

総合科学技術・イノベーション会議評価専門調査会
大綱的指針ワーキンググループ(第2回)

研究開発評価をめぐる国際動向

プログラム(横断型)の評価を中心に



2026年9月7日

田原 敬一郎

目次

CONTENTS

1. 本報告の目的・位置付け
2. 全体的な動向：STI政策の新たな潮流と評価の役割の変化
3. 全体的な動向トランスフォーマティブ・イノベーション政策(TIP)と評価
4. 全体的な動向STI政策の拡張に伴う新たな評価の課題
5. 米国の事例：省庁横断的課題の目標・業績管理
6. 英国の事例：政策階層間を意識した評価の構造
7. 英国の事例：FAにおける評価方法選択の考え方
8. スウェーデンの事例：国家目標に紐づけた共管プログラムの評価
9. まとめ

本報告の目的・位置付け

前回の大綱的指針改定以降の着目すべき国際動向について、特にプログラム(事業等)を横断する評価に焦点をあてて報告。

評価対象 \ 評価時期		事前評価	途上評価／モニタリング	事後評価／追跡評価
研究開発プログラムの評価	プログラム横断型	機関を横断するシステムレベルの評価 同一機関内におけるプログラム横断型(ポートフォリオ等)の評価		
	個別事業型			
研究開発機関等の評価				
研究開発課題の評価			国際的な研究評価改革の動向	
研究者等の業績評価				

