

[SciREX-WP-2021-#02]

研究成果指標における多様性と標準化の両立

－ 人文・社会科学に焦点をおいて －

**Balancing diversity and standardization in research output indicators**

**- Focusing on the humanities and social sciences -**

2021/06

林 隆之 (政策研究大学院大学 教授)

藤光 智香 (文部科学省 大臣官房人事課 長期在外研究員)

秦 佑輔 (文部科学省 科学技術・学術政策局 企画評価課 総括係長)

中渡瀬 秀一 (国立情報学研究所 特任研究員)

安藤 二香 (政策研究大学院大学 科学技術イノベーション政策研究センター 専門職)

HAYASHI Takayuki, Professor, National Graduate Institute for Policy Studies

FUJIMITSU Chika, Japanese Government Long-term Fellow, Personnel Division Minister's Secretariat, Ministry of Education, Culture, Sports, Science and Technology

HATA Yusuke, Unit Chief, Planning and Evaluation Division, Science and Technology Policy Bureau, Ministry of Education, Culture, Sports, Science and Technology

NAKAWATASE Hidekazu, Project Researcher, National Institute of Informatics

ANDO Nika, Program Specialist, Science for RE-Designing Science, Technology and Innovation Policy Center (SciREX Center) National Graduate Institute for Policy Studies



SciREX Center  
WORKING PAPER

# 研究成果指標における多様性と標準化の両立

－ 人文・社会科学に焦点をおいて －

2021年6月

林隆之、藤光智香、秦佑輔、中渡瀬秀一、安藤二香

「研究力向上に向けた新たな測定指標の開発：各研究文化に適合した

分野別指標と組織・ネットワークの機能指標」プロジェクト

「研究成果指標における多様性と標準化の両立 – 人文・社会科学に焦点を置いて –」政策研究大学院大学科学技術イノベーション政策研究センター (SciREX センター)ワーキングペーパー、2021年6月.

林 隆之 (政策研究大学院大学 教授)  
藤光 智香 (文部科学省 大臣官房人事課 長期在外研究員)  
秦 佑輔 (文部科学省 科学技術・学術政策局 企画評価課 総括係長)  
中渡瀬 秀一 (国立情報学研究所 特任研究員)  
安藤 二香 (政策研究大学院大学 科学技術イノベーション政策研究センター 専門職)

*Balancing diversity and standardization in research output indicators -Focusing on the humanities and social sciences-*, National Graduate Institute for Policy Studies, Science for RE-Designing Science, Technology and Innovation Policy Center (SciREX Center) Working Paper, March 2021.

HAYASHI, Takayuki, Professor, National Graduate Institute for Policy Studies  
FUJIMITSU Chika, Japanese Government Long-term Fellow, Personnel Division Minister's Secretariat, Ministry of Education, Culture, Sports, Science and Technology  
HATA, Yusuke, Unit Chief, Planning and Evaluation Division, Science and Technology Policy Bureau, Ministry of Education, Culture, Sports, Science and Technology  
NAKAWATASE Hidekazu, Project Researcher, National Institute of Informatics  
ANDO, Nika, Program Specialist, Science for RE-Designing Science, Technology and Innovation Policy Center (SciREX Center) National Graduate Institute for Policy Studies

本ワーキングペーパーは、文部科学省 SciREX 事業「研究力向上に向けた新たな測定指標の開発：各研究文化に適合した分野別指標と組織・ネットワークの機能指標」プロジェクト（2019～2020 年度）の成果の一部である。

## 目次

要旨 .....	1
Executive summary .....	2
1. はじめに .....	3
2. 研究成果指標の多様性と標準化の関係 .....	4
2.1 日本における研究成果指標の資金配分への活用状況 .....	4
2.2 多様な指標設計のフレーム .....	5
3. 多様な研究成果の測定：人文・社会科学における検討の国際的状況 .....	7
3.1 発表メディアの質管理や格付け .....	7
3.2 ピアレビューにおける研究成果の多様性の推奨 .....	17
3.3 「責任ある研究評価」概念と指標の関係 .....	22
4. 日本における研究成果測定の試行 .....	26
4.1 分析対象 .....	26
4.2 日英における歴史学・経営学の研究成果の種類 .....	27
4.3 日英における出版メディアの状況 .....	32
4.4 ピアレビューにおける代替指標 .....	42
5. 社会的インパクトの測定における多様性と標準化：論点整理 .....	45
5.1 社会的インパクト測定の論点 .....	46
5.2 人文・社会科学におけるインパクト測定の論点 .....	55
6. おわりに .....	57
参考文献 .....	59

## 要旨

研究成果の測定は、資金配分や組織の戦略策定など様々な目的のもとで行われる。研究成果の測定において留意しなければならない点は、全ての研究分野に適応可能な一律の指標群は存在しないにもかかわらず、限られた数の指標が使われやすく、それによって、組織や研究者の行為に望まれない影響が生じることである。本ワーキングペーパーは、大学等の組織を単位とした研究測定における多様な研究成果指標に関する課題について、主に人文・社会科学に焦点を置きながら検討する。

最近の日本における大学等への「業績に基づく資金配分」の文脈の中では、比較可能な少数の標準的な指標を設定し測定することが求められる。しかし、学術界からは、多様な研究活動を奨励するために、できるだけ多種多様な研究成果を認識する重要さが指摘され、それらは比較可能な形で集計することが難しい。そのために議論のすれ違いが生じやすい。このような「多様性」と比較可能な「標準性」とを両立させることが現実の制度設計において課題となる。この関係に対して、指標中心の評価の仕組みと、ピアレビューの中で指標を活用する評価の仕組みとでは対応の仕方が異なる。海外の状況を分析した結果、少数の指標中心の仕組みとしては、ノルウェーモデルと呼ばれるような、英語ジャーナル論文以外も含めて広く定義した「学術出版物」を計測する方法がとられている。そこでは、国内データベースの整備、学術コミュニティによる「学術出版物」の定義と学術出版チャンネルリストの作成、測定による影響のモニタリングの仕組みが必要となっている。他方、ピアレビュー中心の仕組みでは、研究の定義を広く設定するとともに、分野ごとに多様な成果の例示を評価機関等が作成し、定義や記載内容を共通化していくことが必要となっている。

日本でこのような取組を実施しうる可能性について、歴史学と経営学を対象に、大学評価への提出業績、科学研究費補助事業の成果について分析を行い、英国の REF2014 での成果と比較した。結果、日本では研究成果の多様性が英国より高く、ジャーナルや出版社を「学術出版物」として区分するよりは、幅広いオーディエンスを対象とする成果発表を行っている傾向があり、海外のように定義した「学術成果物」の測定をそのまま用いることは現状では難しいことが示唆された。また、補足的に、国際的にも経験が十分に蓄積されていない、研究成果の学術面を超える社会的インパクト測定においても同様に、その多様性と標準化について、一般的論点と人文・社会科学に特有の論点があることを示した。

少数の指標を中心とした評価とピアレビューを中心とした評価の双方の仕組みにおいて、多様性と標準化を追求するためには複数の留意点があり、今後は、測定のみならず人文・社会科学研究の価値についての根本的議論も含めて、大学やアカデミーなどの関係者が協議することが期待される。さらに、今後、社会変革を促進するための人文・社会科学を含めた「総合知」が求められるなかで、その評価のあり方については、こうした論点も踏まえた検討が必要となる。

Research outputs are measured for a variety of purposes, including funding decision and organizational strategy formation. One thing to keep in mind when measuring research outputs is that although there is no one-fit-all set of indicators for all research fields, a limited number of indicators are likely to be used. It causes undesired effects on a researcher's behavior. This working paper examines issues related to various research output indicators, focusing mainly on the humanities and social sciences (SSH).

In the context of "performance-based funding allocation" to universities in Japan, it is required to set a small number of comparable standard indicators. However, the academic community has pointed out the importance of recognizing as many different types of research results as possible in order to encourage diverse research activities, although it is difficult to tabulate them in a comparable manner. Therefore, arguments don't overlap each other. Balancing "diversity" and comparable "standardization" is an issue in actual institutional design. This balancing is dealt with differently between a metrics-centered evaluation and a peer-review-centered evaluation. As a result of analyzing the situation overseas, as a metrics centered evaluation, a method called the Norwegian model is adopted, which measures "academic publications" that are defined including those other than English journal articles. For this model, there is a need for a domestic database, a scholarly community to define "academic publications" and academic publishing channel lists, and a mechanism for monitoring the impact of measurements. On the other hand, in a peer-review-centered evaluation, it is necessary to broadly set the definition of research, and to create examples of various research outputs in each field by evaluation institutions, etc., and to standardize the definition and description for them.

Regarding the possibility of implementing such efforts in Japan, we analyzed the submitted research outputs for university evaluations and the reports of the Grant-in-Aid for Scientific Research focusing on history and management research from SSH and compared them with the achievements at REF2014 in the UK. As a result, the variety of research outputs in Japan is higher than in the UK, and there is a tendency to publish outputs for a wider audience, not classifying journals and publishers strictly as "academic publications". Thus, It was suggested that it is currently difficult to use the defined "academic publication" measurement as in overseas. In the measurement of social impact of research which has not been sufficiently established internationally, the general issues regarding the diversity and standardization are also discussed.

There are several points to keep in mind in balancing diversity and standardization in both the evaluation system. It is expected that stakeholders such as universities and academies will discuss not only measurement but also fundamental discussions on the value of research in SSH. Furthermore, as "converging knowledge" including SSH is required to promote social transformation in the future, it is necessary to consider how to evaluate it based on these issues.

## 1. はじめに

研究成果の測定は、行政機関においては大学等の研究実施組織に配分した公的研究費による成果を確認し、その後の配分の継続に関する意思決定等に活用するとともに、納税者への説明責任を果たすために行われる。研究実施組織内部においても同様に、自組織の活動の成果を確認し、それを踏まえた戦略の策定や組織内部での資源配分、活動実績の説明のために行われる。

このような研究成果の測定において留意しなければならない点は、全ての研究分野に適用可能な一律の指標群は存在しないにもかかわらず、データの入手しやすさ等から、資金配分等の重要な意思決定において、限られた数の指標が使われやすいことである。さらに、少数の指標に基づいて決定が行われることで、組織や研究者の行為に、望まれない影響が生じることである。その典型的な例は、英語で書かれたジャーナル論文を中心に収録したデータベースであるクラリベイト・アナリティクス社の Web of Science (WoS) やエルゼビア社の論文データベース Scopus 等を用いた指標を全分野に一律に適用することにより、本来はこれらのデータベースに含まれない研究成果が多い分野の研究者の研究成果発表や研究活動に影響が生じることである。

このような課題は国際的には認識され、後述するように、DORA やライデン声明をはじめとする複数の提言が公表されており、また、欧州諸国を中心に研究成果の多様性を踏まえた具体的な方策の検討がなされてきた (Bornmann, 2011; Glaser, 2007; Rijcke et al., 2016; Wilsdon et al., 2015)。日本においても一定の問題認識が持たれてきたが、具体的な方策の検討まではほとんど展開していないのが現状である。

本ワーキングペーパーは、大学等の組織を単位とした研究成果測定における多様な研究成果指標に関する課題について、主に人文・社会科学に焦点をおきながら、以下の点を検討する。まず、現時点の日本における大学への資金配分や大学評価の文脈の中で、標準的な測定を求めることと、多様な研究成果指標を設定することの関係をどう整理できるか検討する (2章)。次に、学術的成果の指標について、海外諸国において人文・社会科学に対してどのような取組が行われているのかを示す (3章)。それらと同様の取組を、日本において行うことは可能であるのかについて、日英のデータ比較から検討する (4章)。また、研究成果の学術面を超える社会的インパクト測定の論点について補足的に検討する (5章)。最後に要点をまとめる (6章)。なお、本ワーキングペーパーは研究成果測定の方法に焦点をおき、研究の多様な「質」の評価基準については、直接は扱わない。

## 2. 研究成果指標の多様性と標準化の関係

### 2.1 日本における研究成果指標の資金配分への活用状況

日本では近年、大学や部局といった組織を単位として研究成果を量的に測定し、その結果を、資金配分に直接的に用いることがみられるようになってきている。例えば、2013年の文部科学省「研究大学強化促進事業」では「ヒアリング対象機関選定のための指標」として Top10%論文数の割合など複数の指標を用い<sup>1</sup>、2014年の「スーパーグローバル大学創成支援事業」の「タイプA（トップ型）」の審査基準では、論文の被引用状況などの指標を示した<sup>2</sup>。

一方、国立大学法人評価においては、学部・研究科等を単位とした研究評価（現況分析）が行われ、そのうちの「研究業績水準判定」では、定性的なピアレビューが行われてきた。ここでは、大学が学部・研究科等を代表する研究業績の説明書を提出し、それらを各分野の研究者で構成された委員会が段階判定を行う。しかし、そのように手間をかけたピアレビューによる判定結果は、大学への資金配分には強く影響しない。研究業績水準判定の結果は集計され、学部・研究科等を単位とする研究評価（現況分析）の一情報として使われる。さらに学部・研究科等の研究評価が、国立大学法人を単位とする中期目標・計画の達成度評価の一情報として使われる。この達成度評価の結果が、国立大学の運営費交付金総額のうちの年間30億円分（0.3%程度）の配分に反映される。つまり、ピアレビューによる研究評価結果は、間接的な形で少額に反映されるのみである。

しかし、2019年度配分から国立大学への運営費交付金配分の方式が変更された。「成果を中心とする実績状況に基づく配分」が導入され<sup>3</sup>、5つの共通指標の値によって700億円分（運営費交付金全体の7%程度）が競争的に再配分されることとなった。そのうちの一つの指標が「運営費交付金等コスト当たり Top10%論文数」であった。これはエルゼビア社の論文データベースを用いた測定であり、前述のように、英語ジャーナル論文以外の研究成果が多い分野にとっては不適切な測定である。

そのため、国立大学協会では委員会を設置して適切な指標についての検討を始め、教育・研究実績を11学系（分野）ごとに測定するような指標群の設定を求める提言を行った（国立大学協会 2019）。その中で研究成果に関しては、国立大学法人評価におけるピアレビューが次回は2020年に行われる予定であり、まだ直近の評価結果が存在しないことから、その前提のもとでは暫定的に、「教員あたり研究業績数」等の定量指標のみで構成する提言となった。ただし、これは商用の英語論文データベースにて計測されるものを使うのではなく、学系（分野）別に測定するべき適切な研究業績の種類を検討するように求めた。提言を受けて文部科学省が検討を行った結果、上記の Top10%論文数の指標を残しつつも、教員あたり研究業績数の指標が新たに設定されることとなった（国立大学協会 2019、林 2019）。

しかし、これまでも分野ごとの研究活動や成果の多様性を尊重し、測定や評価でもその多様性を踏まえるべきという指摘は度々なされてきた（日本学術会議 2008、2012、文部科学省 2017）ものの、具体的に分野別にどのような種類の研究成果を測定すべきかについては、これまで学术界を含め、十分な議論の蓄積がないのが実情である。また、何らかの形で資金の配分を行うことが目的である場合には、比較が可能なように、分野内のみならず分野間でもある程度の標準化が必要になるが、この点についての議論の蓄積もなされていない。

<sup>1</sup> [https://www.mext.go.jp/component/a\\_menu/science/detail/\\_icsFiles/afiedfile/2013/06/03/1333816\\_03\\_2\\_2.pdf](https://www.mext.go.jp/component/a_menu/science/detail/_icsFiles/afiedfile/2013/06/03/1333816_03_2_2.pdf)

<sup>2</sup> [https://www.jsps.go.jp/j-sgu/data/download/03\\_sgu\\_shinsakijun.pdf](https://www.jsps.go.jp/j-sgu/data/download/03_sgu_shinsakijun.pdf)

<sup>3</sup> [https://www.mext.go.jp/a\\_menu/koutou/houjin/1417427.htm](https://www.mext.go.jp/a_menu/koutou/houjin/1417427.htm)



実際に、国立大学協会の委員会においても、人文・社会科学においては書籍が主要な成果発表の形態であることは認識されていたが、具体的にどのような書籍を「研究業績」として考え、何を単位として測定すべきであるか、それ以外の成果として、たとえば「紀要」をどう扱うか、芸術学やスポーツ学における公演やパフォーマンスを含めたその他の業績を集計可能な形でいかに測定できるかといった問題に具体的に直面した。

---

## 2.2 多様な指標設計のフレーム

上記の検討を経てわかることは、活用目的を踏まえた上で、指標の多様性と標準化のバランスをとった仕組みをいかに構築できるかの検討が必要ということである。

学術界からは、分野ごとに異なる種類の研究成果を尊重するように、多様性や包摂性の重要性が指摘され、それが研究内容や研究方法の多様性も育てて学問の発展を促進することが主張される。しかし、資金配分へ定量的な指標を利用する場面では、指標が多数あり多様でありすぎれば、それらを比較して活用することが難しくなる。そのため、多様な種類の指標から代表的な指標を選別し、比較可能なように定義等の標準性を有する指標群の設定が求められる。

すなわち、指標の検討においては、多様性と標準化のバランスの問題が生じる。このバランスをいかにとるべきかという問題は、測定や評価の結果をどのような目的で用いるのかを明確にしない限り解決しない。その前提が無い場合に、多様性を求める学術界と、説明責任を伴う意思決定の中で比較可能な指標を求める行政機関を始めとする評価主体との間ですれ違った議論が展開されることになる。

本ワーキングペーパーでは、成果指標の使用目的として組織への経常的資金などの資金配分への活用を主に想定する。しかし、その中でも最適なバランスは付随的な目的や、どの程度の規模の対象を測定するのか、どの程度のコストをかけるかによって変わる。

表1には大きく二種類の利用目的における指標設定の考え方を整理した。

一つの種類は、大学間の比較などを定量指標により行い、分野の特徴は踏まえつつも限られた数の種類の指標を用いる方法である。目的としては、運営費交付金配分や組織単位の資金配分などを、透明かつ客観的で、低コストに行うために、測定を行うものである。ここでは、個人の主観的判断は含まずに、指標のみの設計が重視される。指標については、分野ごとの特徴を考慮しながらも、定義や測定の方法が定まった少数のものとされる。

ただし、ここで論点となるのは、量を測定するだけでなく、ある程度、質の側面を考慮する必要があることである。このような方式の具体的な事例は、学術面では北欧諸国における経常的資金の配分のための測定（ノルウェーモデル等の学術出版物の総合的把握）があげられ、研究による社会インパクトの測定では、オーストラリアの大学評価である Excellence in Research for Australia (ERA)における、研究の応用の指標群や、英国の Knowledge Exchange Framework (KEF: 知識交換フレームワーク)の指標群などがあげられる。

もう一つの種類は、レビューアーによる評価という枠組みの中で、多様な成果指標を用いる方式である。この場合も経常的資金の配分を目的とすることができるが、研究内容の質的側面についての判断を行い、場合によっては被評価者にフィードバックをして改善を促進することも目的として設定される。

レビューアーによる評価では、用いる情報には自己評価文書や事例説明文書などの定性的な記述が含まれ、その中で定量指標についても記述される。それらを踏まえてレビューアーが総合的に判断を

行う。そのため、指標の多様性に制限をつける必要は無いが、一方で、レビューアーが指標を解釈するためには、ある程度共通した指標の定義や測定方法の標準化が求められる。

このような方法は、日本では大学改革支援・学位授与機構が行っている研究業績水準判定において、大学が研究の卓越性を定性的に説明する中で多様な根拠・指標を記載する場合に見られる。また、英国の Research Excellence Framework (REF)、並びに、オーストラリアの ERA やイタリアの大学評価機関 ANVUR による VQR (Research Quality Assessment) における人文・社会科学の評価では多様な研究成果の提出は推奨されている。研究の社会的インパクトの測定についても、英国の REF、オーストラリアの Engagement and Impact Assessment (EI) はケーススタディ手法をとっており、ガイドラインとして示す記載が期待される多様な指標について例示を行っている。

このように、指標の多様性と標準性のトレードオフは大きく 2 つの枠組みによって対応が大きく異なる。次章では、それぞれの方法について各国での取組を確認する。

表 1 活用目的による指標の検討方向の違い

評価の種類	評価の目的	指標の検討方向	評価・指標利用のコスト	指標検討の海外動向	事例
指標に基づく 大学等の比較	経常的予算配分や 組織単位の資金配 分などのために、 透明かつ客観的 で、低コストに研 究成果を測定す る。	指標の数は少数。 機械的に判断が 可能な状態。	指標収集のため の DB 構築を行 う場合はコスト。 指標値による評 価判断のコスト は低い。	質がある程度 は担保された 上で量を測定 する指標の検 討。	学術面：経常的経費配分におけ る研究成果指標（例：ノルウェ ーモデル） 社会インパクト面：産学連携研 究費や特許数等の少数指標 （例：オーストラリア ERA、英 国 KEF）
レビューアー による評価(ケ ーススタディ やナラティブ などの定性的 記述の中で指 標を活用)	経常的予算配分や 組織の自己・外部 評価のために、研 究内容の質的側面 を含めて評価を行 う。	多種類の指標の 例示。一方、その 中でもある程度 の比較可能性を 担保するための 標準化をガイド ライン等で実施。	指標収集を自己 評価の記述で行 う場合にはコス トは低いが、評価 者の評価作業に 多大なコスト。	例示する主要 な指標の検討。 多様な指標に ついて記述形 式の統一定義 などの整理。	学術面：研究成果そのものの評 価（例：英国 REF、オーストラ リア ERA、イタリアの人文社 会）。多様なエビデンス提示に 基づく質的判断（例：日本の現 況分析）。 社会インパクト面：インパクト のケーススタディの評価（例： 英国 REF、オーストラリア EI）

### 3. 多様な研究成果の測定：人文・社会科学における検討の国際的状況

研究成果の多様性を踏まえた指標を検討する上で、中心的な対象として人文・社会科学を取り上げる。上述のように、自然科学分野における研究評価では、WoS や Scopus のような英語ジャーナル論文を主に収録する商業的な論文データベースが用いられることが多い（ただし、自然科学内の分野によって課題があることは 4.2.2 節を参照）。一方、多くの人文・社会科学においては、研究成果は上記のデータベースに収録されるようなジャーナル論文だけではない。人文・社会科学分野の主要な研究成果として考慮すべきものには、国際的ジャーナル、書籍、国内のジャーナル、学術界向けではない一般雑誌などがあり（Hicks 2004）、加えて、分野によっては、歴史資料や政策提言、芸術の実演や作品なども研究成果として扱われる。

このような点から、人文・社会科学における研究成果の測定について、国際的には様々な検討がなされてきた(Reale et al., 2018)。欧州では、最近では EvalHum という人文・社会科学における研究評価、イノベーション、インパクト評価に関するイニシアティブ（フランスにて協会として登録された組織でもある）<sup>4</sup>のもと、EU が資金提供するプロジェクト European Cooperation on Science and Technology (COST) Action の活動が 2016 年から 2020 年まで行われ、40 カ国 150 人の研究者や政策形成者が人文・社会科学の評価について議論を行うネットワーク（European Network for Research Evaluation in the Social Sciences and the Humanities: ENRESSH）が形成されてきた(Jong et al., 2020)。ENRESSH は、その目的を、人文・社会科学における研究の可視性を高め、また、それらの研究が社会的課題に関連する問題に対処する潜在的能力を有することの可視性を高めること、ならびに、人文・社会科学の研究者のコミュニケーションの方法に適合する総合的な評価方法を開発することとしている。

以下では前章の整理を踏まえ、まずは資金配分に直結する定量的測定における人文・社会科学の研究成果の指標について ENRESSH の内容も踏まえて整理する。次に、ピアレビューを中心とする評価の中で、多様な研究成果を認識するための各国の取組を整理する。最後に、近年、提示されている「責任ある研究評価」「責任ある研究測定」の考え方がこれらとどのように関連するかを検討する。

---

#### 3.1 発表メディアの質管理や格付け

##### 3.1.1 全般的状況

研究成果の測定を、大学等の経常的資金の配分に結びつけている国の中で、ピアレビューのようなコストがかかる方法を採用しにくい国では、複数の研究成果の種類を対象に測定を行い、その中で質に関しても代替的な測定を行っている。その典型的な例は、英語ジャーナル論文だけでなく、自国語で書かれた国内ジャーナル論文や書籍等についても測定を行い、それらの発表メディアの質を何らかの方法で設定し、それによる重み付け集計を行う方法である。これはノルウェーが最初に本格的に開始したために、ノルウェーモデルとしばしば称されるが(Aagaard et al., 2015; Sivertsen, 2018, 2016)、デンマーク(Aagaard, 2018)、フィンランド(Pölönen, 2018)などの北欧諸国や、ベルギー(Engels and Guns, 2018)、ポーランド(Kulczycki and Korytkowski, 2018)においても導入されている。

これらの具体的な方法は国により異なるが、このような仕組みが実現されるためには、いくつかの共通した条件がある(Pölönen et al., 2021)。

---

<sup>4</sup> <http://www.evalhum.eu>

## 国内データベースの整備

第一に、国内で英語ジャーナル論文以外の研究成果を入力することが可能な独自データベースが整備されていることである。欧州における国内データベース（国内の全大学の研究成果を対象としたデータベース）を2016年に調査した結果によれば(Sile et al., 2018)、調査対象39ヶ国のうち、21ヶ国に国内データベースが存在し、そのうちの18のデータベースは全分野を対象としている。大学への経常的資金配分に研究成果数を用いる国では、国内データベースは不可欠であり、ノルウェーのCRISTIN<sup>5</sup>（その中で資金配分に用いられるのはNVIという部分である）、デンマークのBFI<sup>6</sup>、フィンランドのVIRTA Publication Information Services<sup>7</sup>、スウェーデンのSwePub<sup>8</sup>、ベルギーのフランデース地方のVABB-SHW<sup>9</sup>などが形成されている。

