平成 18 年度文部科学省委託調査

「研究開発評価の質の向上のための調査・分析」

報告書

平成 19 年 3 月

財団法人政策科学研究所

本報告書は、文部科学省の研究開発評価推進調査委託費による委託業務として、財団法人政策科学研究所が実施した平成 18 年度「研究開発評価の質の向上のための調査・分析」の成果を取りまとめたものです。

従って、本報告書の著作権は、文部科学省に帰属しており、本報告書の全部又は一部の無断複製等の行為は、法律で認められたときを除き、著作権の侵害にあたるので、これらの利用行為を行うときは、文部科学省の承認手続きが必要です。

はじめに

わが国における研究開発評価は第 1 期科学技術基本計画において制度化を要請され、続く「国の研究開発全般に共通する評価の実施方法の在り方についての大綱的指針」によってその枠組みが示されることとなった。

研究開発評価は研究開発を効果的・効率的に推進し、その成果を国民に還元していく上で、重要な役割を担うものであり、第 3 期科学技術基本計画においても、評価システムの一層の発展が重要とされている。

現在、多種多様な研究開発が進められている中、研究開発の特性に応じた評価手法の導入に役立つ研究開発評価テキストの作成等、様々な手段を講じて、研究開発評価の質の向上を図ると同時に、評価を担う人材を養成、確保していく必要がある。

本調査報告書は、文部科学省委託調査「研究開発評価の質の向上のための調査・分析」の成果をとりまとめたものである。本報告書に付属する「研究開発評価テキスト（基礎編）」ならびに「研究開発評価テキスト（応用編）」の作成にあたっては、委員諸氏の研究開発評価の現場からの経験を踏まえ、より実務的な観点から編纂している。本テキストが今後のわが国の研究開発評価の向上や評価人材の拡充に資することを願ってやまない。

平成 19 年 3 月

財団法人政策科学研究所

平成 18 年度文部科学省委託調査

「研究開発評価の質の向上のための調査・分析」

報告書目次

第Ⅰ部　業務報告書　1

　1　本調査の目的　3

　2　本調査の内容　3

　3　本調査の実施体制　9

　参考情報　11

　資料　ワークショップ配付資料　77

第Ⅱ部　成果報告書　97

　1　研究開発評価研修プログラムの概要　99

　2　研究開発評価研修プログラム（基礎編）　教材案　107

　3　研究開発評価研修プログラム（応用編）　教材案　157

※なお、教材案の内部では頁番号を独自にふり直している。

第Ⅰ部　業務報告書

１　本調査の目的

研究開発評価は、研究開発を効果的・効率的に推進し、その成果を国民に還元していく上で、重要な役割を担うものであり、第 3 期科学技術基本計画においても、評価システムの一層の発展が重要とされている。

現在、多種多様な研究開発が進められている中、研究開発の特性に応じた評価手法の導入に役立つ研究開発評価ハンドブックの作成等、様々な手段を講じて、研究開発評価の質の向上を図ると同時に、評価を担う人材を養成、確保していく必要がある。

本調査では、行政機関等における研究開発評価担当者、評価を受ける側の事業推進部署の担当者が学ぶべきトピックスを整理し、体系化することで、より本質的、かつ実効的な研究開発評価への理解を促進し、高度な研究開発評価人材の養成を図ることを目的とする。

２　本調査の内容

本調査では以下の業務を実施した。

（１）研究開発評価研修プログラムのための予備的調査

- 研究開発評価研修プログラムに資する文献情報等の収集
- 研究開発評価を実施している国内機関に対するヒアリング調査
- 研究開発評価に関する海外有識者に対するヒアリング調査

（２）研究開発評価研修プログラムならびに教材開発

- 専門委員会の設置
- ワークショップの開催
- 研究開発評価研修プログラムおよび教材案の作成

２－１　研究開発評価研修プログラムに資する文献情報等の収集

研究開発評価は実施する側の評価目的、経営方針、評価結果の活用の仕方によって種々異なるかたちで展開されている。我が国の研究開発評価に包括的に資する参考文献や参考情報というものは存在せず、個別のトピックスにおいて参照すべき文献類が存在するのみである。

特に、研究開発評価における分析面での文献情報が充実している中、本調査ではなるべく実務的観点からとりまとめている文献等に焦点をしぼり、研修プログラムの参考情報として情報源情報をとりまとめた。

本調査項目の成果は第Ｉ部業務報告書の章末に参考情報としてとりまとめている。

２－２　研究開発評価を実施している国内機関に対するヒアリング調査

本調査では実務的観点からの研修プログラムならびに教材開発を作成するために、研究開発評価実施機関における担当者からの意見収集を行い、教材の記述内容等に反映させている。具体的には、研究開発評価に関連する担当者が集まる会議、セミナー等への参加機会を捉え、現場の声として教材開発のためのアドバイスをいただいた。

主なヒアリングの対象は以下の通りである。

- 独立行政法人科学技術振興機構（JST）
  戦略的創造事業本部，産学連携事業本部開発部，原子力業務室
- 独立行政法人理化学研究所（RIKEN）
- 独立行政法人宇宙航空研究開発機構（JAXA）
  産学官連携部
- 経済産業省
  産業技術環境局技術評価調査課
- 独立行政法人産業技術総合研究所（AIST）
- 独立行政法人新エネルギー・産業技術研究開発機構（NEDO）
  研究評価部，企画調整部

２－３　研究開発評価に関する海外有識者に対するヒアリング調査

本調査では海外における科学技術政策評価、プログラム評価の新たな試みに関する情報収集とネットワーク構築のために、平成 18 年度に実施された主要な国際会議に参加し、有識者との接触機会を得た。

PRIME 2007 Annual Conference

時期：2007 年 1 月 29 日～2 月 1 日

場所：PISA（イタリア）Centro Studi della Cassa di Risparmio

プログラム概要 URL：

http://www.prime-noe.org/index-restricted.php?project=prime&locale=en&level1=menu1_prime_1b8057d059a36720_1&level2=4&doc=2007_Annual%20_Conference&page=2

テーマ：”POLICIES FOR RESEARCH AND INNOVATION IN THE MOVE TOWARDS THE EUROPEAN RESEARCH AREA”

PIRME は欧州における科学技術政策評価のためのネットワーク・エクセレンスであり、その年次大会には欧州の第一線級の評価研究者が集う。参加した年次大会の主なトピックスは以下の通り。

- European universities at a crossroad
- Scientific production and technological performance: beyond the European paradox syndrome
- Beyond the paranoia of rankings. Towards theory-based positioning indicators for universities and research
- New loci for policy making in the European Research Area
- New rationales for policy making in science and technology
- Linking research policy to innovation policy. Conceptual and implementation challenges
- The role of Networks of Excellence in the European research landscape; Perspectives and future agenda

【主な接見者とヒアリング内容】

- Philippe Laredo 氏, ENPC and University of Manchester, PRIME Coordinator
  FP7 における PRIME の動向
- Irwin Feller 氏, Professor Emeritus of Economics, Penn State University, USA
  米国における科学技術政策評価の動向
- Luke Georghiou 氏, PREST, University of Manchester
  欧州イノベーション政策の動向
- Stefan Kuhlman 氏, Twente University
  蘭国 MOT 教育の状況
- Andrea Bonaccrosi 氏, University of Pisa
  伊国におけるプログラム評価の現状

