

I -15.研究開発評価に関する海外動向調査

Cross-Country Analysis of the R&D Program Evaluation System

キーワード Key Word	研究開発プログラム評価、資金配分機関、目的基礎研究
	R&D Program Evaluation, Funding Agency, use-inspired basic research

1. 調査の目的

本調査は、国立研究開発法人科学技術振興機構（JST）戦略的創造研究推進事業の研究開発評価の改善に資することを目的に、海外資金配分機関の研究開発プログラムの評価動向等を調査するとともに、提言のとりまとめを行ったものである。

2. 調査研究成果概要

(1) 調査の構造

本調査は、1)海外ファンディング機関のうち JST 戦略的創造研究推進事業の参考となる研究開発の評価手法などの調査と、2)JST 戦略的創造研究推進事業の国際評価・領域評価などにおける評価手法などのレビューおよび提案のとりまとめの2つで構成されている。

1)について、調査対象は8カ国10機関のファンディング機関(次表)であり、それぞれについて、①対象機関の概要(ナショナル・イノベーション・システムの特徴と位置づけ、事業の構成、組織、予算)、②評価システムの概要(機関評価、プログラム評価、プロジェクト評価)、③プログラム評価の事例(評価対象の概要、評価の実施体制、評価の内容、評価の方法、評価結果およびその活用)といった観点から情報を整理するとともに、④調査対象機関における取組からの示唆をまとめた。

	国	調査対象 FA	着目点(選定理由等)
1	米国	全米科学財団(NSF)	CREST 類似の新興技術支援プログラム(EFRI)を対象に外部有識者による評価を実施
2	米国	エネルギー省(DOE) ARPA-E	採択審査システムを含むプログラム運営にフォーカスした評価を第三者評価で実施
3	欧州連合	欧州委員会欧州研究会議(ERC)	世界最先端の研究を支援する ERC の実務を担う ERCEA の運営を評価、上位のフレームワークプログラムの評価等に活用
4	欧州連合	欧州委員会コミュニケーションネットワーク・コンテンツ・技術総局(DG Connect)	TRL1-5 をカバーする3つのサブプログラムからなる新興技術支援プログラム(FET)について、多様な評価を試行
5	英国	英国研究イノベーション機構工学・物理科学研究会議(EPSRC, UKRI)	英国全体の数理科学分野の国際レビューを実施、ベンチマーク等により国の強み・弱み等を把握し、プログラム改善等に活用
6	ドイツ	ドイツ研究振興協会(DFG)	拠点形成プログラム(エクセレンス・クラスター)の採択審査システムを体系的に評価
7	フランス	国立研究機構(ANR)	FA 後発国ではあるが、世界標準の評価システムを積極的に導入
8	オランダ	科学研究機構(NWO)	プロジェクトの長期的なインパクトを終了5年後、10年後にレビュー、実用化報告書で公表
9	スウェーデン	イノベーション・システム庁(Vinnova)	プログラムの設計から終了数年後のインパクト分析にいたるまで一貫した評価体系を構築
10	カナダ	国立自然科学・工学研究会議(NSERC)	サブプログラムや領域単位ではなく、プログラム全体の評価を法定により5年毎に実施

2)については、研究開発評価の海外動向や評価手法などに詳しい専門家を構成員とする専門家パネル(検討委員会)を設置し、海外の評価動向を参考にしつつ、JST 戦略的創造研究推進事業におけ

