

【平成22年度委託調査報告】

海外政府系研究開発機関における 研究開発評価システムに関する調査・分析

田原 敬一郎

k.tahara@ifeng.or.jp



2012年3月6日(火)

富士ソフト アキバホール

研究開発機関を取り巻く動向

- 研究開発機関の**パフォーマンスに対する評価**の要請
 - 独立行政法人をはじめとする各種科学技術施策は、**投資に見合った成果**が現実に出されているのか。**施策の評価・検証は十分か。**
(行政刷新会議「提言型政策仕分け」)
- 独立行政法人の**制度・組織の見直し**
 - 「高い専門性等を有する研究開発に係る事務・事業を実施し、**公益に資する研究開発成果の最大化**を重要な政策目的とする」成果目標達成法人として位置づけ
 - 研究開発面における国際水準にも即した適切な目標設定・評価の双方に資するため・・・研究開発の専門性を踏まえた**成果重視の実践的な評価**を行う
(「独立行政法人の制度及び組織の見直しの基本方針」2012年1月20日閣議決定)

問題の所在

- 公的な研究開発機関の抱える問題の困難性
 - 科学的価値の創出だけではなく、それが**経済的、社会的な価値の実現にどれだけ役に立つか**のパフォーマンスを問われる。
 - パフォーマンスを発揮するための手段として、**研究開発(への助成)**に限定
 - 多くは**政策レベルの問題**。ただし、多くの評価の矛先は「研究の現場」に向いている。
- **研究開発の不確実性／長期性**に係る問題
 - 挑戦的な課題であればあるほど研究開発の成果が社会的・経済的な価値に転換するまでに時間がかかる。
 - そもそも研究開発の成果が生まれるかどうかは事前には確定できず、本質的な不確実性を持つ。
 - 研究開発の成果が使われる**環境も一定ではなく、変化**。
⇒ **結果としての「業績」を示すだけでは生き残れない。**

諸外国における取り組み

- 諸外国では、評価をどのように用いているのか。どのような工夫を行っているのか。
 - **上位機関に求められる評価の機能**
 - 研究開発機関の自律性を確保した上で、機関のミッションを達成させるために、評価がどのように用いられているか
 - 「**実績の説明**」のための評価
 - 各機関が自律性を保証してもらうために、自らどのような取り組みを行っているか
業績の観点からみた説明責任をどのように果たしているか、そのためにどのようなパフォーマンス指標を設定しているか
 - 「**能力の証明**」としての評価システムの構築
 - ミッションを果たすために、どのような組織活動の改善・見直しの仕組みを具備しているか