AAAS 2007 Annual Meetings

時期：2007 年 2 月 15 日～2 月 19 日

場所：San Francisco（米）Hilton San Francisco

プログラム概要 URL：

http://www.aaas.org/meetings/Annual_Meeting/01_REG/Register.shtml

テーマ："Science and Technology for Sustainable Well-Being"

AAAS（全米科学振興協会）は 1848 年設立の科学技術を支援する非営利団体で、有名な科学雑誌「Science」を発行している。その年次大会は米国のみならず、世界中の科学者が参画し、活発な意見交換が行われている。2007 年の年次大会では 35 のセッションが設けられた。

【主な接見者とヒアリング内容】

- Albert Taich 氏, Director, Science and Policy Programs AAAS, USA
  米国における科学技術政策とイノベーション政策の位置づけ，”Science of Science Policy”と評価の関係性
- Bill Valdez 氏, Director of Planning and Analysis in the Office of Science since DOE, USA
  米国における競争的研究プログラムの特徴と評価の精神，NSF のキーパーソン（Kaye G. Husbands Fealing 女史）の紹介
- Kaye G. Husbands Fealing 女史, Science Advisor for Science of Science Policy, National Science Foundation (NSF), USA
  NSF のプログラムマネジメント、採択評価のマネジメント
- Gerald Hane 氏, Training Program for Research and Development Evaluation, USA
  米国における科学技術予算の決定プロセスとプログラム評価の位置づけ

２－４　専門委員会の設置

研究開発評価研修プログラムのあり方ならびに教材開発について、評価経験者ならびに行政機関・大学等の評価業務担当者から構成される専門委員会を設置した。委員の構成は以下の通りである。

主査　平澤　泠　（東京大学名誉教授）
伊地知　寛博（一橋大学　助教授）
勝本　雅和　（京都工芸繊維大学　助教授）
菊池　純一　（青山学院大学　教授）
調　麻佐志　（東京農工大　助教授）
嶌田　敏行　（茨城大学　助手）
隅蔵　康一　（政策研究大学院大学　助教授）
林　隆之　（独立行政法人大学評価・学位授与機構　助教授）
藤森　英俊　（独立行政法人海洋研究開発機構　チーフ）
柳澤　和章　（独立行政法人日本原子力研究開発機構　研究主幹）
弓取　修二　（独立行政法人新エネルギー・産業技術研究開発機構　主任研究員）
吉田　秀紀　（独立行政法人科学技術振興機構　副調査役）

事務局

川島　啓　　財団法人政策科学研究所　主任研究員

田原　敬一郎　　財団法人政策科学研究所　研究員

野呂　高樹　財団法人政策科学研究所　研究員

上記専門委員会を委託調査期間中に５回実施した。開催実績については以下の通りである。

【専門委員会開催実績】

第１回専門委員会

日時：2006 年 12 月 7 日（木）18:00～21:00

場所：財団法人政策科学研究所会議室

議事次第：

１．委員会主査からのご挨拶（平澤主査）

２．委託元からのご挨拶（文部科学省評価推進室　後藤 裕　室長）

３．委員紹介（自己紹介形式）

４．委託調査の主旨と研修プログラムの全体像に関する検討（事務局　川島）

５．討議

６．スケジュール調整

第２回専門委員会

日時：2006 年 12 月 27 日（水）18:00～21:00

場所：財団法人政策科学研究所会議室

議事次第：

１．テキスト案に関する検討（事務局　川島）

１）「我が国の研究開発評価の制度」

２）「評価の定式化」

２．討議

３．スケジュール調整

第３回専門委員会

日時：2007 年 1 月 17 日（水）18:00～21:00

場所：GRS ビジネスセンター八重洲 annex　会議室２

議事次第：

１．テキスト案に関する検討（事務局　川島）

１）研究開発評価の体系

２）「研究開発マネジメントと評価」

２．討議

３．スケジュール調整

第４回専門委員会

日時：2007 年 2 月 6 日（火）18:00～21:00

場所：財団法人政策科学研究所会議室

議事次第：

１．研修テキストの枠組みと説明項目に関する検討１（事務局　川島）

１）「研究開発評価研修用テキスト基礎編の全体像」

２）「研究開発評価研修用テキスト応用編の全体像」

２．討議

３．出張報告（平澤主査）

PRIME Annual Conferrence，2007 年 1 月 29 日～2 月 1 日，イタリア・ピサ

４．スケジュール調整

第５回専門委員会

日時：2007 年 2 月 26 日（火）18:00～21:00

場所：財団法人政策科学研究所会議室

議事次第：

議事次第：

１．研修テキストの枠組みと説明項目に関する検討２（事務局　川島）

１）「研究開発評価研修用テキスト基礎編の全体像」

２）「研究開発評価研修用テキスト応用編の全体像」

２．討議

３．ワークショップの開催について（事務局　川島）

４．出張報告（平澤主査）

AAAS Annual Meetings, 2007 年 2 月 15 日～2 月 19 日，サンフランシスコ

２－５　ワークショップの開催

専門委員会で検討した研究開発評価研修用テキスト基礎編について、外部の意見を収集するためにワークショップを開催した。

【開催実績】

日時：2007 年 3 月 28 日（水）18:00～20:00

場所：丸ビルホール＆コンファレンススクエア　Room 1
件名：研究開発評価研修ワークショップ基礎編
内容：

１．プログラムの獲得目標
２．テキスト案の全体像
３．基礎編の演習と解説
４．質疑応答＆全体討論

配付資料：別添参考資料２に付す。

２－６　研究開発評価研修プログラムおよび教材案の作成

本調査で得られた各種情報、専門委員会での議論を参考に、研究開発評価研修プログラムの案と研修用テキストの案を作成した。これらは本報告書「第Ⅱ部成果報告書」にとりまとめている。

３．本調査の実施体制

本委託調査の実施体制を以下に記す。

（1）管理体制

（2）業務実施体制

（3）担当研究者リスト

氏　　名	所　属　・　役　職（職　名）
千葉　勝	研究部　研究部長・主席研究員
川島　啓（○）	研究部　主任研究員・評価研究センター主幹
小松　真弓	研究部　主任研究員
藤澤　姿能子	研究部　主任研究員
田原　敬一郎	研究部　研究員
野呂　高樹	研究部　研究員
宮腰　佐和	研究部　研究員

※○印は委託調査主務担当者

参考情報

参考情報1．研究開発評価関連学協会とネットワーク
参考情報2．研究開発評価関連国際会議等の開催状況
参考情報3．研究開発評価関連研修コース・大学院専攻等
参考情報4．研究開発評価関連学協会誌
参考情報5．研究開発評価関連主要文献・基本資料

参考情報1. 研究開発評価関連学協会とネットワーク

研究開発（研究技術、研究イノベーション、科学技術等を含む広義の意味として用いる。以下同様）評価に関連する学協会やネットワークは、評価人材を育成していくうえで重要な機能を担っている。これらの活動を通じて、個人ないし機関レベルでの方法論やノウハウなどの情報を共有、深化させることにより、それぞれの評価機関の活動もまた高度なものに発展していくことを可能にしている。これらはプラクティショナーにとっての主要なスキルアップの場であると同時に、アナリスト等のプロフェッショナルによる成果の発表の場でもあり、両者の情報交換を図る機能を提供している。

