

令和6年度文部科学省委託調査

研究開発評価に関する実態調査・分析業務

報告書

2025年3月



本報告書は、文部科学省の令和 6 年度科学技術総合研究委託事業による委託業務として、公益財団法人未来工学研究所が実施した令和 6 年度「研究開発評価に関する実態調査・分析業務」の成果を取りまとめたものです。

研究開発評価に関する実態調査・分析業務
報告書

目次

1. 調査の概要	3
1.1 調査の目的	3
1.2 調査の内容	3
1.2.1 研究評価改革に関する国際的な動向及び G7 各国の対応状況	3
1.2.2 研究開発評価に関する調査・分析	3
1.2.3 文部科学省研究開発評価人材育成研修の開催	5
1.3 調査の方法	5
1.3.1 文献ウェブ調査	5
1.3.2 インタビュー調査	5
1.3.3 検討会議の開催	5
1.4 調査の体制	6
2. 研究開発評価に関する調査・分析	7
2.1 研究評価改革に関する国際的な動向及び G 7 各国の対応状況	7
2.1.1 研究評価をめぐる主な国際的取組の概要	8
2.1.2 国際動向の比較	29
2.1.3 G7 各国における対応状況	34
2.2 アメリカ合衆国（米国）	40
2.2.1 国・地域における研究開発評価制度の体系とその内容	40
2.2.2 主要機関における研究開発評価システムの体系とその内容	53
2.2.3 評価疲れについての各国・地域の現況	61
2.3 ドイツ連邦共和国（独国）	63
2.3.1 国・地域における研究開発評価制度の体系とその内容	63
2.3.2 主要機関における研究開発評価システムの体系とその内容	74
2.3.3 評価疲れについての各国・地域の現況	82
2.4 イタリア共和国（伊国）	84
2.4.1 国・地域における研究開発評価制度の体系とその内容	84
2.4.2 主要機関における研究開発評価システムの体系とその内容	95
2.4.3 評価疲れについての各国・地域の現況	98
2.5 調査・分析結果の比較と示唆	100
2.5.1 評価制度における研究開発評価の位置づけ	100
2.5.2 国レベルにおける研究開発評価に関する規定の整備状況	102

2.5.3	プログラム評価における自己評価、外部評価、第三者評価の使い分け等	103
2.5.4	研究開発課題の評価（プロジェクト評価）において重視されている評価の時期 ..	105
2.5.5	評価疲れをめぐる議論とそれらへの対応	106
3.	文部科学省研究開発評価人材育成研修の開催	109
3.1	開催概要	109
3.1.1	趣旨・目的	109
3.1.2	日程・会場	109
3.1.3	対象者及び募集方法	109
3.1.4	参加者	110
3.1.5	受講理由	111
3.1.6	プログラムの構成	111
3.2	開催結果	113
3.2.1	講義内容に対する評価及び意見	113
3.2.2	プログラム評価演習に対する評価及び意見	120
3.2.3	研修全体に対する評価及び意見	122
3.2.4	今後とりあげてほしいテーマ	123
3.3	今後に向けての示唆	124
3.3.1	講義について	124
3.3.2	演習について	124

1. 調査の概要

1.1 調査の目的

文部科学省では、「国の研究開発評価に関する大綱的指針」（平成 28 年 12 月 21 日内閣総理大臣決定）（以下、「大綱的指針」という。）が改定されたことを受け、「文部科学省における研究及び開発に関する評価指針」（平成 29 年 4 月 1 日文部科学大臣決定）（以下、「評価指針」という。）の改定を行った。

評価指針は、文部科学省の所掌に係る研究開発の評価を遂行する上での基本的な考え方をまとめたガイドラインであるが、平成 29 年度以来、改定がなされていない状況である。評価指針の最終改定時から研究開発評価の周辺状況は変化しており、現在の研究開発評価を巡る状況や研究開発評価の実態と課題を把握する必要があると考えられる。

このため、本調査では、諸外国の研究開発機関や資金配分機関にかかる研究開発評価に関して、諸外国の種々の取り組みの具体的、詳細な状況等について、文献調査や現地における担当者からの直接の聞き取り調査を行うものである。

さらに、我が国の研究開発機関や資金配分機関の実務担当者等を対象に研究開発評価人材育成研修を行い、研究開発評価を取り巻く状況や研究開発評価の実態を共有するとともに研究開発評価の更なる普及を図る。

これらの結果を報告書に取りまとめ、文部科学省及び研究開発機関等において今後の研究開発評価に関する実務・施策の企画立案に資するものとする。

1.2 調査の内容

1.2.1 研究評価改革に関する国際的な動向及び G7 各国の対応状況

研究評価改革に関する国際的な動向及び G 7 各国の対応状況を把握し、その動向・状況について概況を簡潔にまとめた。

主な国際的動向としては、「研究評価に関するサンフランシスコ宣言（DORA）」、「研究評価の改革に関する合意」及び「同合意を推進するための有志連合（CoARA）」等に加え、評価研究および実務上において着目されているトピックやトレンドを把握するため、ウィーンで開催された国際会議「REvaluation Conference 2024」に参加し、情報収集を行った。

1.2.2 研究開発評価に関する調査・分析

以下大きく 3 項目についての調査・分析を実施した。

なお、調査対象国及び機関は、次表のとおりである。これらは、後述の検討会議の助言を受けて選定したものである。

表 1-1 調査対象

国等	特徴	調査対象機関
米国	政府業績成果現代化法(GPRAMA)等に基づきプログラム評価を中心としたエビデンス重視の業績マネジメントシステムを展開、評価の実効性を高めるための多様なガイドラインも整備。その他、多様なチェック・アンド・バランスの仕掛けが存在。	全米科学財団(NSF)
独国	連邦政府と州政府が分権的に関与し、共同して資金配分を実施。そのため、研究開発評価は、個々の研究開発プログラムの特性に応じて実施されており、評価に関する体系的な規程等も整備されていない。	ドイツ研究振興協会(DFG) DLRプロジェクトレガー
伊国	研究評価はイタリア共和国憲法第97条に謳われている善政の原則を、科学的・学術的領域においても保証する。研究評価では、研究活動の厳密な科学的側面と社会経済的側面の両方が考慮される。公的研究機関であるCNRはANVURから機関評価を受け、自らも研究プログラムの評価を行う。	全国大学システム研究評価委員会(ANVUR) 全国科学研究センター(CNR)

(1) 研究開発評価制度に関する調査・分析

検討会議における助言を踏まえ、大きく以下の2点の調査分析を実施した。その際、特に、「研究開発課題の評価(プロジェクト評価)¹⁾」については分野特性への配慮や事前・中間・事後評価の重み付けに、「研究プログラムの評価」については自己評価、外部評価、第三者評価の区分に着目して調査分析を実施した。

1) 国・地域における研究開発評価制度の体系とその内容

各国・地域政府当局が提示する研究開発評価に関わる法令、規程、ガイドライン等(研究開発評価に対する要求事項)を調査し、政策評価や機関評価等と関係を含めて、研究開発評価がどのように制度化されたり、義務づけられているのかを体系的に明らかにした。

それらの要求事項において、「研究プログラムの評価」及び「研究開発課題の評価(プロジェクト評価)」の枠組みや考え方、具体的な手法や評価フロー等の詳細について、規定しているものはあるか、規定しているとすればそれは具体的にどのようなものかに加え、近年注目されている事項があれば合わせて調査を行った。

2) 主要機関における研究開発評価システムの体系とその内容

各国における代表的な資金配分機関を事例としてとりあげ、国レベルにおける評価制度の枠組みの中でどのような研究開発評価システムの体系を構築しているのか調査を実施した。特に、機関独自に策定しているガイドライン等がある場合、その詳細について取りまとめを行った。

¹⁾ 大綱的指針における「研究開発課題の評価」は、評価論において一般的に「プロジェクト評価」と呼ばれており、諸外国においても基本的にはこの呼称を用いている。本報告書では、特に断らない限り、この両者を無差別に用いている。

(2) 評価疲れについての各国・地域の現況

大綱的指針における「評価疲れ」について、各国の研究開発評価及び関連する政策評価等の評価に係る実施状況や重複状況を確認し、各国・地域における「評価疲れ」の実態を調査分析した。なお、「評価疲れ」の主体には、評価事務局、評価者、被評価者が考えられ、これら主体の違いを意識した上でとりまとめを行った。

1.2.3 文部科学省研究開発評価人材育成研修の開催

委託業務の目的を実現するため、文部科学省研究開発評価人材育成研修を実施した。研修の内容や実施形式、具体的な対象者については、検討会議において検討の上決定した。

1.3 調査の方法

1.3.1 文献ウェブ調査

これまで当研究所が実施、蓄積してきた調査研究の実績をベースに、調査対象国・地域及び機関の公式資料等をウェブサイト等から収集し、情報の整理及び分析を実施した。また、関連する評価研究等における成果があれば、合わせて検討の素材として活用した。

1.3.2 インタビュー調査

文献ウェブ調査からでは明らかにすることのできない評価内容や詳細な運用の実態等について把握するために、2024年12月4日から6日にかけてウィーンで開催された国際会議「REvaluation Conference 2024」に参加し情報収集を行うとともに、同会議に参加の調査対象機関担当者等と意見交換を行った。

1.3.3 検討会議の開催

調査・分析及び研修の実施に必要な事項について検討し助言を得ることを目的として、有識者からなる検討会議を設置した。

検討会議のメンバー構成は、次のようなものである。

伊地知寛博 成城大学 社会イノベーション学部 教授
寫田 敏行 大学改革支援・学位授与機構 研究開発部 教授
前波 晴彦 自然科学研究機構 研究力強化推進本部 特任准教授
祐野 恵 京都大学学際融合教育研究推進センター 特定講師

検討会議は、次のような日程、議事で、全4回開催した。いずれもオンラインでの実施であり、すべての回にメンバー全員が参加し、検討を行った。

表 1-2 検討会議の開催概要

日時	議事
2024年11月15日(金) 11:00~12:30	1) 委員紹介 2) 調査の概要及び進め方について 3) 次回以降の予定について
2024年12月20日(金) 9:00~11:00	1) これまでの調査に関して 2) 研修について 3) 次回以降の予定について
2025年2月13日(木) 9:00~11:00	1) 海外事例調査に関する中間報告 2) 研修の振り返り 3) 「評価疲れ」に関する話題提供 4) 次回以降の予定について
2025年3月17日(月) 11:00~13:00	1) 海外事例調査(米、独、伊及び国際動向)に関する報告 2) 海外調査からの示唆に関する意見交換

1.4 調査の体制

本調査は、以下のような体制で実施した。

田原敬一郎	公益財団法人未来工学研究所政策調査分析センター	主任研究員
安藤 二香	公益財団法人未来工学研究所政策調査分析センター	主任研究員
大竹 裕之	公益財団法人未来工学研究所政策調査分析センター	主任研究員
野呂 高樹	公益財団法人未来工学研究所政策調査分析センター	主任研究員
山本 智史	公益財団法人未来工学研究所政策調査分析センター	主任研究員

2. 研究開発評価に関する調査・分析

2.1 研究評価改革に関する国際的な動向及びG 7 各国の対応状況

本稿で扱う対象は、DORA（研究評価に関するサンフランシスコ宣言）、CoARA（研究評価推進連合）、GRC WG RRA（責任ある研究評価に関するグローバル・リサーチ・カウンシル作業部会）、INORMS（研究管理学会国際ネットワーク）のSCOPEフレームワーク、More Than Our Rank イニシアティブ、Leiden Manifesto（ライデン声明）、The Metrics Tide である。各調査対象の概要については次表のとおりである。

表 2-1 調査対象の概要

対象	概要	関連国・機関など
DORA（研究評価に関するサンフランシスコ宣言）	2012年12月の米国細胞生物学会(ASCBASCB)の年次大会において、学術雑誌の編集者や研究者が議論した内容を元にまとめられた。第一義的には、査読誌に掲載された研究論文に関する実践方法に焦点を当てている。	英国 UKRI、スイス National Science Foundation、Howard Hughes Medical Institute、オランダ NWO、欧州分子生物学機構(EMBO)など
CoARA（研究評価推進連合）	研究、研究者、研究組織を評価する方法とプロセスの改革に取り組む団体の集まり。CoARA のビジョンは、査読が不可欠な評価において、定量的指標の責任ある使用に支えられた定性的判断を重視することにより、研究の質とインパクトを最大化する多様な成果、実践、活動を認識すること。	700を超える研究組織、資金提供者、評価機関、専門家協会、およびそれらの団体
GRC WG RRA（責任ある研究評価に関するグローバル・リサーチ・カウンシル作業部会）	研究評価に対する責任あるアプローチについての理解を促進し、協力を通じて学び、優れた実践例をオープンに共有することにより、RRA の採用を促進し、世界的に支援することをミッションとしている。	英国やドイツ、カナダ、中国、豪州、アルゼンチンなど世界中の 23 の組織から 25 名のメンバーで構成
INORMS（研究管理学会国際ネットワーク）の SCOPE フレームワーク	国際研究管理学会ネットワーク (INORMS) の研究評価グループ (REG) は、研究評価が有意義で責任があり、効果的であることを確実にするための最善の方法を検討するために 2018 年に設立された。この目的のために、グループメンバーは、責任ある研究評価のための SCOPE フレームワークを設計している。	INORMS は、英国の研究管理者協会 (ARMA) や欧州研究管理者協会 (EARMA)、全米大学研究管理者協議会 (NCURA)、中国科学技術政策協会 (CASSSP) など、世界中の研究管理学会の会員ネットワークで構成。
More Than Our Rank イニシアティブ	INORMS の研究評価グループで進めている活動で、世界大学ランキングの問題点や影響に対応するために開発された。このイニシアティブは、学術機関がランキング順位に反映されていない世界への貢献の多様で多様な方法を強調する機会を提供。	キール大学、クイーンズランド工科大学 (QUT) などが署名。支援団体としては、ライデン大学 CWTS、DORA、EUA、サイエンス・ヨーロッパなど。
Leiden Manifesto（ラ	10 項目の原則 (principles) から成り、研究評価	Nature 誌で Dr. Diana

イデン声明)	における計量データの利用についてのベストプラクティスや注意点を示したものであり、研究者、管理者、評価者の全てに対する、計量データに立脚した研究評価のガイドラインと考えられる。	Hicks、Dr. Paul Woutersら5名の著者により発表。
The Metrics Tide	「研究評価と管理におけるメトリクスの役割に関する独立レビュー」の主な所見と提言をまとめたもの。このレビューでは、先行研究の枠を超え、研究評価指標と指標の潜在的な用途と限界について深く考察している。様々な学問領域におけるメトリクスの利用を検討し、研究の卓越性とインパクトの発展に対するメトリクスの潜在的な貢献を評価した。	レビューの議長は SPRU のジェームズ・ウィルストン教授が務め、サイエントロメトリックス、研究助成、研究政策、出版、大学経営、研究管理の専門家からなる独立した学際的なグループがサポート。

出所:「研究評価に関する国際シンポジウム—研究評価改革に関する国際動向—」結果概要²及び各対象のウェブサイトをもとに未来工学研究所作成

2.1.1 研究評価をめぐる主な国際的取組の概要

(1) DORA（研究評価に関するサンフランシスコ宣言³）

研究評価に関するサンフランシスコ宣言（DORA）」は、2012年12月の米国細胞生物学会（ASCB/ASCB）の年次大会において、学術雑誌の編集者や研究者が議論した内容を元にまとめられた。DORAの支援団体については、英国 UKRI、スイス National Science Foundation、Howard Hughes Medical Institute が Visionary として、オランダ NWO や欧州分子生物学機構（EMBO）、The Research Council of Norway などが Sustainer として、Cancer Research UK や F1000Research、Iowa State University Library などが Contributor としてサポートしている。また、慈善財団 Arcadia は、研究評価を促進するツール（TARA）をサポートしている。DORAは2023年時点で21,000人以上の個人と3,000を超える機関が署名をしている。

研究の有効性を評価するにあたり、研究論文以外の成果の重要性は今後増していくにしても、研究の到達点を示す研究成果の中核は査読付き研究論文であり続けるだろうことから、勧告は、第一義的には、査読誌に掲載された研究論文に関する実践方法に焦点を当てている。勧告は、助成機関、学術機関、学術雑誌、数量的指標を提供する機関、そして、個々の研究者に向けたものである。

勧告に一貫している論点は次のとおりである。

- 資金助成、職の任命や昇進の検討の際に、インパクトファクターのような雑誌ベースの数量的指標の使用を排除する必要性

² 研究評価に関する国際シンポジウム—研究評価改革に関する国際動向—結果概要

< <https://www.jst.go.jp/osirase/2023/pdf/20240229-2.pdf> > [2025/3/21 取得]

³ DORA ウェブサイト < <https://sfdora.org/read/read-the-declaration-japanese/> > [2025/3/21 取得]

- 研究が発表される雑誌をベースにするのではなく、研究自体の価値に基づく評価の必要性
- オンライン出版が提供する機会（例えば論文における単語や図表、参考文献の数についての無意味な制限の緩和、および重要性やインパクトに関する新しい指標の探索）を十分に活用する必要性

一般勧告を含む各対象に向けた勧告は次のとおりである。

<一般勧告>

- 個々の科学者の貢献を査定する、すなわち雇用、昇進や助成の決定をおこなう際に、個々の研究論文の質をはかる代替方法として、インパクトファクターのような雑誌ベースの数量的指標を用いないこと。

<助成機関へ>

- 助成申請者の科学的生産性の評価に用いられる判断基準が明示的であること。また、特にキャリアの初期段階にある研究者に対して、出版物の数量的指標やその論文が発表された雑誌がどのようなものであるかということよりも、その論文の科学的內容の方がはるかに重要であることを、はっきりと強調すること。
- 研究評価を行う上で、研究出版物に加えて（データセットやソフトウェアを含む）研究のすべての成果の価値とインパクトを検討すること。また、政策や実用化への影響といった研究インパクトの質的な指標を含む、幅広いインパクトの評価基準を考慮すること。

<学術機関へ>

- 雇用、任期、昇進の決定する際に用いられる判断基準が明示的であること、特にキャリアの初期段階にある研究者に対して、出版物の数量的指標やその論文が発表された雑誌がどのようなものであるかということよりも、その論文の科学的內容の方がはるかに重要であることを、はっきりと強調すること。
- 研究評価を行う上で、研究出版物にくわえて研究の（データセットやソフトウェアを含む）すべての成果の価値とインパクトを検討すること。また、政策や実用化への影響といった研究インパクトの質的な指標を含む、幅広いインパクトの評価基準を考慮すること。

<出版社へ>

- 販売促進手段としてのインパクトファクターの強調を大幅に縮小させること、理想的にはインパクトファクターの宣伝を中止すること、または雑誌のパフォーマンスについてより豊富な視点を与える様々な数量的指標（例、**5-year impact factor**、**EigenFactor**、**SCImago**、**h-index**、編集と出版に要する時間等）の文脈に沿った上でインパクトファクターを提供すること。

- 様々な論文レベルでの数量的指標を利用可能にすること、それによって論文が発表された雑誌についての数量的指標ではなく、論文自体の科学的内容を基にした評価への転換を促すこと。
- 責任あるオーサiershipの慣行と各著者個別の貢献についての情報提供を促すこと。
- 雑誌がオープンアクセスであろうと購読モデルであろうと、研究論文の参考文献リストについての再利用の制限を取り除き、それらをクリエイティブコモンズのパブリックドメインの下で利用できるようにすること。
- 研究論文のレファレンスの数についての制限を縮小、または廃止させること、そして必要に応じて、最初に発見を報告したグループの功績を認めるために、レビューではなく原著論文の引用を義務付けること。

<数量的指標を提供する機関へ>

- すべての数量的指標は、それを計算するために使われたデータと方法とを提供することにより、オープンかつ透明であること。
- 無条件の再利用を認めるライセンス下でデータを提供し、可能な限りコンピュータからアクセスできるようにすること。
- 数量的指標の不正な操作が決して許されないよう明確に示すこと、また不正な操作に相当するものとは何か及びこれに対する措置について明示的に示すこと。
- 数量的指標が使われ、集約され、あるいは比較される際に、論文のタイプ（例、レビュー記事 vs 研究論文）あるいは異なる対象領域において生じる数量的指標の差異について、説明すること。

<研究者へ>

- 研究助成、雇用、任期、昇進について決定する委員会に参加した場合は、出版物の数量的指標ではなく科学的内容を基にして評価を下すこと。
- 認めるべき功績を認めるために、適切である限り、レビュー記事ではなく観察結果が最初に報じられた原著論文を引用すること。
- 個人の発表した論文やその他の研究成果のインパクトの根拠として、自己推薦書では、論文に関する様々な種類の数量的指標を用いること。
- インパクトファクターに不適切に依存している研究評価の慣例を批判し、個別の研究成果の価値や影響に注目するベストプラクティスを推進し、振興すること。

2023年3月には新たな3か年戦略計画を発表した。新たな戦略目標は次のとおりである。

- 不適切な指標に過度に依存する研究評価方法の悪影響と、代替方法の好影響についての認識を高める
- 研究評価改革のための明確かつ具体的な対策の策定を加速する
- 世界中の研究評価改革の支持者を支援する
- DORA の使命を可能な限り効率的かつ迅速に遂行するために必要な資金を確保する。

戦略計画では、DORAの原則、価値観、根拠、および今後3年間で完了したいと考えている具体的な活動の概要が示されている。改革に対する個人および組織の取り組みを示すために署名を集めることを目指しているが、主な目標は、コミュニティのメンバーに、変化を実行するために必要なリソース、コミュニティ サポート、およびツールを提供するための大きな進歩を遂げることである。

また、2024年1月には Reformscape プロジェクトを立ち上げている。Reformscape は、世界中の学術コミュニティが、採用、昇進、終身雇用をより公平に、より強固に、より多様化する方法の例を模索し、共有できるようにする新しいオンライン リソースである。使いやすいオンライン ポータルから無料で使用できる Reformscape は、世界中の大学やその他の学術機関が、学術キャリア評価に対して、より公平で責任ある、より情報に富んだアプローチをどのように導入しているかを示す、数百の実例を集めた、豊富で整理されたデータセットである。Reformscape には、世界中の200を超える機関からのポリシー、行動計画、その他の文書が、傾向や専門家がまとめた洞察とともに掲載されている。Reformscape は、オランダのライデン大学科学技術研究センター (CWTS) や イリノイ工科大学デザイン研究所、インペリアル・カレッジ・ロンドン、DORAの共同プロジェクトである、学術キャリア評価のための新しいポリシーと実践の開発を促進するプロジェクト「Tools to Advance Research Assessment (TARA)」の一環として開発された。

最近の文書としては、2024年には「研究評価における定量的指標の責任ある使用に関するガイダンス」が公表されている。この文書は、研究評価に使用されるいくつかの指標（メトリクスと呼ばれることもある）の使用に関するガイダンスを提供している。ジャーナル・インパクト・ファクターやその他のジャーナルの測定、引用数、h 指数、分野標準化引用指標、およびオルトメトリクスである。これらのメトリクスの使用には、明確であること、透明性があること、具体的であること、文脈的であること、公平であることという5つの原則が指針となる。

2023年11月8日（水）に日本科学未来館にて開催された研究評価に関する国際シンポジウム「研究評価改革に関する国際動向」においては、DORA 運営委員長のステファン・カリー氏より講演「研究評価に関するサンフランシスコ宣言 (DORA) 卓越性の測定方法について再考する」およびパネルディスカッション等が行われた⁴。DORA カリー運営委員長からは、「Excellence (卓越性)」については幅広く使用されているが、その概念についての定義は明確ではない。一般的には、トップジャーナルに論文を発表するか、多くの研究資金を獲得する場合を意味するが、そもそも質の高い研究とは何かを定義しなければ過当競争の中で色々な問題が発生する。アカデミアにおいては、どのジャーナルに発表したかが非常に重要であることから、ジャーナル・インパクト・ファクター (JIF) に集中しすぎると論文発表が遅れるという問題が発生する。DORAは18の提言を行っており、うち17の提言は、助成機関、学術機関、出版社、数量的指標を提供する機関、研究者それぞれに対するものである。研究評価にあたっては、論文の出版だけでなく政策への影響やその実践とい

⁴ 前掲「研究評価に関する国際シンポジウム」結果概要

ったインパクト等の質に関する指標を含めて、全ての研究成果の価値を考慮しなければならない。一般的なプロポーザルの審査を行うと、男性の研究者が採択される。性別、民族・人種の多様化が進まない。変化のための一つのツールとして、世界的に多く採用されているのが Narrative CV である。これまでの研究者の研究履歴を見れば、研究に関する貢献、また、定量的な研究成果も把握できる。また、英国で進めている取組 REF2028 では、評価の重点を、個人の業績から、より健全な研究環境への貢献へと変更しようとしている。研究環境をどのように測定するかについては、これから協議されるとしている。

(2) CoARA（研究評価推進連合）⁵

Coalition for Advancing Research Assessment (CoARA)は、研究、研究者、研究組織を評価する方法とプロセスの改革に取り組む団体の集まりである。CoARA のビジョンは、査読が不可欠な評価において、定量的指標の責任ある使用に支えられた定性的判断を重視することにより、研究の質とインパクトを最大化する多様な成果、実践、活動を認識することである。CoARA の使命は、共通の原則とコミットメントに基づき、一定の期間内に研究評価の体系的な改革を可能にすることである。協調的な行動、知識の交換、相互学習を通じて、CoARA はすべてのメンバーがより包括的で効果的なアセスメントの実践を採用し、改良することを支援することを目指している。

現在の研究評価方法は、引用数のような論文ベースの指標に大きく依存しており、研究者による幅広い貢献を認識できないことが多い。700 を超える研究組織、資金提供者、評価機関、専門家協会、およびそれらの団体が、研究、研究者、研究組織の評価における改革を実施するための共通の方向性と指導原則に合意し、改革と実施の概要を示す 2022 年 7 月に発表された「研究評価の改革に関する合意」にまとめられている。

「研究評価の改革に関する合意」は、研究の質とインパクトを最大化するという包括的な目標を掲げ、研究、研究者、研究実施機関の評価慣行を変えるための共通の方向性を示している。この合意には、改革の原則、約束、時間枠が盛り込まれ、改革を実施するために協力する意思のある組織の連合の原則が示されている。

署名者は、研究、研究者、研究組織の評価において、研究の質とインパクトを最大化する多様な成果、実践、活動を認識するという共通のビジョンを約束する。そのためには、査読を中心とした質的判断に基づいた評価を行い、それを定量的指標の責任ある使用によってサポートすることが必要である。

協定と連合に支えられた改革運動は、より質が高く、よりインパクトがあり、より効率的で包括的な研究システムに向けて共に前進するための、包括的で協力的な場となることを目指している。組織の自主性を十分に尊重しつつ、新しい評価基準、方法、ツールを開発し、共同で批判的な考察を行い、グッドプラクティスを交換し、相互学習するための、試験的、実験的なプラットフォームを提供する。連合の活動に支えられながら、各組織は公約を実行に移すためのステップを決定し、改革の旅のペースを決定する。

⁵ CoARA ウェブサイト < <https://coara.eu/> > [2025/3/21 取得]

コミットメントとしては次の 10 項目が挙げられている。

- ① 研究のニーズと性質に応じて、研究への貢献と研究におけるキャリアの多様性を認識する
- ② 研究評価は、主にピアレビューを中心とした定性的な評価に基づき、定量的な指標を、責任を持って使用して行う
- ③ 研究評価において、ジャーナルや出版物に基づく指標、特にジャーナルインパクトファクター (JIF) や h 指数の不適切な使用をやめる
- ④ 研究評価において研究機関のランキングを使用しない
- ⑤ 組織改革を達成するために必要な研究評価改革に資源を投入する
- ⑥ 研究評価基準、ツール、プロセスを検討し、開発する
- ⑦ 研究評価改革に対する意識を高め、評価基準とプロセスおよびその使用に関する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
- ⑧ 連合内外での相互学習を可能にするために実践と経験を交換する
- ⑨ 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伝える
- ⑩ 確固たる証拠と最先端の研究に基づいて実践、基準、ツールを評価し、証拠収集と研究のためにデータを公開する

また、行動計画 (アクションプラン) の概要については、研究評価の改革に関する CoARA 協定に関連して、署名組織 (加盟組織も含む) は、協定締結後 1 年以内に、マイルストーンを定めたアクションプランに従って、コアコミットメントの実施プロセスをどのように開始したかをコミュニティと共有するよう求められている。期限については、本協定の加盟者は、本協定に署名してから 1 年以内に、マイルストーンを定めた行動計画に従い、中核的公約に沿った基準、ツール、プロセスの見直しや開発のプロセスを、自らの組織がどのように開始したかを、互い、また地域社会と共有することに同意する。この協定の署名者は、2027 年末または協定署名から 5 年以内に、評価基準、ツール、プロセスの見直し、開発、評価に向けた進捗状況を定期的に示すことに同意する。

CoARA のワーキンググループとしては次図に示すように、責任ある研究評価の実践への移行を支援するため、研究機関が専有的なインフラや研究情報から、オープンで相互運用可能な代替手段へと移行できるようにすることをミッションとした **Towards Open Infrastructures for Responsible Research Assessment (OI4RRA)**、排除や優先順位をつけることなく、あらゆる言語での多様な能力、交流、コミュニケーションを評価する学術文化を育むことをミッションとする **Multilingualism and Language Biases in Research Assessment** (研究評価における多言語主義と言語偏見)、新しい取り組みや代替的な方法論を収集し、改良し、試験的に実施する変化を可能にするプロセスを確立することによって、研究評価における実験的なアイデアのインキュベーターを形成することをミッションとする **Experiments in Assessment: Idea Generation, Co-Creation, and Piloting** などがある。

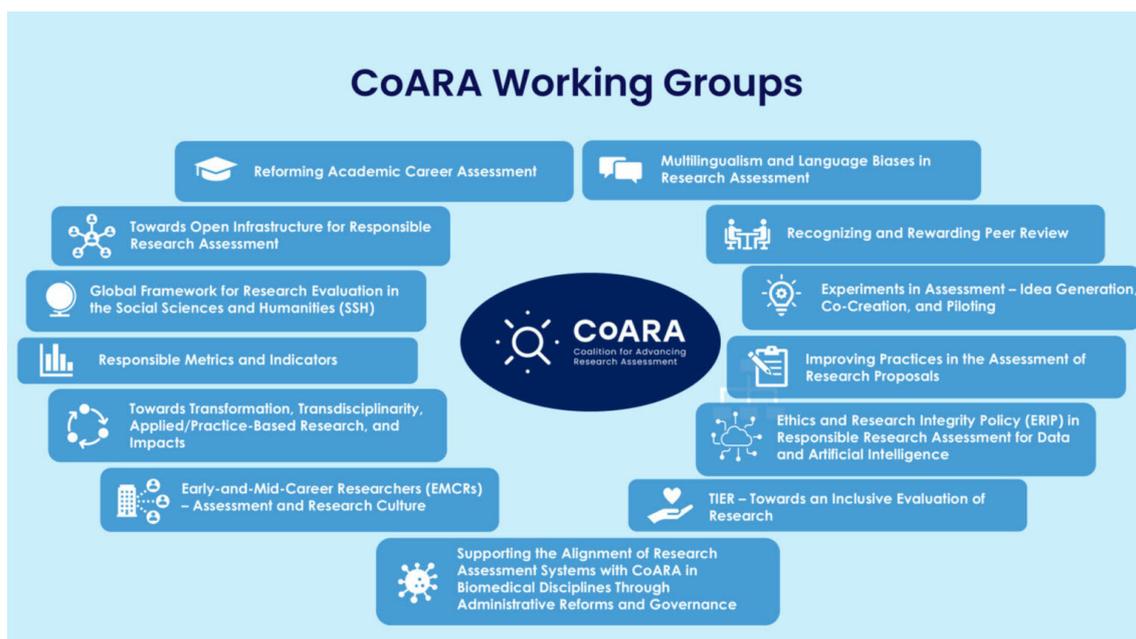


図 2-1 CoARA ワーキンググループの構成

出所: Working Groups ウェブサイト⁶

また、プロジェクトとしては、CoARA Boost プロジェクトなどが展開されている。CoARA Boost プロジェクトは、CoARA の運営能力を強化することを目的としており、欧州科学財団は CoARA ブースト・イニシアチブのコーディネーターを務めている。

このプロジェクトは、研究評価を改革するためのクリティカル・マスを開発し、新規加盟国のための重みを生み出すとともに、研究評価の新しいモデルを調査し、実施するための手段を提供するものである。プロジェクトは 2023 年 10 月に開始され、期間は 36 ヶ月である。プロジェクト全体の予算の半分以上は、カスケード助成金という形で研究評価改革に参加する機関に再配分される。

CoARA Boost プロジェクトの主な活動は以下の通りである。

- 情報・知識の共有、意識向上など、連合に対する運営支援を強化する
- 多種多様な地域にわたる多数の組織への制度改革に貢献するための段階的助成金メカニズム。総予算 275 万ユーロから少なくとも 50 のプロジェクトに資金が提供される
- 合意された 10 のコミットメントの達成に向けた行動計画の取りまとめや実行を支援することを含め、研究評価の改革に向けた制度改革の実施に貢献する
- 研究評価の新たなモデルを模索する作業部会に支援を提供する
- CoARA の会員数を世界的に拡大し、研究評価の進化に関する国際協力を強化することで、研究評価の改革に向けた欧州の取り組みのアウトリーチを拡大する

⁶ Working Groups ウェブサイト <<https://coara.eu/working-groups/working-groups/>> [2025/3/21 取得]

2023年11月8日（水）に日本科学未来館にて開催された研究評価に関する国際シンポジウム「研究評価改革に関する国際動向」においては、CoARA 運営委員会メンバーであるサイエンス ヨーロッパ事務総長のリディア・ボレル ダミアン氏より講演「研究評価改革における国際動向」およびパネルディスカッション等が行われた⁷。同氏は、CoARA として、特に次の4つの重要なコミットメントがあるとしている。

- 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定性的評価に基づき、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（JIF や h index）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

(3) GRC WG RRA（責任ある研究評価に関するグローバル・リサーチ・カウンシル作業部会）⁸

グローバル・リサーチ・カウンシル（GRC）は、世界中の科学および工学の資金提供機関の責任者から構成される仮想組織であり、世界中の資金提供機関間での質の高いコラボレーションのためのデータとベストプラクティスの共有を促進することなどを目的としている。GRCの詳細な目的は以下の6つである。

- 資金提供機関間のコミュニケーションと協力を改善すること
- 質の高い研究協力のためのデータとベストプラクティスの共有を促進すること
- リサーチ・カウンシルのトップによる定期的な会合の場を提供すること
- 研究と教育の支援における共通の関心事に対応し、その機会を提供すること
- 世界トップクラスの研究環境を構築することを望む機関のリソースとなること
- グローバルな科学事業と世界の研究コミュニティを支援するメカニズムを探求すること

責任ある研究評価（Responsible Research Assessment : RRA）とは、多様で包括的な研究文化をサポートするために、高品質の研究の多様な特性を奨励し、反映し、報いる評価アプローチを指す。

Responsible Research Assessment（RRA）ワーキンググループでは、研究評価に対する責任あるアプローチについての理解を促進し、協力を通じて学び、優れた実践例をオープンに共有することにより、RRAの採用を促進し、世界的に支援することをミッションとしている。また、GRC参加組織（および参加組織が資金を提供する組織）は、最高水準で行われる厳密な研究を支援する、より良い、より健全な研究文化を支えるために、優れた研究の

⁷ 前掲「研究評価に関する国際シンポジウム」結果概要

⁸ GRC WG RRA ウェブサイト < <https://globalresearchcouncil.org/> > [2025/3/21 取得]

多様な属性にインセンティブを与え、それに報いる 評価アプローチを導入することをビジョンとしている。多様で包括的な研究文化の構築を支援するため、国際的な研究・イノベーションシステムにおける RRA の推進と実施について、GRC を主導的な声として位置づけることを目標としている。当ワーキンググループのミッション、ビジョン、目標は、当グループの活動の基礎となるアクションプランにまとめられている。

2024 年 5 月には、スイスのインターラーケンで第 12 回 GRC 年次総会が開催されている。この会議では、GRC 参加組織が 3 つの側面から「持続可能な研究」について議論した：持続可能性に関する研究、研究の持続可能性、持続可能性科学の社会への利用である。このトピックは、2023 年 10 月から 11 月にかけて開催される 5 つの GRC 地域会合で発展的に議論された。この中で、RRA の 11 の側面と GRC RRA 調査も発表されている。また、ケーススタディ・フレームワークを共有し、会議のテーマである持続可能性と RRA の関連性を探る機会を得た。

RRA ワーキンググループの設立については、2021 年 5 月、GRC 参加者は、RRA の目標を奨励し実現するための具体的な行動を取ることに合意した。参加組織がこれらの目標に向けて前進できるよう支援するために、新しい GRC RRA ワーキンググループが設立された。RRA ワーキンググループは、英国やドイツ、カナダ、中国、豪州、アルゼンチンなど世界中の 23 の組織から 25 名のメンバーで構成されている。グループのメンバーには GRC の各地域の代表者が含まれており、世界的な視点と世界中の資金提供者への働きかけが保証されている。

これらの背景としては、2020 年 11 月、英国の UKRI は、南アフリカの NRF および英国の責任ある研究指標フォーラムと共同で、責任ある研究評価に関する GRC 仮想会議を開催した。会議に先立ち、Research on Research Institute (RoRI) は「責任ある研究評価における資金提供者の役割の変化：進捗、障害、今後の方向性」に関するワーキングペーパーを発表した。このペーパーでは、責任ある研究評価の現状を示し、その意味を探り、この分野における既存の原則とフレームワークを概説している。また、2020 年 9 月に実施された GRC 参加組織に対する各組織の RRA ポリシーと実践に関する調査結果も紹介している。2021 年 5 月に開催された第 9 回 GRC 年次総会では、会議概要報告書が発表され、GRC は参加組織に既存の RRA 原則を定着させ、RRA の目標を奨励して達成するための具体的な行動をとるよう求める行動要請を承認した。GRC RRA ワーキンググループは、これらの活動に基づいて活動を進めている。

現在の活動としては次の 4 つが挙げられる。

1) 責任ある研究評価の側面（次元）

GRC 参加組織は、責任ある研究評価の枠組みを求め、資金提供者の RRA のビジョンを明確にするために 11 の次元のセットが開発された。資金提供機関はこれを使用して、責任ある研究評価と評価改革の原則をどこに組み込むかを理解できる。

<基本原則>

- 研究の誠実性と責任ある研究実施に対する責任ある取り組みと推進
- オープンな研究を奨励するための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
- 地球規模の課題や緊急事態が研究に与える影響に対する責任ある対応

<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
- 研究評価改革への責任あるアプローチ
- 組織の方針や慣行に対する影響力を責任を持って行使する

<プロセスと方法論>

- 研究評価基準の責任ある使用と普及
- 研究への貢献と成果の責任ある評価
- インパクト・アセスメントに対する責任あるアプローチ
- 査読者とパネルの募集およびトレーニングに対する責任あるアプローチ

2) 調査

Research on Research Institute (RoRI) プロジェクト「A Global Observatory on Responsible Research Assessment (責任ある研究評価に関する世界的観測)」と GRC RRA ワーキンググループは、GRC 参加者間の RRA 実践の範囲を評価するための調査を開始した。調査は 2024 年 5 月 27 日に開始され、2025 年 1 月 6 日まで実施されている。調査結果は GRC 地域会議で共有され、完全なワーキングペーパーは 2025 年の GRC 年次会議で公開され、共有される予定である。

3) ケーススタディ

結果については近日に公開予定だが、RRA の次元に沿って、以下の目的でケーススタディを収集している。

- 資金提供者が RRA をどのように実装し、組み込んだかを共有
- ギャップや障壁が存在する場所で知識ベースを拡張

4) 自己評価ツールとサポートロードマップ

成果については近日に公開予定だが、今後リリースされるツールは、ワーキンググループのメンバーやその他のメンバーの国際的かつ幅広い視点を活用し、ツールが使用される地域の状況に合わせて調整できるさまざまな要素が確実に含まれるようにする予定となる

いる。この自己評価ツールは、資金提供者が体系的な変更を行い、障害を克服し、RRAの実践を実施する方法に関するガイダンスを提供するロードマップによって補完される。

(4) INORMS（研究管理学会国際ネットワーク）SCOPE フレームワーク⁹

国際研究管理学会ネットワーク（INORMS）は、世界中の研究管理学会や協会をまとめるために2001年に設立された。その目的は、会員学会間の交流、優れた実践の共有、共同活動を可能にし、各会員の利益となるようにすることである。