諸外国における取り組み

上位機関に求められる評価の機能

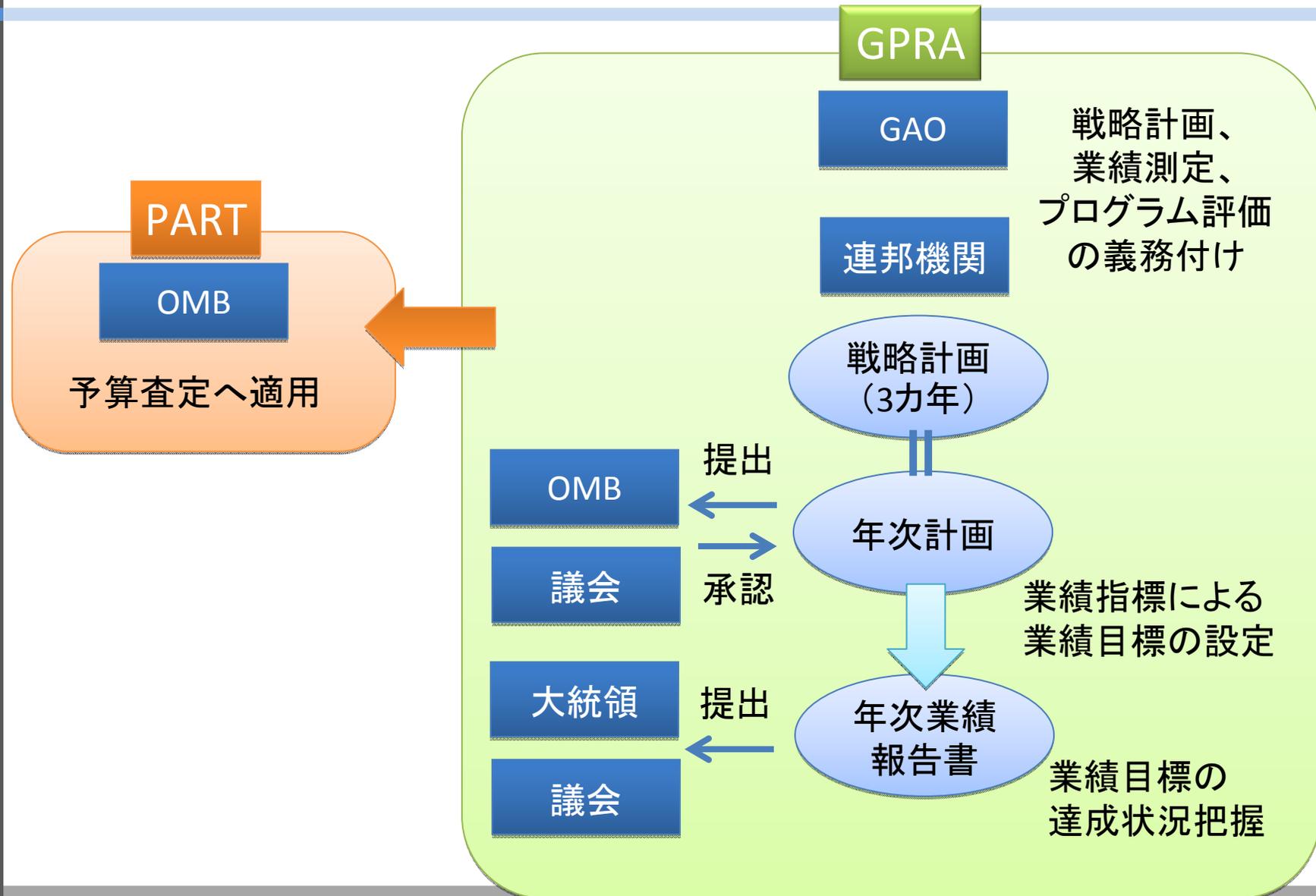
上位機関に求められる評価の機能

- **中央集権と分権の矛盾**をどのように解消しようとしているか
 - － 上位機関側が研究開発機関の多様性を制限しすぎると、環境の変化に対応できない。各機関には自律的に行動する余地が最大限確保されなければならない。
 - － 上位機関側がコントロールしなすぎると、組織は方向を見失い、目標が達成できなくなる。
- 研究開発機関の自律性を確保した上で、機関としてのミッションを達成させるために、評価がどのように用いられているか

事例1: 米国・GPRA／PARTの仕組み

- 米国の政府業績成果法(GPRA)に基づく評価やプログラム評価採点ツール(PART)による評価
 - GPRA: 全ての行政機関に戦略計画の策定とその進行状況を議会へ報告することを義務付ける法律
 - PART: 各機関の実施する事業をプログラムとしてとらえ、プログラムごとにその実施状況や成果を評価し、予算査定に用いる方式
- 主に行政改革等の一環として、研究開発以外の施策や機関にも適用

事例1: 米GPRAとPARTの関係



事例1: 研究開発施策や機関に対する評価の取り扱い

- 研究開発には**長期性にまつわる不確実性**もあり、一律的に規定される評価には適さない場合が多く、共通の枠組みで求められている**定量的指標を用いた評価も困難**である場合が多い
- 研究開発機関や研究開発プログラムの評価にあたっては、評価者である大統領府行政管理予算局(OMB)を納得させることができれば、という前提で、**一般的な枠組みとは異なる評価方法を用いることを認める**
- ただし、プログラムの業績が規準を満たしているかどうかについて、**正確かつ独自に判断できる**ものでなければならない