学協会について、研究開発評価のみを主題にした学会は存在しないが、評価全般を対象とする学協会の一部門として扱われるか、研究開発関連政策課題全体を対象とする学会の一部の活動として扱われる場合はある。

実務的観点から特に重要なのは、永続的な組織として定着する前のネットワークである。評価ネットワークは特定の実務的課題に取り組むための臨時の組織であり、その活動としては、会合やインターネットを通じ当該目的に適した濃密な情報交換が行われている。多くの場合、ノードとコアメンバーを定める他はオープンになっている。

以下では、こうした学協会とネットワークについて、現在活動中のもの、現在は活動を停止しているがこれまで重要な機能を果たしてきたもの、その他、の 3 類型に分けてまとめた。

なお、組織によっては、自ら研修コースの提供を行うなど、単なるネットワークを超えた高密度な活動を展開しているところもあるが、その種の情報に関しては、参考情報 3 でまとめた。

1－1　現在活動中の学協会とネットワーク

1. WREN

http://www.wren-network.net/

WREN (Washington Research Evaluation Network)は、ブッシュ政権が導入した PART(Program Assessment Rating Tool)への対応の必要性から行政組織の評価担当者を中心にして、科学技術に関わる評価マネジメントを改善するための新たなアプローチや方法論を開発することを目的として 2003 年に結成された。実務的評価活動の情報交換に特色がある。ワシントン DC を中心とした連邦政府の研究開発評価コ

ミュニティの実務的フォーラムとして機能している他、「より大きな評価ネットワークの一部である」とウェブサイトに明記されているように、現在ではカナダ、EU、韓国等との国際的なネットワークの強化にも努めている。設立の動機は 2001 年 8 月に行政管理予算局 OMB によって導入されたプログラム評価法への対処方策の検討にあったと言われている。

設立にあたっては、PART の最初の適応事例とされた米国エネルギー省科学局（DOE-SC）が資金提供を行ったが、現在、DOE-SC、ジョージ・ワシントン大学及び AdSTM 社（Advanced Systems Technology and Management, Inc.）の３者が共同スポンサーとなっている。中心人物は、運営委員会メンバーでもある DOE-SC の William. J. Valdez 氏、DOE 傘下のサンディア国立研究所（Sandia National Laboratories）の Dr. Gretchen B. Jordan 氏、ジョージ・ワシントン大学の Prof. Nicholas S. Vonortas 氏などであり、連邦政府の評価実務家だけでなく、現在では大学等の研究者やシンクタンク等の専門家など多様なメンバーが活動に参加している。

WREN は、その活動の焦点として、科学技術への継続的な政府投資に対する全般的な正当性の追求、研究開発のアウトカムをシステマティックに把握する方法、科学技術に対するシステム（特に国家システム）レベルの分析と評価への挑戦等を挙げている。最近では、戦略策定に利用できる評価を行うための新たな方法論を模索している様子がうかがえる[1]。

具体的な活動内容としては、ワシントンの評価関係者を集めて 1 時間半程度の飲食物持参のランチミーティング（Brown Bag Lunch Meeting）を開催するなど個人レベルでの情報交換とスキルアップを行っているほか、2003 年から毎年大規模なワークショップを国際的に展開している[2]。その活動は欧米にとどまらず、直近では韓国 KISTEP（科学技術評価・企画院）との共催で「公的研究開発の評価のための国家モデル－ベストプラクティスと協働の機会を求めて」（2005 KISTEP-WREN Workshop）と題した国際シンポジウムをソウルで 2 日間にわたって開催するなど国際的な連携を強める傾向にある[3]。

なお、WREN のウェブサイトでは、これらのイベント情報をはじめ、研究開発評価に関する国内動向、求人情報等が随時更新されている。また、国際会議や主要学会での WREN によるオーガナイズド・セッション等のプレゼンテーション資料や論文、電子書籍などもダウンロード可能である。

---

[1] セントルイスで開催された 2006 年 AAAS 年次大会では、Valdez 氏がモデレーターを務め WREN の主要メンバーが関わった「科学技術におけるラディカルなイノベーション－マネジメントと測定」と題するセッションが行われた。これは、従来型の評価がどうしても後追いになってしまう、つまり、評価のために利用可能なデータや指標にはタイムラグがあり、戦略策定と結びつけることには自ずと限界がある、という問題意識から出発したものであった。この問題意識は WREN の主要メンバー間で共有されている。

[2] 詳細は、「参考２．研究開発評価関連国際会議等の開催状況」を参照のこと。

[3] KISTEP-WREN ワークショップ開催後、我が国においても、WREN 関係者をコメンテータ、講師として招聘したワークショップを研究・技術計画学会、経済産業省、文部科学省等の共催で開催している。

2. PRIME

http://www.prime-noe.org/

PRIME(Policies for Research and Innovation in the Move Towards the ERA)は、欧州研究圏 ERA(European Research Area)の構築に向けた動きの中で、科学及びイノベーション政策の長期的な研究と共通のインフラを展開することを目的とし、EU 各国の科学イノベーション政策と評価に関する研究者を中心に結成されたものであり、第 6 次フレームワーク・プログラム(2002-2006)によって支援を受ける Network of Excellence（NoE）の 1 つである。第 4 次フレームワーク・プログラム以来の欧州における研究開発評価ネットワーク活動の実績を背景とし、NoE 概念の基となる全欧的活動を前史とし、その間に形成された研究者ネットワークの実態化を図ったものである。また、WREN が個人を中心として展開されているのに対し、PRIME は研究開発評価関連機関を中心としたネットワークであり、欧州 16 カ国の 49 の機関、230 人の研究者、120 人の PhD 学生が参加している。

PRIME では、研究及びイノベーション政策が直面している次の 6 つの課題に対する挑戦を掲げており、その活動は、これらの課題を中心に展開されている。

- 研究のダイナミックスの変化に対する挑戦
- 地域と中小企業の重要性の増大に対する挑戦
- 知識循環の促進に対する挑戦
- 公共部門における研究システムの変化に対する挑戦
- 公衆の関心の急激な増大への対応
- 政策決定における多様なアクターの統合に対する挑戦

主な活動は、次のようなものである。まず、欧州共通のインフラ構築の一環として、研究及びイノベーション政策に関するトレーニングコースの提供（参考 3.で詳述）と科学技術指標開発者のための欧州ネットワーク(the European Network of Indicators Producers; ENIP)の構築支援を行い、また、長期的研究として、上記 6 つの挑戦に関わる研究プロジェクトを展開している。プロジェクトには、入会及びレビュー行動(Initiation and Review Actions)、探索型研究(Exploratory Research)、比較研究プロジェクト(Comparative Research Project)の 3 タイプがあり、競争的プロセスを経て選定される。申請は、科学委員会(Scientific Committee)によって評価され、実行委員会(Executive Committee)により採択される。申請 b に際しては、共同出資で、少なくとも 3 カ国から 3 人以上の参加が要求される。2006 年現在、公共介入の合理的根拠、政策形成及び実施プロセスとアプローチの変革等、5 つのクラスターのもとで 22 のプロジェクトが展開されている。