## 「学術出版物」の定義と学術出版チャネルリストの作成

研究成果データベースがあるだけでは、資金配分のための測定には不十分である。どのような成果が資金配分のための集計の対象になるのかを定義することが必要である。それがなければ、雑多な対象について集計を行うことになり、比較をするのは難しい。そのため研究活動を学術面から評価する際には、「学術的な」成果物を定義することが必要となる（研究成果による社会的インパクトはこの議論の対象となっていない）。発表・講演や実演など、中身が多様で集計を行いにくい成果でなく、出版物に制限されていることが各国の実態である。

多くの国ではピアレビューなどの質的判断を経た学術出版チャネル（ジャーナルや出版社）によって発表された出版物を「学術出版物」と定義している。それらの国では「学術出版物」とみなしうるジャーナルや出版社のリスト（登録簿）を作成し公表している。表2に例示するように、ジャーナルのリストについては、欧州全体の取組、北欧諸国のノルウェー、デンマーク、フィンランド、南欧諸国のスペイン、イタリア等の事例がある。また、フランスやオーストラリアは以前に実施していたが、学術界からの反対があり、現在はフランスでは一部の分野のみで作成され、オーストラリアでは提出対象になりうるジャーナルのリストとして利用されているが最終的には分野単位（学科などに相当する単位）で評価者が指標などを確認して総合的に評点付けを行う。この他にも、ブラジル等の南米諸国、台湾、オランダでも人文・社会科学のジャーナルのリストを作成している(Ferrara and Bonaccorsi, 2016)。なお、フィンランド等では国内データベースには、ピアレビューを経ない一般書等の多様な成果も入力できるが、資金配分のための指標に用いる学術出版物はこれらのリストに含まれるものだけとしている。

さらに、多くの国では、リスト化されたジャーナルや出版社の中でも、学術的質の点から特に優れたものを識別し、集計の際に重みをつけることを行っている。これは、過去にオーストラリアで研究成果数を資金配分のための指標としたことで、平均引用数が低いジャーナルへの論文投稿が増えたため(Butler, 2003)、単なる量だけでなく、質の側面を加味したものとするという意図である(Schneider et al., 2016)。

---

<sup>5</sup> <https://www.cristin.no/english/>

<sup>6</sup> <https://bfi.fi.dk/>

<sup>7</sup> <https://wiki.eduuni.fi/display/cscvirtajtp/VIRTA-julkaisutietopalvelu>

<sup>8</sup> <http://swepub.kb.se/>

<sup>9</sup> <https://www.ecoom.be/en/data-collections/vabb-shw>

表 2 国・地域単位のジャーナルリスト（登録簿）の例

	ジャーナルリスト
欧州	European Reference Index for the Humanities and Social Sciences (ERIH PLUS) <sup>10</sup>
ノルウェー	Norwegian Register for Scientific Journals, Series and Publishers <sup>11</sup>
デンマーク	BFI lists <sup>12</sup>
フィンランド	Finland's Julkaisufoorumi (JUFO) <sup>13</sup>
ベルギーフランダース地方	VABB-SHW Database: Lists of journals and book series、および Publisher lists <sup>14</sup>
スペイン	RESH <sup>15</sup>
イタリア	Italian ANVUR ranking
オーストラリア	ERA Journal List <sup>16</sup>
フランス	HCERES ジャーナルリスト（経済学・経営学および言語学のみ） <sup>17</sup>

ジャーナルだけでなく、「学術出版社」として認められる出版社のリストもいくつかの国で作成されている。出版社の登録簿に関して欧州 19 ヶ国を対象とした調査によれば(Giménez-Toledo et al., 2019)、ノルウェー、フィンランド、デンマーク、ベルギーのフランダース地方、ポーランド、スロバキア、スロベニア、スペインが作成している。また、出版社だけでなく、個々の書籍やシリーズについても質のラベルをつけているのは、フィンランド、ベルギーのフランダース地方、スペインである。他方、チェコ共和国、リトアニア、ポーランドは書籍についてもポイント制を用いているが、これらの国では書籍のポイントはジャーナル論文よりも低く設定されており、出版社のリストは存在しない(Williams et al., 2018)。

#### 学術コミュニティによる選定体制および選定基準の明確化

「学術出版物」と見なされる出版チャネルのリストが学術コミュニティから受け入れられるためには、それを作成する主体は、各分野の研究者から構成された委員会などの組織であることが必要になる。たとえば、後述するようにノルウェーでは学長会議の下に置かれた委員会がおこなっており、各学会が選定した研究者から構成されている。ただし、分野ごとに選定基準が異なる形にはしておらず、共通的・標準的 (universalistic) な基準とすることで、分野を超えた集計に意味があるようにしている(Sivertsen, 2018)。

#### 3.1.2 学術出版チャネルリストの各国事例

以下に、いくつかの学術出版チャネルリストを形成している事例を紹介する。

<sup>10</sup> <https://dbh.nsd.uib.no/publiseringskanaler/erihplus/>

<sup>11</sup> <https://dbh.nsd.uib.no/publiseringskanaler/Forside>

<sup>12</sup> <https://ufm.dk/en/research-and-innovation/statistics-and-analyses/bibliometric-research-indicator/bfi-lists>

<sup>13</sup> <https://julkaisufoorumi.fi/en>

<sup>14</sup> <https://www.ecoom.be/en/data-collections/vabb-shw>

<sup>15</sup> <http://epuc.cchs.csic.es/resh>

<sup>16</sup> <https://www.arc.gov.au/excellence-research-australia/era-2018-journal-list>

<sup>17</sup> <https://www.hceres.fr/fr/guides-des-produits-de-la-recherche-et-activites-de-recherche-0>

## (1) 欧州 European Reference Index for the Humanities and Social Sciences (ERIH PLUS)

欧州全体における、人文・社会科学における質が確認されたジャーナルのリストとしては、2001年に欧州科学財団 European Science Foundation (ESF) が開始したイニシアティブである European Reference Index for Humanities (ERIH) がある<sup>18</sup>。ERIH は 2008 年に最初のリストを公表し(Pontille and Tornø, 2010)、2011-2012 年にその改訂版を公表した。その後、2014 年にノルウェー教育研究省が所管するノルウェー研究データセンター (Norwegian Centre for Research Data: NSD。2016 年までの旧称は Norwegian Social Science Data Services) に移管され、European Reference Index for the Humanities and Social Sciences (ERIH PLUS) に改称した。当初は ESF の人文学常任委員会 (Standing Committee for the Humanities: SCH) によって開発されたために、人文学のジャーナルのみを対象としていたが、移管後は社会科学も含むように拡大されている。ERIH PLUS の目的は、人文・社会科学の可視性、検索可能性、利用可能性を向上させることにある。

ジャーナルの登録には、まず、大学や研究機関の研究者らが申請を行う (商業出版社からの申請は認められない)。登録の要件としては、過去 2 年間の出版歴を対象に、ジャーナルのウェブサイト等において以下の事項が満たされていることが確認できることである。

1. 外部の独立したレビューの手順が明確にされている。
2. 学者によって構成される編集委員会が設置されている。
3. 有効な ISSN コードがある。
4. 公開されたすべての論文に英語、またはその分野の国際言語での要約がついている。
5. 全ての論文の著者の所属機関名が記載されている。
6. 同じ機関の著者が 3 分の 2 以下である。

ERIH PLUS には 2021 年 3 月現在で 9,348 誌が登録されている (書籍や会議録は含まれていない)。日本語によるジャーナルの登録はないが、出版国が日本となっている雑誌は 10 誌登録されている<sup>19</sup>。

## (2) ノルウェー Norwegian Register for Scientific Journals, Series and Publishers

ノルウェーでは、教育研究省から高等教育機関へ配分される経常的資金は、固定要素 (平均して約 70%) と成果に基づく要素 (約 30%) で配分される。成果に基づく要素は①学生の修得単位数、②卒業生数、③海外学生交流数、④博士修了者数、⑤研究成果物数、⑥EU からの資金、⑦ノルウェーリサーチカウンシルからの資金、⑧市や企業などからの公的・私的収入、の 8 つの定量指標によって配分される。この中の⑤研究成果指標による資金配分の方式は、「ノルウェーモデル」と呼ばれており、ノルウェー高等教育機関協会 (Norwegian Association of Higher Education Institutions (UHR) の提言に基づいて 2005 年より実施されている (Aagaard et al., 2015; Sivertsen, 2018, 2016; UHR, 2004)。

ノルウェーでは研究成果を入力する全国共通システムとして CRISIN (Current Research Information System in Norway) があり、WoS や Scopus 等の論文データベースに含まれない研究成果も大学自身により入力されている。ただし、CRISTIN の中で資金配分のための計測対象となる部分は、Norwegian Scientific Index (Norsk vitenskapsindeks : NVI) と呼ばれる学術的出版物の部分のみであ

---

<sup>18</sup> <http://archives.esf.org/hosting-experts/scientific-review-groups/humanities-hum/erih-european-reference-index-for-the-humanities.html>

<sup>19</sup> Cuadernos CANELA、Inter Faculty - Journal of Interdisciplinary Research in Human and Social Sciences、Brain & development (Tokyo, 1979)、Japanese Religions、Japanese journal of religious studies、Monumenta Nipponica、Historia Scientiarum、Asian Ethnology、JALT Journal、The Language Teacher (TLT) の 10 誌である。

る<sup>20</sup>。NVIにおける学術的出版物は、以下の要件を満たすものと定義されている。

1. 新たな洞察を提供している。
2. 成果が検証可能な、あるいは新たな研究へ活用可能な様式で公表されている。
3. 関心を有する研究者の多くがアクセス可能な言語や普及形態になっている。
4. ピアレビューが定期的実施される出版チャンネル（ジャーナル、叢書、出版社、ウェブサイト）に掲載されている。

上記の3と4については具体的には、世界の学術出版チャンネル（ジャーナルや叢書、および出版社）の登録簿である Norwegian Register for Scientific Journals, Series and Publishers に登録されていることが要件となっている。登録は、大学や研究機関の研究者や図書館員が申請を行い、それをもとに、UHR に常設された、各分野代表者で構成される全国学術出版物委員会（National Board of Scholarly Publishing: NPU）<sup>21</sup>が審査をして承認を行う。

ジャーナルや叢書の登録要件は、ISSN を有し、学術的な編集委員会を有し、外部ピアレビューを行い、単一機関の著者が3分の2以下であることである。また、出版社の登録要件は、学術的出版物の定義に沿った編集を行う体制があり、外部アドバイザーを有する学術出版事業を有し、単一機関の著者が3分の2以下であることである。

さらに、ノルウェーでは上記を満たす出版チャンネルの質をレベル1と2にわけ、集計のポイントに差をつけている（上記を満たさない出版チャンネルはレベル0である）。レベル2は国際的なジャーナルやシリーズや出版社など、先導的な出版チャンネルであり、各分野において出版物の20%までとなっている。レベル2のリストは、毎年、全国学術出版物委員会の下の分野別パネル等が推薦を行い、全国学術出版物委員会が承認して公表する。

2021年3月末現在、ジャーナルおよび叢書36,201誌、出版社3,257社が登録されている。ジャーナルおよび叢書については、言語表記が含まれているものが20,044誌あり、うち、英語が15,428誌（77%）、英語以外が4,616誌（23%）である（表3）。分野別では、英語以外は人文学が44%、社会科学が21%であり、社会科学は、自然科学を含む全体の値よりも低い。ノルウェー語は195誌（1%）と多くない。なお、日本語は9誌（うちレベル1以上に分類されたものは5誌）が登録されているのみであり<sup>22</sup>、言語を問わず、日本の出版社から出版されているジャーナルは291誌（うちレベル1以上は261誌）である。

---

<sup>20</sup> <https://www.cristin.no/english/resources/reporting-instructions/>

<sup>21</sup> <https://npi.nsd.no/organisering> <https://npi.nsd.no/organisering/npu?id=1109>

<sup>22</sup> 日本語でLevel1以上となっているジャーナルは、人文学では Japanese Journal of Northern European Studies、Poetica (Tokyo)、社会科学では IAFOR Journal of the Social Sciences、Tokushu kyōikugaku kenkyū、Bulletin of the Hokkaido Museum of Northern People の計5誌である。

表 3 Norwegian Register for Scientific Journals, Series and Publishers のジャーナル・叢書登録数の言語別構成

言語	全分野				人文学				社会科学			
	全登録数 (レベル0 を含む)	レベル 1	レベル 2	レベル2 の割合	全登録数 (レベル0 を含む)	レベル 1	レベル 2	レベル2 の割合	全登録数 (レベル0 を含む)	レベル 1	レベル 2	レベル2 の割合
English	15,428	11,335	1,242	10%	2,964	2,077	488	19%	4,245	3,133	352	10%
Multiple languages	2,052	1,570	92	6%	1,071	858	78	8%	442	361	8	2%
German	493	342	36	10%	272	194	34	15%	109	73	2	3%
French	362	279	17	6%	213	160	15	9%	91	76	0	0%
Spanish	354	272	3	1%	194	157	3	2%	85	77	0	0%
Norwegian bokmål	195	83	0	0%	65	32	0	0%	73	36	0	0%
(略)												
Japanese	9	5	0	0%	3	2	0	0%	5	3	0	0%
(略)												
全合計 (言語表記が ないものを含む)	36,201	25,752	2,192	8%	7,030	5,232	658	11%	8,013	6,204	462	7%

(出典：Norwegian Register for Scientific Journals, Series and Publishers より筆者作成)

出版社については、米国の出版社が 575 社 (19%) と多く、ノルウェーの出版社は 190 社 (6%) である。日本は 36 社 (1%) であり、うちレベル 1 以上となっているものは 17 社である<sup>23</sup> (表 4)。

表 4 Norwegian Register for Scientific Journals, Series and Publishers の出版社登録数の国別構成

出版社の国	全登録数 (レベル0を 含む)	レベル 1	レベル 2	レベル2の割合 (1&2のうち)
United States	575	269	27	9%
United Kingdom	373	183	23	11%
Germany	322	172	19	10%
Norway	190	59	0	0%
France	142	83	2	2%
(略)				
Japan	36	17	0	0%
(略)				
全合計 (言語表記がない ものを含む)	3,257	1,637	85	5%

(出典：Norwegian Register for Scientific Journals, Series and Publishers より筆者作成)

登録されたジャーナル、書籍、書籍の中の章の3種について、資金配分における集計の重みは表 5 のように設定されている。重み付けされた出版物数が機関レベルで集計される。なお、共著については、著者数に応じて分数で計測されていたが、人文・社会科学のほうが自然科学よりも生産性が高いという結果になったため、2015年からは分数のルートで集計する方法がとられている(Sivertsen,

<sup>23</sup> 日本の出版社で Level1 以上となっているものは、Akashi Shoten Co. Ltd.、Aratake Shuppan、Bensei Publishing、Hituzi Syobo Publishing、Hokkaido University Press、Institution of Electrical Engineers of Japan (IEEJ)、International Association for Universal Design、International Power Electronics Conference - ECCE Asia、Iwanami Shoten、Japan Institute of Navigation、Kurosio Publishers、Kyoto University Press、Minerva Shobou、Pacific Asia Conference on Language, Information and Computation、United Nations University Press、東京大学出版会(Tokyo Daigaku Shuppankai)、Yokohama Publishers の計 17 誌である。



2018)。

表 5 ノルウェーの資金配分における出版物の種類とレベルごとの重み

出版物の種類	レベル 1	レベル 2
ジャーナル論文	1	3
書籍	5	8
書籍の中の論文	0.7	1

(3)フィンランド Publication Forum によるリスト<sup>24</sup>

フィンランドでは、1990年代より大学への経常的資金（コアファンディング）の一部を実績に基づき配分している。さらに、2013年よりコアファンディングの13%（2億ユーロ）を出版物指標に基づき配分している。

出版物データはVIRTA publication information serviceと呼ばれるシステムで収集している。ノルウェーでは国レベルの新たなシステムとしてCRISTINを構築したが、フィンランドでは既に各大学が独自の情報システムを有しており、多様なローカル情報を統合するシステムとしてVIRTAを開始した。

VITRAでは出版物を表 6の区分のもとで収集している。大別すれば、①学術出版物、②専門職関連の出版物、③一般向けの出版物である。出版物のほかに、芸術作品、特許、学位論文、公表記事なども入力することはできる。

表 6 フィンランド VIRTA publication information service における研究成果の種類

学術出版物	専門誌
A1 ピアレビューされたジャーナル論文：オリジナル研究	D1 業界誌における論文
A2 ピアレビューされたジャーナル論文：レビュー	D2 専門的な書籍の論文
A3 ピアレビューされた論文や書籍の章	D3 専門的会議の予稿集の論文
A4 会議予稿集のピアレビューされた論文	D4 研究開発レポート
B1 ピアレビューされていないジャーナル論文	D5 教科書、専門マニュアルやガイド
B2 ピアレビューされていない論文や書籍の章	D6 専門書の編集
B3 会議予稿集のピアレビューされていない論文	一般向けの出版物
C1 書籍（ピアレビューされた）	E1 一般向けの論文や新聞記事
C2 書籍や特集号の編集（ピアレビューされた）	E2 一般向けの書籍
	E3 一般向け書籍の編集

(出典：Pölonen, 2018より和訳)

資金配分に用いる出版物指標については、2005年から検討を開始し、2009年にフィンランド学長会議（Finish Council of University Rectors。現在のUniversities Finland: UNIFI）が出版物の質の指標開発のWGを設置し、2010年より250学会の代表者で構成されるフィンランド学会連合（Federation of Finnish Learned Societies: TSV）が23の専門家パネルを設置し、ジャーナルと叢書、および出版社の最初のリストを2012年に公表した。その後、TSVの中にPublication Forum(JUFO)を設置し、学会代表者から構成される体制でリストを継続的に更新している。

<sup>24</sup> <https://julkaisuforum.fi/en/publication-forum>

Publication Forumのリストは、VIRTAのデータのうちA1～A4（ピアレビューされた学術出版物）およびC1～2（ピアレビューされた書籍）を分類するためのものであり、全分野の国際・国内の出版チャンネルについて、その質をレベル0～3で分類してリストしている。

レベル0はレベル1の基準を満たしていないチャンネルである。

レベル1（基礎レベル）は、学術的研究成果を公表することに特化し、その分野の専門家から構成される編集委員会を持ち、ピアレビューが行われ、ISSNあるいはISBNが付いている出版チャンネルである。ただし、著者と査読者の半数以上が単一の組織に属している場合は認められない。大部分の出版チャンネルはレベル1に分類される。

レベル2（先導的レベル）とレベル3（最高レベル）は、Publication Forumの専門家パネルによって高いレベルにあると認定された少数の出版チャンネルとされる。レベル2は厳しいピアレビューと競争の結果として、高いレベルの成果が出版されるチャンネルと定義される。主には国際的な出版チャンネルであるが、人文学と社会科学ではフィンランド語とスウェーデン語のチャンネルも含むことが可能である。レベル3は、各分野での最高レベルの研究を発表し、極めて高い影響を及ぼすものであり、分野を総合的にカバーし、著者・読者は国際的であり、編集委員会も各分野の先導的な研究者によって構成され、国際的な研究コミュニティの間で高く評価されているチャンネルである。レベル2および3の選定の上限については、ジャーナルやシリーズでは、全論文数（ジャーナルやシリーズの数でなく）のうち、レベル2にあたるジャーナルやシリーズの論文数が20%までとなることを上限とする。レベル3は、レベル2の全論文数の中のさらに25%までを上限としている。出版社については、レベル2が全出版社の10%、レベル3はレベル2の中のさらに10%を上限としている。

レベル1については毎年、新しいチャンネルの提案に基づいて分野別パネルが評価を行い、条件を満たせば登録される。レベル2と3については、4年に1度、更新される。

2021年3月末での収録数は表 7の通りであり、実際にはレベル2と3に分類されたものはジャーナル・叢書で10.3%、書籍（出版社）で3.0%である。フィンランド語やスウェーデン語のジャーナルや出版社も含まれている。なお、日本語のジャーナルは38誌収録されており、レベル2以上は無く、レベル1が21誌<sup>25</sup>、レベル0が16誌、未評価1誌である。

---

<sup>25</sup> 21誌は、自然科学が Acoustical science and technology、Bunseki kagaku、Experimental animals、Japanese journal of applied entomology and zoology、Japanese journal of crop science、Japanese journal of hygiene、Japanese journal of physical fitness and sports medicine、Journal of the japanese society for food science and technology-nippon shokuhin kagaku kogaku kaishi、Mokuzai gakkai kaishi、Neurological surgery、Osaka journal of mathematics、Oyo tokeigaku、Rigakuryoho kagaku、Seibutsu-kogaku kaishi、Seikagaku、Shonika rinsho の16誌。人文学が Journal of historical studies、Shakai gengo kagaku、Studies on the inner asian languages の3誌。社会科学が Japanese journal of learning disabilities、Library and information science の2誌である。

表 7 フィンランドの Publication Forum のジャーナルや出版社の構成

		合計	レベル別内訳				
			レベル 3	レベル 2	レベル 1	レベル 0	未評価
総数		34,309	747	2,509	22,367	7,816	600
種類別	ジャーナル・シリーズ	29,857	734	2,352	21,057	5,260	454
	会議	628	0	64	62	502	0
	書籍	3,554	13	93	1,248	2,054	146
言語別 (一部)	英語	21,876	698	2,090	15,664	3,149	275
	フィンランド語	380	0	18	130	192	40
	スウェーデン語	148	0	9	95	42	2
	デンマーク語	0	0	57	8	0	0
	ノルウェー語	0	0	35	2	0	0
	日本語	0	0	21	16	1	0

(出典：www.julkaisuforum.fi より筆者作成)

資金配分における利用については、VITRAに登録された情報を2013年より指標として活用しており、2015年からはPublication Forumによる出版チャンネルの質のレベルで重み付けした指標を用いている。2017-20 年におけるウェイトは以下の表 8のようになっている。

表 8 フィンランドの資金配分における出版物の重み (2017-2020 年)

	レベル 3	レベル 2	レベル 1	レベル 0
ピアレビューされた書籍 (C1)	16	12	4	0.4
ピアレビューされたジャーナル論文 (A1 と A2)	4	3	1	0.1
ピアレビューされた論文や書籍の章 (A3)	4	3	1	0.1
会議予稿集のピアレビューされた論文 (A4)	4	3	1	0.1
ピアレビューされた書籍や特集号の編集 (C2)	4	3	1	0.1
ピアレビューされていない書籍	0.4			
ピアレビューされていない論文や書籍の編集	0.1			

#### (4) イタリア Rating of Scientific Journals<sup>26</sup>

イタリア ANVUR が行う大学の研究評価 (Valutazione della Qualità della Ricerca: VQR) においては、人文・社会科学についてはピアレビューを実施しており、ジャーナルのリストは用いていない。他方で、教授職資格 (National Habilitation) の指標として 2012 年から、また、博士課程プログラムのアクレティテーションのための指標として 2017-18 年から、ジャーナルの格付け(rating)情報が使われている。

ジャーナルの格付け情報を使うのは、論文データベースを用いた指標が不適合である建築学 (分野 08)、古物学、文献学、文学、美術史 (分野 10)、歴史、哲学、教育学 (分野 11)、法律 (分野 12)、経済学および統計 (分野 13)、政治学および社会科学 (分野 14) に限られる。

ANVUR は、「学術ジャーナル」と、質の高い「A クラスジャーナル」の 2 レベルのリストを外部専

<sup>26</sup> <https://www.anvur.it/en/activities/rating-of-scientific-journals/>

<https://www.anvur.it/attivita/classificazione-delle-riviste/classificazione-delle-riviste-ai-fini-dellabilitazione-scientificazionale/elenchi-di-riviste-scientifiche-e-di-classe-a/>

門家 42 名の支援を受けて作成している。学術ジャーナルとして分類される要件は、読者が学界であり、ISSN コードを所有し、少なくともシングルブラインドのピアレビュー方式を採用していることである。さらに A クラスジャーナルに分類される要件は、ダブルブラインドのピアレビュー方式を採用し、研究評価 VQR に提出されたそのジャーナルの論文が同じ学問分野の学術ジャーナルよりも平均して優れた評価を受けている必要がある。また、ジャーナルは、規則的な発行、理事会の構成、学界での普及、アクセスの規則、学術的内容、国際的視点に関する基準を満たしていることも必要である (Cicero and Malgarini, 2020)。2019 年 9 月時点で、学術ジャーナルは 20,000 誌以上、A クラスジャーナルは 5,900 雑誌以上が登録されている。世界中のジャーナルを含んでおり、イタリア語は 13.5% である。

#### (5) スペイン

スペインでは 1980 年代より国内論文の評価の必要性が議論され、1990 年代にジャーナル評価の手法の議論が多くなされた。その結果として多様なジャーナル評価が濫立している (Giménez-Toledo and Román-Román, 2007)。一部を以下に説明する。