事例1: GPRAにおける研究開発プログラム向けの評価項目

評価項目	内容
関連性 (Relevance)	<p>プログラムには完全な計画と明確な目標がなければならない。 プログラムでは社会的利益を明確に説明しなければならない。 プログラムでは特別に考慮されるべき政府の具体的優先事項との関連性を示さなければならない。</p> <p>国民のニーズ、科学技術分野のニーズ、及びプログラム「顧客」のニーズとの関連性を、将来を見通した外部審査により評価しなければならない。 国民のニーズ、科学技術分野のニーズ、及びプログラム「顧客」のニーズとの関連性を、過去に遡った外部審査により定期的に評価しなければならない。</p>
品質 (Quality)	<p><u>競争による、メリットを基準にしたプロセス以外の方法で資金を配分するプログラムは、</u>資金提供方法の正当性を説明し、どのように質が保たれているかを証拠により示さなければならない。 プログラムの品質は、過去にさかのぼった専門家の審査により定期的に評価しなければならない。</p>
業績 (Performance)	<p>プログラムは、関連性のある投入データを毎年追跡し報告する必要がある。 プログラムは、アウトプットとアウトカムの適切な測定法、スケジュール、及び決定ポイントを明確にしなければならない。 プログラムの業績は、毎年過去にさかのぼって評価しなければならない。</p>

事例1: PARTによるプログラム評価

- 4つのセクションごとに、評価のための具体的情報を求める質問がいくつか行われ、各セクションが加重されて、最終評点化。予算査定に適用。
- PARTは各連邦機関による**自己評価の評価(メタ評価)的な側面**もある。

セクション	ウェイト	説明
プログラムの目的と設計	20	プログラムの目的と設計が明確で健全であるかどうかを評価
戦略計画	10	プログラムの有効な長期目標と年間成績数値を定めているかどうかを評価
プログラム・マネジメント	20	機関の財務監視やプログラム改善努力などのプログラム・マネジメントを採点
プログラムの結果／説明責任	50	戦略計画のセクションで審査された数値や目標に基づいたプログラム成績を、他の評価方法で採点

諸外国における取り組み

「実績の説明」のための評価の 実施

「実績の説明」のための評価

- 各機関が上位機関（ひいては利害関係者、国民）に対し、**業績の観点からみた説明責任**をどのように果たしているか、そのためにどのような**パフォーマンス指標**を設定しているか
- 各機関が自律性を保証してもらうために自らどのような取り組みを行っているか。
- 公的研究開発機関が実施、支援する研究開発は、各国とも民間にまかせていては実現が難しいもの（市場価値では測れないものや、効果の発現に時間のかかるものなど）を対象にしていることもあり、上位機関から求められる評価とは別に、各機関が独自に自らの業績をアピールするための活動を展開していることも多い。

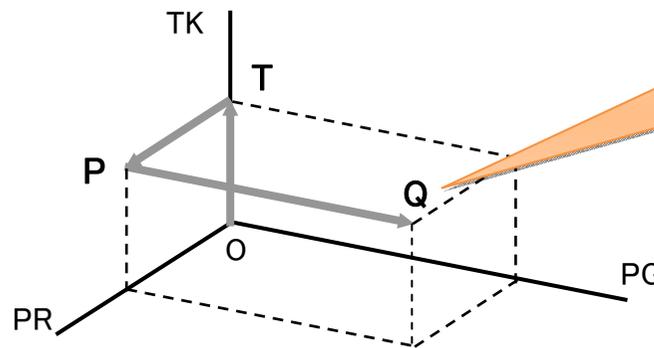
事例2: ATPにおける評価

- 国立標準・技術研究所NISTによる先端技術プログラム(ATP)
 - － 民間企業へ直接研究開発投資を行う、米国においては例外的なプログラム
 - － 目的は、米国に広範囲な経済的利益をもたらす可能性の高い技術を開発するためのハイリスクな科学技術研究に関して米国産業界に協力すること
- ATPにおける評価研究の振興
 - － 存在意義を巡って激しい政治的な争点に。その必然的な帰結として、**自らの業績を説明するための評価**に精力的に取り組む
 - － 内部スタッフに経済性分析の専門家を抱えるとともに、学会、コンサルタント企業のエコノミスト、評価専門家を活用し、**成果の尺度の質や信頼性向上に寄与する検討などを実施**(評価計画、評価モデル化、評価データベース開発、評価調査、評価のケース・スタディ、評価の統計的経済的分析など)
 - － 現在ATPは廃止となったが、そこで得られた評価研究の蓄積、専門人材は現在でもNISTにおける取り組みに活かされる