そのほか、年次大会や国際共同カンファレンスの開催（参考 2.）、書籍の刊行(PRIME series)、各種イベントや公募の情報提供などの重要な活動を行っている。

PRIME における意思決定は次の通りである（図１参照）。統括評議会(Governing Board)は、新メンバーの加入や新たな方向性の検討など、戦略的なことがらを決める役割を持つ。実行委員会は、統括評議会により選出される 12 名の委員からなり[4]、支援を行う活動の選定等、運営上のことがらについて意思決定を行う。また、フランスの非営利組織であるARMINESのメンバーで構成される専門的マネジメントチームのサポートを受ける。科学委員会は 6 名で構成され[5]、提出されたプロポーザルの質の評価と、ネットワーク自体の全体的な方向性についての戦略的な評価を行う。独立した組織である特性グループ(Characterisation Group)は 3 名で構成され[6]、ネットワークのダイナミクスを明らかにし、卓越性と統合に関する進化をモニタリングすること、それらを通じて、他の NoE に対しても一般化可能なフレームワークを開発することをその役割としている。

（出典）PRIME ウェブサイト

図1　PRIME の意思決定機構

[4] CNRS-LATTS の Philippe Laredo 氏、マンチェスター大学 PREST の Luke Georghiou 氏、Fraunhofer –ISI（当時、現トゥエンテ大学教授）の Stefan Kuhlmann 氏、サセックス大学 SPRU の Ben Martin 氏、トゥエンテ大学の Arie Rip 氏等。

[5]ジョージア工科大学の Susan Cozzens 氏、パリ国立高等鉱業学校(Ecole des Mines de Paris)の Michel Callon 氏等。

[6] ELTA の Terttu Luukkonen 氏等。

3. INTEVAL

http://www.inteval-group.org/

政策・プログラム評価の国際的研究グループである INTEVAL (International research group on policy and program evaluation)は、ブリュッセルの International Institute of Administrative Sciences からサポートを受けた 1986 年以来、毎年、各国政府、会計検査機関、大学、民間セクターに属す評価専門家と会合を開催している。第 1 回の公共政策評価のトピックに関する国際会議は、1990 年 12 月初頭にオランダのハーグにて開催された。現在メンバーは 20 ヶ国以上にわたる。

国際的な視野に立って、電子メールやサブグループによる年次会合によって研究をすすめている。結果は、共著の書籍や Comparative Policy Analysis Series (Transaction Publishers)のシリーズとして出版される。

4. Platform Research & Technology Policy Evaluation（FTEval）

http://www.fteval.at/home.php

初めは 1996 年に非公式の cooperation として設立され、2001 年に民法のもとで法人として再設立された。オーストリアにおける RTD 政策の最適な戦略計画のための透明性の高い評価活動を奨励し、技術及び研究政策の領域における意思決定者間で評価の文化を進展させることをミッションとしている。メンバーは、オーストリアの連邦教育科学文化省や交通イノベーション技術省（BMVIT)、経済労働省などの省庁や、ARC systems research GmbH や Austrian Science Fund(FWF)、TECHNOPOLIS などのファンディング機関・シンクタンク等を合わせた 16 の連合体となっている。

活動としては、年に数回、テーマを定めた Newsletter “Plattform”の発行（2006 年 3 月現在で 27 巻）や、研究技術政策評価の標準的ガイドラインの策定[7]、ファンディング機関や技術開発プログラムの評価レポートを作成するなどしている。会合としては、2006 年 4 月下旬に BMVIT の音頭による FWF とのジョイントで”New Frontiers in Evaluation”という国際会議を開催した。また、それ以前の会合としては、OECD との共催で、2005 年 1 月 31 日より 2 日間にわたり”Measuring the Behavioural Additionality Effects of Government Financing of Business R&D: Lessons from Country Studies”と題するワークショップを開催している。2003 年 5 月には、ドイツの ZEW 及びオーストリアの Joanneum Research との共催で”Evaluation of Government Funded R&D Activities”という会議も開催している。

[7] http://www.fteval.at/standards/Standards.pdf.

5. Six Countries Programme –The International Innovation Network

http://www.6cp.net/

1975 年に設立された、イノベーションの研究と政策形成にかかわる専門家、政策形成担当者、政策実施者などからなる国際的なネットワークである。主な目的は、イノベーション・プロセスとその経済・社会へのインパクト、及び有効なイノベーション政策の展開に関するより良い理解をもたらすよう貢献することである。オランダ、ドイツ、フランス、UK、アイルランド及びカナダの 6 カ国で初めは組織化されたが、その後、スウェーデン、オーストリア、ベルギー、フィンランド、ハンガリー、南アフリカが新たなメンバーとして加わっている。

秋と春の 2 回のワークショップが最も重要なイベントで、メンバー外にも開かれた雰囲気の中で、各国の最新情勢に関する情報と意見を交換し、この領域の専門家の研究を刺激することをねらいとしている。第 1 回のワークショップは"Government direct financial assistance to industry, programmes, experiences and trends"というテーマで 1976 年 10 月にロンドンにて開催された。評価関連のトピックでは、1982 年 11 月にロンドンにて開催された"Evaluating the effectiveness of government innovation policies"などがある。

最近の活動としては、2005 年 4 月にロッテルダムにて"The Future of Research: New players, roles and strategies"が、同年 11 月にはマンチェスターにて"Innovation and Procurement"というテーマで開催されている。2006 年においては、5 月下旬にストックホルムにて"Innovation Policy learning: change in thinking – change in doing?"というテーマで開催予定である。

6. The American Evaluation Association

http://www.eval.org/

成績評価を中心とした米国の評価者協会として出発したもので、現在はプログラム評価や技術評価などの部門を加え、評価者の国際的な専門職協会に発展している。米国 50 州及び海外 60 カ国を代表する約 4,000 人のメンバーで構成されている。2005 年度には、我が国からも NEDO 職員 3 名を含む 8 名ほどがプレゼンテーターとして参加しており、経済産業省や民間シンクタンク等からの出席者も数名いた。

学会のミッションは、評価の実践や方法を改善すること、評価の活用を増大させること、専門的職業としての評価を促進すること、有効な人間活動についての理論及び知識の創出に対する評価の貢献をサポートすることである。1993 年より年会を開催し

ており、テーマは評価の地位向上（1993 年）、新世紀の評価（1995 年）、理論と実践（1997 年）、評価の能力構築（2000 年）、方法論（2003 年）などとなっている。最近では、2004 年 11 月 3 日から 6 日までジョージア州アトランタにて"Fundamental Issues"というテーマで開催された。2005 年 10 月 26 日から 29 日には、Canada Evaluation Society と共同で"Crossing Borders, Crossing Boundaries"というテーマにてオンタリオ州トロントで開催された。2006 年は、オレゴン州ポートランドにて 11 月 1 日から 4 日までの開催予定となっており、テーマは「評価の結果（The Consequences of Evaluation）」である。詳しくは下記のサイトを参照されたい。

http://www.eval.org/Training/conferencehistory.asp

7. AAAS

http://www.aaas.org/

全米科学アカデミー協会（AAAS）の事務組織の一部門である科学・政策プログラム部門(Directorate for Science & Policy Program)は、米国の研究政策コミュニティ（research policy community : RPC）を結集する場を提供している。「科学・政策プログラム部門」には、「フェローシップ・プログラム(Fellowship Programs)」、「研究開発予算及び政策プログラム(The R&D Budget and Policy Program)」、「科学技術と議会センター(The Center for Science, Technology, and Congress)」、「研究競争力プログラム(The Research Competitiveness Program)」の 4 つのユニットから構成される「科学・工学政策実践グループ(Science and Engineering Policy and Practice Group)」があり、このうち、「研究競争力プログラム」では、研究開発評価に深く関わる活動を展開している。また、「フェローシップ・プログラム」では、人材育成の一環として、科学者、技術者に対し、ワシントン DC の行政機関と議会において 1 年間の科学技術政策の形成に携わる機会（インターンシップ）を提供している（参考３）。