INORMSは、英国の研究管理者協会（ARMA）やオランダの研究管理者協会、欧州研究管理者協会（EARMA）、全米大学研究管理者協議会（NCURA）、ブラジル研究管理者協会、カナダ研究管理者協会（CARA）、中国科学技術政策協会（CASSSP）、国際研究管理者協会（SRAI）など、世界中の研究管理学会の会員ネットワークで構成されている。

INORMS 研究評価グループ（REG）は、以前は REWG として知られていたが、研究評価が有意義で責任があり、効果的であることを確実にするための最善の方法を検討するために2018年に設立された。（現在の INORMS 研究評価チームの議長は、英国ラフバラー大学研究文化・評価部長のエリザベス・ガッド博士で、研究評価推進連合（CoARA）の副議長も務めている。2022年には、「Harnessing the Metric Tide: Indicators, Infrastructures & Priorities for UK Research Assessment」の共著者となった。）この目的のために、INORMS を構成する研究管理学会を代表するグループメンバーは、2つの作業パッケージに注目した。1つ目は、責任ある研究評価の採用と実践において、大学の上級リーダーと実務家を導くことを目指している。この目的のために、責任ある研究評価のための SCOPE フレームワークと、上級リーダーを関与させるための一連の「5つの議論」を設計した。2つ目のイニシアティブは、世界大学ランキングが大学の行動に及ぼす影響の増大を強調することに焦点を当てており、スコアリングと競争の重視に対する留保を表明している。これらのランキングの基礎データに関する議論が促進され、潜在的な代替案がグループによって検討されている。さらに、世界大学ランキングの特定の問題のある特徴と影響に対応して、グループはランキングの活用を再考することを目的とした More Than Our Rank イニシアティブを導入した。

研究評価のための SCOPE フレームワークは、責任ある評価を行うための5段階モデルである。これは、研究マネージャーや研究評価の実施に携わるすべての人が新しい評価を計画したり、既存の評価を確認したりできるように設計された、実用的なステップバイステップのプロセスである。SCOPE は頭字語で、S は「価値あるものから始める」、C は「コンテキストを考慮する」、O は「評価のオプション」、P は「徹底的に調査する」、E は「評価を評価する」を表す。

⁹ SCOPE Framework for Research Evaluation ウェブサイト<<https://inorms.net/scope-framework-for-research-evaluation/>> [2025/3/21 取得]



図 2-2 SCOPE Framework

出所: INORMS, The SCOPE Framework – A five-stage process for evaluating research responsibly, 2021

<SCOPE の原則>

SCOPE の 5 つのステージは、次の 3 つの主要原則の下で運営されている。

(1) 必要な場合のみ評価する。

評価が常に正しい戦略とは限らない。例えば、インセンティブを与える行動に関しては評価するよりもそれを可能にする方が実り多いかもしれない。

(2) 被評価者とともに評価する。

どのような評価も、評価されるコミュニティによって共同設計され、共同解釈されるべきである。

(3) 評価の専門知識を活用する。

私たちは、学術適用するのと同じ厳密さを評価にも適用すべきである。

S : 自分が大切にしていることから始めよう

- ・評価対象企業について、何を評価するかを明確にすること
- ・他人の価値観（外的要因）ではなく
- ・利用可能なデータソースがない（「街灯効果」）

C : コンテキストに関する考察

- ・文脈に応じた評価を行う

- ・誰を評価しているか？(事業体の規模と分野)
- ・なぜ評価するのか？

O：評価オプション

- ・量的および質的オプションを検討する
- ・品質を示すために量を使用する場合は注意が必要である

P：深く掘り下げて調査する

- ・評価アプローチは誰を差別する可能性があるか？
- ・評価アプローチはどのように悪用される可能性があるか？
- ・予期せぬ結果として何が起こり得るか？
- ・評価の費用対効果を考慮する

E：評価を評価する

- ・評価の目的は達成されましたか？
- ・それは形成的なものだったのか、それとも総括的なものだったのか？
- ・SCOPE を使用して行う

また、SCOPE ワークショップのアプローチについて、評価デザインにワークショップのアプローチを取ることは、「被評価者とともに評価する」良い方法であるとしている。ワークショップは、被評価者と関連するステークホルダー・コミュニティの両方が、何を評価するか、どのようにするかについて発言できるようにする方法を提供する。そこで生み出される議論は、プロセスに対する信頼感を醸成し、進むべき最良の道についてのコンセンサスを得ることができる。しかし、ワークショップに参加できる人数には限りがあるため、オンライン調査によってワークショップを補う方法を選ぶ組織もある。アンケートは、誰もが評価デザインに意見を述べる機会を提供できる。また、「集団思考」の危険性も少ないと指摘している。

(5) More Than Our Rank イニシアティブ¹⁰

More Than Our Rank イニシアティブは、先述した国際研究管理学会ネットワーク (INORMS) の研究評価グループで進めている活動で、世界大学ランキングの問題点や影響に対応するために開発された。このイニシアティブは、学術機関がランキング順位に反映されていない世界への貢献の多様で多様な方法を強調する機会を提供している。これは、ランキング順位に誇りを持ちながらも、使用されている指標の限界を認識している機関、およびランキングが自分の強みや機関の使命を反映していないと感じている機関のためのイニシアティブである。

¹⁰ More Than Our Rank ウェブサイト<<https://inorms.net/more-than-our-rank/>> [2025/3/21 取得]

早期導入した機関（Early Adopter Institutions）は、キール大学、クイーンズランド工科大学（QUT）、Loughborough 大学などがある。署名した大学等はこれらを含めて 20 以上ある。また、支援団体としては、ラテンアメリカ研究評価フォーラム（FOLEC-CLACSO）、コインブラ・グループ、ライデン大学の科学技術研究センター（CWTS）、DORA、欧州大学協会（EUA）、サイエンス・ヨーロッパなどがある。

参加を希望する機関は、Web サイトからサインアップし、More Than Our Rank のロゴを Web ページまたは販促資料に追加するだけであり、参加無料となっている。学術機関は、国内または国際的な大学ランキングでは十分に反映されていない機関の活動、成果、または野望を宣伝する声明を掲載するよう求められる。



図 2-3 More Than Our Rank のロゴ

出所: More Than Our Rank イニシアチブウェブサイト

例えば豪州のスウィンバーン大学¹¹においては、成功の補完的な尺度として、成功とは、支援的なデュアルセクター教育環境を育み、研究成果への多様な貢献を称え、業界のパートナーやコミュニティと有意義に協力するという私たちの取り組みを反映するものとしている。

ランキングでは、スウィンバーン大学がビクトリア州で唯一、子供大学（子供たちが教室外で学び、学業成績を上げる機会を増やす国際的な取り組み）を提供している大学であるという事実を反映できていない。公平性と卓越性のための奨学金や、包括的で歓迎的な AccessAbility サービスを通じて学生の生活をどのように変えているかを認識していない。また、地元のコミュニティと国を結びつけるために、先住民の学生とスタッフと協力してキャンパスを設計していることも認識していないと主張している。

加えて、ランキングには、スウィンバーンの画期的な研究がいかに重要で現実世界に影響を与えているかも反映されておらず、引用数だけでは、スウィンバーン大学の研究が、生活を向上させ、一般の人々の関心を引き、公平性、多様性、包摂性を向上させて、すべての人

¹¹ More Than Our Rank <<https://www.swinburne.edu.au/about/our-university/rankings-ratings/more-than-our-rank/>> [2025/3/21 取得]

にとってより良い結果をもたらす政策にどのように役立っているかを示すことはできないとしている。そして、スウィンバーン大学における現在の研究は次のように紹介している。

- 「Mental Health Online」イニシアティブを通じて、何万人もの人々が無料のオンラインメンタルヘルスサービスや無料のセラピーにアクセスできるよう支援している。
- 「未来の工場」の製造、デジタル、商業の専門家のおかげで、何百ものオーストラリアの中小企業が将来を見据えた事業を展開し、高度な製造業を取り入れ、インダストリー4.0に対応できるよう支援している。
- ヨーグルトから種子まで、実験物を国際宇宙ステーションに送り返すという宇宙技術産業研究所のプログラムを通じて、何百人もの学齢期の子供たちに STEM や宇宙産業への進出を促している。
- デジタル技術とサイバーの安全性に重点を置き、特に女性と子供に対する、テクノロジーによる家庭内暴力や家族内暴力を理解し、防止するために取り組んでいる。
- スウィンバーン神経画像診断施設では、あらゆるキャリア段階のスポーツクラブやアスリートが接触スポーツによる脳震盪のリスクを理解できるよう支援している。

同様に、ランキングにはスウィンバーン大学の数え切れないほどの「初めて」が反映されていない。例えば、オーストラリアで初めて主任科学者を任命した大学であり、学生に職業統合学習を保証した最初の大学であり、1992年には最先端のスーパーコンピューターを導入した最初の大学であるとも指摘している。また同様に、ランキングでは、スウィンバーン大学の先駆的な大学モデルのプラスの効果を捉えることはできていない。この大学モデルは、業界、地域社会、政府、他の大学との強力なパートナーシップの上に構築され、イノベーションと起業家精神を当大学の DNA に組み込んでいるとしている。ランキングは、フランス・オーストラリアエネルギー転換センターやハワイの WM ケック天文台との歴史的なパートナーシップなど、世界をリードする科学的協力関係を認めていない。また、ライセンス使用料からスピンアウト企業の株式、業界向けの短期コース、スウィンバーンが設立したスタートアップ企業まで、研究の商業化の成功を把握することもできていないと主張している。

(6) Leiden Manifesto (ライデン声明)¹²

論文の被引用数等の計量データは、適切に利用されれば専門家（ピア）による評定をより妥当、公正にするための補完となり得るが、データに主導された評価や、指標の意味・性質の不十分な理解による誤用がしばしば見られる。このような状況に対して、研究評価における計量データの利用についてのベストプラクティスを示した「研究計量に関するライデン声明」（“The Leiden Manifesto for research metrics”）が、2015年に Nature 誌上で公表された。

¹² 小野寺ら [1]

ライデン声明は 10 項目の原則 (principles) から成り、研究評価における計量データの利用についてのベストプラクティスや注意点を示したものであり、研究者、管理者、評価者の全てに対する、計量データに立脚した研究評価のガイドラインと考えられる。

ライデン声明の基礎となったのは、2014 年 9 月にオランダのライデン大学で開催された 19th International Conference on Science and Technology Indicators (STI 2014) において、Dr. Diana Hicks (Georgia Institute of Technology) が行った基調講演である。ここで彼女は、研究者、研究機関、研究プログラムの評価への計量データの使用に関して次の七つの原則を示し、Nature や Science のようなトップレベルの雑誌でこれを公表することを提案した。

- ①計量は評価の代替物ではない。
- ②高品質のデータを得るために時間と金を費やせ。
- ③計量は透明かつ受け入れやすいものでなければならない。
- ④データは被評価者により確認される必要がある。
- ⑤研究分野による違いに敏感であれ。
- ⑥分野と時期による違いを考慮してデータを規格化せよ。
- ⑦計量は戦略的目標と連携すべきである。

このスピーチに対して活発な議論がなされた。計量データ・指標の責任ある利用のガイドラインとなる原則を科学計量学コミュニティが共同して発表すべきであるという多くの意見があり、研究評価のための計量データの適正な利用のために科学計量学の研究者は積極的役割を果たすべきという声も上がった。その結果、この Hicks の 7 原則を基にして、ライデン大学科学技術研究センター (CWTS) の Dr. Paul Wouters が中心となって、このコミュニティが合意できるマニフェストをまとめることとなった。こうしてまとめられたのが、Nature 誌で Hicks、Wouters ら 5 名の著者により発表されたライデン声明である。

<ライデン声明-10 の原則>

原則 1 定量的評価は、専門家による定性的評価の支援に用いるべきである。

定量的計量は、ピアレビューで生じやすいバイアスについて異なる見方を提示し、考察を深めるのに役立つ。同業研究者について判定することは広範な関連情報なしには難しいので、これによりピアレビューは強化されるはずである。しかしながら、評定者は意思決定を数字に任せてはならない。指標は情報に基づく判定を代替してはならない。評定者はそれぞれが行う評定に責任を保持している。

原則 2 機関、グループ又は研究者の研究目的に照らして業績を測定せよ。

プログラムの目標はその開始時に明示されるべきであり、また、業績を評価する指標は、それらの目標と明確に関係づけるべきである。指標の選択やその活用に際しては、より幅広い社会経済的及び文化的な状況を考慮すべきである。科学者の研究目的は様々である。学術的知識の最前線を進める研究と、社会的問題の解決を目指す研究とは目的が異なる。

学術的なアイデアの卓越性よりも、政策、産業、あるいは公衆への貢献に基づく評価もある。すべての状況に適用できる単一の評価モデルはない。

原則 3 優れた地域的研究を保護せよ。

世界の多くの地域で、優れた研究は英語で発表されると見なされている。たとえば、スペインの法律は、同国の学者が高インパクトの雑誌に発表することを望ましいとしている。インパクトファクターは、米国中心で、いまだにほとんどが英語である Web of Science 収録の雑誌を対象に計算されている。こうしたバイアスは、国・地域についての研究が多い人文・社会科学において特に問題が大きい。他の多く分野でも、国・地域という側面を持つ。例えば、サハラ以南アフリカにおける HIV の疫学などの例がある。

しかし、このような多元性や社会的関連性は、高インパクトのゲートキーパーたる英語雑誌の関心を得るような論文を創出するために抑制される傾向がある。Web of Science で高引用を得ているスペインの社会学者たちは、抽象モデルに長年取り組んでいるか、米国データの研究を行っている。高インパクトのスペイン語論文では、地域の労働法、高齢者のための家族健康管理、移民の雇用などのトピックについての社会学者の独自性が失われている。優れた地域的研究の発見・それらへの報奨の付与のためには、高品質の非英語文献に基づいた計量が有用であろう。

原則 4 データ収集と分析のプロセスをオープン、透明、かつ単純に保て。

評価のために要求されるデータベースの構成は、明確に表現された規則に従い、研究が終了する前に設定されるべきである。これは、数十年にわたり計量書誌学的評価の方法論を確立してきた学術グループと商業グループに共通の経験である。これらのグループは、査読論文に公表されたプロトコルを参考としてきた。この透明性は精密な検討を可能とした。たとえば、2010 年に、我々のグループのひとつ（ライデン大学の科学技術研究センター(CWTS)) が用いていた重要な指標の技術的性質について公開の討論が行われ、この指標の計算法の改訂に結び付いた。最近参入している商業グループも同様な標準に従うべきである。また、ブラックボックスの評価マシンを受入れるべきではない。

指標が単純であることは、その透明性を増すことであり長所である。しかし、単純化した計量は記録を歪めることもある（原則 7 参照）。評価者は、バランス（研究過程の複雑性に忠実である単純な指標）を得ることに努めなければならない。

原則 5 被評価者がデータと分析過程を確認できるようにすべきである。

データの品質を確かなものにするため、計量書誌学的調査の対象となるすべての研究者が、自分の成果が正確に同定されていることをチェックできるようにすべきである。評価過程の指揮・管理者はすべて、自己確認又は第三者の検査によりデータの正確性を保証すべきである。大学は、その研究情報システムの中にこれを実装することができるだろうし、それは、これらのシステムの提供者の選択の指針であるべきである。正確で高品質なデータの照合・処理には時間と資金を要する。そのための予算を惜しんではならない。

原則 6 分野により発表と引用の慣行は異なることに留意せよ。

ベストプラクティスは、一揃いの指標候補を選び、分野によってその中から選択できるようにすることである。数年前のことだが、欧州のある歴史学者のグループが、その国のピアレビュー評価において比較的低い評点を得たことがあったが、それは、このグループが、Web of Science に収録される雑誌よりもむしろ図書に成果を発表しているためであった。この歴史学者は不運なことに心理学の学科に属していた[歴史学者が心理学の学科に属していたため、雑誌論文によってピアレビュー評価がなされたという意味だと思われる]。歴史学者や社会学者は、成果のカウントに際して図書や自国語の論文が含まれることを要求するし、計算科学者は会議論文がカウントされることを要求する。

分野により引用傾向は異なる。トップにランクされる雑誌のインパクトファクターは、数学ではおよそ 3、細胞生物学ではおよそ 30 である。[この差を埋めるための]規格化した指標が必要である。最も頑健な規格化法はパーセンタイルに基づくものであり、各論文は、それが属する分野の被引用数分布中のパーセンタイル位置(たとえばトップ 1%、10%、20%) に従って重み付けされる。非常によく引用される論文 1 件は、パーセンタイル指標に基づくランキングでは、大学の位置を僅かに上げる程度だが、平均被引用数に基づくランキングでは、中位から一挙にトップまで押し上げることがあり得る。

原則 7 個々の研究者の評価は、そのポートフォリオの定性的判定に基づくべきである。

h 指数は、新しい論文がなくても年齢を重ねるほど高くなる。h 指数は分野によっても異なる。トップレベルの研究者の場合、生物学では 200、物理学では 100、社会科学では 20-30 程度である。

この値は、[h 指数の計算に使う]データベースにも依存する。計算科学分野では、Web of Science では h 指数が 10 前後であるが、Google Scholar では 20-30 である研究者がいる。研究者の成果物を読んで判定する方が、一つの数字に頼るよりもずっと適切である。多数の研究者を比較する場合でも、個々の専門性、経験、活動及び影響に関するより多くの情報を考慮するやり方が最良である。

原則 8 不適切な具体性や誤った精緻性を避けよ。

科学技術指標は、その概念が曖昧で不確かになりがちであり、また、普遍的には受入れられない強い仮定に立っていることがある。たとえば、被引用数の意味も長らく論争されてきている。従って、ベストプラクティスは、より頑健で複眼的な描像を与えるように複数の指標を用いることである。もし不確かさや誤差が定量化出来るのであれば(たとえばエラーバーの形で)、その情報を公表される指標値とともに示すべきである。それができない場合、指標の作成者は少なくとも誤った精緻性を避けるべきである。たとえば、[Journal Citation Reports では]インパクトファクターを小数点以下 3 桁まで表示して同点の雑誌の出現を避けるようにしている。しかし、被引用数の概念上の曖昧さやランダムな変動性を考慮すれば、このような僅かなインパクトファクターの差によって雑誌を区別する意味はない。誤った精緻性は避けよ。小数点以下 1 桁で十分である。

原則 9 評定と指標のシステム全体への効果を認識せよ。

指標は、それがもたらすインセンティブによってシステムを変化させる。これらの効果を予期しなければならない。このことは、一揃いの指標を用いることが常に望ましいことを意味する。単一の指標は、ゲーム化や目標の取り違えを招く（指標の測定自体が目標になる）。たとえば、1990年代のオーストラリアでは、機関からの発表論文数に大きく依拠する数式を使って大学の研究への資金配分を行った。大学は査読制雑誌の1論文あたりの「価値」を計算することができた。2000年時点でのその価値は800豪ドル（当時のレートで約480米ドル）の研究資金に相当した。予想されたように、オーストラリアの研究者が発表する論文数は増加したが、それらは被引用数の低い雑誌に集中し、論文の質の低下を示唆した。

原則 10 指標を定期的に吟味し、改善せよ。

研究の目的と評定の目標は変化し、それに伴って研究システム自体も共進化する。かつて有用であった計量が不適切になり、新しいものが現れる。指標のシステムも見直しが必要であり、適時修正しなければならない。[原則 9 で述べた]単純な数式の影響に気付いて、オーストラリアは2010年に、より複雑で質の面を強調した **Excellence in Research for Australia** イニシアティブを導入した。

(7) The Metrics Tide¹³

本書は、「研究評価と管理におけるメトリクスの役割に関する独立レビュー」の主な所見と提言をまとめたものである。レビューの議長は英国のサセックス大学科学政策研究ユニット (SPRU) のジェームズ・ウィルズドン教授が務め、サイエントロメトリクス、研究助成、研究政策、出版、大学経営、研究管理の専門家からなる独立した学際的なグループがサポートした。また、HEFCEの研究政策チームが運営グループの事務局を担当している。

このレビューでは、先行研究の枠を超え、研究評価指標と指標の潜在的な用途と限界について深く考察している。様々な学問領域におけるメトリクスの利用を検討し、研究の卓越性とインパクトの発展に対するメトリクスの潜在的な貢献を評価した。また、REF (Research Excellence Framework) の次期サイクルを含む、研究評価のプロセスにおけるメトリクスの役割についても分析している。大学が管理システムにおいて定量的指標を用いる方法の変化や、リーグテーブルやランキングの力の増大についても考察。そして、研究文化の様々な側面における、測定基準の否定的な、あるいは意図しない影響についても考察した。

本報告書では、まず英国内外の研究管理・評価におけるメトリクスの歴史をたどっている。異なる研究文化におけるメトリクスの適用可能性を検討し、査読システムとメトリクスの代替案を比較し、両者の間でどのようなバランスを取るべきかを考察している。また、研究機関における研究管理システムの発展過程を描き、定量的指標の利用が、業績を含む研究文

¹³ Wilsdon, J., et al. [2]

化のさまざまな側面に及ぼす影響についても検証している。マネジメント、平等性、多様性、学際性、評価システムの「ゲーム性」などである。このレビューでは、様々な資金提供者が定量的指標をどのように利用しているかを調べ、研究・イノベーション政策における定量的指標の潜在的な役割について考察している。最後に、REF2014で評価指標が果たした役割を検証し、今後の評価貢献についてのシナリオを概説している。

このレビューは、その所見と結論を導き出すために、多様な証拠に基づいている。これには、正式な証拠募集、文献の包括的なレビュー（補足報告書 I として出版）、フォーカスグループ、ワークショップ、伝統的および新しいメディアを通じた利害関係者との広範な協議が含まれる。レビューでは、HEFCEによるREF2014の最近の評価も参考にし、REF2014の得点と評価指標のバスケットとの相関関係についての独自の詳細な分析も依頼している（補足報告書 II として公表）。

このレビューでは、英国の研究システム全体の利害関係者によるさらなる作業と行動のために、20の具体的な提言を特定している。これらの提言は、本レポートで収集され、提示されたエビデンスに基づくものであり、研究ガバナンス、管理、評価を強化するための、より広範な試みの一部とみなすべきものである。提言は以下の通りであり括弧内は対象者である。

<研究文化の効果的なリーダーシップ、ガバナンス、マネジメントの支援>

1. 研究コミュニティは、定量的指標の貢献と限界について、より洗練されたニュアンスのあるアプローチを開発すべきである。
2. 機関レベルでは、HEIのリーダーは、定量的指標の役割を含め、研究管理と評価へのアプローチに関する明確な原則を策定すべきである。
3. 研究マネージャーと管理者は、これらの原則と責任ある評価基準の使用を機関内で支持すべきである。
4. 高等教育機関の人事マネージャーや採用・昇進委員会は、教員の任用・昇進の決定に用いる基準を明確にすべきである
5. 個々の研究者は、自身の経歴を発表したり同僚の研究を評価したりする際に、特定の指標の限界に留意すべきである。
6. 高被引用度論文産出群と同様に、研究助成機関も、研究評価・管理において定量的指標を利用するための、それぞれの状況に応じた原則を策定し、それらが十分に周知され、見つけやすく、理解しやすいようにすべきである。
7. 大学ランキングやリーグテーブルのデータ提供者、分析者、制作者は、異なる測定システム間の透明性と相互運用性の向上に努めるべきである。
8. 出版社は、販促ツールとしてのジャーナル・インパクト・ファクターの強調を減らし、業績のより豊かな見方を提供する、ジャーナルに基づく様々な指標との関連においてのみ、インパクト・ファクターを使用すべきである。

<研究情報管理を支えるデータインフラの改善>

9. 研究データ基盤の透明性と開放性を高める必要がある。オープンで信頼できる研究情報管理をサポートする技術、実務、文化に関する一連の原則を策定すべきである。
10. 英国の研究システムは、ORCIDを一意的識別子の好ましいシステムとして最大限に活用すべきである。次回のREFでは、すべての研究者にORCID iDを義務付けるべきである。
11. ISNIはすでに出版社、研究助成機関、研究組織を十分にカバーしている。したがって、ISNIの使用は、将来のREF提出で参照されるすべての機関をカバーするように拡張されるべきであり、HEI内部および資金提供者の管理より広く使用されるべきである。
12. 出版社は、論文投稿の際にORCIDのiDやISNI、助成金の参照先を義務付け、出版ライフサイクルを通じてこのメタデータを保持すべきである。
13. デジタルオブジェクト識別子(DOI)の使用は、すべての研究成果をカバーするために拡張されるべきである。
14. 研究情報インフラへのさらなる投資が必要である。資金提供者とJiscは、特に研究管理システムの相互運用性を向上させるための戦略的追加投資の機会を探るべきである。

<既存のデータや情報源の有用性を高める>

15. HEFCE、資金提供者、高等教育機関、Jiscは、REFのプロセスを支援するために既存のプラットフォームで保有されているデータをどのように活用するか、またその逆も然りである。
16. BISは、研究関連プラットフォーム(Gateway to Research、Researchfish、REFを含む)から収集したデータを、BISや他省庁の政策プロセス、特に先見性(foresight)に関する政策プロセスに、より直接的にリンクさせる方法を特定すべきである。

<次期REFでのメトリクスの活用>

17. 次回のREFサイクルについて、我々はHEFCEおよび他の高等教育財政機関に対し、以下のような具体的な提言を行う。
 - a. アウトプットの評価においては、定量的なデータ、特に公表されたアウトプットに関する研究の質に関する査読の判断材料として引き続き活用することを推奨する。
 - b. インパクトの評価にあたっては、HEFCEと英国の高等教育助成機関は、REF2014のインパクト事例研究の分析を基に、今後のインパクト事例研究における定量的指標の使用に関する明確なガイドラインを策定することを推奨する。
 - c. 研究環境の評価においては、定量的データの強化する余地があるが、これらのデータには、その解釈を可能にする十分な文脈が提供される必要があることを提言する。

<活動の調整とエビデンスの構築>

18. 英国の研究コミュニティは、本報告書で示されたアジェンダを推進するための仕組みを必要としている。研究助成機関、高等教育機関およびその代表機関、出版社、データプロバイダー、その他を集め、データ標準、相互運用性、公開性、透明性の問題に取り組む「責任ある測定基準のためのフォーラム (Forum for Responsible Metrics)」の設立を提案する。
19. 研究助成機関は、科学政策の科学への投資を増やす必要がある。
20. このレビューの肯定的な側面のひとつは議論が巻き起こったことである。レガシー・イニシアチブとして、運営グループは本報告書で提起された問題を継続的に議論する場として、ブログ (www.ResponsibleMetrics.org) を立ち上げている。

2.1.2 国際動向の比較

研究評価の改革をめぐる議論は、異なる包括的な評価基準の必要性、ピアレビューの役割、オープンサイエンスの活用など、評価のさまざまな側面に焦点を当てている。また、ジャーナルの評価基準に重点を置くのではなく、共同研究、データ共有、コミュニティへの参画を含む、より包括的で質的な研究インパクトの評価へとシフトする必要性を指摘する声もある。

InterAcademy Partnership (IAP)¹⁴、International Science Council (ISC)¹⁵、GYA and the National Young Academies¹⁶の3団体が共同で発行した「THE FUTURE OF RESEARCH EVALUATION: A SYNTHESIS OF CURRENT DEBATES AND DEVELOPMENTS」(Rijcke et al. [3])は、研究評価制度の現状をレビューし、世界各地のいくつかの事例を通して、さまざまなステークホルダーが取った行動、対応、取組について論じている。

¹⁴ IAPは150の国、地域、および世界のアカデミーが加盟しており、100カ国以上の主要な科学者、技術者、医療専門家を代表している。

¹⁵ ISCは200以上の国際科学連合および協会、ならびに各国および地域の科学組織(アカデミー、研究機関、財団など)からなる多様な機関ネットワーク。

¹⁶ GYA and the National Young Academiesは、あらゆる分野から選出された200名の若手科学者で、博士号取得後、通常3~10年が経過した30~40歳。会員および同窓会メンバーは、合わせて94カ国から参加し。また、GYAは50を超えるナショナル・ヤング・アカデミー(NYA)およびその他の若手研究者ネットワークへの窓口でもある。GYAはIAPの正会員であり、ISCの準会員でもある。

The Future of Research Evaluation

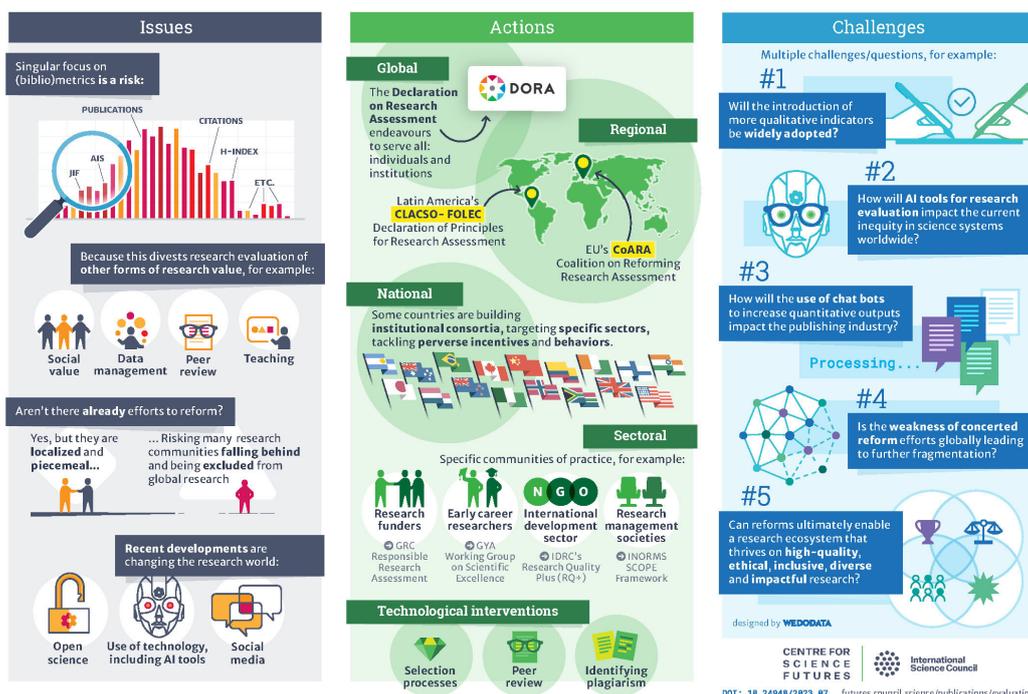


図 2-4 The Future of Research Evaluation 概要

出所: Rijke et al. [3]

当報告では、地域の展望と発展の観点より、ほとんど定量的な評価システムによって生み出される問題は、「北半球」の視点から見られ、診断されることがほとんどであり、「南半球」はキャッチアップの危機にさらされている。一般化しすぎる危険を冒してでも、「北半球」には、多様性、公平性、インクルージョンの欠如をめぐる大きな制度的問題があり、それが評価制度によって悪化していると指摘している。グローバル・サウスでは、何が「質」と「インパクト」を構成するのかについて、地域的・地方的な定義が欠如しており、評価システムは（同じ学部間でさえ）大きく異なり、現状に対する挑戦は比較的少ない。世界的に見れば、定量的な指標を過度に重視すること、評価と資源配分の関連性、競争の激しい研究助成制度と出版へのプレッシャー、研究や学術生活の定量化しにくい他の側面を軽視することなどが、問題につながっているとしているとし、以下のように地位別に概要を示している。

(1) ヨーロッパ

2022年7月に承認されたEU研究評価推進連合（CoARA）は、研究評価改革に関する世界最大のイニシアティブである。4年がかりで、40カ国（主にヨーロッパ）の350の組織が開発したこのイニシアティブは、欧州大学協会と Science Europe（欧州の科学助成機関

および科学アカデミーのネットワーク)が欧州委員会と協調して、より包括的で責任ある研究評価のための合意または一連の原則(「改革の旅」)を策定したものである。この合意は、機関、研究者個人、研究そのものの3つの評価レベルに焦点を当てている。欧州のパートナーによって管理されているものの、この連合は世界的なものになることを目指しており、DORAとGYAはすでに署名している。署名者は、研究評価の改善、評価のための新しい基準とツールの開発、研究評価に関する意識向上とトレーニング(ピアレビューアーなど)にリソースを投入することを約束している。

EUは、研究評価の改革を支援するいくつかの興味深い新たな取り組みにも資金援助を行っている。特に、複数の研究プロセスと成果物にわたる包括的な指標群を開発し、それによって欧州の研究者にオープンサイエンスの実践を促すことを目的とした「オープン・ユニバーサル・サイエンス(OPUS)」、および、研究評価のための政策改革を支援するオープン・データスペースを構築する「オープンサイエンス評価データスペース GraspOS」などである。欧州研究会議(ERC)は、あらゆる分野における先端研究を支援する機関であり(2021年から2027年の予算は160億ユーロ)、CoARAに署名し、研究コミュニティへの「並外れた貢献」や、より非伝統的なキャリアパスを考慮した説明を盛り込むために、評価フォームとプロセスを修正した。提案は、申請者の過去の業績よりもその価値に基づいて評価されるようになり、今後も一流の学者で構成されるピアレビューパネルが科学的卓越性という唯一の基準に基づいて評価を継続している。

ヨーロッパのいくつかの学術団体も参加している。ヨーロッパ40カ国に50以上ある各国アカデミーのうち9つを代表する、ヨーロッパ人文・自然科学アカデミー連合(ALLEA)理事会は、CoARAの取り組みを支持している。ALLEAは、新しいアカデミー会員の受け入れに関する優れた慣行を収集、交換、推進する専門のタスクフォースを設立し、質、完全性、多様性、開放性の原則に基づき、研究評価の「有意義な文化交流」に貢献することを目的としている。欧州の研究重視型大学もまた、「多面的」な研究キャリアの道筋として研究評価の改革を支援している。大学が研究、教育、社会への貢献の多様性を認識するよう鼓舞し、支援するための共通の枠組みを開発している。

国家レベルでは、現在、複数の国々でさまざまな評価モデルの試験運用が行われている。例えば、ベルギー、オランダ、スイス、英国の国家研究助成機関はすべて「ナラティブ履歴書(Narrative CVs)」を使用している。ナラティブ履歴書は、学術的業績をより総合的に評価するもので、知識の創出、個人の成長、より幅広い研究コミュニティ、より幅広い社会への貢献を評価する。このような履歴書に対する支持が広がる一方で、学術界に万能を強いることになり、万能の地位を追求するあまり、深い専門性が損なわれるリスクがあるという懸念もある。

(2) 北米

北米では、オープンサイエンスのアジェンダによって加速された、純粋に定量的な指標からのシフトが進行している。オープンサイエンスとオープンレビューは、評価の実践をより

透明化し、自己反省の機会を提供し、雇用、昇進、査読委員会における自己引用や縁故主義、生得的なジェンダーやその他のバイアスなどの問題を表面化するのに役立っている。基礎科学（知識の向上）と応用科学（社会的インパクト）に貢献するハイブリッドで収束的な評価モデルの可能性もあり、よりスマートでインテリジェントな指標や混合評価手法を開発する必要性について議論が続いている。また、大学には、「先行者不利」に陥ることなく、現在評価のために使用しているツールから離脱するためのアカデミックな空間と自由が必要であり、知識の有用性、その取り込み、インパクトを測定するために、ユーザーコミュニティも評価プロセスの一部であるべきだという認識もある。しかし、研究エコシステムの上層部から下層部からも、つまり現状から利益を得ている人たちからも、最近そのエコシステムに参入した人たちからも、変化に対する偶発的な抵抗（「意志的盲目」）も存在する。米国で DORA に署名した大学はほとんどなく、新しい DORA プロジェクトは、なぜそうなのかを理解しようと努めている。

なお、米国およびカナダにおける対応状況については、2.1.3 G7 各国における対応状況を参照されたい。

(3) アジア太平洋

競争が激しく、定量的な指標に基づく評価システムがこの地域を支配しており、英語圏の国々が評価の枠組みを形成し、他の国々がそれに追随しているのが一般的である。例えばオーストラリアでは、ピブリオメトリクスと大学ランキングに基づく競争的資金調達システムがある。同様の課題はマレーシアやタイにもあり、他の ASEAN 諸国も追随する可能性が高い。重要な例外は中国であり、そこでは政府がシステム変革に重要な役割を果たしている。勇気づけられることに、この地域の研究コミュニティの間では、現在の研究評価システムの限界と、研究の完全性に対する脅威に対する認識と懸念が高まっている。ナショナル・ヤング・アカデミーを含む ECR（Early career researcher）や、若手科学者の ASEAN ネットワークは、草の根運動とともに、この問題への関与を強めているが、声はなかなか届かない。大学の指導者を含む政府や資金提供者コミュニティは、この議論にほとんど参加していない。実際、相談者の報告によると、量的基準がさらに追加され、研究機関や研究者が制度を利用し始め、研究不正を助長している。

しかし、変革の大きなチャンスもあり、例えばインドでは、科学技術省の政策研究センター（DST-CPR）は、インドにおける研究評価とその改革に関する最近の研究を実施し、主要関係者（国の資金提供機関、研究機関、アカデミー）とのワークショップ、インタビュー、調査を行った。その結果、大学や多くの国家政策上重要な研究機関（農業など）が、ほとんど定量的な評価基準のみに重点を置いている一方で、一部の資金提供機関やインド工科大学のような機関では、より定性的な評価基準も採用していることがわかった。

(4) ラテンアメリカとカリブ海諸国

ラテンアメリカとカリブ海諸国 (LAC) は、他の地域とは多くの点で対照的である。ラテンアメリカ・カリブ海地域では、科学は世界的な公共財であると考えられており、研究・学術出版制度やインフラは公有（資金提供）かつ非商業的である。変化をもたらすことができる主なステークホルダーは、各国の研究評議会、科学省、主要研究大学であり、研究者の60%以上が大学に所属していることを考えると、高等教育機関の役割は極めて重要である。評価制度を SDGs や、この地域で盛んなオープンサイエンスや市民科学の動きとより密接に連携させる可能性がある。研究評価制度は国、地域、機関ごとに分断化が進んでおり、研究は教育、普及、共同研究など他の機能と競合している。中南米カリブ地域の研究評価と研究者表彰制度は一般的に、学術雑誌のインパクトファクターと大学ランキングのみに基づく「北半球」の方法論に根ざした卓越性の概念を支持している。さまざまな形態の知識生産とコミュニケーション、そして学術的キャリアの多様性（教育、研修、指導、市民科学、科学の公共コミュニケーションなど）の認識は、研究評価の慣行にはほとんど見られない。

しかし、より包括的な研究評価スキームへの移行に対処するため、この地域は研究評価の原則とガイドラインを採択した。2022年6月に承認された CLACSO-FOLEC 研究評価原則宣言は、質の高い社会的意義のある科学を保証し保護することを目的としており、DORA とオープンサイエンス、研究成果と研究キャリアの多様性、地域ジャーナルと索引サービスの価値、学際性、地域言語、先住民の原則を包含している。220 を超える賛同者がおり、責任ある研究評価や改革の例において、すでに前向きな傾向が見られる。

(5) アフリカ

アフリカにおける研究奨励金や報奨金制度は、「国際的な」、主に欧米の規範や慣例を反映する傾向がある。アフリカの研究機関は、研究の「質」と「卓越性」に対するアプローチを開発する際、これらに従おうと努めているが、必ずしも現地の知識やニーズに適しているとは限らない。研究の「質」、「卓越性」、「インパクト」は、アフリカ大陸では十分に定義されておらず、「研究のインパクト」という文化に慣れていない研究者もいる。アフリカの評価システムは、社会的利益のための研究、教育、能力開発、研究管理・運営を考慮しない傾向がある。出版モデルは文脈に敏感でなく、APC (Author processing charge) はアフリカの研究成果に対する障壁となっている。研究評価システムの改革は、アフリカの研究が社会的課題に貢献できるという非対称性の是正に役立つだけでなく、アフリカの研究コミュニティの活動を支援するリソースへのアクセスを改善することができる。分野横断的・学際的協力の障壁を取り除くことは、次のような課題解決に不可欠である。

アフリカ大陸では、RRA (Responsible Research Assessment) を中心に強力なパートナーシップが構築されつつある。開発機関の国際コンソーシアムから資金提供を受けた Science Granting Councils Initiative (SGCI) は、アフリカ 17 カ国を巻き込み、グローバル・サウスの視点から科学助成機関と研究者評価を調査し、アフリカにおける研究の性に関する研究を実施した。また、2022 年の世界科学フォーラムでは、SGCI と GRC (Global