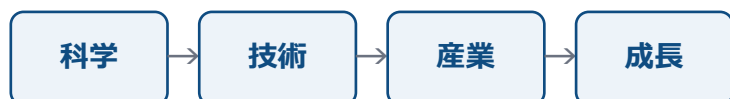
STI政策の新たな潮流と評価の役割の変化

国際的トレンドとして、STI政策の役割・対象範囲が「社会変革」を含むものにまで拡大。こうした文脈の中で、評価に求められる役割なども変化。

フレーム 1

リニア・モデル

「よい研究への投資は、いずれ経済に実る」



考え方

基礎研究への投資が、応用→開発→産業化・経済成長へと一方向に流れていく

政府の介入根拠

「市場の失敗」：企業任せでは基礎研究に資金が回らないため、政府が投資する

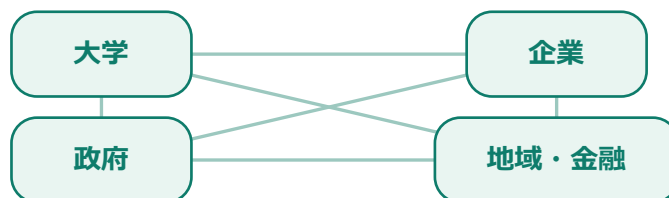
見ている範囲

科学の世界（大学・研究機関が中心）

フレーム 2

ナショナル・イノベーション・システム

「イノベーションはつながりから生まれる」



考え方

大学・企業・政府・地域などの相互作用（つながり）がイノベーションを生む

政府の介入根拠

「システムの失敗」：資金の不足よりも、つながりの機能不全を直すことが課題

見ている範囲

科学+経済（産学官のネットワーク全体）

フレーム 3

トランスフォーマティブ・チェンジ

「技術だけでなく、社会のしくみごとを変える」



考え方

技術単体では解けない課題（気候変動など）に対し、社会のしくみごとを変える。府省・機関を横断した取組が必要。

政府の介入根拠

「方向性・調整・需要・再帰性の失敗」：社会全体でどこへ向かうかを共に定める。不確実性を前提に、試行錯誤を繰り返す

見ている範囲

社会全体（エネルギー・制度・生活・価値観まで）

注:これら3つのフレームはSTI政策の発展形態を表すものではなく、重なりを持ちつつ相互に影響を与え合うものとして位置づけられる。

出所: Schot & Steinmueller(2018)、田原・吉岡・山野・高橋(2025)をもとに作成

トランスフォーマティブ・イノベーション政策(TIP)と評価

政府内組織間にとどまらずセクターを超えた連携が必要な複雑な対象であり、社会変革をもたらす施策のあり方を含め、各国で試行錯誤。

フレーム 3

トランスフォーマティブ・ チェンジ

「技術だけでなく、社会のしくみごとを変える」

科学技術

社会全体（制度・生活・価値観）を変革

考え方

技術単体では解けない課題（気候変動など）に対し、社会のしくみごとを変える。府省・機関を横断した取組が必要。

政府の介入根拠

「方向性・調整・需要・再帰性の失敗」：社会全体でどこへ向かうかを共に定める。不確実性を前提に、試行錯誤を繰り返す

見ている範囲

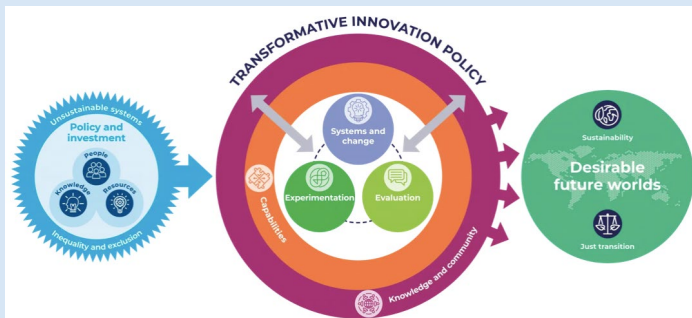
社会全体（エネルギー・制度・生活・価値観まで）

【TIPにおける評価の方向性(例)】

- 評価を政策設計の中核に最初から組み込む
- 形成的 / 発展的 / 現実主義的評価を採用する。柔軟なTheory of Change、リアルタイム観察、参加型手法を用いる
- 反復的・再帰的なモニタリングによって、プロジェクト・プログラム・政策の継続的学習と変更を支える
- 定性・定量を含む複数の方法を組み合わせる
- プロジェクト→プログラム→社会技術システムという入れ子構造(nested)の複数レベルを評価する

出所: JRC, Aligning smart specialisation with transformative innovation policy: Lessons for implementing challenge-led missions in smart specialisation. European Commission, 2023.

【TIPコンソーシアムの構築】



出所:
<https://tipconsortium.net>

- 世界各国で台頭しつつあるTIPに関する実践コミュニティの構築等を目的に、サセックス大学SPRUがユトレヒト大学UGlobe等と協力して運営する国際イニシアチブ。
- TIP Resource Labでは、5つのコンポーネントごとに学習ツール等を提供: 1) 社会技術システムの**変化理論の開発**; 2) TIPのための**実験**; 3) “トランスフォーマティブ・アウトカム”を用いたモニタリング、評価、学習; 4) TIP**実践のための能力**; 5) TIPの**知識とコミュニティ**。

STI政策の拡張に伴う新たな評価の課題

- プログラム及びプロジェクトが実現を目指す価値は、「**科学的・学術的価値**」、「**経済的価値**」、「**社会的価値**」の3つに区分可能。
 - これらの価値創出のベースとなる「研究基盤・エコシステム」の構築自体が重要な価値目標となりうる。
 - これらの価値は本来等価。

価値区分	成果の受け手	価値発現のメカニズム
科学的・学術的価値 (ディシプリン志向)	研究者(ピア)。知識生産者と同じコミュニティ	ジャーナルや高等教育システムを通じて
経済的価値 (ミッション志向)	企業, 消費者等。知識生産者と異なるコミュニティ	市場を通じて
社会的価値 (ミッション志向)	利害関係者、政策決定者、市民等。知識生産者と異なるコミュニティ	公的秩序もしくは市民的秩序を通じて (社会実装・社会的受容、政策過程・政治的受容)

出所:各種資料をもとに未来工学研究所作成

- STI政策は、産業競争力・経済安全保障・地域振興を重視する「**産業への転換(industrial turn)**」が進み、基礎研究や科学的自律性とのバランスが課題となっている。
- STI政策評価には、説明責任や有効性の検証に加え、政策・産業を橋渡しし、「**学習・先見性・内省**」を通じて政策**適応を支える役割**が求められる。

出所:REvaluation Conference 2027
<<https://revaluation2027.eu/components/63265/how-it-works>>

米国の事例：省庁横断的課題等の目標・業績管理

米国では、政府業績成果現代化法(GPRAMA)に基づき、「**政府横断課題の重視**」と「**少数の重要目標への重点化**」の二つの改革潮流を統合。さらに、EBPM法により「**組織学習**」を重視する方向へ。