事例2: ATPにおける評価の三次元

国家的規模での科学的・技術的知識の
増進への寄与の程度

科学技術知識の進歩



Q点の極大化が目的。

最新技術を追求すると同時に、他の2つを満足させるようなプロジェクトを選定。

※ATPが存在しなかった場合に
想定されるレベルとの比較

助成受給者にとってのリターン

研究の結果得られた技術の
実用化可能性の程度等

J.-C. Spender

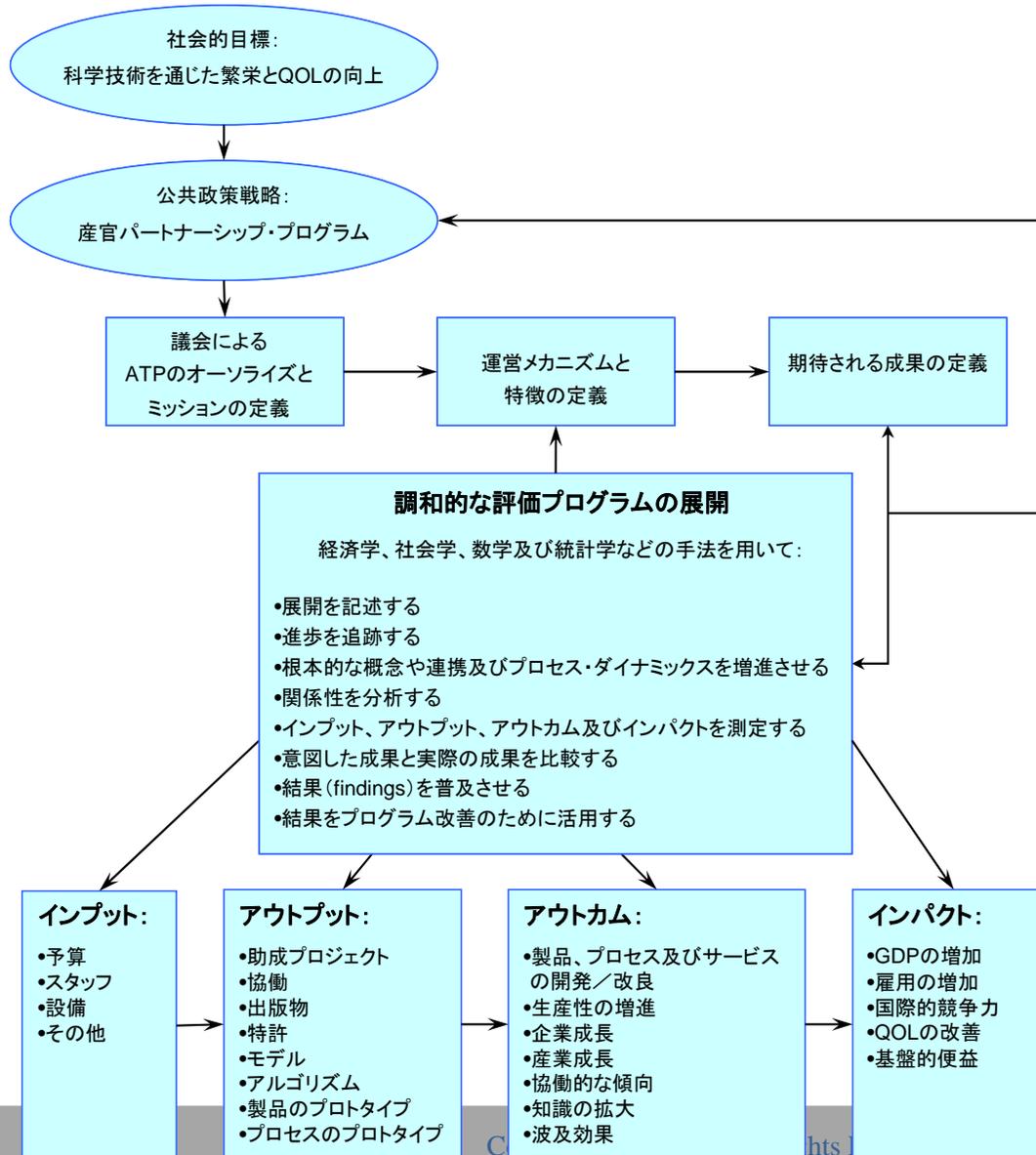
「スピルオーバー」
第三者へのリターン

プロジェクト実施企業に占有されない
全国的な波及効果の程度

機関としての評価の能力を向上させる様々な試み

⇒プロジェクトの進展、段階が変化する時点、影響の拡がるパターンや速度、それらの背景にある因果関係等、プログラムの成果を決める要因の相互関係には未解明な部分が多い、という前提

【参考】ATPにおける評価のロジックモデル



Ruegg & Feller (2003)

【参考】ATPで用いられる評価手法の概要

手法	概説	使用例
分析／概念モデリング (Analytical/Conceptual modeling)	プログラムやプロジェクト、現象をよりよく理解するために、基調をなす概念を調査したりモデルを開発したりする	スピルオーバー効果が起こる経路を概念的に記述する
調査(Survey)	統計的分析のために、一連のそろった質問を複数の関係者に尋ねる	どれくらい多くの企業が自分の開発した技術を他にライセンスしているかを割り出す
ケーススタディ －記述的	プログラムやプロジェクト、テクノロジー、設備を掘り下げて調査する	どれくらいの特定のジョイントベンチャーが形成されたのか、どのように協力しているのか、成功の理由や不足の理由を列挙する
ケーススタディ －経済的評価	記述的なケーススタディに経済的効果の定量化を加える。例えば、費用便益分析を用いる	プロジェクトの便益がコストを上回っているかどうか、どれくらい上回っているかを見積もる
