



オープンアクセスは被引用率を上昇させるのか － Scopus による京都大学の 生命科学系論文の調査から見えるもの－

Investigation about relationship between Open Access and the number of citations in the field of life science in Kyoto University

坂 本 拓*

抄 録

オープンアクセスを推進していくためには研究者の協力が不可欠であるが、オープンアクセスによって論文の被引用数が増加するということが示せれば、協力を得る際に大変有効であると考えた。そのため、京都大学所属研究者が2014年に発表した論文を Scopus で抽出し、DOAJ とマッチングすることによってそこからオープンアクセスになっている論文を特定して、被引用数の数がどのようになっているのかを調査した。また分野間での被引用数の平均値が異なるため、その中から生命科学分野のみに限定した調査も行った。結果としては、生命科学分野は、他分野に比べ平均被引用数が多く、またオープンアクセスになっている割合も高かった。生命科学分野におけるオープンアクセス論文とそうでない論文を比較した場合、明らかにオープンアクセスの論文の方が平均被引用数が多かった。また、一度も引用されない論文の割合もオープンアクセスの論文は、他に比べて低かった。

目 次

1. はじめに
2. 調査方法
3. 調査の結果
4. 結果の分析
5. 最後に

○ 1. はじめに ○

オープンアクセス。昨今、この言葉を耳目にする機会が以前に比べて非常に増えたように思う。これは、国内でオープンアクセスポリシーを制定する機関が増加してきたこともあり¹⁾、図書館以外の分野でもオープンアクセスが話題になることが増えてきたことの表れだと思われる。ただ、

オープンアクセスは、第1には研究者の研究の自由のためのものであり、また研究者の協力無くしては絶対に実現されないものである。日本ではこの肝心の「研究者」にオープンアクセスは果たしてどれだけ理解され協力を得られているだろうか。

近年、研究者（本稿では大学所属の研究者についてのみ言及する）は、研究費を自力で獲得する必要性が以前に比べ各段に高まった。また大学の運営そのものが以前よりも様々な面で厳格さが求められるようになったため、各種の学内委員会が設置され、その委員職を担当することも求められている。このような状況の中で今日の研究者は、各種文書の作成や委員会の出席等、本来の職務である教育・研究以外に多大な時間が奪われている。

*さかもと たく 京都大学附属図書館 2016年6月6日受理

先述のように、オープンアクセスは研究者の理解・協力無くして実現されないが、このように多忙を極める研究者をオープンアクセスに巻き込んでいくためには、オープンアクセスは、本当に研究者のためのものであり、研究者にとって有益である、ということを示す必要があると考える。

研究者、特に自然科学系の研究者においては、自身の論文の被引用数（Citation）が、研究業績の評価において重要な要素となる。また業績評価という観点以外でも、自身が労力を払って執筆した論文が、他の研究者に多く引用されて研究分野の進展に貢献できていると感じられることは、研究者として大きな喜びであろう。

このため、オープンアクセスにすることによって、自身の論文の被引用率が上昇することを示すことができれば、研究者からのオープンアクセスへの協力がより得られると考えた。このオープンアクセスと被引用率の関係については、すでに何本もの論文が書かれているが、対象分野や観点によって、論文内での結論はまちまちとなっている。SPARC Europe が、このオープンアクセスと被引用率について執筆されている論文70本を集約した Web サイトを構築しており²⁾、それによると、オープンアクセスによって被引用率が上昇すると結論付けている論文は70本中46本、上昇しないという結論の論文は17本、どちらとも言えないとする論文は7本であった。つまり、65%以上の論文がオープンアクセスと被引用数増加の関係を認めていることになる。ただこれは全て海外の研究論文である。実際に学内の研究者に向けてよりリアリティを持ってオープンアクセスによる被引用率上昇の効果を感じてもらうには、京都大学の研究者が発表した論文については実際にどのような状況になっているのかを、具体的数字で示す必要があると考えた。本稿ではこの京都大学の研究者が発表した論文の被引用率が、オープンアクセスになっていることで、どのような変化があったのか、調査の手法と結果を紹介するとともに、その結果に対する考察を論ずる。

○2. 調査方法○

具体的な調査方法としては、論文引用データベースである Scopus を用いた。Scopus 内の京都大学の研究者が2014年度に発表した論文のうち、オープンアクセスジャーナルに掲載されたものと、そうでないものの間で、被引用数にどのような差異があるのかを明らかにしようと考えた。オープンアクセスにはゴールド・オープンアクセスとグリーン・オープンアクセスの2種類があり、論文の発表時からオープンアクセスジャーナルに掲載されたものが前者であり、後者は論文の発表時にはオープンアクセスになっていないが、機関リポジトリ等でその後オープンアクセスになっているものである。今回は前述のとおり、ゴールド・オープンアクセスの論文を対象とした。グリーンについては、データベース等で網羅的に把握する方法が現時点では存在しないと思われるためである。