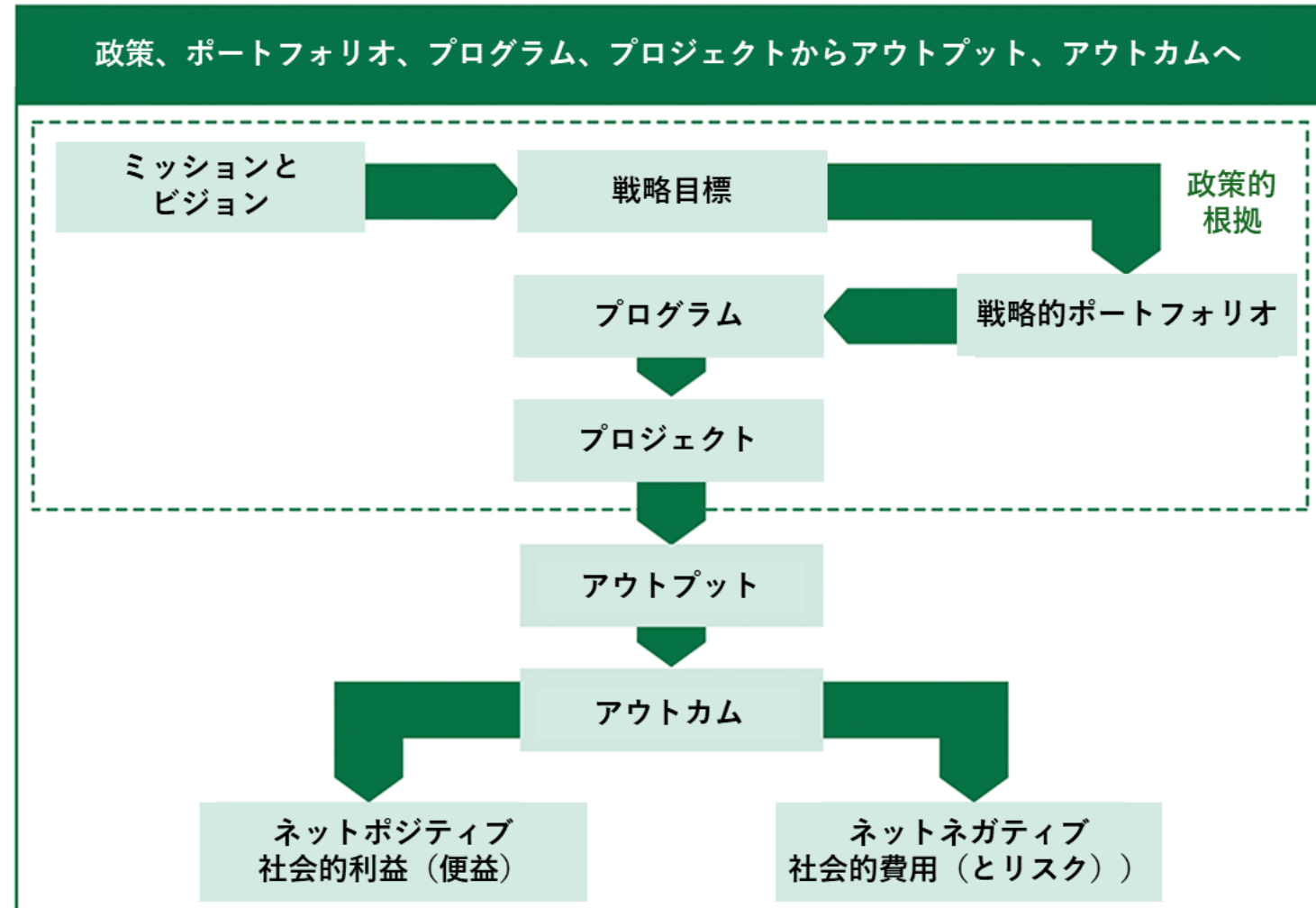


英国の事例：政策階層間を意識した評価の構造

英国では、評価の時期別にガイダンスを整備。政策階層の上位を常に意識した評価の実施を推奨。



- 事前評価を中心とした大蔵省によるガイダンス「Green Book」では、評価対象として、政策-ポートフォリオ-プログラム-プロジェクトといった4つの階層性を明示。
- 特に、共通の**政策目的**と**ポートフォリオレベルのTheory of Changeを設定し**、個別評価を共通質問の下に設計・統合、相互作用とポートフォリオ全体の成果を分析したうえで、事業の継続・中止・追加・資源再配分に結び付ける仕組みを重視。



出所:Green Book 2022 ed.(最新版は2026 ed.)