計量経済学及び 統計的分析	経済と社会的現象のつながりを分析するために、統計や数理経済学、計量経済学を用いる	公的資金配分がどれくらい民間の研究資金配分に影響を及ぼしているかを定める
計量社会学及び 社会ネットワーク分析	社会的/組織的振る舞い及び関連する経済的アウトカムの理解を増進するために結びつきの構造を特定し調査する	結果として生じる知識の拡散が増大できるようにするには、どのようにプロジェクトを構築すればよいかを学ぶ
ビブリオメトリクス －総数(counts)	研究のアウトプットの量を追跡する	プログラムが生成し、研究に使われる資金あたりの出版数を求める
ビブリオメトリクス－引用	他人が出版物や特許を引用している頻度を評価し誰が引用をしているかに言及する	プロジェクトの出版物や特許の普及の範囲やパターンを学ぶ
ビブリオメトリクス －内容分析	共語分析やデータベース・トモグラフィー、テキストデータ・マイニング、視覚化技術を用いてテキストから情報を取る	プロジェクトの貢献度や技術の進展に関連のあるタイミングを特定する
歴史的トレーシング	研究から将来のアウトカムを追跡したり、アウトカムから貢献している開発を追跡したりする	公的研究プロジェクトと重要な後の発生のつながりを特定する
専門家の判断	評価をするために、見識ある判断を用いる	新しい技術の十中八九最初の使用と仮定する

諸外国における取り組み

「能力の証明」としての評価システムの構築

「能力の証明」としての評価システムの構築

- 説明責任をよりよく果たすには、**継続的に成果を生み出せる能力を各機関が保持していることを示す必要がある**
 - 結果のみでは「業績」への寄与が分からない
 - 組織を取り巻く環境は絶えず複雑に変化しており、過去の成功したやり方が将来にわたって機能するかを保証してくれない
- 組織をとりまく環境の変化に対応するために、各機関がどのような**学習の仕組み**を持っているか
 - 業績を生み出す仕組みとしての制度やプログラムに対し、改善・見直しの契機を与える評価システム
 - 評価システム自体も状況の変化に応じて見直される必要