学会活動としては、2 月中旬に開催される年次大会があり、会員からの提案企画セッションでは、研究開発評価を主題とする企画が毎回いくつか取り上げられている。また、次年度予算を議会で審議する途上に合わせて 4 月下旬に開催される科学技術政策フォーラム(AAAS Forum on Science & Technology Policy)は、前述の「研究開発予算・政策プログラム」が主催するものであり、米国の研究開発政策のあり方が RPC によって検討される[8]。これは広い意味での政策評価に相当する活動である。

[8] 以前は、the AAAS Colloquium という呼称が使われていた。

8. 研究・技術計画学会

http://wwwsoc.nii.ac.jp/jssprm/

研究・技術計画学会は、技術経営の向上や科学技術関連政策の立案と推進など、科学技術の経営・政策全般にわたる研究交流と情報交換を図ることを目的として、1985年10月に設立された国内有数の学会である。会員は、科学技術の推進・研究開発・利用に係わる幅広い分野、セクターのメンバーで構成されており、その数は、個人会員962人、法人会員51社にのぼる（2005年9月末現在）。

主な活動内容としては、年次学術大会（秋）や年次シンポジウム（春）、研究分科会の開催、学会誌『研究技術計画』の発行（参考４.）等であるが、ここでは、特に研究開発評価に関係の深い2つの研究分科会の活動について紹介する。

まず、科学技術政策分科会（主査：平澤冷　東京大学名誉教授）では、ウォッチャー・シリーズと題し、主要各国の科学技術政策や評価システム等に関する最新動向を継続的にフォローしているほか（2006年3月現在で計13回開催）、キーパーソン・シリーズとして、文部科学省や経済産業省、科学技術振興機構や産業技術総合研究所をはじめとした我が国の主要な研究開発関係機関のキーパーソンを招いた講演と討論を行っている（2006年現在で計9回開催）。

1991年に活動を開始した研究評価分科会（主査：松井好　社団法人科学技術と経済の会常務理事）では、毎月1回の頻度で研究開発評価に関わる高密度の話題提供と活発な議論を行っている（2006年3月現在で計68回開催）。その内容は、国内外の評価システムの実態やそのあり方をめぐる議論、文献計量学や経済性評価、特許分析等の方法論的検討、概念構築に関する提案など多岐にわたり、話題提供者も、大学や政府機関のみならず、民間の研究所やシンクタンク、海外の研究者など多様である。

なお、評価人材の育成に関して、この2つの研究分科会の両主査による呼びかけによって2003年にはじまった「政策評価相互研修会」は、我が国における産官学の評価人材ネットワークとしても重要な機能を果たしつつある。相互研修会は、シンクタンクを含め研究者・実務家など関心のある主体に広く無料で開放し、協力して、我が国の新たなエキスパート層の人材育成・交流プラットフォーム機能を果たすモデルを構築しようと企画されたものであり、開催当初から財団法人政策科学研究所がそのノードとして継続的に企画・運営等の支援を行っている。2003年度の途上からは、文部科学省との共催で同省や関連機関の職員研修をかねるかたちで実施されている（参考3.）。

１－２　現在活動を休止中のネットワーク

研究開発評価ネットワークとして一時期機能していたが、現在では活動を休止、または先述のネットワークに実質的に吸収されたネットワークもある。そもそもこうしたネットワークは、外部からのプロジェクト資金等を得て組織されているものが多く、プロジェクトの終了と同時に活動を休止する場合が少なくない。

１．ETAN (European Technology Assessment Network)

http://www.cordis.lu/etan/home.html

EU では評価の研究者・実務的専門家ネットワークが形成されてきたが、その代表的な活動として ETAN(European Technology Assessment Network)がある。

ETAN は、EU 第 4 次フレームワーク・プログラムのもとで 1997 年に欧州委員会（European Commission）第 12 総局（科学・研究・開発）の重点社会・経済研究（TSER）プログラムの領域 1「科学・技術政策の選択肢の評価」の活動基盤として形成された。1997 年から翌年にかけての試行期間を経て、欧州 RTD プログラムの社会・経済的インパクト評価の選択肢と限界（Options and limits for assessing the socio-economic impact of European RTD programmes）や科学・技術政策問題に関する専門家と政策立案者間のコミュニケーションの改善：ETAN を事例として（Improving communication between experts and policy-makers on science and technology policy issues: the case of ETAN）など 9 つの研究を実施した。2000 年以降の第 5 次フレームワーク・プログラムでは継続していないが、そのネットワーク基盤となったものである。

ETAN の活動の目的は、EU と国レベルでの科学技術政策立案者とその他の関係者（stakeholder）を支援し、共通の関心についてどのような政策イシューがあり、政策上の選択肢には何があるか、共通の理解を形成することにある。それと同時に、適切な分野においては、その形成を通じて EU および国家にとって、より一貫性があり、調和が取れた補完的な科学技術政策の開発を促進することを最終的な目標とする。そのために、重要な科学技術政策イシューについて、ETAN は全欧レベルで政策研究者および政策立案者のコミュニケーションと討論を促進する活動を行うものである。そのため大学やシンクタンクに所属する欧州の多くの評価研究者が参加した。

政策評価活動の面では、第 3 次フレームワーク計画までの EU の評価体制の改善を目指し、EU 構成各国の評価に関する知見の交換や政策研究者ネットワークの形成に寄与した。

2. World Research Evaluation Network

http://www.prism.gatech.edu/~sc149/reseval/index.html

世界各国の評価研究者と評価機関が経験、方法、データを共有するための組織である。活動の目標は、研究評価（research assessment）の手法に関する最新状況をユーザーのコンテキストを含みつつレビューすること、そしてその分野の研究のベースとなるアジェンダや評価方法の開発について概観することである。

2002 年 9 月に独フラウンフォーファ研究機構システム・イノベーション研究所（FhG-ISI）の主催で開かれた第 7 回国際科学技術指標会議の情報を最後に更新はとまっているが、ウェブサイトには、関連する世界中の研究者、雑誌、研究機関、コンサルタント会社、政府機関へのリンクが張られているほか、過去の会議の記録をはじめとした以下のような大量の重要文献資料が蓄積されている。

・研究者および研究グループ／助成機関等の評価グループ
・国際会議情報
・評価方法に関する重要文献
・当ネットワークが関与した評価活動ならびに評価研究活動
・評価に関連する情報源（S&T データ、連邦政府のプログラム評価関連情報源）
・専門誌“Research Evaluation”について

