を集めている。誰かが興味をもった論文の情報をソーシャルメディアに投稿することで、論文の「再発見」が促されていることが読み取れる。

言及元については、Twitter がおもな情報源であるのは Altmetric スコアと同様であるが、はてなブックマークがそれに次ぐ言及元となっている。最初に、オルトメトリクスは被引用数などで評価対象とならない社会的な影響度も含む「広域性」をその利点に掲げていると述べたが、集計の対象が欧米で主流のサービスにとどまるのであれば、「広域」といっても主として欧米社会に与えた影響を評価する指標にすぎないことになる。それを受けて中国は自国固有のサービスをオルトメトリクスに組み入れるよう働きかけている。日本からも同様に声を上げると同時に、Ceek.jp Altmetrics のような日本独自のサービスを活用することも視野に入れる必要があると思われる。

オルトメトリクスをどう受け止めるか

ここまで、身近な分野のオルトメトリクスの集計結果や上位ランキングを見てきたことで、「あまり使いものにならない」という印象をもたれた読者もいるかもしれない。その印象は正しい。すでに述べたとおり、オルトメトリクスは多くの問題点を孕んでおり、少なくとも現状、オルトメトリクス単体で研究評価を行うことは無謀である。それどころか遠い将来においても、ピア・レビューや被引用数に代わってオルトメトリクスが研究評価に用いられるようになることを考える研究者は少ない。

しかしオルトメトリクスの登場によって、被引用数に基づく指標への偏重が見直されるようになってきたことも確かである。

Altmetric スコア 1 位の論文は被引用数では高く評価されないかもしれないが、日本社会の最重要課題の一つに真正面から向き合ったものである。Ceek.jp Altmetrics での上位論文は、研究成果とはいいがたいが、どれも市民社会と学問の世界の橋渡しを果たすものといえよう。研究や論文を評価する観点はこのようにさまざまにありうるし、研究や研究者の「質」を構成する要素も多様であってよい。“Alternative”という刺激的な名前とは裏腹に、オルトメトリクスはピア・レビューや被引用数とは異なる観点から研究を捉え、従来の評価を補完しうる指標と見なす論調が最近では主流となっている。そのため、オルトメトリクスや被引用数など、複数の指標を参照できる環境を用意し、評価の目的に応じて適切な指標を選ぶようにしようという動きもでてきている。

完璧な研究成果の評価法はない

主流派については以上の通りであるが、最後に私見を述べさせていただきたい。まずは、このオルトメトリクスの登場を契機に、研究業績を量的指標によって評価すること自体の不可能性が、研究者のあいだで意識されるようになることを願う。より正確を期せば、被引用数もオルトメトリクスも、突出した業績を発見する能力をもつことは確かである。Web of Science のデータに基づくトムソン・ロイター社のノーベル賞予測はよく当たるし、Ceek.jp Altmetrics で上位の論文は、どれも何かしらの面白さがある。

しかしそのような突出した業績は、研究活動の営み全体のごく一部にすぎない。加えて、とくに被引用数に基づく場合、指標を用いて示されるまでもなく、分野内ではよく知られた業績の評価が高くなることも少なくない。つまり、評価の分析を行う前から評価が高いことは自明なのである。

一方で、本稿であげたような量的指標によって機械的に評価したいのは、大多数の研究者が書いてほどほどの被引用数を得て、自分と出版社を含めて数回程度 Tweet されただけの普通の業績である。そういう地道な研究の積み重ねの先にこそ科学の発展があるとも考えられるが、量的指標はそれら

の地道な、あるいは普通の業績を評価するには著しく不向きである。1～2回の被引用数の差、あるいは10～20回程度のTweet数やFacebookの「いいね!」の差にどんな意味があるのか。そんなものはほとんど友人の数の差で決まってしまうだろう。

それでも現代の社会状況を考えるに、研究や研究者を「評価しない」とはいいがたいし、評価するにあたってはピア・レビューだけではなく客観的な指標を用いることも要請されてくる。それは仕方のないことではあるが、だからといって指標に踊らされること、ましてや指標上の評価が高くなるような研究を意図的にしたり、自ら指標を上げようと操作することに時間を費やすことがないようにしたいものである。大学ランキングのランク向上を意図して、集計対象となる雑誌に論文を投稿することを要請するという、耳を疑いたくなるような大学の事例も聞く。オルトメトリクスまでもその列

に加わって、「研究の評価を上げるために、Facebookで友達をいっぱいつくって、たくさん『いいね!』を集めましょう」などと大学や雑誌編集部に要求されるようになったりしては目も当てられない。「普通の論文の評価には使いものにならないけど、面白い論文を見つかったり、自分の論文の話題度を知ることはできるな」くらいの位置づけがオルトメトリクスのちょうどよい立ち位置であり、無視すべきものではないが、過度に重視することもないようにしたい。

参考文献

- 1) 逸村 裕, 池内有為, 化学, **68** (12), 32 (2013).
- 2) オルトメトリクスの定義が次のサイトで掲げられている: J. Priem, D. Taraborelli, P. Groth, C. Neylon, "Altmetrics: A manifesto, 2010-10-26" (<http://altmetrics.org/manifesto/>).
- 3) ImpactStoryのホームページ (<https://impactstory.org/>).
- 4) Altmetric社のホームページ (<http://www.altmetric.com/>).
- 5) Ceek.jp Altmetricsのホームページ (<http://altmetrics.ceek.jp/>).
- 6) CiteULikeのホームページ (<http://www.citeulike.org/>).
- 7) Mendeleyのホームページ (<https://www.mendeley.com>).
- 8) L. Bornmann, *J. Informetrics*, **8**, 895 (2014).