る評価手法などのレビューを実施した。また、同事業におけるアウトカムやインパクトの評価の在り方を含めた評価手法などの改善に向けた提案をとりまとめた。

(2) 主な成果

調査結果のうち、調査対象機関における取組からの示唆を抜粋してまとめると次のようなものである。

機関	示唆
NSF	● EFRI は新興領域や新産業の形成、大きな社会的課題の解決といった将来の出口を見据えて変革の機会を拡大することを志向する戦略的な基礎研究支援プログラムであり、研究領域・研究課題の予算規模、複数チームによる研究課題の構成、技術成熟度の観点から CREST と類似。こうしたタイプのプログラムに対しても、外部有識者 (COV) による評価を定期的 (4 年を目処) に実施 ● COV による評価では、必要性・位置づけ、プロセス・セオリー、インパクト・セオリーと幅広く評価を実施。とくに、インパクト・セオリーに関して、著しい成果をあげたプロジェクトの有無や成功率ではなく、新興領域や新産業の形成、大きな社会的課題の解決といった将来の出口を見据えて、変革の機会を拡大する＝適切なプログラムポートフォリオを形成するという入り口を充実させることができているかに焦点をあてることで、プログラムの成果を最大化することを企図。こうしたやり方は、具体的な波及効果があらわれるまでに長期を要する戦略事業のようなプログラムに対して非常に有効 ● 機関レベルの評価に関して、端緒についたばかりであるが、組織学習に主眼を置いた「Learning Agenda」も注目に値
DOE	● 個々の領域に対する評価ではなく、ARPA-E というプログラムの運営に係る評価に重点。プロジェクト等の技術評価の妥当性やプログラム運営に係る評価を実施 ● 多様なメトリクスを用いながら、他の DOE 部局のプログラムと成果を比較、プログラムの成功を評価する試みを実施 ● プロジェクトレベルの目標達成のために設定した内部のメトリクス・データベースからのデータをプログラムレベルの目標のメトリクス (商業・非商業的な短期・長期のアウトカム) につなげる作業を行うこと、これらのプログラム目標を観察可能なイノベーションのメトリクスにつなげること、さらにそのメトリクスを機関レベルのミッションや目標の達成の進捗に翻訳していくというインパクト評価に関するフレームワークの構築に着手している点、インパクト評価に必要なエビデンスデータの収集に商用データベースや機械学習などの新たなツールの導入に積極的に取り組んでいる点も有用 ● ARPA-E で実施される長期的かつハイリスクの研究開発では、当初の計画が達成できないプログラムについて、プログラム内容を変更せずに研究開発の実施チーム自体を変更することや研究開発チームの統合等を実施。研究開発が長期的かつハイリスクを伴うものは一定の失敗を許容することも重要
ERC	● 設定した個々の評価項目に対して、鍵となる質問を複数作成することにより、どのような状況がプログラムの成功と言えるのかを判断できる材料を幅広く揃える。その上で、これらの質問に対する回答を数値データやアンケート調査結果などより明確に提示できるようにするための、データベースなどの基盤的環境の整備も実施 ● ERCEA の運営面に関する評価報告書のデータの一部が、Horizon2020 の中間評価でも用いられるなど上位レベルの評価と効果的に連動。上位政策で設定されている KPIなどを意識したプログラム運営が行われており、全体としての成果の最大化を意識

	した評価の枠組みを体系的に整備
DG Connect	● TRL1-5をカバーする3つのサブプログラムで構成されるFETでは、競争的に研究開発への資金配分を行うと同時に、プログラム運営や評価に資する実践的研究も公募 ● この中で、「研究者がほとんどいない初期の段階の、根本的に新しい未来の技術に資するアイデアに対して資金提供を行う」FET Open 及び「新興のテーマを育成し、多くの有望な探索的研究トピックでヨーロッパの研究者のクリティカルマスを確立すること」を目的としたFET Proactiveで採択された研究開発課題のインパクト評価を目的としたFET Tracesというプロジェクトを実施。採択課題による研究成果の展開状況、インパクトの事例および類型化などに取り組む。具体的には、知識、人材、経済、社会といった多様な観点からアウトカムやインパクトを捕捉。 ● これらの多様な視点からの検証を行うことで、「利用に触発された基礎研究」の意義が明確化されており、戦略的創造研究推進事業においても、こうしたインパクト評価研究自体を振興していくのも一案 ● プロジェクトの提案段階でインパクトに至る計画やロードマップを詳細に書かせる方向性ではなく、インパクトを事後的に追跡できる方法を充実させることの重要性が提言されており、目的基礎研究の評価のあり方を考える上で示唆的
EPSRC, UKRI	● 英国全体における数理科学研究における国際評価を実施し、数理科学研究における問題群をマクロレベルで把握。ただし、国際評価は、リソースとコストの制約及びその効果という観点から、10年前の実施を最後に行われていない ● EPSRCでは現在、国内の評価関係者のリソースを最大化することや、プログラムの投資効果やプロジェクトのアウトカムを把握することに注力。プロジェクト評価は採択審査のみだが、プロジェクト終了後のアウトカムを捕捉するための報告を義務化。プログラム評価は、目的に応じて評価対象を選択し、より詳細で体系的に実施 ● プログラム評価におけるデータサイエンスの専門家が果たす役割が大きくなっており、専門家がプログラム評価の運営に従事
DFG	● プログラム評価として、レビューアの選定プロセスや選定されたレビューア群による影響等を評価、採択審査プロセスの適切性を分析 ● 採択審査プロセスの妥当性を多面的に検証したものであり、レビューアに対するオンライン調査の質問項目は参考となる。質問項目には、プロジェクトの提案内容だけではなく、参加研究者・機関の経験や研究分野の将来に与える影響等が含まれる。さらに、レビューの方法(査読のみ、パネルメンバーによる議論、PIによるプレゼン)がどのようにレビューア個人の評価に影響を与えているかについても自然科学・工学と人文・社会科学分野に分けて分析しており、示唆に富む
ANR	● 国家戦略から各省庁や資金配分機関、研究実施機関が行う事業に至るまで「プログラム化」が行われ、予算と紐づく形で評価を実施 ● 2005年のANRの設立以降、世界標準の評価システムを積極的に取り入れる。ピアレビューのプロセスも、レビューアの選定方法を含めて体系的に構築 ● プロジェクト全体を様々な観点(採択年度、採択プログラム、学術分野、評価委員会、国家計画、戦略分野、地域、機関種別など)からポートフォリオ化できる情報システムにより、機関レベル、プログラムレベル、プロジェクトレベルでの可視化が容易
NWO	● TRLを考慮してプログラムを設計し、公募でも提示 ● すべてのプログラム及び関連公募に対し、標準化された質問票に基づきアカデミアや