Research Council) の後援の下、南アフリカの国立研究財団 (NRF) と科学イノベーション省が国際的および地元のパートナーを招集し、RRA の推進における資金提供機関の役割について議論し、経験を共有し、優れた実践を進め、能力構築と協力の進捗状況を評価している。国家レベルでは、漸進的な変化が始まっている。研究助成機関が主導している他の国には、タンザニア (COSTECH)、モザンビーク (FNI)、ブルキナファソ (FONRID) などがある。GRC の RRA イニシアティブは、大陸における変革のための重要なプラットフォームであることが証明されている。また、国際開発セクターからの学習、特に IDRC (International Development Research Centre) の研究品質プラス (RQ+) 評価フレームワークは、すでに適用され、研究され、改善されているという特徴がある。

2.1.3 G7 各国における対応状況

(1) アメリカ合衆国 (米国) ¹⁷

米国では、全米科学財団 (NSF) が「社会における研究のインパクトの促進」プログラムと、研究者および評価者向けのより広範なインパクトを与えるツールキットを通じて、変革の主導的な役割を担っている。先住民や伝統的に疎外されてきたコミュニティの参加を含む、公平性、多様性、包括性が重要な推進要因となっている。IAP および ISC のメンバーである米国科学アカデミーは、幅広い改革を促進することを目指しており、従来の研究者の履歴書 (CV) の改革に関する情報交換や学習のプラットフォームを提供している (NAS Strategic Council, 2022)。米国アカデミーの取り組みから生まれた「高等教育におけるオープンな研究を推進するリーダーシップ・イニシアティブ (Higher Education Leadership Initiative for Open Scholarship)」は、オープンな研究を推進するための集団行動を誓う 60 以上の大学・カレッジの連合体であり、研究評価の再考によるオープン性と透明性の向上もその取り組みに含まれている。例えば、米国国立衛生研究院 (NIH) は、助成金申請を行う研究者のバイオ・スケッチ (SciENcv) を新たに設計し、系統的な偏見や報告の負担を最小限に抑え、同時に影響力重視の姿勢を打ち出した。

(2) ドイツ連邦共和国 (独国)

ドイツでは、一部の特定研究プロジェクトを除き、州政府が教育・研究の権限を有していたが、2014 年以降、ドイツ連邦基本法 (憲法) の改正により、連邦政府が州政府との研究協定を締結し、大学機関に対して直接投資を行うことが可能となった。一方で、連邦政府が直接投資を行うプログラムの多くは、研究拠点や研究インフラに関するものが中心であり、本調査で対象としている研究開発プログラムは、現状の連邦政府の所掌範囲の中心から外れているところである。このため、国レベルで研究開発評価を体系的に整理する枠組みはなく、各プログラムの特徴に応じた評価を実施している (例えば、2006 年から 2017 年まで

¹⁷ Rijke et al. [3]

実施されたエクセレンス・イニシアティブでは、大学に対する資金提供を実施したプログラムであり、合同科学会議（GWK）は、その評価は独立した国際的な専門家委員会に評価させた）。

このような背景の下、国際動向に対するドイツの動きは、機関レベル、草の根レベルでのものとなっている。近年の動きとしては、2022年に欧州委員会と利害関係者が研究者評価のための新たなガイドラインに関する合意案に署名する準備を進める中で、ドイツの研究機関が欧州における研究評価の中央集権的な改革に警告を発する、というものがあつた。たとえば、ドイツ研究振興協会（DFG）は発表した立場表明書の中で、欧州における研究評価制度のトップダウン改革に反対すると述べている。また、その2週間前には、DFGを含む10のドイツ研究機関の連合が、提案された改革が拘束力を持つものとなった場合、EU加盟国、利害関係者、欧州委員会が共同で作成した協定案に署名しないと警告している¹⁸。

DFGの生命科学部門の責任者であるトビアス・グリム氏は、研究評価の改革はドイツで現在も議論が続いているが、EUの合意に署名した大学ネットワークはごくわずかであり、研究機関は「人々が操作しようとしている」出版指標にあまり頼るべきではないと発言している。また、DFGが査読システムを変更しようとしていること、DORA宣言に署名したこと、そして間もなくプロジェクトの内容により重点を置いた新しい履歴書テンプレートと研究者の応募ガイドラインを発表することを指摘している¹⁹。

また、ドイツ学長会議においては、欧州委員会による欧州における研究評価方法の変更の取組はトップダウン的すぎると指摘している。ドイツ学長会議は、欧州における研究の評価方法を再考することに価値があると考えているが、議論は十分に独立していないと捉えている。「合意自体は、欧州委員会の政策方針に沿って設計され、概念化されたものであり、その根底にあるのは、基本的にトップダウンのプロセスの結果である」と、学長会議の研究・学術キャリアパス・移転担当副会長のウォルター・ローゼンタール氏は語っている。この改革プロセスは、単一の欧州研究圏（ERA）を創設する動きの一環として、欧州委員会によって2022年初頭に開始されたものである。この取組は、EUの研究エコシステムのさまざまな部分で反発に直面しながらも強力な支持を得てきたが、研究評価推進連合（CoARA）が設立され、議論を継続し、改革に取り組むという次の段階に進んだが、ドイツ学長会議は、2023年初頭の段階では、この連合は政治的すぎる機構だとして参加しないことを決定した。ドイツの学長たちは、この改革は、研究の定量的評価から、社会へのインパクトなどの観点から見た研究の価値に対するより総合的な見方へと焦点を移すものとして位置付けられているが、ローゼンタール氏は、この合意には「はるかに広範で野心的ではない品質の概念」が含まれており、これは、試行錯誤された卓越性の原則を研究のベンチマークとして使用することと矛盾していると主張している。この合意は「基準と前提条件を混同する傾向がある」「多様性と協力が研究の卓越性に不可欠な前提条件であることは否定できない。しかし、そ

¹⁸ Research assessment: German funder raises concerns as EU reform talks come to a close
<<https://sciencebusiness.net/news/european-research-area/research-assessment-german-funder-raises-concerns-eu-reform-talks-come>> [2025/3/14 取得]

¹⁹ 同上

の 2 つは、研究提案の科学的内容を評価する基準にはならない」とローゼンタール氏は述べている。

このようなアプローチの懸念は、この取組が研究評価改革にとどまらず、ヨーロッパにおける資金配分方法に影響を及ぼすことだと彼は指摘する。「これは単なる政治的な結果であり、科学的な結果ではないため、（ドイツ学長会議は）これを支持できない」とローゼンタール氏は述べている²⁰。

(3) フランス共和国（仏国）

2022 年 2 月にパリで開催されたオープンサイエンス会議（OSEC）で、EU の研究リーダーたちは、ファンディング機関が個々の研究者の科学的地位を評価する方法を一新するという欧州委員会の計画を支持した。本会議は、フランスの欧州連合理事会議長国（当時）の一環として開催されている。フランスは当時より、欧州委員会や科学・大学協会と協力し、欧州における研究を評価する新たなシステムを推進しており、研究評価に関するパリ・コール（Paris Call）を立ち上げている。フランス政府のオープンサイエンス委員会が起草したこの呼びかけでは、研究者をインパクトファクターの高い学術誌に掲載された引用数の多い論文の数のみで評価することは、他の貢献の価値を過小評価し、研究の再現性を低下させ、研究者が研究を公開して他者と協力することを阻むことになっているとしている。パリ・コールは、欧州研究圏の新たな政策アジェンダに定められた 20 の優先事項のチェックリストに適合しており、リストの最初には科学と研究の評価が位置づけられている²¹。

パリ・コールにおいては、次のような研究評価システムが必要であるとしている²²。

- 研究の質とさまざまなインパクトを評価する
- 研究が倫理と誠実さの最高基準を満たすことを保証する
- 出版物やプレプリント、データ、方法、ソフトウェア、コード、特許などの研究活動と成果の多様性、ならびにそれらの社会的インパクトとトレーニング、イノベーション、公共の関与に関連する活動を重視する
- 研究分野の多様性を尊重する評価基準とプロセスを使用する
- 研究成果だけでなく、適切な研究の実施も評価し、優れた実践、特に可能な限り研究結果と方法論を共有するためのオープンな実践を評価する
- 適切な場合には、共同作業、学際性、市民科学を重視する
- 多様な研究者プロフィールとキャリアパスをサポートする

²⁰ German universities still wary of EU push to reform research assessment <<https://sciencebusiness.net/news/german-universities-still-wary-eu-push-reform-research-assessment>> [2025/3/21 取得]

²¹ France helps Brussels move ahead with ‘disruptive’ plan for research assessment <<https://sciencebusiness.net/news/horizon-europe/france-helps-brussels-move-ahead-disruptive-plan-research-assessment>> [2025/3/21 取得]

²² Paris Call on Research Assessment <<https://osec2022.eu/paris-call/>> [2025/3/21 取得]

研究資金提供組織、研究実施組織、評価機関の連合を結成し、共通の目的、原則、行動（相互学習、文書の共有、共通の監視活動など）に沿って現在の研究評価システムを改革する意志と熱意を持つことを求め、このような連合の成功は、具体的な実施プロセスを提案する能力と、あらゆるレベルの研究者を連携させ、関与させる能力に深く関係するとしている。

このような取組を受けて、フランス最大の政府基礎研究機関であるフランス国立科学研究中心（CNRS）は、2022年9月に欧州委員会が始めた、研究評価改革のための連合を形成する署名者による合意に向けた署名活動に対して、最初の署名機関となった（DORA宣言に対しては2018年にCNRSは署名している）。評価は、2019年11月に公開されたCNRSのオープンサイエンスロードマップの4つの基本的な柱の1つであり、量ではなく質に基づいた基準を実装するように設計されている²³。署名集めは、2021年に欧州委員会が「研究評価システムの改革に向けて（Towards a reform of the research assessment system）」と題する報告書を発表したことから始まったプロセスの集大成と言える。その後、改革を開始するための連合に参加するための関心表明の呼びかけが行われ、CNRSと40か国以上から350を超える組織が関心を示した。2022年初頭、欧州委員会は共通の方向性を示す文書を起草するための委員会を設立した。この委員会は、CNRSを含むさまざまな組織を代表する20人の専門家グループと定期的に交流した。したがって、CNRSは研究評価の改革に関するこの協定の構築に最初から関与していた。このような文脈により、研究評価推進連合（CoARA）が設立された。

(4) イギリス（英国）²⁴

英国研究評価枠組み（REF）は、研究のインパクトを2つの次元で測定している。「重要性（significance）」（プロジェクトがもたらす目に見える違い）と「到達度（reach）」（その違いを定量化できる範囲）（UKRI）。ここでいう「インパクト」とは、「経済、社会、文化、公共政策やサービス、健康、学術界を超えた、経済、社会、文化、公共政策やサービス、健康、環境、生活の質に対する効果、変化、利益」と定義されているが、それ以外については非常に曖昧で、分野によって異なり、おそらく曖昧である。

例えば、社会との関わりについては適切に考慮されていない。英国のREFは、2022年から2023年にかけて、将来の研究評価プログラム（Future Research Assessment Programme）の下で評価され、英国の高等教育研究実績の評価に対する新たなアプローチを模索している。

また、国際的な研究評価の実践の理解も含まれている。次回のREFでは、より多様な成果が考慮される可能性があり、多様な成果を考慮し、それらへの重要性を低減する可能性もある。現在のモデルでは、成果に60%、研究へのインパクトに25%、研究文化・環境に15%

²³ Reforming scientific assessment: the CNRS among the first signatories

<<https://www.cnrs.fr/en/update/reforming-scientific-assessment-cnrs-among-first-signatories>>

[2025/3/21 取得]

²⁴ Rijke et al. [3]

の重要性を割り当てている。これらの比重がより均等になれば、REFは大きく様変わりし、研究文化、研究インテグリティ、チームワークにより重点が置かれることになるだろう。

REFへの提出は、34の評価単位(UoA)で行われる。次回のREF(REF 2029)では、4つのメインパネルと34のサブパネル(UoAごとに1つ)で構成されるREF2021のUoA構造が維持される。サブパネルは以下のとおりである²⁵。

メインパネルA: 医学、健康、生命科学

メインパネルB: 物理科学、工学、数学

メインパネルC: 社会科学

メインパネルD: 芸術と人文科学

タイムテーブルとしては、2023年夏にREF 2029の初期決定をし、2023年夏に人々、文化、環境(PCE)に関する円卓会議を開催するなどして初期決定の次のステップに進んだ。2024年春にはOpen Access consultationを実施するとともに、諮問委員会の募集を開始した。2024年夏には専門家パネルの募集を開始し、2025年秋冬頃までにオープンアクセスポリシーやボリューム測定ガイダンスおよび関連する実践規範ガイダンスを策定する予定である。2026年に最終的な完全なガイダンスが設定され、2028年より評価フェーズが開始して2029年に結果が公表される。

(5) イタリア共和国(伊国)

全国大学システム研究評価委員会(Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca: ANVUR)が設立以来、ベルルスコーニ第4次内閣時の教育改革であるいわゆる「ジェルミニ改革(Gelmini Reform)」によって導入された研究の質の評価システムであるValutazione della Qualità della Ricerca(VQR)を実施している。

近年はイタリアの科学技術イノベーション計画においても、研究評価に関する記述が見られるほか、イタリア大学・研究省(Ministero dell'Università e della Ricerca: MUR)が所管する研究開発プログラムにおいても研究評価のガイドラインが定められている。イタリアにおける研究評価は、政権交代等のたびに新たな制度や機関が発足するなどの改革が行われ、絶え間なく進化しており、確定的な構造となるには至っていない。イタリアにおける研究評価は、イタリア共和国憲法第97条に謳われている「善政の原則(buona amministrazione)」を、科学的・学術的領域において保証するものと位置付けられている。研究評価では、研究活動の厳密な科学的側面と社会経済的側面の両方が考慮されるのが原則である。

最近では、PNR 2021-2027におけるオープンサイエンスのための全国計画が策定され、研究インフラ全国計画(PNIR)とオープンサイエンスのための全国計画が含まれている。このうち、オープンサイエンスのための全国計画では研究評価のあり方についてもオープンサイエンスの促進の観点から規定されている。本規定の性格はイタリアにおける今後の

²⁵ REF2029 ウェブサイト <<https://2029.ref.ac.uk/>> [2025/3/21 取得]

研究評価のあり方の方向性を示す政策文書である。当計画のアクションプランにおける特長としては、評価の方法論としてインパクトファクターや H-Index といった計量書誌学指標の比重を減らしていくことが提案されている点が挙げられる。EU における研究評価の標準への接続の点で、DORA 宣言の原則の順守が掲げられている点も注目に値する。

(6) カナダ（加国）²⁶

カナダでは、DORA を推進力として研究評価の改革に関する複数の議論が行われており、3 つの連邦研究協議会のすべてが署名者となっている。自然科学・工学研究評議会 (NSERC) は、DORA の原則に沿って、書誌学、引用、h 指数を廃止し、研究の質に関する基準を再定義した。質的評価基準には、質の高い研究データやデータへのアクセス管理、公平性、多様性、包含性、研修責任などが含まれる。他の 2 つの研究評議会もこれに追随する可能性が高い。カナダの研究者は、ユーザーコミュニティとの共同生産を通じて研究の社会的影響力を高める意図的な取り組みである「知識の活用」に重点を置く傾向がある。Research Impact Canada は、20 以上の大学からなるネットワークであり、公共の利益のために研究のインパクトを最大限に高めることを目的として、インパクト・リテラシー（「適切なインパクト目標と指標を特定し、インパクト経路を批判的に評価し最適化し、状況に応じてアプローチを調整するために必要なスキルを振り返る」能力）を通じて組織能力を構築することを目指している。注目すべきは、DORA に署名したカナダの大学が非常に少ないことである。変化の主な要因は、おそらく先住民の学問を受け入れることであろう。これはカナダでは道徳的な要請となっている。

【参考文献】

- [1] 小野寺 夏生, 伊神 正貫, 研究計量に関するライデン声明について, STI Horizon, No.4 Vol.2, 2016.
- [2] Wilsdon, J. et al. The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management, 2015.
- [3] Sarah de Rijcke, Clemencia Cosentino Robin Crewe and et al. (the members of the GYA-IAP-ISC Scoping Group), THE FUTURE OF RESEARCH EVALUATION: A SYNTHESIS OF CURRENT DEBATES AND DEVELOPMENTS, 2023.

²⁶ Rijcke et al. [3]

2.2 アメリカ合衆国（米国）

2.2.1 国・地域における研究開発評価制度の体系とその内容

(1) 評価制度の全体像と研究開発評価の位置付け²⁷

米国では、2025年1月のトランプ新政権発足後、大統領令などを通じ、行政運営に関わるさまざまな改革が行われようとしている。その代表的な動きとしては、「政府効率化省（DOGE）」の設置があげられる。これは、大統領主席補佐官の管轄下に置かれている機関であり、政府外部のイーロン・マスクが主導している²⁸。DOGEは、合衆国憲法上歳出削減の権限を持たず、議会に対して勧告を行うことができるにすぎないが、行政裁量の範囲で執行可能な予算削減や契約見直しを行っており、議会の正式な歳出法案を待たずに削減が実施された例もすでに複数確認されている²⁹。その中には、全米科学財団（NSF）や国立衛生研究院（NIH）といった、歴代政権において継続的に重視され、予算の増額が図られてきた機関も含まれる³⁰。

また、トランプ政権における政策形成や評価のあり方に関わる動きとしては、2025年2月19日に署名された大統領令14219号「合法的なガバナンスの確保と政府効率化省による効率的な規制イニシアチブの実施」がある³¹。これは、連邦政府の規制プロセスを抜本的に見直し、規制の簡素化と大統領の政策的優先事項との整合を強化しようとするものであり、これまで規制レビュープロセスの対象外であった連邦取引委員会（FTC）や証券取引委員会

²⁷ 未来工学研究所 [1] をもとに、最新動向を踏まえて加筆修正

²⁸ ロイターの報道によると、トランプ大統領とホワイトハウスの公式見解との間に一部矛盾が生じているなど、DOGEの責任者は必ずしも明確ではない（2025年2月20日午後1:08 GMT+9 配信）。

< <https://jp.reuters.com/world/us/BJVYQFV7VNNT7A23B2SOTURKEE-2025-02-20/> > [2025/3/21 取得]

²⁹ 2025年2月26日にトランプ大統領が署名した大統領令「費用効率イニシアチブ」では、すべての既存の契約および助成金を30日以内に見直し、適切であれば終了または変更すること、新規契約の承認は機関長の事前承認がない限り一時的に停止することなどが盛り込まれている。

White House : IMPLEMENTING THE PRESIDENT'S "DEPARTMENT OF GOVERNMENT EFFICIENCY" COST EFFICIENCY INITIATIVE < <https://www.whitehouse.gov/presidential-actions/2025/02/implementing-the-presidents-department-of-government-efficiency-cost-efficiency-initiative/> > [2025/3/21 取得]

³⁰ たとえば、榎木は、NIH、CDC（アメリカ疾病予防管理センター）、NSF、NASA（アメリカ航空宇宙局）をはじめ多くの機関が人員削減、予算削減等を指示されており、その影響が予算を受け取る大学などにも波及していることを各種報道記事をもとに指摘している。榎木英介、トランプ政権、米連邦政府の科学関連部門に大ダメージを与える、Yahoo!ニュース、2025年2月24日。

< <https://news.yahoo.co.jp/expert/articles/68f0540346cd503bbf855dad92979837067a9419> > [2025/3/21 取得]

³¹ White House : ENSURING LAWFUL GOVERNANCE AND IMPLEMENTING THE PRESIDENT'S "DEPARTMENT OF GOVERNMENT EFFICIENCY" DEREGULATORY INITIATIVE < <https://www.whitehouse.gov/presidential-actions/2025/02/ensuring-lawful-governance-and-implementing-the-presidents-department-of-government-efficiency-regulatory-initiative/> > [2025/3/21 取得]

(SEC) などの独立規制機関も対象とするものである。これに伴い、バイデン政権下で発行された大統領令 14094 号「規制レビューの現代化³²⁾」は廃止されることとなった。「規制レビューの現代化」では、規制政策の包摂性の確保と市民参加の積極的な推進を規定するなど規制の社会的影響や公平性が重視されていたが、トランプ政権では規制の経済的影響や政府の効率性を重視する立場へと大きな転換を行った、と考えられる。

こうした一連の動向からは、小さな政府志向や（評価なき）行政効率の重視、「イデオロギー」の排除（DEI への反発というイデオロギー）、即効性の優先（長期的な成果を見込む基礎研究より、短期的な費用対効果が明確なものを優先）、といった政権の方針をみてとることができる。

一方、2001 年のジョージ・W・ブッシュ政権以降、政権交代後に示されていた評価を含む行政のマネジメントにかかる方針「大統領のマネジメント・アジェンダ (President's Management Agenda: PMA³³⁾」について、政権発足からまもないこともあり、トランプ政権ではまだ公表されていない。上述の傾向から、これまで米国が積み上げてきた評価システムの延長線上にはない方針転換が起こる可能性もあるが、制度変更について公式に言及したものはないため、以下では、前バイデン政権における評価制度を中心にみていくこととする。

前バイデン政権における評価制度は、研究開発評価を含め、基本的にはオバマ政権時に制定された「政府業績成果現代化法 (GPRAMA³⁴⁾」をもとに運用されていた。GPRAMA は、クリントン政権時の 1993 年に成立した政府業績成果法 (GPRA) を改正したものである。GPRA は、個々の連邦政府機関に 5 年間を対象とする戦略計画 (strategic plan) の策定と、その下での年度ごとの業績計画 (performance plan) 及び業績報告書 (performance report) の作成を義務付けるものであった。これは、「業績測定」型、すなわち、目標に対する達成度の測定とモニタリング活動からなる自己評価として、次年度予算案と同時に前年度の年次業績報告と次年度の年次業績計画を公表させる、という枠組みであり、戦略計画は 3 年ごとに更新されていく、というプロセスがとられていた。

GPRA はその後、ブッシュ政権時に大統領のイニシアチブで導入された「プログラム評価・査定ツール (PART)」、「マネジメント・スコアカード」、「業績予算」といった仕組みを併用することにより、予算と行政活動との連結を重視するものへと変質していった。一方、「GPRA・PART の結果、多くの業績指標と、重複する複数の業績マネジメント・システムができたが、その結果、議会・省庁・国民の誰も評価結果を使っていない」、「業績指標のほとんどはプロセス重視でアウトカム・ベースではない。省庁横断的な目標がない。

³² White House: Executive Order on Modernizing Regulatory Review (APRIL 06, 2023)
< <https://bidenwhitehouse.archives.gov/briefing-room/presidential-actions/2023/04/06/executive-order-on-modernizing-regulatory-review/> > [2025/3/21 取得]

³³ 直近の PMA については以下を参照 < <https://bidenadministration.archives.performance.gov/pma/> > [2025/3/21 取得]

³⁴ GPRA MODERNIZATION ACT OF 2010 :
< <https://www.congress.gov/111/plaws/publ352/PLAW-111publ352.pdf> > [2025/3/21 取得]

業績情報は変革をもたらすために使われていない³⁵⁾といった問題点が指摘されていた（新日本有限責任監査法人 [2]， 未来工学研究所 [3]）。

GPRAMA は、こうした背景の中、「政府全体の業績をどう改善するか、あるいはそのためのシステムをどう改革するか」という視点で見直されたものであり（南島 [4]）、業績評価結果を予算編成に反映させるというよりも、諸改革と法制化を通じて、業績評価から得られる情報をマネジメントに活用することがより重視されるようになった（新日本有限責任監査法人 [2]）。

GPRAMA の主なポイントとしては次のとおりである。

表 2-2 GPRAMA の概要

ポイント	概要
戦略計画関連	<ul style="list-style-type: none"> ・ 大統領の任期にあわせて 4 年間以上とする。 ・ 大統領就任年の翌年 2 月第 1 月曜までに各機関の公式ウェブサイトで公開、大統領と議会に通知。関係する議会の委員会と各機関との間で 2 年毎に協議を実施(重大な外部環境の変化がある場合、適宜見直し) ・ 機関の目標・目的並びに関連する政府横断的な優先目標の達成に向けて、他機関とどのように連携するかや、議会からの指摘をどのように組み込んでいるかの記載を新たに要求。 ・ (GPRA に引き続き)目標・目的の設定や修正に用いたプログラム評価の説明と、今後のプログラム評価の予定を記載。
年次業績計画関連	<ul style="list-style-type: none"> ・ 年次業績計画には、政府横断的な取組に係る連邦政府業績計画 (federal government performance plan) と、機関ごとに策定する機関業績計画 (agency performance plan) がある³⁶⁾。 ・ 連邦政府業績計画については大統領府行政管理予算局 (OMB) が各機関と調整の上設定した機関横断型の優先目標 (Cross-Agency Priority Goals: CAPs³⁷⁾) に対し、現年度、次年度の業績目標を設定。また、業績目標の達成に貢献する機関や諸活動を特定するとともに、全体の進捗状況やそれら個別の機関、活動の貢献を測定もしくは評価するための業績指標を 4 半期ごとの目標とともに設定。目標達成に係る調整の責任者についても指定。 ・ 機関業績計画については現年度、次年度の業績目標を設定するとともに、それらが機関の目標・目的もしくは連邦政府業績計画で設定された業績目標に対しどのように貢献するかを記載。また、設定した業績目標のうち、機関としての優先目標 (Agency Priority Goals: APGs) を 2 年毎に特定。その他、特に貢献度の高いプログラムや 4 半期ごとのマイルストーン、目標の責任者、バランスのとれた一連の業績指標、データの正確性・信頼性、主なマネジメント上の課題等を記載。

³⁵⁾ ジェフリー・ジェンツ首席業績担当官 (CPO) による 2009 年 10 月 29 日上院委員会での議会証言

³⁶⁾ これらは、後述の業績報告書とあわせ、「大統領予算教書 (補足説明報告書)」の一部として公表される。

³⁷⁾ Cross-Agency Priority Goals (CAPs) という表現は後述の OMB Circular No. A-11 で用いられているものであり、GPRAMA には登場しない。

OMB Circular A-11, Preparation, Submission and Execution of the Budget, July 2024.

< <https://www.whitehouse.gov/wp-content/uploads/2018/06/a11.pdf> > [2025/3/21 取得]

<p>年次業績報告書 関連</p>	<ul style="list-style-type: none"> ・ 予算年度終了後遅くとも 150 日以内に提出。 ・ 目標達成状況に関するレビューを 4 半期ごとに実施するとともに、オンラインによる公開を義務付け。 ・ 次の事項を明記：業績目標を達成し成功しているかどうかについての検証及び過去 5 年間の実績；達成できた業績と関連させて、現会計年度の業績計画を評価；（業績目標が未達の場合）未達の理由、設定された業績目標を達成するためのプラン及びスケジュール；（業績目標が未達かつ業績目標が非実用的又は実行不可能である場合）そのような事態に陥った理由；データの正確性及び信頼性をどのように確保しているかの解説；対象となる期間中に完了したプログラム評価の成果の要約、等。
-----------------------	--

出所：GPRAMA 及び新日本有限責任監査法人[2]をもとに未来工学研究所作成

米国の評価制度に関して、もう 1 つ重要なものが 2019 年 1 月に成立した「エビデンスに基づく政策形成基盤法（Foundations for Evidence-Based Policymaking Act of 2018 : EBPM 法）」である。ここでいう「エビデンス」とは、狭義には「統計的な目的のための統計活動の結果として生成される情報」のことを指すが、実際の評価制度の運用においては「信念や命題が真または妥当であるかどうかを示す利用可能な事実または情報の総体」として広義に解釈される³⁸。同法では、連邦政府のエビデンス構築活動、オープンガバメントデータ、機密情報保護と統計の効率化を法律で義務付けることで、連邦政府におけるデータとエビデンス構築機能を推進するための協働と調整を促すことを狙いとしている。また、エビデンス構築活動と国民消費のためのアクセスをより促進するために、政府データ管理の体系的な再考を義務付けている。

EBPM 法では、各連邦政府機関に「エビデンス構築計画（evidence-building plan）」を策定することを要求している。これは別名「ラーニング・アジェンダ（Learning Agenda）」と呼ばれているものであり、同法 306 条に基づき、プログラム等に関わる重要な改善課題（組織学習の課題）のリストとそれらに対する調査・分析・評価を行うことになった。

こうした方向性は、前バイデン政権でも強化される形で引き継がれている。具体的には、バイデン政権における PMA において、「すべての国民に結果をもたらす、公平かつ効果的な、説明責任のある政府に向けて」というビジョンの下、1) 連邦政府職員の強化とエンパワーメント、2) 優れた連邦サービスの提供と顧客体験の向上、3) より良い政府を築くための業務マネジメントの 3 つが優先事項として掲げられており、その中で、「連邦政府全体で 10 年以上注力してきたことを踏まえ、連邦政府のデータマネジメント及びデータサイエンス能力を活用する」ことや、PMA を実現するためのツールとして「ラーニング・アジェンダ」を重視していくことが明記されている³⁹。

このほか、プログラム・マネジメントに関する政府全体の基準や方針などを定めた「プログラム・マネジメント改善説明責任法（Program Management Improvement

³⁸ OMB Circular A-11 (2024)

³⁹ The Biden-Harris Management Agenda Vision: Track progress toward building a more equitable, effective, and accountable Government that delivers results

< <https://bidenadministration.archives.performance.gov/pma/> > [2025/3/21 取得]

Accountability Act: PMIAA) 」などの関連法があり、これらの法を一貫性を持って効果的に運用して行くために、OMB が通達や覚書の形でガイダンスを提供している。これらのガイダンスの詳細については後述する。

次図は、こうした GPRAMA や EBPM 法等に基づく米国の評価制度（業績マネジメントシステム）の全体像を示したものである。このうち、「エビデンス、評価、分析及びレビュー」に記載の内容については、(2) 研究開発評価に関する規定等の整備状況で解説する。

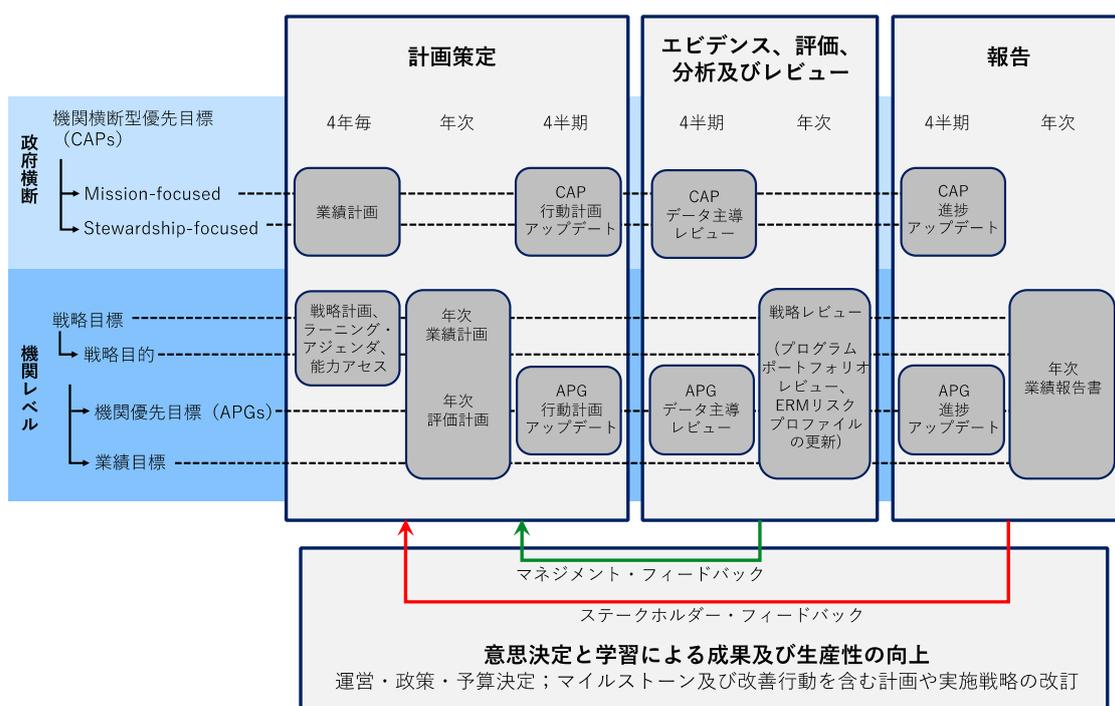


図 2-5 米国の評価制度の全体像

出所: OMB Circular No. A-11 (2024), Page 7 of the Executive Summary をもとに未来工学研究所作成

なお、米国では、GPRAMA で求めている定型的な評価以外にも、多様なチェック・アンド・バランスの仕掛けが組み込まれている。たとえば、会計検査院 (GAO)、議会調査サービス (CRS)、国立科学・工学・医学アカデミー等によってアドホックに評価が行われるほか、米国における主要なシンクタンクも、政府のプログラムと政策オプションに関するアセスメントを実施するなどしている (Wessner [5])。

(2) 研究開発評価に関する規定等の整備状況

前述のように、米国における研究開発評価を含む評価制度は GPRAMA や EBPM 法等に基づいて運用されているが、これらの実施を支援する多様な規定等が整備されている。次表は、このうち主なものを一覧化したものである。

表 2-3 研究開発評価に関する主な規定等の一覧

名称	性格	概要
OMB Circular No. A-11	通達	連邦政府機関の長に向け、OMB が予算編成・執行に関するガイドラインとしてまとめたもの
OMB 長官(代理)による覚書 (OMB M-19-23; OMB M-21-27 等)	覚書	ラーニング・アジェンダからプログラム評価に至るまで4段階でガイダンスを整理

出所:各種資料をもとに未来工学研究所作成

まず、第一に挙げられるのが「Circular No. A-11」である。これは、各連邦政府機関等の長に向けた通達文書として、OMB が予算作成に関するガイダンスや予算執行に関するインストラクションを提供するものであり、最新のものは 2024 年 7 月に公表されている。Circular No. A-11 は、GPRAMA や EBPM 法等を前提に、大統領の示すビジョンや優先事項の実現に資する予算編成・執行を行うための実務的なガイドラインとなっている。なお、ここでいう各連邦政府機関が踏まえるべき大統領の優先事項とは、前述の PMA のような行政管理一般に関わるものだけではなく、OMB と大統領府科学技術政策局 (OSTP) が関係連邦政府機関等の長に宛てて例年 8 月に公表する「研究開発予算優先事項⁴⁰」なども想定されている。

また、OMB では、OMB 長官 (代理) による覚書として、2019 年 7 月 10 日に関連するガイダンス (OMB M-19-23) のとりまとめを行っている。ここでは、「EBPM 法の要件を満たすには、継続的かつ集中的な努力が必要」であり、「(その) 実施は必然的に既に実施中の活動を基礎とし、場合によっては新たな方法でこれらの活動の調整を必要とすることを認識している」との問題意識の下、次のようなことに取り組むこととしている。なお、この覚書自体が下記第 1 段階のガイダンスに対応するものとなっている。

- 関連するデータ・情報政策のガイダンスと、OMB Circular No. A-11 (毎年度の予算作成に関するガイダンスと予算執行に関する指示を提供) のような関連する OMB ガイダンスを整合させること
- 4つの段階に分けて反復的かつ補足的なガイダンスを提供すること
 - 第 1 段階：ラーニング・アジェンダ、人事及び計画策定
 - 第 2 段階：オープン・データ・アクセス及びマネジメント
 - 第 3 段階：統計目的のためのデータ・アクセス
 - 第 4 段階：プログラム評価

さらに、2021 年 1 月 27 日の「科学的インテグリティと EBPM を通じて政府に対する信頼を回復することに関する大統領覚書」を受け、2021 年 6 月 30 日には、OMB M-19-23 や第 4 段階のガイダンス (OMB M-20-12) 等を拡張したガイダンス (OMB M-21-27) が発表された。

⁴⁰ 最新のものは次の通り：Multi-Agency Research and Development Priorities for the FY 2025 Budget (8/17/2023)。なお、2026 会計年度については公開されていない。

(3) 研究開発評価に関する規定等の詳細

1) OMB Circular No. A-11

a. 規定の概要及び構成

OMB Circular No. A-11 は、7 パート 290 セクションで構成され、その分量は 1,000 ページを超えるものとなっている。このうち、パート 6「プログラム及びサービス提供の改善のための連邦政府業績フレームワーク」が評価に直接的に関わる部分である。主にプログラム評価及び機関評価に関わるものであり、その構成は次のようなものである。なお、2023 年度版と基本構成は変わっていない。

表 2-4 OMB Circular No. A-11 PART6 の構成

第 200 条 連邦業績フレームワークの概要
第 210 条 公的な報告書の作成
第 220 条 PMA、機関横断型優先目標 (CAPs) 及びマネジメント・イニシアチブ
第 230 条 機関の戦略計画策定
第 240 条 年次業績計画の策定及び報告書の作成
第 250 条 各機関の優先目標 (APGs)
第 260 条 データ主導業績レビュー及び戦略的レビュー
第 270 条 プログラム及びプロジェクト・マネジメント
第 280 条 カスタマー・エクスペリエンスのマネジメントとサービス提供の改善
第 290 項 評価とエビデンス構築活動

出所: OMB Circular No. A-11 をもとに未来工学研究所作成

b. 規定の詳細

以下では、本通達における主な内容について紹介する。各条項の冒頭では、改訂箇所のサマリーが掲載されているが、ここで取り上げた項目について、2023 年 8 月版からの大きな変更はない。なお、全編にわたって Q&A 形式でまとめられていることに 1 つの特徴がある。

ア) 第 200 条 連邦業績フレームワークの概要

第 200 条では、関連する法や行政文書の解説、各機関における責任体制のあり方、用語の定義や目標体系、タイムライン等についてまとめられている。2023 年 8 月の改訂版からは、組織の健康 (organizational health) や業績に関わる概念・定義、そして、進化する職場環境においてそれら进行评估するためのフレームワークが本ガイダンスに盛り込まれた。ここでいう組織の健康とは、「組織のミッションを支え、プログラムやサービスを提供し、

ステークホルダーのニーズと優先事項を継続的に満たすために、集団としてパフォーマンスを向上させる能力」のことである。また、業績測定のアプローチの一種として「有効性 (Effectiveness)」を定義するなどしている。

イ) 第 230 条 機関の戦略計画策定

第 230 条では、戦略計画の目的や内容、年次業績計画との関係、効果的な戦略目標や目的のあり方などについて解説されている。特に、戦略計画の策定やエビデンスの調整に役立つツールとして、後述の OMB M-21-27 から引用する形でロジックモデルの紹介を行っている (230.10)。

ウ) 第 240 条 年次業績計画の策定及び報告書の作成

第 240 条では、年次業績計画及び業績報告書の目的や内容、戦略計画や議会による予算承認との関係、機関横断型優先目標 (CAPs) の取り扱い、業績指標の変更や適切なエビデンス利用のあり方などについて解説されている。

このうち、GPRAMA に基づいて行う組織の業績マネジメントのためのエビデンスについて、EBPM 法の定義を置き換えるものではないとしつつ、前述のように、「信念や命題が真実もしくは妥当かを示す、入手可能な事実または情報の集合体」としてより広義に捉えている。こうしたエビデンスには量的なものだけではなく質的なものがあり、基礎的なファクト・ファインディング、業績測定、政策分析、プログラム評価を含む多様な情報源や構成要素からもたらされる可能性がある。また、エビデンスの信憑性には様々な程度があり、「一般的に、最も強力なエビデンスは単一の研究ではなく、質の高い信頼できる情報源のポートフォリオから得られる」ものであるとし、OMB M-19-23 の図を引用しながら、これらの情報源がどのように組み合わせられてエビデンスの構成要素になっているかを示している。なお、日本における研究開発評価の文脈では、政策・施策とプロジェクトとの中間にある政策階層としてプログラムを位置付け、それを対象とする評価をプログラム評価と呼ぶことが多いが、米国におけるこれらの政策文書においては、上記の記述からも分かるように、特定の政策階層における評価を意味するものではなく、業績測定等とは異なる様式のエビデンス構築活動のことを指していることに留意する必要がある⁴¹。