3. EPUB Network

http://epub.jrc.es/　※現在は http://projects-2002.jrc.cec.eu.int/にアーカイブ化

EPUB は、欧州委員会 RTD フレームワーク・プログラム（FP）の STRATA Action (Strategic Analysis of Specific Political Issues)のもとでの Thematic Network である。EPUB Network では、公的 RTD 政策の社会経済性評価の領域において蓄積された知識を吟味することで、可能な政策オプションを調査し、強みや限界を示し、公共政策がどのように経済パフォーマンスに影響を与えているかに関する指標を提供している。専門家と政策立案者との交流にも力を入れており、2000 年 3 月よりセビリアやウィーンなどで数多くのワークショップを開催している。それらのアウトプットとして、”Research and Technological Development Evaluation Toolbox –Assessing the Socio-Economic Impact of RTD-Policies”が 2002 年に出された。EPUB Network のメンバーは、RTD 評価の分野や評価と政策立案の間の相互作用のインターフェイスに関する専門家により構成されている。

第 6 次 FP で研究者ネットワーク部分が PRIME として独立し強化された。

１－３　その他の学協会等

研究開発評価を主題とするものではないが、以下のような独自の評価関連の学協会をもつ国や地域も少なくない。

- The European Evaluation Society（ヨーロッパ）
  http://www.europeanevaluation.org/
- The UK Evaluation Society（イギリス）
  http://www.evaluation.org.uk/
- French Evaluation Society (Société française de l'Evaluation)（フランス）
  http://www.sfe.asso.fr/
- German Evaluation Society (DeGEval)（ドイツ）
  http://www.degeval.de/
- 日本評価学会
  http://www.idcj.or.jp/JES/

また、科学技術のあり方を主題にした学会の中で、研究開発評価が扱われている場合もある。

- EASST（European Association for the Study of Science and Technology）
  http://www.easst.net/
- 4S（Society for Social Studies of Science）
  http://www.4sonline.org/

技術経営を主題とする学会の中で、研究開発評価が扱われている場合もある。

- PICMET（Portland International Conference on Management of Engineering and Technology）
  http://www.picmet.org/main/

参考情報2. 研究開発評価関連国際会議等の開催状況

研究開発評価関連の国際会議やワークショップは、これらを通じて養成される評価人材の数量的側面からも大きな効果をもたらしているものである。アナリストや研究者等の評価プロフェッショナルのみならず、評価マネジメント人材の知的交流の場として重要な機能を果たしている。

また、国際会議におけるテーマの変遷をみれば、評価研究や実務のフロンティアでどのような問題意識がもたれているのか、国際動向を探る手がかりにもなる。

1. OECD が主導する国際会議

OECD では、1997 年 6 月にパリで開催された国際会議「イノベーション及び技術の政策評価」での議論をまとめた記念碑的報告書"Policy Evaluation in Innovation and Technology –Towards Best Practices"を皮切りにして、研究開発評価に関わる多様なテーマの国際会議、ワークショップを開催している。

以下、近年開催された主な会議の状況についてまとめた。

テーマ	Workshop on Measuring the Behavioural Additionality Effects of Government Financing of Business R&D: Methodological Approaches and Preliminary Findings
日 時	2004 年 5 月 10～11 日
場 所	Manchester, UK
主催等	OECD 及び PREST 共催
参 照	http://www.oecd.org/document/23/0,2340,en_2649_34269_34509399_1_1_1_1,00.html
概 要	・ 政府による民間企業への研究開発助成に関するBehavioural Additionalityの測定が焦点 ・ 2003 年 9 月 12 日にブリュッセルにて開催された IWT-Flanders 及び OECD 共催の専門家会合により開催の提案 ・ 豪州、オーストリア、ベルギー、フィンランド、ドイツ、アイルランド、ノルウェイ、英国、米国の研究者及び行政官によるプレゼンテーション ・ 11 カ国から 30 名の研究者、行政官、BIAC(Bussiness and Industry Advisory Committee to the OECD)に代表されるビジネス・コミュニティのメンバーが参加
テーマ	Workshop on Measuring the Behavioural Additionality Effects of Government Financing of Business R&D: Lessons from Country Studies
日時	2005 年 1 月 31 日～2 月 1 日
場所	Vienna, Austria
主催等	OECD 及び Platform FTEval の共催
参 照	http://www.oecd.org/document/0/0,2340,en_2649_34273_34538432_1_1_1_1,00.html

テーマ	International Workshop on the Evaluation of Publicly Funded Research
日　時	2005 年 9 月 26～27 日
場　所	Berlin, Germany
主催等	OECD 及び BMBF 共催
参　照	http://www.oecd.org/document/37/0,2340,en_2649_34293_35450213_1_1_1_1,00.html
概　要	・ 機関評価及び大型研究施設/プロジェクトの評価が焦点 ・ 24 カ国から 100 名以上が参加

2. G8 が主導する国際会議

G8 では、2001 年の米国を皮切りにして、研究開発評価に関するワーキンググループの会合を毎年開催している。

主要メンバーは NSF（米）、EPSRC（英）、DFG（独）、NSERC（加）のアナリストであり、フランスは研究技術省の視学官と研究機構 CNRS（仏）のアドミニストレータである。露、伊には適切な人材が存在せず毎回欠席、我が国からは行政機関の評価担当部署の行政官が出席している。

以下、ここ 3 年間の会議情報について掲載した。

テーマ	G-8 Working Group on Research Assessment
日時	2004年10月17日～20日
場所	Berlin, Germarny
主催等	ホスト国はドイツ

テーマ	G-8 Working Group on Research Assessment
日時	2005年11月14～16日
場所	Mancester, UK
主催等	ホスト国はイギリス

テーマ	G-8 Working Group on Research Assessment
日時	2006年9月26～27日
場所	Otawa, Canada
主催等	ホスト国はカナダ

3. EU が主導する国際会議

EU で本格的な評価関係の国際会議が企画されるようになったのは、第 5 次フレームワーク・プログラム（FP5）においてである。FP5 下では、議長国持ち回りで半年に 1 回の定例国際会議が開催され、その成果は、PREST et al. (2002.6)としてまとめられている。

この流れとは別に、ETAN の流れを受ける EPUB の枠組みにおいて、研究開発評価の Toolbox をまとめる作業が行われた。現在では、未来志向の技術アセスメントの会合（FTA）や、PRIME の枠組みのもとで多様な国際会議が開催されている。

テーマ	The European Research Evaluation Network Conference and the RTDN Meeting "Evaluation – Connecting Research and Evaluation"
日　時	2005 年 11 月 17～18 日
場　所	Manchester, UK
主催等	European Research Evaluation Network 及び RTDN の共催
参　照	http://www.mbs.ac.uk/research/centres/engineering-policy/evaluation-network-conference.htm

テーマ	Making monitoring and evaluation of innovation programmes a competitiveness tool
日　時	2005 年 7 月 5 日
場　所	Brussels,
主催等	欧州委員会 Innovation Policy Unit of DG Enterprise of the European Commission
参　照	http://www.eu-innovation.net/innovation_evaluation/htm/workshop.htm

4. PRIME が主導する国際会議

(1)年次大会(Annual Meeting)

