

調査報告書

研究評価の国際動向

エグゼクティブサマリー

2010年前後から、欧米の研究者コミュニティでは研究評価における定量的指標の偏重が強く批判されるようになった。一つの研究成果を意図的に分割して複数の論文にする「サラミ出版」のような戦略的な論文投稿や、研究不正へのインセンティブにつながりかねないという懸念が共有されてきたためである。2013年には「研究評価に関するサンフランシスコ宣言（DORA）」が公表され、従来の研究評価システムの問題点を指摘し改革を求める国際的な枠組みを形成してきた。大学、ファンディング機関、出版社など多様な組織や個人が宣言に署名し、賛同を示してきた。2020年前後を境に、研究評価改革の動きは理念的な訴えにとどまらず、具体的な研究評価手法の開発や新たな仕組みの試行・効果測定へとシフトしている。

同時に、近年では大学などの高等教育・研究機関に対し、教育・研究に加え、社会課題・地域課題へ積極的に取り組むことが求められている。その中で、研究の社会的インパクトをどのように評価するかが課題となり、制度化が模索されている。また、社会的インパクト評価は、研究評価改革と密接に関連した形で推進されてきた。

科学技術振興機構（JST）研究開発戦略センター（CRDS）では、こうした研究評価改革の議論と社会的インパクト評価に関する国際動向、ならびに諸外国や主要なファンディング機関などでの導入状況について調査を行った。本報告書はそれらの情報を整理し、報告するものである。

研究評価改革の議論においては、2013年に公表されたDORAを皮切りに、英国の『メトリクスの潮流（The Metrics Tide）』報告書など、さまざまな組織が、研究評価において既存の定量的指標、特に学術雑誌に基づく指標を利用することに対し警鐘を鳴らす宣言や合意文書を公表してきた。2022年にはサイエンス・ヨーロッパ（ファンディング機関や研究機関の連合）が『研究評価改革についての合意文書（Agreement on Reforming Research Assessment）』を発表し、署名機関の連合体The Coalition for Advancing Research Assessment（CoARA）が発足した。署名機関へ評価改革に向けた実際の活動を求めつつ、人文学や社会科学の評価枠組みの策定、キャリア初期の研究者に適した評価手法の検討など、実務的な改革に踏み込んだワーキンググループでの活動が進展している。さらに、南米では2019年にラテンアメリカ社会科学評議会内に、研究評価を扱うフォーラムFOLEC-CLACSOが立ち上がるなど、研究評価改革の流れは地理的にも広がりを見せている。

研究評価については、主に先進国において、1970年代前後の財政悪化を背景に、公的投資による研究活動が社会にどのような利益をもたらしているのかが問われ続けてきた。近年では、大学等研究機関に対して第三の使命としての社会貢献が強く期待されるようになってきていることから、研究機関や研究活動の社会的インパクト評価が改めて注目を集めている。英国の大学評価制度であるResearch Excellence Framework（REF）など、大学評価制度の中に社会的インパクト調査が導入されている例もある他、研究プログラム、プロジェクトを対象とした評価手法も開発が進んできた。また、評価の利用目的も多様であり、REFのように研究機関の基盤的経費の配分に用いられる場合や、オランダのStrategy Evaluation Protocolのように研究ユニットや研究機関の学習と改善に資する仕組みとして活用されている例もある。

本報告書では、こうした国際動向を踏まえ、さらに具体的な導入や運用状況を把握するために、「研究機関」と「研究課題」を対象とする評価に着目した。「研究機関評価」においては第4章にて、英国、オランダ、イタリア、オーストラリア、日本について事例を記述した。「研究課題評価」については第5章で、まず、米国（NSF、NIH）、ドイツ（DFG）、英国（UKRI-EP SRC）、カナダ（NSERC）など主要国のファンディング

機関における基礎情報をまとめている。また、「実験的ファンディング」や研究課題評価への「AIの活用」といった新しい手法の試行・導入事例についても記述した。

近年、日本でも評価の国際通用性の観点から、DORAをはじめとする研究評価改革に関する国際枠組みへの参加、もしくは参加を検討する組織が増えている。しかし、今回の調査結果を踏まえると、定量的指標の責任ある利用という観点では研究課題評価レベルでの試行・導入が進む一方、研究機関等の評価への導入は依然として限定的であり、国際的な議論や具体的な改革の検討はまだ緒に就いた段階にある。日本の大学等研究機関やファンディング機関において評価改革を実践していくためには、現行の研究評価制度が抱える課題や障壁の検討など、政策的議論を並行して進める必要がある。

社会的インパクト評価については、その目的や評価結果の活用方法が多様であり、資金配分などの意思決定への利用や、研究機関の経営改善・研究の質向上を目的とした学習機会としての利用など、各国の政策目的や制度に応じた幅広い形態があることが明らかとなった。日本では、国立大学評価、すなわち国による研究機関評価の一部として社会的インパクト評価の導入が始まっている。各国の先行事例の調査からは、国の研究評価制度に社会的インパクト評価を導入する際に、評価主体である国等と、被評価者である大学など研究機関、さらに研究活動を担う研究者コミュニティとの丁寧な対話と合意形成が重要であることが示唆された。

本報告書では研究機関と研究課題に関する評価に焦点を置いたが、他にも科学技術・イノベーション政策や施策の評価、研究機関での研究者評価、学術出版社や研究者コミュニティでの学術論文への評価など、研究評価に関する重要なトピックは枚挙に暇が無い。また、研究評価改革に関しても近年オープンサイエンスと深く関わりながら推進されているが、本報告書ではあまり触れていない点にも限界がある。

加えて、本報告書で取り上げた取り組みの多くはここ数年で具体化されたものも多く、効果の検証や取り組みを拡大させていく上での課題が顕在化する可能性が高い。そのため、今後の展開を注視していく必要があるだろう。

目次

1	背景と目的	1
1.1	背景.....	1
1.1.1	研究の「質」の評価へのシフト：研究評価改革の動向.....	1
1.1.2	科学技術の社会的価値の可視化：研究開発の社会的インパクト評価の動向.....	2
1.1.3	新たな研究提案レビュー手法やAIの導入.....	3
1.2	日本の研究開発の評価制度.....	3
1.2.1	国の研究開発評価に関する大綱的指針.....	4
1.3	本報告書における「研究評価」の定義.....	5
1.4	目的と構成.....	6
2	研究評価改革の動向	8
2.1	研究評価改革の背景にある問題意識.....	8
2.1.1	評価への論文数・引用数といった定量的指標の誤った利用への、 研究コミュニティからの問題提起.....	8
2.1.2	研究評価改革、責任ある研究評価とは.....	9
2.2	研究評価改革の拡がり、イニシアチブの立ち上げ.....	10
2.2.1	研究評価に関するサンフランシスコ宣言(DORA).....	10
2.2.2	Coalition for Advancing Research Assessment(CoARA).....	11
2.2.3	ラテンアメリカ社会科学評議会(CLACSO)-ラテンアメリカ研究評価フォーラム (FOLEC).....	15
2.2.4	More than our Rank(MTOR).....	15
2.3	Global Research Councilでの議論と研究評価改革の進行状況調査.....	16
2.3.1	Dimensions of Responsible Research Assessment.....	16
2.3.2	GRC加盟国へのRRA進捗調査(2020、2025).....	16
	コラム1：Research on Research Instituteによる研究評価システムの国際動向調査.....	20
	コラム2：研究評価改革の課題と批判.....	21
	コラム3：研究評価改革の視点からの研究者評価.....	22

3	研究の社会的インパクト評価の動向	24
3.1	社会的インパクト評価の背景	24
3.2	社会的インパクト評価モデル	25
3.2.1	研究機関など評価制度に用いられる社会的インパクト評価モデルの例	25
3.2.2	研究プログラムや研究課題の評価に用いられる評価モデルの例	26
3.3	その他の社会的インパクト評価実施事例	29
3.4	導入する上での課題	29
4	研究機関評価への導入状況	30
4.1	研究評価改革での議論の導入状況	30
4.1.1	英国 REF への導入状況	30
4.1.2	オランダ SEP への導入状況	30
4.1.3	イタリア Valutazione della Qualità della Ricerca (VQR) への反映状況	31
4.2	社会的インパクト評価の導入状況	31
4.2.1	英国 REF Impact Case Study	32
4.2.2	オーストラリア Excellence in Research for Australia (ERA)、Engagement and Impact (EI)	33
4.2.3	オランダ Strategy Evaluation Protocol (SEP)	34
4.2.4	日本 国立大学法人の第4期中期目標期間の評価のうち、ミッション実現戦略分	35
	コラム4：主要国との比較における国立大学法人評価制度	36
5	研究課題評価への導入状況	39
5.1	諸外国の公的ファンディング機関での研究課題評価（事前評価）の状況	39
5.1.1	米国立科学財団(National Science Foundation : NSF)の状況	40
5.1.2	米国立衛生研究所(National Institutes of Health : NIH)の状況	42
5.1.3	ドイツ研究振興協会(Deutsche Forschungsgemeinschaft : DFG)の状況	43

5.1.4	英国工学物理科学研究会議(Engineering and Physical Sciences Research Council : EPSRC)の状況	44
5.1.5	カナダ自然科学・工学研究会議 (Natural Sciences and Engineering Research Council of Canada : NSERC)の状況	45
5.2	研究評価改革の議論を受けた新たな評価手法	46
5.2.1	『Review of Peer Review』報告書	46
5.2.2	実験的ファンディング	49
5.2.3	アルゴリズムによるファンディング —AIの試行的導入—	50
5.2.4	主要国のファンディング機関の申請書にみるCV	52
5.3	社会的インパクト評価の導入状況	54
5.4	研究評価への新たな手法や社会的インパクト評価導入にあたっての課題、留意事項	56
6	まとめ	58
附録	UKRI『Review of Peer Review』報告書に基づくピアレビューへの介入手法一覧	60

1 | 背景と目的

1.1 背景

本報告書は、研究評価について諸外国の①研究評価改革¹の動向、②社会的インパクト評価の観点から調査を行ったものである。

2010年前後から既存の評価システムが研究開発へ負のインセンティブを与えているとの問題意識を背景に、評価のあり方を再検討すべきであるという趣旨の宣言や合意文書²の公表、国際的枠組みの立ち上げが相次いでいる。また、2020年前後から活動の方向性が、「定量的指標の評価への利用を控えるべきである」、「研究成果が掲載される学術雑誌を評価の基盤とするのではなく、個々の研究の価値に基づいた評価をするべきである」といった理念への賛同を募るものから、評価手法の開発や制度への導入といった、実際の評価システムの変革へとシフトしつつある。その一環として、主として民間のファンディング機関を中心に、レビューの不足や評価の公平性の確保といった個別の課題に対する新たな評価手法の実験的導入や効果測定³、評価プロセスでのAIの利用⁴などを試みている。第2章で研究評価改革の国際的動向について記述する。

研究評価のもう一つの動向として、研究評価の対象を学術的な卓越性から社会にもたらした影響へと拡張する、社会的インパクト評価の導入についての議論がある。社会的インパクトを把握するための評価モデルの開発に加え、評価制度への導入が各国で進んでいる⁵。この動向を踏まえ、第3章では研究開発の直接の成果に加え社会へ与えた影響を捉え評価しようとする社会的インパクト評価に焦点を当て記述する。

第4章と第5章では、こうした動向がどのように各国の評価制度へ導入されてきたのか、政策での議論が直接的に影響を与える研究機関など評価と研究課題評価に焦点を絞り整理する。

1.1.1 研究の「質」の評価へのシフト：研究評価改革の動向

近年欧州を中心として、主に研究者の採用、昇進、組織内での評価や助成機関による研究課題採択の場面

- 1 近年では「責任ある研究評価（Responsible Research Assessment、RRA）」と記載されることが多くなってきているが、日本では定着していないことを鑑み、本報告書では「研究評価改革」の語を用いる。
- 2 代表的なものとして、San Francisco Declaration on Research Assessment. 2012.『研究評価に関するサンフランシスコ宣言（DORA）』。https://sfedora.org/read/. アクセス日 2025年12月18日、が挙げられる。
- 3 Bendiscioli, Sandra; Firpo, Teo; Bravo-Biosca, Albert; Czibor, Eszter; Garfinkel, Michele; Stafford, Tom; et al. (2022). *The experimental research funder's handbook* (2nd edition, ISBN 978-1-7397102-0-0). Research on Research Institute. Report. https://doi.org/10.6084/m9.figshare.19459328.v5.
- 4 (RoRI), Research on Research Institute; Newman-Griffis, Denis; Buckley Woods, Helen; Youyou, Wu; Thelwall, Mike; Holm, Jon (2025). *Funding by Algorithm - A handbook for responsible uses of AI and machine learning by research funders* (ISBN 978-1-7397102-2-4). Research on Research Institute. Book. https://doi.org/10.6084/m9.figshare.29041715.v1.
- 5 Reed, M. S., M. Ferré, J. Martín-Ortega, R. Blanche, R. Lawford-Rolfe, M. Dallimer, and J. Holden. 2021. "Evaluating Impact from Research: A Methodological Framework." *Research Policy* 50, no. 4 (May): 104147. https://doi.org/10.1016/j.respol.2020.104147.

で、研究評価改革に向けた知見の共有やツールの開発、研究機関向けのガイドブック⁶の発行などを通して、どのような実践的手法が評価を改革する上で有効なのか検証を行っている。一方、日本でも責任ある研究評価に関する議論が始まっており、日本学術会議が2021年に発出した提言「学術の振興に寄与する研究評価を目指して～望ましい研究評価に向けた課題と展望～」⁷や2025年に発出した「研究の活性化へ向けた研究評価の具体的な改善方策」⁸は、研究の多様性の尊重、評価対象分野の研究者（ピア）や研究成果のユーザーなどによる質的な研究評価を原則とし、定量的評価指標を補助的に活用することが望ましい、とする欧州の責任ある研究評価の議論を踏まえた内容である。また、近年では東京大学や理化学研究所、科学技術振興機構などが新たに「研究評価に関するサンフランシスコ宣言」（The San Francisco Declaration on Research Assessment：以下、DORA）へ署名しており、研究評価改革の理念自体は徐々に浸透してきているように見える。一方、欧州で試みられている、評価手法の試行と効果の検証の動きはまだ見られない。理念の受容という面でも、「定量的指標は補助的に用いればよい」という表面的な対応に留まっている例があり、現状の研究評価システムは研究の多様性や創造性、若手研究者の育成、成果の社会的なインパクトなどの多様な貢献を十分に評価できておらず「出版か死か（Publish or Perish）」の言葉が示す負の文化を生んでいるという、本来の問題意識が共有されているとはいえない状況にある。

1.1.2 科学技術の社会的価値の可視化：研究開発の社会的インパクト評価の動向

英国を中心にさまざまな国で科学技術が社会へ与えたインパクトを捉え評価する試みがなされている。日本でも国立大学法人評価にて運営費交付金のうちミッション実現戦略分を、「社会的インパクト」についての評価を通じて配分することが決まっている⁹他、国際卓越研究大学事業、地域中核・特色ある研究大学強化促進事業（J-PEAKS）事業採択大学においても社会的インパクト評価を実施する¹⁰。このように大学などの研究機関の社会的価値を可視化する動きが日本でも出てきている。

諸外国、特に英国のResearch Excellence Framework（以下、REF）の一部として実施されている社会的インパクト評価については国内の関心も高く、しばしば紹介されてきた¹¹。一方、国内で研究開発の社会的インパクト評価手法の蓄積は進んでいない状況である。また、評価に際しての研究機関側の人的、事務的コストなど日本で実施する際に想定しておくべき論点についても評価制度の企画者、評価者、被評価者の間で共有されていないという課題がある¹²。

6 Allen, Liz, Virginia Barbour, Kelly Cobey, Zen Faulkes, Haley Hazlett, Rebecca Lawrence, Giovanna Lima, Flora Massah, と Ruth Schmidt. 「A Practical Guide to Implementing Responsible Research Assessment at Research Performing Organizations」. Declaration on Research Assessment (DORA), 2025年5月13日.
<https://doi.org/10.5281/zenodo.15000683>.

7 日本学術会議. 2021.『提言「学術の振興に寄与する研究評価を目指して～望ましい研究評価に向けた課題と展望～」』. 東京：日本学術会議. <https://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-25-t312-1.pdf>. アクセス日 2025年12月18日.

8 日本学術会議. 2025.『提言「研究の活性化へ向けた研究評価の具体的な改善方策」』. 東京：日本学術会議.
<https://www.scj.go.jp/ja/info/kohyo/pdf2/kohyo-26-t394-1.pdf>. アクセス日 2025年12月18日.

9 文部科学省. 2021.「第4期中期目標期間における国立大学法人運営費交付金の在り方について 審議まとめ」. 東京：文部科学省.
https://www.mext.go.jp/content/20210624_mxt_hojinka_000016041_2.pdf. アクセス日 2025年12月18日.

10 文部科学省科学技術・学術政策局,「地域中核・特色ある研究大学強化促進事業（J-PEAKS）全体の事業評価実施要領 別紙」.
https://www.mext.go.jp/content/20250228-mxt_sanchi01-000025865_5.pdf. アクセス日 2025年2月28日.

11 国立大学協会. 2022. “英国研究評価制度 REF 及び社会的インパクト評価に係る講演会『大学における社会的インパクトとは？—英国の事例に学ぶ—』を開催（1/20）.” 2022年1月24日.
<https://www.janu.jp/news/9426/>. アクセス日 2025年11月13日.

12 2023年度の日本評価学会全国大会では「大学の社会的インパクト評価を考える」とのシンポジウムが開催され、今後どのような視点で評価を行うべきか議論された。その中でコストも含め阻害要因に関しても取り扱われた。

1.1.3 新たな研究提案レビュー手法やAIの導入

研究助成機関において現在の研究課題評価が抱える課題、中でも研究提案の増加に伴う人的コストの増大と評価結果のバイアスへ対応する形で、新たな研究提案のレビュー手法や評価プロセスへの部分的なAIの導入が提案されつつある。日本でも文部科学省にて部分的ランダム化ファンディングの取り組みが紹介される¹³など、関心が高まりつつある。

1.2 日本の研究開発の評価制度

日本の研究開発の評価制度は政策評価法をはじめとした各種の法令や制度によって評価の目的、実施主体、評価対象、評価の時期が定められている。表1-1で示した評価制度に加えて、各研究機関にて研究者評価を実施している。

表 1-1 研究開発に関連する日本の評価制度¹⁴

評価制度など	評価の目的	評価の実施主体	評価の対象	評価の時期
政策評価 (行政機関が行う政策の評価に関する法律)	効果的かつ効率的な行政の推進、国民への説明責任の徹底	行政機関の長	行政機関の所掌に関する政策	施策を5年ごとに事後評価 総額10億円以上の研究開発事業の開始前(事前評価)
国立研究開発法人評価 (独立行政法人通則法)	「研究開発成果の最大化」「適正、効果的かつ効率的な業務運営」の両立の実現	主務大臣(国立研究開発法人審議会が助言)	業務実績(研究開発の評価だけではなく、法人のマネジメントに関わる業務運営についても評価)	中長期目標期間の毎年度、最終年度、終了時
国立大学法人評価 (国立大学法人法)	質的向上、社会への説明責任、次期以降の中期目標への反映、運営費交付金などの算定に反映	文部科学省国立大学法人評価委員会(教育研究の評価は、大学改革支援・学位授与機構の評価結果を尊重)	業務実績(教育研究活動や業務運営、財務内容などの総合的な達成状況について評価)	中期目標期間(6年間)の4年目、終了時
認証評価 (学校教育法)	大学の教育研究水準の向上に資する	認証評価機関	大学の教育研究など(教育研究、組織運営、施設設備)の総合的な状況	7年以内ごと
行政事業レビュー (行政事業レビューの実施などについて(閣議決定))	事業のより効果的かつ効率的な実施、国の行政に関する国民への説明責任および透明性の確保	各府省庁自ら	原則全ての事業	毎年度(概算要求前に自己点検を実施)

13 文部科学省研究振興局基礎・基盤研究課、「基礎研究の振興について～評価の視点から～」、令和7年7月7日。
https://www.mext.go.jp/content/20251110-mxt_kiso-000045742_4.pdf. アクセス日2025年12月18日。

14 各種法律や政府文書をもとにCRDSが作成。

研究開発評価 (①研究開発プログラム、②研究開発課題評価、③研究者などの業績、④研究開発機関など) (国の研究開発評価に関する大綱的指針 (内閣総理大臣決定))	①政策・施策などの効果的・効率的推進、②研究開発の質の向上、③研究者の処遇への反映、④機関の資源配分への反映など	①プログラムを推進する府省など、②課題を設定・推進する府省など、③④機関の長	国費を用いて実施される研究開発	①②開始前、中間、終了時、追跡、③機関が設定、④3～7年ごとを目安
国家的に重要な研究開発の評価 (内閣府設置法)	国の科学技術政策を総合的かつ計画的に推進。評価結果を推進体制の改善や予算配分に反映	内閣府 総合科学技術・イノベーション会議 (および同評価専門調査会)	国費総額が約300億円以上の研究開発のうち、調査会が評価すべきと認めたものなど	事前、中間、終了後 (事業終了の次年度および調査会が必要と認めた時点)

それぞれの評価制度ごとに課題が指摘されている。評価全般については評価の設計や実施に専門性を持つ人材の不足がここ四半世紀指摘¹⁵されつつも改善の道筋は見えていない。一方で評価の事務的な負担、いわゆる評価疲れについては近年対策が検討されつつあり、2024年に内閣府が行った調査¹⁶では申請書や報告書の執筆負担、申請書への評価フィードバックの要望、事務手続きの簡素化・柔軟化・統一化への要望などが指摘されている。また、2021年には国立大学協会が大学法人評価について、画一的な中期目標、中期計画の策定から個々の法人の特色に応じた柔軟な形へ向けた検討の必要性や、大学のより自律的な経営に向けた目標として機能させるべき、法人評価と他の評価制度が重複しており業務が増加している、などが課題として指摘されている¹⁷。

1.2.1 国の研究開発評価に関する大綱的指針

これらの評価制度の基礎として国としての研究評価の方針を定めたものが『国の研究開発評価に関する大綱的指針』である。この指針では、研究評価を府省、研究助成機関、大学など研究機関の3階層に分け、府省と研究助成機関ではプログラム評価、プロジェクト評価、研究機関では機関評価、研究者評価を行うこととされている。1997年に公表されて以降、科学技術基本計画の策定に合わせ大綱的指針の重点も変遷している。例えば、1997年の初版では、良い研究提案をどのように採択するか、研究機関の評価の2点が強調されており、2001年の改訂では研究提案の採択と研究機関の評価に加えて政府の施策の評価、研究者の業績評価を実施するとしている。2005年の改訂では評価疲れへの対応、「創造への挑戦を励ます」意欲を向上させる評価、が強調されている。2016年に改訂された最新版¹⁸までの変遷を表1-2で示す。

15 内閣府 平成17年度内閣府委託調査、「研究開発評価の人材養成システムに関する調査報告書」、平成18年3月。
<https://www.ifeng.or.jp/wordpress/wp-content/uploads/2012/10/CR-2005-181.pdf>. アクセス日 2025年12月18日.

16 内閣府 科学技術・イノベーション推進事務局 研究に専念する時間の確保に向けて～評価疲れアンケートの結果と対応策～、令和6年6月13日。
https://www8.cao.go.jp/cstp/package/wakate/result_summary_20240613.pdf. アクセス日 2025年12月18日.

17 国立大学協会、第4期中期目標期間へ向けた国立大学法人の在り方について-強靱でインクルーシブな社会実現に貢献するための18の低減、令和3年6月。
<https://www.janu.jp/wp/wp-content/uploads/2021/06/38bd0f7a612a527239d991d72d2a1dae.pdf>. アクセス日 2025年12月18日.

18 内閣府 科学技術・イノベーション推進事務局. 2016.「国の研究開発評価に関する大綱的指針」(2016年12月). 東京:内閣府. アクセス日 2025年12月18日. <https://www8.cao.go.jp/cstp/kenkyu/taikou201612.pdf>. アクセス日 2025年12月18日.