⁴¹ プログラム評価のガイダンスである OMB M-20-12 では、評価及びプログラム評価について、「1つまたは複数のプログラム、政策、組織の有効性と効率性を評価する (assess) ために、体系的なデータ収集と分析を行うプロセス」として定義している。そのため、評価は、プログラム、政策、または組織のレベルを超えて行われる場合があること、また、重要な点として、形成的評価、インパクト評価、プロセス/実施評価等異なる種類があること、プログラム、政策、または組織の実施や設立に関する問いや特定の戦略の有効性や影響に関する問い、有効性の変動に関連する要因についての問い、プログラムを取り巻く文脈要因の理解に関する問いなどの検証に用いられること、これらは、現在のプログラム、政策、組織運営に関する意思決定に対し重要な参照情報を提供するものであること、などが述べられている。なお、本定義の最後では、「評価は学習と改善のためだけでなく、説明責任を果たすためにも使用されるべき」ことが強調されている。<<https://www.whitehouse.gov/wp-content/uploads/2020/03/M-20-12.pdf>> [2025/3/21 取得]

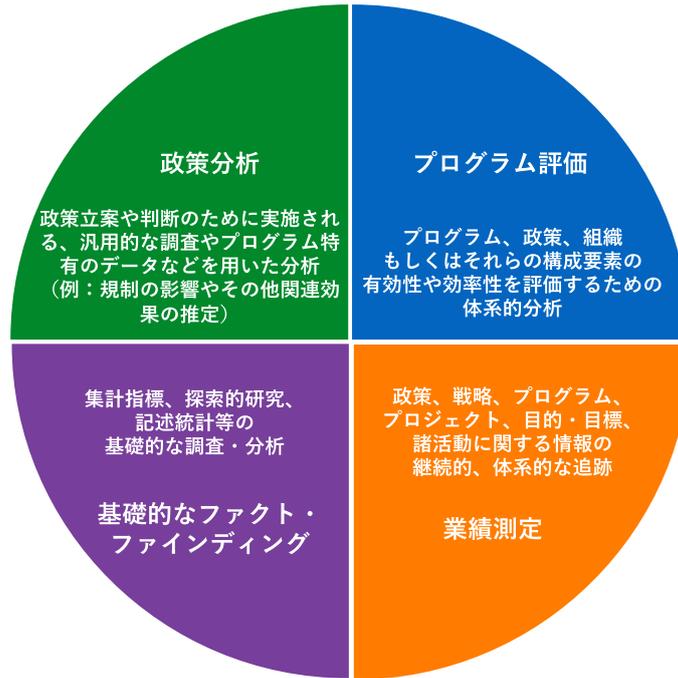


図 2-6 エビデンスの情報源及び構成要素

出所: OMB Circular No. A-11 (元図は OMB M-19-23) をもとに未来工学研究所作成

エ) 第 260 条 データ主導業績レビュー及び戦略レビュー

第 260 条では、「データ主導業績レビュー (Data-Driven Performance Reviews) 」と「戦略レビュー (Strategic Reviews) 」及び後者の一環として実施が義務付けられている「事業リスク・マネジメント (Enterprise Risk Management: ERM) 」について解説している。これらのそれぞれについて、概要をまとめると次のとおりである。

表 2-5 データ主導業績レビュー、戦略レビュー及び事業リスク・マネジメント

構成	概要
データ主導業績レビュー	機関の業績向上のための優先事項の一部について、少なくとも 4 半期に一度データに基づく進捗のレビューを機関のリーダーに課すもの。これにより、目標を進展させるために必要な人材やリソース、分析を結集させるとともに、継続的な学習と改善を行う組織文化の確立を企図。
戦略レビュー	機関の戦略計画で設定され、年次業績計画で毎年更新される機関の各戦略目標について、毎年の進捗状況のレビューを機関のリーダーに課すもの。その結果を戦略的意思決定、予算編成、短期的な機関の行動、年次業績計画及び年次業績報告書の作成に反映。
事業リスク・マネジメント	OMB Circular No. A-123「事業リスク・マネジメント及び内部統制に関する経営者の責任」に従い、事業のリスク評価とマネジメントを課すもの。

出所: OMB Circular No. A-11 をもとに未来工学研究所作成

オ) 第 290 条 評価とエビデンス構築活動

第 290 条は、EBPM 法の文脈にそってまとめられているものであり、各機関における評価責任者の役割や専任方法、ラーニング・アジェンダ及び年次評価計画の詳細や相互の関係性並びにそれらの実施状況のモニタリング、関連する OMB のガイダンス等について解説されている。

まず、ラーニング・アジェンダ (エビデンス構築計画) は、当該機関のプログラム、政策、規則に関連する政策上の問いを特定し、それに取り組むための体系的な計画として位置付けられている。具体的には、短期的・長期的に重要な戦略上の問い (プログラム、政策、規制が個々に、あるいは組み合わせでどのように機能するかなど、機関がそのミッションをどのように果たすかに関する問い)、及び業務上の問い (人事、助成金申請、内部プロセスなど、省庁の業務に関する問い) に答えるためのエビデンスを開発するための戦略を特定し、優先順位を付け、確立しようとするものである。OMB は、これらの問いについて、現在の行政の優先事項や長年抱えている課題、将来出てくる可能性のある優先事項に関わるものが含まれることを期待しており、また、こうした問いに答えるために、プログラム評価、業績測定、政策分析、基礎的な事実探索など、あらゆる種類のエビデンスを用いることを期待している、としている。

また、ラーニング・アジェンダを作成するプロセス (利害関係者の参加、利用可能なエビデンスの検討、問いの作成、エビデンス活動の計画と実施、結果の普及と利用、作成されたエビデンスに基づく問いの精緻化) は、「集団的な学習と継続的な改善のひとつ」であり、「結果として得られる文書そのものと同等かそれ以上に有益である可能性がある」、としている。そのため、ラーニング・アジェンダは「少なくとも毎年見直される、柔軟で反復的な文書であるべき」であり、「すべての機関に有効な単一のアプローチや形式は存在しない」としている。

評価計画は、提出された年度の翌年度に各機関が実施する予定の重要な評価活動を記載したものであり、EBPM 法によって策定が義務づけられているものである。本通達では、この計画には、「ラーニング・アジェンダの優先的な問いに対する回答に役立つ「重要な」評価と、法令で義務付けられているようなその他の「重要な」評価を含めるべき」であり、こうした「法定の評価の定義と各機関の「重要」の定義を満たす活動のみを含めるべきである」、としている。何が「重要」かについては、各機関によって定義されるべきであり、各機関にはその基準を示すことが求められる。その際、機関のミッションに対するプログラムや資金の流れの重要性、資金や対象となる人々の観点から見たプログラムの規模、プログラム、対象となる人々、またはプログラムが対処するために設計された問題に関する重要な知識のギャップをどの程度埋めることができるか、といった要素を考慮しなければならない、とする。

第 290 条の最後には、4 年間の戦略計画とエビデンス計画のサイクルを示した図を OMB による覚書 M-21-27 から引用し、ラーニング・アジェンダと年次評価計画の活動、及びこれらの計画に含まれていないその他の関連するエビデンス構築活動から生み出されたエビ

デンスが、ラーニング・アジェンダの更新と次年度の年次評価計画の起草にどのように用いられるかを示している。これについては、後述する。

c. 特記事項

本ガイダンスでは、業績報告書に記載されるデータの信頼性に関する情報を提示することを求めている。年次実績報告書に含まれる送付状には、提示された実績データの完全性と信頼性に関する機関長による評価、及び必要な場合には完全性、信頼性、及び質を改善するための機関計画についての記述が含まれていなければならない、とする。その中で、ピアレビュー等の外部評価を活用することが有効な手段であると指摘されている。

2) OMB M-21-27

a. 規定の概要及び構成

OMB M-21-27 は、OMB 長官代理による覚書として、2021 年 6 月 30 日に連邦政府機関の長に向けて出されたものであり、「エビデンスに基づく政策形成：ラーニング・アジェンダと年次評価計画」を主題とするものである。

その構成を示すと次のようなものである。28 ページと比較的コンパクトな内容になっている。

表 2-6 OMB M-21-27 の構成

<ul style="list-style-type: none">・ エビデンスに基づく統治・ EBPM 法からの機会<ul style="list-style-type: none">➤ エビデンスを構築し活用するためのリーダーシップ➤ 政府全体にわたる学習とエビデンス文化の構築・ 戦略的エビデンス構築<ul style="list-style-type: none">➤ ラーニング・アジェンダと年次評価計画の概要➤ エビデンス計画プロセス➤ 戦略的エビデンス構築計画・ 機関における重要な機能としての評価・ エビデンス構築の制度化<ul style="list-style-type: none">➤ 評価政策➤ 調査、評価、統計及びその他の分析のための能力アセスメント➤ 評価のための資源➤ 進捗状況のモニタリングと報告・ 結論・ 付録 A: エビデンスの構成要素と方法論的アプローチの概要

- ・ 付録 B: 定義(「OMB M-20-12「EBPM 法の実施第 4 段階:プログラム評価の基準と実践」より)
- ・ 付録 C: FY21 および FY22 における EBPM 法の成果物のタイムライン
- ・ 付録 D: 各機関のための実施リソース

出所: OMB M-21-27 をもとに未来工学研究所作成

b. 規定の詳細

以下では、本ガイダンスのポイントをいくつか紹介する。

まず、前述の OMB Circular No. A-11 でも触れたロジックモデルの活用があげられる。本ガイダンスでは、エビデンスの構築にあたっては、それがどのように利用されるか、機関内外のプログラムやマネジメント、規制、運営上の意思決定にどのように役立つのかを考えるべきであるとしているが、「政府全体にわたる学習とエビデンス文化の構築」の項で触れられているように、ロジックモデルはこうしたことを考える上での基盤となりうるツールであることが次図のように示され、重視されていることが分かる。

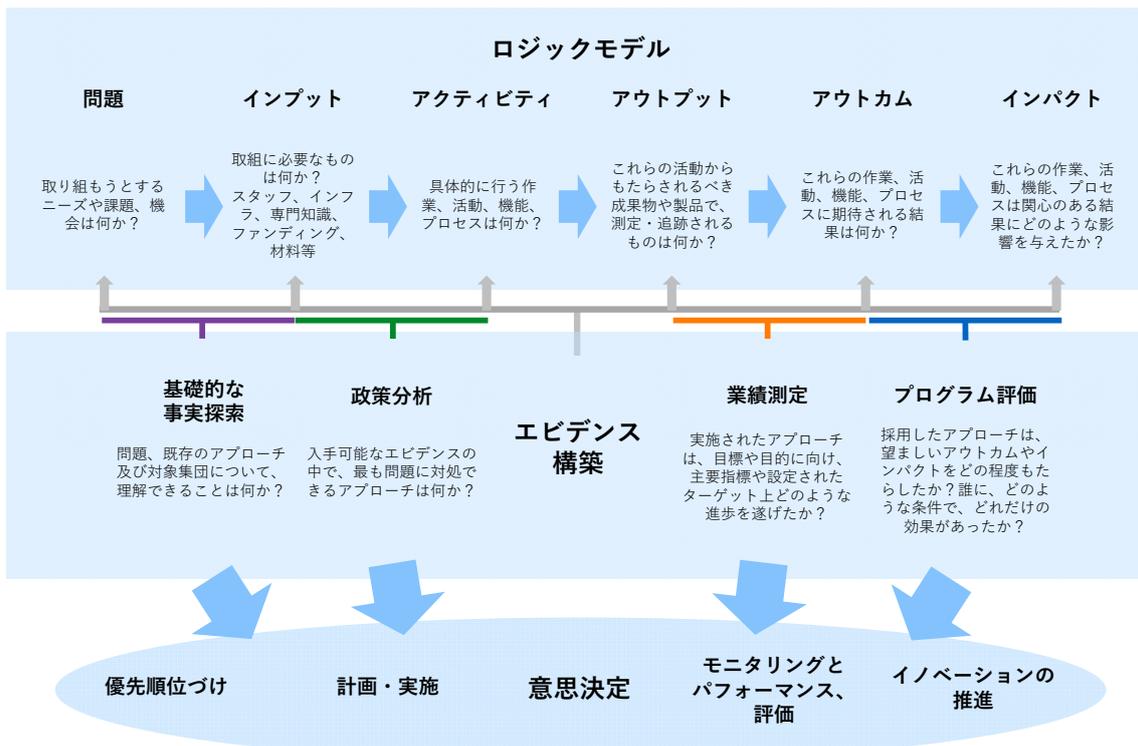


図 2-7 ロジックモデルに基づくエビデンス構築と意思決定への活用

出所: OMB M-21-27 をもとに未来工学研究所作成

また、「エビデンス計画プロセス」の項では、エビデンスが構築、共有、適用され、新たな優先事項が出現するに従って、ラーニング・アジェンダとその関連活動を適応させる必要性が述べられている。

次図は、OMB Circular No. A-11 でも引用されたものであり、4年間の戦略計画とエビデンス計画のサイクルを表したものである。ラーニング・アジェンダや年次評価計画、これらの計画に含まれないその他の関連するエビデンス構築活動を実施することで新たなエビデンスが生み出されると同時に、これらの反復によりエビデンスの基盤が豊かになっていく。OMBでは、こうして蓄積された基盤と毎年度新たに生み出されるエビデンスをもとに、ラーニング・アジェンダの更新と次年度の年次評価計画の作成を行っていくことを理想として掲げている。

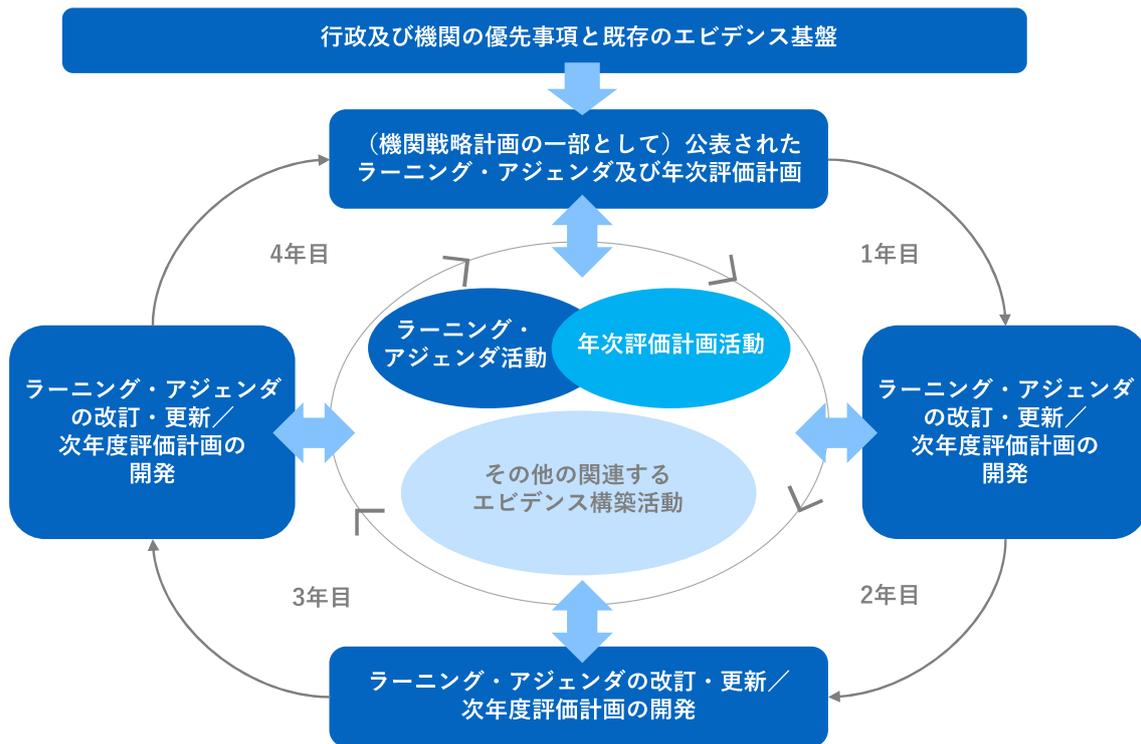


図 2-8 エビデンスのサイクル

出所: OMB M-21-27 をもとに未来工学研究所作成

c. 特記事項

OMBでは、EBPM法に基づく活動を各機関が実施していく上で、追加的なリソースや支援が必要であることを認識しており、その支援の一環として、利用を推奨するいくつかのリソースを付録Dにおいて示している。

そのうちの1つが、MAXという連邦政府職員専用のウェブサイト上に作成された、OMBのエビデンス・チームが運営する「OMBエビデンス及び評価コミュニティ」のページである。これは、エビデンス及び評価に係る実践コミュニティ (Community of Practice) を構築することを企図したものである。職員以外には非公開のものであるが、機関の文書例や、ツールキット及び参照ツール、研修の情報 (過去の研修資料を含む) 等、各機関を支援する

ための精選された資料のライブラリが存在することが、本ガイダンスにおいて示唆されている。

実効的なエビデンス構築活動や評価を実施していくためにはガイダンスのみでは不十分であり、継続的な相互研鑽が欠かせないと言える。

2.2.2 主要機関における研究開発評価システムの体系とその内容

(1) 調査対象機関の概要

全米科学財団 (National Science Foundation: NSF) は、「科学の進展を促進すること、国民の健康と繁栄、福祉を向上させること、国の安全を確保すること」などを目的に、1950年に議会により設立された独立した省レベルの連邦政府機関である。医療科学を除くすべての分野の基礎研究・工学を支援する連邦政府唯一の機関であり、科学・技術・工学・数学 (STEM) 教育の支援も行っている。年間予算は約 101.8 億ドル (2025 年度予算案⁴²) であり、米国の大学で行われる基礎研究に対する連邦政府の支援のうち約 25%の資金源となっている。

NSF では、研究分野ごと (数学・物理科学; 社会・行動・経済科学; コンピュータ及び情報科学・工学; 地球科学; 工学; 生命科学; STEM 教育) に設置された 7 つの部局 (Directorate) 及び 2022 年に約 30 年ぶりに新設された「技術・イノベーション・パートナーシップ局 (TIP)」がファンディングを行っているほか、NSF 長官直属の統合活動室 (OIA) や国際科学・工学室 (OISE) も部局横断的なプログラムも設置し、研究者等の支援を行っている⁴³。なお、評価に関しては、OIA 内に設置の評価能力セクション (Evaluation and Assessment Capability: EAC) が機関全体におけるエビデンスの生成及び利用に関する能力向上を支援するために、エビデンス構築に関わる諸活動を実施するための集中的な技術支援やツール、リソースの提供を行っている⁴⁴。

⁴² NSF ウェブサイト <<https://www.nsf.gov/about/budget/fy2025>> [2025/3/21 取得]。なお、2023 年度予算比で 3.1% 増の要求額となっているが、全米科学ファンディング連合 (CNSF) は、実質的に 2024 年度比で「8% 減という歴史的な予算削減」になっているとし、2023 年度の予算水準である 99 億ドルの予算付与を求めて、下院歳出委員会委員宛に書簡を送付している。CNSF Requests at Least \$9.9 Billion for NSF in FY26 (March 13th, 2025) <https://www.aau.edu/key-issues/cnsf-requests-least-99-billion-nsf-fy26?utm_source=chatgpt.com> [2025/3/21 取得]

⁴³ これらの支援には、いわゆる個人や研究チームを対象とした研究助成 (research grant) だけではなく、機関や設備・施設を対象にしたもの、奨学金の授与なども含まれる。

⁴⁴ Evaluation and Assessment Capability (OIA/EAC) <<https://www.nsf.gov/od/oia/eac>> [2025/3/21 取得]

(2) NSF におけるプログラム評価

1) NSF における代表的なプログラム評価制度：COV による外部評価

前述のように、GPRAMA に基づく評価は基本的にはプログラムを単位とした評価がベースとなっており、各連邦政府機関ではそれぞれ独自のプログラム評価の方法を構築している（未来工学研究所 [1]）。

NSF の場合、代表的といえるものが、Division（課）レベルで実施される外部有識者からなる委嘱審査委員会（Committee of Visitors: COV）によるプログラム評価（外部評価）であり、バイデン政権下で行った昨年度調査では、COV は次のような特徴を持つことが NSF のウェブサイト上で開示されていた（未来工学研究所 [1]）。

- NSF は、助成の決定のための評価・勧告の質を維持するため、統合活動室(OIA)が作成した指針に従い、COV を招集する。
- COV は、各プログラムを約 4 年ごとに審査する。加えて、NSF が GPRAMA に基づいて策定した目標をどの程度達成しているか、専門的視点から評価する。
- COV は、①提起された評価の過程の健全性と効率性、②NSF の投資の結果の、質その他を含めた効果、の 2 つを検討する。
- この作業の後、COV は、勧告や指摘を含め、評価の結果を報告書にまとめて提出する。
- NSF は、提出された勧告に関してどのように対応するかを検討し、COV による報告書に対し書面で回答する。
- COV の評価を形骸化させないために、監査室（OIG）がその活用実態についての監査を行い、その改善についての勧告を行う。
- COV の結果は外部諮問委員会（External Advisory Committee）⁴⁵へフィードバックされる。外部諮問委員会は、各 Directorate や Office（部局）ごとに設置されているものであり、それぞれ 15 名程度の外部専門家（主にアカデミア）で構成される。半年に 1 回の頻度で 1 日半程度かけて開催され、既存プログラムの改善や研究分野への投資などの方向性について議論する場となっている。

一方、2025 年 3 月現在、ウェブサイトで公開されている情報は非常に限定的であり、科学・工学の研究及び教育の実施や支援に関わるプログラムや部局を対象に 4 年ごとに定期的実施すること、①プログラム運営の品質及びインテグリティの評価と②プロポーザルの採否に関するプログラムレベルの技術上および管理上の問題といった 2 領域において外部専門家の判断を提供するものであること、問い合わせれば、法に基づいた情報公開が可能なが記されているのみである⁴⁶。なお、2025 年 2 月 19 日に署名された大統領令「連邦

⁴⁵ NSF ウェブサイト：Advisory Committees <https://www.nsf.gov/about/performance/dir_advisory.jsp> [2025/3/21 取得]

⁴⁶ NSF ウェブサイト：Committees of Visitors <<https://www.nsf.gov/od/oia/activities/cov/>> [2025/3/21 取得]

官僚機構の縮小を開始する⁴⁷」では、各連邦政府機関に設置された諮問委員会など、大統領が不要と判断した官僚機構を廃止する方針が示されており、外部諮問委員会はその対象となる可能性がある。

2) ラーニング・アジェンダに基づくプログラム評価

EBPM 法では、各連邦政府機関に対し、「4年間の戦略計画及びエビデンス構築計画（ラーニング・アジェンダ）」と「年次評価・エビデンス計画」の作成を義務付けているが、ここでは、後者について、調査時点における最新のものとして2025年度版を取り上げ、その内容を紹介する。

a. 評価課題の選定基準

「年次評価・エビデンス計画 2025年度版⁴⁸」では、次の5つの基準に基づき、ラーニング・アジェンダとして評価すべき課題の選定を行っている。なお、これらの基準は、EBPM法の改訂や関連法規の成熟度が高まった場合、NSFが対応すべき優先事項や外部要因が変化した場合、見直される可能性がある。また、選定に当たっては、①から③の個々の基準は必要条件ではあるが、十分条件ではないこと、①から④の基準をすべて満たす課題はリソースの制約がない限り優先される可能性が高いこと、⑤が十分条件として最も重視されることが方針として掲げられている。

表 2-7 評価課題の選定基準

課題選定の基準	概要
①知識のギャップを埋めるものであること	求める情報は、既存の資料(学術文献や他の機関が実施する類似の取組に基づく評価など)から入手できないものであること
②リーダーシップの支援が得られるものであること	業務に必要なスタッフの時間とリソースを優先的に確保し、コミットメントが得られるものであること
③今後の意思決定を支援する可能性があるものであること	実行可能かつ有用なエビデンスを適切なタイミングで得られる可能性が高いものであること
④広範な影響力があるものであること	多様なステークホルダー、プログラム、または組織にとって有用な知見をもたらす可能性が高い
⑤NSFのリーダーシップにとって優先度の高いものであること	変化する要件、議会の指示、および国家的な長期戦略の優先事項に対応していること

出所: Annual Evaluation and Evidence Plan FY 2025 をもとに未来工学研究所作成

⁴⁷ White House: Commencing the Reduction of the Federal Bureaucracy (February 19, 2025)
 < <https://www.whitehouse.gov/presidential-actions/2025/02/commencing-the-reduction-of-the-federal-bureaucracy/> > [2025/3/21 取得]

⁴⁸ NSF, Annual Evaluation and Evidence Plan FY 2025, January 2024 <<https://nsf.gov/resources.nsf.gov/files/NSF-FY25-Annual-%20Evaluation-and-Evidence-Plan%20FINAL.pdf>> [2025/3/21 取得]

b. 選定された評価課題

これらの基準に基づき、選定された優先度の高い評価課題は次の5つである。このうち、③から⑤は2022年8月にバイデン政権下で成立した「CHIPS及び科学法（CHIPS and Science Act of 2022）⁴⁹」に関わるものとなっている。

表 2-8 選定された評価課題の一覧

評価課題	概要	エビデンスの構成要素
①NSFにおけるハイブリッド型労働	パンデミック以降導入され、改訂されてきたハイブリッド型の勤務形態が、新規人材を惹きつける要因となっているか、職員の職務や組織に対する態度にどのような影響を与えているか、定着率や離職率とどのような関係があるか、NSFの多様性や組織文化にどのような影響を与えるか等を検討	業績測定、政策分析
②メリットレビュープロセスにおける採択率の差異に関する精査	PIに人口統計データの提出と標準化されたPIの経歴概要の使用を義務付けた以降のプロポーザルについて、データ分析を実施。採択結果に差異があるか、あるとすれば、PI個人や研究チームの特徴、研究分野、プロポーザルのタイプ、メリットレビュープロセスと関連している可能性はないかを検討	基礎的なファクト・ファインディング
③評価基準「広範囲の影響」の実施に関するプロセス評価	異なる組織単位や関心のある外部関係者において、採択審査基準の1つである「広範囲の影響」がどの程度理解されているのか、基準の適用効果を向上させるためにどのように改善すれば良いかを検討	プログラム評価
④農村地域におけるプレK-12 STEM教育と人材育成	現在の連邦政府プログラム及び研究の質と量について評価し、農村地域における幼稚園から高校3年生までの生徒を対象としたSTEM教育及び人材育成に関する取組を分析、等	基礎的なファクト・ファインディング、政策分析
⑤NSF地域イノベーションエンジンに関するニーズ評価とベースライン評価	地域イノベーションエンジンプログラムについて、NSF Engineを有する地域と、同様の地域でNSF Engineの助成金を受けていない地域との間で、社会経済的及び技術的成果がどのように異なるかを検討	基礎的なファクト・ファインディング

出所: Annual Evaluation and Evidence Plan FY 2025 をもとに未来工学研究所作成

年次評価・エビデンス計画では、これらの評価課題のそれぞれについて、背景・理由、技術的アプローチ、データソース、課題と対応策、利用及び普及といった項目を2ページずつコンパクトにとりまとめられている。例えば、③評価基準「広範囲の影響」の実施に関する

⁴⁹ 同法は、2022年8月9日にバイデン大統領が署名し成立した米国の連邦法であり、国内の半導体製造能力の強化と科学技術分野の研究開発促進を目的とするものである。

議会ウェブサイト < <https://www.congress.gov/bill/117th-congress/house-bill/4346> > [2025/3/21 取得]

プロセス評価については、次のような記載内容となっている。なお、「広範囲の影響」を含むメリットレビュープロセスについては、続くプロジェクト評価の項で詳細に取り上げる。

表 2-9 「③評価基準「広範囲の影響」の実施に関するプロセス評価」の記載内容

項目	概要
背景・理由	<ul style="list-style-type: none"> 1997年に独立した採択審査基準として「広範囲の影響」が正式に導入されたが、依然として懸念や混乱を表明するPIやレビューアが存在
技術的アプローチ	<ul style="list-style-type: none"> NSF職員、経営層、NSBメンバー、外部関係者(PI、レビューア、専門団体等)に対するインタビュー及びフォーカスグループを通じて、NSFの「広範囲の影響」評価基準に関する認識と経験を理解する NSFが資金提供する研究(医療以外の基礎研究、ユーザーに触発された研究、STEM教育)について、プロポーザルの価値(知的メリット及び広範囲の影響または関連概念)を評価するためのエビデンスに基づく実践に関わる、分野横断的な体系的文献レビューを実施(類似の審査基準を有する機関の実践例、有効性に係るエビデンス、潜在的なバイアスに関する懸念及び考慮点、バイアスに関する懸念やレビューアの主観を軽減するための効果的な戦略について重点的に検討)
データソース	<ul style="list-style-type: none"> NSFの管理データ; アンケートやインタビュー、フォーカスグループからの情報; 出版物データベースからの文献計量データ; COVが実施したメリットレビュープロセスの評価報告書; メリットレビュープロセスを直接的な対象とした学術文献及びグレイ・マテリアル
課題と対応策	<ul style="list-style-type: none"> インタビュー及びフォーカスグループから得られるエビデンスの質は、作業チームのスキルに依存。そのため、経験豊富なインタビュー調査員からなるチームの編成、半構造化インタビューの採用、分析のための標準化されたルーブリックを用いて回答をコード化する小規模チームの活用を実施 回答者の記憶バイアスに関連する課題に対応するため、1)対象者に事前に質問を提供、回答を検討する十分な時間を確保する、2)インタビュー及びフォーカスグループの対象者を、過去5年間以内にNSFのメリットレビュープロセスに参加した個人及び過去3年間に参加した若手研究者から選択 可能な限り、NSFの管理データやメリットレビュー調査などの他のデータソースを活用、インタビュー回答者の解釈を支援し、定性的に分類されたデータの信頼性を高める
利用及び普及	<ul style="list-style-type: none"> CHIPS及び科学法が求める「NSF全体において「広範囲の影響」基準がどのように適用されているかを評価し、アメリカCOMPETES法第526条で定められた目標を達成するために、その提供効果を向上させるための提言を行う」ことに対応

出所: Annual Evaluation and Evidence Plan FY 2025 をもとに未来工学研究所作成

(3) NSF におけるプロジェクト評価

1) NSF における標準的なプロジェクト評価制度⁵⁰

NSF におけるプロジェクト評価、特に採択審査の方式は、メリットレビューシステムと呼ばれる共通の仕組みが用いられている。具体的には、「知的メリット(Intellectual Merit)」と「広範囲の影響(Broader Impact)」の2つの基準をすべてのプロジェクトの採択基準として基本的に用いている。プログラムによっては追加的な基準を設ける場合もある。

NSF では、1997 年、全米科学理事会 (NSB) 等による議論を受け、1981 年以来用いてきたレビュー基準の大幅な見直しを行った。その背景には、NSF の事業がより幅広い教育イニシアチブやセンター・プログラムを含むものとして拡大してきたことや、組織目標・戦略と投資の結果との関連性を強調する政府業績成果法 (GPRA) が 1993 年に導入されたことがある。「広範囲の影響」もこの見直しの際に導入されたものである。その後、2007 年の改訂では、原則的にはこれを踏襲した上で、「知的メリット」の基準にトランスフォーマティブ研究の概念を取り入れた。具体的には、「提案する活動は、創造的、独創的もしくは潜在的にトランスフォーマティブな概念をどの程度提示し、探求するものとなっているか」といった記述がメリットレビューにおける考慮すべき事項として盛り込まれた。つまり、トランスフォーマティブであることがすべてのプロジェクトの採択において「積極的に考慮される事項 (positive consideration)」として位置づけられ、2008 年 1 月以降に提案されるすべてのプロジェクトに適用されている。

2つの基準の関連性をより明確化し、総体として社会的目標をより重視しようとする現在の基準は 2013 年 1 月から導入されたものである。その見直し作業は、NSB が 2010 年 2 月に設置した「メリットレビューに関する特別委員会 (Task Force on Merit Review⁵¹)」を中心に進められ、その過程では、民間シンクタンクによる主要な利害関係者からのインプット分析や外部有識者からなる委嘱審査委員会 (COV) 報告書の分析、パブリックコメントなどが行われた⁵²。このように、現在の基準は非常に緻密な分析と幅広い利害関係者を巻き込んだ議論を通じてまとめられたものであり、改訂の際の強力なエビデンスとなっている。

標準的なメリットレビューのプロセスは図 2-9 のとおりである⁵³。

フェーズ I は公募からプロポーザルの受領に至るプロセスであり、まず、公募を周知する (1)。公募期間は 90 日間であり、プロポーザルは電子申請システムを通じて提出することになっている (2)。その後、プロポーザル処理ユニットのスタッフが適切なプログラムへの割り振りを行うとともに、ページ数制限などの形式を満たさないプロポーザルを申請者に返却する (3)。

⁵⁰ NSF ウェブサイト及び未来工学研究所 [6] をもとに執筆。

⁵¹ NSF ウェブサイト < https://www.nsf.gov/nsb/committees/archive/task_force/tskforce_mr.jsp > [2025/3/21 取得]

⁵² この改訂のプロセスについては、遠藤 [7] に詳細が紹介されている。

⁵³ NSF ウェブサイト < <https://www.nsf.gov/funding/overview> > [2025/3/21 取得]

フェーズⅡは実際にレビューを行うプロセスであり、6か月の期間がこれに割り当てられる。レビューアの選定にあたっては、プログラムオフィサーの知識やプロポーザルに記載の参考文献、近年の学会プログラム、プロポーザルに関連する学術雑誌のコンピュータ検索、そして、申請者からの提案を参照する。申請者は、レビューアとして最適な人物に加え、不適切な人物についても提案できる(4)。レビューアによるメリットレビュー基準に基づく評価結果を踏まえ(5)、プログラムオフィサーがプロポーザルを採択すべきか否かについて部門長に対して推薦(6)、部門長が採否の決定を行う(7)。

フェーズⅢでは、予算・財政・助成管理室・契約課(DGA)が契約に関わる検討や調整を行い(8)、最終的に契約の締結となる(9)。この期間として約30日間は費やされる。

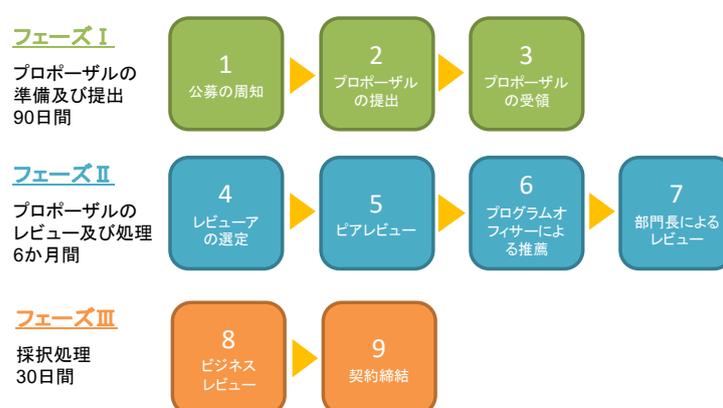


図 2-9 メリットレビューのプロセス

出所: NSF ウェブサイトの情報をもとに未来工学研究所作成

なお、課題(プロジェクト)に対する中間評価及び終了時の評価は基本的には実施されない(拠点形成を支援する工学研究センター(ERC)プログラムなど予算規模の大きなプロジェクト等では、例外的に中間・事後評価が実施される)。その代替として、プロジェクト実施側の義務として、研究代表者(PIもしくはCo-PI)に対し、研究の進捗及び研究費の使途に関して定期的に報告書をまとめることを要求している。具体的には、年次報告、中間報告、最終報告、プロジェクトアウトカム報告書(Project outcomes report)の4種類である。公開データの範囲であるが、実態として、論文以外のデータを記載するプロジェクトは少なく、自己申告のため外部から公式に検証されることはない⁵⁴。公開データを見る限り、Project Outcome Reportsは、科研費の報告書や戦略的創造研究事業の年次報告書、終了報告書よりも文量も少ない。

⁵⁴ Research Gov < <https://www.research.gov/research-web/content/rpprdemosite> > [2025/3/21 取得]

2) GCR プログラムにおけるプロジェクト評価⁵⁵

ここでは、研究開発分野の特性に配慮した採択審査や例外的に中間評価を行っている事例として、「Growing Convergence Research (GCR) プログラム」を取りあげる。

GCR プログラムは、「特定の切実な問題に取り組むことを目的とした収束的アプローチの最も初期段階の基礎を育成し、成長させること」を目的に、「意図的に多様な研究者や関係者を集め、研究課題の設定、分野やセクターを超えた効果的なコミュニケーションの方法、解決のための共通のフレームワークに加え、必要に応じて、新しい科学語彙を開発することを目指す」ものである。すなわち、我が国における総合知のモデルの1つにもなっていると思われる「コンバージェンス研究」を支援するものであり、OIA による調整の下、TIP を除く 7 局が関わる部局横断的なプログラムとなっている。支援対象は、必然的に学際的な研究チームとなる。

GCR プログラムは、フェーズ I (1~2 年目、予算総額最大 120 万ドル)、フェーズ II (3~5 年目、予算総額最大 240 万ドル) の 2 段階に分かれており、提案者は最大 5 年間の研究計画を提案時に提出し、提案が採択されるとまずはフェーズ I に対して 2 年間の資金が提供される、という仕組みになっている。各チームの進捗状況は、2 年目の終わりに評価され、コンバージェンス研究の最前線で更なる進歩を遂げる計画を明確にしたチームのみが最長 3 年間の追加資金を受けることができるというステージゲート方式を採用している。ただし、切るための評価というより、研究チームをエンカレッジするための評価という色合いが強いものとなっている⁵⁶。

このことは、見方を変えれば、採択審査システムが有効に機能している証左とも言えるが、GCR プログラムでは、通常のメリットレビュー基準に加え、次のような 6 つの観点からなる固有の審査基準を採用している。

表 2-10 GCR プログラム固有の審査基準

項目 ^注	GCR において追加される審査基準
目的の妥当性	<ul style="list-style-type: none"> この提案の動機となるビジョンは、新しいコンバージェンス研究の成長に投資することを正当化するに足る十分な説得力と野心を持っているか。 このプロジェクトは、基礎的な科学的理解を変革し、新たな研究の展望を開く可能性があるか。
プログラムとの関連性	<ul style="list-style-type: none"> 提案された研究は本募集にふさわしいか。 提案されたアイデアは、学問分野を超えて深く統合され、NSF の他のプログラム、募集、又は助成制度によって支援される研究とは著しく異なるか。
目標の新規性	<ul style="list-style-type: none"> 研究計画の 2 つのフェーズに概説された目標は、新しいパラダイムやアプローチを開発し、提案のきっかけとなった問題への対処に向けて科学を前進させるのに十分な新規性があるか。
アプローチの妥当性	<ul style="list-style-type: none"> 提案された研究活動は、革新的で将来性があり、コンバージェンス研究の発展にふさわしいか。

⁵⁵ 未来工学研究所 [1] をもとに執筆。

⁵⁶ 昨年度調査では、GCR プログラムによって採択された 108 件のプロジェクトのうち、フェーズ I で終了したものは 7 件であり、約 94%がフェーズ II に進んでいる実態が明らかになっている。

	<ul style="list-style-type: none"> これらの活動は、コンバージェンスを構築し、現在進展を妨げている科学的・技術的課題に取り組むのに適しているか。
研究計画の妥当性	<ul style="list-style-type: none"> 提案されたマネジメント計画は、効果的なコンバージェンス・チームを育成し、意図されたコンバージェンスと研究成果を促進するために適切か。
研究体制の妥当性	<ul style="list-style-type: none"> 集まったプロジェクト参加者とパートナーのチームは、計画されたプロジェクトにとって適切かつ不可欠なものか。 パートナー組織と参加者は有意義に統合されているか。

注) 項目名は未来工学研究所による

出所: PROGRAM SOLICITATION (NSF 24-527), Additional Solicitation Specific Review Criteria, pp.10-11.をもとに未来工学研究所作成

2.2.3 評価疲れについての各国・地域の現況

評価疲れに関して、各種評価制度間の調和が図られており、評価結果の有効活用に向けた制度改善なども行われていることなどから、「評価疲れ」を指摘する公的文書などは見当たらない。ただし、トランプ政権では、評価なきままに予算や人員の削減等が行われており、これまである種の防衛装置として機能していた評価に対する無力感や、長年培われてきた評価文化の衰退につながっていくことも懸念される。

評価疲れを防ぐための制度としては、連邦政府から求められる書類作成等の情報提供に関する報告者（個人、中小企業、教育・非営利機関、連邦政府との契約者、州・地方政府）の負担を最小化することを目的とした Paper Reduction Act がある。OMB では、既存の情報収集に関する活動の検証やエビデンスデータ取得のための情報収集に関する活動についての許認可を行っており、エビデンスデータの収集とペーパーワーク削減のバランスへの配慮が必要となっている。

例えば、NSF では、前述のように、プロジェクト実施側の義務として、研究代表者（PI もしくは Co-PI）に対して、研究の進捗及び研究費の使途に関して定期的に報告書をまとめることを要求しているが、これらの報告書の記入・提出などのやり取りは、研究課題の公募申請時から利用されている NSF 独自のオンラインシステム（Research.gov）⁵⁷が用いられている。オンラインシステムに登録されている成果論文の情報などは、科学技術イノベーションに関与する連邦政府機関の情報を一括検索できる科学情報ポータルサイト（Science.gov）⁵⁸とも連動している。最近の NSF 独自のオンラインシステムに関する取組として、同じ研究者が所属先の異動や兼務で複数持っていた NSF の研究者 ID をひとつにまとめるなどの試みも始められており、研究者に負担をかけずに公的研究費の助成を受けた研究者やプロジェクトのデータを追跡するシステムが整備されつつある。

【参考文献】

[1] 未来工学研究所, 2.2. アメリカ合衆国（米国）, 研究開発評価に関する実態調査・分析業務（2023 年度文部科学省委託調査）, 2024 年 3 月.