一つは、RESH (Spanish Journals of Social Sciences and Humanities) <sup>27</sup> という、グラナ大学とスペイン国立研究評議会 (CSIC) が開発した、人文・社会科学を対象とするジャーナルリストである。このリストは、既に大学評価機関など複数の組織で行われたジャーナル評価の結果を含めた各種の指標と、研究者へのサーベイ調査等の情報を総覧できるものである。具体的な情報としては、①編集の質の指標として、CNEAI、ANECA、および Latindex の 3 つのジャーナル評価で使われている評価基準群への適合・不適合の情報、②普及の指標として、各種のジャーナルデータベースへのジャーナルの登録状況、③評判の指標として、1 万人以上の専門家へ自己の分野における主要な 3 つの国内ジャーナルを調査した結果、④科学的インパクトの指標として、引用数である。

別には、スペイン科学技術財団 (FECYT) が 2008 年より ARCE プロジェクトを開始し、その中でスペインの学術ジャーナルに対して、公募と評価を通じて、「質のシール (quality seal)」を授与している。スペインには 2019 年時点で、国内ジャーナルが 1,818 誌存在すると推定されており、そのうちの 396 誌 (22%) がシールを獲得しており、うち多数を占める 336 誌は人文・社会科学のジャーナルである。また、このような質の認定がなされた後には、評価機関から、人文学と社会科学における認定ジャーナルを対象に、最高の質にあるジャーナルのリストの作成が求められた。現在、分野別にジャーナルを指標によって 25% ずつの 4 グループにランクしたリストが公表されている <sup>28</sup> (De Filippo et al., 2020)。

また、出版社については、スペイン大学出版社協会 (UNE) が、大学評価機関である Agencia Nacional de Evaluación de la Calidad y Acreditación (ANECA) と FECYT との共同により、学術書籍の質のシール CEA-APQ を授与する取組も始めている <sup>29</sup>。さらに、科学研究高等評議会 (CSIC) の Research Group on Scholarly Book (ÍLIA) により、Scholarly Publishers Indicators として出版社のランキングが公表されている <sup>30</sup>。

---

<sup>27</sup> <http://epuc.cchs.csic.es/resh/>

<sup>28</sup> <https://calidadrevistas.fecyt.es/ranking>

<sup>29</sup> <http://www.selloceaapq.es/>

<sup>30</sup> <http://ilia.cchs.csic.es/SPI/index.html>

## (6) フランス

2008年に、当時存在した大学評価機関である Agence d'Evaluation de la Recherche et de l'Enseignement Supérieur (AERES) は、ビブリオメトリクス指標を考慮して3つのクラス (A、B、C) にランク付けした6,305の人文・社会科学のジャーナルリストを公開した。

しかし公開後に、リスト化の作業に参加した専門家名を公表しなかったこと、研究者コミュニティに意見照会を行わなかったこと、ランク付けの基準が不透明であること、低く評価されたジャーナルに悪影響が生じる懸念があることなどの批判がなされ(Pontille and Torny, 2010)、2011年に3つのクラスの評価は取り下げられた。AERESの後継組織である HCERS によって、現在は、経済学・経営学、および言語学のみジャーナルリストが公表されている<sup>31</sup>。

## (7) オーストラリア

オーストラリアでは、大学の研究評価である Excellence in Research for Australia(ERA)を2009年より試行的に開始し、2010年から2~3年おきに実施している。ERAが始まる前の政権では Research Quality Framework (RQF)という別の評価が検討されており、その検討段階の2007年にジャーナルリストの取組が開始された。2010年にはジャーナル20,712誌をA\*、A、B、Cにランク付けし、1,952件の国際会議をA、B、Cにランク付けを行い評価に用いていた。しかし、2011年に、新たに任命された研究大臣はランク付けを撤回し、それ以降はランク付けの無い、評価の対象となりうるジャーナルのリストとして継続されている。撤回の要因としては、トップのA\*の格付けが一部の分野によって不公平になること、引用指標とランクに強い相関があり冗長であること、大学が研究者を評価するためにランクを誤用していることなどの批判がなされたことがあげられる(Ferrara and Bonaccorsi, 2016; Vanclay, 2011)。

---

## 3.2 ピアレビューにおける研究成果の多様性の推奨

ピアレビューは、各分野の研究者が、自身の分野における研究成果の種類などの特徴を踏まえ、また、分野に適合した評価基準のもとに、総合的な判断を行うことができる方法である。

国レベルで実施する研究評価をピアレビューにより行っている国としては、英国、フランス、オランダなどがある。また、イタリアやオーストラリアは、自然科学に対しては指標を中心とした評価を行うが、人文・社会科学ではピアレビューを実施する方法をとっている。このことは、人文・社会科学は、自然科学と比して、ピアレビューが評価のために重要な方法であると考えられていることを反映している。

そもそも、人文・社会科学において論文数や引用数などの指標を使いにくい理由として、自然科学のように支配的なパラダイムが存在せずに、競合するパラダイムが共存するという特徴がある(Reale et al., 2018)。引用数は、支配的なパラダイムが存在するという前提のもとで、特定の論文に引用が集中することが、論文の影響力の大きさを示していると考えられている。異なる概念やアプローチを用いる「学派」が並立する中では、優れた研究成果に引用が集中するという前提は成立せず、複数のピアレビューアーによる多面的で複合的な視点が必要となる。

ピアレビューにおいて、研究成果の多様性を尊重し、推奨する方法としては、被評価側に多様な成果

---

<sup>31</sup> <https://www.hceres.fr/fr/guides-des-produits-de-la-recherche-et-activites-de-recherche-0>

が評価されるという情報を評価資料作成の前段階で提示することである。そのために、各国では、まず①研究を広く包括的に定義し、多様な研究成果が公平に評価の対象になることを明言し、②多種類の研究成果の例示を行っている。

まず①研究の定義に関しては、たとえばオーストラリアの大学研究評価である ERA2018 では、OECD の Frascati Manual における「研究および試験開発 (R&D)」の定義である、「知識の蓄積を増すためや、利用可能な知識の新たな応用の考案のために行われる創造的で体系的な仕事。そこには、人類、文化、社会に関する知識も含む」を引きながら、以下のように定義している。

「新しい知識の創造、および/または新しい概念、方法、発明および理解を生み出すための新しい創造的な方法での既存の知識の利用。これは、新しく創造的である限り、既存の研究の統合と分析を含みうる。」(ERA 2018 Submission Guidelines、1.8 節)

同様に、英国の REF2021 では、以下のように研究を定義している。

「研究とは、新たな洞察が生まれ効果的に共有されるような調査(investigation)のプロセスと定義される。これには以下のものを含む。商業、産業、文化、社会のニーズや、公的セクターおよび第三セクターに直接関連する活動。学術的活動。新たな洞察やその改善へとつながるような、アイデア、イメージ、パフォーマンス、デザインを含む芸術品の発明と形成。新たな材料、デバイス、製品、プロセス、デザインや建設やそれらの改善につながる試験開発における既存知識の利用。一方で、新たな分析技術の開発には当たらない、国の標準の維持などの材料、コンポーネント、プロセスの定期的な試験や分析は含まない。また、オリジナルな研究を具体化するものではない教材開発も含まない。」(“Guidance on submissions’(REF2019/01) Annex C: Definitions of research and impact for the REF”)

次に、②測定対象の研究成果の種類具体的な提示については、特に人文・社会科学においていくつかの例示を行っている。オーストラリア ERA では、研究成果の評価について 22 大分野および 157 細分野ごとに、指標のみを用いて簡素な評価を行うか、ピアレビューを実施するかを決定している。ピアレビューの対象となる分野は人文・社会科学や情報学であり、研究成果の種類には、「伝統的な成果」(書籍、章、ジャーナル論文、学会予稿)以外の「非伝統的な研究成果」を含むとされ、全体としては表 9 オーストラリア ERA における伝統的および非伝統的研究成果のように例示されている。

**表 9 オーストラリア ERA における伝統的および非伝統的研究成果**

【伝統的研究成果】	【非伝統的研究成果】
・ 書籍	・ オリジナルな創造的作品
・ 書籍の章	・ 創造的作品の実演
・ ジャーナル論文	・ 記録や表示された創造的作品
・ 学会予稿	・ 監督や実施された公開展示会やイベント
	・ 外部機関への研究報告書
	・ ポートフォリオ (一連の研究成果のかたまり)

(出典：Australian Research Council(2017) ERA 2018 Submission Guidelines をもとに筆者作成)

「非伝統的な研究成果」の場合には、研究の説明書を 2,000 文字 (約 250 語相当) 以内で提出する。

そこには研究背景（研究分野、文脈、研究目標）、研究の貢献（イノベーション、新しい知識）、研究の重要性（卓越性の証拠）を記すことになっている。

他方、英国 REF では、「REF の根底にある原則は、あらゆる種類の研究成果が公正かつ平等に評価されることである。サブパネルは、特定の種類の成果を他の種類よりも質が高いまたは低いとは見なさない」（パラグラフ 217）と述べた上で、整理のために以下の表 10 のように研究成果の種類を列挙している。メインパネル D（人文学）については、各種類に関する追加説明を別途行っている。

表 10 英国 REF2021 における研究成果の分類

書籍（あるいはその部分）	その他文書
A - 著書	F - 特許・特許出願
B - 編著	J - 作曲
C - 書籍の一章	K - デザイン
R - 解説書	N - 外部組織向けの研究報告
ジャーナル論文	O - 外部組織向けの非公開の研究報告
D - ジャーナル論文	デジタルの人工物
E - 会議報告	G - ソフトウェア
U - ワーキングペーパー	H - ウェブサイトのコンテンツ
物理的な人工物	Q - デジタルまたはビジュアルメディア
L - 美術品・工芸品	S - 研究データ、データベース
P - 装置・製品	その他
展示・実演	V - 翻訳
M - 展示	T - その他
I - 実演	

（出典：REF2021 Panel criteria and working methods より一部抜粋して和訳）

フランス HCERES が行う研究ユニット評価では、分野ごとに、提出する研究成果および研究活動の情報について、種類やその定義に関してガイドラインを作成している（分野数は人文・社会科学 6、科学工学 6、生命科学・地球科学 6。ただし分野ごとに統合あるいは分割してガイドラインを作成している）<sup>32</sup>。表 11 は人文・社会科学の中の「歴史学・美術史学・考古学」のガイドラインに記された研究成果の項目の一覧である。オーストラリアや英国の事例と比べても、教育向けや一般市民向けの成果物を含めて、多様な研究成果を提示している。

<sup>32</sup> <https://www.hceres.fr/en/guides-research-output-and-activities-0>

表 11 フランス HCERES の「歴史学・美術史学・考古学」における研究成果のガイドライン

<p><b>I. 著作物</b></p> <ol style="list-style-type: none"> <li>1. モノグラフ、著作物、批評書、翻訳</li> <li>2. 学術書の編集</li> <li>3. その他（刊行・出版された博士論文、考古学調査報告書、発掘記録、史料編纂、書誌ツールの制作、芸術解題目録、展覧会カタログなど）</li> </ol> <p><b>II. 雑誌</b></p> <ol style="list-style-type: none"> <li>1. 学術論文</li> <li>2. 総説論文および学術報告書</li> <li>3. その他の論文（ピアレビューなしの論文など）</li> </ol> <p><b>III. シンポジウム、学術会議、研究セミナー</b></p> <ol style="list-style-type: none"> <li>1. シンポジウム・学術会議のプロシーディングス論文、学術的著作物の章、学術辞典への寄稿</li> <li>2. シンポジウム・学術会議および研究セミナーで発表したその他の成果（口頭発表や掲示物による発表、大学の夏期講習や博士課程ワークショップにおける教育など）</li> <li>3. 大学の夏期講習や博士課程ワークショップにおける講演</li> </ol> <p><b>IV. 手段および方法の開発</b></p> <ol style="list-style-type: none"> <li>1. プロトタイプやデモンストレーション用モデル</li> <li>2. プラットフォームおよび観測所</li> <li>3. 史料、コーパス、調査フィールド</li> </ol>	<p><b>V. デジタル成果物およびツール</b></p> <ol style="list-style-type: none"> <li>1. ソフトウェア</li> <li>2. データベース・地理情報システム</li> <li>3. コーパス</li> <li>4. 大学のブログおよび研究ノート</li> </ol> <p><b>VI. 特許、ライセンスおよび発明申請</b></p> <p><b>VII. 教育活動の成果物</b></p> <ol style="list-style-type: none"> <li>1. 刊行物（教科書など）</li> <li>2. eラーニング、MOOCs、マルチメディア講義など</li> </ol> <p><b>VIII. 一般市民向け成果物</b></p> <ol style="list-style-type: none"> <li>1. ラジオ・テレビ番組、定期刊行物</li> <li>2. 普及目的の成果物</li> <li>3. 科学的仲介物</li> <li>4. 学術と社会に関する議論（ポッドキャストやイベント）</li> </ol> <p><b>IX. 特定学問分野固有のその他の成果物</b></p> <ol style="list-style-type: none"> <li>1. 芸術作品</li> <li>2. 舞台演出</li> <li>3. 映画およびドキュメンタリー作品</li> <li>4. 展示（キュレーションまたは参加）</li> <li>5. 発掘の活用</li> </ol>
---	---

（出典：HCERES, Guide des produits de la recherche et des activités de recherche - Sous-domaines SHS 6 : Mondes anciens et contemporains - Disciplines : Histoire, histoire de l'art, archéologie より筆者作成）

オランダでは、人文学を対象に、人文学部の学長らによる組織が中核となり、人文学における大学研究評価(SEP)での自己評価の方法と多様な指標を例示した Quality and Relevance in the Humanities (QRiH)というプロジェクトがある<sup>33</sup>。SEPでは自己評価において、学術面だけでなく、社会的関係性（レレバンス）の面から、「ピア向けの成果物」「ピアによる利用」「ピアからの認知」「社会向けの成果物」「社会による利用」「社会からの認知」の6領域での指標の記載を求めている。QRiHでは、人文学の中の17の分野ごとに、領域ごとの多様な研究成果の種類について検討を行い、指標群全体を表12のように示している。各指標については、その定義等を明記している。2018年に行われた各大学の人文学の研究評価では、多くの大学にて活用されている(Honingh et al., 2019)。これらの指標に加えて、QRiHでは、前節のノルウェーなどの取組と同様に、人文学におけるジャーナルや出版社のリストを、読者層に応じてわける方式で（分野、詳細な専門分野、ハイブリッドなど）作成している。

<sup>33</sup> <https://www.qrih.nl/en>



表 12 オランダ QRiH における指標 (出典: <https://www.qrih.nl/en>)

ピア (同分野研究者) 向けの成果物	社会向けの成果物
<p>【パネルにより認証された指標】</p> <ul style="list-style-type: none"> <li>・ 学術書、歴史的資料、展覧会カタログ</li> <li>・ ジャーナル論文・レビュー論文</li> <li>・ 本の章</li> <li>・ 編著・特集号の編集</li> </ul> <p>【妥当な指標】</p> <ul style="list-style-type: none"> <li>・ 学会予稿</li> <li>・ デジタルインフラストラクチャ、データベース</li> <li>・ ウェブサイト</li> <li>・ 映画、ドキュメンタリー、展示会、その他の AV 成果</li> <li>・ ソフトウェア</li> <li>・ デザイン</li> <li>・ 委託報告書</li> <li>・ 研究集会での講演</li> <li>・ 学術的な会議の開催</li> <li>・ 科学コラム、ブログ、フォーラム</li> <li>・ その他の研究成果</li> </ul>	<ul style="list-style-type: none"> <li>・ 専門職に関わる成果</li> <li>・ 一般向けの成果</li> <li>・ (ハイブリッドな研究成果は、ピア向けの成果物欄に記載)</li> </ul>
ピアによる利用	社会のステークホルダーによる利用
<p>【パネルにより認証された指標】</p> <ul style="list-style-type: none"> <li>・ 書籍、編著、展示、その他の研究成果の書評・批評</li> </ul> <p>【妥当な指標】</p> <ul style="list-style-type: none"> <li>・ データセット、データベース、ソフトウェアツールまたは研究施設の使用</li> <li>・ 論文、書籍、その他の出版物の引用 (適切な場合のみ。分野により標準化された値)</li> <li>・ その他のピアによる利用</li> </ul>	<ul style="list-style-type: none"> <li>・ 社会における団体等と連携したプロジェクト</li> <li>・ 受託研究</li> <li>・ 教育における製品の使用</li> <li>・ 研究施設、データベース、デジタルインフラストラクチャ、ツール、ソフトウェア、アプリ、デザインの研究施設の社会による使用</li> <li>・ メディアでの成果物の批評</li> <li>・ 論文、書籍、レポート、論文・書籍・レポートのウェブサイト、その他の成果物の、専門職分野や社会での引用</li> <li>・ 研究と関係がある専門職やその他ユーザーのネットワークへの研究者の参加</li> <li>・ 社会による使用に関するその他の説明</li> </ul>
ピアからの認知の証拠	社会のターゲットグループからの認知の証拠
<p>【パネルにより認証された指標】</p> <ul style="list-style-type: none"> <li>・ 個人に授与される研究助成金</li> <li>・ 主要な研究プロジェクトに授与される助成金</li> </ul> <p>【妥当な指標】</p> <ul style="list-style-type: none"> <li>・ 研究プロジェクトに授与される助成金</li> <li>・ ジャーナルまたは出版社の編集委員会のメンバーシップ</li> <li>・ 著名な学術評議会・委員会のメンバーシップ</li> <li>・ 賞</li> <li>・ 招待講演</li> <li>・ 他の学術機関や研究機関への併任</li> </ul>	<ul style="list-style-type: none"> <li>・ 一般の賞</li> <li>・ 社会における理事会、評議会、諮問機関のメンバーシップ</li> <li>・ 社会の組織・機関への併任</li> <li>・ 公開講演会・公演への招待</li> <li>・ 社会の基金・組織・機関による研究に対する財政的・物質的支援</li> <li>・ 社会・専門職組織・機関によるその他の認知の証拠</li> </ul>

日本の場合には、国立大学法人評価の中で大学改革支援・学位授与機構が行う現況分析および研究業績水準判定において、いくつかの例示がなされている。第二期評価 (2016 年) においては、「学系列の教育・研究水準の評価にかかる参考例」として、人文科学、社会科学を含めて、分野ごとに測定を行うことが予想される研究成果の種類、ならびに、研究業績水準判定において、個々の研究業績が学術面、あるいは社会・経済・文化面の貢献の面で卓越していることを大学が説明するための根拠・資料の

例示を参考文書として行っていた（林、土屋 2016、林 2018）<sup>34</sup>。第三期評価（2020年）では、「現況調査表ガイドライン」として、学系ごとに現況調査表に記載する項目を定め、その中で研究成果を測定する項目を示している。ただし、たとえば人文科学系では、「著書数」（日本語、外国語）、「招待論文数」（日本語、外国語）、「査読付き論文数」（日本語、外国語）、「その他」の4種であり、他国の状況と比べれば限定された種類となっている。

以上のように、海外諸国では、研究を広く定義し、多様な研究成果を評価することを表明している事例がみられた。しかしながら、そのような取組をしたとしても、たとえば英国では、研究成果発表がジャーナル論文に偏ってきているという現実がある（次章参照）。そのため、英国では2020年10月に科学担当大臣 Amanda Solloway が、REF が研究成果をどこに出版するかを重視し、研究目的の多様性を制限していると述べ、REF の再検討を要請している<sup>35</sup>。また、大学人により”hidden REF”<sup>36</sup>として REF では把握が不十分な取組（教材、シチズンサイエンス、標準化、芸術活動等）および、それを支える研究者以外の人材（支援者等）を把握するという取組も行われるなど、多様な研究成果の把握については問題意識が継続している。

---

### 3.3 「責任ある研究評価」概念と指標の関係

人文・社会科学に限定されることなく、全ての分野について、多様な研究成果を測定して評価することが重要という考え方は、近年、さらに広まりを見せており、同趣旨の提言が国際的に複数、発出されている。

最初のまとまった提言として参照されているのは、米国細胞生物学会年次会議での会合を発端とする「研究評価に関するサンフランシスコ宣言(DORA)」<sup>37</sup>であり、ジャーナル・インパクト・ファクターの限界を指摘し、科学者を評価する際に代替指標として用いないこと等を勧告した。2020年12月現在で、2,103組織、16,727個人が署名をしている。

その後、科学計量学者らによりまとめられ、2015年に Nature 誌に発表された「ライデン声明(Leiden Manifesto)」では、研究評価における計量データ利用について包括的にあるべき姿を示した10項目の原則を述べた<sup>38</sup>。また、英国 HEFCE からの委託により行われた、研究評価における指標活用に関する調査レポートである「The Metrics Tide」が2015年に公表された<sup>39</sup>。これを受けて、UK Forum for Responsible Research Metrics が設立される<sup>40</sup>など、「責任ある研究測定 Responsible Research Metrics」という概念が形成された。

さらに、Curry et al.(2020)は、測定に限らず、より広く「責任ある研究評価 Responsible Research Assessment」概念の整理を行った。同レポートによれば、「責任ある研究評価」とは「多様で包摂的な研究文化のもとで、複数の異なる特性を有する質の高い研究を促し、把握し、報奨するような評価のア

---

<sup>34</sup> [http://www.niad.ac.jp/n\\_shuppan/project/\\_icsFiles/afieldfile/2016/01/20/no09\\_nr15-1-0903.pdf](http://www.niad.ac.jp/n_shuppan/project/_icsFiles/afieldfile/2016/01/20/no09_nr15-1-0903.pdf)

<sup>35</sup> <https://www.gov.uk/government/speeches/science-minister-on-the-research-landscape>

<sup>36</sup> <https://hidden-ref.org/about/>

<sup>37</sup> <https://sfdora.org/read/read-the-declaration-japanese/>

<sup>38</sup> <https://www.nistep.go.jp/wp/wp-content/uploads/NISTEP-STIH2-4-00050.pdf>

<sup>39</sup> <https://re.ukri.org/sector-guidance/publications/metric-tide/>

<sup>40</sup> <https://www.universitiesuk.ac.uk/policy-and-analysis/research-policy/open-science/Pages/forum-for-responsible-research-metrics.aspx>

プローチ」とされる。表 13 に示すように現時点で、世界各地域・各国にて同様の趣旨の提言が少なくとも 15 編存在している。

これらの「責任ある研究評価」に共通している考えは、定量的評価は定性的な専門家による評価を支援すること、評価基準は評価対象の研究ミッションに沿って検討する必要があること、指標は分野による多様性を考慮すること、評価委員会は評価対象の研究の質を客観的に評価できる関連専門家で構成されることなど、があげられる<sup>41</sup>。

提言では人文・社会科学を含めた多様性の問題だけでなく、欧州委員会の「Next-generation metrics: Responsible metrics」(Wilsdon et al., 2017)や欧州大学協会の「Roadmap on Research Assessment in the Transition to Open Science」(European Association of Universities, 2018)、などにみられるように、オープンサイエンスやオープンデータが普及しているという文脈の中でデータ公開などの、論文等とは異なる研究成果の形態も評価することの必要性が指摘されていることが最近の特徴である。

このような「責任ある研究評価」と同様の趣旨の提言書が、日本でも日本学術会議により準備されており、研究の多様性の尊重や定性的評価の重要性が指摘されている<sup>42</sup>。

表 13 責任ある研究評価に関連する提言 (出典: Curry et al. 2020 より作成)

1.	DORA: The San Francisco Declaration on Research Assessment (2012)
2.	The Leiden Manifesto for research metrics (2015)
3.	The Metric Tide (2015)
4.	Science in Transition (2013)
5.	Hong Kong Principles for Assessing Researchers (2019)
6.	HuMetricsHSS: Humane Metrics Initiative (2016)
7.	INORMS Research Evaluation Working Group の取組 (SCOPE model 等) (2018)
8.	EC Open Science Policy Platform による”Next Generation Metrics” (2017)
9.	Science Granting Councils Initiative (2015)
10.	Global Young Academy Working Group on Scientific Excellence による”Publishing models, assessment, and open science” (2018)
11.	Helsinki Initiative on Multilingualism in Scholarly Communication (2019)
12.	FOLEC: Latin American Forum on Research Assessment (2019)
13.	Science Europe による”Position Statement on Research Assessment Processes” (2020)
14.	European University Association (EUA) による”Roadmap on Research Assessment in the Transition to Open Science” (2018)
15.	Wellcome Trust’s campaign to Reimagine Research (2019)

これらの提言は、ピアレビュー等の定性的評価の中で、多様な種類の研究成果情報を評価者に提供する方や、そのようにして得られたピアレビュー結果を資金配分等のための一指標として扱うという考えとは整合的である。ピアレビューでは、その分野の中で重要な研究成果をその発表形態に拠らずに評価することができるし、オープンデータのような新たな成果も認識することができる。また、学

<sup>41</sup> <https://custom-eur.cvent.com/9565b735100f4e0cbf1035e1de6b3240/files/event/7ca86a3d6e6f4d1198e9f01fe69fdf46/bfab7226fb2e416db349fa0a572d9edc.pdf>

<sup>42</sup> 日本学術会議(2020)「学術フォーラム：学術振興に寄与する研究評価を目指して」趣旨説明資料 <http://www.scj.go.jp/ja/event/pdf2/287-s-0829-t0.pdf>

術面の視点だけでなく、より広い視点からも評価することができる。

一方で、ノルウェーモデルのような研究成果発表メディアのリスト化や格付けについては、この提言との関係をどう整理するべきか、丁寧な議論が必要となる。「責任ある研究評価」概念との整合性を図るため、発表メディアのリスト化を行う側はいくつかの留意点を検討している(Pölonen et al., 2021)。以下に主要な論点を検討する。

