

指標の注意点

清家 弘史 (東北大学 研究推進・支援機構 特任准教授)

I 背景

研究力の分析、および評価のための定量的な指標 (メトリクス)が注目を集めている。メトリクスの中でもとりわけ、計量書誌学 (ビブリオメトリクス)¹⁾が取り上げられることが多い。その背景には、英国の REF²⁾に代表される政府、助成金機関による大学の評価や、様々な世界大学ランキング³⁾において、発表論文に関するメトリクスが取り入れられていることがあげられる。日本においても科学技術政策研究所 (NISTEP)⁴⁾が定期的に報告する、「研究論文に着目した日本の大学ベンチマーキング」⁵⁾や「サイエンスマップ」⁶⁾などは多くの大学や研究機関が、評価の対応や戦略の立案のために必ず目を通す資料となっている。ここでは、メトリクスは客観的な (objective) エビデンスに基づいた評価や意思決定を行うための重要な要素の一つに位置付けられている。一方、技術的な革新もメトリクスの使用を拡大し、さらに加速させている。実際、書誌情報を基にしたビブリオメトリクスを用いた研究生産性の議論の試みは 1920 年代にはすでに行われているようであるし、現代ビブリオメトリクスの父と呼ばれる Eugene Garfield が Science Citation Index の発行を開始したのは 1964 年であり、ビブリオメトリクスにはすでに長い歴史があるといえる。実際、我々が現在、頻繁に議論している「問題点」についてもすでに発見され、指摘されてきたものも少なくない。⁷⁾ 今日見られる発表論文に関するメトリクスに対する関心の大きな広がりとはまりは、Web of Science⁸⁾や Scopus⁹⁾のようなオンライン抄録・引用文献検索データベースが研究の場に導入され、大量のデジタルデータを迅速に扱えるようになった 2000 年以降の動向といえよう。大量のデジタルデータを PC 等で迅速に扱えるようになったことで、抄録・引用文献検索用のデータをもとに、メトリクスの計算や引用一被引用の解析を自分の手で行えるようになったからだ。さらに、上述のデータベースに収録されているデータを基にした、InCites¹⁰⁾や SciVal¹¹⁾のような分析ツールが開発され、メトリクスへのアクセスがより迅速で容易になったように見える。InCites や SciVal のような分析ツールは、研究機関や研究分野のようにあらかじめ標準設定された論文集合 (あるいは、論文リスト) に関するメトリクスのデータベースであるだけでなく、分析者が作成した論文集合に対しても同様のメトリクスを計算してくれる計算機でもあり、さらに論文を絞り込む篩でもある。したがって、メトリクスのデータベースである分析ツールを購読すればワンクリックで客観的なエビデンスを入手することができ、エビデンスに基づく意思決定が可能となると考える向きも少なくないが、自信をもってそう言い切れる方ばかりではないと個人的には推測する。私が 2015 年 6 月に英国ブライトンで開催された

ARMA 年会のメトリクスに関するセッションのパネルディスカッションに参加した際には、会場にいた参加者の半分くらいがメトリクスに関してはその意味や使い方、今後の用いられ方に対して確信が持てないでいる (skeptical) と挙手で表明したようにパネル席からは見えた。ピアレビューの伝統が確立された英国で、初めてビブリオメトリクスを参考として使った REF2014 の準備、提出、そして結果の発表と一連の作業が終わった後の、「さあ、次の REF にむけてどうするか？」という様々な思いの率直な表れであると感じた。それから二年経った、2017 年 6 月に英国を訪問した際には、「メトリクスはなくなる。真剣に取り組む必要がある。」という雰囲気が変わっていることを感じた。SciVal のような分析ツールを契約した大学が増えていることや、メトリクスを使って分析する体制をこれから作ろうとしている大学も増えていることを意見交換の際に見聞きした。Snowball メトリクスや REF2014 でのメトリクスの使用でメトリクス先進国のように言われる英国でさえも、必ずしもすべての大学でメトリクスを用いた分析や評価が自信をもって行われているわけではない。むしろ、REF2014 の結果を受け、Snowball メトリクス¹²⁾の編集にかかわった先進的な八大学だけでなく、助成金機関も含めた幅広いメトリクス使用者 (stakeholder) が、本来の目的に合致した、分別ある賢いメトリクスの使用法 (responsible use of metrics) についてこれから議論を進めていこうとしているというのが現状と言えよう。