英国の事例：FAにおける評価方法選択の考え方

英国の資金配分機関（FA）では、予算希規模等に応じて、適切な評価方法を選択することを推奨



Innovate UKの「評価フレームワーク」では、評価にあたって、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実施すべきであり、例えば、新規性があるもの、予算の多いもの、注目度の高いもの、効果や影響について重大なリスクや不確実性があるものについては、評価のために多くのリソースを割く必要がある、としている。

予算とプロファイル	高	多額の予算を伴う大規模なプロジェクトや、メディアや一般市民の関心を集め、高い影響力を持つ可能性のある注目度の高いプロジェクト	レベル2	レベル3	レベル3
	中	中程度の予算と、ある程度のメディアや一般の人々の関心を集め、大きな効果が期待できる中規模プログラム	レベル2	レベル2	レベル3
	低	低予算および/または限定的な一般の人々またはメディアの関心で、比較的低い影響力	レベル1	レベル2	レベル2
			低	中	高
リスクと不確実性					

* 予算基準値
低 <£1m
中 £1-10m
高 £10m+

レベル1：ライトタッチ、前後のモニタリングを含む内部評価を推奨

レベル2：適切な予算配分で外部委託を検討する

レベル3：詳細な外部委託評価、予算はプログラム総額の1～5%を推奨

スウェーデンの事例：国家目標に紐づけた共管プログラムの評価

スウェーデンでは、国家目標の実現に向けて、複数機関を横断する共管プログラムを設定、下位階層の評価結果を活用しながら、全体的、統合的な評価も実施

- スウェーデンでは、同国にとって戦略的に重要な分野における主体主導の協働を通じて、地球規模の社会課題に対する持続可能な解決策と国際競争力の向上に向けた条件整備を図る長期プログラム群「戦略的イノベーションプログラム(SIP)」を推進。

- Vinnova、Formas、Energimyndighetenが共同で資金提供

- 企業、大学・研究機関、公的機関その他の主体が参画

- SIPの評価では、自己評価を材料として活用しつつ、外部評価・専門家評価・メタ評価を組み合わせ、**プログラム群全体の学習と制度改善につなげる**点に特徴。

- 各プログラムは最長12年を上限として開始から3年ごとに評価を受け、3年評価は立上げ段階の点検、6年評価は設定目標への到達状況、資金配分されたプロジェクトの結果及び初期的効果を中心として評価。9年評価は、結果や早期効果の把握、6年評価後の行動計画のフォローアップ、終了段階への提言及び継続資金判断のための資料提供を主目的として実施、総括的性格と形成的性格を併せ持つ評価として設計。

- 9年評価では、各プログラム側の自己評価票を用いつつ、文書調査、レジスター分析、半構造化インタビュー、ウェブ調査、ケーススタディ、専門家評価等などを実施。17件の個別SIP評価や年次メタ評価をもとに、SIP全体の総合メタ評価も実施されており、個別プログラム評価を横断的に統合する仕組みがある。

まとめ

- プログラムの評価では、**評価を学習・適応・方向修正のインフラと位置づけ**、各国で様々な試行錯誤を展開。
- 重要な国家的目標に対し、政策ミックス(資金配分政策だけではなく、規制や税制、公共調達等を含む)の評価など**システムレベルの評価も様々に試行**。上位政策や関連施策における当該プログラムの位置付けを明確にし、有機的に関連づけていくことを重視。
 - STI政策は、第3のフレームである「トランスフォーマティブ・イノベーション政策」へと拡張、特に近年では、産業競争力・経済安全保障・地域振興を重視する「産業への転換(industrial turn)」が進み、基礎研究や科学的自律性とのバランスが課題となっている。
- すべての事業等に対して一様に同負荷の評価を求めるのではなく、**国家的目標などに重点化**し、評価疲れを回避。
 - 政策判断上重要な問い、予算配分や制度改善に直結する問い、現場で不確実性が高い問いを優先的に設定し、それに応じて評価手法を選択。