

Title	研究者の国際移動の追跡 : 高被引用論文生産形態の特徴
Author(s)	山下, 泰弘; 吉永, 大祐
Citation	年次学術大会講演要旨集, 27: 735-738
Issue Date	2012-10-27
Type	Conference Paper
Text version	publisher
URL	http://hdl.handle.net/10119/11125
Rights	本著作物は研究・技術計画学会の許可のもとに掲載するものです。This material is posted here with permission of the Japan Society for Science Policy and Research Management.
Description	一般講演要旨

研究者の国際移動の追跡 —高被引用論文生産形態の特徴—

○山下 泰弘, 吉永 大祐 (山形大学)

1. はじめに

計量書誌学は、科学技術政策研究における重要なツールであるが、研究者の経歴（移動歴）や取得した助成金と研究アウトプットの関連性の把握にはこれまで十分に活用されてこなかった。それは研究者のプロフィールや、謝辞などの情報が書誌事項に含まれないことによる。前者は ResearcherID 等の普及により長期的に解消される可能性があるが、まだ利用者が多くないのが現状であり、後者も近年 Web of Science のレコードにおける謝辞が索引されるようになったことにより、新しいデータについては研究に使用された助成金の分析が可能になっているが、現時点では古いデータに対する網羅的な付与はなされていない。

そこで、本研究では、上記課題に取り組むために、ウェブ上で公表される CV や Google Scholar で提供される論文情報を取り入れた計量書誌学分析手法を試行し、その結果を評価する。高被引用論文著者の情報をウェブ上の CV より抽出して実施する分析については、2011 年にすでに報告しているが、引用されない論文にまで対象を拡大したときに同様の精度が得られるかは定かではない。本研究では、研究者の国際移動の論文生産への影響と当該研究に使用された助成金の把握を、高被引用論文と引用されない論文について試み、両者の比較を行う。それを通じて、研究者の経歴及び助成状況についての分析方法の開発・評価を行うことを目的とする。

2. 使用するデータと方法論

本研究では、2011 年 1 月に” Web of Science ” よりダウンロードした” Computer Science, Artificial Intelligence ” 分野の 2004～2006 年の論文データに基づき、被引用上位 1%論文と被引用数 0 の論文（ゼロ被引用論文）を抽出して使用する。前者は対象論文すべてを使用するが、後者については数が多いため、全論文の 1%に相当する数をランダムに抽出して使用する。文献種は” Article ” のみを対象とする。それぞれの論文について、インターネット上で公開されている CV 等を活用して著者の経歴を抽出する。論文数の計算には、各論文を著者数で按分する分数カウントを使用する。論文数は、A 国の機関に属する研究者による論文数 $N_i(A)$ 、A 国出身の研究者による論文数 $N_o(A)$ 、B 国出身者が A 国で生産した論文数 $N_c(A, B)$ について集計を行う。著者の出身は、学士号ないし相当学位取得国ないし出生国とする。また、欧州大陸国については、学士号取得国が記載されておらず、修士号ないし相当学位の取得国のみ記載されているケースは、修士課程までの一貫課程と見なす。一方、研究で使用された助成金についての情報は、Google Scholar ないし、山形大学及び早稲田大学で利用可能な電子ジャーナルに個別に当たって抽出し、資金助成者については whole カウントを用いて集計する。

3. 分析結果

3.1 研究者の国際移動

(1) 著者出身の判明状況

高被引用論文 140 件のうち分数カウントで 116.4 件 (83.2%) について著者出身国が判明したが、ゼロ被引用論文 137 件については 85.4 件 (62.3%) に止まった。

著者の所属機関に基づく高被引用論文の発表国は、上位から米国、中国、英国、台湾、フランスの順であり、中でも米国は 60 件超と量的にも突出している (図 1)。一方、ゼロ被引用論文の発表国を見ると、上位からロシア、米国、カナダ、イタリア、中国の順となり、高被引用論文生産の場合と比較して、特定の国による寡占度は低い。また、著者所属の内訳をみると、高被引用論文は、いずれの国の論文も大半において著者所属が判明しているが、ゼロ被引用論文では、概して低く、特にロシアについては大半の著者についての情報が得られていない。ただし、米国、カナダ、イタリア等については、比較的良