表 1-2 『国の研究開発評価に関する大綱的指針』の変遷¹⁹

発行、改訂年	発行、改訂時の重点
1997年	良い研究開発提案の採択、研究機関評価の実施
2001年	上記に加え施策と研究者の業績評価を実施
2005年	評価疲れへの対応 「創造への挑戦を励ます」意欲向上のための評価
2008年	評価結果に基づく資源配分の実施 資源配分を通じた研究成果の社会還元効率化、迅速化
2012年	アウトプット重視からアウトカム重視への転換 評価単位として「プログラム」概念の導入
2016年	実効性のある研究開発プログラム評価のさらなる推進 挑戦的な研究開発の推進 評価負担の軽減

表 1-2 で示したように、改訂を重ねるにつれ研究開発評価の対象が拡大し、重視する点や評価目的も変遷している。評価疲れは一貫した課題として示される他、定量的な指標は不完全であり補助的に用いるべきとの言及も各版で触れられている。一方で、現在国際的に研究評価改革の文脈で議論されている、既存の評価システムへの省察や多様な学術活動を捉えた評価のあり方のトーンは薄い。また、2012年の改訂以降、研究開発のアウトカムを重視する傾向は見られるものの、社会的な価値の可視化、社会的インパクト評価の観点は、科学技術政策が科学技術・イノベーション政策へ拡大している近年の政策環境と比して言及は少ない。また、なにより大綱的指針自体の趣旨は責任ある研究開発や社会的インパクト評価を含んでいるものの、実際に評価を行う際に適切な手法を検討できる評価の専門性を持った適切な支援者を得られることは極めて稀であり、評価実践への影響に乏しいことが大きな課題である。

1.3 本報告書における「研究評価」の定義

日本で研究開発評価といった場合、事前評価や事後評価といった評価タイミングに加え、評価対象として科学技術・イノベーション政策、大学などの研究機関、研究プロジェクトの採択、採択後の中間評価やステージゲート、研究者まで含み、評価目的もそれぞれ異なるため、立場によって想起するものは異なる。英語の場合 Assessment、Monitoring、Evaluation と分かれている概念が全て「評価」の一語に入れられるため、評価に関して議論する際にどの評価を指しているのか分かりづらい部分がある。

本報告書では、「研究評価」を緩やかに「研究開発に関する活動（政策、研究提案、成果など）に対する調査活動と価値判断」と定義する。下記の表 1-3 では、評価対象と評価時期の 2 軸で評価活動を示した。この 2 軸に加え、評価目的としておおまかに①資金配分や事業の継続といった意思決定、②改善・学習、③対外的な説明責任があることに留意されたい。

19 各版の「国の研究開発評価に関する大綱的指針」から CRDS が作成。

表 1-3 評価対象と評価時期の例²⁰

	事前評価	中間評価、モニタリング	事後評価、追跡評価
STI政策評価	事業の事前評価、予算査定など	事業の中間評価、年度報告など	事業事後評価、フォローアップ評価、社会的インパクト評価
研究機関評価	(中期目標中期計画策定などの国立大学法人評価) ²¹	中期計画中間評価など国立大学法人評価	中期目標中期計画、J-PEAKSに伴う社会的インパクト評価
課題評価	研究プロジェクト採択	年度報告書、ステージゲートなど	事後評価、フォローアップ評価
研究者評価	研究者採用、昇進	人事業績評価	賞や称号の授与

1.4 目的と構成

本報告書の目的は、現在の研究評価に関する議論の中で①研究評価改革、②社会的インパクト評価、の2点に焦点を絞り、研究機関評価と研究課題の評価へどう適用されているか調査することである。研究評価改革については、現行の評価制度が研究活動へ負のインセンティブを与えており、研究活動の一部しか評価対象にできていない、との国際的に共通した問題認識がある。この問題意識に対して公表されてきた宣言や国際枠組みの動向に加え、実際にどのような実践やツール開発を通じて責任ある研究評価を実現しようとしているのか、欧州を中心とした潮流を把握に努めた。研究の社会的インパクト評価については、社会的インパクト評価が求められた背景や諸外国での評価枠組みに加え、これまでに試みられてきた社会的インパクト評価モデル、手法の把握に努めた。また、これらを日本の評価システムへ導入する上での留意点についても記載した。

本報告書の構成は以下である。まず、欧州を中心とした研究評価改革についての近年の動向を、国際的枠組みの取り組みに重点を置き記述した(第2章)。次に研究開発の社会的インパクト評価の歴史的な背景や、これまでに開発されてきた評価モデル、手法について記述した(第3章)。そして、研究評価改革での議論と社会的インパクト評価が、研究機関評価と研究課題評価へどのように導入されつつあるのか、国際動向を中心に整理した(第4章、第5章)。

研究機関評価については、諸外国の評価制度ガイドラインなどへの明示や評価の利用目的、社会的インパクトの把握手法について取り上げる。研究課題評価については、研究助成の効率化や効果の最大化、バイアスの低減を目的とした新たな評価方式の事例とともに、研究を評価する基準として社会課題やニーズが導入されている事例を取り上げた。

20 各種評価制度に関する文書からCRDSが作成。

21 国立大学法人等の中期目標、中期計画の策定には評価委員会が意見することが可能であり、本報告書では事前評価として扱う。

表 1-4 本報告書の対象

評価時期 評価対象	事前評価	中間評価、モニタリング	事後評価、追跡評価
STI政策評価 (含プログラム評価)	事業の事前評価、 予算査定など	事業の中間評価、 年度報告など	事後評価、フォローアップ 評価、社会的インパクト評価
研究機関等評価	研究機関等の法人の 計画策定	研究機関等の年次評価、 中間評価 4章の対象	機関活動の事後評価、 社会的インパクト評価 など
研究課題評価	研究プロジェクト の採択	プロジェクトの中間評価、 ステージゲート評価 など 5章の対象	プロジェクトの事後評価、 フォローアップ評価 など
研究者評価	研究者の採用、昇進	人事業績評価 など	賞や称号の授与

2,3章の対象

2 | 研究評価改革の動向

2.1 研究評価改革の背景にある問題意識

2.1.1 評価への論文数・引用数といった定量的指標の誤った利用への、研究コミュニティからの問題提起

これまで研究評価の場面で論文数や被引用数、ジャーナル・インパクト・ファクター（Journal Impact Factor：以下、JIF）、h指数といった定量的な指標の限界やこれらの指標を偏重することへの批判は、度々報告書や宣言として公表されてきた。例えば、本章で取り上げる研究評価に関するサンフランシスコ宣言（The Declaration on Research Assessment: DORA）や2015年に発表されたライデン宣言²²、UKRIの『メトリクスの潮流』報告書²³では、そもそもこれらの指標が個々の論文や個人を直接的に評価する指標ではないことや、研究歴の長さ大きく影響されること、そして何よりも研究自体に負のインセンティブを与えてきたことへの問題意識を表明している。つまり、論文以外の学術への貢献が見落とされていることや、JIFの高い雑誌への投稿が可能かとの観点が研究活動の基準となっていることについて、研究現場の危機感がある。その上で、研究評価のあり方を変革することで研究活動に正のインセンティブをもたらそう、というのが研究評価改革の趣旨の大きな要素とされる。

表 2-1 研究評価に関する主要な提言や宣言²⁴

提言など	DORA (2013年)	ライデン宣言 (2015年)	香港原則 (2019年)	ラテンアメリカ研究評価フォーラム (2019年)	CoARA (2022年)	Dimensions of Responsible Research Assessment (2024年)
主な問題意識・論点	<ul style="list-style-type: none"> 評価における定量的指標の過度な利用 個々の研究ではなく発表された雑誌に基づいた研究評価が行われている 	<ul style="list-style-type: none"> 計量が評価の代替として用いられている 英語以外の言語での研究の軽視 データ収集、分析のブラックボックス化 	<ul style="list-style-type: none"> 疑わしい研究慣行の横行 幅広い研究活動が評価されていない 不当なインセンティブの最小化 	<ul style="list-style-type: none"> 評価における定量的指標の過度な利用 ジェンダーの不均衡 文化、民族、研究テーマといった面での多様性の考慮 	<ul style="list-style-type: none"> 評価における定量的指標の過度な利用 大学ランキングの評価への利用 新たな研究評価ツールやプロセスの開発 	<ul style="list-style-type: none"> 責任ある研究活動の評価と報奨 研究から社会的インパクトに至る経路の多様性 評価者へのトレーニング

一方で、研究の「卓越性」という概念も曖昧であり、研究者間で共通した明確な定義や評価基準が存在し

22 Hicks, D., Wouters, P., Waltman, L. et al. Bibliometrics: The Leiden Manifesto for research metrics. *Nature* 520, 429–431 (2015). <https://doi.org/10.1038/520429a>.

23 Wilsdon, J., L. Allen, E. Belfiore, P. Campbell, S. Curry, S. Hill, R. Jones, R. Kain, S. Kerridge, M. Thelwall, J. Tinkler, I. Viney, P. Wouters, J. Hill, and B. Johnson. 2015. *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. London: HEFCE. DOI:10.13140/RG.2.1.4929.1363.. <https://www.ukri.org/wp-content/uploads/2021/12/RE-151221-TheMetricTideFullReport2015.pdf>.

24 UK Research and Innovation (UKRI). 2015. *Review of metrics in research assessment and management*. から CRDS が作成。

ない²⁵のが実情である。研究の卓越性を構成する要素は、研究自体の信頼性や学術への貢献といった研究活動自体²⁶や、研究活動の蓄積によって生じるレピュテーション、国際的なネットワーク、資金獲得能力など、多くの側面が挙げられる。しかし、研究の卓越性を測る代理指標として定量的指標を用いることにはさまざまな問題が伴うというのは、研究評価改革の議論の中で共通している。例えば、JIFは論文の引用数との相関が弱く、論文の卓越性の代理指標としては不適切であるとされる。また、FWCI (Field-Weighted Citation Impact) をはじめとするさまざまな計量書誌学的指標も、研究の質を直接評価することはできない²⁷。これらの問題に対処し、研究文化をより健全な方向に導くためには、定量的指標を責任ある形で用いる研究評価へ向けた改革が不可欠であると考えられている。

この問題意識は、論文以外の研究成果の評価にもつながっている。例えば、プレプリント、オープンデータ、研究に用いるプログラムのソースコード、学会での指導的立場といった、従来の定量的指標では捉えにくい学術的な貢献をどのように評価すべきか、という課題がある。さらに、研究機関のミッションが社会課題への対応へと拡大する中で、研究が社会に与えるインパクトや、ステークホルダーとのエンゲージメントをいかに評価システムに組み込むかという点も、研究評価改革の重要な課題となっている。

2.1.2 研究評価改革、責任ある研究評価とは

近年、研究評価改革に関する活動は「責任ある研究評価 (Responsible Research Assessment : RRA)」に向けた活動の一環として取り扱われてきつつある。RRAは、「責任ある研究・イノベーション (Responsible Research & Innovation : RRI)」の考え方をその基盤としている²⁸。RRIは、「現在における科学とイノベーションの集合的な管理を通じた未来に対するケアを意味する」とされ、科学技術・イノベーションを進めるプロセスの正統性、妥当性、透明性を向上させることにより、応答責任の明確化、倫理的な受容可能性、社会的要請への応答、潜在的危機への洞察深化などを促すことを目指している。研究者側の問題意識と呼応する形で、欧州を中心に研究機関間のネットワークや国による政策の一部としても推進されてきた。

英国 UKRIによる『メトリクスの潮流』報告書では、RRAの基本的な観点には以下の5つがあると指摘している。

- 頑健性 (Robustness) : データが正確で、その範囲が適切であることを重視する。
- 謙虚さ (Humility) : 定量的評価は、質的評価や専門家による判断を支援する補助的な役割を果たすべきであるという考え方。
- 透明性 (Transparency) : 評価に用いられるデータと分析プロセスが公開され、その結果の妥当性が検証可能であることを保証する。

25 Nature Editorial. 2018. "Science Needs to Redefine Excellence." *Nature* 554, no. 7693 (February 21): 403-404. <https://doi.org/10.1038/d41586-018-02183-y>.

26 Pär Mårtensson, Uno Fors, Sven-Bertil Wallin, Udo Zander, and Gunnar H. Nilsson. "Evaluating Research: A Multidisciplinary Approach to Assessing Research Practice and Quality." *Research Policy* 45, no. 3 (2016): 593-603. <https://doi.org/10.1016/j.respol.2015.11.009>.

27 Declaration on Research Assessment (DORA). 「Guidance on the Responsible Use of Quantitative Indicators in Research Assessment」. DORA, 2024年9月26日. <https://doi.org/10.5281/zenodo.13844873>.

28 Wilsdon, J., L. Allen, E. Belfiore, P. Campbell, S. Curry, S. Hill, R. Jones, R. Kain, S. Kerridge, M. Thelwall, J. Tinkler, I. Viney, P. Wouters, J. Hill, and B. Johnson. 2015. *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. London: HEFCE. DOI:10.13140/RG.2.1.4929.1363.. <https://www.ukri.org/wp-content/uploads/2021/12/RE-151221-TheMetricTideFullReport2015.pdf>.

- 多様性 (Diversity) : 研究分野の特性や、研究者個々のキャリアパスの多様性を反映し、支援するような指標の使用を促す。
- 省察性 (Reflexivity) : 指標がもつ潜在的かつシステム上の効果を常に認識し、それに応じて評価システムを更新していく姿勢を指す。

これらの観点は、日本学術会議が2021年に発表した提言「学術の振興に寄与する研究評価を目指して-望ましい研究評価に向けた課題と展望-」にも通じるものである。また、サイエンス・ヨーロッパやGlobal Research Councilが公表してきた責任ある研究評価の原則や側面にもこれらの観点が含まれている。

2.2 研究評価改革の拡がり、イニシアチブの立ち上げ

2.2.1 研究評価に関するサンフランシスコ宣言 (DORA)

DORAは、2013年5月に発表された研究評価改革の重要なマニフェストである。その主眼は、ジャーナルインパクトファクター (JIF) をはじめとする計量書誌学的指標の不適切な使用を批判し、研究成果の質を評価する際には、その内容と影響をより包括的かつ多角的に評価する必要があると主張する点にある。DORAは、論文数やJIFといった単一の指標に過度に依存することなく、研究の真の価値を認識するための評価システム構築を目指している。

DORAの重要なポイントは、論文以外の多様な成果を評価システムに含める必要性を強調している点だ。プレプリント、オープンデータ、プログラムのソースコード、学会での指導的立場、社会貢献活動など、従来の定量的指標では捕捉しにくい学術的貢献を適切に評価する枠組みを求めている。これは、研究機関のミッションが社会課題への対応へと拡大する現代において、社会的インパクトやエンゲージメントといった要素をいかに評価に組み込むかという課題にも通じる。

DORAは、定量的指標の利用に関する具体的なガイドラインも提供している。JIF、h指標、FWCI (Field-Weighted Citation Impact)、オルトメトリクスといったさまざまな定量的指標について、それぞれの強みと課題を詳細に指摘し、評価に用いる場合の注意点を具体的に列挙している。例えば、h指標は研究歴に大きく影響を受けるため、シニア層に有利に働きやすいといった点を指摘している。

さらにDORAは2021年からアルカディア財団より資金提供を受け、研究評価の進歩へ向けたツールキット開発プロジェクト (Project Tools to Advance Research Assessment : Project TARA) を実施し、評価改革へ向けた研究機関を対象とする評価ルーブリック²⁹や研究機関向けの実践ガイドの発行を通じて、研究評価改革を具体的な組織の活動に落とし込むためのツールキットを提供している。このガイドは、改革を成功させる上で重要な要素として、研究コミュニティ全体の構築、研究機関の幹部層のサポート、既存の評価システムを漸進的に変化させていく継続性を挙げている。具体的な導入事例として、東京大学でのコミュニティ構築や、中国の国家自然科学基金委員会 (NSFC) における若いレビュアー向けのトレーニングの提供などが紹介されている。

2025年までに計3,500余りの機関が組織としてDORAへ署名しており、その多くが大学や研究機関、出版社である。その他、ファンディング機関も署名している。日本ではJSTを含め18の研究機関、学会などが

29 Anna Hatch, Ruth Schmidt. SPACE to Evolve Academic Assessment: A Rubric for Analyzing Institutional Conditions and Progress Indicators. Zenodo, 2021年6月11日. <https://doi.org/10.5281/zenodo.4927605>.

DORAに署名している。署名していない研究機関においても、例えば京都大学は責任ある研究評価の文脈に沿った形で学内評価システムの確立を研究推進戦略の一つとして位置づけている³⁰。研究者や研究機関による取り組みが始まりつつある。

2.2.2 Coalition for Advancing Research Assessment (CoARA)

CoARAは、研究評価改革を推進するための最も重要な国際的イニシアチブの一つである。2022年7月、サイエンス・ヨーロッパ（研究助成・実施機関の連合）、欧州大学協会、欧州委員会が中心となり、研究評価改革に関する合意文書の最終版³¹が公開され、同年9月28日の研究・イノベーション・デイにCoARAが正式に発足した。CoARAの主要な目標は、研究の質とインパクトを最大化することであり、そのためにピアレビューを中心とした質的評価を重視するアプローチを採っている。

CoARAには、研究助成機関、研究機関、大学、国際機関など、多様な組織が参加可能であり、各参加機関は以下の原則に合意するとともに、10のコミットメントを実施することとされている。

CoARA全体原則

包括的な条件のための原則

- 倫理・インテグリティに関する規則と慣行の遵守
- 科学的研究の自由の保護
- 研究機関の自律性の尊重
- 研究評価と研究のインパクト決定に必要なデータ・インフラ・基準の独立性と透明性を確保

評価基準・プロセスに関する原則

- 質・インパクト
- 多様性・包摂性・協働

CoARA署名機関に求められる10のコミットメント

- 研究の必要性や特性に応じて、研究への貢献やキャリアに多様性があることを認識する。
- 研究評価はピアレビューを中心とした質的評価に基づき、定量的指標の責任ある利用によりサポートされる。
- 研究評価において、ジャーナルや出版物に基づく評価基準の不適切な使用を放棄する。
- 研究評価における研究機関ランキングの利用を避ける。
- コミットされた組織的な変化を達成するために、必要な研究評価の改革に資源を投入する。
- 研究評価基準、ツール、プロセスを見直し開発する。
- 研究評価改革に対する認識を高め、評価基準や評価プロセス、その使用方法について、透明性のあるコミュニケーション、ガイダンス、トレーニングを提供する。
- 実践と経験を交換し、連合内外の相互学習を可能にする。
- 原則の遵守とコミットメントの実施に関する進捗状況を報告する。
- 確固たるエビデンスと最先端の研究の研究（research on research）に基づき、実践・基準・ツールを評価し、エビデンス収集と調査のためにデータをオープンにする。

30 京都大学 COMON: 対話と価値の共創による京都大学独自の研究評価システム、<https://www.research.kyoto-u.ac.jp/comon/>. アクセス日2025年12月18日。

31 Arentoft, Michael, Stephane Berghmans, Lidia Borrell-Damian, Silvia Bottaro, Jean-Emmanuel Faure, Vinciane Gaillard, Kostas Glinos, ほか. 「Agreement on Reforming Research Assessment」. Zenodo, <https://doi.org/10.5281/zenodo.13480728>.

2023年には組織の機能強化や国際的なネットワークの拡大、研究評価改革へ向けた試行的取り組みを支援するため、EUの枠組みプログラム Horizon Europe から3年間で約9億2600万円³²（500万ユーロ）の資金を獲得した。この予算によって、後述するワーキンググループの活動やCoARA Boostプロジェクトなどの取り組み³³が支援されている。

2025年11月時点で、欧州を中心にファンディング機関や大学など800超の組織がCoARAに署名しており、ドイツ、フランスなど16ヶ国で具体的なコミットメントに向けた活動を行う国別の支部（ナショナル・チャプター）が活動中である。これらの署名機関は、各々がアクションプランを策定し、その進捗を共有している。

CoARAの研究評価改革へ向けた具体的活動

個別課題に対応したワーキンググループ³⁴

CoARAは、その理念を実際の評価活動に落とし込むため、評価プロセスにおけるさまざまなバイアスへの対応や、新たな評価手法の実験などの目的に応じてワーキンググループを設置しており、具体的なツールの開発にも取り組んでいる。例えば、「責任ある研究評価のためのインフラ構築」ワーキンググループは、個別機関ごとの研究情報インフラから、持続可能で相互運用性のあるインフラへの移行に関する検討を進めている。また、「学術キャリア評価」ワーキンググループは、研究歴によるバイアスをどのように排除可能かを検討している。さらに、「研究申請書の改善、進化」ワーキンググループは、研究採択の評価基準や応募者に提出を求める情報の取捨選択などについて議論している。

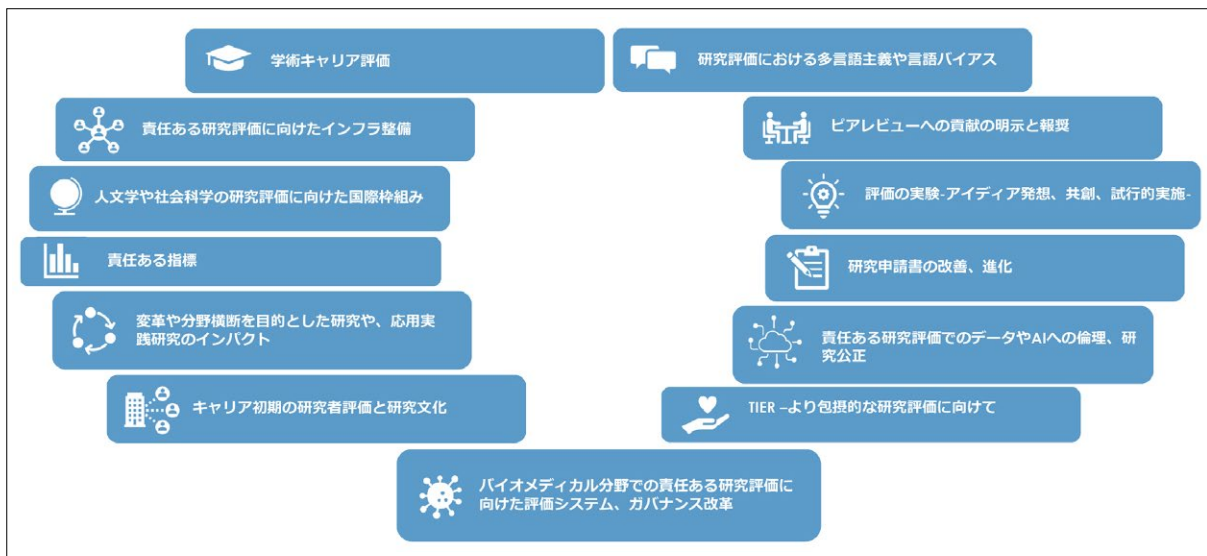


図 2-1 CoARA ワーキンググループ一覧³⁵

32 2026年1月21日時点での、1ユーロ=185.22円で換算。

33 EUが資金提供するプロジェクトが母体となり活動助成の公募を行う、カスケードファンディングも行われている。

34 Coalition for Advancing Research Assessment (CoARA). 2025. "Working Groups + National Chapters." <https://www.coara.org/wg-nc/>. アクセス日 2025年12月18日.

35 Toth Cziffra, e., & Coalition for advancing research assessment (CoARA) i. (2025, 10月). CoARA generic presentation からCRDSが作成

各国の実情に合わせた署名機関ネットワークの構築（ナショナルチャプター）³⁶

ナショナルチャプターとは、同一国に所属するCoARA加盟機関が緩やかに連携し、自国の制度状況や政策文脈に合わせた研究評価改革の推進を図るための、半自律的なネットワークである。各チャプターはCoARA本体の下部組織ではなく、むしろ加盟機関同士のボトムアップな協働体として形成される。

CoARAが目指す改革を実践するには、国や機関によって成熟度や政策環境からもたらされる障壁も大きく異なる。研究評価の制度は、大学の自治、国の高等教育政策、研究ファンドの仕組み、学問分野の慣行などの影響を受けるため、CoARAとして共通原則を掲げるだけでは十分ではない。そこで、加盟機関が互いに知識を共有し、自国の状況に応じた改革ロードマップを描く必要が生じる。ナショナルチャプターは、このギャップを埋める役割を担い、国内の大学、研究機関、ファンディング機関、アカデミー、学会などが一体となって議論・実験・評価の循環を作り出す場として期待されている。

ナショナルチャプターの活動は各国でやや異なるが、一般的には次のようなものが挙げられる。

- 国内における研究評価改革の情報共有：CoARA本体の作業部会（WG）の成果、欧州・国際的な動向、政策レベルの議論を国内で共有し、署名機関の理解促進を図る。
- 実践例・ガイドラインの相互学習：指標の使用方針、査読プロセスの改善、質的評価の導入方法、研究者キャリア評価の改革など、各機関の試行事例を共有し、横断的な学習コミュニティを形成する。
- 政策対話の場の創出：国の科学技術・イノベーション政策当局やファンディング機関と連携し、制度レベルの改革課題を整理する。科学技術・イノベーション政策を担当する府省やナショナルアカデミーがチャプターと連携する事例が増えている。
- CoARA年次報告書や評価サイクルへのインプット：各国の研究評価改革の進捗を国際コミュニティに共有し、各国の状況に応じたフィードバックを受ける。
- 共同プロジェクトの形成：評価手法開発、新たな評価手法の試験的導入、博士人材評価の改革など、国内で必要性の高いテーマについての共同プロジェクトを形成する。

ナショナルチャプターの設立には以下の要素が求められる。

- 国内加盟機関による合意形成（CoARAに参加していない機関もオブザーバとして関与可能）
- チャプター内のコーディネーターまたは代表機関の選出
- 年次活動計画とガバナンス方針の策定
- CoARA Steering Board への設立申請と承認

承認後は、各チャプターが独立したネットワークとして活動し、CoARAとの接点を保ちながら国内活動を展開する。

2023年から2025年にかけて、フランス、オランダ、ドイツ、イタリア、スペイン、北欧諸国などで続々とナショナルチャプターが設立されている。国によっては、研究評価改革を国家施策として進める動きと強く結びつき、CoARAチャプターが政策コミュニティの核となっている例もある。特にオランダやノルウェーでは、研究文化改革や人事・キャリア評価改革と一体的に進められている。

36 Coalition for Advancing Research Assessment. 2025. "National Chapters."
<https://www.coara.org/wg-nc/national-chapters/>. アクセス日2025年12月18日.