第一の論点は、指標の活用目的を明確化し、研究者個人の評価などでの不適切な利用がなされないことである。「責任ある研究評価」においては、ジャーナルのインパクトファクターのようなジャーナル単位の指標を、個人を単位とする研究評価に利用することは否定されている。フランスやオーストラリアでジャーナル格付けに批判的な指摘がなされた背景には、フランスでは「出版している研究者」と「していない研究者」の区別に使われ、オーストラリアでは研究ユニットという小さい単位の評価における情報として使われたからという指摘もある(Pölonen et al., 2021)。ノルウェーでは、指標は個々の研究者の研究成果の尺度としては適切でなく、「個人レベルで出版物の定性的評価を置き換えたり、シミュレーションすることはできない」(UHR, 2004)と明確に指摘されており、大学単位の集計値として使うことのみを想定している。あくまでも大きな組織単位で使う場合には、そのような測定が、個々の研究成果のピアレビューよりも効率的で有効であるということである。ただし、当然、組織のミッションも多様であり、研究成果はその一つの指標でしかない。フィンランドでも Publication Forum のユーザーガイドを公表し、DORA 等を引用しつつ、研究ユニットや研究者の評価においては、ピアレビューを支援する目的で Publication Forum の分類を使うことを求めている<sup>43</sup>。

第二の論点は、発表メディアの格付けは、インパクトファクターのような連続量でなく、最低限の「学術的」要件を設定するとともに、名声ある少数のジャーナルを区別している程度ということである。自然科学においては、インパクトファクターへの批判は存在しながらも、研究者がインパクトファクターの高いジャーナルへ論文を掲載する競争が生じており、結果として質の高い論文が集まるという状況も生まれている。しかし、人文・社会科学では多くの著者は必ずしも平均引用数の高いジャーナルに論文を送るのではなく、トピックの関連性の高いジャーナルに論文を送る傾向があるという指摘もされている(Ferrara and Bonaccorsi, 2016)。そのため、引用数により連続量で評価を決めることは、特に不適当となる。しかし、学術的であるという要件を満たすジャーナルと、特に秀でた少数のジャーナルを、専門家間で識別するだけであれば、それには意味を見いださうという考えである。

第三に、そのような判定はアカデミアにより行われるということである。人文・社会科学では研究の質を構成する概念は複数のものがあり(Hug et al., 2013; Ochsner et al., 2013)、ジャーナルの質についても研究者により意見の相違がある可能性はある。ただし、上記のように、細かく格付けを行う場合にその中間的な位置に整理されるジャーナルについては、研究者によって付ける順位の違いがあったとしても、トップに位置する少数のジャーナルについては、研究者の判断は比較的的一致するとされている(Ferrara and Bonaccorsi, 2016)。そのため、アカデミア自身が、その質の感覚を踏まえて作成した格付けであれば、特に上位(非常に優れているもの)や下位(学術的ジャーナルではない)とした評価については信頼が置けるという考えである。

第四に、このような格付けによる影響の継続的モニタリングをあわせて導入することである。ノルウェーでは、ジャーナル格付けとそれによる資金配分による影響についての調査がなされてきた(Aagaard et al., 2015; Schneider et al., 2016)。結果では、出版のタイプ(書籍、論文)、言語、論文の長

---

<sup>43</sup> <https://julkaisuforum.fi/en/user-guide>

さに大きな変化は生じておらず、共著率の変化も他国と同様であり、研究者の行為変化が生じているとは見られないというものであった。一方で生産性は向上しており、レベル 2 にあたる出版物の数も増えている。この他にも、ジャーナルの格付けは北欧諸国では、研究成果の出版行為に対して、ピアレビューを伴う出版物の増大などの変化はもたらしたが、予想外の悪影響はもたらしていないことが分析では主張されている (Hammarfelt, 2018; Ingwersen and Larsen, 2014)。仮に問題があれば、指標利用を見直すことができる仕組みを入れておくことが必要となる。

以上のような留意点を十分に学界等の関係者と検討し、継続的に指標の影響を確認して見直すのであれば、出版メディアの情報を使っていることをもってすぐに「責任ある研究評価」概念と対立するとはまでは言えないという見方はできよう。しかし、それでも「出版物」という形態以外の成果は測定対象になりにくく、オープンサイエンスの取組を含めて「多様で包摂的な研究文化」を推進することができるかという点では限界がある。少なくとも、測定や評価の仕組みを設計する際には、事前に関係者との十分な検討と対話を行うことは必須となる。

## 4. 日本における研究成果測定の実行

前章でみたように、海外においては人文・社会科学の研究成果の測定の仕組みが構築され、また、評価における多様な指標の例示と定義がなされている。このような仕組みは、現時点の日本の状況下においてどの程度利用できるであろうか。

以下では、人文学から歴史学、社会科学から経営学を例にとり、日本の研究成果の特徴を分析する。4.1 節で分析対象を説明した後、4.2 節では、そもそも日本における研究成果の多様性がどのような状態であるかを確認する。次に 4.3 節でノルウェーモデルのような出版チャンネルについてリスト（登録簿）を作ることが可能か予備的な分析を行う。4.4 節では、ピアレビューにおいて用いる指標とその標準化の可能性を検討する。

### 4.1 分析対象

分析対象は、日本の 2008 年実施、ならびに 2016 年実施の 2 回の国立大学法人評価における研究水準の現況分析における研究業績水準判定において、国立大学法人および大学共同利用機関法人の学部・研究科等から提出された研究業績説明書に記載されている研究成果（論文、書籍など具体的な成果物名称）を対象にする<sup>44</sup>。2 回で提出の方法は異なるが、提出された研究成果数は、2008 年は 19,626 本（重複提出もあり、延べでは 20,242 本）であり、2016 年は 13,475 件の提出研究業績書類中の 32,453 本である。その中でも以下の分析では、「史学」と「経営学」の細目に提出された研究成果を用いる<sup>45</sup>。

また、比較として、北欧等の具体的な研究成果データは入手不可能であったため、評価への提出資料が公開されている英国の大学研究評価 REF（2014 年実施）を用いる。英国はピアレビューによる大学評価(REF)を実施している国であるが、英国との比較から、日本の歴史学や経営学における研究成果発表にはどのような特徴があるのかを確認する。ただし、大学が提出する研究成果について、日本では学部・研究科等の専任教員数の半数程度を上限として厳選して提出するが、英国では教員 1 人あたり 4 編を提出する。そのため、英国のほうが数が多く、多様な成果を含む可能性がある。

研究成果提出数に大きな違いがあるため、日本については参考に、科学研究費助成事業における研究成果報告書に書かれた研究成果についても分析対象とする。科学研究費助成事業金に採択された研究課題は、事前評価（採択審査）を経て採択されているため、一定の質を事前には担保されている。しかし、報告書に研究成果の記載数の制限はなく、原則的には、研究課題に関係する成果が選定されずに、全てが記載されていると考えられる。データは「科学研究費助成事業データベース」から入手し、科学研究費助成事業の種目のうち、博士課程学生やポスドク向けの「特別研究員奨励費」と職員向けの「奨励研究」を除く全種目を対象とした。また、国立大学法人評価と評価対象大学や、対象期間の条件を合わせるために、研究代表者が国立大学法人または大学共同利用機関法人に所属しており（研究実施期間中の一部期間のみに所属していた場合を含む）、2003～2015 年の間に採択された全研究課題に

<sup>44</sup> 2008 年は研究業績説明書は、各学部・研究科等の専任教員の半数の数を提出数の上限として一つの研究業績説明書が一つの研究成果として提出された。また、大学単位の重点研究分野についての研究業績説明書も提出された。2016 年は専任教員数の 20%を提出数の上限とし、一つの研究業績説明書には最大 3 件までの研究成果を記載する方法がとられた。

<sup>45</sup> 大学改革支援・学位授与機構と政策研究大学院大学科学技術イノベーション政策研究センターの共同研究「イノベーション・システムを推進する大学の研究・教育指標の研究」に基づき実施した。

よる 2004 年～2015 年に出版された研究成果を対象とした<sup>46</sup>。

分析対象は歴史学と経営学である。ただし、歴史学について、英国 REF では考古学は歴史学とは別の「古典学」に分類されている。そのため、日本の国立大学法人評価、ならびに科学研究費助成事業のデータについても、分科「史学」の中の「史学一般」「日本史」「アジア史・アフリカ史」「ヨーロッパ史・アメリカ史」のみを用い、「考古学」は除外した。

---

## 4.2 日英における歴史学・経営学の研究成果の種類

### 4.2.1 英国に関する予備的分析

まず、英国の状況を確認する。英国では先述のように、REF において研究成果の定義を広くとり、研究成果の多様な種類を例示していた。ただし、現実には多様な研究成果がでているか否かは別の問題である。

英国では 1992 年以降の評価について、提出された研究成果の集計値(1992 年)あるいは全ての研究成果のリスト (1996 年以降) を公表しており、その内容を第三者が分析することが可能である(たとえば、Digital Science, 2016、佐藤 2018)。以下の図 1 には、まずは全分野を合計した推移を示す。ただし、年により研究成果の分類の数や内容は異なる。そのため、共通している、①著書・編著、②書籍の章、③ジャーナル論文、④会議への貢献、と⑤その他に大きくわけた。なお、RAE1996 以降は一人当たり 4 件までの研究成果を提出するが、RAE1992 はそれ以降とは提出方法が異なり、2 件までの出版物と 2 件までの出版物以外の様式の成果を提出するようになっていた。そのため、主には RAE1996 以降のみが、時系列で比較可能である。

図 1 では、全分野合計で見た場合に、ジャーナル論文の割合が RAE1996 の 62%から RE2014 では 81%へと増加しており、それ以外の研究成果種類はみな減少している。つまり、ジャーナル論文が、評価において提出することが望ましい研究成果として認識される傾向が進んできたことがうかがえる。

---

<sup>46</sup> 出版年の記載がない研究成果は対象に含めていない。研究成果については、同じ成果を複数年の成果報告書に重複登録している場合が多くあるため、タイトルが同じ研究成果は一つとカウントした。また、「研究成果報告書概要(欧文)」に掲載された情報は、「研究成果報告書概要(和文)」の英訳であるため、対象に含めていない。

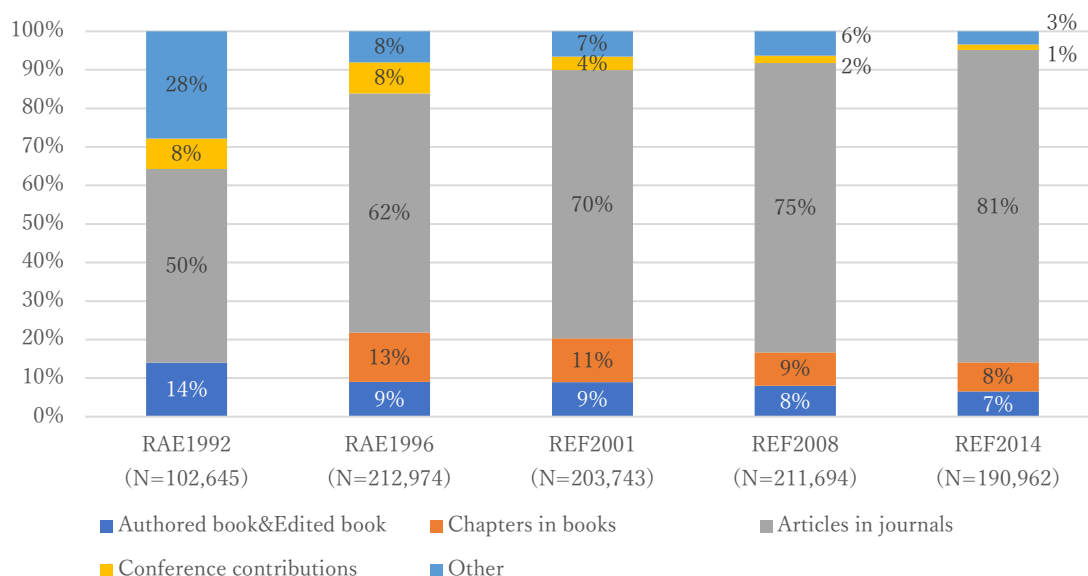


図 1 英国の研究評価における提出研究業績の種類（全分野合計）

※RAE1992は提出件数や対象が他と異なる。また、RAE1992では「Chapters in books」という種別がなく、その一部が「Edited book」および「Other」（うち「Short works」等）に入っている可能性がある。

次に図 2 では歴史学の推移を示す。歴史学においては書籍やその中の章が提出研究成果の半分以上を占めている状況は、RAE1996 から変わらない。しかし、全体傾向と同様に、徐々にジャーナル論文の割合が増加しており、RAE1996 の 31%から RAE2014 では 44%となり、歴史学でも半数弱が既にジャーナル論文となっている。

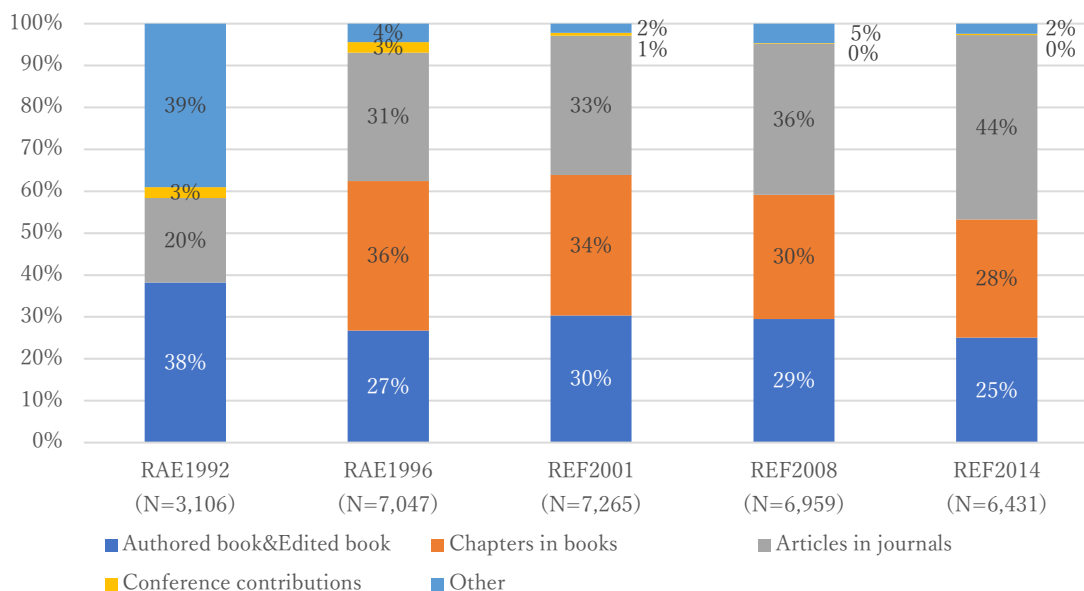


図 2 英国の研究評価における提出研究業績の種類（歴史学）

※上記の注参照

次に図 3 に経営学の結果を示す。経営学の変化は顕著であり、ジャーナル論文の比率が RAE1996 の



59%から RAE2014 では 96%へと増加しており、評価に提出する研究成果として望まれる形態は、ジャーナル論文であるという考えが普及してきているように見える。そもそも経営学においては欧州各国で研究スタイルの米国化が進んでおり（浅川 2019）、トップジャーナルへの掲載を第一に考える傾向が進んでいる。経営学においては、英国の Association of Business Schools Academic Journal Quality Guide や米国・欧州諸国・豪州などで作成されているリストなど、ジャーナルランキングが国際的に少なくとも 18 件構築されており (Vogel et al., 2017)、ジャーナル論文掲載を重視する傾向が加速している（佐藤 2018）。

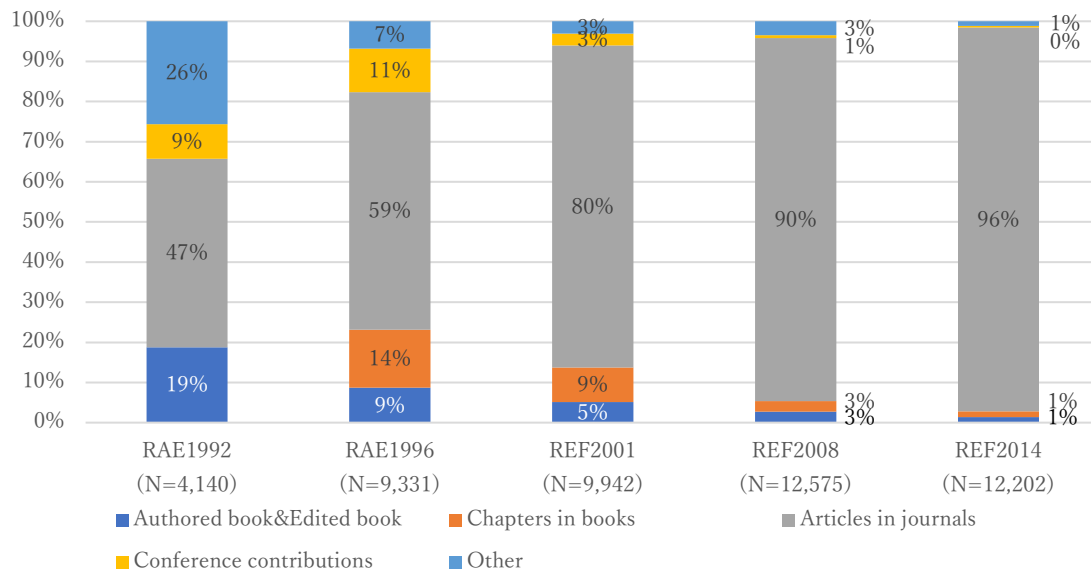


図 3 英国の研究評価における提出研究業績の種類（経営学）

※上記の注参照

以上のように、人文・社会科学といっても、英国では歴史学と経営学では研究成果の種類には大きな違いがある。しかし、どちらもジャーナル論文の割合が増加している傾向にある。

#### 4.2.2 日本の大学評価において提出された研究成果の Scopus 収録率

日本の場合には、英国と同様の形で分析できるような研究成果分類は研究業績説明書には記載されていない。そのため、後述の歴史学と経営学については、目視にて分類を行う。

一方、提出された研究業績説明書に書かれた研究成果が論文データベースに収録されているジャーナル論文か否かは、書誌情報に書かれた、著者名、タイトル、ジャーナル名、巻号・頁数等のデータから論文データベースと突合することで機械的に確認できる。図 4 に 2016 年実施の国立大学法人評価において提出された研究業績説明書に記載された研究成果 32,454 本の Scopus 収録率を分野別（科学研究費補助事業における「分科」レベル）で示している。

全分科合計では69%がジャーナル論文である。同様の分析を、第一期法人評価を対象に WoS 収録率を測定した結果は65%であり（林、土屋 2016）、データベースの違いなどがあるために明確には言えないが、大きな変化は見られないと考えられる。一方で、分野別に見れば、医学分野などでは収録率が高いが、人文学、社会科学では低い。他方で、社会科学で経済学、心理学、経営学は半数近い値であり、また、工学分野にも半数、ないし半数以下の分野（建築学、土木工学など）があり、詳細な分野によって Scopus に収録されるようなジャーナル論文が研究成果である割合は異なる。

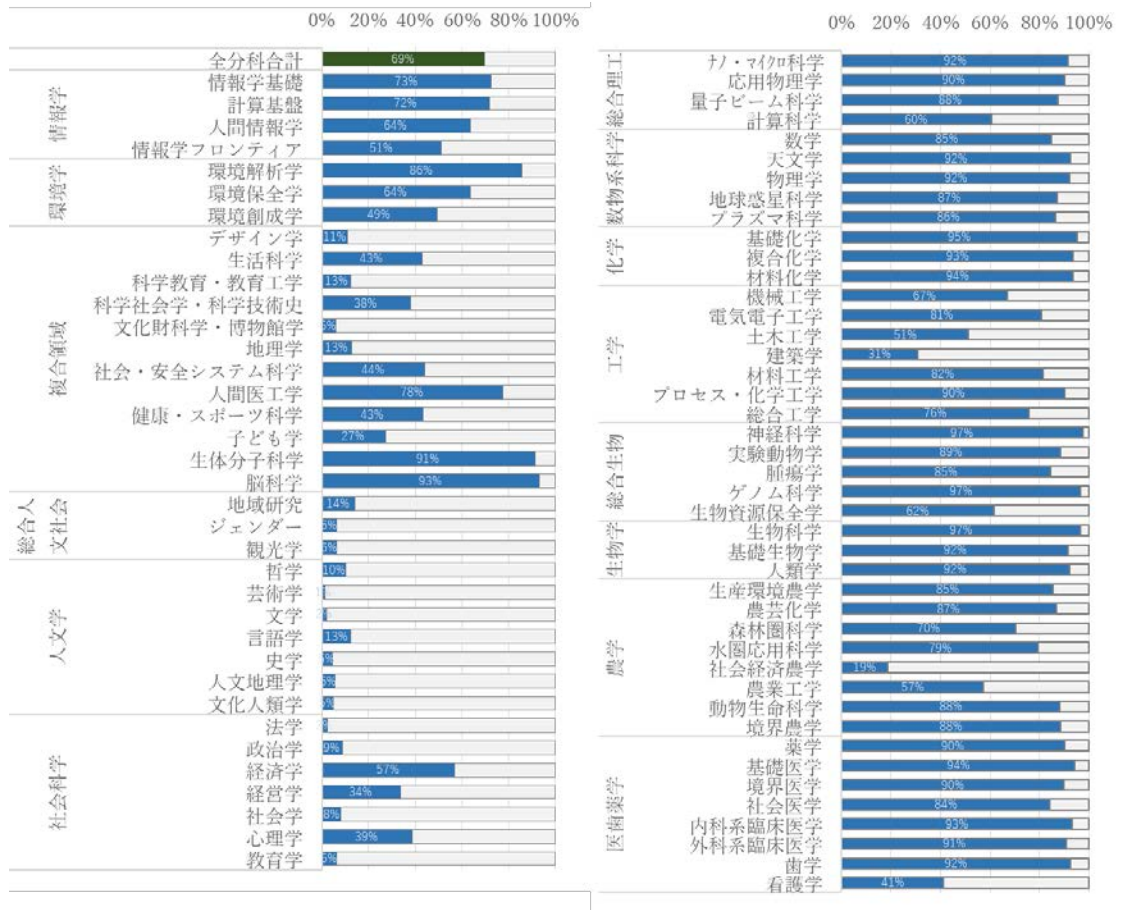


図 4 第二期国立大学法人評価において提出された研究成果の Scopus 収録状況

#### 4.2.3 歴史学における研究成果の種類の比較

次に日本の歴史学と経営学について、過去 2 回（2008 年、2016 年）の国立大学法人評価の研究業績水準判定にて提出された研究成果の種類を、英国 REF2014 と比較する。

歴史学の結果は表 14 である。ジャーナル論文が、英国 REF2014 では 44%を占めていたのに比べ、日本では 2008 年に 22%、2016 年に 26%であり、低い。母比率の検定の結果、日本は 2 時点とも英国 REF と比べてジャーナル論文の比率が有意に低く(p<0.01)、また、日本の 2 時点の間ではジャーナル論文の比率が若干上がっているが、統計的に有意な差ではない(p>0.1)。日本では、ジャーナル論文の代わりに、書籍や編著が多い。一方で、英国は本の章が日本より多い。日本では、歴史資料や報告書などの多様な研究業績も提出されている。

表 14 歴史学における研究評価に提出された研究成果の種類

	NIAD2008 (N=329)	NIAD2016 (N=373)	REF2014 (N=6,431)
著書	35.6%	31.9%	20.5%
編著	19.5%	12.3%	4.5%
本の章	7.6%	14.5%	28.2%
ジャーナル論文	21.0%	26.3%	44.0%
会議	1.5%	4.0%	0.3%
報告書、ワーキングペーパー	7.3%	2.9%	0.9%
データベース	0.6%	1.9%	0.2%
解説・批評書(史料やその解説など)	6.1%	2.7%	0.7%
展示	0.0%	2.1%	0.0%
ほか(学位論文等)	0.9%	1.3%	0.2%
合計	100%	100%	100%

解釈において注意が必要であるのは、上述のように、日本については 300 件程度の、大学や学部・研究科等によって厳選されて提出された研究成果が対象であるが、英国 REF は、教員一人当たり 4 件の提出による 6,000 件以上の研究成果が対象となっていることである。そのため、日本では厳選された過程で、論文よりも著書が優先された可能性はある。

参考に、科学研究費助成事業データベースにおける研究成果を確認する。科学研究費助成事業では、研究成果は「図書」、「雑誌論文」、「学会発表」、「産業財産権」と大きく分けて入力するようになっているが、「産業財産権」については歴史学ではほぼ無いこと、ならびに、「学会発表」については科学研究費助成事業データベース中で 2012-2014 年間の「学会発表」収録数が明らかに他年よりも少なくなっておりデータに問題があると思われることから、「図書」「雑誌論文」のみでの比を示す。ただし、この二つの定義は明瞭でなく、書籍の中の章について、「図書」に記入している場合と「雑誌論文」に記入している場合の双方が見られることから、参考値程度として解釈する必要がある。

結果、表 15 のように、雑誌論文の割合が全年の合計で 72%と大きい。一方で、英国でジャーナル論文が増えていた状況とは逆に、図書の割合が増してきている状況がみられており、日本では雑誌論文（主にジャーナル論文）に集中してきている状況はみられない。