好に情報が得られている。

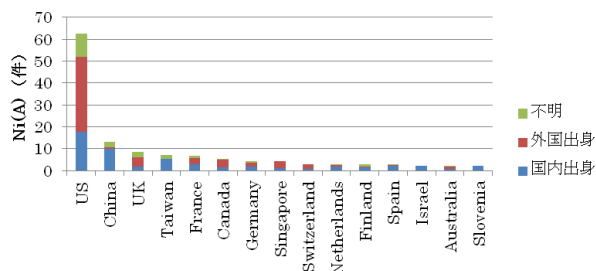


図1 各国の高被引用論文の著者内訳

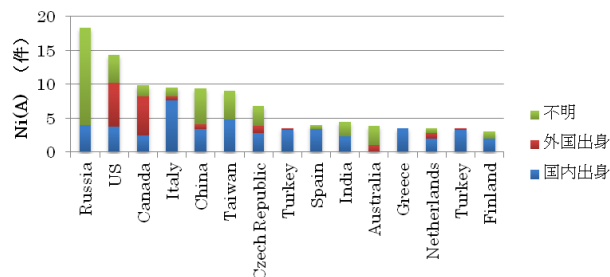


図2 各国のゼロ被引用論文の著者内訳

(2) 高被引用論文著者とゼロ被引用論文著者の国際移動状況

次に、高被引用論文とゼロ被引用論文の著者について、国際移動の状況を概観する。前述の通り、ゼロ引用論文著者についてのデータは捕捉率が低く、かつ系統的に捕捉されにくい国が存在するため、必ずしも正確なデータとは言えないが、捕捉可能な範囲での結果に基づき考察する。

高被引用論文著者の国際移動が各国の論文生産に与える影響を見ると、中国・インド出身者が米国に渡航して発表した論文が多いことがわかる（図3）。両国とも欧米を中心に多くの研究者を供給する人材大国であるが、中国は自国においても米国に次ぐ高被引用論文を生産しているのに対し、インドは国内での高被引用論文生産は極めて限定的である。

一方、ゼロ被引用論文生産への研究者国際移動の影響を見ると、高被引用論文と比べて異動した研究者による論文数も少なく、移動パターン（線）の数も図3と比べて少ない（図4）。Ni(A)は、図2にも示した通り、高被引用論文を生産していないロシアが最多であるが、他国出身者の流入は、出身情報が把握されている範囲では見当たらない。また、高被引用論文生産においては、他国への研究者流出が目立つインドは、ゼロ被引用論文については大半を自国で生産している。また、日本、ブラジル、イランなど高被引用論文を生産していない多くの国がゼロ被引用論文生産に関与している。