CoARA Boostプロジェクト³⁷

CoARA Boostプロジェクトは、評価ツールの開発や機関間の相互学習を促進するため、Horizon Europeの枠組みの中で、2023年から2026年にかけて総額500万ユーロの予算で実施されている。

CoARA カスケードファンディング採択機関とプロジェクト概要（抜粋）

- **チームプロジェクト：先進的な評価改革に取り組んでいる組織から、その知識と経験を他の組織へと移転することを目的としている。**
 - ➔ キプロス大学、リエカ大学（クロアチア）、キプロス欧州事務所：研究キャリアの向上と政策革新のためのオープンサイエンスコラボレーション。
 - ➔ オウル大学（フィンランド）、ルール大学ポーフム（ドイツ）、タンペレ大学（フィンランド）：CoARA ExPECT – 評価基準の変更を共同学習し促進する。
 - ➔ カルタゴ大学（UCAR） – 国立応用科学技術研究所（INSAT）（チュニジア）、カタルーニャ生物工学研究所（IBEC）（スペイン）：カルタゴ大学における研究評価の改革（質的アプローチ、学際的モデル、持続可能なプロセスなど）へ向けた連携を行う。
- **機関パイロットプロジェクト：個々の研究機関が、新たな評価アプローチやツールを試行し、その有効性を検証。**
 - ➔ スウォンジー大学（英国）：研究評価改革の影響と学術キャリア促進におけるナラティブCVの影響評価。
 - ➔ ソフィア大学（ブルガリア）：研究評価をオープンサイエンスに向けて推進する手法の開発。
 - ➔ バルセロナ自治大学（スペイン）：研究評価のための新しいアプローチと指標の試行。
- **制度改革プロジェクト：各組織が自らの評価制度を改革するための具体的活動を支援。**
 - ➔ トリノ工科大学（イタリア）：FEDRA – 透明性、包括性、再現性のある研究評価のために、FAIR原則（見つけやすく -Findable-、アクセスしやすく -Accessible-、相互運用可能で -Interoperable-、再利用可能 -Reusable-）に則った研究機関向けデータレイクを整備。
 - ➔ グラーツ大学（ドイツ）：研究評価改革に伴う学内文化の移行管理。
 - ➔ チェコ科学大学（チェコ）：内部プログラムのための新しい評価ツール利用。

署名機関によるアクションプラン、タイムフレームの公開と共有³⁸

CoARAに署名した機関は、そのアクションプランとタイムフレームをZenodo（研究成果のオープンアクセスリポジトリ）で公開し、共有している。これは、各機関がCoARAのコミットメントをどのように具体的に実行していくかを示すものであり、他の機関との相互学習を促進する目的も持つ。

37 Coalition for Advancing Research Assessment. 2023. “CoARA Boost Project: A Catalyser in the Coalition's Ambitions.” October 2023.
<https://www.coara.org/coara-boost-cascade-funding/coara-boost-project/>. アクセス日 2025年12月18日.

38 アクションプランの策定状況は署名機関ごとにCoARAウェブサイトにてリストが公開される。アクションプランの策定が延期されている署名機関も散見される。 Coalition for Advancing Research Assessment. 2024. “Action Plans: Submitted & Pending to Date.”
<https://www.coara.org/action-plan/action-plans-submitted-pending-to-date/>. アクセス日 2025年12月18日.

2.2.3 ラテンアメリカ社会科学評議会（CLACSO）-ラテンアメリカ研究評価フォーラム（FOLEC）³⁹

CLACSOは1967年に立ち上げられた社会科学に関する学術ネットワークであり、現在ラテンアメリカの680の研究機関が加盟している。このネットワーク内にラテンアメリカ研究評価フォーラムが設けられており、主に人文学や社会科学分野での研究評価改革を主導している。主な目標として、以下の4点を掲げている。

- 研究評価プロセスを変革し、ラテンアメリカとカリブ海の地域のニーズに向けた研究をリードすること。
- オープンサイエンスに関する国際的議論へラテンアメリカの学術コミュニティの参加を支援すること。
- 地域の複雑さや学術的貢献多様性を評価するための議論をすること。
- ラテンアメリカとカリブ海諸国での評価に関する推奨事項や評価ポリシーを作成すること。

特に英語が主体である論文データベースを基にした学術的評価に対して、いかにスペイン語やポルトガル語での研究成果の発信が不利にならないようにできるか、といった言語的な多様性がテーマとなっている。

また、CLACSO-FOLECは2024年からCoARAステアリングボードにオブザーバ参加しており、イニシアチブ同士の連携が図られている。

2.2.4 More than our Rank（MTOR）⁴⁰

MTORは、2022年にURAの国際組織International Network of Research Management Societies（INORMS）が、高等教育機関の評価が大学ランキングへ過度に依存している状況を是正するために発足した国際ネットワークである。現在27の大学が参加しているほか、DORA、CLACSOなど14機関が協力機関として参加している。

大学が世界大学ランキングの順位向上を目的化し、研究者に過剰な論文生産を求める例や、国際共同研究を「戦略的KPI」として扱うなど、本来の学術的価値と乖離した行動を誘発する問題も指摘されている。MTORはこうした「ランキング依存の行動変容」の弊害を批判し、大学が持つ本来の公的価値を取り戻すべきだと主張し、多様な価値・貢献・ミッションを可視化することに焦点を当てている。DORAやCoARAが掲げる責任ある研究評価（RRA）の原則と親和性が高く、特に「大学ミッションの多様性」「地域・社会への価値」「教育・人材育成」「研究文化・学内環境」など、従来のランキングでは評価されにくい側面を尊重する点に大きな特徴がある。

MTORの基本理念は次の3点である。

- ミッション・多様性の尊重：大学は研究だけでなく、教育、人材育成、地域連携、社会実装、文化的貢献など多様な役割を担う。ランキングでは測れない独自のミッションを明確化し、それに基づく成果を示すことが求められる。
- 質的評価の重視：指標の数値に依存するのではなく、ケーススタディ、ナラティブ、自己評価、外部レビューなど、複数の質的手法を組み合わせる。
- 透明性・説明責任の確保：大学が自らの「強み・弱み・改善点」を明示し、どのように社会に貢献し、どのような評価方法を用いるかを公開する。過度な“順位競争”ではなく“価値の多元的提示”を促す。

39 Latin American Council of Social Sciences (CLACSO). 2024. "Latin American Forum on Research Assessment (FOLEC)." <https://www.clacso.org/folec/>. アクセス日 2025年12月18日.

40 International Network of Research Management Societies (INORMS) Research Evaluation Group. 2022. "More Than Our Rank." <https://inorms.net/more-than-our-rank/>. アクセス日 2025年12月18日.

2.3 Global Research Councilでの議論と研究評価改革の進行状況調査

Global Research Council (GRC) は、世界の主要なファンディング機関からなる組織である。現在の研究評価改革についてもワーキンググループを形成し、国際的な議論や調査活動を行っている。

2.3.1 Dimensions of Responsible Research Assessment

GRCは、2024年5月に『Dimensions of Responsible Research Assessment』⁴¹を発行した。この中で、責任ある研究評価における重要な要素を基本原則、ガバナンスと戦略、プロセスと方法論の3つの観点から以下のように提示している。

● Guiding Principles (基本原則) :

- 研究の誠実性と責任ある研究行動に対する責任ある取り組みと推進。
- オープンリサーチの奨励における責任あるアプローチ。
- 研究における公平性、多様性、包括性への責任ある取り組み。
- グローバルな課題や緊急事態が研究に与える影響への責任ある対応。

● Governance & Strategy (ガバナンスと戦略) :

- 研究評価プロセスの責任ある管理と監視：持続可能な組織的枠組みを確立する。
- 研究評価改革における責任あるアプローチ：改革は責任ある堅牢な方法論に従うべきである。
- 機関の政策および実践に対する影響力の責任ある行使：機関の自主性と学問の自由を尊重する。

● Process & Methodology (プロセスと方法論) :

- 研究評価基準の責任ある使用と普及：多様な役割、活動、成果を促進するための基準を普及する。
- 研究の貢献と業績の責任ある評価を行う。
- 社会的インパクト評価での責任あるアプローチ：社会的インパクトの多様性を認識すること。
- レビューワーとパネルの採用およびトレーニングにおける責任あるアプローチ。

2.3.2 GRC加盟国へのRRA進捗調査(2020、2025)

GRCは、加盟するファンディング機関を対象に、研究評価改革の進捗状況を把握するための調査を定期的実施している。2024年に行った調査結果は『Transforming Assessment -The 2025 Global Research Council survey of funder approaches to responsible research assessment-』⁴²として2025年5月に公表された。この調査は、Research on Research Instituteとの共同で行われ、GRCのRRAワーキンググループが担当している。調査は2024年5月28日から2025年1月6日まで実施され、GRC加盟の全機関(調査実施時点で117機関)の代表者を対象とし、50機関から回答を得ている。

41 Benamara, Abdelmajid; Fahal, Ahmed; Kowaltowski, Alicia; Trinh, Anh-Khoi; Cody, Anne; Firth, Catriona; et al. (2024). Dimensions of Responsible Research Assessment (full report and summary). figshare. Online resource. <https://doi.org/10.6084/m9.figshare.26064223.v3>

42 Alexander Rushforth et al., Transforming Assessment: The 2025 Global Research Council Survey of Funder Approaches to Responsible Research Assessment (Research on Research Institute, 2025), <https://doi.org/10.6084/m9.figshare.28856480.v5>.

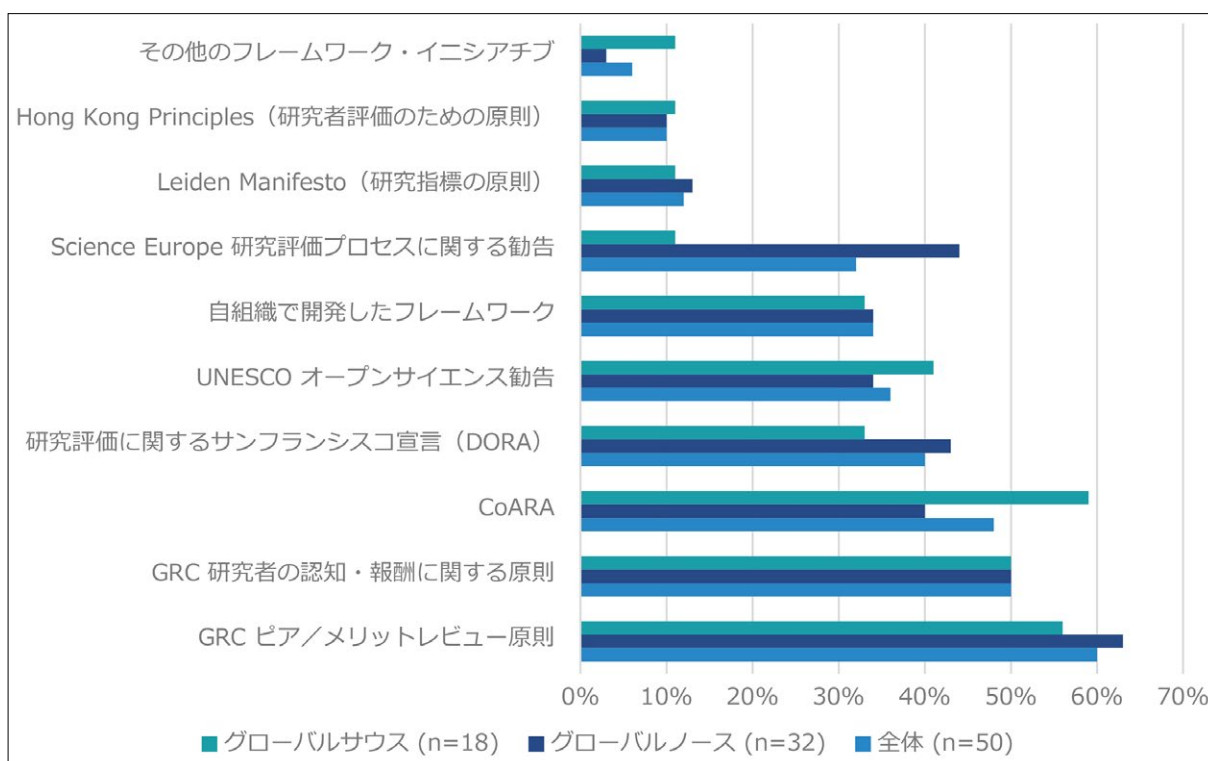


図2-2 GRC参加機関の研究評価改革に関する各種枠組みへの賛同状況⁴³

主な調査項目は、責任ある研究改革の宣言や枠組みへの参加状況、提案書レビューへのさまざまな手法の導入状況、レビューで考慮に入れている事項、ファンディングプロセスでのAIや機械学習の導入状況などである。この調査結果からは、CoARAなどの既存の枠組みには多数のFAが賛同しているものの、それが実際の評価プロセスにどの程度影響を与えているかはまだ不明な点が多いことが示唆されている。

43 Alexander Rushforth et al., Transforming Assessment: The 2025 Global Research Council Survey of Funder Approaches to Responsible Research Assessment (Research on Research Institute, 2025), <https://doi.org/10.6084/m9.figshare.28856480.v5> からCRDSが作成。

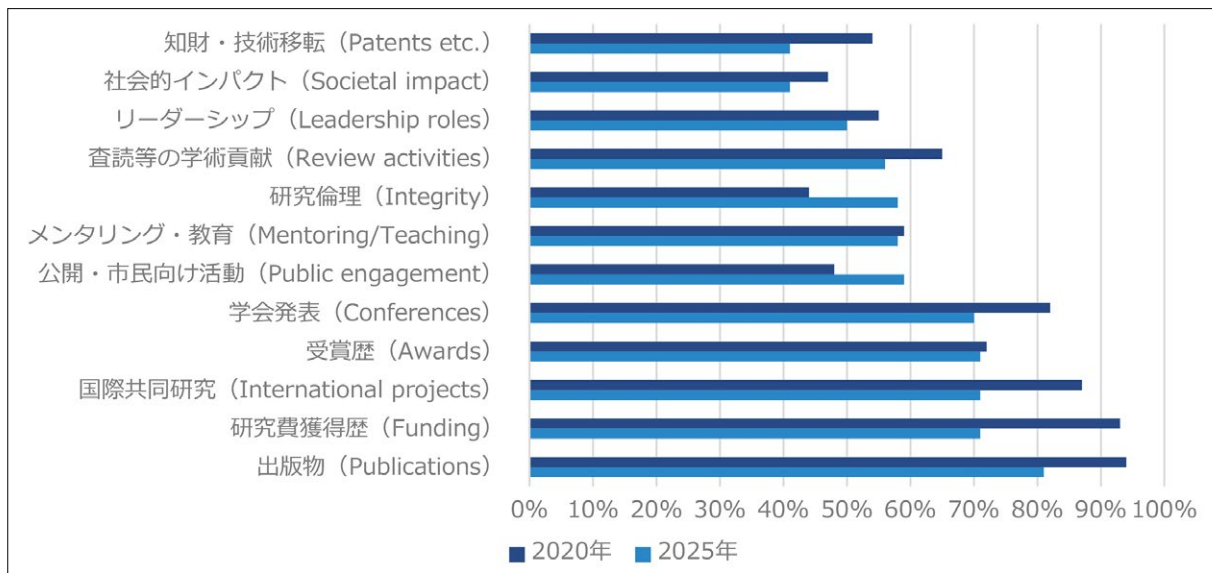


図2-3 研究課題の事前評価で考慮する要素 2020年調査と2024年調査との比較⁴⁴

研究提案の事前評価において考慮される要素は、これまで論文、特許、外部資金獲得、引用数といった計量書誌学的指標が主流であった。しかし、近年の調査では、オープンアクセス、オープンデータへの貢献、パブリックエンゲージメントなども評価基準として組み込まれつつあることが明らかになっている。これは、RRAの趣旨が研究提案のレビューに反映されつつあることを示唆する。2020年にも同種の調査が行われており、『The changing role of funders in responsible research assessment: progress, obstacles & the way ahead』⁴⁵として発行された。その結果と比較すると、2024年調査ではPublic Engagementやオープンサイエンスといった要素の利用が増加していることが確認できる。

44 Alexander Rushforth et al., Transforming Assessment: The 2025 Global Research Council Survey of Funder Approaches to Re-sponsible Research Assessment (Research on Research Institute, 2025), <https://doi.org/10.6084/m9.figshare.28856480.v5>. と⁴⁵Curry, Stephen, Sarah de Rijcke, Anna Hatch, Dorsamy (Gansen) Pillay, Inge van der Weijden, and James Wilsdon. 2020. The Changing Role of Funders in Responsible Research Assessment: Progress, Obstacles & the Way Ahead (RoRI Working Paper No. 3). November 2020. Research on Research Institute (RoRI) からCRDSが作成。

45 Curry, Stephen, Sarah de Rijcke, Anna Hatch, Dorsamy (Gansen) Pillay, Inge van der Weijden, and James Wilsdon. 2020. The Changing Role of Funders in Responsible Research Assessment: Progress, Obstacles & the Way Ahead (RoRI Working Paper No. 3). November 2020. Research on Research Institute (RoRI). https://sfedora.org/wp-content/uploads/2020/12/GRC_Responsible_Research_Assessment_RoRI_WP03_finalpdf.pdf. アクセス日2025年12月18日。

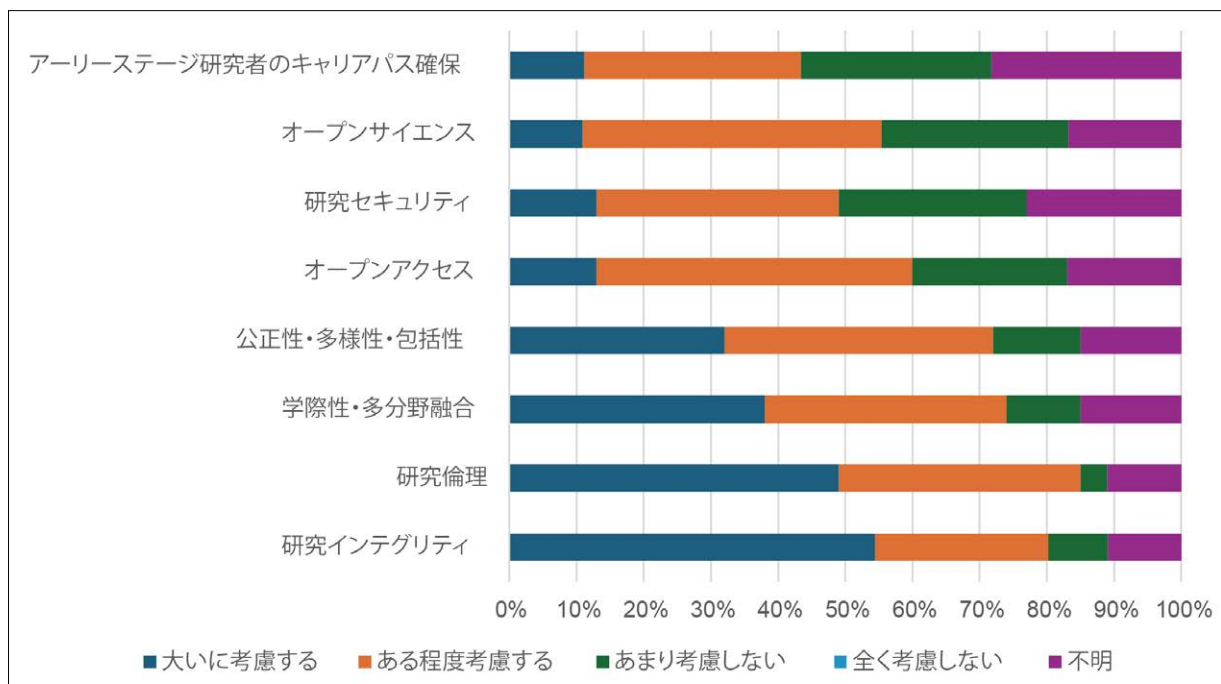


図 2-4 研究課題の評価で評価者に考慮するよう指示している要素（2024年調査）⁴⁶

また、バイアスの排除に関しては、研究の卓越性以外にレビューで考慮する点として、研究公正、倫理、学際性、公平性・多様性・包摂性（EDI）が突出して重要視されている。特に2021年から2024年の間に、EDI、オープンサイエンス、論文へのオープンアクセス、若手研究者への配慮といった要素が研究提案のレビューでの評価ウェイトを増加させており、若手研究者やマイノリティへのバイアスというRRAでの懸念点がフォローされている状況が見て取れる。

46 Alexander Rushforth et al., Transforming Assessment: The 2025 Global Research Council Survey of Funder Approaches to Re-sponsible Research Assessment (Research on Research Institute, 2025), <https://doi.org/10.6084/m9.figshare.28856480.v5>. からCRDSが作成。

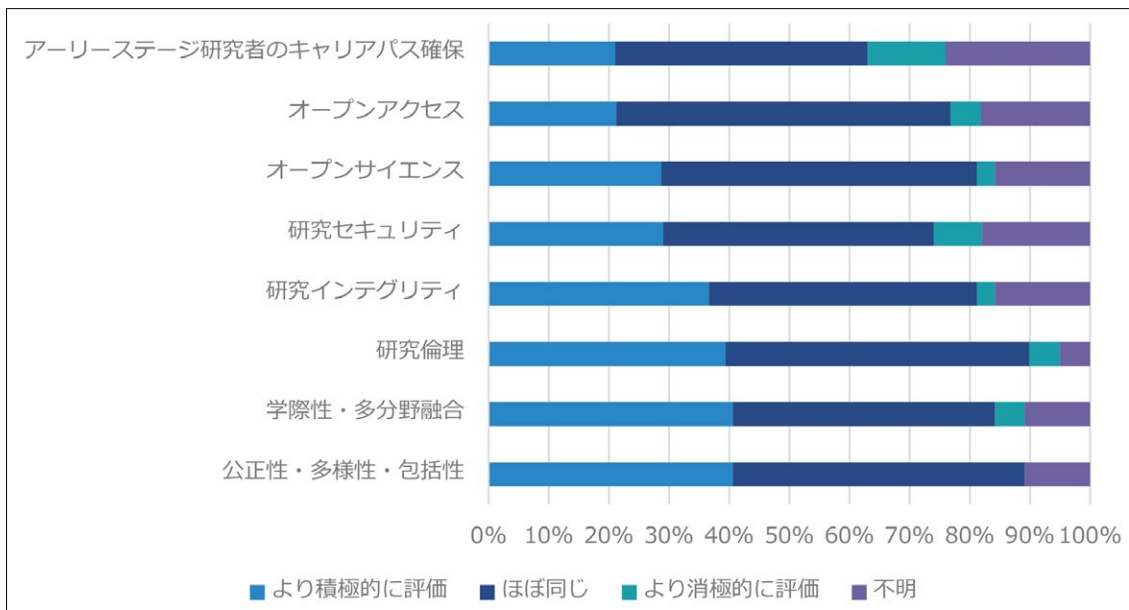


図2-5 研究提案で考慮する要素 2021年と比較した際の重みづけの変化（2024年調査）⁴⁷

コラム1

Research on Research Instituteによる研究評価システムの国際動向調査

Research on Research Institute (RoRI) は英・ユニバーシティ・カレッジ・ロンドンに拠点を置く、メタサイエンスに関する国際ネットワークである。RoRI はプロジェクトのテーマとして研究評価改革も扱っており、2025年に発行されたワーキングペーパー『A New Typology of National Research Assessment Systems: Continuity and Change in 13 Countries』⁴⁸では、アルゼンチン、オーストラリア、ブラジル、チリ、中国、コロンビア、インド、イタリア、メキシコ、オランダ、ノルウェー、ポーランド、英国の13ヶ国の研究評価システムを比較分析している。この調査は、評価目的、評価単位、資金配分とレピュテーションへの影響度合い、評価の観点から各国のシステムを分類し、研究評価のパラダイムシフトを論じている。

現在、研究評価のパラダイムとしては、専門家による分野内での評価、研究の卓越性と競争の促進が支配的であり、卓越性は計量書誌学的なインパクトと見なされ

47 Alexander Rushforth et al., Transforming Assessment: The 2025 Global Research Council Survey of Funder Approaches to Re-sponsible Research Assessment (Research on Research Institute, 2025), <https://doi.org/10.6084/m9.figshare.28856480.v5> からCRDSが作成。

48 Rushforth, Alexander; Sivertsen, Gunnar; Wilsdon, James; Bin, Adriana; Firth, Catriona; Fraser, Claire; et al. (2025). A new typology of national research assessment systems: continuity and change in 13 countries. RoRI Working Paper No.15. Research on Research Institute. Preprint. <https://doi.org/10.6084/m9.figshare.29041787.v4>

ることが多い。成果ベースの資金配分や、評価基準への政策目標の組み込みが特徴で、アカウントビリティと競争を促進するシステムとしての評価が行われている。

一方で、責任ある研究評価（RRA）の概念が台頭し、評価を通じてEDIや社会課題解決の推進を目指す動きが強まっている。RRAは評価を通じた活動の改善や学習を目的とする形成的評価を志向しており、資金配分とは必ずしも直接結びつかない評価システムと適合的であるが、その場合、被評価者の行動変容を促す力が弱いという課題も存在する。また、定量的指標の高度な利用を通じてRRAを導入するという経路も考えられている。全体として、評価対象が研究のアウトプットから、インプット、プロセス、そして社会的インパクトを含むものへと徐々に変化していることが示されている。

コラム2

研究評価改革の課題と批判

さまざまなイニシアチブやツールの開発が進む一方で、研究評価改革の課題も指摘されている。主な課題として以下の点が挙げられる。

- 評価実務とのギャップ：機関は宣言に署名はするが、採用や人事評価、研究課題評価といった実務へは反映されない。
- 資源的制約：質的評価は時間や専門知識、インフラを必要とする。
- 文化的惰性：研究者自身がJIFの高いジャーナルをキャリア構築の道具として使い、研究コミュニティも評価指標として引き続き用いることがある。
- 国や地域間の不均衡：多くの低・中所得国は研究評価改革を行うための人的、時間的資源に乏しい。
- 意図しない影響：新しいツール自体がパス依存性や特定の学術的、社会的貢献を排除する可能性がある。

これらの批判は、RRAがキャリアや研究資金といった直接的な報酬制度の変革に結びつけられない限り、単に追加的な負担になる可能性があるとの指摘もある⁴⁹。

49 Rushforth, Alexander. 2025. "Research Assessment Reform as Collective Action Problem: Contested Framings of Research System Transformation." *Minerva* (February 19): 1-21. <https://doi.org/10.1007/s11024-025-09573-3>.