事例3: 英国・Alvey Programmeにおける試み①

- Alvey Programme
 - 1983～1987年、UKにおいて実施の情報通信分野の研究開発事業
- リアルタイム評価制度の採用
 - ①プログラム開始直後から、研究開発推進側とは別の外部評価者が定期的に主題毎の報告書を管理者にフィードバック
 - 知財に係る問題等、多くはプログラムに付随する複雑な問題
 - プログラムの抱える問題点を専門性を持った外部の第三者が常時把握、プログラム・マネジメントの改善に活用。
 - ②プログラム終了の2年後に最終評価報告書を提出
 - アウトカムを含めた実績や、プログラム実施期間中には顕在化しない長期的視点からの問題点を把握するとともに、次の取り組みのための教訓を導出。(単なるアカウントビリティのための評価ではない)
 - プログラムを管理する担当課の所掌範囲を超える問題の存在。プログラム・マネジメントの改善だけでは追いつかない問題。

Georghiou, L. & D. Roessner (2000), "Evaluating technology programs: tools and methods," Research Policy 29: 657-678.

事例3: 英国・Alvey Programmeにおける試み②

- 評価結果
 - 技術的目的の達成と産学間の共同研究文化の育成には成功したが、IT産業の再活性化という**商業目的に対しては不成功**(UKのIT産業の市場におけるポジションが実質的に後退)
- 得られた教訓
 - ①**設定された目的自体に問題**があった、つまり総合的に対応が本来必要な産業政策の代替制度として、研究開発制度を設定した。
 - 「発見－発明－[死の谷(1)]－開発－[死の谷(2)]－事業化－[ダーウインの海]－産業化」の各フェーズを考慮し、研究開発投資、政府調達、税制、規制等の政策手段を組み合わせる必要があった。
 - ②このようなプログラムは多様で、時に矛盾する目的を持つが、評価のために行う調査に適した形で表現されていなかった。**評価可能な形で目的を再構築する必要**があった。
- 評価の経験に基づく評価システムの改善
 - **ROAME(or ROAMEF)システム**の導入:プログラムの事前評価の評価項目として次の項目が詳細に記述されていることを求める。
 - プログラム設定の理由・位置づけ(rationale)、検証可能な目的(objectives)、プロジェクトの事前評価(appraisal)、途上評価(monitring)、事後評価(evaluation)のための計画、それに加えてプロジェクトの評価結果のフィードバック(feedback)手順の設定。

事例4: NSFによる外部評価の活用と実質化

- 全米科学財団NSFでは、委嘱審査委員会(COV)等の外部評価を活用することにより、助成の決定に用いられる**評価・勧告の質を維持するシステム**を具備
 - COV: 大学、産業界、政府、公共部門からの専門家で構成される独立の委員会
 - 1) **評価の過程の健全性と効率性**、2) NSFの投資の結果の質その他を含めた効果の2つを検討
- 勧告を形骸化させない仕組み
 - 監査室がCOVによる勧告の活用実態についての監査を行い、その利用改善に向けた勧告を実施
 - COVの会合の運営方法や、報告書の活用方法に関して改善の余地が無いのか、NSFがCOVによる報告書をNSFの業績報告へのインプットとして適切に活用しているか、等