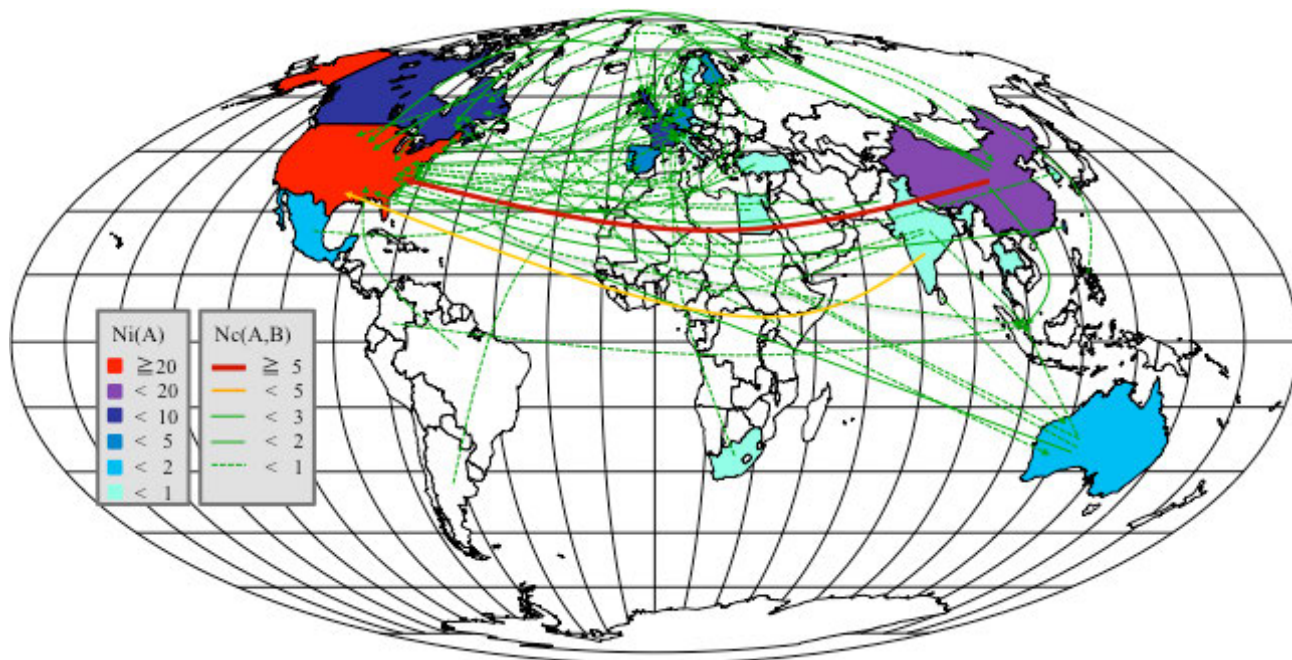


図3 高被引用論文生産への研究者国際移動の影響

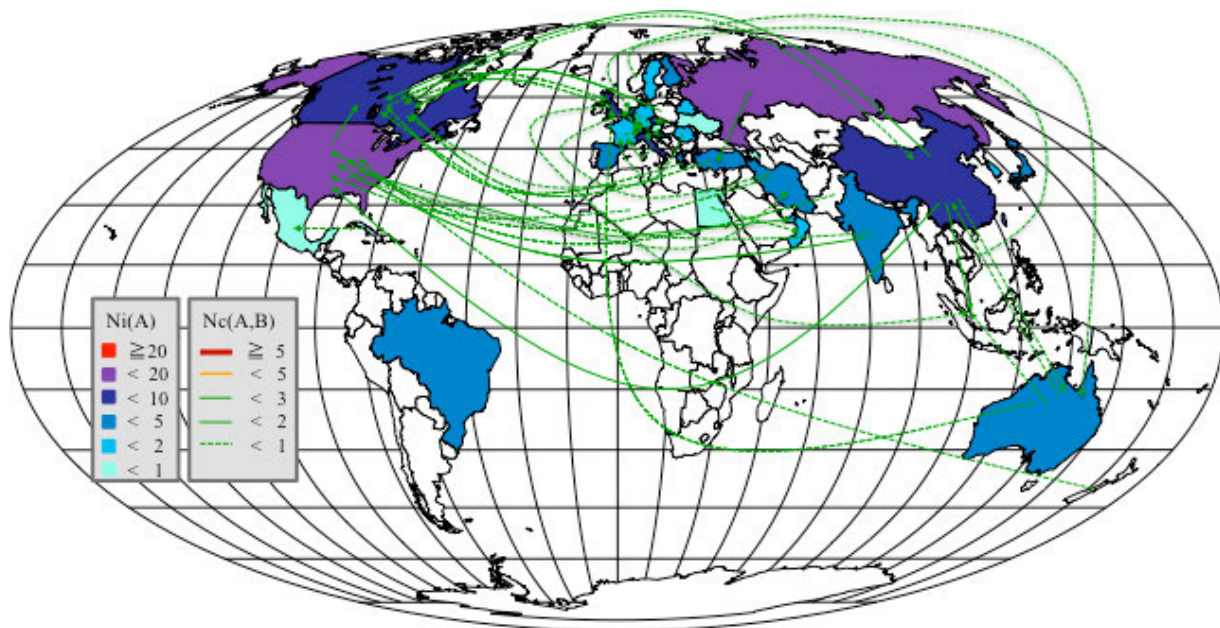


図4 論文生産への研究者国際移動の影響

(3) 各国出身者の状況

次に各国出身者がどこで論文を生産しているかを分析する。ここでは、各論文のうち著者出身の判明したもののみを集計するため、出身不明者の動向次第で結果は大きく変わる可能性がある。

高被引用論文では、米国・台湾など少数の例外を除き、人材輩出上位国の多くが国外に人材を供給している。一方、ゼロ被引用論文では、中国・インドを除き論文の大半が国内で生産されている。中国は高被引用論文もゼロ被引用論文もほぼ同比率で国内生産しているが、インドは高被引用論文のほとんどを国外生産する一方で、ゼロ被引用論文は半数程度国内生産している。双方ともアジアの人材供給大国であるが、中国は質の高い研究を国内で実施する体制がより整っていると考えられる。米国は自国出身者のほとんどが高被引用論文を国内で生産しているが、ゼロ被引用論文は一定の割合で国外生産されている。すなわち、米国では、国内外の優秀な研究者を集積してインパクトの高い研究を実施する体制が確立されていることが示唆される。