コラム3

研究評価改革の視点からの研究者評価

研究者の評価に関しては、近年 Springer Nature 社が研究者個人を対象に⁵⁰、また International Science Council (ISC) が研究機関や学協会を対象とした⁵¹、評価実態の調査結果を公表している。Springer Nature 社の報告では、現行の研究者評価方法の実態と、研究者が望む理想的評価像とのギャップを浮かび上がらせている。一方、ISC 報告では、科学組織（アカデミー、研究機関、学術団体など）が評価改革にどう取り組んでいるか、また改革を推進するための“レバー（てこ）”は何か、という観点から整理されている。

Springer Nature 社の報告によれば、現在、多くの研究者が評価制度において「メトリクス（定量指標）中心」であると感じており、55%が「評価は全て、または主にメトリクスを用いている」と回答している。その中でも、出版件数・被引用数・ジャーナル・インパクト・ファクター等が突出しており、依然として研究成果の“数”や“影響度（定量）”が主要な評価軸であることが明らかである。

また、定量的指標以外の評価対象——例えば社会的・経済的インパクト、教育・査読・研究コミュニティへの貢献など——も“使われてはいる”ものの、実際には評価ウェイトが低い、あるいは使われる状況もあるが一般的ではない、といった回答が多い。

研究者自身は評価手続きそのものに対して一定の肯定的な見解を持っているとのこと。しかし同時に、「制度自体への信頼性に疑問がある」「評価プロセスが不透明、主観・バイアスを含んでいる」「社会的影響を捉えるのが難しい」「自分ではコントロールできない指標で判断されている」といった批判的な声も散見される。

さらに、理想的な評価像を尋ねたところ、現在 60% 超を占める“研究成果”の比重を、将来的には 46% 程度まで引き下げ、社会公共への貢献や研究文化・コミュニティへの貢献の比重を高めるべきだと回答している。このように、研究現場の実感として「量的指標への依存」「評価対象の限定」「制度運用・透明性の課題」「理想とのギャップ」が浮かび上がっている。

50 Farr, Una. 2025. “The State of Research Assessment: Insights from a Survey of 6,600+ Researchers.” The Link, April 8. Springer Nature. <https://www.springernature.com/gp/librarians/the-link/research-management-blogpost/research-assessment-insights-6600-researchers/27769018>. アクセス日 2025 年 12 月 18 日.

51 Global Young Academy; InterAcademy Partnership; International Science Council. 2024. Snapshots of Reform: Researcher Evaluation within Science Organizations. May 2024. <https://council.science/publications/snapshots-of-reform-researcher-evaluation-within-science-organizations/>. アクセス日 2025 年 12 月 18 日.

International Science Councilによる報告書では、改革の実践例として「ナラティブ形式のCV」「広範なインパクト評価」「柔軟な評価枠組みの試験的導入」などが紹介されているものの、制度・文化・ランキング・出版システムなど既存構造との強い相互依存性により、「改革が容易ではない」ことも指摘されている。

さらに、調査対象となった研究機関や学協会の多くが「評価改革に取り組んでいる」「改革を議論している」と答えているものの、その実践・運用が追いついていないという実情が浮かび上がっており、特に「文脈に応じた設計が必要である」との認識が繰り返されている。

このように、組織視点からは、制度改革を進めるための戦略的な“てこ”や支援構造が十分でないという点が明らかになっており、研究者の評価は研究者側の志向・理想と組織側の実践可能性・障壁が絡み合っている状況と言えるだろう。

3 | 研究の社会的インパクト評価の動向

20世紀後半以降、研究活動への公的投資に対する説明責任が高まり、研究成果が社会に与える影響を示すことが求められるようになってきた。この潮流に伴い、研究評価の対象も研究活動の直接の成果物である学術的なアウトプットから、社会的なインパクトへと拡大しつつある。

3.1 社会的インパクト評価の背景

研究の評価は、これまで主に学術的成果、つまり出版物や引用、分野での認知に焦点を当ててきた。しかし、科学は学術の世界に閉じたシステムではなく、社会や政策、経済と深く関わっているものである。1990年代後半から政府やファンディング機関は、研究の社会的インパクトをパフォーマンス評価のひとつの視点として重視するようになった⁵²。これは説明責任、公的投資による便益の証明、社会課題との関連性を求める動きである。

加えて、研究活動自体の変化も背景として指摘される。従来の好奇心駆動型ではなく社会課題に対応した形での研究活動や、学術的な新規性以上にイノベーションを志向した研究活動が行われるようになった結果、これら进行评估する手段として社会的インパクトに着目した評価が行われるようになった経緯もある。

研究における社会的インパクト評価は、こうした流れを受け実践されてきた。インパクトを単なる偶発的な副産物としてではなく、評価システム自体が明示的に測定し奨励するよう変化してきたといえるだろう。しかしながら、インパクト評価の定義、方法および結果についてはいまだ議論や模索が続いている⁵³。

また、研究の社会的インパクト評価には長い歴史的背景がある。特に基礎研究に関して、1970年前後、先進国では財政状況が悪化する一方で研究活動は巨大化・高コスト化し、「研究は社会にとって何の役に立っているのか」という疑問が強まった。こうした問題意識のもと、米国の国立科学財団（National Science Foundation、NSF）や国防総省では、基礎研究がどのようにイノベーションへ貢献してきたのかについて、過去数十年にわたる追跡調査が実施されるようになった。

1980～2000年代に入ると、研究活動の経済効果を数量的に示そうとする分析や、テクノロジー・アセスメントの取り組みが各国で始まった。英国では、サッチャー政権下の1986年に高等教育機関への資金配分を目的とした研究評価制度「Research Assessment Exercise（RAE）」が開始され、社会的インパクトの事例を収集し評価する枠組みが生まれた。後に現在のResearch Excellence Framework（REF）へと発展するこの仕組みは、社会的インパクト評価の方向性を大きく形づくることになる。

2010年代以降には、社会的インパクト評価の手法が確立・多様化し、欧州を中心にREFをモデルとした評価制度が各国で制度的に取り入れられ始めた。同時に、新たな評価モデルの研究開発も進み、研究開発プログラムや研究機関単位で研究の社会的価値をより精緻に把握しようとする試みが広がっている。

52 Donovan, C. 2011. "State of the Art in Assessing Research Impact." *Research Evaluation* 20, no. 3 (September) : 175-79. <https://doi.org/10.3152/095820211X13118583635918>.

53 Bornmann, Lutz. 2013. "What Is Societal Impact of Research and How Can It Be Assessed? A Literature Survey." *Journal of the American Society for Information Science and Technology* 64, no. 2 (February) : 217-233. <https://doi.org/10.1002/asi.22803>.

3.2 社会的インパクト評価モデル

社会的インパクトを概念化し、測定するための評価モデルはこれまでに複数開発されている。本項では代表的なモデルについて整理する。

表 3-1 社会的インパクト評価のモデル例⁵⁴

評価モデル 調査手段例	研究機関等評価		研究課題評価		
	英 Research Excellence Framework (事後評価、1986-)	蘭 Strategy Evaluation Protocol (事後評価、1994-)	Payback Framework (事前、事後評価評価、2011年に英ブルネル大学で開発)	SIAMPI (事前、事後評価、2009-11にEUのプロジェクトが開発)	RAPID outcome mapping approach (事前評価、2014年に英ODIが開発)
インタビュー					
ケーススタディ					
サーベイ調査					
ピアレビュー					
関連統計					
商業化の統計					
書誌情報					
インパクト計画					
ロジックモデル					
ワークショップ					
インパクトレポトリ					
オルトメトリクス					
インパクト追跡					

3.2.1 研究機関など評価制度に用いられる社会的インパクト評価モデルの例

REF⁵⁵ Impact Case Study (英国) :

REFの一部としての社会的インパクトの評価は、主として「インパクト・ケーススタディ」に基づいて行われる。大学は評価期間に生じた特定のインパクト事例を選定し、研究成果とインパクトの因果関係、インパクトの規模・範囲・持続性、利害関係者との関与、社会的変化の内容を詳細に記述し提出する。ケーススタディには、政策文書、法改正、技術導入、医療改善、文化活動の拡大、企業活動への影響、社会意識の変化など、インパクトが生じたことを示す多様なエビデンスの提出が求められる。提出されたケーススタディは、分野別評価委員会によるピアレビューで評価される。分野別評価委員会は研究の学術的背景とともに、社会的変化の実質性、エビデンスの妥当性、研究者・機関の貢献度、さらにはインパクトが持つ公共的価値を総合的に審査する。パネルは比較的標準化された評価基準に基づき、ケーススタディに対して1～4のスケールで採点を行い、この結果が高等教育機関への基盤的研究資金の配分に直接影響を与える。

54 David Budtz Pedersen, Jonas Følsgaard Arønvad, Rolf Hvidtfeldt, Methods for mapping the impact of social sciences and humanities—a literature review, research evaluation, volume 29, issue 1, January 2020, pages 4–21, <https://doi.org/10.1093/reseval/rvz033> から CRDS が作成。

55 REFの詳細については第4章にて記述する。

Strategy Evaluation Protocol⁵⁶ (オランダ) :

戦略評価プロトコル (Strategic Evaluation Protocol : SEP) は、オランダにおいて公的研究機関が実施する研究の「質」「関連性」「実行可能性」を体系的に評価するために設計された枠組みである。1993年に限られた分野でパイロットが行われ、1994年に正式導入されて以来、全ての公的資金を受ける研究実施機関は6年ごとにSEPの基準に沿った評価を受けることが義務づけられてきた。SEP導入の根本的な目的は、社会的インパクト評価を通じた研究機関に対する説明責任の確保に加え、組織の学習と戦略的發展を促すことにあった。

社会的インパクトの評価は、形成的評価を基調とした質的審査として実施されている。SEPは英国REFのようなインパクト・ケーススタディ方式を採用せず、研究ユニットが自ら作成する自己評価報告書と、外部専門家によるピアレビュー、現地視察を通じて、研究活動の社会的意義や社会との関係性を総合的に把握する点に特徴がある。

3.2.2 研究プログラムや研究課題の評価に用いられる評価モデルの例

Payback framework⁵⁷ (事前、事後評価) :

ペイバック・フレームワーク (Payback Framework) は、1990年代に英国のヘルスサービスの分野で開発された、研究の成果と社会的便益を体系的に把握するための評価モデルである。政策決定者に対する説明責任を強化し、研究投資がどのような形で社会に還元されているかを可視化することを目的として設計された点が特徴である。フレームワークはその後、保健医療分野に限らず、教育、社会政策、環境研究など幅広い領域で応用され、国際的に最も影響力のある社会的インパクト評価モデルの一つとして位置づけられている。

ペイバック・フレームワークの中心の特徴は、研究プロセスからアウトカム、さらには社会的便益に至るまで、研究活動を多段階モデルとして捉える点にある。具体的には、「研究の投入」「プロセス」「アウトプット」「中間的成果 (知識、能力、実践改善)」「最終的な社会的便益」といった階層で構造化され、研究がどのような経路を通して社会に影響を与えたのかを因果的に理解するための枠組みを提供する。研究成果の量的指標だけでなく、政策利用、臨床実践の改善、医療システムの効率化、健康アウトカムの改善など、広範な便益を評価対象とする点がこのモデルの特徴である。

ペイバック・フレームワークは、質的情報と定量的データを統合する評価手法として設計されている。研究事例の分析では、文献レビュー、インタビュー、ドキュメント分析、研究データの統計的把握を組み合わせ、研究成果がどのように利用され、どのような影響を生じたのかを多角的に検証する。また、政策文書、ガイドライン改訂、臨床プロトコルの変化、技術導入やビジネス化、健康指標の改善など、多様なエビデンスが統合される。評価にあたっては、研究成果が直接の社会変化に結びつかない場合でも、能力構築やネットワーク形成など中間成果を正当に評価する視点が重視される。

ペイバック・フレームワークは、知識利用の経路を明示的にモデル化した点にも特徴がある。研究から社会への影響は直線的な「成果の伝達」ではなく、政策担当者や専門職組織、産業界、患者団体など多様なステークホルダーとの相互作用を通じて形成されると捉える。このため、フレームワークは研究の価値を単一の最終成果ではなく、複数の経路と段階を通じた価値創出として評価する。これにより、短期的なアウトカムだけで

56 Verenigde Vereniging van Samenwerkende Nationale Onderzoeksuniversiteiten (VSNU); Royal Netherlands Academy of Arts and Sciences (KNAW) ; Netherlands Organisation for Scientific Research (NWO). 2020. Strategy Evaluation Protocol 2021-2027 (SEP 2021-2027). The Hague. https://storage.knaw.nl/2022-06/SEP_2021-2027.pdf. アクセス日 2025年12月18日。

57 Buxton M, Hanney S. How can payback from health services research be assessed? J Health Serv Res Policy. 1996 Jan;1 (1) :35-43. PMID: 10180843. <https://journals.sagepub.com/doi/10.1177/135581969600100107>.

なく、組織能力の強化や制度変革への寄与といった長期的影響の把握も可能になる。

政策的な意義として、ペイバック・フレームワークは研究資金配分の最適化に資するだけでなく、研究者・研究機関が自身の活動を振り返り、戦略的方向性を改善するための形成的評価ツールとしても活用されている。また、研究と政策・実践の接続を強化するための分析モデルとして国際的に採用され、研究利用の阻害要因や促進要因を可視化する枠組みとして機能している。

SIAMPI⁵⁸ (事前、事後評価) :

SIAMPI (Social Impact Assessment Methods for Research and the Production of Knowledge)は、2010年前後に欧州委員会の支援を受けて開発された、研究の社会的インパクトを評価するためのモデルである。従来の研究評価が学術的成果や最終的なアウトカムに焦点を置いてきたのに対し、SIAMPIは研究が社会に与える影響を「生産的相互作用」の観点から捉える点に特徴がある。研究と社会の関係性をプロセスとして評価し、研究が社会のアクターや制度とどのような関係性を築き、変化を生み出す基盤を形成したのかを重視するモデルである。

SIAMPIの中心概念は「価値創造のための生産的相互作用」である。これは、研究者と社会的アクター（政策担当者、企業、市民組織、専門職団体など）との間で、相互に有意義な知識交換が発生し、それが社会的変化の契機となるような関係性を指す。このアプローチでは、研究成果が社会に影響を与えたかどうかという最終的アウトカムを必ずしも前提とせず、価値創造につながる関係性や協働プロセスそのものを重要な評価対象として扱う。

SIAMPIでは、生産的相互作用を三つの類型に整理している。

- 直接的相互作用：会議、ワークショップ、共同研究、非公式な助言など、人間同士の直接的な接点を評価。
- 間接的相互作用：政策文書、報告書、新技術、データセットなど、研究成果が媒介となって生じる関係性を評価。
- 制度や経済的な相互作用：アドバイザリーボードの設置、共同プログラム、長期的パートナーシップなど、組織レベルで構築される協働体制を評価。

評価方法は、主に質的アプローチに基づいている。評価者は文書分析、インタビュー、ケーススタディを通じて、研究が社会のステークホルダーとどのように関わり、社会的価値創出の基盤を形成してきたかを分析する。重要なのは、研究と社会の関係が「双方向」であるという点であり、研究が社会を変えるだけでなく、社会的ニーズや課題が研究の方向性に影響を与えるという相互作用のプロセスが重視される。したがって、SIAMPIは研究の利用や成果の一方向的な追跡よりも、関係性の密度や構造、持続性に焦点を置く。

SIAMPIの意義は、研究の社会的インパクトを「結果」ではなく「プロセス」として評価する枠組みを提示した点にある。このアプローチは、研究が必ずしも短期的な政策変更や経済成果につながらない場合でも、社会的アクターとの協働が価値創造の基盤となり得ることを示す点で、ペイバック・フレームワークなどアウトカム中心のモデルとは異なる補完的な視点を提供する。また、学術界と社会の関係性が多様化し、学際的・超学際的研究が重要性を増す中で、研究が社会的課題にどのようにコミットしているかを把握するうえで有効な方法論として評価されている。

SIAMPIはその後、オランダや英国を中心に科学技術・イノベーション政策の議論に影響を与え、形成的評価や研究文化・エンゲージメントを重視する制度設計の流れに生かされた。生産的相互作用の概念は、社会的インパクト評価を行う際に、研究と社会の協働関係を理解するための基本的枠組みとして現在でも広く参

58 Spaapen, Jack, and Leonie van Drooge. 2011. "Introducing 'Productive Interactions' in Social Impact Assessment." The Hague: SIAMPI Project. http://www.siampi.eu/Content/Introducing_Productive_Interactions.pdf. アクセス日 2025年12月18日.

照されている。

このアプローチは、研究と社会の関係性の質を中心に据えた評価手法として、研究の社会的価値を短期的成果に還元せず、協働プロセスと知識の循環を重視する点に大きな特色がある。これは、研究の社会貢献を理解するための観点を提供し、今日の責任ある研究評価（RRA）やエンゲージメント重視の政策潮流を先導した枠組みとしても位置づけられる。

ROMA（Rapid Outcome Mapping Approach）^{59,60}（事前評価）：

Rapid Outcome Mapping Approach（ROMA）は、英国のシンクタンク Overseas Development Institute（ODI）が開発した、政策へのインパクトや社会的変化を伴う研究活動を評価するための実践的フレームワークである。ROMAは、研究が直接的に政策変更を引き起こすという直線的なモデルではなく、政策プロセスが複雑で非線形であるという前提に立ち、研究が社会や政策環境にどのような変化をもたらしたのかを、ステークホルダー行動変容に着目して評価する点に特色がある。研究評価の目的を、成果の測定というよりも、戦略立案と学習支援として位置づける形成的アプローチである。

ROMAは「成果」を政策・社会システムの中で生じる行動変容として捉える点にある。ここでいう行動変容とは、政策決定者、行政機関、専門職組織、市民社会団体、国際機関など、特定のステークホルダーが研究成果をどのように理解し、活用し、実践に反映したかという変化を指す。ROMAは最終的な政策の成果よりも、政策過程における中間的な変化を重視し、研究がどの段階でどのように影響を及ぼしたかを可視化する。

評価プロセスは以下の7段階で行われる。

- ①研究や政策課題の問題構造を明らかにする。
- ②影響を与えるべき主要なステークホルダーを特定する。
- ③ステークホルダーに期待する行動変容を設定する。
- ④行動変容に至る変化の過程をマッピングする。
- ⑤研究組織やプロジェクトが取り得る介入戦略を設計する。
- ⑥評価指標を設定する。
- ⑦モニタリングを実施する。

評価手法は質的アプローチが中心であり、文書分析、インタビュー、ワークショップ、ステークホルダーの行動観察などを組み合わせ、研究が引き起こした変化を複数の経路から検証する。評価では、研究成果そのものではなく、研究が社会的アクターにどのような「認知」「態度」「スキル」「関与」「実践」へ変化をもたらしたかが主要な論点となる。ROMAは因果関係を厳密に同定することよりも、影響が生じた「合理的説明」を可能とする点に特徴がある。

ROMAの意義は、研究の影響を短期的アウトカムや政策成果に限定せず、行為主体の行動変容を通じた多段階の変化として把握する点にある。このアプローチは、政策形成における研究の利用は段階的・累積的であり、単一の研究成果が直接政策転換をもたらすわけではないという現実に適合している。また、ROMAは研究機関やシンクタンクが自らの戦略を改善し、効果的な活動を設計するための実践ツールとしても用いら

59 Pellini, Arnaldo. 2011. "The RAPID Outcome Mapping Approach and Project Management for Policy Change." ODI Opinion. London: Overseas Development Institute. <http://www.odi.org/documents/6339/7173.pdf>. アクセス日2025年12月18日.

60 Young, John, Enrique Mendizabal, Louise Shaxson, and Julius Court. 2014. Rapid Outcome Mapping Approach (ROMA): A Guide to Policy Engagement and Influence. London: Overseas Development Institute.

れている。

研究の社会的・政策的影響を評価するうえで、アウトカムよりも行動変容、因果関係の同定よりも合理的説明、最終成果よりもプロセスの構造化を重視する評価モデルといえ、研究機関の戦略的学習を促進する形成的評価手法として用いられている。

3.3 その他の社会的インパクト評価実施事例

- オーストラリア：2018年に実施されたEngagement and InnovationはREFをモデルにしており、ケーススタディをもとにした社会的インパクト評価を実施した。
- 北欧諸国：デンマーク、ノルウェー、スウェーデンは論文指標とインパクト・ケーススタディの混合モデルを試行している。
- 欧州連合：ホライズンヨーロッパはファンディングプログラムの設計に「ミッション」と「キー・インパクト・パスウェイ」を組み込み、研究開発とSDGsとの調整を図っている。

3.4 導入する上での課題

一方で社会的インパクト評価には以下の課題⁶¹も指摘されている。

- 複雑なシステムで特定の研究を社会変化に結び付けるのは困難。
- インパクトが現れるまでには数十年かかることがあり、短期の評価サイクルと矛盾する。
- 分野によってインパクトへ至る道筋が異なる。芸術と人文学は長期的なインパクトを生み出す。
- ゲーミングと道具化：影響を示すための誇張や保守的な研究に拍車をかけるリスクがある。
- 測定可能なインパクトを強調しすぎると、好奇心に基づく研究を排除してしまう可能性がある。

一般に社会的インパクト評価は資源配分や監査ツールとして用いた際の弊害が大きく、反省や学習、対話を促進に利用するべきであるとの主張が根強い⁶²。

61 Bornmann, Lutz. 2013. "What Is Societal Impact of Research and How Can It Be Assessed? A Literature Survey." *Journal of the American Society for Information Science and Technology* 64, no. 2 (February): 217-233. <https://doi.org/10.1002/asi.22803>. では社会的インパクト評価を行う上での理論上、実務上困難となる点を指摘している。

62 Derrick, Gemma Elizabeth. 2018. *The Evaluators' Eye: Impact Assessment and Academic Peer Review*. Cham: Palgrave Macmillan. <https://doi.org/10.1007/978-3-319-63627-6>.