PRIME では、2005 年以来、以下のような年次大会を開催している。

【第 1 回】

テーマ	The future of reseach in Europe – A knowledge base for policy and manageme
日時	2005年1月6～9日
場所	Manchester, UK
主催等	PRIME主催
参照	http://www.prime-noe.org/index.php?project=prime&locale=en&level1=menu1_prime_1b8057d059a36720_1&level2=6&doc=Annual_Conference

【第 2 回】

テーマ	Monitoring and analysing integrative dynamics in research and research policy in Europe
日時	2006年2月6～10日
場所	Paris, France
主催等	PRIME主催
参照	http://www.prime-noe.org/index-restricted.php?project=prime&locale=en&level1=menu1_prime_1b8057d059a36720_1&level2=6&doc=2006_Annual_Conference

【第 3 回】

テーマ	Building the research agenda of the speciality which we now label as "science and
日時	2007年1月29～2月1日
場所	Pisa, Itary
主催等	PRIME主催
参照	http://www.prime-noe.org/index-restricted.php?project=prime&locale=en&level1=menu1_prime_1b8057d059a36720_1&level2=6&doc=2006_Annual_Conference

(2)国際会議

2006 年には米国アトランタにてジョージア工科大学との共同で以下のような国際会議を開催することになっている。

【The Atlanta Conference on S&T Policy 2006】

日　時	2006 年 5 月 18～20 日
場　所	Atlanta, USA
主催等	The Georgia Tech School of Public Policy, The Business School at Georgia Tech, 及び PRIME 主催
参　照	http://www.spp.gatech.edu/conference2006.php
概　要	・ イノベーション促進政策のあり方が焦点 ・ WREN のウェブサイトからもリンクが貼られており、主要メンバーも発表者等として参加 ・ WREN の Valdez 氏(DOE)がチェアを務める「評価のフロンティア(Frontiers of Evaluation)」と題するセッションは、以下のような演題とメンバーで構成; 1. A Strategic Balanced Scorecard for Publicly Funded Science. Gretchen Jordan (Sandia National Laboratories)他 2. Impact Assessment as a Management Tool: Experiences in Finnish Research Organizations. Jari Konttinen/ Kirsi Hyytinen(VTT(Technical Research Centre of Finland)) 3. R&D Assessment: Global Trends in Metrics and Measures Julia Melkers/ Eric Welch (University of Illinois)他

5. WREN が主導する国際会議

(1)年次ワークショップ

参考１．で言及したように、WREN では、2003 年の設立以来、以下のような国際ワークショップを毎年開催している。

【第 1 回】

日　時	2003 年 12 月 4～5 日
場　所	Washington, DC
主催等	DOE の主催
参　照	http://www.wren-network.net/resources/WREN2003DecWorkshopReport.pdf 等
概　要	・ 行政管理予算局 OMB によるプログラム評価法 PART に対応するための能力をいかに改善するかが焦点（全体のテーマは"Planning for Performance and Evaluating Results of Public R&D Programs: Meeting the OMB PART Challenge"） ・ 27 以上の連邦政府機関及び海外 7 カ国から 200 名以上が参加

【第 2 回】

日　時	2004 年 6 月 17～18 日
場　所	Brussels, Belgium
主催等	欧州委員会 EC 及び WREN の共催
参　照	http://www.wren-network.net/resources/2004eu.htm
概　要	・ RTD プログラムの実績をいかに把握するかが焦点（全体のテーマは"Performance Assessment of Public Research, Technology and Development (RTD) Programmes"） ・ 欧州各国、米国、カナダ、豪州、ニュージーランド、日本、韓国等より延べ 115 名が参加

【第 3 回】

日　時	2005 年 5 月 30～31 日
場　所	Seoul, Korea
主催等	韓国 KISTEP 主催、WREN 及び欧州委員会 EC 共催
参　照	http://www.wren-network.net/resources/2005kistep.htm
概　要	・ システムレベルの評価が焦点（全体のテーマは"National Models for Public R&D Evaluation: In Search of Best Practices and Collaborative Opportunities"） ・ 欧州各国、米国、カナダ、インド、日本、韓国、中国のスピーカーとファシリテータ

（2）WREN Logic Model Working Group

2004 年には、春秋の 2 回、以下のようなロジックモデルのワーキンググループによるワークショップが開催されている。

日　時	2004 年 6 月 30 日及び 9 月 30 日
場　所	Washington, DC
主催等	WREN による共催
参　照	http://www.wren-network.net/events/2004-67.htm

（3）その他

上記のほか、WREN の主要メンバーが運営に深く関わるものとしては、以下のものがある。

【R&D Evaluation Workshop in Japan】

日　時	2005 年 6 月 2～3 日
場　所	東京
主催等	研究・技術計画学会、経済産業省、文部科学省、産業技術総合研究所及び科学技術政策研究所による共催
参　照	http://www.wren-network.net/resources/2005RD.Japan/2005.RD.Japan.htm
概　要	（タイトル：Evaluation of Science and Technology Policy System in Japan: Country case Study）

【2005 AEA / CES Evaluation】

アメリカ評価学会 AEA(American Evaluation Association)の年次大会には、毎年のように WREN の主要メンバーが発表者やオーガナイザーとして名を連ねており、WREN のウェブサイトからもリンクが貼られている。

テーマ	2005 AEA / CES Evaluation:
日　時	2005 年 10 月 26～30 日
場　所	Toronto, Canada
主催等	American Evaluation Association 及び Canadian Evaluation Society の共催
参　照	http://www.wren-network.net/resources/2005AEA_CES.htm
概　要	（タイトル：Crossing Borders, Crossing Boundaries）

6. FTEval 等によるオーストリアを中心とした国際会議

【2003 年】

テーマ	Evaluation of Government Funded R&D Activities
日時	2003 年 5 月 15 日～16 日
場所	Vienna, Austria
主催等	ZEW（ドイツ）、JOANNEUM RESEARCH、PLATFORM の共催
参　照	http://www.fteval.at/papers/

【2006 年】

テーマ	New Frontiers in Evaluation
日時	2006 年 4 月 24 日～25 日
場所	Vienna, Austria
主催等	BMVIT、FWF 及び FTEval の共催
参　照	http://www.fteval.at/conference06/

7. 我が国において開催の国際会議

我が国においては、上述の WREN のメンバーを招いた会議とあわせて、これまで以下のような評価関連の国際会議が開催されている。

テーマ	国の研究開発プログラム＆プロジェクトの評価システムのあり方
日時	1999 年 3 月 5 日
場所	東京
主催等	財団法人政策科学研究所（委託元：通商産業省工業技術院）

テーマ	公的研究開発の社会・経済へ及ぼすインパクト
日時	2002 年 4 月 18 日～19 日
場所	東京
主催等	経済産業省及び新エネルギー・産業技術総合開発機構の共催

参考情報3. 研究開発評価関連研修コース・大学院専攻等

我が国においては、資金配分機関に限らず、研究開発・科学技術・イノベーション政策を担う専門人材が不足しており、また、系統的な育成プログラムが未整備であり、ようやく分散的でないプログラムが 2004 年に政策研究大学院大学に設置された段階であるが、欧米では、以下に挙げたような高等教育機関や研修プログラムが長年運用されている。

これらのコースや専攻の多くは、政策研究というより広い枠組みの中で運用されており、必ずしも評価を直接的にとりあげるものではないが、幅広いスキルを求められる難度の高い人材育成という観点からも、我が国への示唆に富む内容であると言えよう。