⁵⁷ Research Gov < <https://www.research.gov/research-web/> > [2025/3/21 取得]

⁵⁸ Science.gov < <https://www.science.gov> > [2025/3/21 取得]

[2] 未来工学研究所, 2.1. アメリカ合衆国 (米国), 海外の評価及び追跡調査等に関する最新動向調査 (2022 年度 NEDO 委託調査), 2023 年 3 月.

[3] 新日本有限責任監査法人, 「アメリカの政府業績成果現代化法 (GPRAMA) 等の運用から見た我が国の政策評価の実施及び会計検査」に関する調査研究 (平成 26 年度会計検査院委託業務報告書), 2015 年 2 月.

[4] 未来工学研究所, 2.アメリカ合衆国 (米国), 主要国における科学技術・イノベーション政策の動向等の調査・分析 (令和元年度「第 5 期科学技術基本計画のレビュー及び次期科学技術基本計画の策定に関する調査・分析等の委託」), 2020 年 3 月.

[5] 南島和久, 米国 GPRAMA にみる制度改革への視座 —日本への示唆と業績マネジメント—, 評価クォーターリー, 38, pp.45-60, 2016 年 7 月.

[6] Wessner, Charles W., RIO Country Report 2015: United States (EUR 28134 EN), 2016. doi:10.2791/089107

[7] 未来工学研究所, 3.2 米国: 全米科学財団 (NSF), 研究開発評価に関する海外動向調査 (2019 年度科学技術振興機構委託調査), 2020 年 1 月.

[8] 遠藤悟, 米国国立科学財団 (NSF) の評価基準の改訂—基礎科学研究活動が潜在的に持つ社会的インパクトに関する新たな概念の提示—, 科学技術動向, 2013 年 3・4 月号: pp.13-19, 2013 年.

2.3 ドイツ連邦共和国（独国）

2.3.1 国・地域における研究開発評価制度の体系とその内容

(1) 評価制度の全体像と研究開発評価の位置付け

1) 連邦と州の科学・イノベーション政策の所掌範囲（連邦基本法）

ドイツでは、連邦政府と州政府から構成される連邦制がとられているが、ドイツ連邦基本法（Grundgesetz：GG〈以下、連邦基本法〉）⁵⁹で定められているように、各州の権限が非常に大きいことにその特徴がある。ドイツの研究開発評価を考える場合、この連邦制と連邦基本法が、教育・研究における連邦政府と州政府の役割を規定し、研究開発評価の各機関の所掌範囲を示すものとなっている。連邦基本法では、研究政策を含む教育・文化に関する事項について、州に大きな自治権を与えている。他方、同法の第 74 条では、連邦政府と州政府の双方が立法権限を有する「共同立法分野」として、教育・研究（科学研究の振興、国際的な教育交換活動等）も対象として規定し、科学研究の振興等では州法よりも連邦法が優先される条件等が示している⁶⁰。このため、州も一定の権限を持ちながら、連邦政府が基本的な枠組みを設定することが可能となっている。具体的に、連邦政府が主たる責任を有する領域として、公的研究開発資金の配分や科学技術イノベーション政策の立案、大規模な科学プロジェクト（航空、宇宙、海洋、原子力等）の実施等である。一方で、教育については、州の権限となっており、大学の運営は州の権限と責任において行われている。

連邦基本法の改正は適宜行われてきたが、研究に関する改正は、2006 年の連邦制改革（Föderalismusreform I）で、教育・研究分野における連邦政府と州政府の役割が再定義され、連邦政府が実施する国際的な研究協力や大型研究プロジェクトへの資金提供が容易になった。また、連邦政府と州政府が共同で取り組むプロジェクトの範囲が再定義さえ、特定分野（環境・持続可能性、医療・バイオテクノロジー、交通インフラ・技術革新〈自動運転技術、スマート交通システム〉、教育とデジタル化等）の研究協力が促進された。2009 年の連邦制改革（Föderalismusreform II）では、連邦政府の財政支援の拡大（特定の研究プ

⁵⁹ ドイツ連邦基本法は、ドイツの憲法にあたる。1949 年 5 月 8 日に制定され、5 月 23 日に交付された。連邦と州の権限分掌は第 70 条に規定され、州はこの基本法が連邦に立法権限を付与しない限りにおいて、立法権を有すること、連邦と州間の権限分掌は、排他的立法権および共同立法権に関するこの基本法の規程に従うとしている。また、第 71 条では、連邦の排他的立法権に属する事項については、各州は連邦法によって明示的に権限を与えられた場合に限り、またその範囲内でのみ立法権を有としている。なお、ドイツ連邦基本法は、これまで適宜改正されてきており、1990 年の東西ドイツ統一後も改正を経ている。直近では、2025 年 3 月に改正が行われ、財政規律が緩和された（例：防衛費が GDP 比 1%を超える場合の債務ブレーキの対象外となる等）。

〈https://www.gesetze-im-internet.de/englisch_gg/englisch_gg.html#p0364〉 [2025/3/21 取得]

⁶⁰ 第 74 条第 1 項の共同立法権の 9 号に「教育および訓練助成金の規制および研究の促進」がある。また、25 号に「州の責任」がある。同条第 2 項では第 1 項の 25 号から 27 号に基づき制定される法律は、連邦参議院の同意を必要とする。

プロジェクトの競争力強化)、連邦政府と修正府が共同で取り組む研究プロジェクトの範囲の拡大(国際研究協力、大型研究プロジェクトが優先)、研究インフラの整備(研究施設、技術インフラの整備に関する連邦政府の役割の強化、州政府との連携)が図られるようになった。また、2014年の連邦基本法の改正では、基本法第91b条⁶¹にて連邦政府と大学との関係に大きな変化をもたらした。これまで大学の教育と研究は主に州政府の責任であり、連邦政府が大学に対して助成できる範囲は施設建設と期間限定のプロジェクトのみとされ、大学に対して直接資金援助(Direct funding)を行うことは制限されていた。当該条項により、連邦政府は特定のプロジェクト(研究・イノベーション関連)に対して、大学に直接、財政支援を行う権限を確保した⁶²。条項の改正理由は、連邦政府と州政府間の協力関係を促進し、ドイツの科学技術の競争力強化を図るためである。これらから、改正後は州政府の同意(州政府と協定を締結)があれば、基盤的経費の交付が可能となり、高等教育インフラや研究イニシアティブへのより充実した投資も可能になった⁶³。2019年から助成を開始した「エクセレンス・ストラテジープログラム」⁶⁴で初めて制度的な基盤経費の拠出を行った。

2) 科学・イノベーションシステムと研究開発評価の担い手

ドイツの科学技術・イノベーションシステムは、図 2-10 に示すとおりである。連邦政府レベルで、教育・研究政策を所掌するのは、連邦教育研究省(Bundesministerium für Bildung und Forschung : BMBF)であり、研究資金配分機関にドイツ研究振興協会

⁶¹ 基本法第91b条は、教育プログラム及び研究の振興に関するもので、①科学、研究、教育の振興に関し、地域を越えた重要性を有する場合には、連邦と各州は協定に基づき協力することができる。高等教育機関に主として影響を及ぼす協定は、すべての各州の同意を必要とする。この規定は、大型科学施設を含む研究施設の建設に関する協定には適用されない。②連邦と各州は、教育システムの国際比較における実績評価、および関連する報告書や勧告の作成において、相互に協力することに合意することができる。③費用負担の割合は、関連する協定で規定するとした。

⁶² 山岸 [1] によると、連邦基本法91aから91c条までの規定は、連邦基本法30条と70条以下及び基本法104a条第1項に対する特別規定とされる。連邦基本法91b条は、指定領域において連邦と州の特別な協働・協力関係が促進され、同規定は連邦国家における混合行政禁止の例外となっている。

⁶³ ドイツ基本法第91条bに基づき行われた高等教育インフラ、研究イニシアティブとして、DFGが実施した大型研究機器プログラム(Major Research Instrumentation Programme)があり、高等教育機関の大型研究機器に50%を資金提供するものである。審査はDFG合同委員会により選出された科学機器及び情報技術委員会委員により評価が行われる。

⁶⁴ エクセレンス・ストラテジーは、2016年6月16日に連邦政府と州政府によって採択された戦略で、エクセレンス・イニシアティブの後継のプログラムである。大学強化策を継続・発展させることを目的に学術的トップレベルの成果、プロフィールの構築、連携の促進により、ドイツを科学の中心地として持続的に強化し、国際競争力を向上させるためのものである。連邦政府と州政府は、2025年までに毎年総額5億3300万ユーロ、2026年以降は毎年総額6億8700万ユーロ(プログラム、大学への助成金、管理費、段階的廃止のための資金を含む)を拠出し、連邦政府と受入する各州政府が75:25の割合で分配する。エクセレンス・ストラテジーの資金提供方法は、卓越クラスター(大学・大学連合に対するプロジェクトベースの資金提供)に対するものと、卓越大学(大学・大学間ネットワークの強化)に対するものの2つがある。卓越大学への資金提供は、連邦基本法第91b条の改正を活用した取組。〈<https://www.gwk-bonn.de/en/themen/foerderung-von-hochschulen/exzellenzstrategie-exzellenzinitiative>〉[2025/3/21 取得]

(Deutsche Forschungsgemeinschaft : DFG) がある。また、研究実施機関として、大型研究機関を擁する各研究協会 (マックスプランク学術振興協会、ヘルムホルツ協会、フラウンホーファー応用研究促進協会等) がある。

研究資金の配分については、連邦政府と州政府の分担等の取り決めを行う合同科学会議 (Gemeinsame Wissenschaftskonferenz von Bund und Ländern: GWK) と、科学施策への助言・評価機能として科学審議会 (Wissenschaftsrat : WR) がある。

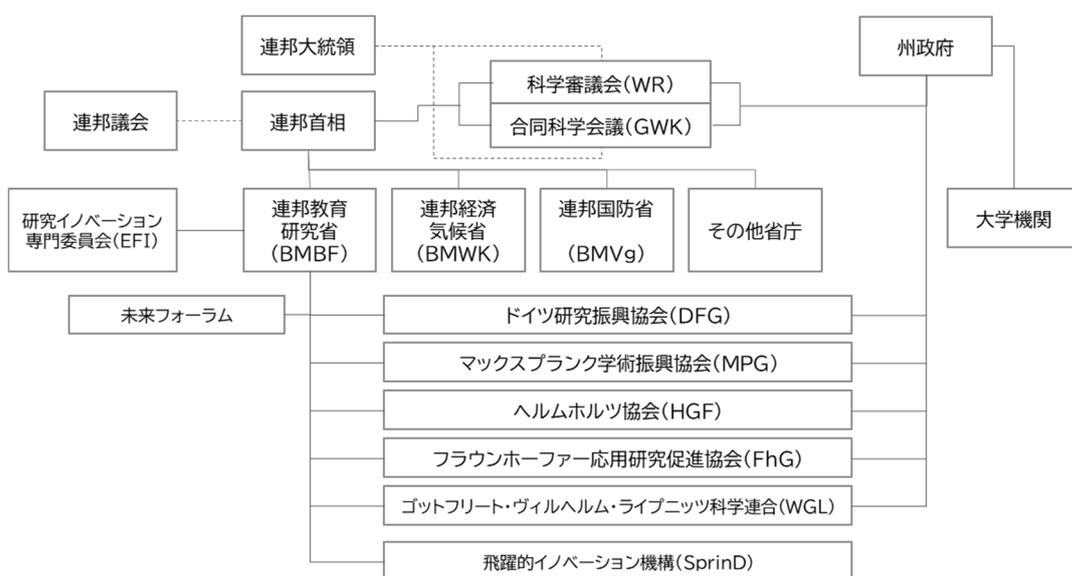


図 2-10 ドイツの科学技術・イノベーションシステム

出所:各種資料に基づき未来工学研究所作成

a. 連邦政府と州政府の合同実施体制と評価の担い手 (合同科学会議、科学審議会)

ア) 合同科学会議 (GWK)

合同科学会議 (GWK) は、連邦政府と州政府が共同で実施する科学・研究政策の戦略立案、科学への資金提供、科学システムに関する共通の関心事項に取り組む組織であり、連邦政府と州政府の調整の場である⁶⁵。GWK のメンバーは、連邦政府と州政府の両方の代表者 (17 名) で構成され、連邦政府側は連邦教育研究大臣 (BMBF) が代表であり、他の連邦省庁 (経済省、厚生省等) も関連する議題は参加可能となっている。州政府側は 16 の連邦州の科学大臣と財務大臣が代表して参加する。なお、GWK の議長については、連邦政府と州政府からそれぞれ 1 名ずつ、共同議長の立てており、議長は毎年、両者間が交代で行っている。GWK のメンバーは、それぞれの権限を維持しながら、国際的な科学・研究政策分野

⁶⁵ GWK は、2007 年 6 月に連邦政府と州政府の長により設立が合意され、2008 年 1 月 1 日より業務を開始した。

における共通の関心事項について緊密な連携を図り、当該分野のドイツの地位を強化することを目指している。

GWK の責任範囲は、主にマックス・プランク協会、フラウンホーファー協会、ヘルムホルツ協会等の大規模研究機関への資金援助及び関連政策、大学インフラへの資金援助、共同研究プログラム、エクセレンス・ストラテジーのような戦略的研究イニシアティブ及び関連の合意に関する責任を負っている。

GWK の法的な役割については、2014 年の連邦基本法第 91b 条の改正と 2007 年の「連邦政府と州による合同科学会議の設立に関する行政協定」（GWK 協定）により明確になっており、連邦・州政府間の合意に準拠した連邦・州政府による共同ファンディングの調整・連携促進を行うことである（研究イノベーション協定、エクセレンス・ストラテジープログラムの施策、連邦と州が合同で基盤経費を拠出する公的研究機関等への各種施策）⁶⁶。なお、GWK は、連邦大統領府（Bundespräsidialamt）に属し、中立的な立場を維持し、合議の結果報告は、連邦教育・研究省の大臣に対して行うことになっている。また、GWK で全会一致で下した決定については、連邦ならびに州政府の長による決定と同等のものとみなされる。

表 2-11 と表 2-12 は、連邦基本法第 91b 条第 1 項に基づく、連邦政府と州政府の共同出資事業の分担を示したものである。大学支援事業では、エクセレンス・ストラテジーは、連邦政府と受入する各州政府の分担割合は 75:25 となっている。また、非大学支援事業では、DFG は連邦政府と州政府の分担割合は 58 : 42 となっている。

表 2-11 連邦基本法第 91b 条第 1 項に基づく連邦政府と州政府の共同出資事業
(大学支援)

No.	施設／資金調達策	分担 (連邦政府)	分担 (州政府)	州負担分の増加
1	教育による進歩	100%	0%	各国又はスポンサー組織が全体的な資金調達を確保する
2	エクセレンス・ストラテジー	75%	25%	位置する州
3	応用科学大学の研究開発	100%	0%	ホスト州は基本設備の資金調達枠組み内でプロジェクト関連支出総額に寄与する
4	研究用建物、大型設備、国立高性能コンピューティング	50%	50%	研究用建物および大型設備：居住国、全国高性能コンピューティング：FGH 手続き原則に準ずる
5	大学協定Ⅲ（プログラムライン：追加 1 年生の受入）	50%以下	50%以上	連邦政府は追加学生一人当たり、算出された費用に対して 13000€を拠出する。各州は全体的な資金調達を確保し、連邦政府と同等の拘束力のある財政的貢献を行う
6	高等教育におけるイノベーション * 2021～2023 年 * 2024 年～	100% 73.33%	0% 26.66%	※ケーニヒシュタイン・キーに基づく全州
7	革新的な大学 (Innovative University)	90%	10%	位置する州

⁶⁶ 連邦政府と州政府による共同出資の可能性は、連邦基本法第 91b 条第 1 項に記載されているが、2007 年の「連邦政府と州による合同科学会議の設立に関する行政協定」（GWK 協定）で明確になっている。
< <https://www.gwk-bonn.de/die-gwk/aufgaben> > [2025/3/21 取得]

No.	施設／資金調達策	分担 (連邦政府)	分担 (州政府)	州負担分の増加
8	応用科学大学スタッフ	71%	29%	位置する州
9	教授プログラムⅢ	50%	50%	位置する州
10	教員養成のための質的向上策	100%	0%	—
11	教育のための質的協定	100%	0%	それぞれ位置する州が全体的な資金調達を確保する
12	若手科学者育成プログラム	100%	0%	それぞれ位置する州が全体的な資金調達を確保する
13	研究と教育を強化する未来契約	50%	50%	それぞれ位置する州

注: ケーニヒシュタイン・キーは、連邦と州の共同資金調達分担金の配分を規定するもの。主に DFG、MGP、ライプニッツ協会の各機関への助成金における連邦州の分担金の配分に用いられる。ケーニヒシュタイン・キーは、科学研究機関への資金調達手段として 1949 年に導入されたケーニヒシュタイン合意に由来。ケーニヒシュタイン・キーは、各州の税収と人口に基づいて算出される。

出所: GWK ウェブサイト⁶⁷より未来工学研究所作成

表 2-12 連邦基本法第 91b 条第 1 項に基づく連邦政府と州政府の共同出資事業
(非大学研究機関)

No.	施設／資金調達策	分担 (連邦政府)	分担 (州政府)	州負担分の増加
1	Acatech	1/3	2/3	バイエルン自由州が 50%、ケーニヒシュタイン・キーに従って全州が 50%
2	ドイツ研究振興協会 (DFG)	58%	42%	ケーニヒシュタイン・キーに従って全州 ※大学協定Ⅲの管理協定の付属文書 2 に従うプログラム
3	DZHW	70%	30%	ケーニヒシュタイン・キーに従って全州
4	フラウンホーファー協会 (FhG)	90%	10%	州に所在する FhG 機関の資金ニーズの割合に基づく。1/3 はケーニヒシュタイン・キーに基づく
	FhG 拡張措置 (特別資金)	50%	50%	
5	HGF	90%	10%	通常、当該州が負担する。個々のセンターについては異なる規定がある
6	ライプニッツ協会	50%	50%	主要研究プロジェクト: 設置される州 さらに • 科学インフラの任務を相当程度担う機関; 位置する州 = 75%、ケーニヒシュタイン・キーに基づく全州 = 25% • その他機関: 位置する州 = 75%、ケーニヒシュタイン・キーに基づく全州 = 25%
7	レオポルディナ (Leopoldina)	80%	20%	位置する州
8	マックス・プランク協会 (MPG)	50%	50%	位置する州が 50%、ケーニヒシュタイン・キーに基づく全州が 50% ※別に若手アカデミー助成金は 90 : 10 の割合で、州負担分はザクセン＝アンハルト州とベルリンで均等に分割
9	ベルリン科学アカデミー	50%	50%	位置する州

出所: GWK ウェブサイト⁶⁸より未来工学研究所作成

⁶⁷ GWK ウェブサイト: Finanzierungsschlüssel < <https://www.gwk-bonn.de/themen/finanzierung-von-wissenschaft-und-forschung/finanzierungsschlüssel> > [2025/3/21 取得]

⁶⁸ 前掲 GWK ウェブサイト

GWK では、連邦政府と州政府の合意と予算配分が必要な分野に限り議論を行い、政策提言や評価等はあまり実施していないとされるが（CRDS [2]）、いくつかの共同研究プログラム／プロジェクトでは、モニタリング評価等の評価が、各プログラム等の担当機関・組織レベルで行われている⁶⁹。GWK では、研究・イノベーション協定（Pakt für Forschung und Innovation：PFI）モニタリング報告書は、毎年公表している。研究・イノベーション協定とは、①ダイナミックな発展の促進、②産業界及び社会への移転の強化、③ネットワークの深化、④優秀な人材の誘致と確保、⑤研究インフラの強化を行うためのプログラム等を組織の包括的な研究政策の目標として掲げ、フラウンホーファー協会、ヘルムホルツ協会、ライプニッツ協会、マックス・プランク協会等の研究機関及び研究助成機関のドイツ研究振興協会（DFG）への資金提供の枠組みを示した施策である。同協定ではこれらの機関に独自の財政計画の安定性を保証し、その代わり、これらの機関は共通の研究政策目標の追求を約束する形態をとっている。この共同事業は、2005年から実施され、2021年に第4期（10年間）が始まり、2030年に終了予定である。

同協定による運営組織は、目標達成に向けた進捗状況を連邦政府及び州政府に毎年報告することになっており、様々な指標を用いてミッション別に進捗状況を評価し報告書に取りまとめている。研究・イノベーション協定モニタリング報告書 2024 は、3巻で構成され、現行の協定期間の中に達成されたことを評価（中間評価に相当）し、第2巻では具体的な実施状況が示した（第3巻はPFI組織の個々の報告書を記載）⁷⁰。

また、エクセレンス・ストラテジーについては、ドイツ科学審議会（WR）が大学の選定と評価を行い、プログラムのモニタリングを実施している。WRは、GWKに定期的に報告書を提出し、2027年6月にプログラムの進捗状況と影響評価に関する報告書を公表予定である⁷¹。

イ) 科学審議会（WR）

科学審議会（WR）は、連邦政府と州政府の協定に基づき 1957年9月5日に設立され、両政府への科学的な審議・評価（助言）等を通じて、a) 教育の拡大と高等教育システムの拡張、b) 科学と研究における有効性と効率性の問題への対応、c) ドイツ統一とその影響、d) 科学システムの差別化と国際化の傾向等の重要な科学政策の推進を担っている⁷²。

⁶⁹ GWKで調整した共同研究プログラム/プロジェクトの評価として、i) テニューア・トラック・プログラム評価、ii) 全国教育パネル調査（NEPS）の評価、iii) エクセレンス・ストラテジープログラムのモニタリング評価（現行）、iv) エクセレンス・イニシアティブの評価（終了済）等がある。

⁷⁰ 研究イノベーション協定モニタリング報告書 2024

< https://www.gwk-bonn.de/fileadmin/Redaktion/Dokumente/Papers/PFI-Monitoring-Bericht_2024_Bd_I_barrierefrei.pdf > [2025/3/21 取得]

⁷¹ WR ウェブサイト

< https://www.wissenschaftsrat.de/EN/Fields-of-Activity/Competitive_review_processes/Excellence-Strategy/Programme_monitoring/program_monitoring_node > [2025/3/21 取得]

⁷² WRの担当分野は、科学と高等教育システム、評価（Evaluation）、競争力の評価（Competitive assesment）、認定（Accreditation）、WR業務計画の策定（1月、7月に更新）からなる。

WR では、評価業務の運営組織として「評価委員会」を設け、評価業務を遂行するための作業グループを設置している。WR の評価委員会は、主に機関評価を担当しており、機関評価に関連して、個々の研究分野の横断的な評価、個々のテーマのシステム評価、構造研究 (Structural studies) 等を実施している。機関評価、施設評価に係る基準には、研究業績の評価 (研究プログラム、パブリケーション、第三者資金、品質保証、国内外の共同研究、教育、若手研究者の支援)、インフラ及び移転サービス、組織及び設備等が含まれる。また、プログラム志向の資金提供については、プログラム評価を実施している⁷³。評価のためのガイドラン・規則は、前述のとおり、それぞれの業務を担当する委員会レベルで策定されており、狭義の研究業績だけではなく、評価対象に応じて組織の本質的な側面や科学環境への統合も考慮し、プログラムの特性を考慮した評価が行われている。

表 2-13 WR の評価ガイドランの例⁷⁴

評価ガイドライン
● 大学医学部評価のためのガイドライン (Guidelines for the Assessment of University Medical Institutions (Drs. 1968-24), July 2024 Guidelines 2024 medicine)
● 連邦機関の研究業務部局に関する機関評価ガイドライン (Guidelines for the institutional evaluation of federal institutions with departmental research tasks (Drs. 1010-23), January 2023 Guidelines 2023 Departmental research)
● 学術機関の機関別評価に関するガイドライン (Guidelines for the institutional evaluation of scientific institutions (Drs. 8823-21), January 2021 Guidelines 2021 Evaluations)
● 大学医学部施設評価ガイドライン (Guidelines for the Evaluation of University Medical Facilities (Drs. 6867-18), January 2018 Guidelines 2018 medicine)
● WR 評価委員会の任務、基準および手続き (Tasks, criteria and procedures of the Evaluation Committee of the Science Council (Drs. 4205-14), as of October 2014 Guidelines 2014 Evaluations)
● 連邦機関の研究業務部局を担う機関評価委員会の評価基準 (Criteria of the Evaluation Committee for the Assessment of Institutions with Departmental Research Tasks of the Federal Government (Drs. 3078-13), July 2013 Guidelines 2013 Departmental research)
● WR 評価委員会の業務、評価基準および手順 (Tasks, criteria and procedures of the Evaluation Committee of the Science Council (Drs. 8922-09), as of January 2009)

⁷³ WR ウェブサイト

<<https://www.wissenschaftsrat.de/DE/Aufgabenfelder/Evaluation/Leitfaeden/leitfaeden>> [2025/3/21 取得]

⁷⁴ WR ウェブサイト

<https://www.wissenschaftsrat.de/SiteGlobals/Forms/Suche/Publikationssuche_Formular.html?nn=24018c45-88fa-435f-be1f-1d2805f8c8fc&cl2Categories_Typ=download-typ-1f&submit=Senden&cl2Categories_Themen=sachgebiet-evaluationen> [2025/3/21 取得]

WR では、エクセレンス・ストラテジーの枠組みのうち、大型研究インフラの評価、私立大学の認定、研究施設の評価（Assessment）と、各州の高等教育制度の構造評価の枠組みのうち、個々の研究機関や大学付属医療施設の評価（Evaluation）等を行っている⁷⁵。2025年1月から7月までのWRの作業プログラムでは、WRが卓越大学支援プログラム等の作業手続きを策定している。WRに設置されたエクセレンス・ストラテジー委員会では、プログラムの効果を検討し助言を行っている。また、当該委員会では、2021年に卓越大学を評価するための評価ガイドを作成した（2024年に更新）。

b. 国レベルの科学技術・イノベーション政策評価（研究イノベーション専門委員会）

国レベルの科学技術・イノベーション政策（連邦政府が実施する戦略等）の助言・評価は、研究イノベーション専門委員会（Expertenkommission Forschung und Innovation（Expert Commission for Research and Innovation）：EFI）が担っている⁷⁶。EFIは、連邦教育研究省により設置された専門委員会で、6名の専門委員（科学・イノベーション政策に関する大学教授等）で構成され、ドイツの科学技術・イノベーションシステムの強みと弱みに関する包括的な分析を行うとともに、研究・イノベーション拠点としてドイツの将来性や政策全般の評価（Assessment）を実施し、毎年、首相に対して報告書を提出し政策提言を行っている。専門委員会のタスクは、①経済学及び社会科学、教育経済学、工学及び自然科学におけるイノベーション研究、並びに技術予測に関連する学際的な議論をまとめること、②科学政策提言からなる。科学政策提言は、a) ドイツの研究・イノベーションシステムの構造・傾向・実績・見通しの分析（時間的及び国際比較の観点）、b) 研究・イノベーションシステムに関連する主要問題の評価（Assessment）、c) 研究・イノベーションシステムのさらなる発展に向けた行動の選択肢と提言の策定からなる。年次報告書は、毎年3月1日に首相に提出される。

参考：年次報告「Opinion（GUTACHTEN）2025」

- 本報告書は、退任する連邦政府（2024年11月6日に社会民主党（SPD）、緑の党、自由民主党（FPD））の連立政権の崩壊に提出された後に公開され、次期連立協議の推進力となることを目的に作成したものである。
 - ・ 主な内容：政治への信頼回復、経済と社会の不安軽減、有望な見通しを打ち立てることができる新政権の樹立を希求。
- EFIでは、イノベーション経済および研究・イノベーション政策の観点から特定された課題の解決に貢献することを目指している。（デジタル化と脱炭素化課程における構造転換の分析、量子技術等）

⁷⁵ エクセレンス・ストラテジーの内、エクセレンス・クラスター支援プログラムはDFGに委託し実施している。

⁷⁶ 最初の科学技術・イノベーション政策（科学政策）の提言は、2008年2月に提出され、年次報告書「オピニオン」では国際比較及び歴史比較におけるドイツのイノベーションシステムの強みと弱みに関する包括的な分析が含まれる。

- 第20期議会（2021-2025）：多くの面で目標が達成されず、目標の実現に必要な一貫性を欠いていた。戦略面では「未来研究・イノベーション戦略」が導入され、包括的かつ省庁横断的な枠組みが提示された。主要な政策として、研究開発投資の拡大／デジタル化イニシアティブ／新興技術への支援／教育・研究機関の強化／気候変動と持続可能性に関する研究等が行われた。
- 第21期議会（2025-）への提案：適切なガバナンス構造の確立（スタートアップやイノベーションに関する責任のBMBFへの再統合等）

本報告では、連邦政府の欠陥（遅さ）により、連立合意で発表された研究イノベーション政策のプロジェクト（ドイツ技術移転・革新機関（DATI）、研究データ法、データ研究所、学術分野の有期雇用契約法、科学ジャーナリズムのための独立財団の設立等）の多くは未実施であると報告した。また、将来戦略（「研究とイノベーションのための未来戦略（Umsetzung der Zukunftsstrategie Forschung und Innovation）」）で策定された6つの主要分野⁷⁷は広範にわたり定義されており、省庁横断的な目標を導出した施策を策定することができていないとした。ミッション志向の施策を実施するための具体的な運営や飛躍的イノベーション庁（SPRIND）とDATIを補完する「移転ブリッジ」、「イノベーション地域」の進展はわずかと評価した。



EFIでは、ミッション指向型研究開発をEUと協力して推進しているが、これらの評価もEFIが担い、プログラムを所管する府省が研究プログラムの目的の達成度、プログラムの効果や効率性等の情報収集・分析を担っている。

府省レベルの科学技術・イノベーション政策の担い手（連邦教育・研究省）

連邦政府レベルで、教育・研究の所管省庁は、連邦教育研究省（BMBF）であり、新たな研究テーマに取り組み、イノベーションを促進するための制度を導入している。BMBFは、ドイツのSTI政策の主要省庁で、連邦政府の研究開発予算の約50%を管理している。BMBFでは、社会イノベーションを包含し、研究・イノベーションプロセスにおける社会を主要なアクターとして位置づけた「拡大されたイノベーション」と定義し、有望な技術の研究開発を支援している⁷⁸。また、BMBFのタスクには、教育、科学、研究における国際交流に加え、合同科学会議（GWK）における各州との協力も含まれる。

⁷⁷ 6つの分野には、①資源効率に優れ、循環経済志向の競争力のある産業と持続可能なモビリティの実現、②気候保護、気候適応、食糧安全保障、生物多様性の保全の推進、③全ての人々の健康の改善、④ドイツ及び欧州のデジタル及び技術的主権（技術主権）の確保とデジタル化の潜在能力の活用、⑤宇宙旅行、宇宙探査、宇宙及び海洋の保護と持続可能な利用の強化、⑥社会の回復力、多様性、結束の強化等である。

⁷⁸ BMBF, Ministry - Task and structure

研究イノベーション政策に関する評価（他の法的枠組みとの関連）⁷⁹

政策評価については、ドイツ連邦予算法（Bundeshaushaltsordnung：BHO）では、連邦予算に影響を与えるすべての施策が「経済的実現可能性調査」の対象となることが義務付けられている。行政規則（連邦予算法典〈VV-BHO〉一般管理規則）では、当該調査にパフォーマンス評価を含めることが義務付けられている。具体的には、政策介入の意図した目的がどの程度達成されたか、経済効率性はどの程度か、政策介入が実際に目的達成の原因となったかどうかを判断することを目的としている。

2009年から2023年までに連邦教育研究省（BMBF）と連邦経済・気候行動省（BMWK）が委託した政策評価研究レポート（81件）では、レポートの59件（73%）が政策介入の因果関係を分析しまとめている。

ドイツの研究イノベーション政策は多種多様な資金調達アプローチを用い、極めて状況に特化している特徴がある。研究イノベーション政策は、基礎研究、応用研究、製品開発等で技術、セクター、地域の対象が異なるほか、近年はミッション志向型政策に対する新たな政策設計が必要になっている。

(2) 研究開発評価に関する規定等の整備状況

ドイツでは、国家レベルで研究開発評価に関する体系的な規程等は整備されていない。このため、研究開発評価は、研究開発プログラム単位でそれぞれ行われている。研究開発評価に関する規程等が整備されていない背景として、ドイツでは、長らく、教育・研究は主に州政府の所掌範囲であり、連邦政府が特定プロジェクト（航空、宇宙、海洋、原子力等の大規模な科学プロジェクト）に限定したものであった。2014年の連邦基本法の改正（第91b条）により、州政府の合意に基づき、連邦政府が研究・イノベーション関連の特定プロジェクトの大学に直接財政支援を行う権限を確保した一方で、特定プロジェクトの実施内容は、研究インフラ、研究イニシアティブに対する投資の中心となっており、研究開発プログラム／プロジェクトは限られる。また、連邦基本法の改正により、連邦政府が直接投資できるようになったプログラムは、2019年から助成を開始した「エクセレンス・ストラテジープログラム」（2005年に導入されたエクセレンス・イニシアティブの後継）が初めて制度的な基盤経費の拠出を行った例であり、体系的な研究開発評価が必要とされる機会がまだないものと考えられる^{80,81}。

< <https://www.bmbf.de/DE/Ministerium/AufgabenUndAufbau/aufgaben%20und%20aufbau.html> >

[2025/3/21 取得]

⁷⁹ Büchele [3]

⁸⁰ エクセレンス・ストラテジーは、連邦政府による研究開発プログラムであるが、クラスターの構築（国際的な評価の高い競争力のある研究を領域横断的に実施可能なネットワークの構築）、拠点大学に対する助成に係る資金である。

⁸¹ 2005年に連邦政府と州政府は、科学研究の中心地としてのドイツを長期的に強化し、国際競争力を向上させ、トップレベルの大学研究をより可視化することを目的として、エクセレンス・イニシアティブを立ち上げた。同プログラムは、2006/2007年から2011年、2012年から2017年の2段階で行われ、ドイツの大学に約46億ユーロの資金を提供した。2009年より合同科学会議（GWK）が連邦政府と州政府に代わ

それ以外の背景要因として、連邦政府が所掌するプログラムが教育、施設、研究インフラに関するものが中心であることのほか、それらプログラムの資金が連邦一州の分担割合もそれぞれ異なるため、トップダウン型で体系的な評価規定等の策定が難しいことも考えられる。

(3) 研究開発評価に関する規定等の詳細

国レベルの府省統一の研究開発プログラム等の評価規程は見当たらない。WR はエクセレンス・ストラテジーの卓越大学プログラムの評価を行っているが、WR では各評価対象に応じて個々の委員会で評価基準を設定している。

一方で、ドイツにおける評価研究の基盤として、1997年に設立された評価学会「DeGEval (Gesellschaft für Evaluation e. V.)」⁸²がある。DeGEval は、評価の専門化を推進し、評価に関する様々な視点を集約し、情報交換や意見交換の場を提供し、2001年に評価基準 (DeGEval Standards) を策定 (2016年改訂) している⁸³。また、DeGEval では、個々の評価分野内及び評価分野間の評価者間の交流を促進し、様々なワーキンググループの活動が行われている。研究・技術・イノベーション政策ワーキンググループ (約 200名の評価専門家、関係者が参加) では、研究・技術・イノベーションプロジェクトや関連機関の科学的な質、成果を検証し、公的資金による措置の様々な効果を調査している。同 WG の研究会議 (年次大会) では、連邦政府のインパクト志向と評価文化に関する各省庁の現在の方法論に関する議論やインパクト志向のモニタリングシステムの構築、オーストリアのプログラム事例に基づくトランスフォーマティブ・イノベーション政策手段の評価等を議論している。DeGEval は、個人会員が 519名、法人会員が 212機関である。法人会員の中には、フラウンホーファー協会システム・イノベーション研究所 (FhG-ISI)、プロジェクトトレーガーの VDI/VDE イノベーション&テクニク社 (VDI/VDE Innovation + Technik GmbH)、ユーリッヒ研究センター (Research Center Jülich GmbH)、ドイツ航空宇宙センター (DLR) 等が含まれている。また、連邦経済協力開発省 (BMZ) の包括的開発評価プログラムでは、OECD 開発援助委員会 (DAC) の関連原則・基準とともに、DeGEval の

り、エクセレンス・イニシアティブの効果を独立した国際的な専門委員会 (エクセレンス・イニシアティブ評価国際専門家委員会 (International Commission of Experts for the Evaluation of the Excellence: IEKE) に評価させることを決定した。評価結果を踏まえ、GWK ではエクセレンス・イニシアティブと同程度の継続賛成を表明し、フォローアッププログラム (エクセレンス・ストラテジー) の要件を勧告した。

⁸² DeGEval ウェブサイト <<https://www.degeval.org/en/about-us/>> [2025/3/21 取得]

⁸³ DeGEval の評価基準の作成にあたっては、会員の多くが米国評価協会 (1999年/1995年)、カナダ評価学会のガイドランと同様のものもしくは追加したものを求めていた。2000年のベルリン総会で DeGEval の評価基準の策定決議を可決し、理事会に作業計画の作成と委員会の設置を委任している。作業計画では、教育評価基準合同委員会 (JCSEE 基準 (2000/1994) およびスイス評価学会の関連基準 (SEVAL 基準 2001) に整合させる指示が含まれていた。これらの検討結果、2001年10月4日の総会で満場一致で決定し、施行された。評価基準は、有用性、実現可能性、公平性の3つの分野からなり、有用性では関係者及び影響を受ける当事者の特定、評価の目的の明確化、評価者の能力及び信頼性、情報の選択及び範囲等の8つの項目で構成される。実現可能性は、適切な手続き、外交的アプローチ、評価効率性等の3つの項目、公平性は正式な合意、個人の権利の保護、包括的かつ公平な審査等の5つの項目からなる。

政策横断的基準に基づいた評価を行うとされる。これらから、ドイツ国内で研究開発評価を行う際の知識基盤として、DeGEval の評価基準の影響も考えられる。

2.3.2 主要機関における研究開発評価システムの体系とその内容

(1) 連邦政府レベルの研究開発評価（連邦教育研究省のプログラム評価の例）

BMBF が助成した研究プロジェクトの有効性と影響力を評価するために、体系的な評価を実施している。例えば、BMBF が実施している「持続可能な開発のための研究促進プログラム (FONA) のフレームワークプログラム」⁸⁴の評価では、目的達成度、影響、支援策の有効性及び費用対効果に重点が置かれる一方で、フレームワークプログラムの成果の側面（さらなる発展に向けた展望の特定、方向転換等の戦略的側面）からの評価も実施した（Bührer et al. [4]）。FONA 評価では、①ドイツ及び国際的な持続可能性研究にどのような貢献をしたか、②持続可能な開発目標の達成にどのような貢献をしたか、③ドイツの科学システムにどのような変化をもたらしたかの 3 つの側面に焦点を当てている。評価対象のプログラムは、2004 年から 2015 年までの期間と、FONA3 までを考慮すると 2018 年初頭までが対象となり、多様性と経時的な変化を適切に評価するため、インパクト分析では理論に基づく評価設計を行い、貢献度分析 (Contribution analysis) を用いている。貢献度分析については、介入ロジックのより包括的な形態である「変化の理論 (Theory of Change : ToC)」が中核となっている。