4 | 研究機関評価への導入状況

4.1 研究評価改革での議論の導入状況

本報告書で取り上げた国の多くは、基盤的資金（日本では国立大学に配分される運営費交付金など）と競争的研究資金による、いわゆるデュアルサポートシステムが採られている。競争的研究資金は主に第5章に記す、個々の研究プロジェクトを対象とした研究課題評価を通して支援が決定される。これに対し基盤的資金の大きな割合は前年度までの実績や学生数などの定型的な算定方法を用い配分額が決定される。また、基盤的資金の一部については業績や評価結果などを反映させた資金配分が行われる。その手法は前年度の業績を資金配分に反映させるものと、中期的期間の研究活動に対し評価を行うものに大別可能である。以下においては、大学等研究機関の中期的な期間における研究活動の評価の事例について記述する。

4.1.1 英国 REF⁶³ への導入状況

REFの評価方法はデータの支援を受けたピアレビュー（informed peer review）が中心であり、引用データを補助的に使用されることが認められている。ただし、ジャーナルインパクトファクター（JIF）や学術雑誌のランキングなど、研究の質を直接反映しない指標の使用は禁止されている。評価は34の分野ごとに専門の小委員会が担当し、これらをメインの委員会（A：医・生物系、B：数物工学、C：社会科学、D：人文・芸術）が統括する。また、2029年に実施される予定のREF 2029では、多様な役割やキャリア段階を支持する「人材・多様性パネル」、学際的研究を含む多様な研究形態の公平な扱いを促進する「研究多様性諮問パネル」が支援組織として設置される予定である。

加えて、REF 2029では、個人評価からチーム・部署・機関レベルの評価へのシフトが進み、研究成果のみを卓越性の指標とする重視を緩和し、研究文化・研究環境など、優れた研究を生み出すための条件そのものを重点的に評価する方向へと改革が進んでいる。これにより、研究とそのプロセスに対するより包括的な理解と評価を実現する制度へと進化しつつある。

4.1.2 オランダ SEPへの導入状況

2014年以降の改訂を通じて、国際的な研究評価改革における主要論点を段階的に取り込み、評価の高度化を図ってきた。これらの改革は、DORAやCoARAで主張されている国際的な研究評価改革の方向性軌を一にするものである。また、SEPの形成的評価としての性格を強化する方向で進められている。以下では、近年議論されている改革要素がSEPにどのように反映されているかを整理する。

まずは責任ある指標（responsible metrics）の導入が挙げられる。SEPは2014年改訂以降、ジャーナル・インパクト・ファクターなどの単一指標に依存しない評価を明確に打ち出した。2021年の改訂では「データの支援を受けたピアレビュー（informed peer review）」を制度の中核原則として位置づけた。これにより、研究の評価は定量指標ではなく、研究プロセス、学術的意義、社会的関連性を総合的に判断する方向へとシ

63 2029年に実施が予定されているREF 2029に関しては、<https://2029.ref.ac.uk/> . アクセス日2025年12月18日 . にて適宜情報が更新されているため、最新の状況についてはこのウェブサイトを参照されたい。

フトした。

次に、研究文化の重視が制度に組み込まれた点が重要である。2021年の改訂では学術文化が新たな政策分野として追加され、研究の誠実性、透明性、包摂性、健全な人材育成環境などが評価の対象となった。これは単に研究成果の質を評価するのではなく、優れた研究を可能にする基盤としての組織文化を評価するという、国際的な研究評価改革の方向性を反映している。

加えて、形成的評価の強化がSEPの特徴として明確に示されている点も研究評価改革の議論を反映したもののと言えるだろう。SEPは資金配分とは切り離されており、比較やランキングを目的とせず、研究ユニットの戦略的改善、組織学習、研究文化の向上を促すことを主目的としている。この点は、過度な競争を避け、研究の質と持続性を高める評価制度の構築をめざす国際動向と合致している。

4.1.3 イタリア Valutazione della Qualità della Ricerca (VQR) への反映状況

イタリアの研究評価制度であるVQR⁶⁴ (Valutazione della Qualità della Ricerca) は、高等教育機関および公的研究機関が実施した研究活動の「質」を全国的に評価するための制度として、2010年代以降整備されてきた。VQRは研究成果の卓越性を評価し、その結果を政府による資金配分に反映させることを目的とする点で、英国REFと類似する性格を持つと言える。

VQRは、イタリアの国家評価機関であるAgenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca (ANVUR) が実施主体となる。通常5年周期で評価が実施され、対象期間中に各大学・研究機関が生み出した研究成果のうち一定数を提出し、ピアレビューとビブリオメトリクスとを組み合わせ評価を行う。評価結果は、政府による研究機関への基盤的資金の配分に影響を与えるため、大学の戦略にとって重要な位置づけにある。

VQRの特徴は、提出方式と評価基準の組み合わせにある。特に自然科学・工学系では引用指標とジャーナル評価指標が補助的に使用され、人文・社会科学ではピアレビューが中心となる。こうした方式は、分野間の慣行差を考慮に入れた評価を可能にするために設計されている。

VQRが開始された当初は定量的な指標に偏った評価が行われていたが、2011–2014年を対象とした評価実施以降、研究者コミュニティとの協議を重ねて制度改善を図ってきた。最近の改革では、研究成果の量ではなく質を重視する方針が強化され、研究者単位ではなく機関単位の評価へと重点が徐々に移りつつある。これにより、過度な個人間競争を避け、大学全体としての研究環境強化を促す制度設計が進められている。

評価結果を資金配分に直接反映するという特徴から、大学における研究戦略策定に強いインセンティブを与える制度であり、イタリアの研究システムにおける重要なメカニズムとして機能している。

4.2 社会的インパクト評価の導入状況

研究機関に対する社会的インパクト評価の導入は、近年、国際的に広がりを見せている。英国REFではインパクト・ケーススタディによるアウトカム評価が制度に取り入れられているほか、オーストラリアのExcellence in Research for AustraliaやEngagement and Impact、オランダのSEPでも、各国の研究

64 National Agency for the Evaluation of Universities and Research Institutes (ANVUR). 2025. "VQR 2020-2024 – Research Evaluation Experts." <https://www.anvur.it/en/research/research-evaluation-experts/vqr-2020-2024>. アクセス日 2025年12月18日.

文化や制度環境に応じた形で社会的意義の把握が実施されており、国際的にも参照される枠組みとなっている。

また、社会的インパクト評価の対象範囲も拡張しつつある。従来は研究活動の結果として得られた「最終的なインパクト」に焦点が当てられてきたが、近年では、研究成果が社会に価値をもたらすまでの過程、すなわち「インパクトを生み出すための活動」そのものを評価対象に組み込む事例が増えている。政策担当者や産業界、市民社会との協働、研究者による知識移転活動、対話型エンゲージメントなど、研究が社会とどのように関係性を築き、社会的価値の創出に向けた基盤を形成しているかが重視されるようになりつつある。

このように、評価対象が社会的インパクトそのものから、プロセス（エンゲージメント）も含む形へと拡張している。研究が社会とどのように関わり、どのような環境を整えながら価値創造を進めているのかを理解し、評価を通じて組織の戦略と学習を促進するという視点が、各国の制度設計にも浸透しつつあると言えるだろう。

また、日本でも国立大学法人の第4期中期目標期間の評価のうち、ミッション実現戦略分を対象として社会との共創に関する評価を行うことが決定されている。加えて大学へ社会課題や地域課題への関与を求める潮流の中で、地域中核・特色ある研究大学強化促進事業（J-PEAKS）に採択された大学の中には社会的インパクト評価を試みる法人も出てきている。

以下、諸外国の研究機関評価での社会的インパクト評価の導入状況について整理する。

表 4-1 大学評価の一部に社会的インパクト評価を取り入れた事例⁶⁵

国、名称	英 Research Excellence Framework* (REF, 1986-)	豪 Excellence in Research for Australia (ERA, 2010-2022)	豪 Engagement and Impact Assessment (EI, 2018)	蘭 Strategy Evaluation Protocol (SEP, 1994-)	日 国立大学法人 第4期中期目標期間の評価「ミッション実現戦略分」の評価 (2025-?)
評価対象	<ul style="list-style-type: none"> 2029年に実施が予定のものでは評価項目毎に以下の割合。 「知識と理解への貢献」50%、「エンゲージメントとインパクト」25%、「人、文化、環境」25% 	<ul style="list-style-type: none"> 「研究の質」、「研究活動の量（研究者数、成果数など）」、「研究の応用」の観点から評価。 「研究の応用」には研究成果の商業化による収入や特許などの経済的インパクトを含む。 	<ul style="list-style-type: none"> 大学外との連携活動（エンゲージメント）、と社会的インパクトを促進した組織的なアプローチを個別に評価。 	<ul style="list-style-type: none"> 6年毎に大学が研究ユニット毎の自己評価、外部評価を実施。 評価基準：「研究の質」、「社会的関係性」、「実現能力」。 形成的評価を重視。 	<ul style="list-style-type: none"> 運営費交付金のうち「ミッション実現戦略分」の中間・期末評価で社会的インパクトを評価予定。 教育を含む。
目的	<ul style="list-style-type: none"> 大学の卓越性分析 大学の自己改革へのインセンティブ 基盤的資金の配分 	<ul style="list-style-type: none"> 研究機関の質の向上 意思決定のための参考情報収集 国内研究機関の比較 	<ul style="list-style-type: none"> ERAを社会的なインパクトの側面から補完 	<ul style="list-style-type: none"> 研究機関の組織的学習。 	<ul style="list-style-type: none"> 社会への説明責任 大学の自己改善 一部の基盤的資金の算定

*2014年まではResearch Assessment Exerciseと呼称

4.2.1 英国 REF Impact Case Study :

REFは、英国における学術研究の質を評価するために設計された専門家審査（ピアレビュー）中心の評価制度である。現在のREFは2014年に初めて実施されたが、その起源は1986年に開始された全国的研究評価制度に遡る。当初は研究選抜試験（RSE）、その後は研究評価試験（RAE）として運用されてきたが、2014年以降はREFとして体系化され、高等教育機関に対する研究資金配分と密接に結びついた制度へ発展した。

REFが社会的インパクト評価を導入した背景には、公的投資の説明責任の強化がある。研究活動が社会に

65 各国の研究評価に関する文書をもとにCRDSが作成。

もたらず価値を明確に示すことが求められる中、REF2014以降、大学は研究成果の学術的インパクトだけでなく、研究が社会に与えたアウトカムを体系的に示すことが求められるようになった。REFは、研究の質・インパクト・環境の三要素を包括的に評価するが、そのうち社会的インパクトは25%という高い重み付けをもつ。これは、研究の社会的価値を実証することが英国の科学技術・イノベーション政策において重要な位置づけにあることを示している。

社会的インパクトの評価は、主として「インパクト・ケーススタディ」に基づいて行われる。大学は評価期間に生じた特定のインパクト事例を選定し、研究成果とインパクトの因果関係、インパクトの規模・範囲・持続性、利害関係者との関与、社会的変化の内容を詳細に記述することが求められる。ケーススタディには、政策文書、法改正、技術導入、医療改善、文化活動の拡大、企業活動への影響、社会意識の変化など、研究活動の結果社会的インパクトが生じたことを示す多様な証拠の提出が求められる。

提出されたケーススタディは、分野別サブパネルによるピアレビューによって評価される。サブパネルは研究の学術的背景とともに、社会的変化の実質性、エビデンスの妥当性、研究者・機関の貢献度、さらにはインパクトが持つ公共的価値を総合的に審査する。パネルは比較的標準化された評価基準に基づき、ケーススタディに対して1～4のスケールで採点を行い、この結果が高等教育機関への基盤的研究資金（QR funding）の配分に直接影響を与える。

REFにおける社会的インパクト評価は、学術成果の引用指標とは異なる次元で研究の価値を把握する点に特徴がある。すなわち、研究の社会的波及効果を可視化し、研究者が学術界外のステークホルダーと協働し、現実の課題解決に貢献するプロセスそのものを促す制度的装置として機能している。また、REFの評価は機関間の競争を通じて、大学が自らの強みを明確にし、社会的価値の創出に向けた組織戦略を強化する動機付けにもなっている。

さらに、REF 2029では「エンゲージメントとインパクト」要素として従来のインパクト評価を再編し、研究者・大学が社会と持続的に関わる能力をより重視する方向で制度設計がすすめられている。これにより、インパクトの範囲は従来のアウトカム中心のアプローチから、協働の質やプロセス、文化的・社会的価値の長期的創出へと視野が拡大しつつある。

4.2.2 オーストラリア Excellence in Research for Australia (ERA)、Engagement and Impact (EI) :

オーストラリアでは、大学の研究活動を全国的に評価する枠組みとしてExcellence in Research for Australia (ERA) とEngagement and Impact Assessment (EI) が実施されていた⁶⁶。ERAは学術的卓越性や経済的インパクトに焦点を当てた研究評価制度であり、EIは研究の社会的インパクトと社会へのエンゲージメントの評価を目的とした枠組みである。両者は相互補完的関係にあり、研究成果の質と社会的価値の両面からオーストラリアの大学研究を把握することを意図して設計されている。

ERAは2010年に正式に導入され、自然科学、人文社会科学、医学、工学など各分野における研究の質を査定することを目的としている。評価は、出版物、引用データ、学術誌の格付け、研究者数、外部資金、特許数といったデータに基づき、大学や学科単位で行われる。ERAは学術的卓越性の把握に重点を置くことから、英国REFのように社会的インパクトを独立した評価要素とはしていない。そのため、ERAは長らく研究の「学術的インパクト」を中心とする制度として機能してきたが、研究の社会的価値を包括的に把握するためには新たな評価枠組みが必要であるという議論が高まった。これを受けて2018年に実施されたのがEIである。

66 Australian Research Council. 2021. ERA & EI Review Final Report 2020-2021. Canberra: Australian Research Council. https://www.arc.gov.au/sites/default/files/era_ei_ac_report.pdf. アクセス日2025年12月18日.

EIは、大学研究の社会的インパクトをオーストラリア全国に渡って評価する初の制度として設計された。EIは、研究成果がどのように社会、経済、文化、政策に影響を与えたかだけでなく、大学が社会的価値を創出するための活動をどのように構築しているかについても評価を行った。これは社会的インパクトそのものに加え、プロセス（エンゲージメント）を評価対象に含める点で先進的だったと言えるだろう。

EIの評価は、文字どおり「エンゲージメント（Engagement）」と「インパクト（Impact）」の二つの主要要素からなる。エンゲージメント評価では、大学が産業界、政策機関、地域社会、市民団体などどのような関係性を構築し、協働の仕組みを整えているかが審査される。共同研究の仕組み、連携プログラム、知識移転活動、地域関係の構築などが対象となり、研究活動が社会課題に応答するための組織能力を評価する要素として扱われる。

一方、インパクト評価では、大学が実施した研究が生み出した具体的な社会的成果を示すケーススタディが中心となる。ケーススタディには、政策の改善、ビジネスの効率化、医療技術の普及、環境保全、文化活動の振興など、多様な社会的変化が証拠とともに提示される。大学は研究成果とインパクトとの因果関係を明確に示す必要があり、その実質性、重要性、持続性が評価される。この点は英国REFとも共通するが、EIではプロセスとしてのエンゲージメント評価が独立した要素として位置づけられている点に特徴がある。

EIは、学術成果の評価を主とするERAと対照的に、研究の社会的価値を組織レベルで把握する制度として機能している。EIが示す評価結果は大学の実績比較や政策審議に用いられるほか、大学内部の研究戦略策定や社会連携の改善にも活用されている。また、EIで評価された事例は全国で公開され、優れたエンゲージメント実践や社会的インパクト創出の仕組みが可視化されることで、大学間の学習が促進される仕組みが整えられていた。

ERAとEIは、学術的卓越性と社会的価値創出という二つの側面から大学研究を包括的に評価する枠組みとして機能するよう設計されていた。しかし、この評価制度は被評価者である大学のコストが大きいことに加え、評価結果が基盤的資金の配分に影響を及ぼさない立て付けになっていたことで大学が評価に参加するインセンティブに乏しく、2018年に一度実施されたのみで、現在は実施されていない⁶⁷。

4.2.3 オランダ Strategy Evaluation Protocol（SEP）：

戦略評価プロトコル（SEP）は、オランダ大学協会、オランダ王立芸術科学アカデミー、オランダ科学研究機構が共同で作成した、公的研究機関が実施する研究の「質」「関連性」「実行可能性」を体系的に評価するための手順を示した文書であり、これを用いた評価制度である。1993年に限られた分野でパイロットが行われ、1994年に正式導入されて以来、全ての公的資金を受ける研究実施機関は6年ごとにSEPの基準に沿った評価を受けることが義務づけられてきた。SEP導入の根本的な目的は、社会的インパクト評価を通じた研究機関に対する説明責任の確保に加え、組織の学習と戦略的發展を促すことであった。

社会的インパクトの評価は、形成的評価を基調とした質的審査として実施されている。SEPは英国REFのようなインパクト・ケーススタディ方式を採用せず、研究ユニットが自ら作成する自己評価報告書と、外部専門家によるピアレビュー、現地視察を通じて、研究活動の社会的意義や社会との関係性を総合的に把握する点に特徴がある。

第一に、社会的インパクトに関する基礎資料となるのは、研究ユニットが提出する自己評価報告書である。この報告書では、研究の社会的関連性、政策・産業界・市民社会との関係、利害関係者との協働、社会的成果、研究が社会課題にいかに応答しているかといった項目がナラティブ形式で記述される。SEPでは固定

67 “Research assessment exercises are necessary — but we need to learn to do them better.” 2023. Nature 617 (7961) : 437. <https://doi.org/10.1038/d41586-023-01611-y>.

化された指標が設定されておらず、ユニット固有の目的や戦略に最も適した指標やエビデンスを自ら選択することが認められている点が大きな特徴である。

第二に、評価プロセスの中心には、独立した専門家パネルによるピアレビューがある。パネルは自己評価報告書の内容を検討し、社会的意義の明確性、協働プロセスの質、研究戦略との整合性、学術文化・研究誠実性との関連などを踏まえ、研究ユニットの社会的貢献を質的に判断する。SEPは目的適合性（fitness for purpose）の原則を重視しており、他機関との比較ではなく、当該ユニットが掲げる目的・戦略に照らして妥当性を評価する。

第三に、SEPでは現地視察が重要な役割を担う。専門家パネルは研究ユニットを訪問し、プレゼンテーション、研究者やリーダー層へのインタビュー、社会連携担当者へのヒアリング、必要に応じたステークホルダーへの聞き取りなどを通じて、社会との協働の実態や組織文化を確認する。現地視察は、文書だけでは把握しにくい、社会的インパクトのプロセス面や組織的基盤の有無を検証するうえで不可欠である。

最終的に、専門家パネルは評価報告書をまとめ、強み、改善点、リスク、今後の戦略に関する助言を提示する。この報告書は公開され、大学経営層の意思決定や組織改革に活用されることが多い。SEPは点数化やランキングを伴わず、資金配分とも直接連動しないため、評価の本質は診断と改善支援に置かれている。したがって、社会的インパクトの評価も、成果の量的把握よりも、社会との関係性の質、協働の持続性、組織としての姿勢や文化といった側面が重視される。

以上のように、SEPの社会的インパクト評価は、指標主導ではなくナラティブとピアレビューを基盤とした形式的評価として構築されており、研究組織が長期的な社会的価値を創出する能力を中心に捉える仕組みとなっている。

4.2.4 日本 国立大学法人の第4期中期目標期間の評価のうち、ミッション実現戦略分：

国立大学法人の第4期中期目標期間では、運営費交付金を大きく3つの枠組みに区分している。①基盤的な経費を賄う「運営分」、②各大学が担う特有のミッションを支える「教育研究組織整備・共通政策課題分」、そして③「ミッション実現戦略分」である。この「ミッション実現戦略分」は各大学がミッションを踏まえて社会変革や地域課題へ対応する戦略的な取り組みを推進するための枠組みである。

ミッション実現戦略分は、その評価結果が次期（第5期）に向けた運営費交付金の配分に反映される仕組みとなっている。文部科学省資料によれば、6年間の中期目標期間を通じて支援を継続し、4年目（中間）および6年目（期末）にそれぞれ評価を実施する旨が示されている。評価対象としては「大学のミッションと創出された社会的インパクトとの関係性」が明らかにされている点が海外の類似制度（例えば英国のREF）と異なる特長である。また、評価にあたっては大学関係者だけではなく、産業界・地域社会・公共セクターなど多様なステークホルダーの視点を取り入れるよう求められており、外部の視点による信頼性・説明責任の確保が重要視されている。

国立大学法人評価は大学の教育、研究、社会との共創といった包括的な活動を評価するものであることから、REFなど本報告書で取り上げた研究のみを対象とした事例とは異なり、教育などの社会的インパクトも含まれる。また、評価対象として大学が提出する必要があるケーススタディの数が少ないことも特徴と言える。具体的には、常勤教員数500人以下の大学では最大3件、この規模を超える大学は教員500人ごとに1件加算した数の提出が求められており、諸外国の類似制度のように大学として社会的インパクトを最大化することを目的とした部署の設置など、学内全体への浸透には限界があるとの指摘がある。加えて提出するケーススタディが少ないことにより、特に目立つ分野や成果が評価対象になりやすくなるため、比較的目立たない分野や領域が軽視されかねないとの指摘もある。また、この評価枠組みを通じて傾斜配分される運営費交付金の規模が現状不明であり、大学に社会的インパクトの最大化を志向させるだけのインセンティブとなり得るか、と

いった点も今後社会的インパクト評価を研究機関評価へ入れ込んでいく上での検討課題といえる⁶⁸。

コラム4

主要国との比較における国立大学法人評価制度

(1) 各国の大学の基盤的資金配分における評価の構造

多くの国々において大学に対する公的資金配分は、基盤的資金と競争的資金のいわゆるデュアルサポートシステムにより行われるが、ここでは日本、米国、ドイツ、英国、イタリア、カナダ、オーストラリアの各国の基盤的資金配分にかかる評価制度を概観する（日本は国立大学運営費交付金にかかる評価、米国は公立大学への州政府などの基盤的資金配分にかかる評価）。それぞれの評価時期における具体的な評価メカニズムは表4-2に示す。各国の制度は以下の二つの区分に大括りすることができる。

A. 中期的期間の研究面の業績評価の結果を資金配分に反映：日本、英国、イタリア

B. 中長期的な資金配分メカニズムの中で、一部単年度の業績評価の結果を資金配分に反映（PBF）：日本、米国、ドイツ、カナダ、オーストラリア

日本の国立大学法人評価の特徴的な点は、日本のみが上記A. B. の双方のメカニズムを導入しており、他国に比べ重層的な評価構造となっていることである。

表4-2 各国の基盤的資金配分に関わる研究機関評価⁶⁹

評価時期	日本：国立大学	米国：公立大学	ドイツ	英国	イタリア	カナダ	オーストラリア
事前評価	大学が中期計画を策定（国立大学法人評価委員会の意見を聴く）	州により異なる	算定式に基づく配分の他、一部業績協定を締結し資金配分	知識交換など一部の枠組みにおいて大学の申請に対し計算式により配分	配分額は算定式により決定	州により異なる（大学と州政府の間の合意に基づく資金配分など）	大学と政府の間でミッションを基盤とした協約が締結され、資金配分に反映
中間評価、モニタリング	PBFを導入（研究面の指標は外部資金、業績数など）	一部の州でPBFを導入（教育面の指標が中心）	州によりPBFを導入（教育面の指標が中心で、研究面の指標は限定的）	-	-	州によりPBFを導入（主に教育面の指標）	PBFを導入

68 押海 圭一. 2025.「大学における社会的インパクト評価と評価学的アプローチの接続」. 第19回情報共有会「社会的インパクト」を「評価学」の視点から考える: 39-68. 北海道大学.
<https://eprints.lib.hokudai.ac.jp/dspace/bitstream/2115/94736/1/039-068.pdf>. アクセス日 2025年12月18日.

69 各国の研究機関評価に関する文書をもとにCRDSが作成

事後評価、 追跡評価	中期目標期間 評価を実施し、 次期中期目標 期間の資金 配分に反映	-	業績協定の 結果を評価 (以後の資金 配分への反 映無し)	REFを実施し、 その結果を資 金配分に反 映	VQRを実施 し、その結果 を資金配分 に反映	-	ERAを実施 するが、その 結果の資金 配分への反 映は無し
---------------	---	---	---	----------------------------------	----------------------------------	---	--

事前評価

国立大学法人評価においては、大学が示す中期計画を文部科学大臣が認可する。その際にはあらかじめ国立大学法人評価委員会の意見を聴かなければならないとされており、この手順が事前評価と言うことも出来る。しかしこの手順は形式的な側面が大きく、本報告書で扱う研究評価の枠組みとは異なる。

他方、米国、ドイツ、カナダといった州政府などにより基盤的資金配分が行われる国々においては各州異なる制度が導入されている。大学が中期的な計画を示し、州政府などがその計画を評価し、その結果に基づき追加的資金を配分したり配分額を減額したりする例なども見られる。

また英国において基盤的資金の配分は、REFの結果を反映する部分とは別に、知識交換や基盤形成などの一部の資金については大学の申請に基づき配分する。計算式を参照するなど本報告書で扱う研究評価とは別の手法により追加的資金を配分する。

中間評価、モニタリング

国立大学法人評価においては、毎年度国立大学法人運営費交付金「成果を中心とする実績状況に基づく配分」を実施する。これは、米国、ドイツ、カナダ、オーストラリアで導入されている業績に基づく資金配分（Performance Based Funding：PBF）と類似の制度といえる。各国のPBFの指標の多くは学生の就学に関する達成度や学位の授与数など、教育面の指標で、研究面の指標は外部資金の獲得額など一部に限定されている（オーストラリアにおいては大学の戦略に応じた指標を設定）。これに対し、日本においては若手研究者比率、運営費交付金などコスト当たりTOP10%論文数（一部のグループの大学のみ対象）、常勤教員当たり研究業績数、常勤教員当たり科研費獲得額・件数、常勤教員当たり受託・共同研究など受入額と研究に関する指標が多く含まれており、その比率も1,000億円の配分対象経費のうち470億円（令和7年度/2025年度）と大きい。

PBFは本報告書で扱う海外における研究評価改革の取り組みの対象ではない。他国のPBFにおいては研究面の指標が余り見られないことがその背景にある。本報告書において提示する研究評価改革の観点からも、評価と予算の紐付けについて検討することも必要と考えられる。

事後評価、追跡評価

国立大学法人評価は、中期目標期間の業績に対し行われ、その結果が次期中期目標期間の資金配分に反映される。類似の制度は英国REF、イタリアVQRにみられる（オーストラリアERAの評価結果は資金配分に反映されない）。REFやVQRは第4章で説明した研究評価であるのに対し、国立大学法人評価は大学の教育、研究および社会との共創に関する包括的な評価であり、研究評価はその一部となっている。国立大学法人評価の中で研究に関する評価は、「（1）各中期目標の達成状況」について評価が行われる他、「（2）学部・研究科および研究組織などの現

況分析」の結果が活用される。

「(1) 各中期目標の達成状況」の評価は、各大学が中期計画で定めた評価指標を参照するが、その指標の多くは事業の実施件数、発表論文数などの定量的な指標である。このような中期的な期間における業績を評価するメカニズムは他国においてもみられる。各国のメカニズムはさまざまであるが、日本においては中期計画に数多くの事項がその指標とともに示されていることからその指標を達成したか否かに評価の関心が向きがちであるのに対し、他国においては一般に大学の包括的な戦略の下での取り組みの達成状況を評価することにより、適切な資金配分を行うことを目的とする傾向がみられる。