※国立衛生研究院NIHでも、全米科学アカデミーに対し、新規研究センター設立の可否を決定する**基準と手続**の検討を依頼

事例5:オランダイノベーション機構に おける戦略策定と評価

- オランダイノベーション機構 (NL Agency)
 - 2010年1月に政策実施機関であるSenterNovemと経済省傘下の外国貿易庁、オランダ特許庁との統合により誕生
 - 研究助成の他、技術、エネルギー、環境、輸出、国際協力の分野における助言、ネットワーク(提携相手のサーチ)、情報、金融の支援
- NL Agencyにおける戦略策定
 - 評価枠組みの策定とセットで戦略を策定
 - 「ロジカル・フレームワーク・アプローチ」等の活用と人材養成
 - 各部門においてバランスト・スコア・カード(BSC)に通じたスタッフと政策アドバイザーが戦略の素案を作成
 - BSC: 戦略・ビジョンを4つの視点(財務の視点・顧客の視点・業務プロセスの視点・学習と成長の視点)で分類し、その企業の持つ戦略やビジョンと連鎖された財務的指標、及び非財務的指標を設定する方法
 - 多くのスタッフに一定期間のトレーニングを受講させ、戦略策定ならびに評価のフレームワークのための資料作成能力を養成
 - 内部横断的なプラットフォームを公式・非公式に展開、ストラテジストや評価の経験者も多く参画

事例6: 英国・NESTAの試み①

- NESTA(国立科学・技術・芸術基金)
 - 1998年、全国くじの基金をもとに設置。現在、300百万ポンド(350億円相当)超の基金を有し、その運用益や宝くじ収益等によりその活動資金を賄う。**政府からの資金供与は一切受けておらず、高い独立性を保持。**
 - ミッションは、UKをよりイノベティブなものにすること
 - もともとアーリー・ステージの企業への投資および融資を行う機能を有していたが、近年では、**イノベーションの新しいモデルを構築し、それを試すためにその資源を活用する種々の活動を実施**(伊地知2009)。
 - プロジェクトの最後における単純な定量的な評価ではなく、**それぞれのプログラムの性質に応じた評価を実施**
 - **政府が「どのようにしてイノベーションを支援するか」が重要なポイント。** NESTAは、プロジェクトへの投資を通じて、政府がビジネスやコミュニティをより支援するためにはどうすればよいかを学習。
 - 投資を行うチームや助成プログラムを運営するチームのほか、**イノベーション政策の開発を担うユニット(政策及び研究ユニット)**を設置。それぞれのチームが有機的に連携。
 - 分析や研究を行う15名のフルタイム研究者が在籍。専門は、政策研究、ビジネス投資、高齢化社会、気候変動など多岐にわたる。外部専門家の知見も活用

事例6: 英国・NESTAの試み②

- ラーニング・パートナー制度
 - 外部にしながらプロジェクトと協働・連携するラーニングパートナー(LP)を常に確保し、機能させる仕組みを導入(Neighbourhood Challenge プログラム等)
 - 信頼できるデータをもとに語る経済専門家や、社会的イノベーションをよく理解している社会学者などで構成
 - LPとの連携を継続させることは、ある種の支援的評価を受け続けていることに相当。プロジェクト内に不足する専門性を効果的に補完する役割
 - 「イノベーションは特殊な方向にエッジを利かせるもの」との認識の下、個人や組織の専門性としても分野ごとに分類をすることはせず、人々が領域横断的に交流できる空間を創出することを重要視。そのため、多くのイベントを実施、また、それを期待して多様な人々が集まる好循環を生み出す
 - 影響力のある人物を探しては、関与者との定期的な会合を開催。連携する可能性のある様々な人々と、長期的な視点で良い関係を構築

事例6: 英国・NESTAの試み③

- 「Big Green Challenge」におけるイノベーション・モデルの創出
 - 地域コミュニティにおけるCO2エミッションの大幅削減のためのアプローチを開発、実施しようとする非営利グループや組織のインセンティブを高めることが目的。
 - 1年間でどれほどのCO2削減を実現したかを競わせ、その多寡によって総額100万ポンドの賞金を授与
 - ファイナリストには1年間の活動資金として2万ポンドを供与。
 - 350のエントリーのうち10団体をファイナリストとして選定、1年間で成果を競い合わせる。
- 「社会イノベーションを促進するメカニズムについて、実践ベースの研究開発プログラムを通じたシステムティックな方法で調査を行う必要がある」との問題意識