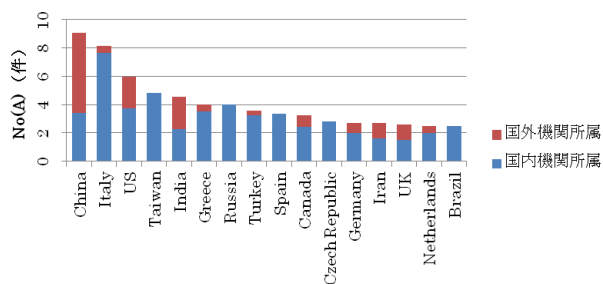
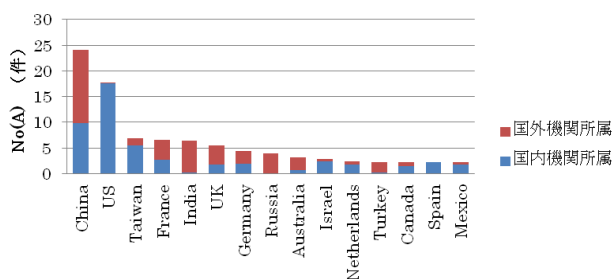


図5 高被引用論文における各国出身者の所属 図6 ゼロ被引用論文での各国出身者の所属

3.2 助成金情報の把握

Google Scholar と 2 大学図書館の併用により、高被引用論文については 99%、ゼロ被引用論文については 70% の論文を収集することができた。助成金情報の記載状況については、ゼロ被引用論文原報の 30% が未入手であるにも関わらず、非記載割合が高被引用論文と 10% ポイント程度しか差がないことから、未入手論文の情報まで含めれば両者は逆転するものと思われる。一方、国外から流入した研究者と国内出身研究者の、助成金の記載状況を見ると、国内出身者は高被引用論文とゼロ被引用論文で記載率にあまり差がないが、国際移動者については高被引用論文 46% に対し、非移動者 24% とほぼ倍の開きがある。すなわち高被引用論文を生産するプロジェクトでは、国際移動者にも直接的・間接的な支援がなされて