「(2) 学部・研究科および研究組織などの現況分析」は、独立行政法人大学改革支援・学位授与機構に設置された国立大学教育研究評価委員会の下で、大学が提出する「現況調査表」と「研究業績説明書」に基づき行う。前者においては科研費、共同研究・受託研究、論文数など定量的な指標を参照し、後者においてはピアレビューにより評価する。ピアレビューの対象は学部・研究科などを代表する優れた研究業績のみであり、REFやVQRと比較した場合、対象が限定された負担の小さい評価手法であるといえる。

これら国立大学法人の中期目標期間の業績に対する評価は他国に類例をみない制度であることから、本報告書の研究評価改革の論議の枠組みには含まれない面も多い。しかし、評価にかかる負担の軽減や定量的指標の活用などの課題の改善に向け、これら海外事例を参考として検討を加えることは有用と考えられる。

5 | 研究課題評価への導入状況

競争的研究資金などの公的資金の配分における研究課題評価の改革は主にファンディング機関の責任とも言えるが、各国のファンディングシステムやファンディング機関の独立性や審査手順はさまざまである。このため、本章ではまず主要国のファンディング機関の中核的なプログラムについて記述する。ここでいう中核的なプログラムとは、研究者が提案を行うボトムアップ型のファンディングプログラムで、主にピアレビューにより評価し採否を決定するものを指す。本章の5.2.4「主要国のFAの申請書にみるCV」のコラムでは責任ある研究評価に関連した比較検討を行った。研究評価改革の流れを受けた新たな課題評価手法やAIの活用について触れた上で、本章5.3では「主要国のFAの中核的なプログラムにおけるインパクト評価に関する指標」では社会的インパクトに関して記述した。

5.1 諸外国の公的ファンディング機関での研究課題評価（事前評価）の状況

第2章では、研究評価の国際動向について、DORA、CoARAをはじめとした研究評価改革の国際的動向や、それらの観点も含めたGlobal Research CouncilやUKRIの取り組み、さらにCoARAなどの下でのさまざまな取り組みの事例を記述した。表5-1では米国、ドイツ、英国、カナダの公的ファンディング機関での研究課題評価の状況、学術的卓越性「以外」の評価基準、研究評価改革への署名状況などをまとめた。

表5-1 各国の公的ファンディング機関の主要な研究助成プログラムでの研究課題評価の状況⁷⁰

機関名	米国 国立科学財団 (NSF)	米国 国立衛生研究所 (NIH)	ドイツ ドイツ研究振興協会 (DFG)	英国 工学物理科学研究会 (EPSRC)	カナダ 国立科学工学研究会 (NSERC)
本資料の対象の研究助成プログラム	PAPPG (申請・審査関連ガイド) で規定される研究助成プログラム	Research Project Grant (R01)	Individual Research Grants	Standard Research Grant	Discovery Grants Program
申請数・採択数	<ul style="list-style-type: none"> 申請数 32,129件 採択数 8,535件 採択率 27% (2023年) 	<ul style="list-style-type: none"> 申請数 33,139件 採択数 5,385件 採択率 16.3% (2024年) 	<ul style="list-style-type: none"> 新規・継続 14,776件 新規採択数 3,201件 個人助成 新規採択率 29.5% (2024年) 	<ul style="list-style-type: none"> EPSRC全体での新規採択率 約22% (プログラム別は非公開) (2022-2023年) 	<ul style="list-style-type: none"> 申請数 2,802件 採択数 1,635件 採択率 58% (2023年)
事前評価プロセス	概ね共通して、以下の事前評価プロセスを踏む。申請から採否の決定まで6ヶ月~9ヶ月程度を要する。 ①申請 → ②事務的確認 → ③少数のレビュアーによる審査 → ④審査委員会による順位付け → ⑤採否の決定				
学術的卓越性「以外」の評価基準	“Broader Impacts”として社会における利益も評価対象。	R01の「臨床研究枠」では、医療との接続を重視。 社会的インパクト評価	(社会的影響は評価基準として明示されず。)	社会課題や国家としての二重への貢献などを含む。	「高度人材育成」を学術的卓越性と同等の重みを持って評価。
DORAへの署名	X	X	○	○	○
CoARAへの署名	X	X	研究評価改革への対応		X
評価に関する近年の動向	NSBよりBroader ImpactからSocietal Benefitへの名称変更などを提案。	2025年1月より、簡素化のために評価の観点を5つ→3つに集約。	CoARA署名に対応して、評価実務の改革に向けたアクションプランを2024年に策定。	メタサイエンスの観点から、課題評価プロセスの合理化も視野に入れた研究助成を開始。	国内FAで協働し、課題評価プロセスの改善も視野に入れた研究助成を開始 (2024年~)。

*申請数の増加に伴うファンディング機関や評価者の負担増加、レビュアーの枯渇、バイアスの低減などは共通した課題。

70 各ファンディング機関の評価ガイドライン等をもとにCRDSが作成。

各国のファンディング機関が先行事例を参照しつつDORAやCoARAに示された理念の実現に向け取り組むことが理想とも言えるものの、これらの取り組みは小規模なプログラムや民間財団が行う研究助成資金を用いた研究助成での導入が先行している状況である。諸外国でもフォルクスワーゲン財団、ウェルカムトラストなどが採用している。一方で、幅広い学術研究を支援するミッションを持つ公的ファンディング機関での導入は、研究活動への影響の範囲も広範になるので、極端な変更で混乱が生じないよう慎重に検討する必要がある。

具体的に新たな研究課題の評価手法を検討する上で、UKRIは、『Review of Peer Review』報告書⁷¹にて、研究課題評価のさまざまな課題に対して検討されてきた介入手法を列挙し、導入事例や効果の有無に関する検討を行っている。しかし、『Review of Peer Review』はあくまでも英国における大学などへのファンディング環境下での検討であり、他国においては環境に応じた検討が必要である。

本節では諸外国（米国、ドイツ、英国、カナダ）の公的なファンディング機関の中核的なプログラムについて、さまざまな観点から比較し記述する。

5.1.1 米国国立科学財団（National Science Foundation：NSF）の状況⁷²

表 5-2 米国国立科学財団⁷³

機関の概要	<ul style="list-style-type: none"> 1950年に設立された独立の政府機関。医学および人文学分野を除く、科学・工学の全分野を支援。
ファンディングの概要	<ul style="list-style-type: none"> NSFの個々の課等が設置するプログラムと、プログラムに関わらず自由な研究計画を提出する形が存在。 申請・審査は「申請および資金配分ポリシーおよび手順のガイド（Proposal & Award Policies & Procedures Guide (PAPPG)）」に基づき実施。 申請32,129件、採択8,535件、採択率27%（2023年度）。 1件あたりの1年間の配分額は中央値約2,300万円、平均値約3,100万円。
課題評価プロセス	
評価基準	<p>(1) 知的メリット (Intellectual Merit) : 研究面での卓越性 (2) より幅広いインパクト (Broader Impacts) : 社会的利益の潜在性</p>
評価に関する近年の動向	<ul style="list-style-type: none"> DORA、CoARAともに署名していない。 2025年にNSFの審議会である国家科学審議会 (NSB) において評価基準についての検討結果を公表。2つの評価基準を維持した上で、ファンディングと社会的利益や国家としての優先事項との関係を明確にするとのこと。推奨事項として、Broader ImpactsからSocietal Benefitへの名称変更や、評価プロセスへの参加者を拡大することなどを示した。

NSFは、医学および人文学を除く多領域を支援するファンディング機関であり、1950年の創設以来、米国の基礎研究を支える中核的機関として機能してきた。NSFの研究課題評価を特徴付けるのは、

71 UK Research and Innovation. 2023. “Review of Peer Review”. July 3, 2023. London: UK Research and Innovation. A <https://www.ukri.org/wp-content/uploads/2023/07/UKRI-060723-Review-of-peer-review-Final-report-revs-v2.pdf>. アクセス日2025年12月18日. なお、検討対象となっている介入手法の名称と概要を本報告書の附録に記載している。

72 National Science Foundation (NSF). 2024. Proposal & Award Policies & Procedures Guide (PAPPG), NSF 24-1. <https://www.nsf.gov/policies/pappg/24-1>. アクセス日2025年12月18日.

73 National Science Foundation (NSF). 2024. Proposal & Award Policies & Procedures Guide (PAPPG), NSF 24-1. アクセス日2025年12月18日. もとにCRDSが作成

「Intellectual Merit (知的貢献)」と「Broader Impacts (より広範な波及効果)」を同等に重視する審査原則である。Broader Impacts は、研究成果が教育、ダイバーシティ、人材育成、地域社会、産業・政策への貢献など、学术界外の社会にもたらす影響を評価する基準として1990年代に導入され、世界的に見ても早期から社会的インパクトを制度として組み込んだ例であった。

Broader Impactは、単なる付加要素ではなく、研究者に社会的説明責任と公共価値創出を求める評価改革の一環としてNSFが積極的に提示してきたものである。近年は、PAPPG (NSF 24-1)においてBroader Impactsの具体例や評価観点がより詳細に示され、研究者の理解促進が図られている。また、NSFの政策的方針を決定する組織である国家科学審議会 (National Science Board : NSB) が2025年12月に発表した報告書『National Science Foundation Merit Review for a Changing Landscape 2025』⁷⁴では、研究活動に関する環境の変化を背景に、検討委員会からNSBとNSFへの推奨事項として以下の6項目が示された。

1. NSBはBroader ImpactsをSocietal Benefitsへ名称変更すること。
2. NSBは現在Intellectual MeritとBroader Impactsと呼ばれている2つの審査基準の構造は維持し、基準の定義を精緻化すること。
3. NSBは科学技術の進歩と研究・教育を通じた社会的利益の確保の双方を重視するよう、課題評価の方針を改善すること。
4. NSBとNSFは課題評価の方針とプロセス全体で参加機関や地域、専門分野、業種、人口動態を踏まえた多様性を拡大すること。
5. NSBは、知的貢献と社会的利益の確保の透明性と説明責任を向上させ、研究助成ポートフォリオがNSFの戦略と整合していることを明確にするため、課題評価の方針を改訂すること。
6. NSBとNSFは透明性と説明責任を向上し、継続的な改善を行うために課題評価の方針とプロセスを更新すること。

審査プロセスは数多くのプログラムにおいて異なる手順において行われるが、その多くは書面審査の後に評価委員会において審議される。さらに、科学政策上の優先課題 (AI、環境、量子、EDI など) が Broader Impacts とどのように接合するかも議論されている。プログラムオフィサーは科学分野の将来展望を踏まえて採否案の検討を行い、科学研究の公共的価値を最大化することが求められる。以上のように、NSF の評価制度は、卓越性の追求と社会的責務の両立を重視する評価体制を取っている。

74 National Science Board. 2025. *National Science Foundation Merit Review for a Changing Landscape 2025*. <https://nsf.gov-resources.nsf.gov/files/Merit-Review-Changing-Landscape-2025.pdf>. アクセス日2025年12月18日.

5.1.2 米国国立衛生研究所（National Institutes of Health：NIH）の状況⁷⁵

表 5-3 米国国立衛生研究所⁷⁶

機関の概要	<ul style="list-style-type: none"> 米国保健福祉省設置の中核的な生物医学研究機関であり、民生研究における最大のファンディング機関
ファンディングの概要	<p>「Research Project Grant R01」</p> <ul style="list-style-type: none"> プログラムの目的：幅広い基礎・臨床研究の支援。 プログラムの規模：1件あたり9,000万円（約60万ドル）程度 R01の申請数は33,139件、採択数は5,385件（採択率16.3%）（2024年）
課題評価プロセス	<p>（NIHの各研究所・センターにおいて審査を行う場合もあるが、CSRに集約することが検討されている）</p>
評価基準	<ol style="list-style-type: none"> 研究の重要性（卓越性とイノベーション）：1～9のスコアで評価 頑健性と実現可能性（アプローチ）：1～9のスコアで評価 専門性およびリソース（研究者および環境）：適切・十分か、不十分かの二択で評価
評価に関する近年の動向	<ul style="list-style-type: none"> DORA、CoARAともに署名していない。 2025年1月に簡素化されたピアレビュー枠組み（Simplified Peer Review Framework）に転換。それまでの5つの評価基準（卓越性、研究者、イノベーション、アプローチ、環境）を上記3つにまとめ、簡素化を実施。 評価基準(3)を二択で評価するには、レピュテーションバイアスを低減させる意味もある。

NIHは、生命科学・医学・公衆衛生分野の民生研究に対する最大規模のファンディング機関として、研究支援だけでなく科学政策全体に大きな影響力を持つ。NIHの審査は科学的厳密性を徹底することで知られる一方、近年は研究の再現性確保、オープンサイエンス、患者・市民参画、研究倫理、社会的価値の明確化といった観点が強調されるようになった。

NIHの評価は「スタディセクション」と呼ばれる専門的な評価委員会で行われ、Significance（重要性）、Innovation（革新性）、Approach（方法論）、Investigator（研究者）、Environment（環境）の5基準で評価されてきた。2025年には簡素化を目的に研究の重要性、頑健性と実現可能性、専門性およびリソースの3基準へまとめられている。また、NIHはこれまで研究の公平性（equity）やインクルージョン（inclusion）を促す政策を進めており、研究者の多様性に配慮した審査ガイドライン、患者・市民を巻き込んだ研究デザイン、社会的弱者を対象とした研究などを積極的に評価する姿勢を示してきた。

審査でスコアがつけられた後に、NIH全体のミッションとの整合性を保たせつつ各研究所が政策的優先性と予算状況を踏まえて採択判断を下す。こうした構造は、科学的卓越性の確保と社会的価値の最大化を両立させようとする制度設計だと言える。

75 National Institutes of Health (NIH). 2024. *Grants Process: Review*. <https://grants.nih.gov/grants-process/review>. アクセス日 2025年12月18日.

76 National Institutes of Health (NIH). 2024. *Grants Process: Review*. <https://grants.nih.gov/grants-process/review>. アクセス日 2025年12月18日. からCRDSが作成.

5.1.3 ドイツ研究振興協会（Deutsche Forschungsgemeinschaft：DFG）の状況⁷⁷

表 5-4 ドイツ研究振興協会⁷⁸

機関の概要	<ul style="list-style-type: none"> 大学、研究機関、学術団体、科学アカデミーが加盟する私法上の協会で、ドイツの学界最大の自治組織として基礎研究を支援。資金は、連邦政府、州政府、EU、個人寄付等。
ファンディングの概要	<p>「個人向け研究助成（Individual Research Grants）プログラム」</p> <ul style="list-style-type: none"> 対象：人文学、自然科学の幅広い分野 個人向け研究助成の新規・継続件数 14,776件、採択数 3,201件（個人研究助成を含む個人向け助成の新規採択率は29.5%）（2024年）
課題評価プロセス	
評価基準	<ol style="list-style-type: none"> 申請プロジェクトの質（科学的重要性・独自性、目標の明確さ） 目的と研究計画（目標の明確さ、リスクの明示、長期的ビジョン） 申請者の質（これまでの学術への貢献） 業務および研究の環境（*社会的価値は評価基準外）
評価に関する近年の動向	<ul style="list-style-type: none"> 2021年にDORA、2024年にCoARAに署名。 CoARA署名に対応し、研究評価においてジャーナルや出版物に基づく評価基準の不適切な使用を放棄する旨のアクションプランを2024年に策定。

DFGは政府から独立性を持った、自律性の高いファンディング機関である。幅広い分野へ支援を行う Individual Research Grant プログラムでは、申請書類の事務的な確認の後、外部専門家の匿名レビュー、専門委員会による分野別議論、委員会による最終的な承認という多段階構造から成る。この仕組みは、科学的卓越性の厳密な評価を実現しつつ、分野の多様性や長期的発展を考慮するコミュニティ主導型の判断を可能にする。

社会的影響に関する点では、DFGは米国 NSF での Broader Impacts のような社会的インパクトの明示は行っておらず、評価は学術的な卓越性に基づいて行われている。

研究評価改革へは積極的に取り組んでおり、DORA、CoARA 双方へ署名している。また、CoARA に対しては具体的なアクションプランの策定や国内のナショナルチャプターへの参加などに取り組んでいる。

⁷⁷ German Research Foundation (DFG). 2024. *Individual Research Grants*. <https://www.dfg.de/en/research-funding/funding-opportunities/programmes/individual/research-grants>. アクセス日 2025 年 12 月 18 日.

⁷⁸ German Research Foundation (DFG). 2024. *Individual Research Grants*. <https://www.dfg.de/en/research-funding/funding-opportunities/programmes/individual/research-grants>. アクセス日 2025 年 12 月 18 日. から CRDS が作成.

5.1.4 英国工学物理科学研究会議（Engineering and Physical Sciences Research Council：EPSRC）の状況⁷⁹

表 5-5 英国工学物理化学研究会議⁸⁰

機関の概要	・英国研究・イノベーション機構（UKRI）の中で工学、物理学分野の研究助成担当する研究会議。
ファンディングの概要	「Standard Grant」 ・分野：エネルギー、情報通信、量子技術等9つの領域での研究開発を幅広く支援。 ・2024年にスタートしたプロジェクトは162件、プログラムごとの採択率は公開されていないが、EPSRC全体での採択率は約22%（2022-2023年）
課題評価プロセス	
評価基準	申請を以下の基準から6段階で評価。 (1) 申請プロジェクトの質（新規性や将来性、手法の適切さ、社会的インパクトへ向けたアプローチの適切性など）：最優先 (2) (社会的) 重要性（世界や英国の社会的課題への貢献可能性、国家的ニーズとの合致など）：優先 (3) 申請者、グループの能力（経験、スキル、研究実行能力など） (4) 資源とマネジメント（利用可能な機器や所属機関から提供される資源とそのマネジメント）
評価に関する近年の動向	・UKRI全体で2019年にDORA、2022年にCoARAに署名。 ・評価パネル向けのガイダンスに、ジャーナルベースの定量的指標を使うべきではない旨を明記している。 ・新たな課題評価手法の構築・試行も対象とするファンディングプログラム “Metascience Research Grant” を開始（2024年～）。

EPSRC は UKRI の一部として工学・物理科学分野の研究支援を担う機関であり、英国では研究評価改革（RRA）とインパクト評価（Impact Assessment）が早期から制度化された歴史を持つ。特に、課題評価に関しては社会的インパクトへ至る経路を申請段階で明示するよう求めていたこともあり、国際的に仕組みが模倣される例も散見される。

EPSRCの審査は、①外部レビュー、②評価委員会による審査、③プログラム判断という構造を持ち、評価原則として科学的卓越性は維持されつつ、「社会・経済的インパクトの潜在性」や「産業界・公共セクターとの連携」「研究環境と文化」「若手研究者支援」といった観点が重視されている。

英国全体で研究機関に対して研究文化改善（Research Culture）を求める政策を強めており、EPSRCも応募者に対し、研究データ管理、人材育成、チームの多様性、研究の再現性確保などを明示的に記述することを求める傾向が強くなっている。

評価委員会での審査では、外部レビューの科学的評価に加え、提案課題の「社会的必要性」や「国の戦略との整合性」が議論される。技術・産業政策との接点が高い EPSRC の特性から、AI、量子、クリーンエネルギー、材料科学など、社会実装に直結する研究領域での期待値が高い。総じて EPSRC は、英国が牽引してきた研究評価改革の理念を反映し、社会的価値と研究文化の健全性を重視する審査制度への転換を進めている。

⁷⁹ EPSRC. 2023. *EPSRC panel member guidance*. <https://www.ukri.org/publications/panel-member-guidance/>. アクセス日 2025年12月18日。

⁸⁰ EPSRC. 2023. *EPSRC panel member guidance*. <https://www.ukri.org/publications/panel-member-guidance/>. アクセス日 2025年12月18日。から CRDS が作成。

5.1.5 カナダ自然科学・工学研究会議（Natural Sciences and Engineering Research Council of Canada：NSERC）の状況⁸¹

表5-6 カナダ自然科学・工学研究会議⁸²

機関の概要	・カナダの自然科学、工学分野を支援するファンディング機関
ファンディングの概要	「Discovery Grant」 ・対象：国内の自然科学、工学分野の研究者、学生 ・規模：平均支援額 約480万円/年 ・申請数2,802件、採択数1,635件（採択率58%）（2023年）
課題評価プロセス	<pre> graph LR A[研究者による研究計画] --> B[申請] B --> C[申請書類の受理] C --> D[レビュアーへの依頼] D --> E[書面審査] E --> F[合議審査] F --> G[最終的な採決の決定] G --> H[応募者への通知] H --> I[採択の通知] H --> J[不採択の通知
（評価コメント等を含む）] </pre> <p>所要期間：約5ヶ月</p>
評価基準	以下の3要素を等しく評価。それぞれ6段階。 (1) 研究者個人の卓越性 (2) 申請書の価値（新規性、分野内へのインパクト、長期的なビジョンなど） (3) 高度専門人材の育成への貢献
評価に関する近年の動向	<ul style="list-style-type: none"> ・2019年にDORAに署名、CoARAは署名していない。 ・2024年、国内の他のファンディング機関（SSHRC、CIHR）と共同で、研究ファンディングのプロセス改善も視野に入れた研究助成 “Research on Research Joint Initiative” を開始（2,000万円規模）。

NSERC は、自然科学・工学分野を支える主要ファンディング機関として、基礎研究の継続的支援と若手研究者育成を制度の中心に据えている。Discovery Grants プログラムは長期的・安定的な支援を特徴とし、研究者のキャリア発展を促す仕組みとして国際的にも評価されている。

NSERC の評価では、研究プログラムの質、研究者の実績、訓練環境の育成が主要基準である。近年はこれらに加え、EDIや地域社会との連携、先住民コミュニティとの協働、産業界との共同研究など、社会的価値の観点により明確に組み込まれるようになった。

NSERC の審査では、評価委員会向けのガイダンスにて、評価の際に定量的な指標を用いるべきではないと命じており、研究評価改革の議論と整合的な取り組みを進めている。また、パネルが研究者をカテゴリに分類する独自方式を採用しており、この仕組みは透明性が高く、過度な量的指標に依存しない点で RRA の理念と整合的である。また、2024年からは新たな過大評価手法の構築や試行も対象とする Metascience Research Grant を開始している。

81 NSERC. 2025. *Discovery Grants peer review manual 2025-26*. https://www.nserc-crsng.gc.ca/_doc/Reviewers-Examineurs/CompleteManual-ManualEvalComplect_eng.pdf. アクセス日 2025年12月18日。

82 NSERC. 2025. *Discovery Grants peer review manual 2025-26*. https://www.nserc-crsng.gc.ca/_doc/Reviewers-Examineurs/CompleteManual-ManualEvalComplect_eng.pdf. アクセス日 2025年12月18日。から CRDS が作成。

5.2 研究評価改革の議論を受けた新たな評価手法

これまで多くのファンディング機関では、研究提案に対し複数の査読者を割り当て、そのスコアをもとに専門家の審査委員会にて議論を行うことで研究助成の意思決定を行ってきた。ピアレビューに基づいたこのプロセスには、いくつかの課題があると指摘されている。まずは評価者、被評価者、ファンディング機関それぞれに大きな負担がかかる点である。次に、評価結果に人種やジェンダー、所属機関の権威といったバイアスがかかる可能性がある点も指摘されている。加えて、ピアレビューは保守的な判断を下しがちである点、研究者のみを評価者とした場合に学術面での卓越性以外の観点を導入しにくい、といった課題もある。また、研究提案自体が国際的に増加傾向にある中、適切な査読者が枯渇するという実務的な課題もある。こういった課題に対して各国のファンディング機関でさまざまな手法の導入や工夫が行われている。

5.2.1 『Review of Peer Review』 報告書

UKRIが2023年に発行した「Review of Peer Review」⁸³ 報告書では、科学技術・イノベーション分野の資金配分における査読プロセスの改善策を検証している。主な目的は、研究費の配分における査読プロセスへのさまざまな介入策を検証することであった。具体的には、申請書のセクション短縮のような軽微な変更から、部分的ランダム化や査読の完全な省略といったより大幅な変更まで、38種類の介入策（介入策のリストは附録を参照）が評価対象となった。それぞれの介入策の潜在的な用途、欠点、関連する危険性を特定し、その介入策を裏付ける証拠の強さを評価することで、十分に研究されている分野とそうでない分野を区別することを目的とし、文献レビューやインタビュー調査を行った。

報告書では、38種の介入策が対応している課題を以下の7つに分類している。

- ①時間の短縮：資金配分までの期間短縮を含む。
- ②資金提供の目的に対する関連性の向上。
- ③挑戦的な研究への投資促進。
- ④関係者全員の負担軽減：申請者、査読者、運営側の負担を減らす。
- ⑤申請量の管理：膨大な申請数を効率的に処理する。
- ⑥バイアスの低減と包摂性の促進：公平性を高め、多様な研究者を支援する。
- ⑦査読の質の向上：査読プロセスの有効性を高める。

例えば、④関係者全員の負担軽減に対しては以下の介入手法が評価対象とされている。

83 UK Research and Innovation. 2023. "Review of Peer Review". July 3, 2023. London: UK Research and Innovation. <https://www.ukri.org/wp-content/uploads/2023/07/UKRI-060723-Review-of-peer-review-Final-report-revs-v2.pdf>. アクセス日2025年12月18日。