表 15 科学研究費助成事業データベースにおける「史学」の研究成果

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	合計
雑誌論文	354	861	1,179	1,341	1,366	1,472	1,689	1,527	1,558	1,645	1,222	1,266	15,480
	78%	80%	78%	77%	75%	72%	71%	71%	72%	68%	63%	70%	72%
図書	101	218	339	400	463	562	685	625	617	789	707	540	6,046
	22%	20%	22%	23%	25%	28%	29%	29%	28%	32%	37%	30%	28%

#### 4.2.4 経営学における研究成果の種類比較

同様の分析を経営学についても実施する。表 16 に日本の 2 回の評価と英国 REF2014 に提出された研究成果の種類を示している。経営学においても、ジャーナル論文の割合は日本は英国より有意に低く ( $p < 0.01$ )、日本の 2 時点ではジャーナル論文の割合が若干上がっているが統計的に有意な差ではな

い。

表 16 経営学における研究評価に提出された研究成果の種類

	NIAD2008 (N=177)	NIAD2016 (N=297)	REF2014 (N=12,201)
著書	20.3%	13.1%	1.3%
編著	16.4%	2.7%	0%
本の章	1.7%	7.1%	1.5%
ジャーナル論文	57.6%	63.6%	95.6%
会議	2.8%	11.4%	0.4%
報告書、ワーキングペーパー	0.6%	1.3%	1.1%
データベース	0.6%	0.3%	0%
ほか(学位論文等)	0%	0.3%	0.0%
合計	100%	100%	100%

また科学研究費助成事業データベースにおける成果については、全年合計では雑誌論文が 85%と高い (表 17)。経営学では雑誌論文の割合には明確な増減傾向は見られない。

表 17 科学研究費助成事業データベースにおける「経営学」の研究成果

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	合計
雑誌論文	316	593	702	637	823	756	960	1,015	931	923	791	853	9,300
	88%	85%	84%	79%	86%	82%	81%	87%	81%	83%	82%	85%	83%
図書	42	106	136	166	139	168	218	149	213	192	179	156	1,864
	12%	15%	16%	21%	14%	18%	19%	13%	19%	17%	18%	15%	17%

以上の結果をまとめると、日本の大学評価では英国 REF2014 と比べて、提出される研究成果は書籍及び、それ以外の多種類のもが含まれており、英国よりは、多様な研究成果の形態やそれにつながる多様な研究活動が認知されている状況と言える。

ただし、日本の大学評価は英国よりも厳選された研究成果に関するものであり、記載数の上限がない科学研究費助成事業の研究成果をみれば、評価に提出された研究成果よりも「雑誌論文」のカテゴリーが多い。つまり、歴史学、経営学ともに、評価に提出する研究成果では著書を優先している傾向があると見られる。また、科学研究費助成事業の成果においては、両分野ともに「雑誌論文」の割合が増えているような傾向はなく、逆に歴史学では若干、雑誌論文の割合が減少している。つまり英国のようにジャーナル論文に研究成果が集中していく傾向は現在ない。

#### 4.3 日英における出版メディアの状況

次に、提出された研究成果の出版メディアの状況を分析し、ノルウェーモデルが対象とする「学術出版物」の認定や格付けを行える可能性があるかを検討する。そのため、まず 4.2.1 節でジャーナルについて、4.2.2 で出版社について分析する。

### 4.3.1 ジャーナルの傾向

#### ① 歴史学

表 18 には日本の 2 回の国立大学法人評価において提出された研究業績の合計のうち、ジャーナル論文 167 編について、ジャーナル名を集計し、上位 10 誌を示している。ただし、各大学から提出された個々の提出研究業績は非公表であるため、表 18 には (大学との関係が明示されていない集計値の情報だけであるが念のために)、ジャーナル名は匿名とし、国内学会が刊行しているジャーナルか、学会以外の雑誌か、海外のジャーナルかなどの種別を示している。あわせて、後述する科学研究費助成事業による上位 10 誌にも含まれているかを示している。また、研究業績水準判定において 2 人の評価者の少なくともどちらかが「SS (卓越)」と判定した成果の各ジャーナルにおける割合を示している。

さらに、ノルウェーモデルなどの出版チャンネルのリストの収録要件では、査読があることが条件となっており、ERIH PLUS などでは各出版チャンネルのウェブサイトにて査読手順の情報が有るかを確認していたため、表の最後の列には、ジャーナルのウェブサイトにおいて査読の情報があるかを示している (ただし、ウェブサイトに掲載されていなくても、現実には実施されている場合もありうることは注意が必要である)。

結果、歴史学は上位 10 誌のうち、国内学会によるジャーナルが 7 誌を占めており、国内および海外の外国語雑誌が 1 誌ずつ、学会ではなく一般出版社が発行するジャーナルが 1 誌である。10 誌によって 27.5% の研究成果となっている。学会によるジャーナルの多く (全てではない) では査読が行われていることはウェブサイトから確認できたが、投稿規定等が掲載され、どのような体制やプロセスで査読が行われるのかの説明が公開されていたものは日本語のジャーナルではなかった。また、学会でない一般出版社の学術ジャーナルでは査読の有無は確認できなかった。

表 18 日本の大学評価における史学の提出研究業績のジャーナル

	ジャーナルの種別	言語	論文数 (N=167)	割合	累積割合	科研費成果 10 位以内と 同一	SS と判 定され た割合	ウェブサイト での査読情報 の有無 <sup>*1</sup>
1	国内学会のジャーナル	日本語	10	6.0%	6.0%	○	60%	有
2	国内学会のジャーナル	日本語	6	3.6%	9.6%	○	0%	無
3	国内学会のジャーナル	日本語	5	3.0%	12.6%	○	40%	有
3	国内学会のジャーナル	日本語	5	3.0%	15.6%	○	60%	有
5	国内学会の英文ジャーナル	英語	4	2.4%	18.0%		0%	有(説明有)
5	国内学会のジャーナル	日本語	4	2.4%	20.4%	○	25%	有
7	海外大学出版会による外国語 ジャーナル	外国語	3	1.8%	22.2%		0%	有
7	一般出版社による学術ジャー ナル	日本語	3	1.8%	24.0%		100%	無
7	国内学会のジャーナル	日本語	3	1.8%	25.7%		0%	有
7	国内学会のジャーナル	日本語	3	1.8%	27.5%		33%	無

※ジャーナルのウェブサイトにて査読について何らかの言及があれば「有」。サイトで公開された投稿規定などに査読の体制やプロセス、スケジュール等について何らかの説明があれば「有 (説明有)」。

参考に科学研究費助成事業データベースにおける歴史学の研究成果における出現数が多い上位 10 誌を表 19 に示す。国内学会のジャーナルが同様に多いが、大学の研究所や大学院研究科による刊行物、

ならびに特定の学会ではなく出版者が発行者となっている学術雑誌が見られる。表 18 に示したように、このうちの 5 誌は、大学評価における 10 誌、特に頻度上位のジャーナルと重複している。つまり、組織が厳選した成果（大学評価の提出物）とそうでない成果（科学研究費助成事業の成果）で頻度の高いチャンネルの半数は同じであり、研究者が研究成果を数多く発表するジャーナル群と、評判が高いジャーナルが別のグループとして存在している状況が極端に存在するわけではない。また、表 18 にあるように、SS と判定された割合も、言語、学会の雑誌か否か、査読情報の有無によって明確な傾向は見られない。

表 19 科学研究費助成事業の史学における研究成果のジャーナル

	雑誌名	発行者	論文数	割合	累積割合
1	歴史学研究	歴史学研究会	296	2%	2%
2	歴史評論	歴史科学協議会	206	1%	3%
3	東京大学史料編纂所研究紀要	東京大学史料編纂所	207	1%	5%
4	史学雑誌	史学会	197	1%	6%
5	日本歴史	吉川弘文館	147	1%	7%
6	日本史研究	日本史研究会	120	1%	8%
7	東洋史研究	東洋史研究会	109	1%	8%
8	東京大学史料編纂所附属画像史料解析センター通信	東京大学史料編纂所	103	1%	9%
9	アジア遊学	勉誠出版	96	1%	10%
10	史淵	九州大学大学院人文科学研究院	83	1%	10%

なお、発表言語については、大学評価における研究業績では 24%（2008 年は 23%、2016 年は 24%）が英語あるいは外国語である。科学研究費助成事業では全年を平均して 10%が英語あるいは外国語であり、時系列での割合の増減傾向はない。つまり、大学評価へ提出した厳選した研究成果では英語・外国語を提出している割合が若干高くなっているが、時系列で外国語のジャーナル論文を執筆することが増している傾向は存在していない。

表 20 科学研究費助成事業の史学における雑誌論文の外国語割合

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	合計
雑誌論文の数	47	94	110	140	121	102	154	159	192	168	116	161	1,564
うち外国語の割合	13%	11%	9%	10%	9%	7%	9%	10%	12%	10%	9%	13%	10%

他方、表 21 には英国 REF でのジャーナル論文において出現頻度が高いジャーナル 10 誌を示している。英国の場合には大学出版会が出版するジャーナルが多くを占めており、それ以外の商業学術出版社は 2 誌のみである。これら全てにおいて、査読実施の情報を確認することができた。

表 21 には各誌の 2019 年のインパクトファクター（以下 IF）を Journal Citation Reports(JCR)より示しているが、JCR における History 分野のジャーナルの平均 IF（2019）は 0.546 であるため、必ずしも IF が平均よりも高いジャーナルが多いわけではない。

表 21 英国 REF の History 分野における提出研究業績のジャーナル

	ジャーナル名	発行者	JIF2019	論文数	割合	累積割合	ウェブサイトでの査読情報の有無
1	Historical Journal	Cambridge Univ Press	0.791	91	3.2%	3.2%	有(説明有)
2	Past & Present	Oxford Univ Press	0.831	88	3.1%	6.3%	有(説明有)
3	English Historical Review	Oxford Univ Press	0.618	76	2.7%	9.0%	有(説明有)
4	Economic History Review	Wiley	1.063	59	2.1%	11.1%	有(説明有)
5	Historical Research	Oxford Univ Press	-	58	2.0%	13.1%	有(説明有)
6	Journal of British Studies	Cambridge Univ Press	0.479	46	1.6%	14.8%	有(説明有)
7	Modern Asian Studies	Cambridge Univ Press	0.488	40	1.4%	16.2%	有(説明有)
8	History	Wiley	-	39	1.4%	17.5%	有(説明有)
9	Twentieth Century British History	Oxford Univ Press	-	36	1.3%	18.8%	有(説明有)
10	Social History of Medicine	Oxford Univ Press	0.754	36	1.3%	20.1%	有(説明有)

## ②経営学

経営学は歴史学とやや異なる。表 22 に大学評価へ提出された研究成果のジャーナル上位 10 位（同順位があるため 12 誌）を示す。結果、国内学会のジャーナルが経営学でも多いが、特定の大学が設立した学会・研究会が発刊しているジャーナルがみられるとともに、出版社や協会による雑誌、海外学術出版社によるジャーナルなど、多様である。上位 12 誌が提出ジャーナル論文の 27%を占めており、特に一番目のジャーナルが 6%を占めている。

また、査読については、学会以外の出版社や協会によるジャーナルを除けば、査読実施の情報をウェブサイトから確認することができ、さらに、多くでは投稿規定などにおいて査読のプロセスや体制についての説明も記載されていた。

表 22 日本の大学評価における経営学の提出研究業績のジャーナル

	ジャーナルの種類別	言語	論文数 (N=290)	割合	累積割合	科研費成果 10 位以内と同一	SS と判定された割合	ウェブサイトでの査読情報の有無※1
1	国内学会のジャーナル	日本語	18	6%	6%	○	11%	有(説明有)
2	海外学術出版社による英語ジャーナル	英語	8	3%	9%		63%	有(説明有)
3	国内学会のジャーナル	日本語	7	2%	11%		14%	有(説明有)
3	特定大学主体の学会のジャーナル	日本語	7	2%	14%	○	0%	有(説明有)
5	国内学会のジャーナル	日本語	6	2%	16%		0%	有
5	国内出版社の学術誌	日本語	6	2%	18%	○	33%	無
7	国内学会のジャーナル	日本語	5	2%	20%		0%	有
7	国内協会の学術誌（機関誌）	日本語	5	2%	21%	○	0%	無
7	国内学会のジャーナル	日本語	5	2%	23%		0%	有(説明有)
10	国内学会のジャーナル	日本語	4	1%	24%		0%	有(説明有)
10	海外学術出版社による英語ジャーナル	英語	4	1%	26%		50%	有(説明有)
10	海外学術出版社による英語ジャーナル	英語	4	1%	27%		75%	有(説明有)

※ジャーナルのウェブサイトにて査読について何らかの言及があれば「有」。サイトで公開された投稿規定などに査読の体制やプロセス、スケジュール等について何らかの説明があれば「有（説明有）」。

一方、科学研究費助成事業の成果を確認すると（表 23）、大学評価の提出業績よりも、特定の大学が設立した学会・研究会が発行するジャーナルの論文が多い。また、一般出版社による専門雑誌も見られる。この 10 誌のうち、4 誌が大学評価での頻度が高い 12 誌と一致している。また、表 22 にあるように海外学術出版社による英語ジャーナル論文が「SS（卓越）」と判定された割合は、国内ジャーナルより高い。

表 23 科学研究費助成事業の経営学における研究成果のジャーナル

	雑誌名	発行者	論文数	割合	累積割合
1	国民経済雑誌	神戸大学経済経営学会	269	2.7%	2.7%
2	會計	日本會計學會、森山書店	256	2.6%	5.3%
3	企業会計	中央経済社	166	1.7%	6.9%
4	組織科学	組織学会	154	1.5%	8.5%
5	一橋ビジネスレビュー	一橋大学イノベーション研究センター、東洋経済新報社	153	1.5%	10.0%
6	神戸大学大学院経営学研究科ディスカッション・ペーパー	神戸大学大学院経営学研究科	133	1.3%	11.4%
7	経済学研究(北海道大学)	北海道大學經濟學會	122	1.2%	12.6%
8	赤門マネジメント・レビュー	東京大学大学院経済学研究科、特定非営利活動法人グローバルビジネスリサーチセンター	121	1.2%	13.8%
9	横浜経営研究	横浜国立大学経営学会	118	1.2%	15.0%
10	産業経理	一般財団法人産業経理協会	105	1.1%	16.1%

また、言語については、英語あるいは外国語（ほとんどが英語である）の割合が、大学評価の研究業績では 2008 年は 35%、2016 年は 57%（2 回を合計して 49%）である。表 24 の科学研究費助成事業の研究成果では全年を平均して 29%が英語あるいは外国語であり、2008 年が 21%と最も低く、それ以降増加しており、2014 年には 34%となっている。つまり、厳選した研究成果では英語の割合が高くなっており、また、どちらも増加傾向にあり、経営学においては研究成果を英語で発表する傾向が高まり、それが評価で提出されるほど質が高いと認識されている成果において顕著である。

表 24 科学研究費助成事業の経営学における雑誌論文の外国語割合

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	合計
雑誌論文の数	102	159	190	156	176	192	286	296	291	302	267	291	2,708
うち外国語の割合	32%	27%	27%	24%	21%	25%	30%	29%	31%	33%	34%	34%	29%

表 25 には英国 REF でのジャーナル論文において出現頻度が高いジャーナル 10 誌を示している。全てが国際的な学術出版社によるジャーナルであり、全てについて査読実施の情報をウェブサイトで確認することができる。各誌の IF については、JCR における Management 分野の平均 IF は 3.288 であり、やはり必ずしも高いジャーナルのみが提出されているわけではない。



表 25 英国 REF の Management 分野における提出研究業績のジャーナル

	ジャーナル名	発行者	JIF2019	論文数	割合	累積割合	ウェブサイトでの査読情報の有無
1	European journal of operational research	Association of European Operational Research Societies (EURO), Elsevier	4.213	278	2.4%	2.4%	有(説明有)
2	British Journal of Management	British Academy of Management, Wiley	3.023	200	1.7%	4.1%	有(説明有)
3	Journal of marketing management	Taylor & Francis	1.719	200	1.7%	5.8%	有(説明有)
4	International Journal of Production Economics	Elsevier	5.134	193	1.7%	7.5%	有(説明有)
5	Human Relations	SAGE	3.600	189	1.6%	9.1%	有(説明有)
6	Journal of Banking & Finance	Elsevier	2.269	183	1.6%	10.7%	有(説明有)
7	Organization studies	SAGE	3.926	178	1.5%	12.2%	有(説明有)
8	International journal of human resource management	Taylor & Francis	3.040	172	1.5%	13.7%	有(説明有)
9	Journal of business ethics	Springer	4.141	164	1.4%	15.1%	有(説明有)
10	Work, employment and society	SAGE	3.171	162	1.4%	16.4%	有(説明有)

以上の結果をまとめると、歴史学と経営学の2分野で、日本、英国ともに傾向が異なる。歴史学については、日本では大学評価において厳選して提出する研究成果で頻度が高いジャーナルが、科学研究費助成事業の研究成果が多く出版されているジャーナルと上位で一致していることから、少数の質の高い研究成果のみを掲載するジャーナルが明確に区分されている状況とは言いにくい。また、査読実施が不明な、出版社による学術雑誌も厳選された研究成果のチャンネルとなっており、事前の査読があるチャンネルのみに優れた研究成果が発表されているわけでもないと考えられる。一方で英国でも大学出版会のジャーナルが多く、商業学術出版社の関与が少ない。

経営学では、日本の科学研究費助成事業においては大学が発刊するジャーナルが多いが、大学評価において厳選されて提出された成果は国内学会のジャーナルか海外のジャーナルが多い。両方で査読の情報の無い一般出版社によるジャーナル2誌を含めて、4誌が重複している。つまり、学術雑誌としては、大学が発行するジャーナルなどの数多くの成果が出されるジャーナルと優れた少数の成果が出されるジャーナルの違いがある程度存在する。海外のジャーナルリスト(登録簿)では、特定の大学が発刊し、その大学の著者が多いジャーナル(いわゆる紀要など)は、学術出版物の対象から除外されていることが多く、日本においても、発表ジャーナルの使い分けがなされていると見られる。

さらに、査読の存在が明確な、英語ジャーナルへの成果発表が増加する傾向も見られており、また、評価結果でも英語ジャーナルの研究成果が高い結果を得ている傾向が見られる。このような傾向は経営学の研究者の間で、米国等のトップジャーナルへの掲載を第一優先と考える傾向が生じているという指摘(浅川 2019、藤本 2020)と合致している。その中では、国際標準レベルの厳密性と関連性をクリアした研究論文の形成が課題と認識されている。

他方で、上記の分析では、学術界以外の読者も想定しうるような一般出版社による雑誌も、大学が評

価に提出するような質が高いものが掲載されているという状況も同時に見られる。上記の経営学研究者の検討においてもグローバルな発信とローカルな発信の共存が必要という主張もされており、現時点でもそのように両者が存在している状況が反映されていると見られる。

そのため、日本の現状においては、歴史学においては、査読が無いが評価において提出されるようなジャーナルが、ノルウェー等におけるジャーナルリストの基準では落ちてしまう可能性があり、また、格付けも行いにくいと見られる。一方、経営学では日英ともに査読のあるジャーナルのリストなどを作り、特に質の高い論文が掲載されるものを区分することも可能であると思われるが、他方で、一般出版社による質の高い成果が抜け落ちる可能性がある。

#### 4.3.2 出版社の傾向

##### ①歴史学

次に出版社について同様に分析を行う。出版社についても大学評価の提出研究業績については念のために、出版社名は匿名として扱う。

日本の場合は、学術書刊行に限定しない大手の総合出版社に加え、大学出版社、歴史書を中心とする人文書専門出版社や、人文・社会科学書の出版社がならんでいる<sup>47</sup>。上位の10社で47.7%を占めている。いずれの出版社も、査読に関する情報はウェブサイトが存在しなかった。評価で「SS（卓越）」と判定される割合にも明確な傾向は見られない。

表 26 日本の大学評価における史学の提出書籍の出版社

	出版社名	研究成果数 (N=447)	割合	累積割合	科研費成果 10位以内 と同一	SSと判 定された 割合	ウェブサイ トでの査読 情報の有無
1	総合出版社	44	9.8%	9.8%	○	50%	無
2	歴史書出版社	42	9.4%	19.2%	○	45%	無
3	大学出版会	35	7.8%	27.1%	○	51%	無
4	人文書出版社	17	3.8%	30.9%	○	18%	無
5	歴史書出版社	16	3.6%	34.5%	○	19%	無
6	人文・美術書出版社	13	2.9%	37.4%		8%	無
7	総合出版社	12	2.7%	40.0%	○	25%	無
7	大学出版会	12	2.7%	42.7%		42%	無
9	人文・社会科学書出版社	11	2.5%	45.2%	○	45%	無
9	大学出版会	11	2.5%	47.7%		45%	無

科学研究費助成事業の研究成果を多く出版している出版社は表 27 である。大学評価の成果の出版社と多くが重複している。また、表 28 では、英語等の外国語の書籍の割合は6%であり、増減の傾向は見られない。

<sup>47</sup> 研究業績判定においては、提出者（大学）が学術面および／あるいは社会経済文化面の貢献の2つの評価基準を選ぶことができる。本集計は全ての提出業績を対象としているが、学術面の基準への提出だけを集計しても上位10社のうち9社は同一となる。

表 27 科学研究費助成事業の史学における研究成果の書籍出版社

	出版社名	研究成果数	割合	累積割合
1	吉川弘文館	385	6.5%	6.5%
2	山川出版社	371	6.2%	12.7%
3	岩波書店	307	5.2%	17.9%
4	東京大学出版会	249	4.2%	22.0%
5	ミネルヴァ書房	164	2.8%	24.8%
6	勉誠出版	137	2.3%	27.1%
7	昭和堂	137	2.3%	29.4%
8	明石書店	103	1.7%	31.1%
9	汲古書院	95	1.6%	32.7%
10	講談社	93	1.6%	34.3%

表 28 科学研究費助成事業の史学における書籍の外国語割合

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	合計
図書の数	2	13	23	36	17	34	47	35	30	44	36	60	377
うち外国語の割合	2%	6%	7%	9%	4%	6%	7%	6%	5%	6%	5%	11%	6%

英国の場合は表 29 に示すように、上位に大学出版社および国際的に有名な総合学術出版社が並んでおり、さらに、歴史学専門や、人文・社会科学を中心とする学術出版社が数社ある。分析対象とした書籍数は日本の大学評価より多いが、上位 9 社で半数以上を占めており、日本よりも集中度が高い。発刊する書籍の査読については全てにおいて査読がなされることがウェブサイトにて示されている。例えば Oxford University Press は「全ての書籍は刊行決定前に編集スタッフと外部の査読者により評価される」と明記されている。

表 29 英国 REF の History 分野における提出研究業績の書籍出版社

	出版社名	研究成果数	割合	累積割合	ウェブサイトでの査読情報の有無
1	Oxford University Press	370	11%	11%	有(説明有)
2	Palgrave Macmillan	320	9%	20%	有(説明有)
3	Cambridge University Press	256	7%	27%	有(説明有)
4	Ashgate	223	6%	34%	有(説明有)
5	Routledge	143	4%	38%	有(説明有)
6	Manchester University Press	142	4%	42%	有(説明有)
7	Boydell & Brewer	135	4%	46%	有
8	Brill	124	4%	49%	有(説明有)
9	Yale University Press	59	2%	51%	有(説明有)
10	Brepols	57	2%	53%	有

## ②経営学

経営学では、社会科学における学術出版書を中心としている出版社が多いが、学術書だけでなく一

般向けのビジネス書も発刊する出版社、学術書以外も発刊する総合出版社も含まれ、また、海外の学術出版社が1社含まれている。歴史学よりも特定の出版社への集中が高く、上位8者で半数を占めている。出版前の査読については、ウェブサイト情報が掲載されていた出版社は、海外の学術出版社以外には無かった。

表 30 日本の大学評価における経営学の提出研究業績の書籍出版社

	出版社名	研究成果数	割合	累積割合	科研費成果 10位以内と 同一	SSと判 定された 割合	ウェブサイト での査読情報 の有無 <sup>※1</sup>
1	ビジネス書出版社	21	16%	16%	○	33%	無
2	社会科学・人文書出版社	13	10%	25%	○	31%	無
3	社会科学出版社	7	5%	31%	○	29%	無
3	ビジネス書出版社	7	5%	36%	○	57%	無
5	社会科学・人文書出版社	6	4%	40%	○	50%	無
5	社会科学出版社	6	4%	45%		33%	無
5	総合出版社	5	4%	49%		60%	無
8	海外の総合学術出版社	4	3%	51%	○	50%	有(説明有)
8	総合出版社	4	3%	54%	○	0%	無
10	社会科学・人文書出版社	3	2%	57%		0%	無

表 31 に示すように、科学研究費助成事業の研究成果についても、社会科学を中心とする出版社に、一般向けビジネス書も発刊する出版社が並んでいる。経営学においても、上位10社の出版社は重複が多い。表 32 から、英語は13%であり、明確な増減傾向は見られない。

表 31 科学研究費助成事業の経営学における研究成果の書籍出版社

	出版者名	研究成果数	割合	累積割合
1	中央経済社	293	16%	16%
2	有斐閣	181	10%	26%
3	白桃書房	111	6%	32%
4	同文館出版	67	4%	35%
5	ミネルヴァ書房	59	3%	39%
6	日本経済新聞出版社	49	3%	41%
7	Springer	40	2%	44%
8	東洋経済新報社	39	2%	46%
9	税務経理協会	33	2%	47%
10	朝倉書店	33	2%	49%