この章では、東北大学研究推進・支援機構 URA センターにおいて 2013 年 7 月から発表論文と外部研究資金の獲得状況について分析を行った経験を踏まえながら、「分析における Golden Rule」(エルゼビア社) や 2015 年に発表された二つのレポート: “The Metrics Tide”¹³⁾ と “The Leiden Manifesto for Research Metrics”¹⁴⁾ から、メトリクスを用いる際の注意点を解説する。

II メトリクスへの期待、および疑問: メトリクスの実際

研究力の現状の分析を客観的なエビデンスを用いて行い、将来的な研究戦略を立案し、PDCA サイクルをまわすことで研究力の向上を目指すことが強く望まれている。従来は教学と財務に関するデータに主眼を置いた情報収集・分析活動を指す IR という言葉を派生させて研究 IR という言葉が作られ、研究大学コンソーシアムにおいても研究力分析の WG が編成されるなど、研究力の分析や評価における客観的なエビデンスの使用に対する期待、および興味は大きい。一方、「数字には踊らされたくない。」など、メトリクス自体に対する不快感や個別のメトリクスに対して疑問を呈する声も多い。この相反する二つの現象は、メトリクスを用いて研究力の現状を議論することに対する戸惑いを示すもので、その要因として、日本国内において研究活動をカバーする標準的なメトリクスの一覧や、メトリクスの使用に関する原則のようなものが存在しない、あるいは浸透していないからであろう。海外においても似たような状況にあり、そのよ

うな状況を打破するべく、様々な活動が繰り広げられている。そのような動きの一つとして、英国における **Snowball Metrics Project** を紹介する。

Snowball Metrics Steering Group は、英国の八つの大学（ユニバーシティ・カレッジ・ロンドン、オックスフォード大学、ケンブリッジ大学、インペリアル・カレッジ・ロンドン、ブリストール大学、リーズ大学、クイーンズ大学ベルファスト、セント・アンドルース大学）によりボトムアップの形式で結成されたグループで、エルゼビア社が協力する形で、大学間のベンチマーキングを可能とするような標準的なメトリクスセットを作成することを目標としている。**Snowball Metrics Steering Group Chair** の **John Green** 氏の話では、八大学が協力関係を結び、データの収集から実際のメトリクスの定義と計算を共同で行いながら標準的なメトリクスセットを編集し、レシピブック¹⁵⁾として公表している。**Snowball Metrics Steering Group** のまとめたレシピブックからメトリクスのテーブルを見てみると、いくつか興味深いことがある。（表1、章末を参照）まず、ビブリオメトリクスだけがメトリクスではないということである。研究活動サイクルの四つのステージ（**Input, Process, Output, Outcome**）に注目し、研究、産学連携活動、大学院学生の活動の三つのカテゴリーについて、多彩なメトリクスを候補として挙げている。研究カテゴリーにおいてはビブリオメトリクスと獲得研究資金の件数と金額のような伝統的なメトリクスを筆頭に、**Altmetrics** のような比較的新しいメトリクスやアカデミーのメンバーの数のように人材や評判に関わるものなどもメトリクスとして挙げている。興味深いのは、大学院生に関する情報をカテゴリーの一つとして設けていることである。**Snowball Metrics Steering Group** が大学院生を研究戦略の重要な柱の一つとして認識していることを見て取ることができる。研究活動を担うのは、多数の学生や博士研究員たちであり、今回の我々のプロジェクトでも **Active Author** として、研究成果に貢献する人的リソースに関するメトリクスをとりあげた。人事データと組み合わせることで、新たな知見が得られる可能性がある。

さて、**Snowball Metrics Steering Group Chair** の **John Green** 氏は2016年10月25日に開催された「研究力を測る指標の抽出と大学の研究力可視化に関する基礎的研究 国際キックオフシンポジウム」において、“**First define the question; then pick the metrics to answer them.**” と述べている。明確な課題を設定したうえで、適切なメトリクスを選択し、ベンチマーキングの手法を戦略的に使って研究力を向上させた例として、**Academic Ranking of World Universities (ARWU)** を挙げるができる。1998年5月4日の北京大学100周年記念の式典での当時の国家主席・江沢民のスピーチを受け、211工程、985工程と中国の大学を **World Class University** に引き上げる政策が進められた。この政策に対応し、独自のビジョンを描くために、上海交通大学は **World Class University** の特徴についてまずは調査をしたという。上海交通大学は、**World Class University** と言われる大学の特徴と、中国の大学の当時の状況を比較し、発表論数が著しく少ないことを認識し、まずは発表論文を増加させる必要を感じた。この分析