図 2-11 は、FONA の資金提供によるインパクトモデルであるが、ミクロレベルでの効果は、助成金交付先レベルでの効果を検証している。これには、助成対象の研究プロジェクトの従来成果測定（論文数）が含まれる。また、行政や市民社会の関係者とのネットワークでは、方向性を示す知識提供が中心的な成果として評価された。（資金提供を受けた関係者間の国内外の）ネットワーク化と協力行動では、科学分野だけでなく、科学者と非科学分野の関係者間の協力も含まれた。若手科学者の育成は、当該プログラムが強く推進する目標となっており、プロジェクト終了後も、科学分野の内外で研究テーマの確立に貢献する重要な知識の担い手として検証の対象としている。マクロレベルでの効果については、助成対象プロジェクトに関与する組織への効果（成果の活用と移転）に焦点を当てるとともに、研究プロジェクトの開始による研究環境の変化、研究に対する新たな需要により助成活動や助成形態への変化にも焦点を当てている。科学システムレベルでは、FONA によりドイツの持続可能性研究がどの程度確立され強化されたか（例えば、研究テーマの変遷等）が検証されるとともに、これらの変化がどの程度、個々の FONA の施策の優先事項に起因するものであるかも検証された。これらの検証には、例えば、企業の新製品や新サービスの開発や行政のガイドライン、マニュアルへ適用等の影響に焦点を当てた分析が行われた。最後に資金調達条件の経年変化が分析され、FONA の枠組みで求められる省庁間の連携の増加、他の連

⁸⁴ FONA の評価は、2005 年から 2009 年にかけて FONA1 と、2010 年から 2014 年までの FONA2 を対象に実施した。

邦省庁との調整の協力状況等が評価の対象とした。マクロレベルでの評価は、経済全体、社会、政治レベルでの影響が対象となるため、科学から応用への移転プロセスは不均一性と複雑性により、厳密な定量化は不可能としている。

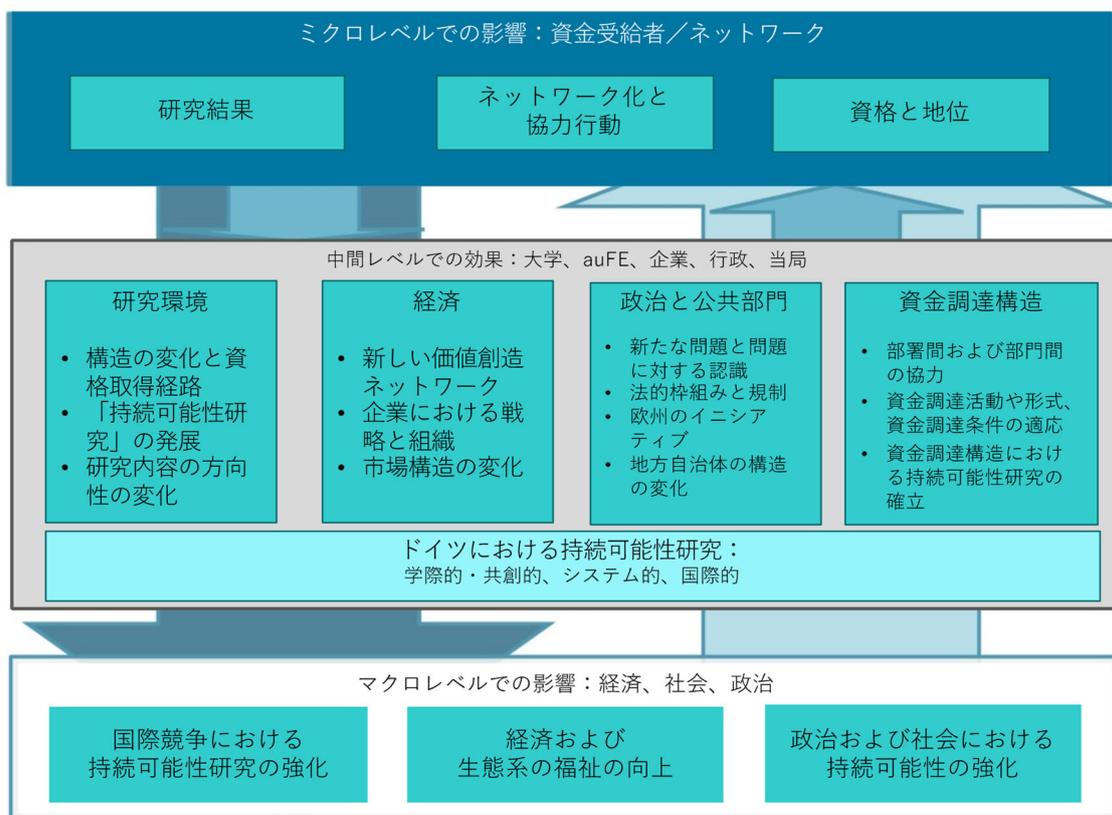


図 2-11 FONA の資金提供によるインパクトモデル

出所：Bührer et al. [4]を元に未来工学研究所作成

Bührer et al. [4]によると、FONA の評価結果は、ドイツの研究システムへの影響と、持続可能な開発に向けた経済・政治・社会の変革プロセスを刺激したと評価されている。

参考：マリーキュリーアクションプログラムの評価(人材育成プログラムの評価例)

BMBF が実施しているプログラムに、EU で推進しているマリーキュリーアクション (Marie Skłodowska-Curie Actions : MSCA) がある。MSCA は、研究者のキャリア発展と国際的な研究協力支援を行うための研究助成プログラムである。助成スキームには、博士課程ネットワーク、ポスドクフェローシップ、スタッフ交換、共同資金提供、科学を一般市民に近づける活動等がある。当該プログラムの評価では、ピアレビューのプロセスで、①卓越性 (研究の科学的な質と革新性の評価)、②インパクト (社会、経済、科学における潜在的な利益の評価)、③実施 (プロジェクトの実行可能性と効率性の評価) の基準の 3 つの区分に分けて採

択段階の評価を実施している⁸⁵。採択評価の実施にあたっては、少なくとも3名の専門家が提案書類の審査を行う(3名の評価者はそれぞれの評価結果は知らせずにバーチャル会議で最終評点を合意する)。

(2) ドイツ研究振興協会 (DFG) における研究開発評価

1) ファンディングプログラムの概要

ドイツ研究振興協会 (DFG)⁸⁶は、主に大学の基礎研究に対してボトムアップ型で研究者主導の研究資金を助成している。DFGの資金配分のポートフォリオは、研究プロジェクトに対するものから研究インフラに対するものまで幅広く資金配分を行っている。

DFGのファンディングプログラムは、個人向け、機関向け、科学的受賞の3つがある。個人向けファンディングプログラムは、研究者別(研究者のキャリア支援)、テーマ別、フォーラム、インフラの4つに区分される。研究者別のプログラムでは、ポストドクや研究キャリアの浅い研究者向け等、若手研究者の育成に係る資金支援が行われている。テーマ別の主要プログラムでは、研究開発プログラムに係る資金として「個人研究助成金」、「重点プログラム」等がある。また、研究者が極めて革新的でリスクの高いプロジェクトを推進するためのものとして、ラインハルト・コゼレック・プロジェクトがある(最長5年間)。

機関向けファンディングプログラムは、テーマ別、フォーラム別、インフラ別の3つに区分され、テーマ別では共同研究センター・プログラム(複数の大学が共同で提案し拠点を形成)があり、フォーラム別ではエクセレンス・ストラテジーのプログラムとして、「エクセレンス・クラスタープログラム」についての資金配分を行っている。インフラ別では、「大型研究機器助成プログラム」のように連邦基本法第91b条に基づき連邦政府予算分の資金(助成額の50%)を配分するプログラムから、DFGが州に代わり審査を実施する「州の主要研究機器助成プログラム」等がある。

本調査では、研究開発プログラム評価の対象として「重点プログラム」を取り上げる。

⁸⁵ BMBF: Evaluation Procedure

<https://www.nks-msc.de/en/Evaluation-Procedure-2290.html?utm_source=chatgpt.com> [2025/3/21 取得]

⁸⁶ DFGは、法的には私法上の協会である。DFGのような資金配分機関に対する連邦政府の機関資金のほとんどは、連邦教育省の予算を通じて配分されている。

表 2-14 DFG ファンディングプログラム一覧（個人向け）

タイプ	区分	プログラム	概要
個人	研究者	ウォルター・ベンヤミン・プログラム	博士研究員研修段階の研究者（ポストドクター）が、希望する場所で独自の研究プロジェクトを独自に実施することを可能にする。〈2年間〉
		研究フェローシップ	キャリアの浅い研究者が、ドイツ国外の希望する場所で、決められたプロジェクトを実施し、新しい研究方法に慣れる機会、あるいは大規模なプロジェクトを終結させる機会として利用（海外手当、旅費、出版費用、子ども・育児手当）。〈2年間〉
		エミー・ネーター・プログラム	優れたキャリアを持つ若手研究者に、6年間、独立した若手研究グループを率い、大学教授のポストを得る機会を与える（研究キャリアの初期段階にあるポストドクおよび臨時契約の若手教授を対象）〈6年間〉
		ハイゼンベルグ・プログラム	すでに永続的な教授職に任命されるための条件をすべて満たしている場合（教授、W2/W3 給与等級）、研究活動を継続しながら、将来の上級研究職を目指す準備をし、学術的な名声を高めるための資金〈最長5年間〉
	テーマ	個人研究助成金	学術研修を修了した個人が、テーマを問わず、テーマと期間を明確に定めた研究プロジェクトをいつでも実施できる助成金。〈原則数年間〉
		ラインハルト・コゼレック・プロジェクト	科学的実績のある優れた研究者が、極めて革新的でリスクの高いプロジェクトを推進するためのもの。大学で教授職又は教授職に就く資格を有する研究者が対象。50万€から125万€までの助成金を5年間、25万€単位で申請。
		パッケージ提案	テーマに関連する個々のプロジェクトは、互いに独立して実施される場合、パッケージ・プロポーザルとして一緒に提出することができる。個々のプロジェクトは別々に審査されますが、テーマが近いため、互いに参照し合う。
		国際共同研究（開始時）	国際的な研究協力の開始を支援するための助成プログラムで、予備的ワークショップ、プロジェクト関連の海外出張、プロジェクト関連の来客訪問の3つの要素で構成される。他方、国際会議や共同研究プロジェクトの実施の支援ではない。
		研究ユニット	特定のテーマに焦点当てた研究ユニットへの助成（個人助成プログラムや重点プログラムの資金調達オプションを超える共同研究プロジェクトを対象）。〈最長8年間：4年毎の2つの助成期間〉
		臨床研究ユニット	疾患または患者を対象とした臨床研究の分野から、特別な研究プロジェクトにおいて緊密かつ中期的な協力をを行う機会を提供。
		臨床試験	学術研修を修了した個人が、一時的なプロジェクトの中で、いつでも患者を対象とした臨床研究を実施。本プログラムは、フィージビリティ・スタディ（第II相）やインターベンショナル・トライアル（第III相）を含むインターベンショナル臨床研究に資金を提供。
重点プログラム	参加研究者間の全国的な協力体制で取り組むことを前提としたプログラム。学際的な提案を重視し、異なる拠点との連携による付加価値の創出、プロジェクトに関与するキャリア初期段階の研究者支援、重点プログラムに合わせた公平性対策を推進することを目的としている。DFG 評議員会（Senate）は、協調的な支援により、特に科学的な成果が期待される場合、優先プログラムを設置することができる（年1回）。〈6年間〉		

出所：DFG ウェブサイトをもとに未来工学研究所作成

表 2-15 DFG ファンディングプログラム一覧（機関向け）

タイプ	区分	プログラム	概要
機関	テーマ	共同研究センター／トランスレギオ	共同研究センターは、大学を拠点とする長期的な研究機関であり、最長 12 年間の期間で設立され、研究者が学際的な研究プログラムのもとで共同作業を行う。このプログラムにより、機関の優先分野の開発と構造的な発展を可能とする。大学以外の研究機関との協力を奨励している。目的は、グループ内の優れた研究の実現、申請大学における重点分野の構造開発、学際的協力、若手研究者の支援、研究における男女平等の達成が掲げられている。最長 12 年間。1 つの助成期間は 4 年間。 ※CRC との違い：本プログラムは複数の大学が共同で提案し実施する。
	フォーラム	研究トレーニンググループ	大学が若手研究者の育成（博士号取得者の育成促進）を目的に設立するもので、最長 9 年間にわたって資金援助を行う。集中的な研究プログラムと体系的なトレーニングの枠組みにより博士研究者の資格を認定する。最初の助成期間は 5 年間（2022 年 5 月以降の設置計画の場合）。2 回目の助成期間は 4 年間。
		DFG 研究センター	国際的に認知された競争力のある研究センターを設立するため、教授職と独立した若手研究グループに対して初期の資金と支援を提供するプログラム。エクセレンス・イニシアティブやエクセレンス・ストラテジーで資金提供されているエクセレンス・クラスターとは異なり、DFG がテーマ別募集に応じて DFG 研究センターが設立される。DFG 研究センターとして設立されたセンターがその後、エクセレンス・クラスターになるケースもある。DFG 研究センターへの助成金は年平均で約 500 万ユーロ、最長 12 年間にわたり受け取ることができる。
		エクセレンス・クラスター	エクセレンス・ストラテジーのプログラムの一つ。2016 年 6 月の連邦政府と州政府が締結した行政協定に基づき、DFG と WR が 2 つの助成プログラムを実施。DFG が資金配分を行うプログラムは、エクセレンス・クラスター（Clusters of Excellence: EXC）で、大学又は大学コンソーシアムによるプロジェクトベースの資金援助を目的としている。提案は、学術主導の競争プロセスで審査・決定される。最初の期間（2019 年 1 月～2025 年 12 月）は、総額 3 億 8500 万ユーロの年間予算で 57 のクラスターに助成。2026 年以降の助成期間は総額 5 億 3900 万ユーロの年間予算で最大 70 のクラスターに助成予定。
	研究インパルス	応用科学大学（UAS）の研究志向型大学の研究コンソーシアムへの資金援助を行う枠組み。1 つの助成年度につき、最大 100 万ユーロ（プロジェクトの間接費用を含む）を助成。助成金は、人件費、プロジェクトの直接費、機器費用等に利用可能。	
インフラ	大型研究機器助成プログラム	連邦基本法第 91b 条に基づくプログラム。研究施設及び大型研究機器に関する連邦予算から DFG は各種大学における大型研究機器の投資費用の 50%を拠出する。助成を受けるには、州又は大学による 50%の共同出資が必要となる。	
	州の主要研究機器助成プログラム	DFG が州に代わり、大学及び大学病院における主要な機器に関する審査を行う。ただし、DFG は本プログラムの機器の調達費用を負担しない。助成プログラムの対象となる機器は、各州及び大学が資金提供し、研究・研修・教育・臨床ケア又はそれらの組合せでの使用を目的としたもの。	
	大型研究機器イニシアティブ	優れた革新的技術を搭載した高額な大型研究機器を申請することができ、特定の科学技術的問題解決の促進を目的としている。大型研究機器の導入構想の設立に関するアイデアや提案を研究コミュニティから構想案の形で提出し、DFG が選定し決定する。	
	コア施設助成プログラム	コア施設の運営と管理の専門化を促進し、これら施設の利用のための安定した構造の開発を支援する（利用と管理）。このプログラムの目的は、他の機関の外部研究者を含め、利用者にとり、より優れた施設と機器ベースの研究インフラの良好なアクセスを提供することにある。申請者は大学であり、年間 15 万ユーロを上限に申請することができる。最長 5 年間。 ※大型研究機器取得のための資金として利用できない。	

タイプ	区分	プログラム	概要
機関	インフラ	オープンアクセス出版助成金	オープンアクセスによる研究成果の出版に対して、一定の助成金を支給することを目的としている。
		全国研究データ基盤(NFDI)	科学的及び研究データを体系的に管理し、長期的なデータ保存、バックアップ及びアクセスを提供し、国内及び国際的なデータをネットワーク化することを目的としている。

出所:DFG ウェブサイトをもとに未来工学研究所作成

2) DFG の評価活動及び評価のプロセス

a. DFG の評価活動

DFG の助成プログラムや政策等の評価を行っている。DFG の助成プログラムの評価では、プロセスの成功を評価し、DFG の助成政策のための経験的基礎となっている。評価結果は、インフォブリーフ、評価研究等で公表し、助成活動の透明性を確保している。DFG の評価研究として、共同研究プログラムの研究者に関する調査（共同研究センター・プログラム、研究トレーニンググループ・プログラム、エクセレンス・クラスターの研究者を対象に人員構成に関する分析）⁸⁷、COVID-19 パンデミック関連の活動調査、DFG が助成するコンソーシアムにおける博士課程の期間と博士号に関する調査、若手研究者のための資金提供側スタッフ職に関する調査、研究助成とキャリアパス等の報告がある。

エクセレンス・ストラテジー関連では、2018年に「連邦及び州政府のエクセレンス・ストラテジー」とする研究が行われている。

b. 研究プログラム／プロジェクトの評価

DFG の採択評価は、審査、評価、決定の3段階の手順に沿って行われている。審査は専門家による初期審査（レビュー）であり、評価は審査委員会（レビューボード）による評価、決定はDFG 合同委員会における採択の決定である。個人助成（Individual Grants）は、ピアレビューに基づく書面審査にて評価が行われ、審査委員会にて各分野から提出されたすべての提案を比較評価し、審査結果（提案の優先順位、提案に対する助成金の支給／不支給等）を取りまとめる。採択の最終決定は、DFG 合同委員会にて学際的な比較に基づき、承認又は却下の最終決定を行われる。また、共同研究プログラム（研究ユニット〈個人向け〉、重点プログラム〈個人向け〉、共同研究センター〈機関向け〉、研究トレーニンググループ〈機関向け〉の4プログラム）の審査では、審査委員会による審査（口頭審査）を実施し、審査委員会には少なくとも関連するDFGの委員1名が含まれるよう、研究評価基準に基づいたバランスの取れた評価を実施している。平均助成率は30%。審査委員会委員は4年毎に投票が実施され、研究分野区分に対して2名の代表者が選出される。

⁸⁷ 2004年に「共同研究プログラムの研究者に関する調査(Statistical Reports on Researchers in the Coordinated Programmes)」が行われているが、DFG 評価研究では「共同研究プログラムの年次調査」を実施している。共同研究センター、研究トレーニンググループ、研究インパルスで何名の博士号取得者がいるか、エクセレンス・クラスターで何名の教授職が創設されたか等の調査を毎年実施している。

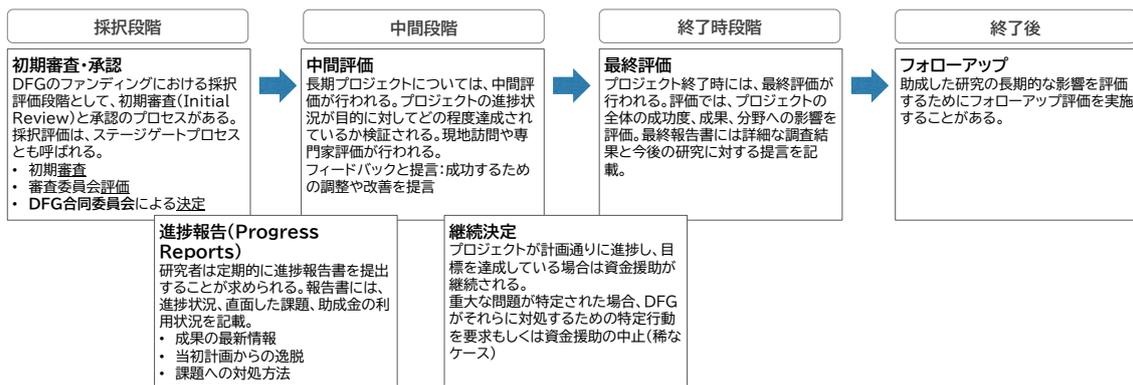


図 2-12 DFG のプロジェクトの採択から終了時までの評価プロセス

出所: DFG ウェブサイト⁸⁸より未来工学研究所作成

3) プログラムレベルの評価 (重点プログラム)

重点プログラム (Priority Programmes) は、重要な新しいテーマに対する複数拠点からの協調的な資金提供を通じて、科学・人文科学の進歩に寄与することを目的とした研究資金である。助成の範囲は、個人プロジェクトに助成金を支給 (最大 30 件) し、助成の期間は最長 6 年間で 2 つの助成期間 (3 年毎) に分かれる資金配分が行われる。評価プロセスは、下記のとおりである。

- 審査プロセス: 提案内容 (重点プログラムの設立提案) は、各専門分野の専門家と審査委員会により審査される。審査委員会は、科学的価値、実現可能性、潜在的な影響力を評価し、審査委員会は、勧告案を DFG の評議員会に提出することになっている。
- 中間評価プロセス: 進捗報告書 (プログラムの成果、課題、今後の計画を記載) を提出することになっている。その後、専門家による外部評価を実施し、進捗報告書を審査し、追加情報や見識を得るため、現地訪問調査の実施。審査に基づき、研究者へのフィードバックを実施する。
- 終了時評価: 助成期間終了時に重点プログラムの最終報告書を提出する (成果、影響、全体的な達成状況をとりとめ)。評価は、独立した専門家や審査委員会により包括的なレビューを実施し、審査では、プログラムの科学的寄与、革新性、より広範な影響の評価が実施される。

(3) プログラム実施機関による評価 (プロジェクトトレーガー)

プロジェクトトレーガー (プロジェクト運営機関) は、大学や公的研究機関が実施する研究開発プログラムの運営・管理・評価を担う。プロジェクトトレーガーは、主に連邦政府の

⁸⁸ DFG ウェブサイト < <https://www.dfg.de/en/research-funding/proposal-funding-process/decision-making-process> > [2025/3/21 取得]

プログラム、プロジェクトを運営し、資金配分のための準備作業、プログラムのモニタリング、プロジェクトの管理を行う。また、プログラムの概念設計や設計段階でのステークホルダーとのミーティング等も行っている。

DLR プロジェクトトレーガー (DLR Projektträger) は、ドイツ最大のプロジェクト管理組織として、そのテーマの広範性を活かし、学際的かつ超学際的な方法で調査を実施している⁸⁹。同社では、大規模な資金調達戦略やプログラム、特定的手段や対策に重点を置き分析し、資金調達プログラムの当初より関わり、現在の課題に最適に対応するための戦略、行動コンセプトをクライアントに提供している。

DLR プロジェクトトレーガーの主なサービスは、①トレンド評価のための有意義なパラメータとデータの特定、収集、分析、②教育、研究、イノベーションにおける国際的な動向の観察と評価、③顧客の政治戦略とターゲットグループおよびニーズとの整合、④国内外の資金調達手段の機会と潜在能力の分析、⑤公的および民間資金調達プログラムのための付随研究の実施、⑥施策、プログラム、機関の評価等である。このうち、分析・研究・評価については、資金調達プログラムの開発に初期段階から関与し、その実施を管理している。顧客から、「新しい研究資金調達イニシアティブ、プログラム等で成功するために必要なこと」、「考慮すべき政治的・技術的・法的要件」、「政治・経済・科学・社会の各レベルの発展が資金調達にどのような影響を与えるか」、「緊急対応が必要な分野はどの分野か」等の質問に答えられる支援を行う。

DLR プロジェクトトレーガーには、分析・調査・評価コンピテンズセンター (Kompetenzzentrum Analysen, Studien und Evaluationen : KS-ASE) が設置されている⁹⁰。KS-ASE は、資金調達施策は意図したとおりに機能し目的を達成しているか、目標達成、有効性、経済効率の観点からどのように評価できるか等、社会からの疑問に対して、戦略や組織の助言、資金調達プログラムの評価、各種調査報告書を通じて支援している。2023年には、新たな調整グループとして「メソッドハブ」と「モニタリング・評価」を設けた。「メソッドハブ」調整グループは、DLR プロジェクトトレーガー (管理機関) 全体と、省庁や財団・協会等の外部顧客の両方に対して包括的なサービスプロバイダーとして、トレンド分析、環境分析、ネットワーク分析等の技術分析、特許分析や文献分析等のカスタマイズしたサービスを提供している。また、戦略的予測、Tableau (BI ツール) による視覚化、機械学習や AI を取り入れた新しいアプローチも活用している。「モニタリング・評価」調整グループは、DLR プロジェクトトレーガーの部門横断的な窓口となり、モニタリング・評価・調査のあらゆる側面を支援している (2023 年の新規プロジェクト「メクレンブルク＝

⁸⁹ DLR プロジェクトトレーガーの資金調達プロジェクトの管理数は 2023 年で 14500 件以上とされる。EU の Horizon Europe の資金調達も行っており、14 のナショナルコンタクトポイントを設置し、ドイツからの申請者を高めることを実施している。

⁹⁰ DLR プロジェクトトレーガー「年次報告書 2023 (Geschäftsbericht 2023)」

(<https://projektraeger.dlr.de/sites/default/files/2024-04/documents/DLR-PT-Geschaeftsbericht-2023.pdf>) [2025/3/21 取得]

フォアポンメルン州における都市気候適応プロジェクトの評価」では、DLR プロジェクトトレーガーの環境・持続可能性部門と KS-ASE が共同で実施している)。⁹¹

参考:DLR プロジェクトトレーガーが扱う連邦政府の研究開発プログラムの資金提供ガイドラインの例(Guideline for the Promotion of Interdisciplinary Pilot Projects on the Topic of “Neurobiologically Inspired Artificial Intelligence”)

この資金提供ガイドラインは、パイロットプロジェクトにおける神経科学と AI の相互インスピレーションと統合の可能性と機会を特定することを目的としたものである。なお、資金調達措置は、人工知能に関する BMBF 行動計画の行動分野「健康分野における AI:すべての人に社会的利益をもたらす」に寄与するもので、ガイドラインは健康研究フレームワークプログラムと併せて適用される。連邦政府の研究開発プログラムの資金調達の窓口を担っている。
(※本プログラムは公募中で 2025 年 3 月 6 日提出締切)

2.3.3 評価疲れについての各国・地域の現況

ドイツにおける研究者の評価の負荷について、直接的なデータはないものの、研究開発評価の管理上の負荷が全体的な業務量に影響を与えていると考えられている。このため、評価プロセスの合理化、作業に係る時間的負担を認めることで、研究者の潜在的な負荷を軽減できる可能性がある。また、ドイツの研究者の中には、研究開発評価プロセスが、論文文化(論文に係る各種指標や被引用数等)に依存しているため、研究者に過剰なプレッシャーを与えているとされる。将来の研究者の評価疲れに関連する問題として下記が挙げられている。

- COVID-19 以降の論文の科学的品質基準の重要性: COVID-19 以降、査読を経ないままプレプリントサーバー上で公開される論文が増加している。論文公開に係る競争を排除し、研究データを即時に共有する動きであるが、プレプリントには従来のジャーナル出版と同等の品質基準が適用されるため、関係者全員の高度な責任が求められる。また、プレプリント論文の受け手側も、講評されたデータと文書化された品質保証を適切に処理する必要があり、研究者の新たな負荷となっている⁹²。
- 研究機関の評価文化の変化が必要: 学術論文の出版と評価に関して、研究者が質を第一に考えることが容易になるよう、この変化に対する信頼を促進したいと考えられている。(⇒論文の出版数や引用数、ジャーナルの権威度によって研究者を評価するのではなく、ジャーナル評価指標への依存度を低減しながら、研究の質を重視すべき)⁹³

⁹¹ KS-ASE センターのメンバーは、ドイツ評価学会 (DeGEval) の年次イベント等の専門化会合に定期的に参加し、セッションへの参加を通じて研究・科学政策の主要関係者との協力関係を強化している。

⁹² DFG betont Bedeutung von wissenschaftlichen Qualitätsstandards bei Publikationen, 2020 年 4 月 23 日<<https://www.dfg.de/resource/blob/174202/e531140697fc59fba0fec9f6ce12abdd/200423-statement-dfg-preprints-data.pdf>> [2025/3/21 取得]

⁹³ Science| Business-Bringing together industry, research and policy, “Research assessment: German funder raises concerns as EU reform talks come to a close”, 2022 年 5 月 19 日。

EFI が実施した委託研究プロジェクトの評価 (BMBF と BMWK が委託した 81 件の政策研究) では、分析対象プロジェクトの大多数 (73%) は政策介入の因果効果があることが解釈された。他方、因果分析を行うには、利用可能なデータ量とデータの質の不十分さを抱えている。政策措置の開始時にデータが原則入手可能な場合であっても、必要な時に利用できない場合がある (Büchele [3])。近年のミッション志向型研究開発の評価においては、新たな政策設計が必要と考えられている。

- 研究者は、プレプリントの公表 (もしくはプレプリント情報の活用等) が増えてきており、当該情報の発信や活用に係る丁寧な対応が必要となる。また、ミッション志向型研究開発等が進む中で、研究プロジェクトの評価では、政策介入の因果関係が評価できるよう、各種情報の提出が求められる可能性があると考えられる。
- なお、政策評価の実施の根拠として、ドイツ連邦予算法にて連邦予算に影響を与えるすべての施策は適切な経済的実現可能性調査の対象となることが義務付けられていることが挙げられる。

【参考文献】

[1] 山岸喜久治, ドイツ検討と連邦国家原理—立法と行政の協働的連邦主義—, 人文社会科学論叢, 宮城学院女子大学, 2020年3月.

[2] JST-CRDS, 科学技術・イノベーション動向報告書: ドイツ編, 2023年3月.

[3] Büchele, Stefan et al., Commissioned Project Evaluations of Research and Innovation Policy in Germany: A Review, Studie zum deutschen Innovationssystem, No.11-2024, EFI, 2024.

[4] Bühner, S., Rainer Walz, R., Sarah Seus, S. (Fraunhofer ISI), Michael Astor, M., Stehnen, T., Malik, F. (Prognos), Evaluation der BMBF-Rahmenprogramme Forschung für die Nachhaltigkeit FONA 1 (2005-2009) & Forschung für Nachhaltige Entwicklungen FONA 2 (2010-2014), Januar 2020.

< [https://www.prognos.com/sites/default/files/2021-](https://www.prognos.com/sites/default/files/2021-01/bmbf_fona_evaluation_abschlussbericht_2020.pdf)

01/bmbf_fona_evaluation_abschlussbericht_2020.pdf > [2025/3/21 取得]

<<https://sciencebusiness.net/news/european-research-area/research-assessment-german-funder-raises-concerns-eu-reform-talks-come>> [2025/3/21 取得]

2.4 イタリア共和国（伊国）

2.4.1 国・地域における研究開発評価制度の体系とその内容

(1) 評価制度の全体像と研究開発評価の位置付け

1) イタリアの研究評価システム

イタリアの研究評価システムの歴史はそれほど古いものではない。イタリアでは長年、公的セクターの研究開発に関する評価システムがないことが、同国の研究開発システムの弱点だと言われてきた、とされる（JST CRDS [1]）。現在、イタリアでは、他の EU 諸国と同様に、公的資金を受けている大学や研究機関が実施する科学研究は、評価の対象となっている。

イタリアでは 20 世紀末から大学における評価手続きが導入されている。研究評価の試みは、1980 年代末から始まっていたが、当初は大学の経営的側面にのみ焦点が当てられていた⁹⁴。その後、1990 年代末になって、研究活動の質的側面に焦点を当てた評価が導入された。具体的には、1998 年に研究評価機関として研究評価委員会（Comitato di indirizzo per la valutazione della ricerca : CIVR）が設立された。CIVR はイタリアにおける最初の研究の質評価の実施を担当し⁹⁵、その後研究評価の役割は全国大学システム研究評価委員会（Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca : ANVUR）に取って代わられた⁹⁶。ANVUR は設立以来、ベルルスコーニ第 4 次内閣時の教育改革であるいわゆる「ジェルミニ改革（Gelmini Reform）⁹⁷」によって導入された研究の質の評価システムである Valutazione della Qualità della Ricerca（VQR）を実施している。こうした体制の整備に伴い、イタリアの評価システムは多様なレベルで行われており、各研究機関内会議による研究機関の自己評価、外部メンバーが構成する評価委員会による評価、ANVUR によるイタリアの研究システム全体を見据えた外部評価、また事前評価・事後評価が行われる（JST CRDS [1]）。近年はイタリアの科学技術イノベーション計画においても、研究評価に関する記述が見られる⁹⁸ほか、イタリア大学・研究省（Ministero dell'Università e della Ricerca : MUR）が所管する研究開発プログラムにおいても研究評価のガイドライン

⁹⁴ Furiosi, E., Research Quality Assessment in Italy and classification of the A-class scientific Journals. Potential distorting effects, 1 April 2021. <<https://ceridap.eu/research-quality-assessment-in-italy-and-classification-of-the-a-class-scientific-journals-potential-distorting-effects/?lng=en>> [2025/3/21 取得]

⁹⁵ 2003 年 12 月に国による最初の研究評価が行われた。（CRDS [1], p50）

⁹⁶ 2006 年には従来の評価機関である大学システム評価委員会(CNVSU)（1999 年設立）と CNVSU が統合し、ANVUR が新しく創設された。

⁹⁷ 2008 年から 2011 年にかけての教育・大学・研究大臣マリアステラ・ジェルミニの任期中に一連の改革法が発効したためこう呼ばれる。

⁹⁸ 全国研究計画（Programma Nazionale della Ricerca : PNR）2021-2027 の付属文書である「オープンサイエンスのための国家計画」

が定められている⁹⁹。イタリアにおける研究評価は、政権交代等のたびに新たな制度や機関が発足するなどの改革が行われ、絶え間なく進化しており、確定的な構造となるには至っていない。

イタリアにおける研究評価は、イタリア共和国憲法第 97 条に謳われている「善政の原則 (buona amministrazione)」を、科学的・学術的領域において保証するものと位置付けられている。研究評価では、研究活動の厳密な科学的側面と社会経済的側面の両方が考慮されるのが原則である。すなわち、公的活動におけるアカウンタビリティ (説明責任) の原則に従い、すべてのプロジェクトリーダーは、投下された公的資源に対して、期待される中長期的効果だけでなく、達成された成果や経済的・社会的利益 (当該分野における知識の増加、生産的イノベーション、特許、製品プロトタイプ等) を透明化しなければならない。他の類似の取り組みとのベンチマーキングを含む研究の評価は、政策決定者にとって多様な分野における公共投資の優先順位を導く上で決定的なものである¹⁰⁰。

以上に見たように、イタリアにおける VQR (および他の研究評価プロセス) の長期的な目標は、政府が公的資金で行う研究への投資との関連において、時間をかけてでも、生産される研究の質を向上させることである¹⁰¹。

2) 研究評価システムを担う主要機関

以下にイタリアの研究評価システムを担う主要機関の概要等を記す。

a. ANVUR¹⁰²

ANVUR は、大学および研究機関に対する国による質評価システムを監督している。主に公的資金を受ける大学および研究機関が実施する活動の質評価を担当している。独立評価機関の活動の運営や、研究・イノベーション活動に対する公的資金提供プログラムおよび奨励プログラムの有効性と効率性の評価も任務とする。具体的には、ANVUR は以下の業務を行う。

- ① 教育機関の経営、教育、研究、技術移転活動の手順、結果、成果を評価する
- ② 教育機関および教育課程 (博士課程、修士課程、卒後医学教育課程を含む) の評価基準と方法論を定め、省による定期的な認証評価を行う
- ③ 大学の独立評価機関による評価活動の運営
- ④ プログラムに対する学生の満足度を収集・評価するための手順の策定 (大学評価ユニットとの協力)

⁹⁹ 国益に資する研究プロジェクト (PRIN) 専門家のためのガイドライン

¹⁰⁰ MUR ウェブサイト: Valutazione della Ricerca <<https://www.mur.gov.it/it/aree-tematiche/ricerca/valutazione-della-ricerca>> [2025/3/21 取得]

¹⁰¹ 前掲 Furiosi (2021)

¹⁰² ANVUR ウェブサイト: Mission <<https://www.anvur.it/en/agency/mission/>> [2025/3/21 取得]

- ⑤ 大学の設置、合併、統合、閉鎖および教育課程の活性化、統合、閉鎖を目的とした量的・質的要件 の策定と省への提案
- ⑥ 大臣の要請に応じて、公的資金配分のベンチマークを提供する。 これには、特定サービスの最低実績水準および標準単価の定義が含まれる
- ⑦ MUR と各機関との間で結ばれたプログラム協定の結果、および期待される結果とあらかじめ定義されたベンチマークに基づく、評価システムの質全体の向上への貢献の評価
- ⑧ 教育・研究・イノベーション活動に対する公的助成プログラムおよび奨励プログラムの有効性と効率性を評価する
- ⑨ 大臣の要請に応じて、さらなる評価実施、標準的パラメータの定義、技術的規制の提供を行う

ANVURが実施主体となるVQRは、MURが行う国家的な研究評価のための手続である。VQR ではイタリアの大学・研究機関が実施する科学研究の成果を評価する。MUR はANVURに、全国レベルでのVQR評価の実施を任務として実施させている。

b. CNVR

全国研究評価委員会 (Comitato nazionale per la valutazione della ricerca : CNVR) は、2021年7月に教育・大学・研究省 (MIUR¹⁰³) の15名のメンバーからなる諮問機関として創設された最新の研究評価組織である。欧州研究会議 (ERC) の評価と助成金管理手続きの執行との関連で、イタリアの研究評価を欧州の優れた取り組みの水準に合わせることを目的としている。CNVRは、国益に資する研究プロジェクト (PRIN) や、2021年に開始された個人助成のためのイタリア科学基金 (FIS) など、いくつかの助成プログラムの評価を担当している。しかし、担当する最新のPRINとFISの公募の評価には遅れが生じている。これは評価を行う外部委員の確保に失敗したためとされている¹⁰⁴。

3) イタリアの科学技術イノベーション計画における研究評価

MURによって策定される全国研究計画 (Programma Nazionale della Ricerca : PNR) は、科学技術戦略や研究システムにおける重点分野・事項を示すものである。

「全国研究計画 2021-2027(PNR 2021-2027)」は、MURが監督し、主な利害関係者と市民に開かれた公開協議を経て策定された。計画はイタリアのニーズとイタリアが置かれた状況の特殊性に焦点を当て、イタリアも重要な役割を果たしている欧州研究システムと一致する計画となるよう企図されている。

¹⁰³ 当時の名称、現MURに当たる。イタリアのいわゆる教育省の名称は政権交代のたびに変更され安定しない。

¹⁰⁴ Nature ウェブサイト : Why the evaluation of Italy's research grants was delayed
<<https://www.nature.com/articles/d43978-023-00106-8>> [2025/3/21 取得]

PNR 2021-2027 は、システムの優先事項、主要な研究・イノベーション分野とそれぞれの介入分野、ミッション、国家計画のパートに分かれている。

システムの優先事項について見ると、PNR 2021-2027 は、以下の9の事項を中心に展開されている。

- ① 研究システムの広範かつ包括的な成長を支援する
- ② 基礎研究および学際的研究を強化する
- ③ イノベーションにおける個人の中心性を確保する
- ④ 研究と生産システムの間での知識と技能の循環を強化する
- ⑤ 新しい世代の研究者や専門家を育成し、知識の移転に協力する
- ⑥ 知識の移転において、新しい世代の研究者や専門家の育成を支援する
- ⑦ 高等教育・研究の国際的側面を促進する
- ⑧ 国内、欧州、国際的な研究のコーディネーションを確保する
- ⑨ 新たな研究の地平を目指す

PNR 2021-2027 の主要研究・イノベーション分野は、健康、人文科学、創造性、社会変革、包括的社会、社会システム安全、デジタル、産業、航空宇宙、気候、エネルギー、持続可能なモビリティ、食品、バイオエコノミー、天然資源、農業、環境である。

研究評価については以下の点が重要である。すなわち、PNR 2021-2027 には、「研究インフラ全国計画 (National Research Infrastructure Plan : PNIR) 」と「オープンサイエンスのための全国計画 (Italian national plan for open science) 」が附属文書として含まれている。このうち、オープンサイエンスのための全国計画は、研究データや研究成果への可能な限り幅広いアクセスを保証し、科学界や社会全体に重要な利益をもたらす、オープンで透明性のある包括的なシステムへの移行を支援するための文書であり、研究評価に関する詳細な行動規定を含んでいる。ここでは、研究評価のあり方についてもオープンサイエンス促進の観点から規定されている。

4) イタリアの研究開発プログラムにおける研究評価

MUR は、国内の14の研究機関、大学、競争的研究助成金への資金提供を直接監督している。これは、イタリアに研究のための国家資金提供機関がないためである¹⁰⁵。

MUR の所管するファンドのうち、PRIN は、3年間のファンディングの公募をアナウンスしている。ファンディングの対象機関は、全てのイタリアの大学や大学の機関、国営もしくはその他、特別法による中等学校や MUR の管轄下の公的研究機関も含まれている。MUR は PRIN において共同ファンディングの配分、グループ研究、評価に関する新しいメカニズムを導入している (JST CRDS [2]) 。