論文引用データベースとしては Scopus 以外にも Web of Science が著名であるが、今回 Scopus を使用したのは、京都大学ではすでに Scopus を用いて APC（Article Processing Charge）の算出を行っていた経験があり³⁾、操作にある程度習熟していた、というのが理由である。

「京都大学の所属研究者が2014年に発表した論文」を抽出した条件を以下に記す。最初に著者の所属機関の名称を“Kyoto university”、そして出版年を“2014-2014”で検索をする。しかしこの検索結果には例えば“Kyoto Prefectural University”など、京都大学以外の研究者の所属論文が含まれている。また実際に京都大学所属研究者の発表論文であっても、所属が“Kyoto University”以外のパターンも存在する。そのため、検索結果一覧の中から、ファセットにより、京都大学の所属に該当する“Kyoto University”、“Kyoto University Faculty of Medicine”、“Kyoto University Hospital”の3つのみに絞り

込む。

ただし昨今の研究スタイルの動向として共同研究が非常に盛んに行われているため、1つの論文が、複数の著者により共同執筆されていることが一般的である。分野によってはこの著者の数が、数百人になることもあり、その中の1人が京都大学の所属研究者である論文を「京都大学所属研究者が発表した論文」と見なすことが妥当かどうかやや疑問に思われた。検索結果の中から「京都大学所属研究者の発表した論文」としてより適切なものを絞り込む方法の1つとして、First Author（第一著者）が京都大学になっている論文を抽出することが考えられる。しかし第一著者は、必ずしもその論文に最も貢献した人物が表示されているとは限らない。論文の執筆に最も貢献したベテランの研究者が、その研究分野の将来を担う若手研究者を立てて第一著者に持ってくることも多いため、あまり妥当性が無いように思われた。次の選択肢として、出版社等と論文の査読等のやり取りの窓口となる、その論文の責任著者とも言える“Corresponding Author”が京都大学の研究者である論文を対象とすることを考えた。これについては特に問題点は考えられず、最も妥当性があると考えられたため、この方法を採用することとした。

上記により「京都大学所属研究者が2014年に発表した論文」を抽出できた。なおこの作業は全て2015年8月16日に行ったので、被引用数等の数値は全てこの時点でのものであるということを断っておく。ここからさらにオープンアクセスジャーナルに投稿された論文を抽出する必要がある。この作業については、オープンアクセスジャーナルの情報が網羅的に収集されているDOAJ(Directory of Open Access Journals)⁴⁾から、格納されているオープンアクセスジャーナルのデータをCSVファイルにて出力し、先ほどのScopusのデータと、ISSNおよびEISSNをキーとしてマッチングを行うことで、オープンアクセスジャーナルの同定を行った。

○3. 調査の結果○

今回の調査の母数である「京都大学所属研究者が2014年に発表した論文」（以後データ群A）の数は、3,176論文であった。これらデータ群Aの被引用数の合計は、6,256件であり、1論文が平均して獲得している被引用数は、1.969であった。それに対し、DOAJとのマッチングで判明したオープンアクセスジャーナル掲載論文（以後データ群B）の数は326件であり、被引用数の合計は779であった。つまりデータ群Bの1論文が平均して獲得している被引用数は、2.389であった。

以上からデータ群Aよりもデータ群Bの方が1論文あたりの平均被引用数は明らかに多いことになる。しかし、このデータ群Bの方が平均被引用数が多い、という結果にオープンアクセスであること以外の要素が関係していないか、今一度慎重に検討する必要がある。引用という行為は学問分野によって様々な文化的差異を持っており、一般的に生命科学分野は他分野に比べて、論文内で文献をよく引用する文化がある。これはインパクトファクターの平均値は生命科学分野が最も高いということにも表れている⁵⁾。そしてまた生命科学分野はオープンアクセスがよく浸透している分野であるとも一般的に言われている。もしこれが正しければ、先のデータ群Bの方が平均被引用数が多いという結果は、被引用数が平均的に高い生命科学分野の占める割合がデータ群Aよりも多いことに起因しており、純粹にオープンアクセスの論文は被引用率が高い、と言えるわけではないのではないだろうか。

この疑念を払拭するため、つまり分野間の平均被引用数の差異という要素を除外するため、データ群Aとデータ群Bといったような全分野間の比較をするのではなく、特定の分野に限定して、その中でオープンアクセスの論文とそうでない論文でどのような差異があるのかを調査する必要がある。

あると考えた。Scopus では、「ライフサイエンス」「理学／化学／工学」「ヘルスサイエンス」「人文社会」の4つの大まかな主題分類がなされている。この中から、「ライフサイエンス」つまり生命科学を対象とすることとした。理由としては一般的に生命科学は頻繁に引用が行われる分野であるため各論文の被引用数が多く、オープンアクセスとそうでない論文との差異がより顕著に表れていると考えたためである。