を踏まえ、発表論文に関係するメトリクスと人材に関するメトリクスを合計 5 つ選び、彼らの描く **World Class University** への道筋を描いた。ARWU は 2003 年に公開されて以来、用いるデータは公開、あるいは商業データのみで公平性を保ち、用いるデータソースや、メトリクスの種類、算出法を変更することなく一貫性を保っているため、経年変化を読み取ることが可能である。2003 年の時点では 200 位に入る中国の大学はゼロであったが、2017 年の時点では九つの中国の大学が 200 位以内に入っている。ARWU の価値観に沿った、戦略的なメトリクスの使用と国際ベンチマーキングによって研究力を向上への道筋をつけることに成功したと言えよう。一方、もしも用いているメトリクスが測定したい研究力を測定していないと感じるならば、メトリクスの選択、計算方法や解釈を含む使い方について再点検をする必要がある。例えば、「国際共著率は本当に研究力の強さを反映しているのか？」という問いに対して、国際化の進んでいない研究機関にとっては海外研究機関との共著論文実績は重要なメトリクスであろうし、すでに国際的なネットワークを構築済みの研究機関にとってはあまり重要ではないメトリクスということになるだろう。国際共著論文数、あるいは率に関しては別章において議論を展開しているのでそちらを参照していただきたい。

それでは、設定した課題を議論するために適したメトリクスがない場合、どうするか？新しいメトリクスを提案したり、既存のメトリクスに新しい解釈を加えて使用をするという解決策が考えられる。実際、今回の我々のプロジェクトにおいて、我々は **H-index** に新たな解釈を適用することで、これまでの発表論文数（量）と被引用数（= 質）という二つの評価軸をバラバラに用いたのではとらえることが難しい研究力の側面を測ることを提案し、厚み指標と呼んでいる。英国人の感覚では **Research Capacity** という英語で表現できるということで、国際的にも通用する概念として今後の発展が期待できる。また、共著論文数を結びつきの強さと解釈し、**H-index** の算出方法を用いて共著機関とのネットワークの広さと結びつきの強さを同時に図る、**Collaborative Network Index (CNI)** という新しいメトリクスも提案した。海外研究機関と十分な量の国際共著論文数を持つ研究機関にとっては、ネットワークの広さと強さを知るうえで有用なメトリクスとなる可能性もある。厚み指標や **CNI** に関する詳細は別章を参照していただきたい。前述の **Altmetrics** も含め、このように新しく提案されたメトリクスや既存のメトリクスを新解釈に基づいて使用する際には、その有用性や有効性については、今後の検証されることになる。しばらくは伝統的な有用性の確認されているメトリクスの土台の上に成り立つと考えるのが適切であろう。

メトリクスを実装し、自信をもって使えるようになるには、英国の **Snowball Metrics Project** のメンバーがやったように、実際に自分の手で一連の分析の作業を経験するしかないだろう。一連の作業とは次のようなものである： グローバル、ナショナル、リージョナル、そしてインスティテューショナルな文脈の中で分析の焦点を定め

て課題を設定する。設定した課題に対応するために、比較可能で信頼性のある連続的なデータを一貫性のある形式で収集する。収集したデータを用いてメトリクスを算出し、解釈を行った後、可視化、および言語化を通じて戦略立案や評価に対応する。もしくは、算出したメトリクスを可視化、および言語化することによって、解釈を試みる。