¹⁰⁵ UK Science & Innovation Network, UK Science & Innovation Network Country Summary Italy, 2024<<https://assets.publishing.service.gov.uk/media/65f41e379d99de001d03df62/Italy-Science-innovation-country-summary-2024.pdf>> [2025/3/21 取得]

5) イタリアにおける行政評価と研究・イノベーション

イタリアでは、全国研究計画およびその他のプログラムを承認する権限を有する経済計画と持続可能な開発のための省庁間委員会（Comitato interministeriale per la programmazione economica e lo sviluppo sostenibile：CIPESS）¹⁰⁶の行政評価パネルが、科学技術イノベーション分野を含む政府の公金支出を伴うプログラム等についてその有効性や効率性を評価している。具体的には CIPESS の公共投資評価・検証パネル（Nucleo di valutazione e verifica degli investimenti pubblici：NUVV）がその任にあっている。

NUVV は、CIPESS に提出される公共投資プログラムおよびプロジェクトに関する提案の評価活動を行い、特に以下の部門に関連し、各部門を所管する機構に対する適切な技術的支援を保証する。すなわち、研究・イノベーション、インフラ・エネルギー・交通、持続可能な開発・気候変動・水管理、環境保護、地方開発・ビジネス促進、保健・社会政策、公共投資計画・監視、財政・公会計の各部門である。

NUVV は、1999 年 9 月 10 日の閣僚理事会議長指令で設立された評価・検証ユニット・ネットワークと統合することにより、公共投資の評価・検証・監視活動を実施し、方法論と実践の共有を促進し、公共投資監視システム（MIP）の開発を支援するために活動している。NUVV は、コーディネーターを含む 15 名以下の専門性の高い専門家で構成されている。また、メンバーは行政の外部から選ぶこともできる¹⁰⁷。このように、イタリアにおいては公金支出を伴う研究・イノベーション活動は行政評価として NUVV の評価・検証を受ける可能性がある。

(2) 研究開発評価に関する規定等の整備状況

1) PNR 2021-2027 におけるオープンサイエンスのための全国計画

上記のように、PNR 2021-2027 には、研究インフラ全国計画（PNIR）とオープンサイエンスのための全国計画が含まれている。このうち、オープンサイエンスのための全国計画では研究評価のあり方についてもオープンサイエンスの促進の観点から規定されている。本規定の性格はイタリアにおける今後の研究評価のあり方の方向性を示す政策文書である。

¹⁰⁶ 2021 年 1 月 1 日付で、いわゆる「気候法令」を改正する 2019 年 12 月 12 日付法律第 141 号によって規定された改革が施行され、旧 CIPE が CIPESS に改称された。これは、欧州の義務である「Do No Significant Harm」の原則に照らし、公共投資事業の持続可能な開発評価を導入することを目的とした重要な変更である。CIPESS ウェブサイト：CIPESS

< <https://www.programmazioneeconomica.gov.it/it/cipess/presentazione/> > [2025/3/1 取得]

¹⁰⁷ NUVV ウェブサイト：Presentazione <<https://www.programmazioneeconomica.gov.it/it/nucleo-di-valutazione-e-verifica-degli-investimenti-pubblici/presentazione/>> [2025/3/21 取得]

2) PRIN における専門家のためのガイドライン

上記のように、MUR の所管するファンドのうち、PRIN は、専門家のためのガイドラインを規定している。本規定の性格は研究評価のためのガイドラインである。

3) VQR

上記のように、VQR は、MUR が行う国家的評価手続である。イタリアの大学・研究機関が実施する科学研究の成果を評価する。MUR は ANVUR に、全国レベルでの VQR 評価の実施を任務として実施させている。

VQR は以下の目的で整備されたシステムである。すなわち、

- 他の研究機関とデータや結果を比較し、強みと弱みの認識を高める
- 社会全体に利益をもたらすために、公的資源の配分を含む今後の政策を導き、政策に関する行動の有効性と効率性を高める

ことを目的とする。

VQR 自体は MUR の定めたガイドラインに基づいて運営されるが¹⁰⁸、VQR に定められた評価規定は財政法の規律と結びつき、経済的支援（政府による補助金交付等）の有無を決定する法的効果を伴う¹⁰⁹。VQR の評価結果は、①MUR により、通常資金（FFO）および研究・知識活用活動に関連するその他の資源の分配額の決定に際して使用されるほか、②評価を受けた大学により、内部資金の分配および独自の戦略目標策定のために使用されるとともに、③研究機関を含む各部局により、研究・知識生産活動に関する自己評価のために用いられる¹¹⁰。

(3) 研究開発評価に関する規定等の詳細

1) PNR での研究評価に関するアクションプラン

PNR 2021-2027 の「オープンサイエンスのための全国計画」では、研究評価のあり方について以下のような具体的な目標とアクションプランが示されている。

¹⁰⁸ 2019 年 11 月 29 日付省令第 1110 号により MIUR が発行し、2020 年 8 月 11 日付省令第 444 号により MUR が補足した VQR2015-2019 のガイドライン

¹⁰⁹ 例えば研究機関等において「卓越した部門」が特定され、そこに追加資金が配分されるのは、常に VQR の結果に基づいている。2016 年の財政法では、MIUR は ANVUR に対し、最新の研究の質評価（VQR）から得られた結果に基づいて、特定の「部局業績の標準化指標」（ISPD）の算出を定義するよう要請した、この指標は、VQR の全国的な分布における各専攻の位置、各科学分野における位置、相対的な ISPD の各専攻への帰属を考慮したものである。この指標は、各部門の VQR 評価と厳密に関連している。前掲 Furiosi (2021)

¹¹⁰ トリノ大学ウェブサイト：Research Quality Assessment (VQR)

< <https://en.unito.it/research/research-quality-assessment> > [2025/3/21 取得]

○目標

- 評価プロセスとそれらの基となるデータをより透明化する
- 評価基準をより信頼性が高く堅牢なものにする
- すべての検証済みの研究結果を含める
- 論文出版に関する計量書誌学指標の影響を低減する
- 機関間および研究者間の協力を促す評価基準を実施する
- 評価活動の対象となる研究結果（例：VQR）をオープンアクセスで公開する
- すべての国家評価活動（例：VQR）にオープンピアレビューの形式を適用する
- 科学文献からの引用データの一般公開を保証するシステムを開発する
- 既存のオープンアーカイブの相互接続を基盤として、イタリアに全国分散型インフラを提供し、全国研究登録簿（ANPREPS）の基盤を提供する

○アクションプラン

これらの目標達成のため同計画では以下のアクションプランが提示されている。

- 評価対象となるすべての出版物をオープンアクセスアーカイブに保管することを義務付けることにより、国家評価プロセスにおける研究成果の提出規則を変更すること
- 評価基準を拡大するために、インパクトファクターや H-Index といった計量書誌学指標の比重を減らし、ERA におけるこれらの基準の進化に準拠して、オープンサイエンスや第三ミッション活動（Terza Missione¹¹¹）への貢献を適切に評価すること
- 科学界の関与のもと、科学的研究成果の影響力を評価する新たな基準を策定する
- 全国研究登録簿を確立する
- 評価方法を整合させることで DORA 宣言の原則を順守すること
- 評価対象となる成果物の中に FAIR データおよびそのサービスの作成と保護を盛り込むことで研究評価基準を適応させること
- 研究機関や学術機関のレベルでも、第三ミッション活動の評価基準としてオープンサイエンスの優れた実践を認め、報いること

特長

研究結果についてのオープンアクセスの推進や地域貢献活動の評価など全体としてオープンサイエンスの促進に向けた取り組みを評価する基準が設定されている。評価の方法論としてインパクトファクターや H-Index といった計量書誌学指標の比重を減らしていくことが提案されている点が特徴的である。EU における研究評価の標準への接続の点で、DORA 宣言の原則の順守が掲げられている点が注目に値する。

¹¹¹ 広義には大学・研究機関の社会貢献活動全般を指す。

2) PRIN における研究評価（“PRIN 2020 GUIDELINES FOR EXPERTS”）

PRIN においては主として採択時の事前評価に重きが置かれており、そのための専門的な評価者に対する評価手順がガイドラインとして示されている。

ガイドラインによれば評価手順は以下の通りである。

プロジェクトの科学的評価は、各評価委員会が、MUR のイタリア科学評価専門家査読者登録（REPRISE）または欧州委員会の専門家データベースから、CINECA¹¹²が管理するウェブベースのアプリケーションを通じて、科学的能力の基準に従って選出した 3 名の匿名の外部査読者に委ねられる。

各外部専門家は、独立した立場で、他の専門家に対して匿名性を保ちながら、CINECA が管理する専用の IT プラットフォームを使用して、各プロジェクトの評価フォームを作成し、プロジェクトの長所と短所を示し、科学的評価に関する 3 つの評価基準ごとに、以下の評点に従って、解説コメントを添えて採点を行う。

○評価基準

1. 研究プロジェクトの質

- a) プロジェクトの目的は、どの程度明確かつ適切に構成されているか？
- b) プロジェクトのオリジナリティや革新的な点は何か？
- c) プロジェクトの目標を達成するために、提案された調査方法と作業体制はどの程度適切か？
- d) その目標は、特定の科学における最先端技術をどの程度超えているのか？

2. 研究グループの構成、プロジェクトの実現可能性と妥当性

- a) PI とユニット長（もしあれば）は、どの程度科学的卓越性を実証したか？
- b) 研究チームの能力、構成、補完性の観点から、プロジェクトはどの程度実現可能で適切か？
- c) そのプロジェクトは、提案された目的、完了までの予想される期間、必要なリソース（ツール、研究チームの規模、管理）とどの程度一致しているか？
- d) 研究チームのメンバーの時間的コミットメントや支出計画は、プロジェクトの目的や活動とどの程度整合性があり、一致しているか？

3. プロジェクトの社会的インパクト

- a) 技術革新、産業への応用、経済成長、社会問題の解決、文化遺産や環境遺産の保護に与える影響など、学際的アプローチを含む研究分野の課題にどの程度取り組んでいるか？

¹¹² イタリアの 69 の大学、27 の国立研究センター、MUR 等からなる非営利コンソーシアム。大学や研究機関向けの ICT サービスを提供している。

b) プロジェクトは、そのアウトカムを普及させるための活動をどの程度提案しているか？EUの研究・イノベーション枠組み計画の目標に照らして科学界と社会にどのような影響を与えることができるのか？

○評点

- 0-5 不十分：致命的、または弱点が多すぎる
- 6-7 十分：プロジェクトのいくつかの部分で弱点がある
- 8 良い：いくつかの中程度の弱点
- 9 非常に良い：非常に説得力がある。
- 10 傑出：十分な説得力があり、弱点がない。

評点に伴う評価者による最小 500 文字から最大 700 文字の採点コメントでは、上記 3 つの基準それぞれについて採点理由を説明しなければならない。得点を正当化するために、評価者は上記の評価フォームに含まれる質問に答えるか、一般的な回答を記す必要がある。

特長

プロジェクトの社会的インパクトが大きな比重を占めている。そこでは経済成長のみならず、社会問題の解決、文化遺産や環境遺産の保護へのインパクトを学際的な研究手法によって創出していくことが企図されている。また、EUの研究イノベーション目標との整合性との文脈において社会的インパクトが位置付けられており、欧州科学界との接続が強く意識されている。

3) VQR

a. 概要

VQR は、MUR が行う国家的評価手続である。イタリアの大学・研究機関が実施する科学研究の成果を評価する。MUR は ANVUR に、全国レベルでの VQR 評価の実施を任務として実施させている。

- 第 1 回 VQR は 2011 年に開始され、2004 年から 2010 年（7 年間）を対象。
- 第 2 回 VQR は 2015 年に開始され、2011 年から 2014 年（4 年間）を対象。
- 第 3 回 VQR は 2020 年に開始され、2015 から 2019 年（5 年間）を対象。
- 第 4 回 VQR は、2020 から 2024 年（5 年間）の期間に関して進行中。

評価手続に関しては、ANVUR の行う VQR は査読を中心とするピアレビューの方法論を採用し、科学分野に応じた適切な計量書誌学情報を併用する。評価対象の研究結果について、その独創性、方法論の厳密性、影響力に関する判断を表明する任務を専門評価者グループ (GEV) に委ねる。

b. VQR 2011-2014¹¹³

近時の評価である VQR 2011-2014（第 2 回 VQR）で VQR の評価の構造は概ね定まった。すなわち、各大学および公的研究機関は、研究スタッフ（教授および研究者）ごとに 2 通の研究成果を提示するよう求められた。提出された研究成果は、16 の評価専門家グループ（GEV）によって評価された。評価は査読方式で行われ、適切な場合には書誌分析の方法が併用された。

評価では、科学的研究の「質」の指標として、研究における独創性が確認されるとともに、方法論の厳密性および潜在的な影響力が挙げられた。それらの基準に基づいて評価された各研究成果物に、excellent、high、fair、acceptable、limited の質的評価レベルが割り当てられた。

各研究成果から得られた評価は集計・修正され、関係機関の適切な「ランキング」を作成することを目的とした様々な指標に変換され、機関の全体的な質のプロファイルが算出され、それに基づいて FFO のボーナス部分が分配された。

c. VQR2015-2019¹¹⁴

最新の VQR である VQR2015-2019（第 3 回 VQR）では、2020 年 9 月から 2022 年 7 月にかけて実施された。VQR の結果は、MUR が FFO のボーナス部分の 80% を配分するために使用されるとともに、2016 年法律 232 号に従って、MUR が卓越した 5 年間の研究開発プロジェクトを特定し支援するために用いられた。

VQR2015-2019 では、130 以上の機関（国立大学、非国立大学、通信制大学、研究機関、その他任意で評価に参加する機関）が評価に参加した。

評価には科学分野ごとの多数の専門家に加え、アシスタント、統計データ分析専門のスタッフ、データ評価のためのプラットフォーム開発のための技術者が参加した。

評価対象となった学術論文（論文、単行本、寄稿文など）が、前回の VQR では約 11 万 8000 件であったのに対し、VQR2015-2019 は約 18 万 3000 件と大幅に増加し、新たに各機関の第三ミッション活動¹¹⁵に関連する 600 件以上のケーススタディが評価対象となった。研究者数は約 6 万 5 千人、評価対象となった大学数は 98 校（前回比 4 校増）、研究機関数は 14 機関（同 2 機関増）と増加したが、自主的に評価申請を行った機関数は 22 機関（同 4 機関減）と減少した。評価作業は 18 の評価専門家グループで分担され、イタリア国内外の機関から約 645 名の研究者が評価者として参加した。ANVUR からは、職員、管理職、

¹¹³ トリノ大学ウェブサイト：Research Quality Assessment (VQR)

< <https://en.unito.it/research/research-quality-assessment> > [2025/3/21 取得]

¹¹⁴ ANVUR ウェブサイト：Rapporto finale VQR 2015-2019 < <https://www.anvur.it/it/ricerca/qualita-della-ricerca/rapporti-di-valutazione/vqr-2015-2019> > [2025/3/21 取得]

¹¹⁵ VQR2015-2019 では「科学普及から学術起業、芸術・文化遺産の制作・管理、臨床試験、技術移転、オープンサイエンスのための革新的ツール、持続可能な開発目標を含む国連 2030 アジェンダに関連する活動まで、さまざまな手段を通じて相互作用しながら、研究機関が社会と接触するすべての活動」と定義されている。

運営審議会のメンバーに加えて、特別に選抜された合計 24 名の GEV アシスタントが評価に関与した。

VQR2015-2019 では質的評価の対象項目として、①研究者のプロファイル、②採用方針のプロファイル、③研究活動のプロファイル、④第三ミッション活動のプロファイルが審査された。前回 VQR からの最大の改正点は、上記の第三ミッション活動¹¹⁶の事例評価が加わったことである。

表 2-16 : VQR2015-2019 における評価基準等

評価項目	評価基準	内容
研究の質	独創性	研究の科学的対象との関連において、その成果物が新しい考え方や解釈を導入し、同じ対象に関するこれまでのアプローチから突出し、革新的であるか
	方法論の厳密さ	目的と文献の現状を明確に提示し、研究対象に適した方法論を採用し、目的が達成されたことを実証しているか
	影響力	国際的な科学界、あるいは適切な分野であれば国内の科学界に影響を、あるいは及ぼす可能性があるか
第三ミッション活動の質	影響の社会的、経済的、文化的側面	経済的、科学的、文化的、社会的、制度的分野において、態度を変え、意識を高め、技能や優れた実践を開発し、理解のプロセスを洗練させたか。これらの項目が実質的に改善された結果をもたらすことへの貢献があったか。
	参照文脈との関連性	地域、地方、国、欧州、国際的な次元との関連における活動の重要性、および参照文脈への影響の量的または質的な増分効果を評価する。「参照文脈」とは、エコシステムに関連する外部文脈を意味するが、組織の制度およびミッションステートメントの特性および実施活動に関連する内部文脈も意味する。
	受益者にとっての付加価値	様々な構成要素における社会の利益のために、活動事例によって誘発された技術的、経済的、社会的、文化的条件の革新と改善を指す。また、技術的な革新だけでなく、社会的な革新の度合いも評価され、さらに、効果的な社会正義を実現するための前提条件として、公正、平等、機会均等、不平等と脆弱性の除去という原則の尊重も評価される。
	提案組織の貢献	質的・量的な貢献と機関側のコミットメントを、人材、組織構造・プロセス、インフラ、および/または新たな専門スキルの開発という観点から評価する。必要な場合は、当該機関の科学研究との関連性を確認する。科学研究との関連性は、間接的、非直線的であったり、あるいはまだ最終的な成果として具現化されていない場合もあるが、明確で実証可能なものでなければならない。

出所: VQR2015-2019 より未来工学研究所作成

以上の基準に従って研究の質および第三ミッション活動を評価し、結果は以下のスコアによって表現される。

a) 優秀かつ極めて適切 (ECR) : スコア 1

¹¹⁶ 評価対象となる第三ミッション活動は、a. 知的財産または工業所有権の強化、b. アカデミックな起業家精神、c. 仲介および技術移転施設 d. 芸術的・文化的商品の生産と管理、e. 臨床試験と健康保護への取り組み、f. 生涯学習とオープンエデュケーション、g. 市民参加活動 h. 社会的・教育的性質の公共財の生産と、インクルージョンのための政策、i. オープンサイエンスをサポートする革新的なツール、j. 国連 2030 アジェンダおよび持続可能な開発目標 (SDGs) に関連する活動の 1 つ以上をカバーするものである。

- b) エクセレント (EC) : スコア 0.8
- c) スタンダード (ST) : スコア 0.5
- d) 十分な関連性 (SUF) : スコア 0.2
- e) 関連性が低い、または許容できない(SR) : スコア 0

特長

最新の VQR においては、若手研究者の登用における質、第3 ミッション活動すなわち社会的課題解決等へのインパクトおよび第3 ミッション活動の一環としてのオープンアクセスの強化が研究評価において重視されている。

2.4.2 主要機関における研究開発評価システムの体系とその内容

(1) CNR における研究開発評価システム

1) 機関概要¹¹⁷

全国科学研究センター (Consiglio Nazionale delle Ricerche : CNR) は、MUR の監督下にある、学際的な研究能力を持つ国立の公的研究機関である。1923 年に設立され、主要分野における科学研究プロジェクトを実施し、その成果を国の発展のために応用し、技術革新を主導するとともに、研究システムの国際化を促進し、産業システムの競争力を育成することを任務としている。

CNR は現在、人類と地球の健康、環境とエネルギー、食糧と持続可能な農業、輸送と生産システム、ICT、新素材、センサー、航空宇宙など、さまざまな分野で現代の課題に取り組んでいる。その研究活動は、イタリア全土で活動する約 8,500 人の所員（うち 7,000 人以上が研究および研究支援活動に従事）を通じて実施されている。

CNR の科学ネットワークは、88 の研究機関と 7 つの科学部門から構成されている。大学や企業の研究者との国際的な共同研究は CNR の果たす重要なネットワーク機能である。

CNR は、大学を除けばイタリアの研究者養成において博士課程に関与する最大の組織であり、高い科学的知名度、革新性、研究の質を特徴とするコースを通じて、若者の高度な教育訓練に貢献することを使命としている。

¹¹⁷ CNR ウェブサイト : Chi siamo <<https://www.cnr.it/it/chi-siamo>> [2025/3/21 取得]

2) 機関評価

a. 自己評価

2011年に実施されたCNRによる研究評価についての報告書では、自機関の評価手法は他の主要な欧州研究機関の評価手法と比べて、①評価が継続的に行われていない、②研究者が評価結果について議論する機会がない、③量的な側面ばかりが強調されている、④戦略的な勧告に欠けるといった課題が挙げられており、従前の評価にはインパクトがないと総括されている (Giuffrida et al. [3])。

こうした自己評価結果を受け、2022年、CNRは、研究評価システムの改革を含む「再出発計画 (Relaunch Plan)」を承認した。2022年11月、改革の手始めとしてCNRはCoalition for Advancing Research Assessment (CoARA) 協定 (The Agreement on Reforming Research Assessment : ARRA) ¹¹⁸に調印し、CNRの研究者と技術者のキャリアアップのための評価設計全体が変更されることになった¹¹⁹。

関連し、2023年のCNRの研究者と技術者のキャリアアップのための選考は、試験的にARRAで示された評価手順の基本原則に基づいて行われた。

試験的選考の目的は、2020年の選考で適用された基準と、2023年に導入された基準の両方を分析し、後者の選考がCoARA協定の原則をどの程度遵守できているかを把握することである。さらに、試験的選考が終了した後、評価者と被評価者の双方から、新しい手続きに関するフィードバックを収集することが予定されている。

今回の試験的選考でCNRが採用した基準を精査することで、副次的に機関レベルでのオープンサイエンスや責任ある研究評価 (RRA) への移行の度合いを定量的・定性的に評価し、改革への道筋を明確にすることも企図されている。今後予定されている試験的選考の事後分析は、将来の評価設計を担当するCNR事務局にとって重要な情報となり、現在実施されている評価プロセスの調整と改善に役立つものと期待されている。

b. 外部評価

機関における研究の質および第3ミッション活動に関してANVURによる外部評価 (VQR) を受けている。上記のようにVQRは、評価対象機関における研究の質の向上を促

¹¹⁸ Coalition for Advancing Research Assessment (CoARA)は、研究、研究者、研究組織が評価される方法とプロセスの改革に取り組む団体の連合である。現在の研究評価方法は、被引用数といった書誌情報指標に大きく依存しており、研究者による幅広い貢献を認識できないことが多いとの問題意識がCoARA発足の背景にある。加盟する700を超える研究組織、資金提供者、評価機関、専門家協会が、研究、研究者、研究組織の評価における改革を実施するための共通の方向性と指導原則に合意し、2022年7月に改革と実施の概要を示すCoARA協定 (The Agreement on Reforming Research Assessment : ARRA) を発表した。欧州科学財団 (European Science Foundation (ESF)) が主催するCoARA事務局が同団体の日常業務を管理している。CoARAウェブサイト : What is CoARA? <<https://coara.eu/>> [2025/3/21 取得]

¹¹⁹ grasposウェブサイト : CNR, a reform in the making - analysing and supporting career progressions < <https://www.graspos.eu/cnr-a-reform-in-the-making-analysing-and-supporting-career-progressions>> [2025/3/21 取得]

進し、実績に応じた資金配分を行うために、様々な科学分野における研究の最新の状況を評価するものである。CNR の最新の評価結果は概ね、評価の全項目において平均をやや下回るスコアとなっている¹²⁰。

3) プログラム評価

「研究プロジェクト@CNR」プログラム (programma "Progetti di ricerca @CNR") は CNR 理事会で承認された 2020-2022 年度 3 年活動計画「復興に向けた研究」の指針(2020 年 6 月 17 日決議 第 157/2020 号)において、CNR の研究者のプロジェクト活動に特化した資金措置として導入されたものである。

プログラムは、CNR の科学コミュニティに資金を提供し、若手研究者の成長を支援し、経験豊富な研究者に従来の公的および民間の研究資金では提供できない機会を与えることを目的とする。

総額 400 万ユーロ (1 テーマ/80 万ユーロ) のプログラム資金が用意され、上記指針で特定された 5 つの科学分野 (下記) に配分される。

- 最も脆弱なグループを支える技術：若者と高齢者
- イタリアの未来のための自然資本と資源
- ポスト COVID-19 社会の産業転換とレジリエンス
- パンデミックの時代における新たな生物医学的優先課題
- COVID-19 パンデミック後の農業食品チェーンの再プログラミング

プロジェクトの予算は 1 チームあたり 5 万~20 万ユーロで、評価パネルによる採択審査に合格したプロジェクトには全額が出資される。プログラムでは研究提案を段階的に評価し、第 1 フェーズ終了後、提案ごとのランキングリストが発表され、第 2 フェーズに進んだチームは 30 日以内に最終プロジェクト案を提出する。当初 757 件あった提案は第 1 フェーズで 122 件に絞られ、最終的に第 2 フェーズを通過した 28 のプロジェクトが採択された¹²¹。

プログラムではプロジェクト採択後もステージゲート評価を行う。プロジェクトの初年度については、まず拠出金総額の 50%が、プロジェクト開始時に主任研究者の所属施設に支払われる。残りの 50%は、プロジェクト初年度の終了から 60 日以内に提出される中間的な技術・科学的報告書および財務報告書が主任研究者の関連組織に提出され、評価を担当する専門家パネルに到達した時点で支払われる。2 年度目以降の助成金は、中間報告書の評価結果が良好であり、経費の妥当性が確認された場合に支払われる。部分的に否定的な評価が

¹²⁰ ANVUR ウェブサイト : Rapporti di Istituzione_CNR < <https://www.anvur.it/it/ricerca/qualita-della-ricerca/rapporti-di-valutazione/vqr-2015-2019> > [2025/3/21 取得]

¹²¹ CNR ウェブサイト : Primo Bando Progetti@CNR < <https://www.cnr.it/it/primo-bando-progetti-cnr> > [2025/3/21 取得]

下された場合、あるいはプロジェクトの進捗に重大な欠点があった場合、助成金は減額されるか、次年度以降は授与されないこともある¹²²。

2.4.3 評価疲れについての各国・地域の現況

イタリアでは研究評価が定着してそれほど時間が経っておらず、欧州他国との比較では研究評価の立ち遅れや実施上の課題も各種文献において多く指摘されているところである。また、上記のように CNR においてもインパクトの高い評価と研究実施のサイクルを実現するための組織改革である「再出発計画」が実行中であり、いわゆる評価疲れの声はイタリアにおいてはそれほど大きくはないのではないかと考えられる。

【課題の指摘】

実際に、CNR による最新のイタリアの科学技術イノベーションプログラムについての報告書は、以下の諸課題及を指摘して研究評価の強化を提案している¹²³。

○課題

- 他の欧州諸国に比べ、ホライズン 2020 フレームワークプログラムまで欧州の研究プログラム 4 の主軸であった KETs (Key enabling technologies) や GSCs (Grand Societal Challenges) に関連する目標への志向はほとんどない。KETs や GSCs は、一部の産業研究公募でのみ言及されている。持続可能性に関連するトピックに研究目標を広げることには出遅れている。これらトピックは実際には、PNRR の開始とともに、研究開発のための公的政策により明確に導入された。
- 資金調達プログラムの設計と実施のための閣僚機関中心の管理は、しばしば施策の実施を遅らせ、欧州のプログラム外でも、専用リソースのプールによる戦略的目標に関する共同研究プログラムのための政府機関間の協力ネットワークの形成を妨げている。
- 入札募集の開始が不連続であり、イニシアチブの発表から助成金の交付が開始されるまでの待ち時間が長い。これらの要因は、公的資金調達手段の利用可能性に依存して計画できる活動に関して、科学コミュニティに有害な不確実性を生み出している。これらの特徴は、法的な規定によって正式に設立されたにもかかわらず、資金が提供されないという意味で、「紙の上」にしか存在しない制度の存在と相まって、潜在的な受益者にさらなる不確実性をもたらしている。

○改善提案

財政的な観点だけでなく、プロジェクト・ポートフォリオの多様性の検討という観点も重要である。投資の重複を避けるために、研究開発支援のための公共政策ミックスの管理に関

¹²² CNR ウェブサイト : Guida Operativa Rendicontazione dei Progetti@CNR
< <https://www.cnr.it/it/primo-bando-progetti-cnr> > [20253/21 取得]

¹²³ CNR [4], pp51-54

する考察を開始することが最も重要であると思われる。この考察は、科学的・財務的報告の要件への準拠に限定されるものではなく、体系的な観点からプロセスと結果を監視することを目的とした、実施された業務の絶え間ない評価活動に基づいていなければならない。

【VQR に対する批判】

これに対して、「評価疲れ」の語を直接に使っているわけではないが、現在のイタリアにおける研究評価制度の主流である VQR に対する批判的な見解として以下がある。

すなわち、MIUR の高等教育研究局長（当時）であった Prof. Giuseppe Valditara は、2019 年初頭に「書誌学的方法で行われる現在の評価システムは、アルゴリズムの算出結果をよく見せたり、あらかじめ設定された閾値を克服する目的のみに有用な論文生産活動をもたらすことがあり、小規模な分野や学際的な分野の発展の可能性を減らしている」と指摘している。また、VQR によって ANVUR 自身が「A-class Journals」と認定したジャーナルで出版することを研究者に事実上強制し、自身の研究分野のジャーナルでのみ出版することを要求しているため、本来望まれていたはずの学際性の実現を自ら否定しているとの見解もある¹²⁴。

【参考文献】

- [1] JST CRDS, 科学技術・イノベーション動向報告～イタリア編～, 2009年3月.
- [2] JST CRDS, 科学技術・イノベーション動向報告～イタリア編～ (2016年度版), 2017年3月.
- [3] Stefania Giuffrida et al, Research evaluation of the Italian CNR institutes: a missed opportunity?, 2011.
- [4] CNR, Relazione sulla ricerca e l'innovazione in Italia, Napoli, 2023.

¹²⁴ 前掲 Furiosi (2021)

2.5 調査・分析結果の比較と示唆

以下では、昨年度調査の対象である EU、フランス、イギリス及びオランダを含め、比較分析を行う。なお、昨年度対象国・地域については、1年前の調査結果であることに留意されたい。

2.5.1 評価制度における研究開発評価の位置づけ

各国・地域の評価制度における研究開発評価の位置づけを比較すると次のようなものである。

表 2-17 評価制度における研究開発評価の位置づけ

国・地域	概要
米国	<ul style="list-style-type: none"> 現在の評価制度は、研究開発評価を含め、予算制度とも連動した「政府業績成果現代化法(GPRAMA)」の下、事業等の循環的な改善を図る仕組み。 GPRAMAと調和させた「エビデンスに基づく政策形成法(EBPM法)」により、政策立案から実施、評価に至るまでの組織としての業務マネジメント能力の向上に向けた様々な取組を義務付け。
EU	<ul style="list-style-type: none"> 欧州連合機能条約において、評価をめぐる措置に関して加盟国と欧州委員会等に「何が可能か」を規定。 評価の詳細はプログラムごとに個別規則で規定。Horizon Europe 規則では、プログラムの継続的なモニタリングを義務付けているほか、プログラム評価の枠組みやプロジェクトの事前評価に関する具体的な方法、手続き等を規定。
独国	<ul style="list-style-type: none"> ドイツの科学技術・イノベーションシステムは、連邦政府と州政府が分権的に関与し、研究機関や研究プロジェクトに対して共同して資金配分を行っている。 評価について府省統一のガイダンスはないが、DeGEval(評価学会)の政策横断的基準に基づき、評価を実施している(同学会の会員機関は 212 機関。FhG-ISI のほか、プロジェクトトレーガー、連邦政府機関が含まれる)。また、開発協力分野を所掌する連邦経済開発省(BMZ)では、OECD 開発援助委員会(DAC)の関連原則・基準、DeGEval の基準に基づき実施している。
仏国	<ul style="list-style-type: none"> 評価の基本方針、考え方を定めたものとして、予算法に関する組織法律(LOLF)と研究法が存在。LOLF では、研究開発評価を含む評価一般に関する枠組みや方法などの「やるべきこと」について、研究法では関連する政策形成・実施機関や研究開発実施機関の責務、研究に関わる国家戦略の評価や機関評価等についてそれぞれ規定。 LOLF の下、毎年秋に策定される次年度予算のための法案 PLF において、プログラムごとにプログラム評価に利用するための目標、指標、実績分析等が書かれた年次業績計画(PAP)が付される。 公共政策評価(EPP)の制度化に向けた議論が進行中であり、2008 年の共和国憲法

	改正時に評価の最終責任が国会にあることが明示化された。
英国	<ul style="list-style-type: none"> • 1979年のサッチャー政権以降、国主導による行財政改革を実施し、各種の政策評価システムを導入。大蔵省によるスペンディング・レビュー(SR)や規制評価、会計検査院(NAO)のレビューなど、多様な評価を実施。 • 大蔵省がプログラム評価のガイダンスとして「グリーンブック」を公表。評価結果が大蔵省による予算査定に用いられるため、事前評価が重視されている。
伊国	<ul style="list-style-type: none"> • 研究評価はイタリア共和国憲法第97条に謳われている善政の原則を、科学的・学術的領域において保証する。研究評価では、研究活動の厳密な科学的側面と社会経済的側面の両方が考慮される。 • 公的研究機関である全国科学研究センター(CNR)は国家(全国)大学システム研究評価委員会(ANVUR)から機関評価を受け、自らも実施する研究プログラムの評価を行う。 • 全国研究評価委員会(CNVR)は大学・研究省(MUR)の所管する研究ファンドである国益に資する研究プロジェクト(PRIN)、イタリア科学基金(FIS)のプログラム評価を行っている。
和蘭	<ul style="list-style-type: none"> • 説明責任法により、各省庁が政策の有効性と効率性について定期的に評価することが一般的に義務付けられている。また、一般行政法では、補助金制度について少なくとも5年に1回評価することが、財務省の「国家予算規則」では、税支出スキームについて少なくとも5年に1回評価することがそれぞれ定められている。さらに、少なくとも7年に一度、包括的な政策レビューの実施義務がある。 • オランダでは伝統的に大学をはじめとする教育研究機関の自律性が高く、研究開発評価については、「戦略評価プロトコル(SEP)」のような学術研究の質を保証する制度が長年にわたって安定的に存在。SEPは、オランダ大学協会(VSNU)、オランダ科学研究機構(NWO)、オランダ王立芸術科学アカデミー(KNAW)が整備、運用しているものであり、オランダの公的機関における研究の質、妥当性、実行可能性を、機関ごとの戦略面から6年周期で評価。評価の対象はさまざま、分野全体を対象とした全国的な評価から、単一組織や組織内の分野全体の評価、単一のセンターや研究グループの評価まで、多岐にわたる。

出所:調査結果をもとに未来工学研究所作成

すべての国・地域において、研究開発評価、特にプログラム評価は、一般的な政策評価制度の一部として実施されており、制度間で重複的な作業が発生することはない。また、これらの政策評価制度は、アカウントビリティのためというより、政策改善やマネジメント改善を主目的として運用されており、政策立案や実施の実効性を高める上で欠かせないツールとして関係組織の隅々に浸透している様子が窺える。このことは、我が国における評価制度そのものの見直しが必要であることを示唆している。

2.5.2 国レベルにおける研究開発評価に関する規定の整備状況

国レベルにおける研究開発評価に関する規定の整備状況をまとめると、次のようなものである。

表 2-18 研究開発評価に関する規定等の整備状況

国・地域	概要
米国	<ul style="list-style-type: none"> • GPRAMA や EBPM 法等に基づく評価活動を支援するための規定を大統領府行政管理予算局(OMB)が通達や覚書といった行政文書として整備。 • 通達文書 OMB Circular No. A-11 では対象となる機関や方針の詳細を、OMB M-21-27 などの覚書では評価のための実践的なガイダンスを提供。 • OMB では、EBPM 法に基づく活動を各機関が実施していく上で、追加的なリソースや支援が必要であることを認識しており、連邦政府職員専用のウェブサイト上に相互研鑽を促すためのプラットフォームを構築。
EU	<ul style="list-style-type: none"> • それまであったインパクト・アセスメントや事後評価、実施などのガイドラインを統合し、EU の政策や法律の目的・目標を最小のコストで達成できるように、政策のデザインや準備から評価・改訂までのサイクル全体をカバーする Better Regulation Guideline を 2015 年に策定(ステークホルダー協議に関するガイダンスを含む)。 • FP のレビューを含め、EU におけるすべての政策はこのガイドラインに基づいて実施されており、ガイドラインの情報を補完・補強するためにツールボックスも提示。
独国	<ul style="list-style-type: none"> • 国レベルで、研究開発評価に関する体系的な規程等は整備されていない。研究開発評価は、個々の研究開発プログラムの特性に応じて実施されている。 • 評価について府省統一のガイダンスはないが、DeGEval(評価学会)の政策横断的基準に基づき、評価を実施している。 • ドイツの研究・イノベーション政策の評価については、研究・イノベーション審議会(EFI)が担っている。EFI は、科学技術・イノベーション政策全般を評価し、政府に対して助言を行っている(報告書の提出先は首相)。EFI は、年次報告書のほか、調査研究レポート、ポリシーブリーフ等を公表している。
仏国	<ul style="list-style-type: none"> • 次年度予算のための法案 PLF に付される年次業績計画(PAP)に関し、その実務的なとりまとめを行っている経済・財務・産業・デジタル主権省(MEFSI) 予算局が詳細なパフォーマンス・ガイドを策定。
英国	<ul style="list-style-type: none"> • 研究開発関連政策に特化した規定等はないが、大蔵省が評価に関するガイダンスとして、前述の「グリーンブック」及び「マゼンタブック」を策定。 • 前者では、政策、プログラム、プロジェクトの各階層を対象に、事前、中間、事後といった各時期における評価の設計と使用に関する詳細な情報を、後者では、プログラムの実施期間中及び事後の評価方法に関して詳細な情報をそれぞれ提供。
伊国	<ul style="list-style-type: none"> • 研究開発関連政策に特化した規定等はないが、MUR によって策定される全国研究計画(PNR)の付属文書中で、オープンサイエンスの実現のための研究評価のあり方に

	<p>ついて規定されている(PNR2021-2027 のオープンサイエンスのための全国計画)。当該計画では、研究評価のあり方について具体的な目標とアクションプランが示されている。</p> <ul style="list-style-type: none"> • MUR の所管するファンドのうち、国益に資する研究プロジェクト(PRIN)では研究提案の科学的評価及び社会経済的評価についての詳細なガイドラインが整備されている
和蘭	<ul style="list-style-type: none"> • 機関評価やプログラム評価に関し、SEP では、最小限の労力で研究評価の効果を最大化することを目的としており、機関側は自己評価書を作成し外部委員会の評価を受けるが、評価に使用する指標は画一的なものではなく、被評価者が自らの研究を評価するのに適していると考えられる指標を自律的に決めることができるとするなど、柔軟性に富んだ規定になっている。 • プロジェクト評価に関し、代表的資金配分機関である NWO では、NWO 補助金規則において、補助金が授与される際に適用される条件と義務を規定。特定の助成プログラムには追加条件が適用される場合があり、その場合は当該プログラムの提案募集に詳細な条件と手続きを記載。

出所：調査結果をもとに未来工学研究所作成

諸外国で整備されている規定と比べて、我が国における「研究開発評価に関する大綱的指針」や「文部科学省における研究及び開発に関する評価指針」の内容に大きな過不足があるわけではない。

一方、調査対象としてとりあげた国・地域では、評価において何をなすべきかをまとめた規定に加え、より良い評価の実践を支援するためのガイダンスを整備している。さらに、米国では、実践コミュニティの構築を意識した関係者間での相互研鑽・学習を促進するためのプラットフォーム作りも進められている。今後は、ガイダンスの整備を含めたこうした支援体制を充実していく必要がある。

2.5.3 プログラム評価における自己評価、外部評価、第三者評価の使い分け等

プログラム評価における自己評価、外部評価、第三者評価の使い分け、及び評価者について比較すると次のようなものである。

表 2-19 自己評価、外部評価、第三者評価の使い分け等

国・地域	概要
米国	<ul style="list-style-type: none"> • GPRAMA 及び EBPM 法等に基づくプログラム評価は、基本的には自己評価により実施。ただし、評価のガイドライン(OMB Circular No. A-11)ではピアレビュー等の外部評価を活用することが評価の信頼性を高める有効な手段であると推奨。そのため、外部専門機関に評価のための調査分析を委託したり、それらの機関による評価研究を参照にしたメタ評価を実施することが多い。 • これらの法に基づく定型的な評価以外にも、多様なチェック・アンド・バランスの仕掛け