表 32 科学研究費助成事業の経営学における書籍の外国語割合

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	合計
図書の数	7	3	18	14	21	19	29	22	41	23	30	22	249
うち外国語の 割合	17%	3%	13%	8%	15%	11%	13%	15%	19%	12%	17%	14%	13%

一方、表 33 に示すように、英国では、国際的に有名な総合学術出版社と大学出版会が並んでおり、

上位の 5 社だけで半数以上を占めており、集中がさらに強い。いずれもウェブサイトにおいて査読に関する情報が掲載されている。

表 33 英国 REF の Management 分野における提出研究業績の書籍出版社

	出版社名	研究成果数	割合	累積割合	ウェブサイトでの査読情報の有無
1	Routledge	48	14%	14%	有(説明有)
2	Oxford University Press	43	12%	26%	有(説明有)
3	Palgrave	38	11%	37%	有(説明有)
4	Edward Elgar	26	8%	45%	有(説明有)
5	Cambridge university press	23	7%	52%	有(説明有)
6	Emerald	18	5%	57%	有(説明有)
7	Ashgate	7	2%	59%	有(説明有)
8	Springer	7	2%	61%	有(説明有)
9	Sage	6	2%	63%	有
10	Wiley	6	2%	64%	有(説明有)

以上のことをまとめると、日本の場合は大学評価において上位に上がっている出版社は、学術書以外を出版する総合出版社も含まれている。また、学術書が中心の出版社であっても、海外の大手学術出版社が書籍についてもピアレビューを行うことをウェブサイトで明確に示しているのに対して、日本ではそのような情報は全く確認できない。仮に単純に海外での学術出版社リストの収録基準をあてはめれば、いずれも適合しないことになる。

しかし、このことから日本の研究成果に問題があるとは言えない。それは日本の出版界において、そもそも学術出版社と一般の出版社という区分が明確でないことが、この結果に影響していると考えられるためである。名古屋大学出版会の橘(2016)によれば、日本では、専門書とそれ以外の書籍のマーケットがはっきり区分されていない。加えて、書籍の査読に関しても、日本では、たとえ大学出版部でも専門家による査読を厳密に制度化しているところは少なく、アメリカのように徹底して行っているところはほとんど無いという。他方で、このことは日本では「学術書」が必ずしも学者のみを対象に限定しない幅広いものとなっていることを示している(佐藤ほか 2011)。専門家による査読は視野が専門領域に閉じる傾向があるため、マーケットが区分されていない状況では、査読を厳密に行うことは一般読者を含む出版としては不適合になる。実際には日本では、査読という形式でなく、編集者が読んで原稿を改善させる作業を行っているという(橘 2016)。

書籍のこのような傾向は実際には日本だけでない。Hicks(2005)も人文・社会科学の研究者は、学術的な読者だけでなく、一般の人々のためにも書いており、欧州各国の人文・社会科学のリサーチカウンシル等の連携組織である Humanities in the European Research Area (HERA) による調査でも(Dolan, 2008)、人文学の書籍は「幅広いオーディエンス(wider audience)」、すなわち、一般読者をターゲットの一つとして設定しているとしており、また、学術出版社以外でも出版する伝統があるとしている。オランダでは学術的な研究者だけでなく、一般読者もあわせて対象とする書籍を「ハイブリッド出版」と称している(Honingh et al., 2019; Williams et al., 2018)。

以上のように、日本のみならず国際的にも人文・社会科学の学術成果は学術界以外もあわせて対象

としている場合があり、特に日本の学術書籍出版市場がその前提のもとで構築されているとすれば、日本の人文・社会科学を対象に、学術出版社の登録簿や学術的な格付けを行うことは適合しないと考えられる。

---

#### 4.4 ピアレビューにおける代替指標

##### 4.4.1 代替指標の種類

上記のように、研究成果の形態が多様であり、研究成果を出版する出版社も学術出版社に限らない中では、出版チャンネルに基づく質の確保や格付けは難しい。その場合には、ピアレビューで評価者が研究成果の情報を見た上で、総合的な判断を行うことが求められる。

日本の国立大学法人評価では、評価者は一つ一つの研究成果を改めて読んでレビューを行うのではなく、大学側が提出書類において、研究成果が卓越していると考えられる根拠・指標を示す方式となっており、評価者はその説明を読んで判断を行う。しかし、このような根拠・指標を評価者が読むという方式であっても、多様な根拠・指標が定義も測定方法も多様なままに記載されていれば、評価者はそれらを見て解釈することは難しい。ある程度の標準性を有する形で根拠・指標等の情報を大学が記載することが求められる。そのためには、まず、どのような指標が評価者にとっては参考情報となりえるものであり、さらに、それらについて標準的な測定方法を定めることが可能かを検討する必要がある。

以下では、第二期国立大学法人評価における研究業績水準判定を行った際に、評価者が大学からの提出物に記載されていた根拠情報の何を「SS（卓越）」という最高評価を下すときの参考に活用したか分析を行った（第一期評価を対象とした分析結果と方法については林、土屋（2016）を参照）。

表 34 には人文学及び社会科学において、少なくとも一人の評価者が SS と判定した研究業績のうち、少なくともどちらかの分野において 5% 以上において用いられていた指標を示している。ここでは件数が少ないために、歴史学、経営学に限定せずに、人文学、社会科学の集計値で示している。

人文学分野全体でみると、「賞」や、学術誌・専門誌及び新聞・一般紙における「書評・紹介」が指標として多く用いられている。一方、5% 以下であるため表にはないが、「出版社名」を指標としているものは人文学で 2%、社会科学でも 2% であり、高い割合ではない。これは、上述のように出版チャンネルによる質管理が難しく、出版チャンネル名では卓越性の説明にできる場合が多くないことを示している。他方で、「post-publication peer-review」と呼ばれるような、研究成果が発表された後の研究成果に対する研究者（ピア）らによる主観的判断（認知や批評）としての書評や賞を重視していることが現れていると考えられる。

表 34 評価における代替指標の活用状況

	人文学 (N=309)	社会科学 (N=398)
賞	51%	39%
新聞・一般紙での書評・紹介	25%	13%
学術誌・専門書での書評・紹介	24%	12%
著名な学術雑誌への掲載	5%	30%
インパクトファクター	0%	13%
引用	5%	9%
被引用数	0%	5%
著名な叢書の一つ	5%	1%
学会発表、芸術講演	5%	3%
招待講演・基調講演・プレナリー	11%	16%
講習会・研修会、一般講演会	5%	4%
本件の研究費	7%	1%
研究による影響（文化的貢献）	5%	0%
研究による影響（教育への影響）	1%	6%
研究による影響（審議会等委員）	0%	5%
研究による影響（政策への貢献）	1%	6%

#### 4.4.2 代替指標としての書評の活用可能性

では、このような post-publication peer-review は、資金配分等のための情報として、個別的な情報だけでなく、何らかの標準的方法による定量指標化が現実的に可能であるだろうか。

最も指標としての利用頻度が高かった「賞」については、評価者が SS と判断した情報に用いた賞は、著名な学術的な賞もあれば、個別の学会賞、学会の中の若手研究者賞、地域内での文化面の賞など種類は多様であり、また、同じ賞を受賞している者は通常少ないために集計値をとることも難しい。各賞の受賞候補となる母集団の大きさから賞の重要性を検討することも考えられるが、母集団を実際に測定することは難しい。

そのため、次に指標としての頻度が高い「書評・紹介」に焦点をおいた分析を行う。以下では、「書評・紹介」に関して指標として外部から標準的方法で入手できる値と、ピアレビュー結果がどの程度相関するのかを確認する。

上記の第二期国立大学法人評価において史学に提出された研究業績のうちの書籍を対象として、学術雑誌・一般雑誌の記事（論文等含む）のタイトルのデータベースである「国立国会図書館雑誌記事索引データベース」において、タイトルに当該書籍名を含む書評・書籍紹介と考えられる記事を検索してその件数を計測した。また、各研究業績に対する 2 名のピアレビューの評価結果を、SS を 2 点、S を 1 点、S 未満を 0 点とし、その合計をとり（0～4 点）、それを個々の書籍の評点とした。この 2 つの作業により、書評・紹介の件数と評点との関係を分析した。

結果は図 5 である。横軸を評点、縦軸を書評の件数として、頻度（研究成果数）をプロットの大きさにしている。結果からは、雑誌（主に学術雑誌）に 4 回以上書評が掲載された書籍は、2 以上の評点を示すなどの、ある程度の関係はみられる。しかし、書評件数が多い書籍が最も高い評点を獲得しているとは限らない。また、書評が存在しなくとも評点が高い研究業績も多い。つまり、多くの書評が書か

れることは学術的あるいは社会的に優れた研究成果であることを代替的に示す可能性はあるが、その情報のみによって判断ができるほどのカバー率も精度も全くない。

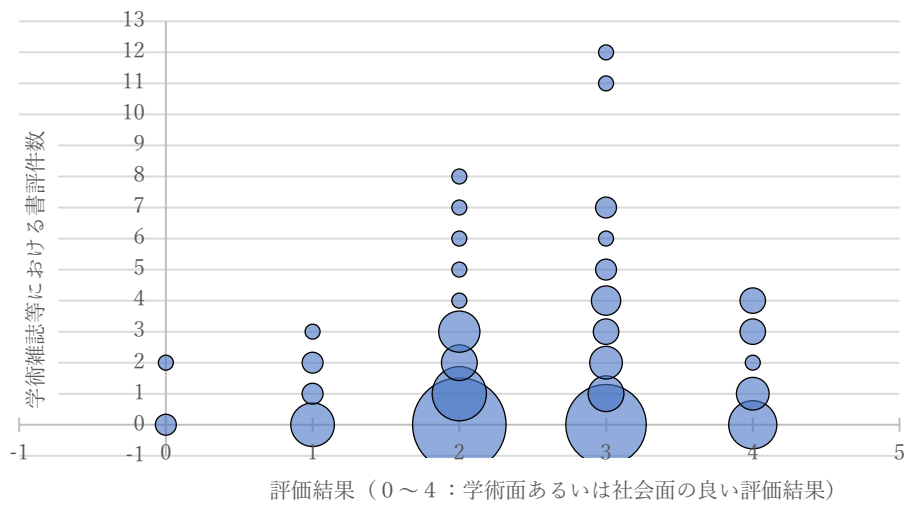


図 5 歴史学における書評数と研究評価結果の関係

経営学の結果は図 6 であり、歴史学と同様の傾向であり、評点が高い研究成果には書評の件数が多いものも見られ、評点が低いものは総じて書評の件数が少ないという結果であるが、明確に区分できるものではない。

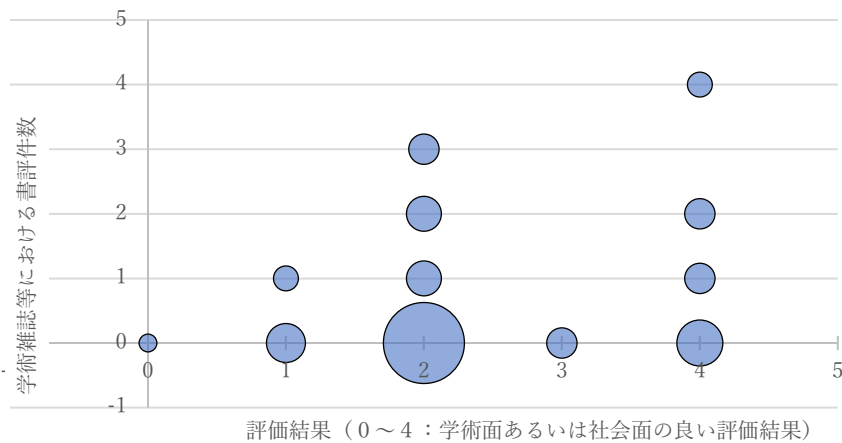


図 6 経営学における書評数と研究評価結果の関係

以上の限られた分析結果からは、評価において当該分野の研究者やレビューアーが用いた指標について統一された方法で測定を行ったとしても、ピアレビューを代替できるような情報を提供することは難しいことが示唆される。一方で、ピアレビューアーが総合的な視点から判断を行う際に、今回のように標準的な測定を行い比較可能性を高めることが、根拠・指標の解釈可能性を高めるかどうかは本分析からは不明であり、更なる検討が求められる。



## 5. 社会的インパクトの測定における多様性と標準化：論点整理

これまでの議論においても、人文・社会科学では、学術コミュニティだけでなく、「幅広いオーディエンス」を対象として研究成果が形成されていることを議論した。表 11 に示したフランス HECRES による歴史学等における研究成果のガイドラインや、表 12 のオランダ QRiH の指標には、教育向けの成果や一般向けの成果、さらにその利用状況に関する指標が含まれていた。一般に、研究成果の学術界以外への効果・影響を把握することは、「社会的インパクト」の測定として、近年、多くの国において研究評価に組み込まれるようになってきている。研究活動への公的投資を、ピアレビューによる学術的価値の説明だけでなく、より直接的な社会への効果から説明することが求められるようになってきているためである (Bornmann and Haunschild, 2019)。

研究の社会的インパクトの評価は、大学等の組織、研究プロジェクト、研究資金制度・プログラムなど、様々な対象の評価において行われるが、大学という組織を対象とする評価では、学術面の評価と同様に、多様な分野の研究活動による多種類の社会的インパクトを、多数の組織を対象に測定することが必要となり、設計が難しい。

大学評価における研究の社会的インパクト評価は、英国の大学研究評価 Research Excellence Framework(REF)では 2014 年から評価項目として位置づけ、インパクトの事例説明 (インパクトケーススタディ) に対する評価という方法で実施しており、国際的にも注目されてきた。一方で、それ以外の国でも取組はなされており、オーストラリアでは 2008 年に大学の研究評価 (Research Quality Framework: RQF と呼ばれた) の検討をはじめ、そこにはインパクトの評価が含まれていた。この RQF 案は REF のインパクト基準の設計でも参照されたが、実際にはオーストラリアでの政権交代により RQF は実現せず、2011 年より始まった Excellent Research in Australia(ERA)では「社会的応用の指標」として特許、登録意匠、研究の商業化による収入、植物育成者権、研究収入のなどの定量指標に基づいて評価がなされている。しかし、2018 年より ERA とは別に Engagement and Impact assessment(EI)が開始され、インパクトのケーススタディ評価が行われている (Williams and Grant, 2018)。他にも、REF を参照して香港の Research Assessment Exercise でもインパクト評価が行われており、また、ポーランドでも導入予定である<sup>48</sup>。

他方で、「インパクト」という語ではなく、オランダでは研究評価のプロトコル Strategy Evaluation Protocol (SEP)において「社会への関連性 Societal Relevance」を一つの評価基準として評価が行われている。また、日本でも前述のように、国立大学法人評価における研究業績水準判定において、社会・経済・文化面の貢献の観点があり、大学が提出した簡単な説明を評価者が読む形での簡素な評価が行われている。

社会的インパクトの測定は、REF で行われているような、事例等の説明文書に対する定性的な評価の方法が徐々に拡大しつつあるが、その方法については、いまだ試行錯誤が続いている。本章では社会的インパクト測定の論点をまとめる。5.1 節で、研究成果の社会的インパクトの測定の論点を整理し、5.2 節で人文・社会科学における社会的インパクト測定について検討する。

<sup>48</sup> Błażej Skoczeń(2020), New solutions in Polish research assessment exercise, AESIS Impact of Science, Poland.

## 5.1 社会的インパクト測定の論点

研究による社会的インパクトの測定には、いくつかの課題がある。以下では、論点を確認する。論点は大きく分ければ、(1) インパクトの定義、(2) 研究とインパクトの関係づけ、(3) インパクトの領域設定、(4) インパクトの根拠・指標の標準化となる。

### 5.1.1 インパクトの定義

#### ① インパクトの段階的発現：エンゲージメントやレlevanceとの関係

評価や測定において、「インパクト」をどのような範囲のものとして定義するかが第一の論点となる。社会的インパクトの定義については、REF2021 では「インパクト」とは「学術界を超えて、経済、社会、文化、公共政策・サービス、保健、環境、生活の質への効果、変化、便益」を指すとされており、詳細には、「地方、国、国際などいずれのかの地域における、聴衆、受益者、コミュニティ、顧客、個人、組織の、活動、態度、認識、行為、能力、機会、実績に対する効果、変化、便益」と説明されている。

ここで具体的に問題となるのは、研究からインパクトが生じる過程は複数の段階があり、どの段階を「インパクト」の対象とするかである。

Bornmann and Haunschild (2019)は、インパクトとアウトプットやアウトカムといった用語を整理した。アウトカム(outcome)を研究による結果の具体的な内容とし、アウトプット(output)は研究結果から生まれる具体的な生産物 (product) であり、出版物や講演などが相当する。インパクト(impact)は、アウトカムに基づいて起こるイベントや生産物や変化である。また、インパクトが幅広い社会変化を起こす場合にはそれを影響(effects)と呼ぶ。

このようにリニアに概念化したとしても二つの問題が生じる。第一の問題は、アウトカムからインパクト (さらには、上記の effects も含め)、どの段階までいけばインパクトと称するかである。Samuel and Derrick (2015) は REF2014 の実施前段階に保健分野 (メインパネル A) の評価者 62 人にインタビューを実施し、インパクトをいかに定義しているかを聞いた。その結果、多くがインパクトとアウトカムを同一視しており、疾病率や経済効果などの「最終」アウトカムでなければインパクトと見なさないと述べる人もいれば、インパクトは段階的であるために、中間的なものも (評価結果が高くない) インパクトと見なす人もいたという。たとえば、ワクチン開発の研究のインパクトの発現は、病原体の特定、ワクチンの臨床での有効性の確認、ワクチンの広範囲の利用、と多段階であり、このように研究成果から最終的なアウトカムが多段階を経る場合に、どこからを「インパクト」と呼ぶかが問題になる。この点は、後述するように、帰属や寄与率の問題とも連動する。多段階を経て発生するインパクトについては、研究成果とインパクトの関係が間接的になるため、どの程度を、特定の研究により生み出されたものであると考えて良いかという問題である。

この点は英国 REF の実際の評価作業においては、4 つのメインパネルおよび 36 の分野別サブパネルごとに、各評価者に対して評価担当分の書類を割り当てる前に、複数のインパクトケーススタディを対象とした評価の議論 (予行演習に相当) を行い、評価基準の解釈の摺り合わせを行っている。また、複数名の評価者による評価結果が大きくずれていなければ平均値を計算するなどの方法によって運用している。

第二の問題は、より前の段階に焦点が遡り、アウトカムが発生する前の取組段階 (プロセス段階) をどれほど積極的にインパクトの評価や測定の対象に入れるかである。結果として実際に現れたインパ

クトのみを評価するという考え方と、インパクトが生じるには長期間を要するため、研究の実施や計画策定の段階で研究者とステークホルダー（研究成果の期待されるユーザー等）との連携関係を構築していることを、インパクト評価にて積極的に評価するという考え方である。このようなプロセス段階を重視する見方は、インパクト評価のフレームワークである「生産的相互作用」(Molas-Gallart and Tang, 2011; Spaapen et al., 2011)や payback フレームワーク(Buxton, 2011; Donovan and Hanney, 2011; Henshall, 2011; Klautzer et al., 2011)などでも重視されてきた。また、インパクトの発現は研究者のコントロールできない範囲において発生することが多いため、インパクトが起こりうる条件を研究者が形成できているか否かを見ることが重要であり、連携はその重要な要素と考える見方もある(Reale et al., 2018)。

この点では、オーストラリアで実施されている Engagement and Impact Assessment (EI) では、連携の評価を「エンゲージメント engagement」の項目として、「インパクト impact」とは別の項目で積極的に評価している(Williams and Grant, 2018)。オーストラリアでは、これらの用語を以下のように定義している。

研究の「エンゲージメント」(engagement) とは、「研究者と学術界以外の研究のエンドユーザーとの間で、知識、技術、方法、資源を相互に有益な形で移転するための相互作用」を指す。

「研究のエンドユーザー」とは、「研究の結果(output)、成果(outcome)、効果(result)を直接使用する、あるいはそれらから直接利益を得る、学術界以外の個人、コミュニティ、組織を指す。たとえば、政府、企業、非政府組織、コミュニティ、コミュニティの組織などである。

研究の「インパクト(impact)」とは、「学術研究への貢献を超えて、研究が経済、社会、環境、文化にもたらす貢献」を指す。

この論点はさらに、インパクトを生む環境（組織戦略や体制）を測定や評価の対象にするか否かに拡大する。オーストラリアの EI では「エンゲージメント」や「インパクト」の評価項目とは別に、さらに「インパクトへのアプローチ」という評価項目を設定している。組織がインパクトの実現をいかに促進したかという点から、研究者への組織的支援や、人事方針、戦略などが確認される。同様に英国の REF でも REF2014 では「インパクト」の評価項目の中でインパクトを実現するアプローチを確認し、REF2021 では「環境」の評価項目の中でそれを確認している。

### 5.1.2 研究成果とインパクトの関係付け

次の論点は、評価や測定において、特定の研究成果とインパクトの関係をどのように整理するかである。

#### ①インパクトの基になった研究の実施時期（インパクトが発現するまでの期間）

資金配分を測定や評価の結果によって行う場合には、通常、評価や測定の対象となる期間が設定されており、インパクトがその対象期間に含まれていることが必要となる。しかし、その基になった研究成果はどの程度古いものが許容されるべきであるかが論点となる。

研究成果がインパクトを生むまでには数十年かかることがある。そのため、REF2021 では、対象となるインパクトの発生期間は 2013 年 8 月 1 日～2020 年 12 月 31 日であるが、基になった研究は 2000 年 1 月 1 日～2020 年 12 月 31 日に実施されたものでよく、最長で 21 年のタイムラグが見込まれてい

る。また、インパクトは発生したばかりの段階でも、2013年以前から存在したインパクトが継続して成熟段階にあるものでも良いとしている。

実際に REF2014 において提出された各インパクトケーススタディについて、インパクトの基になった研究成果 20,779 件の公表年（出版年等）について分析した結果によれば(Digital Science, 2016)、分野によって研究成果の古さは異なっている。人文・社会科学では、ピークは 2007～2009 年に公表された研究成果である。一方で、自然科学や生命科学では記載が可能な最も古い公表年である 1993 年から 2008 年までがほぼ同等の割合となっている。人文・社会科学のほうが比較的最近の研究成果によってインパクトを説明する傾向があり、自然科学のほうが数十年を経てインパクトが発現する事例が多数示されている。

## ②特定の研究と特定のインパクトの関係：帰属

インパクトが多段階である場合に、研究とインパクトの関係は間接的になるため、どこまでを「特定の研究により生まれたインパクト」と認めるかが論点となる。

REF2021 では、提出対象となるインパクトは、評価対象組織(UoA)が行った卓越した研究（研究成果の質評価で星 2 つ以上の質に相当）に基づくものでなければならないとしている。それは、その研究がなければインパクトが発生しなかった、あるいはインパクトが大幅に減少していたものを対象としている。また、研究とインパクトの関係は、多様なものを受け入れるとしている。研究が直接にインパクトを及ぼしているものもあれば、研究が別の主体による研究を誘引し、その結果がインパクトを及ぼすような間接的なものも対象としている。また、複数の個人やグループによる複数の研究が長期間行われてインパクトが生まれた場合も、評価対象の研究が貢献している限りは対象とされる。パブリックエンゲージメントにより、研究が市民の認識、態度、理解、行為へインパクトを及ぼしているものも対象とされる。研究者が審議会委員等に任命されて、卓越した研究に基づく専門的助言の提供を通じてインパクトを及ぼしている場合も対象とされる。REF では比較的広く関係性を認めているが、別の国で新たに評価・測定を設計する際には論点となる。

### 5.1.3 インパクトの領域設定

インパクトの評価や測定においては、社会、経済、環境、文化など、インパクトの領域をどのように設定するかが別の論点である。インパクトに関しても多様な種類があるため、多様なケースが提出されるように促すことが必要となる。一方で、多様なケースについて全体的な集計を行おうとすると、領域が何らかの形で標準的に分類されていることが望ましい。

英国 REF2021 では表 35 のようにインパクト領域の例示をし、多様なインパクトが提出されることを促している<sup>49</sup>。

---

<sup>49</sup> REF2021 Panel criteria and working methods, REF2019/02, January.  
2019<https://www.ref.ac.uk/publications/panel-criteria-and-working-methods-201902/>

表 35 英国 REF2021 の要項において示されたインパクト領域の例

インパクトの領域の例	
人々の健康と福祉、および動物福祉へのインパクト	個人や公衆衛生活動に関する政策や実践の改良を通じて、健康状態が改善された、生活の質が向上された（または潜在的な害が軽減された）、または権利や利益が保護・擁護された個人やグループ（人間と動物の両方）が受益者である場合のインパクト
創造性、文化、社会へのインパクト	行為、創造的取組、権利、義務、その他の活動が影響を受けた個人、個人のグループ、組織、またはコミュニティが受益者に含まれる場合のインパクト
社会福祉へのインパクト	権利、義務、行為、機会、包摂、生活の質、その他の活動が影響を受けた個人、個人のグループ、組織またはコミュニティが受益者に含まれる場合のインパクト
商業や経済へのインパクト	新規・既存の事業、国民保険サービス、民間の健康管理・社会的介護、農業、その他の種類の富を生み出す可能性のある活動を行う組織が受益者に含まれる場合のインパクト
公共政策、法律、サービスへのインパクト	政策、システム、改革の実施あるいは非実施を通じて影響をうける、政府、非政府組織（NGO）、慈善団体、公共部門の組織、社会（社会全体または個人のグループ）が受益者である場合のインパクト
生産へのインパクト	生産が促進された個人（個人のグループを含む）が受益者である場合のインパクト
実務家と専門的サービスの提供、実績の向上、倫理的な実践へのインパクト	専門的サービスや倫理的実践の開発や提供に関与する、サービスユーザーを含む組織や個人が受益者に含まれる場合のインパクト
環境へのインパクト	自然環境、歴史環境、建造環境が主たる受益者である場合のインパクトや、その結果として利益を得る社会、個人、個人のグループが主たる受益者である場合のインパクト
理解、学習、参加へのインパクト	研究の結果として、認知、理解、参加、関与が強化された個人、コミュニティ、組織が受益者である場合のインパクト