課題設定の重要性はすでに述べたが、質の保証されたデータやメトリクスの入手には困難が伴うことが多い。データのソースとしては、組織内データ、公開データ、商業データが考えられるが、必ずしも先に述べたような望ましいような形でデータが保存されているとは限らないし、データソース間での互換性が期待できない場合も多い。ビブリオメトリクスを行う際の基盤となる **Web of Science** や **Scopus** のような抄録・引用文献検索データベースはそもそも研究力の分析や評価のために作成されたものではなく、データの漏れもある。また、**Web of Science** や **Scopus** からのデータを用いてメトリクスを計算する際も、データベースにより発表論文の収録範囲も異なることから、それぞれのデータベースの世界観が反映されたスペースでの分析となることは認識しておく必要がある。メトリクスの計算方法に関しても、市販のデータベースやツールによってまちまちである。例えば、高被引用度論文（トップ論文）を例にとると、「被引用回数を比較するには分野や文献タイプの考慮が必要である。」というアドバイスに関しては多くの方が同意されるだろう。しかし、実際には、エルゼビア社の **SciVal** でのトップ論文は分野別のノーマライゼーションは行われていない。一方で、クラリベイト・アナリティクス社の **InCites** では分野と文献タイプ別にノーマライゼーションが行われている。これは、それぞれのデータベースやツールの思想や世界観を反映しているからで、どのツールのメトリクスが正しいとは一概には言えない。データベースやツールをはじめとするデータソースの癖を理解することは、得られるデータやメトリクスの正しい取り扱いや解釈に不可欠である。実際に自分の手で一連の分析の作業を経験することを勧める所以である。

REF2014 や各種世界大学ランキングだけでなく、日常の研究力分析や評価においてもメトリクスの用いられる頻度が増すにつれ、メトリクスの使用に関するノウハウや注意点が蓄積されてきた。次の二つの節において、メトリクスの使用における一般的で、基本的な注意点を紹介する。

III 二つの黄金則 —Two Golden Rules—

メトリクスを用いる際に肝に銘じるべき「二つの黄金則」として、次のようなアドバイスをエルゼビア社が提唱している。2016年10月25日に開催された「研究力を測る指標の抽出と大学の研究力可視化に関する基礎的研究 国際キックオフシンポジウム」における、Lisa Colledge氏の発表PPTより転載する。

Two Golden Rules:

Rule 1: Always use both qualitative and quantitative input into your decision.

Rule 2: Always use more than one research metric as the quantitative input:

A basket of metrics.

それぞれについて簡単にコメントを加えておく。まず、Rule1 であるが、定性、定量、どちらの方法にも強みと弱みがあり、どちらか一方が他方を置き換えるわけではない。最終的に結論を出すのは人間であり、メトリクスは専門家の意見を客観的に補完するものである。したがって、定性的なデータの示唆する結論と定量的なデータの示唆する結論が相反する場合、新たな観点を導入し分析や解析を行うことで有益な知見を得ることが期待でき、興味深い。次に Rule2 であるが、研究活動やその成果は多彩な側面を有していることから、それに応じて多彩なメトリクスを用いて研究活動や成果を分析していくことが求められる。それぞれのメトリクスにはそれぞれの強みと弱みがあり、複数のメトリクス同士がお互いを補い合うのが望ましい。また、複数のメトリクスを使用することの効用の一つとして、メトリクスの数値の操作（Gaming）を難しくするということが挙げられる。

IV The Metrics Tide & The Leiden Manifesto for Research Metrics^{13), 14)}

Snowball Metrics Project チームやエルゼビア社とは別に、二つのグループがメトリクスの使用に関する注意を促すレポートを発表している。一つ目は英国のグループが次の REF を念頭に置きながら、広がりを見せつつあるメトリクスの使用について注意を促すものである。二つ目はオランダ、米国、スペインの多国籍のチームが世界のいたるところで見られるメトリクスの誤用がもたらす弊害を危惧して注意を促したものである。エルゼビア社の「二つの黄金則」よりも多くの注意点を挙げており、参考になるので次に紹介する。

Responsible Metrics from “Metrics Tide” Executive Summary¹³⁾

1. Robustness: basing metrics on the best possible data in terms of accuracy and scope;
2. Humility: recognizing that quantitative evaluation should support – but not supplant – qualitative, expert assessment;
3. Transparency: Keeping data collection and analytical processes open and

- transparent, so that those being evaluated can test and verify;
4. Diversity: accounting for variation by field, and using a range of indicators to reflect and support a plurality of research and researcher career paths across the system;
 5. Reflexivity: recognizing and anticipating the systemic and potential effects of indicators, and updating them in response;