	企業での研究経験があるプログラム・コーディネーター(NWO 職員)が毎年度レビューを実施、それらを踏まえて手続きやコミュニケーション、予算などの見直しなどプログラム運営の改善を図る ● プロジェクトの長期的なインパクトについても終了 5 年後及び 10 年後にレビューを実施、その結果を毎年度発行される「実用化報告書」で公表
Vinnova	● プログラムの設計段階から終了数年経過後のインパクト分析にいたるまで一貫した評価体系を構築。評価の思想・文化や枠組みがスタッフに共有され、継続的な改善と実践が蓄積。スタッフのスキル開発にも注力 ● 体系的なデータ収集システムを構築。採択したプロジェクトに対して追跡調査への協力を義務化 ● 複数のプログラムや他の取組等を含めてインパクト分析を実施。イノベーションの専門機関として上位機関への提言にも活用。イノベティブなプログラム(事業)を次々に開発していく際の源泉としても有用 ● プロジェクト実施者にロジックモデル作成を義務化。ただし、「出口」と必ずしも直結しない目的基礎研究の場合留意が必要。活用する場合、プロジェクトに対する評価やマネジメントのためではなく、研究者にインパクトを意識付けるなど限定した目的で用いるべき ● 課題駆動型イノベーション(UDI)イニシアチブでは、採択審査における女性の評価者を全体の 4 割以上にするという目標を課す
NSERC	● カナダでは、5 年ごとにプログラム評価の実施が法で義務づけられており、専門性を持った民間調査会社に委託した上で、体系的に評価を実施 ● 戦略的創造研究推進事業の国際評価と同様、サブプログラムや領域といった単位ではなく、プログラム全体の評価を実施。5 年という周期で継続的かつ専門的に評価を行うことでみえてくる課題もあり、行動計画を通じて具体的な改善につなげることで、プログラムの存在意義を高めることにもつながる ● 評価部署のスタッフがジョブローテーションせずに長期間評価の実務に携わっていること、評価報告書に評価を行う上で直面した課題などを明記することで、評価実施機関が変わっても質の高い評価を実現 ● プログラム評価のためのデータ利用に関しても、他のファンディング機関と組織連携を図ることで、費用対効果分析などの高度な評価を実現 ● 反事実的な状況(Counter-factual)との比較によるインパクト評価については、不採択プロジェクトがプログラム評価に関する調査に対して協力的でないため実施が困難。不採択を通知する際に今後のプログラム評価に関する協力打診を含めるなどの工夫が必要。あわせて、経常的に不採択プロジェクトに関する研究者へのフォローやデータ収集・分析を実施することも重要