いるのに対し、プレゼンスの低いプロジェクトでは国際移動者への支援が十分になされていない可能性がある。

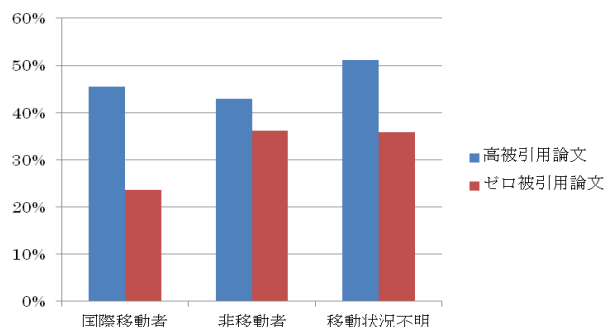
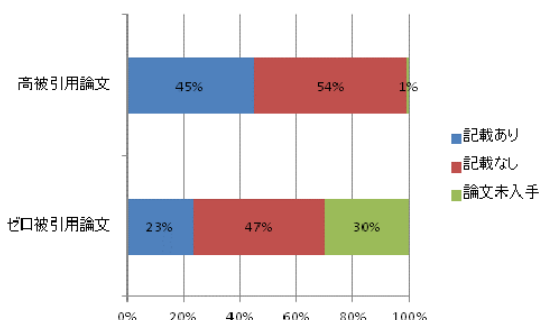


図6 助成金情報の謝辞への記載状況 図7 国際移動状況別の助成金情報記載率 (未入手論文を除いた集計)

4. まとめ

特定の分野についての限定された分析であり全分野で同精度での結果が得られることを保証するものではないが、少なくとも、高被引用論文については著者の経歴に基づく分析がある程度可能であることがわかった。それは、分野の動向に強く影響する一部の論文については、CVに基づく分析が有用であることを示唆するものである。ただし、引用されない論文については、系統的に情報を把握しにくい国やジャーナルがあり、分析結果に強いバイアスを与えている恐れもある。高被引用論文のみではなく、全論文から無作為抽出する場合などには、特に注意が必要となる。

また、Google Scholar と電子ジャーナルを用いることにより、現状の WoS では不可能である遡及的な謝辞分析が、高被引用論文については可能となることも示唆された。したがって、(論文生産の面で)成功したプロジェクトについては、発表論文を検索して複数プロジェクトを比較評価できる可能性があるが、それについては別途検証を行う必要がある。また、本稿では、単に助成金についての記載の有無のみ取り上げたが、今後は国や助成機関のセクタごとの傾向についても分析を進める必要がある。

本研究では、あくまでも高被引用論文の特性を見出すために比較対象としてゼロ被引用論文を取り上げた。それにより高被引用論文の特性把握についてはある程度目的を達したが、特定の国で集中的に生産されないゼロ被引用論文は多様であり、性質については不明な点が多く、その評価を行う際には例えば以下についてより深く理解する必要があることも同時に示唆された。

- (1) 「WoS で引用されない」からと言って当該論文が真に引用されていないのか、収録対象外の雑誌論文で (多く) 引用されているのかは不明であり、その評価には注意が必要である。本研究において、ロシアは、高被引用論文がなく、もっとも多数のゼロ被引用論文を特定のジャーナルで発表していることが判明したが、このようなジャーナルが国内でどのような位置づけにあるのかは、WoS のみで評価することは危険と思われる。
- (2) ゼロ被引用論文の著者情報は相対的に得にくいだが、その要因としては、高被引用論文著者よりもプレゼンスが低く、研究職についていない者が多い可能性もあるが、その一方で著者の居住国の発信情報が少ない、英語でないために発信された情報が捕捉されていない、などの地域・言語バイアスの影響も否定できない。
- (3) ゼロ被引用論文では、助成金情報の記述が相対的に少ないが、助成金を得ていないのか、ブロックファンドで運営されているなどの要因で記載されていないのかは不明である。

著者の経歴に基づく分析は、通常の計量書誌学研究では得られない多くの示唆を与えるが、(a) 多大な人手を要すること、(b) CV を得やすい分野と得にくい分野があること、(c) すべての公開された CV から共通の情報が得られるとは限らない点がボトルネックとなる。極端に論文当たりの著者数が多くなく、かつ研究者情報発信の少ない産業セクタ中心ではない分野を注意深く選択することで、(a) と (b) については多少解決されるが、より大規模な調査を実施するためには、データの体系的整備と処理の自動化が不可欠である。