The Leiden Manifesto for Research Metrics: TEN PRINCIPLES¹⁴⁾

1. Quantitative evaluation should support qualitative, expert assessment.
2. Measure performance against the research missions of the institution, group or searcher.
3. Protect excellence in locally relevant research.
4. Keep data collection and analytical process open, transparent and simple.
5. Allow those evaluated to verify data and analysis.
6. Account for variation by field in publication and citation practices.
7. Base assessment of individual researchers on a qualitative judgement of their portfolio.
8. Avoid misplaced concreteness and false precision.
9. Recognize the systemic effects of assessment and indicators.
10. Scrutinize indicators regularly and update them.

The Metrics Tide のレポートは、Snowball Metrics やエルゼビア社のレポートに比べると、メトリクスに対して消極的な姿勢が感じられるが、The Metrics Tide も The Leiden Manifesto もメトリクスは定性的なデータを補完する立場にあることを明記し、研究活動の多様性に対応できるように促している。そのほかに、データの質保証や分析の透明性、課題設定の重要性についても指摘がされている。特に、メトリクスは組織や人の行動を変える力を有しており、メトリクスのために好ましくない行動をとったりする危険性があることに警鐘を鳴らすとともに、刻々と変動する状況に応じたメトリクスの選定を促している点は注目に値する (The Metrics Tide 5, The Leiden Manifesto 9 & 10)。

V 結語

Snowball Metrics プロジェクトや Metrics Tide、そして The Leiden Manifesto の例をもとにメトリクスの注意点について議論をしてきた。結論としていえるのは、そ

それぞれのレポートにより温度差はあるが、節度を持って使えば、メトリクスはよりよい決断を導くのに有効であるといえ、最も注意すべき点は「二つの黄金則」で言い表されている。すなわち、定性的な手法と定量的な手法を両方使い、研究活動に多彩な角度から光を当てることができるように複数のメトリクス (a basket of metrics) を用いることがカギである。複数のメトリクスの設定の際に重要なことは、明確な課題をまず設定し、その課題に応じた適切な複数のメトリクスを選定することが重要であることをもう一度強調しておく。

分析の作業は総説を書く作業に非常に似ている： まず、課題を選び、視点の切り口やスコープを決める。その際に、課題を取り巻く現状や歴史的な背景に関する知識や理解があれば、総説のアウトラインを描くことは比較的容易である。総説においてあらかじめ、主張したいことや議論の展開方法が決まっていることもある。その場合は、議論の展開の流れに沿った情報を収集し、まとめていくことになる。一方、オープンな立場で課題について調査を行い、まとめようとする場合、まずは課題に関する情報をできるだけ漏れのないように収集する必要がある。視点の切り口やスコープがすでに念頭にあって情報を収集することもあれば、一旦、情報の収集を行って、視点の切り口やスコープを整理し、再度情報を収集することもあるだろう。その次に待っているのは、「収集した膨大な情報（文章、数値、チャートなど様々）を頭脳、あるいは機械で読み取り、編集、および解釈を経て、要点を抽出し、序論で設定した舞台の上で論を展開し、結論へと導く」という一連の知的な作業である。議論を展開し、結論へと導いていく過程で用いる、カギとなる道具には二つある： 可視化された情報と、言語化された情報である。可視化は数値や言語で記述された情報の解釈のための有用な手段であり、言語化は文脈を付与しながら情報を受け手に伝達する。総説の著者は総説の中で用いた情報についてはレファレンスとして開示する。そして、総説の読者は、総説を読むことで短時間のうちに対象課題の現状や問題点を認識し、問題の解決策や次の研究課題について思考する手がかりにしようとする。

上記の一連の作業になじみのある方も多いだろう。データベースやツールからメトリクスを入手し、分析レポートを作成してみるのもいいが、自分のできる範囲で、課題設定からメトリクスの選定、データの収集、メトリクスの算出、可視化、言語化、そして解釈の一連の作業をしてみると、データベースやツールのブラックボックス化している部分の一部が見え、多くのことに気が付き、分析結果の解釈を深化させることができる。その過程で気が付いたコツや注意点をできる範囲で共有することができれば、日本におけるメトリクスのリテラシーが向上し、種々のメトリクスの交換をともなう国際的なベンチマーキングが可能となると考えられる。