結果として、Scopus 内で「ライフサイエンス」に分類されている2014年の京都大学研究者が Corresponding Author になっている論文(以後データ群 C)は1,101件であった。このデータ群 C が獲得した被引用件数の総数は2,819件であり、1 論文あたりの平均被引用数は2.560となった。それに対し、データ群 C の中から DOAJ 収録のオープンアクセスジャーナルに投稿された論文(データ群 D)の総数は149件であり、これらが獲得した被引用数の総数は549であった。ここからデータ群 D の1 論文あたりの平均被引用数は3.561となる。

○4. 結果の分析○

オープンアクセスでないものも含む全分野の論文であるデータ群 A と、同様にオープンアクセスでないものも含む生命科学分野の論文データ群 C を比較した場合、1 論文あたりの平均被引用件数は、1.969と2.560であった。また、オープンアクセスである全分野の論文であるデータ群 B とオープンアクセスである生命科学分野の論文であるデータ群 D を比較した場合は、2.389と3.651となっており、オープンアクセスであるかどうかに関わらず、やはり生命科学分野は、他分野に比べてよく引用される傾向にあることが明らかになった。

またデータ群 A に占めるデータ群 B の割合、つまり全分野の論文にしめるオープンアクセス論文の割合は、10.264%であったのに対し、データ

群 C に占めるデータ群 D の割合、すなわち生命科学分野の論文に占めるオープンアクセス論文の割合は、13.533%であり、やはり生命科学分野は、他に比べてオープンアクセスになっている論文の割合が高いことが確認できた。ここから、データ群 A とデータ群 B の比較ではなく、データ群 C とデータ群 D の比較、つまり特定の分野に限定したオープンアクセス論文とそうでないものの比較に切り替えたことは妥当であると考えられる。

そしてそのデータ群 C とデータ群 D を比較した場合、データ群 C の平均被引用数が2.560であったのに対し、データ群 D の方は3.651となっていた。つまり、生命科学分野においても、オープンアクセス論文は良く引用される傾向にあることが明らかとなった。

また、補足的な分析ではあるが、オープンアクセスになっている論文とそうでない論文で、「1度も引用されない割合」がどのようになっているのかも比較した。生命科学分野の全論文であるデータ群 C は、総数1,101に対し、1度も引用されていない論文の数は456であり、論文が1度も引用されない割合は41.416%であった。それに対し、生命科学分野のオープンアクセス論文であるデータ群 D は、総数149に対し、1度も引用されていない論文数は17であり、1度も引用されない割合は11.409%となり、データ群 C よりも大幅に低い数字となっている。以上からオープンアクセスになっている論文は、一度も引用されないという最悪の事態を回避できる可能性が高いことが明らかになった。この点からもオープンアクセスにすることの有効性は高いと言える。

○5. 最後に○

上記のとおり、今回の調査によって、生命科学分野は他の分野に比べて平均被引用数が多く、またオープンアクセスになっている論文の割合も高いことが判明した。その生命科学分野に限定し

た、全論文とオープンアクセスジャーナル掲載論文の比較を行った場合、後者の方が、平均被引用数が多いことが明らかになった。さらに一度も引用されない論文の割合についても、オープンアクセスになっている論文の方が低いことが判明した。

今後の課題としては、調査結果の有意差をもう少し厳密に検証する必要があることが挙げられる。また今回の調査はあくまでもオープンアクセスジャーナル掲載論文についての調査であり、数が少ないとはいえハイブリッドタイプのオープンアクセス論文についての考察が不足している。ハイブリッドタイプを網羅的に把握することは困難であるが、これについても今後方策を検討する必要がある。また、やはりグリーン・オープンアクセスになっている論文についても、非常に重要な要素になるため、把握するための可能性を今後も追及する必要がある。最後に、今回の調査は、2014年出版分という調査時点での直近の年の論文を対象としたが、時間経過とともにどのような変化が生じるのか、という点も調査する必要があると考える。以上を今後の調査の課題としたい。

注・参考文献

1. 「京都大学オープンアクセス方針」を採択しました. 京都大学図書館機構. <http://www.kulib.kyoto-u.ac.jp/bulletin/11677> (accessed 2016-05-31).
2. "The Open Access Citation Advantage Service". SPARC Europe. <http://sparceurope.org/oaca/> (accessed 2016-05-31).
3. 鈴木秀樹ほか. 平成27年度学内オープンアクセス費支出状況調査報告書. KURENAI. <http://repository.kulib.kyoto-u.ac.jp/dspace/handle/2433/210594> (accessed 2016-05-31).
4. Directory of Open Access Journals (DOAJ). <https://doaj.org/> (accessed 2016-05-31).
5. Journal Citation Reportsにおける "Categories By Rank" のデータによる。