【参考】BGCのスケジュール

期間	活動内容
①2007年11月～12月	Big Green Challengeの広報活動(地方でのイベント)
②2008年1月	ウェブサイトからの第1段階(Stage 1)申請受付開始
③2008年3月3日 (午後11:59)	第1段階申請の締め切り
④2008年早春	第1段階受賞者の決定と第2段階申請への招待
⑤2008年晩春	第2段階申請(詳細計画提出)の締め切り
⑥2008年夏	第2段階申請の受賞者(ファイナリスト)の決定 ※第2段階申請のトップ10の申請者をファイナリストとして選定
⑦2008年10月～ 2009年10月	第3段階:ファイナリストによる活動の実施と、年間を通じた活動の測定と監査
⑧2009年11月	最終受賞者の決定

事例6: 英国・NESTAの試み④

- 活動水準(業績)の測定と審査
 - 気候変動問題、コミュニティ活動、イノベーションといった幅広い領域の専門家から構成されるパネルで審査
 - 決定にあたっては、ファイナリストによって達成された炭素削減量を可能な限り正確に測定するためにNESTA内に設置された炭素監査チーム(carbon audit team (Cred))による報告書を参照。

事例6: 英国・NESTAの試み⑤

- 単なる賞金の授与にとどまることなく、競争のプロセスを注意深く観察、分析することで、**社会問題の解決に必要な知見と普及のモデルを見い出し、国への提言に結びつける**など、それを広く社会に還元しようとする明確な意図を持つ。
 - これらの取り組みを通じて得られた知見は、「政策及び研究ユニット」にもフィードバック、イノベーション研究を質、量の両側面から豊富化することに寄与
 - 政府との関係構築の取り組み
 - NESTAによる政策提言や出版の成果を実社会に反映していくために、国・地方政府両方の政治家との議論を実施。NESTAが政治家に依頼し、毎週に1回程度、NESTAのオフィスでミーティングを実施。
 - 政治家との議論は、NESTAにとっても、政治、現実に政策がどのように作られ、動いていくかを理解する良い機会と捉える。
- ⇒ **エビデンス・ベースで提言を行うこと、また、単なる提言でとどめず、実現のための具体的行動を起こす**

事例6: 英国・NESTAの試み⑥

- 社会問題解決のためには関与者のインセンティブを高める工夫が必須
 - プロジェクトの採択時ではなく実施過程に競争原理を導入する手法や用途を限定しない賞金という助成形態は、実施者の意欲を高めると同時に、プログラムやプロジェクトの社会的認知度を高める、そのことによりさらに質のよい提案が集まる、という好循環をもたらしている。
- 強いモデル志向
 - プロジェクトへの支援を通じて得られた教訓を次のプログラム設計や見直しに活かそうという意図だけではなく、連合王国をよりイノベーティブにするために貢献しようとするものであり、自らの機関の存在意義を高める上で寄与。
 - 独立性があるからこそ得られる内容的・手続き的信頼性
- 内部に高度な知見とスキルを要する人材を抱える

総括

総括

- (1) 組織が直面する環境は多様かつ複雑。価値実現や組織生存のためには**他組織との連携(しようとする意思がみえること)**が重要
 - “上”から降りてくる評価の枠組みが正しいとは限らない
 - 求められる成果に対し、組織の持つ**手段は限定的**
 - 組織による努力だけでは解決できない問題に対する**提言機能としての評価**が決定的に重要
- (2) 成果そのものが実現できたかという評価より、**成果をうみだすための仕組みや成果を価値実現にまでつなげる仕組みを評価**することが必要
 - 実際に研究開発を行うのは研究者であるが、組織に求められるのは成果をうみだすための環境を整備し、それを価値に転換する仕組みを用意すること
 - **被評価者は**(研究者ではなく)**「組織」**であり、研究者の成果を積み上げても組織としての能力を示したことにならない
- (3) 各機関は**インテリジェンス機能**を持つべき
 - 自律性、独立性を与えられる機関の側にとっては、その裏付けとして、期待される専門性を十分に発揮し、**応答していく責任**と、公的資金を扱うことに対する**説明責任**が求められる。そのための**専門性をいかに確保**するか