VI レファレンス

- 1) Bibliometrics のハンドブックの例: Todeschini, R. & Baccini A. Handbook of Bibliometric Indicators (Wiley-VCH Verlag GmbH & Co. KGaA, 2016)
- 2) REF2014 の結果: www.ref.ac.uk/2014/
REF2021 へ向けた最新の情報: <http://www.ref.ac.uk/>
- 3) 三大世界大学ランキングと呼ばれるランキング:
Times Higher Education: <https://www.timeshighereducation.com/world-university-rankings>;
QS World University Rankings: <https://www.topuniversities.com/qs-world-university-rankings>;
Academic Ranking of World Universities (ARWU, also known as Shanghai Ranking):
<http://www.shanghairanking.com/>
- 4) <http://www.nistep.go.jp/>
- 5) 「研究論文に着目した日本の大学のベンチマーキング 2015—大学の個性活かし、国全体としての水準を向上させるために—」 [調査資料—243] : <http://www.nistep.go.jp/archives/25388>;
「研究論文に着目した日本の大学のベンチマーキング 2015—大学の個性活かし、国全体としての水準を向上させるために—」 [調査資料—243] : <http://www.nistep.go.jp/archives/4129>
- 6) 「最新のサイエンスマップ: サイエンスマップ 2014—論文データベース分析 (2009—2014 年) による注目される研究領域の動向調査」—[NISTEP REPORT—169] : <http://hdl.handle.net/11035/3147>
- 7) ビブリオメトリクスとその研究評価への利用の歴史についての総説的なモノグラフ: Gingras, Y. Bibliometrics and Research Evaluation (The MIT Press, 2016)
- 8) <https://clarivate.com/products/web-of-science/>
- 9) <https://www.elsevier.com/solutions/scopus>
- 10) <https://clarivate.com/products/incites/>
- 11) <https://www.elsevier.com/solutions/scival>
- 12) Snowball Metrics Webpage: <https://www.snowballmetrics.com/>
- 13) www.hefce.ac.uk/pubs/rereports/Year/2015/metrictide/
- 14) Hicks, D., Wouters, P., Waltman, L., de Rijcke, S. & Rafols, I. *Nature* **520**, 429-431 (2015).
- 15) Snowball Metrics Recipe Book: <https://www.snowballmetrics.com/wpcontent/uploads/0211-Snowball-Metrics-Recipe-Book-v7-LO-1.pdf>

(せいけ ひろふみ)

| | Research Inputs | Research Processes | Research Outputs and Outcomes |
|--|---|--|---|
| Research | <ul style="list-style-type: none"> ● Applications Volume ● Awards Volume ● Success Rates | <ul style="list-style-type: none"> ● Income Volume ● Market Share | <p><u>Publications & Citations</u></p> <ul style="list-style-type: none"> ● Scholarly Outputs ● Citation Counts ● Citation per Output ● h-index ● Field-Weighted Citation Impact ● Outputs in Top Journal ● Publications in Top Journal Percentiles <p><u>Collaboration</u></p> <ul style="list-style-type: none"> ● Collaboration ● Collaboration Publication Share ● Collaboration Impact ● Collaboration Field-Weighted Citation Impact ● Academic-Corporate Collaboration ● Academic-Corporate Collaboration Impact <p><u>Societal Impact</u></p> <ul style="list-style-type: none"> ● Altmetrics ● Public Engagement ● Academic Recognition |
| Enterprise Activities/ Economic Development | <ul style="list-style-type: none"> ● Academic-Industry Leverage ● Business Consultancy Activities | <ul style="list-style-type: none"> ● Contract Research Volume | <ul style="list-style-type: none"> ● Intellectual Property Volume ● Intellectual Property Income ● Sustainable Spin-Offs ● Spin-Off-Related Finances |
| Post-Graduate Education | <ul style="list-style-type: none"> ● Research Student Funding | <ul style="list-style-type: none"> ● Research Student to Academic Staff Ratio | <ul style="list-style-type: none"> ● Time to Award of Doctoral Degree ● Destination of Research Students Leavers |

(図 1) Snowball Metrics recipes (Snowball Metrics Recipe Book, p.15: Figure 2 を転載)