（出典：REF2021 Panel criteria and working methods）

インパクトの領域やその内容が何らかの形で分類されれば、その結果を分析することが可能になる。King's College London and Digital Science (2015)では REF2014 のインパクトケーススタディのテキスト分析を行い、そこから抽出された 60 の研究インパクトのトピックと、36 の研究分野（UOA）の関係を視覚化し、複雑な関係があることを示した。

同様に、図 7 には、日本の第二期国立大学評価の研究業績水準評価について、卓越した研究と評価された研究業績の研究者に事後的にアンケート調査を行い（N=316）、回答から研究の学問分野と、インパクトが発現した領域との対応関係について示した。ここでの領域は、旧大学評価・学位授与機構が「試行的評価」として 2001～2003 年に行った研究評価で設定されていた領域をもとに修正して設定したものである。この結果からは、領域と学問分野の関係は多対多の関係にあることが分かる。このようにインパクト領域の設定をせずに学問分野別の集計だけでは、各学問分野がどのような効果の広がりを持っているのか、明らかにすることは難しい。

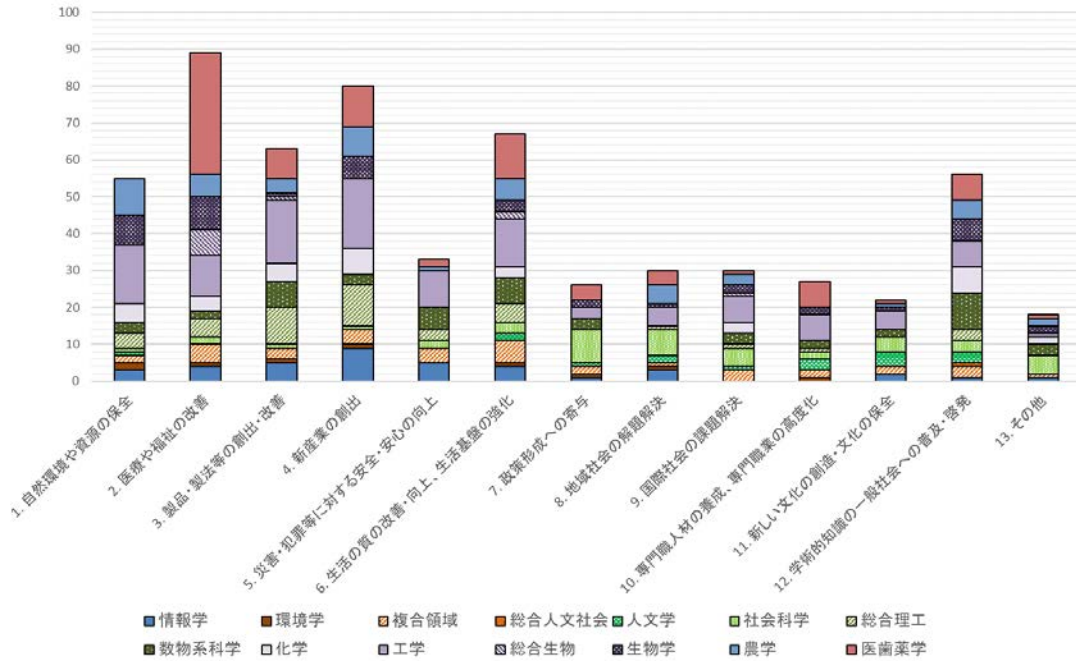


図 7 学問分野とインパクト領域の関係(N=316)

#### 5.1.4 インパクトの根拠・指標の標準化

最後に、インパクトの定性的な説明に記載する根拠や指標などの情報の標準化の問題がある。社会的インパクト評価は、英国、オーストラリアなどでは、インパクトのケースの説明を行い、それを評価者が評価する方法がとられている。オランダの評価でも、文章による記述によって社会的な関連性（レバンス）を説明する。

説明の中にはインパクトの存在や重要性を証明するような根拠・指標を含めることが望まれる。上で既に見たように、学術的な研究成果と同様、インパクトも多様な領域に多様な種類のものがあり、その根拠・指標も多様となり得る。一方で、定義や記載ぶりが多様であれば、評価者はそれを見ても価値判断が難しいために、記載のある程度の標準化が求められる。

##### ①根拠・指標の種類

インパクトの根拠・指標の例示は英国、オーストラリアで行われている。英国では REF2014 における評価基準書にて 4 つのメインパネルごとに、インパクトの例やそこでの指標の例示が行われた。REF2021 でも、表 35 に示した領域ごとに、インパクトの種類や指標を示している。表 36 にそのうちで、人文・社会科学に関連が強い一部の領域について抜粋したものを示している。

表 36 英国 REF2021 におけるインパクトの種類や指標の例示（一部抜粋）

インパクトの領域	インパクトの種類	到達範囲と重要性の指標
<p><b>創造性、文化、社会へのインパクト</b></p> <p>行為、創造的取組、権利、義務、その他の活動が影響を受けた個人、個人のグループ、組織、またはコミュニティが受益者に含まれる場合のインパクト</p>	<ul style="list-style-type: none"> <li>● 美術館の専門家との協働により、美術館やギャラリーの展示など、(文化的) 遺産の保存と解釈の強化。</li> <li>● 映画、小説、テレビ番組などの、新しい文化的作品の共同制作。</li> <li>● 創造的な実践、芸術的質、聴衆の範囲に影響を与える新しい考え方の形成。</li> <li>● 新しい形の芸術的、文学的、言語的、社会的、経済的、宗教的、その他の表現の刺激、共創、支援。</li> <li>● 新しい形の芸術的表現を生み出すような、パブリックアートの場、アーティスト、プログラミングの専門家との協力。</li> <li>● 阻害・周縁化された聴衆、参加度合いが低い聴衆、多様な聴衆の文化的な参加につながるような研究主導のエンゲージメント。</li> <li>● 文化観光への刺激の開発や、観光体験の質への貢献。</li> <li>● 知的財産権を保全するための法的やその他の枠組みの改善。</li> <li>● 発展途上等の特定の文脈での文化保護を強化するような、地元の伝統に対する理解の深まり。</li> <li>● 生活の質の向上をもたらす新しい形の芸術的表現。</li> </ul>	<p>イングランド芸術評議会 (Arts Council England) は、インパクトを評価するためのガイドダンスとツールキットを提供。</p> <ul style="list-style-type: none"> <li>● <a href="https://www.artscouncil.org.uk/quality-metrics/quality-principles">https://www.artscouncil.org.uk/quality-metrics/quality-principles</a></li> <li>● <a href="https://www.artscouncil.org.uk/quality-metrics/quality-metrics">https://www.artscouncil.org.uk/quality-metrics/quality-metrics</a></li> <li>● <a href="https://www.artscouncil.org.uk/measuring-outcomes/generic-learning-outcomes">https://www.artscouncil.org.uk/measuring-outcomes/generic-learning-outcomes</a></li> <li>● <a href="https://www.artscouncil.org.uk/measuring-outcomes/generic-social-outcomes">https://www.artscouncil.org.uk/measuring-outcomes/generic-social-outcomes</a></li> </ul> <p>● 創造的な実践者、キュレーター、メディアの専門家からの証言。</p> <p>● 英国と海外の両方での出版物と売上高の値、聴衆や出席者の数 (関連する場合は人口統計データを含む)、放送データやその他の形式のメディア、ダウンロード数、データベースと Web サイトのヒット数。</p> <p>● メディアでの批評。</p> <p>● 学術文献以外の批評・紹介における引用。オンライン文書を含む、メディアでの第三者による引用。</p> <p>● 展示会、イベント、パフォーマンスでの聴衆や訪問者の数を含む観光データ。</p> <p>● 展示会、パフォーマンス、またはその他の成果の専門家による評価。</p> <p>● 聴衆・訪問者・参加者のフィードバック (例：調査、インタビュー、フォーカスグループを通じて)。</p>
<p><b>社会福祉へのインパクト</b></p> <p>権利、義務、行為、機会、包摂、生活の質、その他の活動が影響を受けた個人、個人のグループ、組織またはコミュニティが受益者に含まれる場合のインパクト</p>	<ul style="list-style-type: none"> <li>● 社会福祉、平等、社会的包摂の改善。公正やその他の機会 (雇用や教育等) へのアクセスの改善。</li> <li>● 研究への関与により、貧困緩和のための政策や実践を強化。</li> <li>● 市民社会グループとの関わりを通じて、社会的、経済的、政治的、法的変化のためのキャンペーンへ貢献。</li> <li>● 研究による情報を得た、社会政策の変更。</li> <li>● 社会政策の変更により、社会福祉、平等、社会的包摂の改善。</li> <li>● 研究により、コミュニティの再生または発展に貢献。</li> <li>● 発展途上等など、特定状況で周縁化されたグループの社会的および教育的包摂の改善。</li> <li>● 難民の受入地域への効果的な統合。</li> <li>● 紛争後状態での和解プロセスにおける被害者のニーズの理解の向上。</li> </ul>	<p>社会的投資収益率 (SROI) を評価するための初心者向けガイドは以下： <a href="http://www.socialvalueuk.org/resource/guidance-on-starting-out-on-sroi-2/">http://www.socialvalueuk.org/resource/guidance-on-starting-out-on-sroi-2/</a>.</p> <ul style="list-style-type: none"> <li>● 社会政策が変更されたことの証拠文書。</li> <li>● 社会的平等、福祉、包摂の改善の指標。</li> <li>● キャンペーン資料の引用 (例：リーフレット)。</li> <li>● 研究によって影響を受けている、メディアや他のフォーラムでの公開討論の証拠。</li> <li>● 社会的包摂の向上の証拠文書 (例：参加者数)。</li> <li>● 市民社会グループおよび政策立案者からの証言。</li> </ul>
<p><b>実務家や専門サービスの提供、性能の向上や倫理的な実践へのインパクト</b></p>	<ul style="list-style-type: none"> <li>● 専門職の基準、ガイドライン、訓練が、研究の影響を受けている。</li> <li>● 専門職の方法、アイデア、倫理が、研究の影響を受けている。</li> </ul>	<ul style="list-style-type: none"> <li>● 専門職の基準、実績、行為の変更の文書。</li> <li>● 優良事例の採用の証拠 (例：教育者や法執行官による)。</li> </ul>

<p>クト</p> <p>専門的サービスや倫理的実践の開発や提供に関与する、サービスユーザーを含む組織や個人が受益者に含まれる場合のインパクト</p>	<ul style="list-style-type: none"> <li>● 研究の結果として、専門家や組織が、変化する文化的価値観に適応できている。</li> <li>● 継続的な人材開発・専門能力開発への貢献。</li> <li>● 実務家・専門家・弁護士が業務を行う際に研究結果を使用した。</li> <li>● 専門職団体や学会が、研究を使用して、優良事例を定義し、政策を策定し、政府やその他の利害関係者に働きかけをしている。</li> <li>● 労働力計画が、研究の影響を受けている。</li> <li>● 初中等教育および高等教育での教育の実践と方法が変化した。</li> <li>● 訓練やコンサルティングの提供を通じて、個人、企業、その他の組織によって、慣行が変更されたか、新しいまたは改善されたプロセスや方法が採用された。</li> <li>● 医学、人事、会計、金融サービスなどの分野のエキスパートシステムの開発が、研究の影響を受けてきた。</li> <li>● 専門的サービスの質、効率、生産性が向上した。</li> <li>● 専門職・法務・法医学の方法が、研究により情報を得てきた。</li> <li>● 法の執行や保安慣行が変更された。</li> <li>● 研究によって効果がないことが示された慣行の中止。</li> </ul>	<ul style="list-style-type: none"> <li>● 新規や変更された専門職の基準や行動規範。</li> <li>● 新規や変更された技術標準またはプロトコル。</li> <li>● 訓練の恩恵を受けた個人の知識、能力、行動の変更の文書。</li> <li>● 態度や行為の発展につながる、専門職間での議論の証拠。</li> <li>● 研究結果とそれらが実際にどのように適用されたかの情報を含む、専門職や顧問からの文献やウェブの情報。</li> <li>● 国内・国際的な業界標準や信頼できるガイドランスに研究が含まれていることを追跡可能な参照。</li> <li>● 専門職による研究の利用やその影響を説明しうるような、研究論文等への追跡可能な参照。</li> </ul>
---	--	--

(出典：REF2021 Panel criteria and working methods より一部抜粋して和訳)

一方、オーストラリアのEIでは、エンゲージメントの指標を示している。標準の指標としては、研究のエンドユーザーからの資金、FTEあたりのHERDC総収入、すべてのHERDCカテゴリー1（オーストラリア政府の競争的資金助成）に対するカテゴリー1特定部分（エンドユーザーとの協力に焦点をおいている助成金）の収入額や件数の割合、研究の商業化による収入を設定している。さらに、これ以外にエンゲージメントの説明文書に追加指標を記入可能としておりその例示がなされている（表 37）<sup>50</sup>。

<sup>50</sup> Australian Research Council(2017), EI 2018 Submission Guidelines  
<https://www.arc.gov.au/engagement-and-impact-assessment/ei-key-documents>



表 37 オーストラリア EI におけるエンゲージメントの説明文書中に記載する追加指標例

<ul style="list-style-type: none"> <li>・研究者の流動性（例：学術界以外で雇用されたり就職した研究者、学術界で雇用されたり就職したエンドユーザー）</li> <li>・特許登録、PCT 出願、日米欧全てに登録された特許</li> <li>・特許における研究成果への引用</li> <li>・インターンシップや就職斡旋に参加した HDR 学生（研究志向の大学院生）</li> <li>・HDR 学生の就職先</li> <li>・産業界との事業継続</li> <li>・スタートアップ企業、スピンアウト企業</li> <li>・研究エンドユーザーとの共著による研究成果</li> <li>・研究エンドユーザーとの共同資金提供による研究成果</li> <li>・研究エンドユーザーからの現物支援</li> <li>・ライセンス契約</li> <li>・機密保持契約</li> <li>・研究、コンサルティング、専門家証言、試験の契約数</li> <li>・ライセンス、譲渡、およびオプションの数</li> <li>・研究ユーザーとの間の形成されたネットワークや関係</li> <li>・しきい値以上の数の契約を持つクライアントの数</li> <li>・重要な組織的連携のエビデンス（例：シドニーヘルスパートナーズ、国際的な研究コンソーシアム、OECD、世界銀行、世界保健機関、国連、ユネスコ）</li> <li>・研究支援と現物支援を伴う慈善団体との関係</li> <li>・本の出版数</li> <li>・外部の諮問委員会への参加</li> <li>・コミュニティグループ、専門家・実務家組織、政府機関へのコンサルテーション</li> <li>・政府へのコンサルテーション、助言</li> </ul>	<ul style="list-style-type: none"> <li>・裁判における専門的証言</li> <li>・産業界の研究に関連する問題に係る公的調査への貢献・資料提出</li> <li>・公開講座、セミナー、一般公開日、学校訪問</li> <li>・専門家コミュニティへのプレゼンテーション</li> <li>・文化施設、セミナー/ワークショップ、インターンシップ、一般市民へのエンゲージメントへのつながり</li> <li>・文化的なイベント/機関の支援（例：文学祭、映画祭、ビビッドシドニー）</li> <li>・公演や展示会の共同設計とコラボレーション</li> <li>・外部の研究パートナーのメンタリング</li> <li>・ユーザー利害関係者や参加者グループとの協力を含め、研究のすべての段階でのユーザーの巻き込み</li> <li>・MOU、協定</li> <li>・公演やさまざまな種類の公的活動に関して収集されたデータ。公的活動において、研究者が自身の研究をコミュニティに報告したり、コミュニティ組織の活動を促進するために研究能力を利用する。</li> <li>・イベント参加統計（公開講座、文化イベント、展示会など）</li> <li>・アウトリーチ活動（公開講座、政策への関与、メディアへの関与、コミュニティイベント）</li> <li>・展示会や新たな作品のメディア報道</li> <li>・ギャラリー、アンサンブル、グループ、その他の専門的な実践団体の設立など、芸術における起業家的活動を把握するための。既存のスピンオフ等の商業化指標の変更</li> <li>・ソーシャルメディア活動を把握するための指標</li> <li>・その他の指標</li> </ul>
---	---

（出典：Australian Research Council(2017), EI 2018 Submission Guidelines より和訳）

## ②記述方式の標準化

一方で、多様な指標の例示とは別に、分野横断的に、根拠となる指標の種類や記述方式の標準化の検討も進められている。

Vertigo Ventures and Digital Science (2016)は REF2014 のケーススタディに記載された、「インパクトを実証しうる根拠」を分析している。多様な根拠が提出されているが、たとえば人文・社会科学では第三者による証言 (testimonials)、表彰 (awards)、メディアの順に多い。一方で各評価パネルの委員長に対してインタビューをした結果、ケーススタディやエビデンスの記述の質が雑多であることや、ウェブサイトが根拠とする場合に容易にアクセスできない場合があったことが指摘されている。そのため、インパクトのエビデンスの種類や、その記述のガイダンスを示している（表 38）。

表 38 REF2014 のインパクト指標のガイド (Vertigo Ventures and Digital Science, 2016)

エビデンスの種類	ガイダンス	関連するエビデンス
賞	<ul style="list-style-type: none"> <li>・いかに研究がステークホルダーに認識されているかを示すような審査員や専門家のコメントを含める。</li> </ul>	<ul style="list-style-type: none"> <li>・賞の授与団体からのコメント</li> <li>・審査委員からの証言</li> </ul>
商業的影響	<ul style="list-style-type: none"> <li>・これらは独立で、時間枠内のものとする。研究がいかに価値の変化を生んでいるかを明確にする。</li> <li>・スピニアウトに関する証拠は、その企業が登記されているだけでなく、売上高、資金調達、従業員数などから商業的な活動が行われていることを示す。</li> </ul>	<ul style="list-style-type: none"> <li>・知財に関する記録。特に研究が事業を創造した記録。</li> <li>・会社の創業者からの証言</li> <li>・収益や投資額の変化</li> <li>・仕事量の変化</li> <li>・投資収益</li> <li>・実施された研究プロジェクトとその結果に関するレポート</li> </ul>
新たなスキルの修得	<ul style="list-style-type: none"> <li>・新しい資格を取得した人数により、専門職の能力への影響を示すことができる。</li> </ul>	<ul style="list-style-type: none"> <li>・証言</li> <li>・訓練を受けなかった対照群との比較</li> </ul>
知財：特許、ライセンス	<ul style="list-style-type: none"> <li>・特許は、研究が革新性を有していることの証拠として使用できる。ライセンスや売上は、商業的能力を示す。</li> </ul>	<ul style="list-style-type: none"> <li>・研究の活用を支援するための資金獲得の証拠</li> <li>・売上高の増加</li> </ul>
法的	<ul style="list-style-type: none"> <li>・研究成果や研究者の助言による、法律の変化。</li> </ul>	<ul style="list-style-type: none"> <li>・法律の変化に伴うインパクトについての、ステークホルダーのレポート</li> <li>・研究が変化に結びついたことの証言</li> </ul>
メディア	<ul style="list-style-type: none"> <li>・最良の例は、研究とそれによる変化について、言及していたり、直接的に関係づけているもの。</li> <li>・メディア報道は、研究が一般社会での議論にいか情報提供したかを示すことができる。</li> <li>・受益者に焦点を置くこと。</li> </ul>	<ul style="list-style-type: none"> <li>・例えば技術購入額の増加の市場データなどの、定量的なレポート</li> </ul>
公共政策	<ul style="list-style-type: none"> <li>・研究者の貢献を具体的に言及した政策文書。あるいは、研究者の助言の結果として起きた変化を示す一連の政策文書。</li> </ul>	<ul style="list-style-type: none"> <li>・法律の変更などのインパクトに関するレポート</li> <li>・研究と変化の関係についての証言</li> <li>・陳情のデータ</li> <li>・政治運動のデータ</li> </ul>
実務的なガイドライン	<ul style="list-style-type: none"> <li>・ガイドライン作成に研究が情報提供したことを示す定性的な説明をつける。</li> </ul>	<ul style="list-style-type: none"> <li>・ガイドラインが使用されていることを示すデータ</li> </ul>
組織が発行するレポート 例：企業レポート、統計レポート	<ul style="list-style-type: none"> <li>・独立した組織によるレポートであり、研究について、研究が利害関係者にどのように影響したかについて直接言及する。</li> <li>・レポートに聴衆の数が記載されている場合は、聴衆に及ぼした変化を示す証拠もあるのが有用。</li> </ul>	<ul style="list-style-type: none"> <li>・研究と報告されたインパクトの関係を示す証言が必要。</li> </ul>
ソーシャルメディア	<ul style="list-style-type: none"> <li>・普及、フォロワー、印象、シェアなどの統計は、特定のトピックが特定の聴衆に関与したことを示しうる。</li> </ul>	<ul style="list-style-type: none"> <li>・例えば技術購入額の増加の市場データなどの、定量的なレポート</li> </ul>
証言	<ul style="list-style-type: none"> <li>・独立の、広く尊敬されている人が、研究活動および、それが彼らにいか影響したかを直接言及したもの。可能であれば、インパクトの定量的な例を含めること。</li> </ul>	<ul style="list-style-type: none"> <li>・生じた変化を示す定量的レポート</li> </ul>
Web リンク	<ul style="list-style-type: none"> <li>・最良の例は、独立したものであり、ウェブサイトへの訪問者数など到達範囲が分かるメタデータがあるもの。</li> </ul>	<ul style="list-style-type: none"> <li>・閲覧数のメタデータや、可能であれば、購入データや特定の行為に結びついたことを示すデータ</li> </ul>

同様に、RAND は REF2014 のインパクト評価において大学より提出された書類の中で、定量的デ

ータを含むものを選択し、テキストマイニングを実施することで、「スタイルガイド（書き方）」を提案している(Parks et al., 2018)。ここでは、定量データの標記揺れの削減のため、推奨される記載様式を提案するとともに、エンゲージメント、メディアメンション、雇用、資金、有害物質の除去の5つの領域についての記載様式例をガイドラインとして示している。このように定性的なインパクト評価における標準化の取組が進められているのが現状である。

---

## 5.2 人文・社会科学におけるインパクト測定の論点

自然科学に限らず人文・社会科学についても、研究活動への公的投資が学術面以外にどのような価値をもたらしているのかの実証が求められるようになってきている。前章で議論したように、人文・社会科学は研究成果を、一般の幅広いオーディエンス向けにも作成する傾向が高い。そもそも人文・社会科学は、人間やその文化・思想と、人間集団としての社会の特性について研究を行うものであり、そこで得られる研究成果は人間・文化や社会の自己理解を支え、顕在あるいは潜在的な課題の解明とその解決を支援するものである。そのため、人文・社会科学がいかなる社会的・経済的・文化的なインパクトを与えているか測定することは重要となる。実際に、日本の国立大学法人評価の「研究業績水準判定」においても、人文・社会科学分野にて、文化面や政策面での貢献が高い事例が把握されている(林 2018)。

しかし、インパクト評価においても、人文・社会科学は自然科学とは異なる特性を有している。Reale et al. (2018)は、人文・社会科学は、自然科学のように新たな発見や理論を重視するよりは、新たなアプローチに独創性を見いだしていることを踏まえ、人文・社会科学は社会のために研究を発展させることを重視しているにもかかわらず、研究者の関心は発見や開発によって「利用可能な」結果を生み出すことではなく、新たなアプローチに基づいて、社会に方向付けを行ったり、文化遺産を維持したり、市民の自己理解の能力を生むことに置かれており、それゆえに、投資に対するリターンという視点から評価できないことを指摘している。また、人文・社会科学の内部での多様性ゆえに、インパクト評価においても標準的な方法を見つけ出すのは難しいとしている。

同様に、de Jong and Muhonen (2020)は欧州 17 カ国からの 60 のインパクトケースを収集し、「生産的相互作用」モデルを前提に分析を行った。結果、人文・社会科学では研究のユーザーや研究成果活用の文脈が多様であり、自然科学のように特許、ライセンス収入、スピンオフなどの標準的指標を使用することは困難であると指摘している。人文・社会科学では社会的インパクトは間接的で非線形な経路であるために、定量的な指標では測定しにくい。また、ステークホルダーとしては政府、メディア、一般市民、博物館などの文化セクター、その他の医療セクターなどが典型的であるが、社会における公的組織への信頼が低く、あるいは、市民参画の風土が低い国では、研究者とステークホルダーが新たな連携関係を形成すること自体も難しく、そのような文脈が社会的インパクトの発現しやすさにも影響していると述べている。