表 5-7 負担軽減を目的として提案された評価介入手法群⁸⁴

カテゴリ	手法	概要	デメリット	導入事例や効果を検証した関連研究の蓄積(4段階)
公募デザイン	公募事前告知、Expression of Interest	事前告知した上で簡便な提案を受付(プロジェクト概要+CV程度)。簡易審査をとって本提案を提出。	助成の決定までの時間が長くなる。目標や基準に合わない、多数の関心表明(EOI)が提出される。意思決定に関する情報が制限される	★★★
公募デザイン	提案書のシンプル化	ページ数や語数の制限。プロジェクト説明のみ求めるなど。	意思決定に関する情報が制限される 提案者負担軽減につながらない場合もある	★★★
選考プロセスデザイン	2段階選考	2段階目のみフルプロポーザルを求めることでレビュアーの負担が軽減される	提案者へのフィードバックが少なくなる	★★★★
選考プロセスデザイン	AIによる自動レビュアーマッチング	-	試行事例が少なく技術的課題も不明確	★★★
選考プロセスデザイン	レビュアーの人数変更	通常1提案に2-3名レビュアーを割り付けるが、この人数を変更する。少数だとハンドリング作業が軽減	減らす場合：評価の信頼性低下、バイアスの増加 増やす場合：1人でも低評価をつけるレビュアーがいると流される	★★★
選考プロセスデザイン	レビュースコアのみによる採択決定	レビュー後のパネル審査を省略、採点の合計値で決定	採択すべき提案を逃すなどリスク コントロールの幅が狭くなる	★★
選考プロセスデザイン	Peer allocation	研究提案者が他の提案のレビューへ参加し採択を決定	提案者による悪用 提案者の負担増加	★★★
選考プロセスデザイン	オンラインパネル審査	-	パネル間での議論が深まらない可能性	★★★
公募前	研究機関向けのデマンド管理	各研究機関が提案できる件数や再提案数を制限	研究機関側での負担増加	★★★★
公募前	研究者向けのデマンド管理	不採択だった提案の再提案を一定期間禁じる	他のファンディング機関への提案が流れる場合がある議論を呼ぶことがある	★★
意思決定	部分的ランダム化	採否のボーダーライン上にある提案をランダムに採択	提案者が評価に対して持つ信頼性	★★★

この報告書が強調するのは、どの課題も単一の介入策で解決できるわけではない点である。どの介入策も上記の課題を解決できる場合がある一方で、それぞれのファンディング機関やプログラムによって異なる文脈

⁸⁴ UK Research and Innovation. 2023. "Review of Peer Review". July 3, 2023. London: UK Research and Innovation. <https://www.ukri.org/wp-content/uploads/2023/07/UKRI-060723-Review-of-peer-review-Final-report-revs-v2.pdf>. アクセス日 2025年12月18日. からCRDSが作成。

や目的に介入策を整合させることが極めて重要であるとしている。加えて、ファンディング機関が介入策を継続的に評価・監視し、その結果を他の組織と共有する相互学習の重要性も指摘している。

さまざまなレビュー手法の導入状況

2025年のGRCの調査⁸⁵によると、『Review of Peer Review』報告書に紹介されている新たな課題評価手法のうち、国際レビューやオンラインでの審査会開催といった手法は多くのファンディング機関で導入されているものの、部分的ランダム化（くじ引き）、分散型ピアレビュー、プログラムマネージャーの権限強化といった手法を導入する機関はごく一部に限られている。

この要因として、専門家のピアレビューが正当性の根拠であるという制度的な慣性や公共的資金を扱うファンディング機関の説明責任、新しい制度の導入を検討する際の国際比較や長期のデータに基づく実証研究がないことなどが挙げられている。

	主な導入目的	採用割合
国際レビュー	バイアスの軽減	80%
オンラインでの審査委員会開催	委員会運営、評価者の負担軽減	78%
グループレビュー	合意形成の促進	78%
DEIを評価基準へ明示	バイアスの軽減	80%

§

AIによるレビュー割り当て支援	作業効率化	35%
くじ引きファンディング	効率化、バイアス低減	18%
マネーの虎形式	革新的アイデアの採択	14%
申請者の匿名化	バイアス低減	13%
分散型ピアレビュー	レビューアの確保	9%

図5-1 公的ファンディング機関での各種レビュー手法の導入状況⁸⁶

⁸⁵ Alexander Rushforth et al., Transforming Assessment: The 2025 Global Research Council Survey of Funder Approaches to Responsible Research Assessment (Research on Research Institute, 2025), <https://doi.org/10.6084/m9.figshare.28856480.v5>.

⁸⁶ Alexander Rushforth et al., Transforming Assessment: The 2025 Global Research Council Survey of Funder Approaches to Re-sponsible Research Assessment (Research on Research Institute, 2025), <https://doi.org/10.6084/m9.figshare.28856480.v5>. からCRDSが作成。

5.2.2 実験的ファンディング

英国やドイツ、カナダなどの公的ファンディング機関や財団を中心として、ピアレビューのみに頼った資金配分から脱却し、部分的ランダム化を含む実験的手法を取り入れ、その結果と影響を評価する試みが行われている。従来のファンディングプロセスに内在する透明性・公平性・多様性・効率性の課題に対し、ユニバーシティ・カレッジ・ロンドンやライデン大学などの研究機関と連携し、理論的枠組みと実際の設計手順、複数の導入事例、評価方法の整理が進められてきた。本項では代表的な手法として「部分的ランダム化」の事例を紹介するが、その他に提案されている手法については附録に記載した。

「部分的ランダム化」(Partial Randomisation) は従来の評価で差がつきにくい、採否のボーダーライン上の研究提案に対してくじ引きによる選抜を導入することで、評価者の主観的判断に依存せず、透明性と公平性を担保する手法として機能することが期待されている。

欧州を中心とした試行事例：

British Academy「Leverhulme Small Research Grants」：

全申請のうち、中位の50%から60%を対象としてランダムに採択を行う。評価負担の軽減と評価の公平性担保に効果。

Volkswagen Foundation「Experiment!」：

形式審査と外部評価により「トップ」「くじ候補」「不採用」に分類。中間層をくじで抽選し、成果や多様性に良好な影響。審査員負担も軽減。

Swiss National Science Foundation (スイス)「Postdoc.Mobility」：

中位層をBayesianランクで抽出し抽選対象とした。受容性高く、他制度にも拡張可能と判断。

Austrian Science Fund (オーストリア)「1000 Ideas」：

革新的提案向けに部分的ランダム化を導入。抽選は物理からデジタルへ移行。

The Health Research Council of New Zealand (ニュージーランド)「Explorer Grants」：

初期段階研究に対し、ランダム抽選を導入。63%が支持、40%が制度拡大を希望。

実験的ファンディングを行う上でのチェックリスト：

フェーズ1：組織としての課題や目的の明確化

- 介入対象：応募者、提案、採択者、審査者など
- 目的の明確化と解決課題の特定

フェーズ2：設計と準備

- 部分的ランダム化が妥当か評価
- 併用する改革（匿名化など）検討
- 法的確認、評価者数と抽選手段の決定
- 理事会承認と説明文書の整備
- 抽選方式（物理/デジタル）、記録方法の決定
- 選抜方法の通知有無の決定

フェーズ3：評価と共有

- 評価期間・方法・指標の設計
- フィードバック収集と結果公開

5.2.3 アルゴリズムによるファンディング —AIの試行的導入—

ファンディング機関でのAI利用動向

2025年のGRCの加盟機関を対象とした調査によると、ファンディング機関が評価プロセスでAIを利用している割合は30%程度に留まっている。調査した利用目的は図5-2で示すように①審査員/パネルの割り当て・リクルート、②研究提案・研究成果の分類、③戦略分析、④ポートフォリオ分析・マネジメント、⑤研究のインパクト評価、⑥研究者の貢献・業績評価や成果の事前評価であり、いずれの項目でも「現在は利用していないが将来的な利用を検討している」との回答が4割程度を占めている。

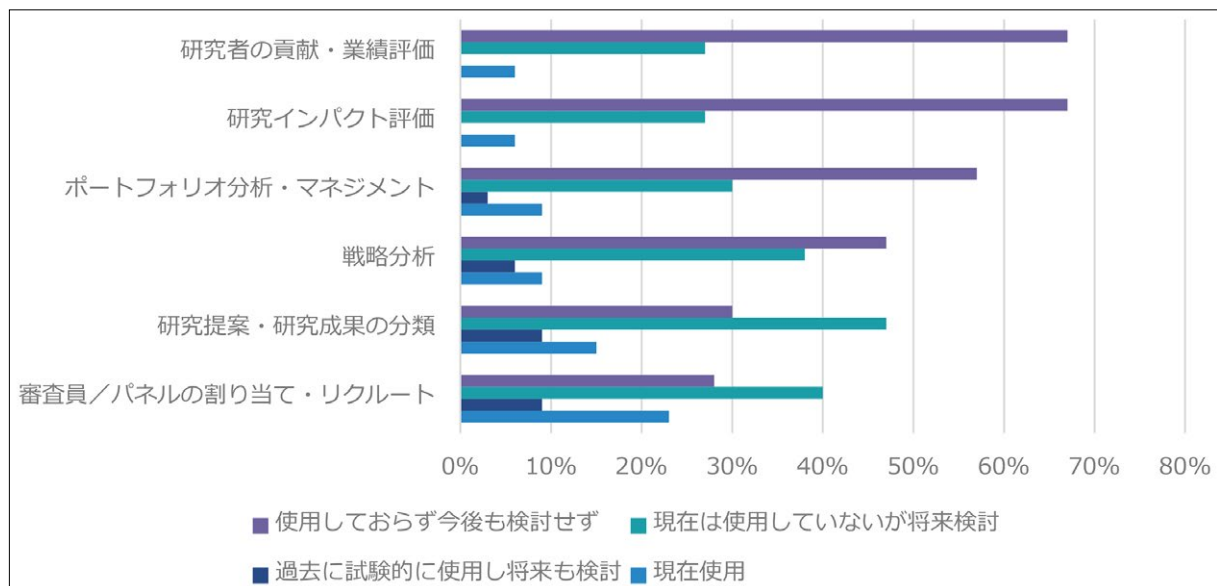


図5-2 研究課題評価プロセスでのAIや機械学習の使用状況⁸⁷

どのプロセスに導入するかによってムラがあるものの、総じてAIの利用が有益だと回答したファンディング機関が多く、特に評価プロセスの管理に有益だと回答する機関が多い。また、AIの利用について議論する際によく挙がるトピックとして、AIを利用した際の影響、適切な使用データの選択、研究者コミュニティへの影響、研究公正への影響などが挙げられている。

⁸⁷ Alexander Rushforth et al., Transforming Assessment: The 2025 Global Research Council Survey of Funder Approaches to Re-sponsible Research Assessment (Research on Research Institute, 2025), <https://doi.org/10.6084/m9.figshare.28856480.v5>. からCRDSが作成。

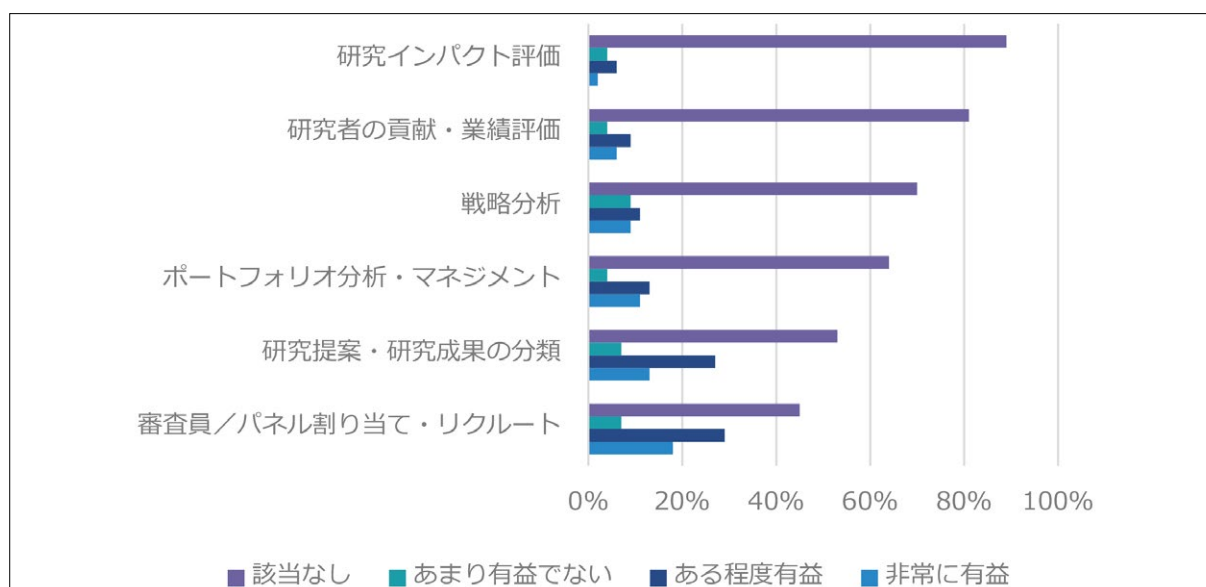


図5-3 研究課題評価プロセスの中でAIや機械学習の使用がどれだけ有益だったか?⁸⁸

AI 利用事例

2025年にRoRIがまとめた報告書『Funding by Algorithm: A Handbook for Responsible Uses of AI and Machine Learning by Research Funders』⁸⁹では、ファンディング機関によるAIと機械学習（ML）の利用事例を基に、AIの責任ある活用を支援するための実践的な知見を提示している。以下に各利用事例の目的、実装プロセス、成果、課題といった概要を示す。

スイス国立科学財団（SNSF）：査読者マッチングのための機械学習

SNSFは、査読者と研究提案の適切なマッチングを行うために、MLを活用したシステムを導入した。従来の手作業によるマッチングは時間とリソースを要していたが、MLを活用することで、申請書のキーワードや内容に基づき、数千人規模の査読者データベースから適切な査読者を自動的に抽出できるようになった。このシステムは、従来のプロセスを補完する形で導入され、査読の質を保ちながら効率性の向上に寄与している。

ラ・カイシャ財団：事前審査のためのAIによるスコアリング

スペインのラ・カイシャ財団は、膨大な数の助成申請の中から有望な提案を抽出するため、AIによる事前スコアリングシステムを導入した。AIは過去の採択・非採択データを学習し、新規の提案に対して予測スコアを算出する。このスコアは、人による審査の前段階で活用され、初期選別を効率化している。一方で、同財団はこのAIの利用を透明かつ説明可能な形で行うことを重視しており、アルゴリズムによる判断が恣意的にならないよう注意を払っている。

ノボ・ノルディスク財団：マッチファンド研究成果の特定

デンマークのノボ・ノルディスク財団は、他機関と共同出資した研究プロジェクトの成果を特定するため、

⁸⁸ Alexander Rushforth et al., “Transforming Assessment: The 2025 Global Research Council Survey of Funder Approaches to Re-sponsible Research Assessment” (Research on Research Institute, 2025), <https://doi.org/10.6084/m9.figshare.28856480.v5>. からCRDSが作成。

⁸⁹ Research on Research Institute (RoRI) et al., “Funding by Algorithm: A Handbook for Responsible Uses of AI and Machine Learning by Research Funders” (2025), <https://doi.org/10.6084/m9.figshare.29041715.v1>.

AIを用いて研究成果（論文、特許など）と資金提供記録との関連性を分析するモデルを構築した。この結果、資金提供記録に記載されていないにもかかわらず、実質的には該当プロジェクトによる成果である可能性のある研究成果も識別できた。資金の効果測定や影響分析がより網羅的かつ正確に行えるようになった。

ノルウェー研究会議（RCN）：研究の社会的影響評価

RCNは、資金提供された研究の社会的インパクトを測定するため、AIベースのテキスト分析を用いた新たなフレームワークを開発した。研究成果が政策文書、ニュース、法制度などに引用されたかどうかを分析し、定量的指標として用いている。社会的影響の可視化は研究成果の正当化や将来的な資金配分の正統性を高めるうえで重要であり、このシステムはそのための有力なツールとなっている。

英国のブロックファンディング制度：学術論文の品質評価

英国においては、大学への助成金の配分に関わる研究評価にAIを用いる試みがなされている。具体的には、査読報告や論文本文に対するAIによる品質評価スコアを導入し、人間による評価の補助とするものである。しかし、このアプローチは学界からの抵抗も強く、AIによる評価結果の正当性や透明性への懸念も根強い。結果的に、本試験運用は本格導入には至っていないが、今後の研究評価改革の一環として継続的に検討されている。

全体として、AI/MLは研究資金配分業務の多くの場面において効率性や客観性を向上させる可能性を示している。一方で、導入にはデータの質、バイアスの問題、アルゴリズムの透明性、組織内の理解と信頼の醸成など、多くの課題が伴う。特に、完全な自動化ではなく「人間が関与するシステム（human-in-the-loop）」の設計が鍵であるという認識が、全てのケーススタディで共通して見られた。

5.2.4 主要国のファンディング機関の申請書にみるCV⁹⁰

ナラティブ CV

欧州を中心に、定量的指標の偏重や研究キャリアやその中断に対する評価のバイアスへの対応として、研究歴や学術への貢献を文章として記述するナラティブ CVが有力なツールの一つとして考えられている。標準化されたフォーマットとしては、UKRIが提示したRésumé for Research and Innovation (R4RI)がある。このフォーマットは以下の4つのモジュールで構成される。

モジュール1：アイデア、ツール、方法論、知識の送付への貢献（論文の他データセット、ソフトウェア、ポリシーブリーフなど）

モジュール2：人材育成や職場への貢献（教育活動、グループ研究での専門知識の提供など）

モジュール3：研究コミュニティへの貢献（管理職、レビュアー、委員会など）

モジュール4：社会への貢献（公共部門や地域、政策立案への貢献など）

加えて、必要に応じてキャリアの中断やボランティア活動についても記載できるようになっている。

業績および研究遂行能力をどう把握するか

本報告書においては、ナラティブ CVの使用を推奨する事例が多く示されているが、DORAなどにおける主張は論文数やジャーナルのインパクトファクターといった定量的な指標の不適切な使用を批判するものであり、申請書の研究計画を評価する際に研究業績を参照することまで否定したものではありません。このことに関する研究

90 ラテン語の「Curriculum Vitae（カリキュラム・ビタエ）」（学歴・職歴を記した履歴書）の略。

評価改革の主要な論点は、申請書において支援を受ける研究者が研究計画を遂行する能力について評価するためにどのような情報が必要かという点であり、ナラティブCVはその一つの手法を提示したものとも言える。

一般にレビュアーが研究者の研究遂行能力を評価するために参照する情報には過去の業績が含まれる。主要国のファンディング機関の申請書類においても論文などを記載することが可能となっており、これらを廃しナラティブCVに置き換えることは現実的ではない。このため、本報告書においては、現行のFAの中核的なプログラムの申請書における経歴や業績の記載内容を比較することにより、どのように研究者の当該研究計画における研究遂行能力を評価しようとしているかを比較する。

表5-8は、各ファンディング機関の中核的なプログラムの公募要領などにおいて記された経歴や業績に関する情報である。それぞれのプログラムの書式（入力項目）は大きく異なるが、次ページの「主要国のFAの中核的なプログラムにおけるインパクト評価に関する指標」に記した評価指標と併せて考えると、いずれのプログラムにおいてもDORAにおいて批判されている論文数やジャーナルのインパクトファクターといった定量的な指標に基づく評価を促す内容は見られない。しかし、文献リストが存在することにより評価プロセスで過度に業績が重視される懸念は存在する。

表5-8に記した6つのファンディング機関の中核的プログラムの申請書において求められる経歴・業績の記載はさまざまであるが、以下の(1)論文などの業績を記載、(2)申請者の裁量により研究遂行能力を記載（業績の記載可）、(3)共通のCVの参照、に区分できると考えられる。

(1) 論文などの業績を記載

過去、多くのプログラムにおいて申請書に文献リストを中心とした業績欄が設けられており、NSF、DFGにおいては現在もリスト化された形の業績欄が含まれている。ただし、いずれも当該研究計画に関連したものに限定するなど、業績数などに影響されないための配慮がみられる。

(2) 申請者の裁量により研究遂行能力を記載（業績の記載可）

日本学術振興会（Japan Society for the Promotion of Science, JSPS）、EPSRCにおいて導入されている。JSPSの「2 応募者の研究遂行能力及び研究環境」は業績を網羅的に記載するものではないとし、その記載は申請者の裁量による部分が大きく、レビュアーは申請者が適切と考える業績などの情報に基づき評価を行うこととなる。他方、EPSRCは事実上論文などの業績を中心とし、ナラティブCVは補完的な位置づけとなっている。

(3) 共通のCVの参照

個々のプログラムごとに申請書に業績を記載するのではなく、ファンディング機関などが提供する共通の経歴・業績を記載したサイトを参照する方法である。NIHにおいて採用され、NSERCにおいても導入が検討されている。また、このコラムの対象とするプログラムには含まれないが、日本においてはResearchmapが定着しており、JSTをはじめとして活用されている。このメリットは、研究者が申請の都度、経歴・業績などを記載する負担が軽減されることと考えられる。他方、共通のCVには当該の研究に関連しない情報も含まれていることから、レビュアーの意識が問われる制度という事もできる。

表 5-8 文献リスト/ナラティブ CVと関連性のある情報⁹¹

	日本	米国 (NSF)	米国 (NIH)	ドイツ (DFG)	英国 (EPSRC)	カナダ (NSERC)
プログラム	科学研究費助成事業	PAPPGで規定される研究助成プログラム	R01	Individual Research Grants	Standard Research Grant	Discovery Grants (individual) program
要件・内容①	「2 応募者の研究遂行能力及び研究環境」において記載。2ページ以内に、(1) これまでの研究活動(主要な研究業績を含む)、(2) 研究環境(研究設備・資料など)を記述。	Senior/Key Personnel Documentsのうち、(i) Biographical Sketchと(v) Products。関連する業績5件以内+その他の業績5件以内を提出。	共通経歴書 (Common Form for Biographical Sketch) 導入予定	研究計画25ページ内に出版物リストを記載。	R4R1で能力・潜在性を示す。個人業績明示可だがチーム全体を示す必要。	カナダ共通CV (CCV) の導入を検討中。
要件・内容②	研究業績は網羅的でなく、実行可能性の根拠となる主要なもののみ記載。	業績は論文、会議録、プレゼン、ウェブサイト、技術成果、発明、特許、DB、AV、ソフトウェアなどを含む。	—	出版物数に制限なし。ただし関連が強いものは10件以内。	ナラティブCV利用可。	—

5.3 社会的インパクト評価の導入状況

主要国のファンディング機関はさまざまなプログラムを設置している。その中には社会的インパクトを主要な目的としたプログラムも多く含まれており、それらの評価では社会的インパクトに関する指標が含まれる。一方、本章で取り上げた主要国のファンディング機関の中核的なプログラムは、基礎から応用まで幅広い研究計画を支援するものである。つまり、必ずしも全てのプロジェクトにおいて学術面以外のインパクトを生み出すことが期待されているわけではない。このことから、各国のファンディング機関のプログラムの評価基準では、学術面でのインパクトから社会的なインパクトまで多様な成果を評価するための工夫がなされている。本節では、各ファンディング機関の中核的なプログラムにおける評価基準を抽出し、学術的価値以外のインパクトについて以下の3つの区分により整理した。：(1) 評価基準に含まれない、(2) 独立した評価基準ではないが、評価基準の中にその要素を組み込む、(3) 独立したインパクトに関する評価基準を設定

(1) 評価基準に含まれない

NIHとDFGにおいては、学術的価値以外のインパクトにする明確な評価基準はない。NIHは疾病などの予防や治療などのインパクト(社会的インパクト)を最終的な目的としている中で、表5-9で取り上げたR01はそのインパクトに向けた研究開発プロセスの初期段階に位置づけられることが、学術的価値を主眼において評価基準となっていると理解することができる。他方、DFGにおいてはIndividual Research Grantに学術

91 各ファンディング機関の中核的なプログラムの公募要領等において記された経歴や業績に関する情報を元にCRDSが作成。

的価値以外の貢献への期待を設けていない。これは、ドイツの研究ファイディング全体の構造の中で、このプログラムが最も基礎的な研究を支援する位置づけにあることなどがその背景にあると考えられる。

(2) 独立した評価基準ではないが、評価基準の中にその要素を組み込む

EPSRC、NSERCにおいては、「社会的インパクト」といった評価基準はないが、個々の評価基準の中に、社会的な面を包含するインパクトについて評価する記述が見られる（表5-9の下線部）。多様な研究プロジェクトを支援するプログラムであることから、社会的インパクトを記載することが困難ないわゆる純粋な知的探求を行う基礎研究についても不利とならないように配慮されており、個々の研究計画に対しどの程度社会的インパクトを重視するかはレビュアーの裁量に委ねられる部分が多い。

(3) 独立したインパクトに関する評価基準を設定

NSFは二つの評価基準のうち一つをBroader impactとしている（ただし、改訂を検討中）。Broader impactの記載の無い申請書は受理されないため、いわゆる純粋な知的探求を行う基礎研究などにおいて内容が伴わない形式な記述に留まることや、レビュアーがもう一つの評価基準であるIntellectual Meritばかり重視する傾向があるといった批判もあった。2026年12月に報告書が取りまとめられ、Broader ImpactをSocietal Impactと名称を変更するなど、今後の方向性が示された。