	<p>が存在。たとえば、ナノテク・イニシアチブのような、議会からの注目度の高い省庁横断的なプログラムについては、個別法の規定により全米科学・工学・医学アカデミーによる第三者評価や大統領科学技術諮問会議(PCAST)による外部評価が定期的に行われる場合もある。</p> <ul style="list-style-type: none"> • 機関独自に外部評価を活用するケースもあり、NSF では、外部有識者からなる委員会(COV)が 4 年ごとに各プログラムの外部評価を行い、その結果を採択審査の仕組み改善などに活用している。
EU	<ul style="list-style-type: none"> • 調査分析(評価研究)は外部専門機関に委託して実施し、それらの結果をもとに内部でメタ評価(自己評価)を実施。また、分野やテーマに沿って設置される独立したハイレベル専門家グループによるレビュー(外部評価)も並行して行われる。
独国	<ul style="list-style-type: none"> • プログラム評価は、対象プログラムの性格に応じて行われており、特定の評価体制が推奨されているわけではない。例えば、エクセレンス・イニシアティブのプログラム評価では、国際的な専門委員会を設置し、外部評価を実施している。 • 海外からの研究人材の受入プログラム等では、プログラムを所掌する府省のプログラム評価法が用いられている。
仏国	<ul style="list-style-type: none"> • 評価の最終責任は国会にあり、国会には評価を関係機関に遂行させる義務があることを憲法で明記。 • LOLF や PLF などの法で評価のあり方を規定する一方、評価に用いられる具体的な目標や指標はプログラムごとに主管省・機関が主導して決定。
英国	<ul style="list-style-type: none"> • UKRI では、プログラム評価にあたって、多くの場合、外部の独立した研究者(シンクタンク等)に委託を行い、それを内部の評価運営グループが監視する、という体制で実施(外部評価の活用)。内部で評価作業を行う場合であっても、調査は独立した組織に委託を行うことを推奨。いずれにせよ、評価結果をプログラムの設計と実施に確実に反映させるために、被評価者であるプログラム関係者を最初から評価に関与させる仕組みになっており、自己評価的な側面も持ち合わせている。
伊国	<ul style="list-style-type: none"> • 機関を対象とした研究の質評価(ANVUR の行う VQR)ではピアレビューの方法論を採用し、科学分野に応じた計量書誌学情報を併用する。評価対象の研究結果について、その独創性、方法論の厳密性、影響力に関する判断を表明する任務を専門評価者グループ(GEV)に委ねる。 • PRIN における研究評価では、採択時の事前評価は欧州を含む地域の傑出した外部評価者による評価に委ねられている。
和蘭	<ul style="list-style-type: none"> • SEP に基づく研究プログラムの評価では、研究機関の自律性に配慮し、自己評価を基点として、外部評価委員会(多くの場合国際的な人選)の評価をもとにした評価案を、研究ユニットと外部評価委員会や執行部との協議の中で練り上げていくという評価者と被評価者間のコミュニケーションを重視した評価手法をとっていることが特徴。評価における具体的な評価指標を被評価者であるユニット自体が開発し選定することができるなどユニークな仕組みを有する。

出所:調査結果をもとに未来工学研究所作成

調査対象国・地域では、政策改善やマネジメント改善が評価の第一義的な目的となっており、プログラムの運営に責任を持つ主体による自己評価が多くの場合取り入れられている。その意味で、評価の独立性はさしたる論点ではなく、目的達成の観点から、国や組織の実情に照らして、より適切な形式が選択されていると言える。

ただし、こうした方式が有効に機能しているのは、施策やプログラムの立案段階において、事後的に検証・確認すべき事項やその方法などを明確に示した評価計画が合わせて立案されているからである。すなわち、仮説・検証の方法として、自己評価が有効であることがあらかじめ明確になっているからであり、こうした評価計画を持たずに自己評価を行った場合、結果の信頼性を大きく損なうことになる。

なお、自己評価が基本であっても、評価のための調査分析を外部の専門機関に委ねるなど、自組織に不足する専門性を適宜補っていることに留意する必要がある。

2.5.4 研究開発課題の評価（プロジェクト評価）において重視されている評価の時期

プロジェクト評価において重視されている評価の時期について比較すると、次のとおりである。

表 2-20 プロジェクト評価において重視されている評価の時期

国・地域	概要
米国	<ul style="list-style-type: none"> プロジェクト評価は、専門性を有する資金配分機関の裁量に基本的に委ねる。 NSF では、COV による採択審査システムの厳正な評価が定期的に行われているという前提もあり、課題(プロジェクト)に対する中間評価及び終了時の評価は、予算規模の大きなものなど一部の例外を除き基本的には実施されない。その代替として、プロジェクト実施側の義務として、研究代表者に対し、研究の進捗及び研究費の用途に関して定期的に報告書をまとめることを要求。 「Growing Convergence Research (GCR)プログラム」のようにステージゲート方式を採用するプログラムもあるが、約 94%が次の段階に進んでおり、切るための評価というより、研究チームをエンカレッジするための評価という色合いが強い。
EU	<ul style="list-style-type: none"> 事前評価については詳細な定めがあるものの、中間評価や事後評価については規定もなく、基本的には実施されない。 ただし、プログラム評価のための情報源として、プロジェクトから収集される各種情報が活用される。
独国	<ul style="list-style-type: none"> DFG における研究助成プロジェクトの採択から実施までのプロセスとして、初期審査(承認)、進捗報告、終了時の評価がある。 事前評価が重視されるが、プロジェクトの種類により書面審査のみで行われる。長期のプロジェクトでは、採択後も進捗報告をもとに中間評価を行う。
仏国	<ul style="list-style-type: none"> プロジェクト評価は、基本的には、プログラムの運営機関がプログラムごとに評価の実施要項を規定して実施。

	<ul style="list-style-type: none"> • ANR では、近年、事後評価を強化する方向で検討が進む。ただし、評価結果は、機関の方針決定や ANR のプログラム編成のレビューに活用することが想定されており、個別の課題を査定するためのものではない。
英国	<ul style="list-style-type: none"> • 事前評価、すなわち採択審査が重視され、採択後もモニタリングを行うが、プロジェクト評価というより、プログラム評価の参照情報として活用される。
伊国	<ul style="list-style-type: none"> • 事前評価、すなわち採択審査が重視され、採択後もモニタリングを行う。数年にわたるプロジェクトではステージゲート評価が行われている例がある。
和蘭	<ul style="list-style-type: none"> • NWO では、事前及び事後評価を規定中に明記し、必ず行われるものとして位置付け。これらの評価にあたっては評価委員会または審査委員会といった諮問機関が組織される(外部評価)。中間評価については研究に関するデータの報告を求め進捗を監視することが基本となっており、本格的な中間評価は所管大臣の指示がある場合などに限られる。

出所:調査結果をもとに未来工学研究所作成

研究開発課題(プロジェクト)の評価では、一般的に事前評価が最も重視され、中間評価や事後評価は行われないことが多い。今回の調査対象国・地域もその例外ではない。ステージゲート方式のような中間段階での評価を導入する事例もいくつかみられるが、予算には限りがあるため、どの提案が最もよくプログラムの目的に貢献しうるかを見極めるための事前評価がまずは重視されることになる。したがって、プロジェクトが結果として失敗に終わった場合であっても、その原因をプロジェクト実施者に帰するのではなく、採択審査の機能不全に求めることがほとんどである。その意味で、採択後にプロジェクトから収集される情報は、プログラムの仕組み改善に活かされるものでなければならないと言える。

2.5.5 評価疲れをめぐる議論とそれらへの対応

評価疲れをめぐる各国・地域における議論とそれらへの対応をまとめると次のようなものである。

表 2-21 評価疲れをめぐる議論とそれらへの対応

国・地域	概要
米国	<ul style="list-style-type: none"> • 各種評価制度間の調和が図られており、評価結果の有効活用に向けた制度改善なども行われていることなどから、「評価疲れ」は指摘されていない。 • また、連邦政府から求められる書類作成等の情報提供に関する報告者(個人、中小企業、教育・非営利機関、連邦政府との契約者、州・地方政府)の負担を最小化することを目的とした Paper Reduction Act があり、エビデンスデータの収集とペーパーワーク削減のバランスへの配慮が法的に義務付けられている。 • 機関レベルでも研究者の負担をかけずに公的研究費の助成を受けた研究者やプロジェクトのデータを追跡するシステムの整備が進んでいる。

EU	<ul style="list-style-type: none"> • EU では「評価」が文化となっており、「評価疲れ」に関する指摘はみられない。一方、モニタリングにあたっては、情報収集にかかるコストは重要性に照らして適正化すること、何が必要かを事前に十分に分析した上で必要最低限のデータを収集することが求められている。 • 既存の法律行為の事後評価と改正案の事前インパクト評価が並行して行われる「バック・トゥ・バック」のプロセスが近年急増しており、それぞれで義務付けられているステークホルダー協議に関して、「協議疲れ」を引き起こす可能性があることが課題として認識されている。
独国	<ul style="list-style-type: none"> • ドイツにおける研究者の研究プロジェクト等の評価の負荷について、直接的なデータはないものの、研究開発評価の管理上の負荷が全体的な業務量に影響を与えていると考えられている。 • 研究者は、プレプリントの公表（もしくはプレプリント情報の活用等）が増えてきており、当該情報の発信や活用に係る丁寧な対応が必要となる。また、ミッション志向型研究開発等が進む中で、研究プロジェクトの評価では、政策介入の因果関係が評価できるよう、各種情報の提出が求められる可能性があると考えられる。
仏国	<ul style="list-style-type: none"> • 「評価疲れ」に言及した文書等は見受けられないが、プログラムやプロジェクトの各種データの管理システムの一元化、デジタル化によって評価業務も含めた一連のプログラム管理やプロジェクト管理などの業務を合理化、簡素化するための活発な動きが見られる。
英国	<ul style="list-style-type: none"> • 評価に要する膨大な事務経費と機会費用が発生することについては日本と同様の問題意識が指摘されているが、資金配分や組織的な学習、研究戦略の策定やマネジメントへの活用等と評価が不可分のものと捉えられており、評価への不満は評価制度改善のためのある種のエネルギー源となっている。 • Innovate UK では、評価実施者や被評価者、評価のための情報提供者の負荷を軽減するために、プログラム予算の多寡や社会的関心の程度、リスクや不確実性の程度などを踏まえて、レベル分けを行った上で評価を実施することを推奨。
伊国	<ul style="list-style-type: none"> • イタリアでは研究評価が定着してそれほど時間が経っておらず、欧州他国との比較では研究評価の立ち遅れや実施上の課題も各種文献において多く指摘されているところである。いわゆる評価疲れの声はイタリアにおいてはそれほど大きくはない。 • CNR においてもインパクトの高い評価と研究実施のサイクルを実現するための組織改革が実行中である。 • CNR は、研究投資の重複を避けるために体系的な観点からプロセスと結果を監視することを目的とした、実施された業務の絶え間ない評価活動実施を提言している。
和蘭	<ul style="list-style-type: none"> • 公的文書等において「評価疲れ」に言及したものは見受けられないが、SEP の評価に携わる大学研究者の現場の声として、評価の実効性に対する懸念が一部で表明されている。 • オランダにおける科学技術政策の最上位戦略である「科学ビジョン」の評価を担うラテ

	ナウ研究所の担当者からは、「研究者に過剰にヒアリングを行ったり、評価に関連する書類作成をさせて疲弊させる事態を避けるには、評価には常に制約が伴うことを評価者自身が自覚し、評価したいことの優先順位を明確にすることが重要である」との指摘がある。
--	--

出所：調査結果をもとに未来工学研究所作成

我が国における「評価疲れ」の問題は、その結果が有効に活用されないことに起因しており、評価者、被評価者双方の作業負荷の問題と相まって、評価自体の必要性に対する疑問へと繋がっている。

一方、今回の調査対象国・地域においては、評価のための情報提供にかかる作業負荷に対する不満などは見られるものの、評価自体の必要性を疑う声は聞こえてこない。評価は、施策やマネジメントの改善のために欠かせないツールであるという認識がまずあり、そのようなツールをいかにブラッシュアップしていけるか、という観点で評価に関わる問題が捉えられていると言える。こうした問題は、作業負荷の低減によって解消できるものではなく、評価制度の実効性を高めていくことでしか解決できないものである。

3. 文部科学省研究開発評価人材育成研修の開催

3.1 開催概要

3.1.1 趣旨・目的

研究開発評価にあたっては、評価に関する制度・枠組みに対する理解や知識が必要であるが、行政や研究助成機関、大学、研究機関等では、評価の実務担当者がジョブローテーションや任期のため数年ごとに変わり、評価の経験・知識が組織内に蓄積されにくく、業務の継承がなされていないことが課題となっている。

このような背景の下、文部科学省では、研究開発評価の効率化及び質の向上に資するとともに、各機関における研究の推進や評価に関わる教職員の質の向上を図る観点から研究開発評価人材育成研修を開催してきた。今年度も、大学、大学共同利用機関法人、研究開発法人及び行政において研究開発評価の業務に従事している教職員を対象に、研究開発評価の業務に関する知識等の習得や実際の事例に係る情報共有を目的として「研究開発評価人材育成研修」を開催した。

3.1.2 日程・会場

開催日程及び会場は次のとおりである。

<開催日>

2025年1月21日(火)～1月22日(水)

<開催場所>

新虎安田ビル NIKAI カンファレンス

3.1.3 対象者及び募集方法

対象者及び募集方法は次のとおりである。なお、受講料は無料とした。

<対象者>

大学、大学共同利用機関法人、研究開発法人及び行政において、研究開発評価に関わる業務¹²⁵に1年以上従事している教職員を対象とした。なお、過年度の同研修の受講者も再受講可能とした。

¹²⁵ 研究開発評価に関わる業務とは、各機関における組織的な研究戦略の策定、戦略に基づいて実施される研究活動及び研究評価活動等に関連する業務を広く意味する。具体的には、以下のような業務を含む：競争的資金のプログラム設計やプロジェクトの募集選考、マネジメント、プログラム評価業務；競争的資金の獲得支援や採択後のマネジメント、学内ファンドのマネジメント・評価業務；研究機関の戦略立案、IR業務。

<募集方法>

科学技術振興機構サイエンスポータルで広く一般に募集をかけるとともに、大学、大学共同利用機関法人、研究開発法人及び行政といった対象ごとに情報展開先を選定した上で周知を行い、未来工学研究所ウェブサイトに申し込み専用ページを置いて受付を行った。

具体的な情報展開先は次のようなものである。申し込み状況をみながら、検討会議委員に個別に関係者への周知を依頼したり、対象を拡大したりした。

表 3-1 情報展開先（対象別）

対象	情報展開先
大学、大学共同利用機関法人、研究開発法人	リサーチ・アドミニストレーション協議会
大学、大学共同利用機関法人、研究開発法人、行政	大学評価コンソーシアム
大学、大学共同利用機関法人、研究開発法人、行政	研究・イノベーション学会
大学、大学共同利用機関法人、研究開発法人、行政	JST プログラムマネージャー研修受講生・修了生メーリングリスト
大学、大学共同利用機関法人	研究大学コンソーシアム
行政	文部科学省担当課等から連絡
研究開発法人	文部科学省担当課等から連絡

出所：未来工学研究所作成

3.1.4 参加者

定員 30 名に対し、25 名からの参加登録があった。このうち、2 日間とも欠席した者は 2 名だった。事後アンケート（後述）の結果から受講者 23 名の属性等を整理すると次のようなものである。

(1) 所属機関

回答者の所属機関は国立大学法人が半数を占めている。属性の割合は、昨年度とほぼ同様であった。

表 3-2 回答者の所属

①国立大学法人	52.2%	(12名)
②公立大学	0%	(0名)
③私立大学	21.7%	(5名)
④国立研究開発法人	13.0%	(3名)
⑤その他(大学共同利用機関法人、文部科学省、省庁)	13.0%	(3名)

出所：未来工学研究所作成

(2) 主な職務内容

主な職務内容としては、「競争的資金（学内ファンド含む）のプログラムの設計やプロジェクト募集選考、マネジメント、評価業務」と「競争的資金獲得支援（プレアワード）」が、それぞれ5名で同数であった。「その他」は、「法人評価業務」、「教育支援部署の長」、「法人評価業務」等であった。

昨年度と比較すると、「競争的資金（学内ファンド含む）のプログラムの設計やプロジェクトの募集選考、マネジメント、評価業務」が半数に減り、その分、「競争的資金獲得支援（プレアワード）」の人数が増えた。「その他」は4名増加した。

表 3-3 主な職務内容

①競争的資金(学内ファンド含む)のプログラムの設計やプロジェクトの募集選考、マネジメント、評価業務	21.7%	(5名)
②競争的資金獲得支援(プレアワード)	21.7%	(5名)
③競争的資金獲得後の支援・マネジメント(ポストアワード)	4.4%	(1名)
④産学等連携支援業務	13.0%	(3名)
⑤その他	39.1%	(9名)

出所:事後アンケート結果をもとに未来工学研究所作成

3.1.5 受講理由

受講理由としては、「職務上必要と感じたから」とする者が約4割、「研修内容（講義）に興味があったから」とする者が約半数であった。「その他」は、「組織評価、法人評価、研究戦略策定業務担当者である」というものであった。

昨年度と比較すると、「研修内容（講義）に興味があったから」が大きく増加し、今年度の受講理由の第1位となった。

表 3-4 受講理由

①職務上必要と感じたから	39.1%	(9名)
②研修内容(講義)に興味があったから	47.8%	(11名)
③上司に勧められたから	8.7%	(2名)
④その他	4.4%	(1名)

出所:事後アンケート結果をもとに未来工学研究所作成

3.1.6 プログラムの構成

プログラムは、初日の講義（座学形式）と2日目の演習（ワークショップ形式）からなる。演習では、プログラム評価の事例を取り上げ、ロジックモデルの作成や評価を実施する上での問い、分析手法の検討などを通じて、研究開発評価において着目すべき基本的な観点等について学ぶ内容とした。

具体的なプログラムは次のとおりである。

表 3-5 プログラム（第 1 日目）

開始	終了	時間	内容
13:00	13:10	10 分	趣旨説明
13:10	13:35	2 分	【講義 1】研究開発評価を巡る現在の状況 講師: 祐野 恵(京都大学) 内容: 国の研究開発評価の全体像や文部科学省の評価指針等を紹介
13:35	14:20	45 分	【講義 2】研究開発評価の意義と枠組み 講師: 伊地知 寛博(成城大学) 内容: 評価の意義や基本的な概念等、習得すべき基礎的な知識を解説
14:20	14:30	10 分	休憩
14:30	15:15	45 分	【講義 3】研究開発課題(プロジェクト)の評価 講師: 安藤 二香(未来工学研究所) 内容: プログラムにおけるプロジェクト評価について事例に基づき紹介
15:15	16:00	45 分	【講義 4】研究者等の業績評価、研究開発機関等の評価 講師: 鳶田 敏行(大学改革支援・学位授与機構) 内容: 研究者の評価や機関評価、IR について事例を交えて紹介
16:00	16:10	10 分	休憩
16:10	16:55	45 分	【講義 5】プログラム評価 講師: 田原敬一郎(未来工学研究所) 内容: プログラム評価について事例に基づき紹介
16:55	17:00	5 分	事務連絡

出所: 未来工学研究所作成

表 3-6 プログラム（第 2 日目）

開始	終了	時間	内容
9:30	9:35	5 分	事務連絡
9:35	9:45	10 分	演習の目的・タイムスケジュール等の説明
9:45	10:00	15 分	自己紹介(各グループ): 評価業務の現状や課題認識、参加目的等の共有
10:00	10:10	10 分	事例の説明
10:10	10:50	40 分	【ワーク 1 課題の設定】個人及びグループワーク
10:50	11:00	10 分	【ワーク 1 課題の設定】全体共有と解説
11:00	11:05	5 分	【ワーク 1 課題の設定】理解の深化
11:05	11:25	20 分	【ワーク 2 有効性の検討】個人及びグループワーク
11:25	11:35	10 分	【ワーク 2 有効性の検討】全体共有と解説
11:35	11:40	5 分	【ワーク 2 有効性の検討】理解の深化
11:40	12:40	60 分	昼休憩
12:40	13:40	60 分	【ワーク 3 ロジックモデルの作成】個人及びグループワーク
13:40	13:50	15 分	【ワーク 3 ロジックモデルの作成】解説
13:50	13:55	5 分	【ワーク 3 ロジックモデルの作成】理解の深化
13:55	14:15	20 分	【ワーク 4 仮説および検証方法】個人及びグループワーク

14:15	14:30	15分	休憩
14:30	15:10	40分	【ワーク4 仮説および検証方法】全体共有と解説
15:10	15:20	10分	【ワーク4 仮説および検証方法】理解の深化
15:20	15:40	20分	振り返り
15:40	15:55	15分	講評・閉会挨拶
15:55	16:00	5分	アンケート協力要請、事務連絡

出所：未来工学研究所作成

なお、昨年度の研修と比較すると、以下の点で異なっている。

- ・ 講義1「研究開発評価を巡る状況」については、昨年度の振り返りを踏まえ、講義パートの全体像と各講義の位置付けに関する説明を加えた。また、DORA や CoARA など国際的な研究開発評価改革の動向についても合わせて紹介を行うこととした。
- ・ 研修時間について、昨年度の振り返りを踏まえ、初日の座学は各講義の時間を5分ずつ延長した。2日目の演習については、各ワークの解説後に5分から10分間グループ内で理解を深める時間を設けた。

3.2 開催結果

研修終了後に、出席した計23名に対して事後アンケート（オンラインアンケート）を送付し、全員から回答があった。以下、結果の概要を示す。

3.2.1 講義内容に対する評価及び意見

事後アンケートでは、講義別に満足度と参考の程度をその理由とともにたずねている。以下では、それぞれについて結果をまとめた。

(1) 講義1：研究開発評価を巡る現在の状況

講義1では、祐野恵氏（京都大学）に、国の研究開発評価の全体像や文部科学省の評価指針等を紹介してもらった。

1) 満足度

講義1に対する満足度は次のとおりである。

表 3-7 満足度（講義1）

①とても満足した	73.9%	(17名)
②やや満足した	26.1%	(6名)
③やや不満であった	0%	(0名)
④不満であった	0%	(0名)

出所：事後アンケート結果をもとに未来工学研究所作成

主な意見としては次のようなものである。

「とても満足した」「やや満足した」に回答

- プログラム評価の必要性と重要性、ロジックモデルの使い方を理解することができた。
- 評価の三層構造や、国の評価の指針の経緯等、体系的な部分の説明が良かった。
- 俯瞰して、制度的なことや心構え、向き合い方を理解することができた。
- 現在、職場で不足している内容の理解がすすんだ。職場へのフィードバックが多く出来そう。

2) 参考の程度

講義1に対する「参考の程度」への評価は次のとおりである。

表 3-8 参考の程度（講義1）

① すぐに活用できる部分が多い	39.1%	(9 名)
② 将来的には役立ちそう	60.9%	(14 名)
③ あまり役立ちそうにない	0%	(0 名)
④ まったく役立ちそうにない	0%	(0 名)

出所:事後アンケート結果をもとに未来工学研究所作成

主な意見としては次のようなものである。

「すぐに活用できる部分が多い」「将来的には役立ちそうだ」に回答

- 法人評価業務の理解に役立つ。
- 学内担当部署への説明に役立つ。
- 自身の研究活動や大学評価の観点で、さっそく研修内容に紐づいた思考実験を始めている。
- 第7期科学技術・イノベーション基本計画に備え、今後の情報収集に役立ちそう。
- 政策から研究開発プロジェクトまでの位置付けを体系的に理解でき、プレアワード支援への重要視点を得られた。
- 参考になる可能性はあると思うが、当面そのような業務の予定が無い。
- 現在、この内容を直ちに取り扱う業務を行っていない。

(2) 講義2：研究開発評価の意義と枠組み

講義2では、伊地知寛博氏（成城大学）に評価の意義や基本的な概念等、習得すべき基礎的な知識を解説してもらった。

1) 満足度

講義 2 に対する満足度は次のとおりである。

表 3-9 満足度 (講義 2)

① とても満足した	78.3%	(18 名)
② やや満足した	21.7%	(5 名)
③ やや不満であった	0%	(0 名)
④ 不満であった	0%	(0 名)

出所:事後アンケート結果をもとに未来工学研究所作成

主な意見としては次のようなものである。

「とても満足した」「やや満足した」に回答

- 評価の意義を根本的なところから理解できた。
- 自分の考えていた「評価」について見つめなおす良い機会となった。
- 法人評価に役立ちそう。
- 学内説明に活用したい。
- 研究開発評価について、これほど体系的に学ぶ機会には他に無い。

2) 参考の程度

講義 2 に対する「参考の程度」への評価は次のとおりである。

表 3-10 参考の程度 (講義 2)

① すぐに活用できる部分が多い	39.1%	(9 名)
② 将来的には役立ちそうだ	60.9%	(14 名)
③ あまり役立ちそうにない	0%	(0 名)
④ まったく役立ちそうにない	0%	(0 名)

出所:事後アンケート結果をもとに未来工学研究所作成

主な意見としては次のようなものである。

「すぐに活用できる部分が多い」「将来的には役立ちそうだ」に回答

- 法人評価に役立ちそう。
- 担当部署や理事への説明に活用したい。
- 資料が豊富で、現在の業務の点検・棚卸に活用できそう。
- 講義の内容を踏まえて、いろいろと考えるという点では役立つ。
- 今現在担当している業務とは直接関連しない。
- 担当業務に直結しないため、日常的に意識しづらい。
- 背景知識として非常に重要な反面、実務ですぐ使うというものではない。

(3) 研究開発課題（プロジェクト）の評価

講義 3 では、当研究所の安藤二香がプログラムにおけるプロジェクト評価について事例に基づき講義を行った。

1) 満足度

講義 3 に対する満足度は次のとおりである。

表 3-11 満足度（講義 3）

① とても満足した	82.6%	(19 名)
② やや満足した	17.4%	(4 名)
③ やや不満であった	0%	(0 名)
④ 不満であった	0%	(0 名)

出所:事後アンケート結果をもとに未来工学研究所作成

主な意見としては次のようなものである。

「とても満足した」「やや満足した」に回答

- 法人評価に役立ちそう。
- 学内での説明会で活用したい。
- プロジェクト評価の方法について理解することができた。
- 実事例との関係で、プログラムにおけるプロジェクト評価の観点を得られた。
- 具体的になじみのあるプログラムが、どのように作られたかの理解が深まった。
- 具体的な事例を用いて説明されており、大変わかりやすかった。
- 所属機関に所属する研究者の研究分野は、人文学・社会科学が中心であり、RISTEX は公募申請の候補として注視している。RISTEX を例として、研究開発課題を学ぶ良い機会となった。

2) 参考の程度

講義 3 に対する「参考の程度」への評価は次のとおりである。

表 3-12 参考の程度（満足度）

① すぐに活用できる部分が多い	65.2%	(15 名)
② 将来的には役立ちそう	34.8%	(8 名)
③ あまり役立ちそうにない	0%	(0 名)
④ まったく役立ちそうにない	0%	(0 名)

出所:事後アンケート結果をもとに未来工学研究所作成

主な意見としては次のようなものである。

「すぐに活用できる部分が多い」「将来的には役立ちそうだ」に回答

- 法人評価に役立ちそう。
- 学内担当者などへの説明に活用したい。
- 公募申請準備に活用できそう。
- 申請支援などでも、プログラムの意図を洞察して研究者に伝えるということが、URAとして行うべきことだと思った。
- 具体的なロジックモデルの例も示してあり、自分の取組でロジックモデルを考える際には、すぐに活用できる。
- 学内の施策を企画する際に、事後評価の方法論を知っておくことは重要であると感じた。
- 担当業務では、なかなか活用場面がなさそう。

(4) 講義 4：研究者等の業績評価、研究開発機関等の評価

講義 4 では、寫田敏行氏（大学改革支援・学位授与機構）に研究者の評価や機関評価、IR について事例を交えて紹介してもらった。

1) 満足度

講義 4 に対する満足度は次のとおりである。

表 3-13 満足度（講義 4）

① とても満足した	87.0%	（ 20 名）
② やや満足した	8.7%	（ 2 名）
③ やや不満であった	4.4%	（ 1 名）
④ 不満であった	0%	（ 0 名）

出所：事後アンケート結果をもとに未来工学研究所作成

主な意見としては次のようなものである。

「とても満足した」「やや満足した」に回答

- 実際の評価に携わっており、かつ、現在は大学評価改革支援・学位授与機構の先生の講義で、非常に良かった。
- 具体的かつ実践上の課題が共有されていた。
- 評価疲れという抽象的なテーマについて、定量的なアプローチで検証されている点が勉強になった。
- 評価に付随する負のイメージ（不信感、不透明感、負担感）を越え、支援的な評価が重要であり、それは実践可能であることが、先生のご経験から伝わった。
- 教員評価の問題点について勉強になった。
- きちんと理解できればとても役に立つと思うが、自分の理解が不十分であった。

「やや不満であった」「不満であった」に回答

- ・ 教員評価をどのように位置付けるのかという点はよく理解できた。もう少し、しっかりと話を聞きたかった。

2) 参考の程度

講義 4 に対する「参考の程度」への評価は次のとおりである。

表 3-14 参考の程度（講義 4）

① すぐに活用できる部分が多い	60.9%	(14 名)
② 将来的には役立ちそう	39.1%	(9 名)
③ あまり役立ちそうにない	0%	(0 名)
④ まったく役立ちそうにない	0%	(0 名)

出所:事後アンケート結果をもとに未来工学研究所作成

主な意見としては次のようなものである。

「すぐに活用できる部分が多い」「将来的には役立ちそうだ」に回答

- ・ 法人評価に役立ちそう。
- ・ 現在取り組んでいる業務が、まさに講義のなかで紹介された「ありがちな例」に陥っていたため、業務の見直しに活用できる。
- ・ 教員評価の課題について議論の材料となる。
- ・ 評価の目的に沿った適切な「問い」の設定は実践していきたい。
- ・ 表員評価を担当する場合には、役に立つと思う。
- ・ 現在の業務で評価に関連する内容があまりないため、すぐには活用できない。今後、学内の支援事業など設計等をする際には、参考にしつつ取組みたい。

(5) 講義 5 : プログラム評価

講義 5 では、当研究所の田原敬一郎がプログラム評価について事例に基づき解説した。

1) 満足度

講義 5 に対する満足度は次のとおりである。

表 3-15 満足度（講義 5）

① とても満足した	73.9%	(17 名)
② やや満足した	26.1%	(6 名)
③ やや不満であった	0%	(0 名)
④ 不満であった	0%	(0 名)

出所:事後アンケート結果をもとに未来工学研究所作成

主な意見としては次のようなものである。

「とても満足した」「やや満足した」に回答

- 仮説検証としての評価、ロジックモデルの用途について理解できた。「やらなければならない」「やらされるもの」というイメージを持っていたので、それを払拭できた。
- ロジックモデルについて理解が深まった。これまでに漠然と考えていたものを体系的に学ぶ良い機会となった。
- よく理解できたが、他の講義とのオーバーラップがあった。ロジックモデルの活用やその意義などに特化しても良かった。
- プロジェクト、プログラム、アウトプット、アウトカムをシステム論的にコミュニケーション基盤として使えるように整理されていた。
- 翌日のワークの予習として、ボリューム・内容ともに満足できた。
- 実務において、どのレベルで活用するか、もう少し時間をかけて考えたかった。

2) 参考の程度

講義 5 に対する「参考の程度」への評価は次のとおりである。

表 3-16 参考の程度（講義 5）

① すぐに活用できる部分が多い	73.9%	(17 名)
② 将来的には役立ちそう	26.1%	(6 名)
③ あまり役立ちそうにない	0%	(0 名)
④ まったく役立ちそうにない	0%	(0 名)

出所：事後アンケート結果をもとに未来工学研究所作成

主な意見としては次のようなものである。

「すぐに活用できる部分が多い」「将来的には役立ちそうだ」に回答

- 法人評価に役立ちそう。
- ワークの内容と合わせて業務に活用できる。
- 申請書等のレビュー支援で使えそう。
- 学内担当部署への説明に活用したい。
- 今後、業務においてロジックモデルを考える機会がある見込みなので、その際に今回学んだ内容を活用したい。
- 取り組んでいる事業の見直し、改善に使っていききたい。
- 課題の整理として役に立つ方法を学べた。

3.2.2 プログラム評価演習に対する評価及び意見

演習では、プログラム評価の事例として「元素戦略プロジェクト<研究拠点形成型>」を取り上げ、ロジックモデルの作成や評価を実施する上での問い、分析手法の検討などを通じて、研究開発評価において着目すべき基本的な観点等について学ぶ内容とした。

演習は、「課題の設定」「有効性の検討」「ロジックモデルの作成」「仮説および検証方法」の4パートで構成され、それぞれについて、5グループに分かれてのグループワークと全体での解説をセットにして行った。また、各グループでのワークの様子を観察し、適宜全体でコメントを行うようにした。

(1) 満足度

演習に対する受講者の満足度は次のとおりである。

表 3-17 満足度（演習）

① とても満足した	81.8%	(18 名)
② やや満足した	18.2%	(4 名)
③ やや不満であった	0%	(0 名)
④ 不満であった	0%	(0 名)

出所：事後アンケート結果をもとに未来工学研究所作成

主な意見としては次のようなものである。

「とても満足した」「やや満足した」に回答

【開催形式／ファシリテーション】

- コンパクトに全体を体験できる組み立てで良かった。
- ファシリテーターが、グループワークで意見や振り返りを上手に引き出していた。
- ファシリテーターの方の適切な交通整理がありがたかった。
- 内容や流れについては非常に学ぶことが多く、よかった。
- 時間的な制約が厳しいと感じた。
- 2日と長い研修ではあるが、ワークショップについては、もう少し時間があると望ましいと感じた。
- 講義の内容を実践でき、理解が深まった。

【参加者の交流】

- 他大学の取り組み、気づきなどがあり役に立った。

【研究戦略への応用】

- 研究戦略策定に役立つと思われる。

【ロジックモデル】

- ロジックモデルの考え方は、どの業務にも活かせると感じた。
- ロジックモデルという言葉自体の知識はあったものの、実際の研究開発プログラムでどのように実装されているかという点を、実例をもって学ぶことができた。
- 実際にチームメンバーと議論しながらロジックモデルを作成するのは、座学での理解とはまた違う気づきがたくさんあった。
- 具体的な例に対して、WS を実施することで、具体的な活用法を理解することができた。

(2) 参考の程度

演習に対する「参考の程度」への評価は次のとおりである。

表 3-18 参考の程度（演習）

① すぐに活用できる部分が多い	77.3%	(17 名)
② 将来的には役立ちそう	22.7%	(5 名)
③ あまり役立ちそうにない	0%	(0 名)
④ まったく役立ちそうにない	0%	(0 名)

出所:事後アンケート結果をもとに未来工学研究所作成

主な意見としては次のようなものである。

「すぐに活用できる部分が多い」「将来的には役立ちそう」に回答

【活用可能な業務】

- プレアワード、ポストアワードといった個別の業務だけでなく、大学の方針決定など大枠の業務にも応用できそう。
- 日常の業務の取り組みを改善したい。
- 学内担当部署への説明や、URA が実施する説明会で活用したい。

【個人や組織での活用・課題】

- 研究開発評価システムの有効的な導入が必須な状況だが、内部調整に時間を要する。

【ロジックモデル】

- 様々な角度から同じ目的について議論することが重要。その時に言葉だけで議論せずにロジックモデルを目の前に共有するのはとても効果的で、これはどんな業務にも応用できる。
- 今後の業務で、ロジックモデルを考える機会がある見込みである。今回の経験を参考にしたい。
- ロジックモデルの考え方は、各種の施策企画に活かすことができる。

3.2.3 研修全体に対する評価及び意見

(1) 満足度

研修全体に対する満足度は次のとおりである。

表 3-19 満足度（研修全体）

① とても満足した	91.3%	(21 名)
② 満足した	8.7%	(2 名)
③ やや不満であった	0%	(0 名)
④ 不満であった	0%	(0 名)

出所:事後アンケート結果をもとに未来工学研究所作成

主な意見としては次のようなものである。

「とても満足した」「満足した」に回答

- 系統だった。
- 講義だけでなくワークもあり実践的だった。
- 直ぐに活用できる。
- 研究開発評価に係る一通りの知識・経験をする、非常に良い機会となった。
- 講師がフラットでリラックスした雰囲気を作ってくれたことで、学びを深めることができた。
- 「評価」人材の育成を狙いとしているが、方法論としては、共通性が高い。

(2) 研修実施希望時期

事後アンケートでは、希望する研修の実施時期についてもたずねている。その結果、6割以上が「1月～3月」に集中し、「4月～6月」の年度はじめは希望者がゼロであった。

表 3-20 研究実施希望時期

① 4月～6月	0%	(0 名)
② 7月～9月	21.7%	(5 名)
③ 10月～12月	13.0%	(3 名)
④ 1月～3月	65.2%	(15 名)
* 希望なし	0%	(0 名)

出所:事後アンケート結果をもとに未来工学研究所作成

主な意見としては次のようなものである。

「7月～9月」

- 年度はじめと終わりは業務が立て込む。また、第二四半期に講義を受けられると、その後の業務に活かしやすい。

「10～12月」

- 次年度の取り組みに活用するために良い時期である。

「1月～3月」

- 出張しやすい。
- 申請支援のピークから外れている。
- 翌年度業務に対する意識・姿勢の検討につながるタイミング。
- 年度末で予算の余剰が認識される時期。
- ストーブリーグのタイミング。

3.2.4 今後とりあげてほしいテーマ

今後とりあげてほしいテーマとしては次のようなものである。

【評価手法・プログラム評価の深化】

- 評価指標や手法そのものの評価。
- 研究開発評価の事例紹介。
- プログラム評価をより深めても良いと思う。

【研修のカリキュラムや応用】

- 文科省の初任者研修や管理職研修に今回の内容を組み込んで欲しい。
- 評価設計によって被評価者の行動などがどう変容するかについても知りたい。
- 今回学んだ知識や方法論については、“研究開発”に限らない分野でも使えるように感じた。“研究開発”に固有の難しさや注意事項があれば、そこを重点的に学ぶ機会があるとありがたい。

【研修対象者の拡大・改善提案】

- 来年は、事務方（研究部）の課長にも参加をすすめている。
- URA でなくても、事務方も知っておかなければならない内容も多い。
- 事務だからこそ通常業務の中で知っておかなければならないことも多い。そのような人にも有益になるような講義内容があればよい。

【研究開発以外への応用】

- JICA の国際援助に関する評価も一つの応用例として興味がある。同じスキームによる評価の実例として紹介してもらえれば。

3.3 今後に向けての示唆

事後アンケートの結果から、研修全体として満足度の高いものとなっており、基礎的な知識の獲得や事例に基づく講義に加え、ワークショップ形式の演習という構成は適切であったといえる。

ここでは、検討会議における議論をもとに、今回の変更点に対する振り返りや今後に向けての示唆をまとめる。

3.3.1 講義について

- すべての講義について受講者の満足度も高く、特段の改善点はみられない。

3.3.2 演習について

- ワークショップであれだけ付箋が出されるのは、受講者が自主的に参加したというエビデンスの 1 つである。受講者がワークに没入できており、満足度の高いものになったと言える。
- 全体としては、ワークの完成度を高めたいとの意向もあるだろうが、この種の研修は少し物足りないくらいがよいのではないか。もう少しやりたかったと思うことで、次の学習につながる。
- 各ワークの後に 5 分間の振り返りの時間を設けたことは、他のグループからのコメントに対して皆で考えることができ、良かった。これにより、当該グループでは出てこなかった意見が他のグループから出ているということや、ファシリテーターとしてグループの傾向とは別の視点を投げかけ、理解を深めていくことができた。
- アウトプットとアウトカムについて、先にアウトカムを議論するのはなぜか、時系列で考えるとアウトプットを先に議論すべきではないかといった質問が複数名からあった。ワークに入る前のインストラクション等で明確に伝えるようにすると良いのではないか。
- プログラム評価について学んだことを組織に持ち帰って実施しようとした場合、コストがかかる。これは実務者にとっては深刻であり、ワークの中で説明があるとよいのではないか。こういうことを実施するのはよいが、かかるコストをどう捉えるべきか、議論できるとよい。また、学内ファンドを担当している受講者もいたが、今後さらに増える可能性がある。1 日目でもよいが、どこかで明示的に触れるとよいのではないか。