そのため、Donovan and Gulbrandsen (2018)は、人文・社会科学では自然科学を中心とする one-size fits all とは異なる質的なインパクト測定が必要であると主張している。それは対象の複雑さを縮減させず、経済的だけでなく社会的・文化的インパクトも含み、研究のインパクトについての長期ビジョンが重要であり、慣行に対する批評や対抗という人文科学が有する役割を踏まえたものとなるべきと指摘している。

現時点では、人文・社会科学のインパクト評価の具体的な方法は確立していない。人文・社会科学のインパクトについてレビューした Pedersen et al. (2020)は、これまで表 39 に示すような方法論が提

唱されてきたと整理しているが、方法論は多様であり、すべてを網羅する指標や普遍的な指標を求めることは難しく、特定のミッションに向けたインパクトの評価ツールを確立することが望まれるとしている。あわせて、実際に用いられている手法を整理した結果では、インタビュー(42.4%)、ケーススタディ(42.0%)、サーベイ調査(39.6%)、ピアや専門家レビュー(36.0%)が中心であり、その次に統計データベース(35.0%)や商業化の統計(29.3%)と続き、人文・社会科学では定性的手法が中心となることを示している。

表 39 人文・社会科学におけるインパクト測定の方法(Pedersen et al., 2020)

大学評価 ・ 英国 REF ・ オランダ standard evaluation protocol プロジェクト評価 ・ IMPACT-EV ・ 医療分野の Payback Framework の医療経済学への応用 ・ The flow of knowledge, expertise, and influence model ・ 生産的相互作用(SIAMPI) ・ Research contribution framework ・ Contribution mapping ・ Linkage and exchange framework (LE) ・ RAPID outcome mapping approach(ROMA)
--

このように人文・社会科学におけるインパクト測定においては、自然科学とは異なる課題が様々な指摘されているが、さらに新たな論点も生まれている。日本では人文・社会科学の「知」と自然科学の「知」の融合による「総合知」により社会変革につながるイノベーションが求められていることに見られるように、「トランスフォーマティブな変化」(Schot and Steinmueller, 2019)の中で人文・社会科学にも総合性が求められるようになってきている。このような展開を Sörlin (2018)は「変容の人文科学」「統合的人文科学」と称しており、環境人文科学のように社会的課題に対応した研究へと人文・社会科学が拡大していると指摘している。その中ではインパクト評価も、社会の中での知識の価値創出、社会変革といった包括的な視点、長期的で体系的なインパクトの測定が今後いっそう求められるようになる。このような測定をいかに行っていかは更なる課題である。

## 6. おわりに

本ワーキングペーパーでは、人文・社会科学を中心にしながら、大学等の組織単位を対象とする研究成果指標の多様性と標準化の両立のあり方を検討した。英国 REF との比較に基づけば、日本では、評価に提出されるような研究成果について、現時点では多様性は比較的に高い状況にある。また、人文・社会科学の主要な研究成果発表メディアとなっている書籍の出版市場は、学術出版と一般出版とが海外の例のように明確に区分されているわけではなく、人文・社会科学の成果は学術界だけでなく一般向けにも形成されている状況にある。このような背景も影響してか、研究成果の多様性を尊重する声はあっても、さほど重要視した議論にまでは展開していない。海外では「責任ある研究評価」として提言が複数出され、DORA にも 2,000 以上の機関（大学や資金配分機関等）が署名している状況からは日本はほど遠い。

しかし、大学への経常的資金配分のために定量的指標が要求されるようになると、このような日本の文化は、比較可能な形での集計を行いづらいと取られかねない。比較可能性を担保するために少数の指標を使うことと、できるだけ多様な研究成果を認識し奨励することの両極端の間で、どのようなバランスのとれた設計を行えばよいのか。海外を参照すれば、そのような状況の中で取り得る選択肢は大きく分けて二種類ある。

一つは、北欧諸国などの非英語圏でみられるように、自国語による出版を含めたデータベースのもとで、学術界が中心となって「学術出版物」（あるいは狭い意味での「学術」を越えるオーディエンスを対象とする出版物や成果物を含めて）の定義について検討し、関係者の理解が十分に得られる定義を定め、あわせて国内データベースを測定に耐えられるような信頼性と網羅性のもとに整備し、測定を進めることである。この方法を取れば、「学術出版物」の定義に合致する研究成果に対する研究者の意識付けがなされ、その数が増すことは予想できる。ノルウェーなどの海外諸国では、測定を始めてから、対象とした学術出版物数や高いレベルと定義した学術出版物数は増加しており、自国語での出版割合の低下やジャーナル論文への発表形態の集中化などの悪影響は観察されない一方で、これまで対象の学術出版物を産出してこなかった研究者が産出するようになったことが示されている (Aagaard et al., 2015; Schneider et al., 2016)。さらに、論文データベース WoS（そのうちの社会科学のデータである SSCI、人文学のデータである AHCI）の論文数も同時に増加しており、この指標から定義される意味での、人文・社会科学の「国際競争力」も向上している。一方で、日本については、佐藤（2020）は、SSCI における東アジア諸国の論文数を分析し、韓国、台湾、中国がその論文数を増して国際順位を上げているのに、日本は低下傾向にあることを示している（ただし、SSCI 論文数の増加は、図書等の他成果の停滞と表裏一体であることを述べ、問題意識をあわせて示している）。仮に、日本の人文・社会科学の何らかの「競争力」向上を政策的に課題として考えるのであれば、「学術出版物」（あるいは、より広く測定対象とする研究成果物）を定義し、その定義に合致する研究成果を測定して、産出を支援していくことは一つの方法としては考えられる。ただし、このような方法をとる場合には、その定義に合致しない成果や、定量的測定がしづらい成果を生む研究活動への悪影響が生じていないかなど、モニタリングや測定方法の見直しの仕組みもあわせて導入することは必須である。ここでは「責任ある研究評価」概念を、日本においても切実なものとして議論しなければならない。

もう一つは、ピアレビューを重視して、学術面だけでなく社会的インパクトを含めた評価を行い、指標化することである。その際には、ピアレビューに提出される多様な根拠・指標の定義や記載事項を、ある程度標準化することにより、比較可能な質の高い評価結果を生むことが重要となる。このような

方法は、研究成果の種類の多様性が比較的高く、また、幅広いオーディエンスも対象に含めて研究を行っているという、現在の日本の人文・社会科学の研究文化には適合しやすい可能性がある。ただし、ピアレビューは評価者の主観に強く依拠するという内在的特徴があり、評価結果に対する外部からの信頼が課題になる。そのため、指標の定義や例示を行いつつ、記載方式の標準化なども行い、ピアレビューに用いる根拠・指標の多様性の担保と比較可能性の向上を進めることが望まれる。また、人文・社会科学以外も含めて、研究成果の社会的インパクトの評価や測定が新たな課題になっており、そこではインパクトのケースに対して評価者が評価を行うという方法が徐々に普及しつつある。ここでも「インパクト」の定義や根拠指標例の提示、記載方式の標準化を進めることが今後の課題である。ピアレビューはコストがかかる方法であるが、根拠・指標の定義や記載を標準化していくことで、大学が自己評価を行う際の手間や、評価者が判断をする際の迷いを減らし、効率的な実施の模索も進むことも期待される。

これら二つの方向は背反ではなく組み合わせることも可能である。今後、いかなる方法により多様化と標準化のバランスをとっていくかは、単に測定だけの問題でなく、人文・社会科学の研究の価値をどのようなものとして捉え、どのような成果の産出を促進していくかに依存する。ここまで度々参照してきた欧州の ENRESSH では、人文・社会科学の評価の設計は、人文社会科学のコアにある価値を支援するように、より創造的になることを提言している(Jong et al., 2020)。また、より具体的には、Sigurðarson (2021)は、人文学では特定の政策課題の解決ではなく、「コミュニティの認識論的能力 epistemic capacity」へのインパクトが重要であると指摘し、人文学研究を通じて、特定のコミュニティの自己理解が進み、自信が醸成されるといったエンパワーメントが重要と指摘している。このような議論は、評価や指標の議論の前に、人文・社会科学の研究の基本的な価値は何かという問題が重要であることを指摘する。日本においても各国と同様に、大学、大学協会、アカデミー、大学評価機関などが積極的に検討を行うことが期待される（最近では、東京大学(2020)がその例である）。

さらに現在、第6期科学技術・イノベーション基本計画では社会全体の再設計に向けて、人文・社会科学を含めた「総合知」の活用が求められている。こうした中では研究成果の測定や評価も、学術面や各種の社会面を含む多様な側面から行い、あるいはオープンサイエンスの潮流の中で多様なステークホルダーからの価値判断を情報として用いるような包摂性や、そこで得られた測定や評価結果から研究というもののあり方自体を、前提を置かずに検討するような反射的な議論が重要となる。このような新たな文脈における学問のあり方に関する定性的議論と、意思決定に利用可能な信頼できる情報を定量・定性の指標を精緻化していくことにより形成していくことの両立を図ることが、今後の課題である。また、本ワーキングペーパーは大学等の組織を対象とした資金配分の場面での指標利用を検討したが、研究者個人やプロジェクト評価においては、研究内容の価値に関する質的な視点がより必要となり、そのような検討は別途行う必要がある。

## 参考文献

- Aagaard, K., 2018. Performance-based Research Funding in Denmark: The Adoption and Translation of the Norwegian Model. *J. Data Inf. Sci.* 3, 20–30. <https://doi.org/10.2478/jdis-2018-0018>
- Aagaard, K., Bloch, C., Schneider, J.W., 2015. Impacts of performance-based research funding systems: The case of the Norwegian Publication Indicator. *Res. Eval.* 24, 106–117. <https://doi.org/10.1093/reseval/rvv003>
- Bornmann, L., 2011. Mimicry in science? *Scientometrics* 86, 173–177. <https://doi.org/10.1007/s11192-010-0222-8>
- Bornmann, L., Haunschild, R., 2019. Societal Impact Measurement of Research Papers, in: Glänzel, W., Moed, H.F., Schmoch, U., Thelwall, M. (Eds.), *Springer Handbook of Science and Technology Indicators*. Springer, pp. 609–632. [https://doi.org/10.1007/978-3-030-02511-3\\_23](https://doi.org/10.1007/978-3-030-02511-3_23)
- Butler, L., 2003. Modifying publication practices in response to funding formulas. *Res. Eval.* 12, 39–46. <https://doi.org/10.3152/147154403781776780>
- Buxton, M., 2011. The payback of “Payback”: challenges in assessing research impact. *Res. Eval.* 20, 259–260. <https://doi.org/10.3152/095820211X13118583635837>
- Cicero, T., Malgarini, M., 2020. On the use of journal classification in social sciences and humanities: evidence from an Italian database. *Scientometrics* 125, 1689–1708. <https://doi.org/10.1007/s11192-020-03581-8>
- De Filippo, D., Aleixandre-Benavent, R., Sanz-Casado, E., 2020. Toward a classification of Spanish scholarly journals in social sciences and humanities considering their impact and visibility. *Scientometrics* 125, 1709–1732. <https://doi.org/10.1007/s11192-020-03665-5>
- de Jong, S.P.L., Muhonen, R., 2020. Who benefits from ex ante societal impact evaluation in the European funding arena? A cross-country comparison of societal impact capacity in the social sciences and humanities. *Res. Eval.* 29, 22–33. <https://doi.org/10.1093/reseval/rvy036>
- Digital Science, 2016. Publication patterns in research underpinning impact in REF2014 A report to HEFCE by Digital Science.
- Dolan, C., 2008. Feasibility Study: the Evaluation and Benchmarking of Humanities Research in Europe, Arts and Humanities Research Council, AHRC.
- Donovan, C., Gulbrandsen, M., 2018. Introduction: Measuring the impact of arts and humanities research in Europe. *Res. Eval.* 27, 285–286. <https://doi.org/10.1093/reseval/rvy019>
- Donovan, C., Hanney, S., 2011. The “Payback Framework” explained. *Res. Eval.* 20, 181–183. <https://doi.org/10.3152/095820211X13118583635756>
- Engels, T.C.E.E., Guns, R., 2018. The Flemish Performance-based Research Funding System: A Unique Variant of the Norwegian Model. *J. Data Inf. Sci.* 3, 44–59. <https://doi.org/10.2478/jdis-2018-0020>
- Ferrara, A., Bonaccorsi, A., 2016. How robust is journal rating in Humanities and Social Sciences? Evidence from a large-scale, multi-method exercise. *Res. Eval.* 25, 279–291. <https://doi.org/10.1093/reseval/rvv048>
- Giménez-Toledo, E., Mañana-Rodríguez, J., Engels, T.C.E.E., Guns, R., Kulczycki, E., Ochsner, M.,

- Pölönen, J., Sivertsen, G., Zuccala, A.A., 2019. Taking scholarly books into account, part II: a comparison of 19 European countries in evaluation and funding. *Scientometrics* 118, 233–251. <https://doi.org/10.1007/s11192-018-2956-7>
- Giménez-Toledo, E., Román-Román, A., 2007. From experimentation to coordination in the evaluation of Spanish scientific journals in the humanities and social sciences. *Res. Eval.* 16, 137–148. <https://doi.org/10.3152/095820207X220409>
- Glaser, J., 2007. Chap.12 The social orders of research evaluation systems. *Chang. Gov. Sci.*
- Hammarfelt, B., 2018. Taking Comfort in Points: The Appeal of the Norwegian Model in Sweden. *J. Data Inf. Sci.* 3, 85–95. <https://doi.org/10.2478/jdis-2018-0023>
- Henshall, C., 2011. The impact of Payback research: developing and using evidence in policy. *Res. Eval.* 20, 257–258. <https://doi.org/10.3152/095820211X13118583635873>
- Hicks, D., 2005. The Four Literatures of Social Science, in: Moed, H.F., Glänzel, W., Schmoch, U. (Eds.), *Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems*. Springer Netherlands, Dordrecht, pp. 473–496. [https://doi.org/10.1007/1-4020-2755-9\\_22](https://doi.org/10.1007/1-4020-2755-9_22)
- Honingh, N.V.D.B., Spaapen, J., Prins, A., 2019. Evaluating academic research at the crossroads of academic and societal demands. *RESSH*.
- Hug, S.E., Ochsner, M., Daniel, H.-D., 2013. Criteria for assessing research quality in the humanities: a Delphi study among scholars of English literature, German literature and art history. *Res. Eval.* 22, 369–383. <https://doi.org/10.1093/reseval/rvt008>
- Ingwersen, P., Larsen, B., 2014. Influence of a performance indicator on Danish research production and citation impact 2000–12. *Scientometrics* 101, 1325–1344. <https://doi.org/10.1007/s11192-014-1291-x>
- Jong, S. de, Balaban, C., Holm, J., Spaapen, J., 2020. Redesigning research evaluation practices for the social sciences and humanities: perspectives from the European network for research evaluation in the social sciences and humanities (ENRESSH). *Deeds Days* 73, 17–35. <https://doi.org/10.7220/2335-8769.73.1>
- King's College London, Digital Science, 2015. The nature, scale and beneficiaries of research impact : an initial analysis of Research Excellence Framework (REF) 2014 impact case studies 92.
- Klautzer, L., Hanney, S., Nason, E., Rubin, J., Grant, J., Wooding, S., 2011. Assessing policy and practice impacts of social science research: the application of the Payback Framework to assess the Future of Work programme. *Res. Eval.* 20, 201–209. <https://doi.org/10.3152/095820211X13118583635675>
- Kulczycki, E., Korytkowski, P., 2018. Redesigning the Model of Book Evaluation in the Polish Performance-based Research Funding System. *J. Data Inf. Sci.* 3, 60–72. <https://doi.org/10.2478/jdis-2018-0021>
- Molas-Gallart, J., Tang, P., 2011. Tracing “productive interactions” to identify social impacts: an example from the social sciences. *Res. Eval.* 20, 219–226. <https://doi.org/10.3152/095820211X12941371876706>
- Ochsner, M., Hug, S.E., Daniel, H.D., 2013. Four types of research in the humanities: Setting the stage



- for research quality criteria in the humanities. *Res. Eval.* 22, 79–92.  
<https://doi.org/10.1093/reseval/rvs039>
- Parks, S., Ioppolo, B., Stepanek, M., Gunashekar, S., 2018. Guidance for standardising quantitative indicators of impact within REF case studies, Guidance for standardising quantitative indicators of impact within REF case studies. <https://doi.org/10.7249/rr2463>
- Pedersen, D.B., Grønvad, J.F., Hvidtfeldt, R., 2020. Methods for mapping the impact of social sciences and humanities—A literature review. *Res. Eval.* 29, 4–21. <https://doi.org/10.1093/reseval/rvz033>
- Pölönen, J., 2018. Applications of, and Experiences with, the Norwegian Model in Finland. *J. Data Inf. Sci.* 3, 31–44. <https://doi.org/10.2478/jdis-2018-0019>
- Pölönen, J., Guns, R., Kulczycki, E., Sivertsen, G., Engels, T.C.E., 2021. National Lists of Scholarly Publication Channels: An Overview and Recommendations for Their Construction and Maintenance. *J. Data Inf. Sci.* 6, 50–86. <https://doi.org/10.2478/jdis-2021-0004>
- Pontille, D., Torny, D., 2010. The controversial policies of journal ratings: evaluating social sciences and humanities. *Res. Eval.* 19, 347–360. <https://doi.org/10.3152/095820210X12809191250889>
- Reale, E., Avramov, D., Canhial, K., Donovan, C., Flecha, R., Holm, P., Larkin, C., Lepori, B., Mosoni-Fried, J., Oliver, E., Primeri, E., Puigvert, L., Scharnhorst, A., Schubert, A., Soler, M., Soós, S., Sordé, T., Travis, C., Van Horik, R., 2018. A review of literature on evaluating the scientific, social and political impact of social sciences and humanities research. *Res. Eval.* 27, 298–308. <https://doi.org/10.1093/reseval/rvx025>
- Rijcke, S. de, Wouters, P.F., Rushforth, A.D., Franssen, T.P., Hammarfelt, B., 2016. Evaluation practices and effects of indicator use—a literature review. *Res. Eval.* 25, 161–169. <https://doi.org/10.1093/reseval/rvv038>
- Samuel, G.N., Derrick, G.E., 2015. Societal impact evaluation: Exploring evaluator perceptions of the characterization of impact under the REF2014. *Res. Eval.* 24, 229–241. <https://doi.org/10.1093/reseval/rvv007>
- Schneider, J.W., Aagaard, K., Bloch, C.W., 2016. What happens when national research funding is linked to differentiated publication counts? A comparison of the Australian and Norwegian publication-based funding models. *Res. Eval.* 25, 244–256. <https://doi.org/10.1093/reseval/rvv036>
- Schot, J., Steinmueller, W.E., 2019. Transformative change: What role for science, technology and innovation policy?: An introduction to the 50th Anniversary of the Science Policy Research Unit (SPRU) Special Issue. *Res. Policy* 48, 843–848. <https://doi.org/10.1016/j.respol.2018.12.005>
- Sigurðarson, E.S., 2021. Capacities, capabilities, and the societal impact of the humanities. *Res. Eval.* 29, 71–76. <https://doi.org/10.1093/RESEVAL/RVZ031>
- Sile, L., Pölönen, J., Sivertsen, G., Guns, R., Engels, T.C.E., Arefiev, P., Dušková, M., Faurbæk, L., Holl, A., Kulczycki, E., Macan, B., Nelhans, G., Petr, M., Pisk, M., Soós, S., Stojanovski, J., Stone, A., Šušol, J., Teitelbaum, R., 2018. Comprehensiveness of national bibliographic databases for social sciences and humanities: Findings from a European survey. *Res. Eval.* 27, 310–322. <https://doi.org/10.1093/reseval/rvy016>
- Sivertsen, G., 2018. The Norwegian Model in Norway. *J. Data Inf. Sci.* 3, 3–19.

- <https://doi.org/10.2478/jdis-2018-0017>
- Sivertsen, G., 2016. Publication-Based Funding: The Norwegian Model, in: *Research Assessment in the Humanities*. Springer International Publishing, Cham, pp. 79–90.  
[https://doi.org/10.1007/978-3-319-29016-4\\_7](https://doi.org/10.1007/978-3-319-29016-4_7)
- Sörlin, S., 2018. Humanities of transformation: From crisis and critique towards the emerging integrative humanities. *Res. Eval.* 27, 287–297. <https://doi.org/10.1093/reseval/rvx030>
- Spaapen, J., van Drooge, L., Drooge, L. Van, 2011. Introducing “productive interactions” in social impact assessment. *Res. Eval.* 20, 211–218.  
<https://doi.org/10.3152/095820211X12941371876742>
- UHR, 2004. A bibliometric model for performance-based budgeting of research institutions. Oslo.
- Vanclay, J.K., 2011. An evaluation of the Australian Research Council’s journal ranking. *J. Informetr.* 5, 265–274. <https://doi.org/10.1016/j.joi.2010.12.001>
- Vertigo Ventures, Digital Science, 2016. *Collecting Research Impact Evidence: Best Practice Guidance for the Research Community*.
- Williams, G., Basso, A., Galleron, I., Lippiello, T., 2018. More, Less or Better: The Problem of Evaluating Books in SSH Research, in: Bonaccorsi, A. (Ed.), *The Evaluation of Research in Social Sciences and Humanities*. Springer International Publishing, pp. 133–158.  
[https://doi.org/10.1007/978-3-319-68554-0\\_6](https://doi.org/10.1007/978-3-319-68554-0_6)
- Williams, K., Grant, J., 2018. A comparative review of how the policy and procedures to assess research impact evolved in Australia and the UK. *Res. Eval.* 27, 93–105.  
<https://doi.org/10.1093/reseval/rvx042>
- Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., Jones, R., Kain, R., Kerridge, S., Thelwall, M., Tinkler, J., Viney, I., Wouters, P., Hill, J., Johnson, B., 2015. *The metric tide : report of the Independent Review of the Role of Metrics in Research Assessment and Management*.  
<https://doi.org/10.13140/RG.2.1.4929.1363>
- 浅川和宏(2019)「経営研究の国際標準化時代における質の高い論文の条件：日本からのアプローチ」*組織科学*、52 巻 4 号 pp.4-12.
- 国立大学協会(2019)「第 4 期中期目標期間における国立大学法人の教育・研究に関する客観的指標等の在り方について (論点整理)」
- 佐藤郁哉(2018)『50 年目の「大学解体」 20 年後の大学再生: 高等教育政策をめぐる知の貧困を越えて』京都大学学術出版会.
- 佐藤郁哉、芳賀学、山田真茂留(2011)『本を生みだす力 学術出版の組織アイデンティティ』新曜社.
- 佐藤幸人・編(2020)『東アジアの人文・社会科学における研究評価 制度とその変化』アジア経済研究所
- 橘宗吾(2016)『学術書の編集者』慶應義塾大学出版会.
- 東京大学人文社会科学振興ワーキンググループ(2020)「東京大学における人文社会科学の振興とその展望」
- 日本学術会議 研究にかかわる「評価システム」の在り方検討委員会 (2008)「(対外報告) 我が国における研究評価の現状とその在り方について」  
<http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-20-t51-3.pdf>

日本学術会議 研究にかかわる「評価システム」の在り方検討委員会 (2012)「(提言) 我が国の研究評価システムの在り方 ～研究者を育成・支援する評価システムへの転換～」

<http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-22-t163-1.pdf>

林隆之(2018)「大学評価の現場における人文・社会科学の研究評価の現状」学術の動向, 23 巻, 10 号, p. 16-23.

林隆之(2019)「国立大学運営費交付金の配分と指標・評価」IDE、2019 年 11 月号.

林隆之、齊藤貴浩、水田健輔、米澤彰純、川村真理、安藤二香(2020)『大学評価と運営費交付金配分の一体的改革の在り方』SciREX ワーキングペーパー 2020 年 10 月.

林隆之、土屋俊 (2016)「学問分野による「卓越性」指標の多様性—多様な研究成果への報償の必要」石川真由美『世界大学ランキングと知の序列化: 大学評価と国際競争を問う』京都大学学術出版会.

藤本隆宏(2020)「発信せんとや生まれけむ —ジャーナル点数主義と日本の経営学—」組織科学、53 巻 4 号、pp. 18-28.

文部科学省(2017)「文部科学省における研究及び開発に関する評価指針」

[https://www.mext.go.jp/a\\_menu/kagaku/hyouka/1260346.htm](https://www.mext.go.jp/a_menu/kagaku/hyouka/1260346.htm)

※. 本ワーキングペーパーの著作権は、著者もしくは政策研究大学院大学科学技術イノベーション政策研究センターに帰属しています。本ワーキングペーパーに含まれる情報を、個人利用の範囲を超えて転載、またはコピーを行う場合には、政策研究大学院大学科学技術イノベーション政策研究センターによる事前の承諾が必要となりますので、以下までご連絡ください。

**【連絡先】**政策研究大学院大学科学技術イノベーション政策研究センター

TEL: 03-6439-6329 / E-Mail: [scirex-center@grips.ac.jp](mailto:scirex-center@grips.ac.jp)



# SciREX Center



GRIPS

政策研究大学院大学

NATIONAL GRADUATE INSTITUTE  
FOR POLICY STUDIES

科学技術イノベーション政策研究センター

Science for RE-Designing Science, Technology and Innovation Policy Center (SciREX Center)

〒106-8677 東京都港区六本木 7-22-1 / Tel 03-6439-6329 / Fax 03-6439-6260

7-22-1 Roppongi, Minato-Ku, Tokyo 106-8677 JAPAN

Tel +81-(0)3-6439-6329 / Fax +81-(0)3-6439-6260