表5-9 各国のファイディング機関の課題評価基準⁹²

	米国 (NSF)	米国 (NIH)	ドイツ (DFG)	英国 (EPSRC)	カナダ (NSERC)
プログラム	PAPPGで規定される研究助成プログラム	R01	Individual Research Grants	Standard Research Grant	Discovery Grants (individual) program
評価基準①	(1) 知的メリット (Intellectual Merit): 研究面での卓越性	(1) 研究の重要性(卓越性・イノベーション): 1~9のスコア	(1) 申請プロジェクトの質 (科学的重要性・独自性、目標の明確さ)	(1) プロジェクトの質 (新規性・将来性、方法の妥当性、社会的インパクトへのアプローチ): 最優先	(1) 研究者個人の卓越性 (6段階)
評価基準②	(2) より幅広いインパクト (Broader Impacts): 社会的利益の潜在性	(2) 頑健性と実現可能性 (アプローチ): 1~9のスコア	(2) 目的と研究計画 (明確な目標、リスク、長期ビジョン)	(2) (社会的) 重要性: 英国や世界の課題への貢献度 (優先)	(2) 申請書の価値 (新規性、長期的ビジョン) (6段階)
評価基準③	—	(3) 専門性・リソース (研究者/環境): 十分 or 不十分	(3) 申請者の質 (これまでの貢献)	(3) 申請者・グループの能力 (4) 資源とマネジメント	(3) 高度専門人材の育成への貢献 (6段階)
評価基準④	—	—	(4) 研究環境 (社会的価値は基準外)	—	—

92 各ファイディング機関の評価ガイドライン等からCRDSが作成。

5.4 研究評価への新たな手法や社会的インパクト評価導入にあたっての課題、留意事項

本報告書において示した研究評価改革や社会的インパクト評価の取り組みは、時に負の影響をもたらす可能性もある。以下では、特に各ファンディング機関の中核的なプログラムへ導入した際に想定される負の影響について記述する。

(1) 申請者およびレビュアーの負担増

『Review of Peer Review』報告書の他にも多くの研究評価改革の検討において言及されているとおり、申請者およびレビュアーの負担は多くのファンディング機関で課題となっている。研究評価の改革により負担が増えれば、研究活動の低下に結びつくだけでなく、研究評価の面においても負の影響が生じる懸念がある。例えば、評価の頑健性を高めるために申請書記載内容を増やしたり、レビュアーの数を増やしたりすることは、評価者、被評価者とも大きな影響を受ける。このため、規模の大きなプログラムにおける研究評価改革は、改革することによりメリットと申請者やレビュアーの負担とのバランスを考慮することが重要である。

(2) 申請書受理から採否決定までの長期化

プログラムの目的や性格、支援の規模によっても異なるが、基本的には研究者にとって、研究計画の立案から資金の獲得までの期間は可能な限り短いことが望ましい。よって、一般に研究評価改革により資金配分までの時間が延長することは回避すべきである。各ファンディング機関の審査システムは大きく異なるが、申請書受理から採否決定通知まで6-9か月程度が一つの目安であることを踏まえると、事前申請・本申請といった多段階の審査プロセスの導入には課題が残る。他方、大規模なプロジェクトの審査は評価の観点も多くなりがちであり、審査期間を長期化させることにより評価の質を高めることに一定の意味があるだろう。また、特に新規性の高い発想を支援するプログラムでは、本申請に先立ち、発想のみを評価する手順も有効と考えられる。

(3) 研究の多様性への配慮

本章で取り上げた各国のファンディング機関の中核的なプログラムは、多様なキャリア段階の研究者による、いわゆる真理の探究から実用的な利用を目的としたものまで幅広い研究計画を支援するものである。よってその評価は多様性を包含する基準・手法を備える必要がある。申請書の記載項目の柔軟性担保と同時に、評価の基準や指標は多様な研究計画に対応できるものとすべきである。加えて、社会的・経済的効果と直接的に結びつくことが考えづらい研究計画にも、これらの観点からの記述を強いるといった画一的な仕組みにならないよう配慮が必要である。

(4) 申請者の意欲の喪失・挑戦的な申請の回避

レビュアーの負担を軽減しつつ研究評価の質を高めるため、1機関当たりの申請数を制限するといった需要管理（デマンドコントロール）を行うことは、規模が大きく業績のある研究者が申請するプログラムなどにおいては意味があると考えられる。しかしここで取り上げた中核的なプログラムの場合、需要管理は研究者の意欲を削ぎ、挑戦的な研究の申請を回避する原因となる可能性がある。

(5) 研究評価の質の向上を目的とした1件当たりの研究助成規模の拡大と採択件数の減少の負の影響

日本の場合、多くの私立大学は研究活動に充当できる資金規模が小さく、また、国立大学では法人化後の一定期間運営費交付金が削減されており、大多数の研究者にとって競争的研究資金の獲得は研究の継続に大

きな意味を持つ。その獲得機会が減少することは多くの研究者にとって死活的な問題となると考えられる。研究助成件数は各国において基盤的資金などの安定的な財源との関係も含め最適な数字が導き出されるべきである。例えば、研究助成の採択件数を減少させることは評価の質の向上には有効かも知れないが、その国の研究エコシステムにとっての是非も合わせて検討する必要がある。

(6) 申請者の研究評価への不信

研究評価改革において最も重要なことは、研究者からその改革についての信頼を得ることと言える。特に中核的プログラムでは、評価のプロセスと観点を明確にし、さまざまな取り組みが研究者の不信を招かないよう留意する必要がある。研究評価における重要な基準は「メリット」にあると言えるが、申請者が考えるメリットと申請書の評価におけるメリットとは必ずしも一致しない。例えばNIHにおいては3つの評価指標のうち、「専門性およびリソース (Expertise and Resources)」について適切か否かの2区分の評価のいずれかとし、適切でない場合にのみその理由を記すこととなっている。評価が所属機関の著名度などに影響される潜在的な評判バイアスを低減させることを目的としているが、他の指標と異なり適否の二択で評点が付されない。

(7) 被評価者の負担増

一般に社会的インパクトは研究活動からある程度時間が経ってから生じるものであり、把握や評価には被評価者に大きな人的、時間的負担がかかる。オーストラリアで実施されたEngagement and Impactは評価を受ける側の研究機関の負担が大きいことから1回のみの実施に留まった。また、英国Research Excellence Frameworkも回を重ねると大学側の負担が増えており、簡素化が大きな課題となっている。

6 | まとめ

(1) 研究評価改革の議論の導入

研究評価改革の議論は、2010年前後から顕在化した既存の評価システムへの深い問題意識に基づいて拡大してきた。特に、論文数や引用数、ジャーナルインパクトファクター（JIF）といった定量的指標の過度な利用が、研究者に対しサラム出版やトレンド追随を促すなど、研究文化に負のインセンティブを与えてきた点が広く批判されている。これらの課題認識を背景として、2013年のDORA、2015年の『メトリクスの潮流（The Metrics Tide）』など、既存の指標依存を見直す国際的フレームワークが整備されてきた。2020年前後からは、理念的な賛同表明の段階を越え、評価手法の実験、ツールの開発、制度への試験的導入といった具体的実装へと議論が進展している点に特徴がある。

近年欧州で展開されている責任ある研究評価（Responsible Research Assessment：RRA）は、定量的指標の限界を踏まえ、ピアレビューを中心とする質的評価、多様な研究貢献の評価、評価指標の透明性確保を志向している。また、RRAは頑健性、謙虚さ、透明性、多様性、省察性という五つの中核原則を明確化しており、評価が研究行動に与える影響まで踏まえた包括的な視点が示されている。2022年に発足したCoARAは、これらの理念を実現するための最も重要な国際的イニシアチブとなっており、研究助成機関・大学・研究機関が共通原則に基づき評価改革を進める枠組みを提供している。特に、ピアレビューの強化、ランキング指標の排除、多様な研究活動の評価、透明性の高い評価基準の整備など、具体的コミットメントを求めている点に特徴がある。

この議論はファンディング機関にも影響を与えており、Global Research Councilでは加盟機関を対象にRRAの導入状況を定期調査している。近年の調査結果からは、オープンサイエンス、EDI、公正性、若手研究者への配慮など、従来の定量的指標に依存しない要素が研究課題評価に取り入れられつつあることが確認されている。

日本でも、東京大学や理化学研究所、JSTなどがDORAに署名するなど、理念レベルでの受容は進んでいる。しかし、欧州のような試行的取り組みやツールの開発、組織全体での改革プロセスの設計など、実務レベルの改革はまだ萌芽的な段階にある。既存の評価制度の下で、定性的な評価を重視し、定量的指標は補助的に用いられるべきであるとの理解に留まり、研究の多様性・創造性を可視化し報奨する、といった根本課題に十分に迫っていないとの指摘も存在する。今後は政策レベルの議論と研究機関レベルの改革を結びつけ、制度的・文化的な変革を促す取り組みが求められる。

(2) 社会的インパクト評価の導入

社会的インパクト評価の導入は、1970年代以降の財政制約下での研究投資の効果説明要求に端を発し、近年では大学など研究機関の社会的役割の拡大とともに再び重要性を増している。特に、研究機関に対して教育・研究に加え、社会課題・地域課題への貢献が強く求められるようになったことから、研究成果が社会にもたらした価値を体系的に捉える必要性が高まっている。こうした動きの中心的事例として、英国のResearch Excellence Framework（REF）が挙げられる。REFでは、研究の卓越性に加えて「インパクト事例（ケーススタディ）」が評価対象となり、研究がもたらした政策、産業、文化、地域社会などへの影響が体系的に評価される。この枠組みは国際的に大きな影響を与え、オーストラリア、オランダ、北欧諸国などでも社会的インパクト評価が制度的に導入される契機となった。オランダのStrategy Evaluation Protocol（SEP）は、社会的インパクト評価を研究ユニットの「学習と改善」に結びつけた代表例であり、研究機関のマネジメント改善に有効なフレームワークとして注目されている。

また、社会的インパクト評価は研究課題評価にも導入が進んでいる。ファンディング機関では、研究助成の

事前評価において社会的課題との関連や研究成果の潜在的インパクトを評価基準に組み込む例が増えている。さらに、助成終了後に追跡評価を行い、研究成果がどの程度社会的価値を創出したかを検証する取り組みもみられる。

一方、社会的インパクト評価は評価目的が国によって大きく異なる。英国のREFのように資金配分に直接結びつける国もあれば、学習や改善を目的とし、資金配分とは切り離す国もある。また、研究分野によって成果の現れ方や評価可能性が大きく異なるため、インパクト評価には評価負担、評価者専門性、データ収集コストなどの問題も指摘されている。さらに、研究成果が社会に浸透するには長期の時間軸が必要であり、短期的な評価制度との整合性をどう確保するかも重要な課題である。

日本でも導入は進みつつある。国立大学法人評価の「ミッション実現戦略分」では社会的インパクト評価が本格導入され、J-PEAKSや国際卓越研究大学事業でも社会的インパクト評価が求められている。しかし、国内では評価モデル・事例の蓄積がまだ少なく、評価業務にかかる人的・事務的負担も十分に議論されていない状況である。制度導入にあたっては、評価目的の明確化、評価手法の適切な選定、データ収集コストの抑制、研究者コミュニティとの合意形成が不可欠である。

社会的インパクト評価の導入は研究評価改革と密接に関連しており、従来の学術中心の評価枠組みを超えて研究の社会的価値を可視化する試みとして、今後の評価制度設計のテーマとなるだろう。

(3) 今後の課題

本報告書では研究機関と研究課題に関する評価に焦点を置いたが、他にも科学技術・イノベーション政策や施策の評価、研究機関での研究者評価、学術出版や研究者コミュニティでの学術論文への評価など、研究評価に関する重要なトピックは枚挙に暇が無い。また、研究評価改革に関しても近年オープンサイエンスと深く関わりながら推進されているが、本報告書に記載するには限界があった。

加えて、本報告書で取り上げた取り組みの多くはここ数年で具体化されたものも多く、効果の検証や取り組みを拡大させていく上での課題が顕在化する可能性が高い。そのため、今後の展開を注視していく必要があるだろう。

附録 UKRI 『Review of Peer Review』⁹³ 報告書に基づくピアレビューへの介入手法一覧

本報告書では実験的ファンディングの代表的な事例として「部分的ランダム化」を紹介したが、他にもさまざまな新規の研究課題評価手法が考案、検討されており、一部は国際的に試験的導入が行われている。そこで附録では、新規性の高い研究課題評価手法を一定程度網羅的に扱った、UKRIの『Review of Peer Review』報告書に基づいた手法の一覧を掲載する。各手法の詳細や導入機関などは報告書原本を参照されたい。

1. 事前（Pre-call）介入

評価基準の明確化：従来の基準に加え、新たな評価基準（例：必須/推奨）を導入する。透明性・一貫性を高め、特定の資金スキーム目標への関連性を最適化し、非学術的インパクトを持つプロジェクト支援に有効。レビュー担当者が基準に従わないリスクや複雑化の懸念もある。

需要管理：個人の制限（1）：研究者が同時に主担当者として関わるプロジェクトや申請数を制限する。申請数を減らし、少数の研究者に資金が集中するのを防ぐ。他のファンディング機関への負担転嫁や、限定的な効果に留まる可能性あり。

需要管理：個人の制限（2）：不採択だった申請者に対し、翌年の申請を一定期間（例：1年間）禁止する「タイムアウト」期間を設ける。申請数を抑制し、資金提供者と審査員の負担を軽減する。他のファンディング機関への申請流出や、申請者の意欲を損なうリスクがある。

需要管理：機関の制限：単一機関からの申請数や再提出数を制限する。資金提供者や審査側の負担軽減に寄与し、申請成功率を高める。選定負担が機関側に移り、機関の審査プロセスによっては新たなバイアスが生じる可能性もある。

過小評価グループへの支援：ファンディング機関のポートフォリオで過小評価されているグループに対し、申請奨励と支援を行う。申請者や採択者の多様性を高めることを目指す。効果が出るまでに時間がかかる可能性や、追加の事務負担が生じる懸念がある。

2. 申請書設計とパラメータ（Application design & parameters）介入

申請者の行動を促す設計：申請書フォームやプロセスを設計し、申請者間のポジティブな行動（例：コンソーシアム構築や連携）を奨励する。多様性を高め、特定の申請者グループによるゲートキーピング（支配）を抑制する。既知のデメリットは報告されていないが、研究は限定的。

関心表明/事前提案：簡易版の申請書（プロジェクト概要や履歴書など）を提出させ、選別後に一部の研究者のみに本申請を求める。審査負担や申請負担を軽減し、本申請の質と関連性を高める。資金配分までの時

93 UK Research and Innovation. 2023. "Review of Peer Review". July 3, 2023. London: UK Research and Innovation. <https://www.ukri.org/wp-content/uploads/2023/07/UKRI-060723-Review-of-peer-review-Final-report-revs-v2.pdf>. アクセス日2025年12月18日.

間が長くなる可能性や、対象外の関心表明が流入するリスクがある。

申請書の長さ短縮/セクション削減：申請書フォームのページ数や文字数制限を設ける。申請者と審査員の負担を軽減し、効率性を高める。情報が不足し審査判断を難しくする可能性や、申請者の負担が必ずしも減らない場合がある。

3. 評価プロセス設計（Process design）介入

「サンドピット」/マッチングイベント：申請候補者をイベントに招き、アイデアを議論し、共同研究チームを形成させる。学際的・分野横断的な研究、新たな連携、変革的な研究を促進する。アクセス性や公平性、多様性、包摂性（EDI）に関する課題が生じる可能性があり、遠隔イベントで一部解決可能。

2段階申請プロセス：査読やパネル審査を2段階で行い、最初の段階で一部の申請を選別する。審査員やプログラム担当者の負担を軽減し、第2段階の提案の質と関連性を高める。第1段階の申請者へのフィードバックが少なくなるという懸念がある。

申請者匿名化：審査員やパネルメンバーが申請者の身元を知ることができないようにする。バイアスを減らし、革新的・変革的なアイデアに資金提供を促す。プロジェクトの実現可能性を判断する能力が制限される可能性がある。

AIによるレビュアー割り当て支援：AIや自然言語処理を用いて、申請書への審査員割り当てを支援する。審査員割り当ての効率を高め、負担を軽減し、審査員と申請書のより適切なマッチングを実現する。十分にテストされていないアルゴリズムに問題があるリスクがある。

「マネーの虎」形式のプレゼンテーション：申請者がパネルの前で提案をプレゼンし、パネルが質疑応答を行う。他の形式の証拠（書面申請や外部審査）は使用しない。ステークホルダーの関与を高め、斬新で変革的なアイデアへの資金提供を促す。プレゼンテーション能力の高い申請者に有利に働く可能性や、アクセス上の問題が生じる懸念がある。

外部レビューのみ（パネルなし）：外部のレビュアーのみが提案を評価し、その評価スコアを単純に合計して最終スコアとする。パネルによる保守的な決定を減らし、負担とコストを削減し、専門知識と申請のマッチングを改善する。リスク管理の層が減るため、リスクコントロールの層が減る懸念や、透明性が低下する可能性がある。

グループレビュー：複数の提案について同じ審査員がコメントする。合意形成を促進し、審査の多様性を高める。グループバイアスのリスクがある。

審査員数の変更：申請書に対する外部審査員の数を標準の2～3名から、1名に減らすか、大幅に増やす。審査員数を増やす場合、審査の頑健性/信頼性を向上させる。減らす場合、時間/負担/コストを節約する。審査員数を増やすと、単一の悪い審査が申請を却下するリスクを高め、労働集約的になる。減らすと、頑健性が低下し、バイアスが増加する可能性がある。

面接：申請代表者（またはチームメンバー）がプレゼンテーションを行い、パネルメンバーや審査員、ファンディング機関の担当者から申請内容について質問を受ける。審査の質を高め、精査を深め、批判への回答機

会を提供する。資源集約的で、特定のグループにバイアスや不利益を与える可能性がある。

審査のモデレーション：審査をファンディング機関の職員が内部で処理し、十分な質がある場合にのみ外部パネルに提出する。審査の一貫性/質を確保する。管理者の時間消費が多く、管理者が十分な専門知識を持たない場合に問題が生じる可能性がある。

モデレーションパネル：評価パネルが外部審査に加え、自身の専門知識を用いて提案を評価する。モデレーションパネルは自身の専門知識を使わず、審査スコアに情報を与えるために外部審査のみを用いる。審査の一貫性を確保し、専門知識と審査の頑健性を高める。有効性は不明。

パネルのみ（書面/外部審査なし）：専門家パネルのみが提案を評価する。意思決定の迅速化、効率性向上、審査の一貫性確保、戦略的視点の導入を目的とする。パネルに必要な専門知識を網羅することが困難な場合や、追加の審査が必要になる可能性、潜在的なバイアスがある。

分散型レビュー：申請者自身が他の申請者の提案を審査し、資金配分を決定する。管理負担を軽減し、審査員選定のプレッシャーを減らす。悪用や不正行為の可能性があり、申請者への負担が増える場合がある。

プログラムマネージャーの裁量：申請が直接プログラム/スキームマネージャーに送られ、資金提供を推奨、あるいは決定する。ピアレビューやパネル審査を完全に省略する。資金提供までの時間を短縮し、全体的な負担を軽減し、緊急事態に対応し、査読で不採択となるリスクが高いハイリスク・ハイリターンプロジェクトに資金提供する。プログラムマネージャー自身がリスクを回避する傾向がある場合、活用が不足する可能性や、透明性の欠如、特定の「勝者」に偏る「勝者のゲーム」になる可能性が懸念される。

常設パネル vs. ポートフォリオパネル：常設パネルは毎年同じメンバーで構成される一方、ポートフォリオパネルは受け取った提案に基づいて構成される。常設パネルは一貫性を確保し、長期的な学習や学際的な対話の場となる。常設パネルは制度的なバイアスにつながる可能性もある。

国際審査員の活用：ファンディング機関の本国以外の国の審査員にクォータを設ける。必要な専門知識を確保し、ギャップを埋め、利益相反を避ける。特定の国に特化した専門知識をもたらす。一部のパネルメンバーに追加のガイダンスや研修が必要になる可能性がある。

指標の活用：意思決定の根拠として、指標や計量書誌学的情報を用いる。申請者に関する追加情報を提供し、審査の頑健性を高める。極めて議論を呼ぶ可能性のある介入策であり、卓越性の不適切な指標、バイアスや悪用の可能性がある。また、DORAをはじめとした責任ある研究評価の趣旨に反する可能性がある。

非学術的審査員（産業界、政策立案者、患者など）の活用：非学術的審査員にクォータを設ける。社会的な関連性やインパクトを高める。基礎研究の概念を希薄化させる可能性があり、そのような文脈には推奨されない。

バーチャルパネル：パネル会議を対面ではなくオンラインで開催する。コストと炭素排出量を削減し、より多くの国際的なパネルメンバーの参加を可能にし、参加への障壁を取り除く。議論の頑健性や詳細さが低下する可能性は不明。

4. 意思決定（Decision-making）への介入

ワイルドカード：各パネルメンバー（または他の意思決定者）が、パネルのランキングや他の意思決定プロセスに関わらず、1つの提案に資金提供を保証できる「ゴールデンチケット」または「ジョーカー」のような権限を持つ。よりリスクが高く、変革的なアイデアに資金提供し、パネルでの議論時間を短縮する。利益相反が厳密に監視されない場合、悪用の可能性があり、匿名審査が必須となる。

部分的ランダム化：採択される提案をランダムに選ぶ。ほとんどの場合、部分的であり、例えばスコアリングによってバンド分けし、採択のボーダーラインにある提案のみをランダム化する。バイアスを排除し、パネルの負担を軽減する。申請者の評判への影響が懸念される。

スコアリングメカニズム：スコアの調整、コンセンサスvs投票、重み付けなどを含む。採択されるプロジェクトの目的への関連性を高め、審査の質/信頼性を向上させる。ハイリスク・ハイリターン申請に不利に働く可能性がある。

基準の段階的適用：提案を1つの基準セットで評価し、ランキングと足切りスコアを決定。その後、足切りスコアを超えた提案を別の基準セットで再評価し、最終的な採択リストを決定する。全ての基準が適用されることを確実にし、プログラムの目的に対する関連性を高める。

カテゴリごとの割り当て活用：ランキングを付けた後、EDI、所属機関の立地、初回申請者など特定のカテゴリで十分な数が確保されるように提案を審査する。非常に抜本的なアプローチだが、他の方法で多様性を確保できる場合もある。

5. 研修とフィードバック（Training and feedback）への介入

若手研究者（ECRs）の審査員登用とメンタリング：パネルメンバーや審査員は経験豊富な研究者/イノベーターである傾向があるが、ECRsに追加の研修を提供し、審査やパネルに参加させる。レビューの質を向上させ、審査員の多様性を高める。

評価へのEDIの組み込み：審査担当者が無意識のバイアスを認識し、審査プロセス中に相互に指摘し合うよう研修や支援を提供する。バイアスを減らし、採択者の多様性を高める。効果のない研修が誤った自信を生む可能性がある。

不採択申請者へのフィードバックの拡大/削減：不採択となった申請者に対しフィードバックを提供する。透明性を高め、不採択申請者の学習を促進する。ファンディング機関の負担が増え、フィードバックの質にばらつきが生じる可能性がある。

審査パネルへのファンディング機関代表者の参加：ファンディング機関の代表者がパネルに参加し、議論を誘導したり、プログラムの目的について説明したりする。純粋な事務機能を超えた役割を担い、議長役などを務める場合もある。ガイダンスが遵守されることを確実にし、決定がプログラムの目的と関連することを保証するのに役立つ。

審査の質の向上：研修、優秀な審査員の確保/維持、評価などを通じて、審査の質を向上させる。審査の質を高め、研修を簡素化し、審査依頼への応答率を向上させる。

公開審査/反論の機会提供：資金決定前でも審査が公開され、申請者が閲覧・回答できる。説明責任と審査の質を向上させる。資金提供者にとって負担が増える可能性（反論プロセスによる）や、タイムラインが長くなる懸念がある。

作成メンバー

総括責任者	山本 里枝子	上席フェロー	
メンバー	菊地 乃依瑠	フェロー	STI 基盤ユニット
	遠藤 悟	フェロー	STI 基盤ユニット

調査報告書

CRDS-FY2025-RR-04

研究評価の国際動向

令和 8 年 2 月 February 2026

ISBN 978-4-86829-035-3

国立研究開発法人科学技術振興機構 研究開発戦略センター
Center for Research and Development Strategy, Japan Science and Technology Agency

〒102-0076 東京都千代田区五番町7 K's 五番町

電話 03-5214-7481

E-mail crds@jst.go.jp

<https://www.jst.go.jp/crds/>

本書は著作権法等によって著作権が保護された著作物です。
著作権法で認められた場合を除き、本書の全部又は一部を許可無く複写・複製することを禁じます。
引用を行う際は、必ず出典を記述願います。
なお、本報告書の参考文献としてインターネット上の情報が掲載されている場合、当該情報はURLに併記された日付または本報告書の発行年月の1ヶ月前に入手しているものです。
上記以降の情報の更新は行わないものとします。

This publication is protected by copyright law and international treaties.
No part of this publication may be copied or reproduced in any form or by any means without permission of JST, except to the extent permitted by applicable law.
Any quotations must be appropriately acknowledged.
If you wish to copy, reproduce, display or otherwise use this publication, please contact crds@jst.go.jp.
Please note that all web references in this report were last checked on the date given in the link or one month prior to publication.
CRDS is not responsible for any changes in content thereafter.

FOR THE FUTURE OF
SCIENCE AND
SOCIETY



CRDS

<https://www.jst.go.jp/crds/>