なお、以下でもいくつかをとりあげているが、米国国内の関連大学院コースについては、AAAS のウェブサイトに、「科学・工学及び公共政策に関する大学院教育ガイド(the AAAS Guide to Graduate Education in Science, Engineering and Public Policy)」としてまとめられているので参照されたい。

http://www.aaas.org/spp/sepp/index.htm

3－1. 高等教育機関の専門コース等

1. マンチェスター大学 PREST

http://www.mbs.ac.uk/research/engineering-policy/index.htm

英国マンチェスター大学の PREST では、5 日間の集中講義の形式で研究評価の教育コースを開講している。このコースは PREST の 1 年間の修士課程(MSc)の中の授業の一つとして位置づけられているものであり、それを外部にも公開したものとなっている。とは言っても、この修士課程には各国の科学技術行政を担当する省庁から留学生が多く派遣されており、また、英国以外の国の省庁や研究施設からも実際にコース受講者が来ている。そのため、研究評価の実務を行っている現場でも役立ちうるように、実習を多く含めた内容となっている。

教育コースは毎年 1 月の半ばの週の月曜日から金曜日までとなっており、ほとんどの日は朝 10:00 から夕方 17:00 頃というスケジュールになっている。なお、PREST の学生以外の受講料については、EU 圏内であれば割引があるという。

講師は、研究評価研究の第一人者である Luke Georghiou 教授をはじめとして主に

PREST 所属の教員で構成されている。これは後述するオランダのトゥエンテ大学が各国の研究評価研究者を集めて行っているのとは異なる。その理由としては、このコースが PREST 内部の授業でもあること、PREST が研究評価研究を主要テーマの一つとする研究機関であるために評価研究・調査の経験や知識を有する教職員が揃っていることが挙げられる。ただし、実際の評価事例の紹介については、外部からゲスト講師を招いている。

また、このコースで対象としている受講者は、上述のように 2 種類の受講者が実際に受講していることからも分かるように、行政組織や資金配分機関あるいは研究所などで評価に当たる実務者、および、研究評価研究や科学技術政策研究の分野のアカデミックな研究者や民間調査機関の調査研究者・コンサルタントなどを目指す者である。そのため、教えられる内容はこれら 2 者に共通的に必要な基盤的内容となっている。

より具体的に見ていくと、修士課程におけるプログラムには「科学、技術及びイノベーションのマネジメント（Management of Science, Technology and Innovation）」がある。これは、6 ヶ月を研究とコースワークに、もう半年を研究論文にあてる 1 年間のフルタイムのプログラムである。専門分野を反映したトピックの選択としては以下のものが挙げられている。

・科学、技術及びガバナンス
・イノベーション及び技術マネジメント
・フォーサイトと未来
・持続可能性
・科学、技術及び情報社会

短期コース及びセミナーは 5 日間のものが主だが、1－2 日間の集中講義の形式もあり、顧客の要望を考慮してモジュール化されている。内容としては、科学技術政策の評価、鍵となるイシュー及び戦略（Key issues and strategy）、フォーサイト、欧州レベルでの研究開発機関がある。以下にいくつか概略を示す。

科学技術政策の評価コースでは、以下の要素が含まれている。

・評価の編成
・評価のデザイン
・ピアレビューとパネル
・ケーススタディ
・研究の社会及び経済インパクト
・量的指標とビブリオメトリクス
・研究評価のインパクト
・機関の評価
・鍵となるイシュー及び戦略（Key issues and strategy）
・科学技術における優先順位とフォーサイト

・Ｒ＆Ｄの評価
・Ｒ＆Ｄにおけるコラボレーション
・イノベーションのための戦略
が含まれている。

2. サセックス大学 SPRU

http://www.sussex.ac.uk/spru/

SPRU(Science Policy Research Unit)は、サセックス大学の科学技術関連政策研究を主題とする大学院研究科である。SPRU は設立当初、科学政策研究のメッカとして機能していたが、対象分野を拡大し相対的に科学政策研究の比重が低下してきた。現在では社会科学系の学術研究に比重が移行し、実務的な政策分析や評価に取り組む有力な研究者はいない。この点、PREST と好対照になっている。これは英国における大学評価制度の顕著な帰結でもあり、資金源を実務的委託研究にシフトした PREST と、リサーチカウンセルからのアカデミックな公的研究助成に依存する SPRU との戦略の違いがその歴史的経緯に如実に現れている。この間、有力な実務的研究者は SPRU からスピンアウトし、コンサルタントやシンクタンカーとなっている。

大学院の修士学位プログラムには毎年約 40 名の学生が参加し、そのうち自国である UK 出身の学生は約 3 分の 1 で多くは他国からの学生である。プログラムは以下の 5 つで構成されており、1 年間のフルタイムか 2 年間のパートタイムで履修する。

- 科学、技術及びイノベーションの公共政策：Public Policies for Science, Technology and Innovation (PPSTI)
- 産業及びイノベーション分析：Industry and Innovation Analysis (IIA)
- 技術及びイノベーション・マネジメント：Technology and Innovation Management (TIM)
- 持続可能性のための科学技術：Science and Technology for Sustainability (STS)
- 国際的マネジメント：International Management (IM)

このうち、IIA と STS はそれぞれ 2003-04 年、2004-05 年に提供された新しいプログラムである。IM は 2006-07 年から開始する予定になっている。

表 1 に例として、科学、技術及びイノベーションの公共政策：Public Policies for Science, Technology and Innovation (PPSTI)のプログラム構造を示す。

表1　PPSTI のプログラム構造

1 学期 秋 2005 年	科学技術研究の紹介 技術及びイノベーションシステム イノベーションに関する経済的観点 イノベーションの機構 科学技術に関する社会学的観点 政策、ガバナンス、規制 研究及び調査スキル
2 学期 春 2006 年	※以下から2つのオプションコースを選択 持続可能性のためのイノベーション 技術的リスクのマネジメント ICT 政策と戦略 科学政策の政治経済学 複雑な製品及びシステムにおけるイノベーションのマネジメント グローバル経済での競争
3 学期 夏 2006 年	統計的研究方法 研究デザイン、計画、マネジメント 論文（20,000 字）

（出典）SPRU ウェブサイト

PPSTI は、イノベーションの原因と結果の関係、科学・技術と社会の関係、公共政策等の役割を明らかにすることを目的としている。科学と技術、社会の関係を誘導するのが政策であるという位置付けである。また、TIM では、創造的なマネジメントの実践と技術経営の研究活動に必要な問題解決能力とイノベーションのための組織論を学ぶことを目的としている。4－5 人のチームで長期間、企業でコンサルティング・プロジェクトも行う。

このようにして、SPRU では科学技術政策、政策評価に関しても、実務的というより、政策研究者の視点で体系化された multi-disciplinary な教育を行っている。

3. トゥエンテ大学

http://www.utwente.nl/en/

オランダのトゥエンテ大学においても、PREST と同様に 1 週間ほどの教育コースが開講されている。このコースは 15 年前から開催されているものである。トゥエンテ大学の場合には、研究評価のみならず科学技術政策研究や STS 研究の第一人者である Arie Rip 教授が中心となり教育コースを開講している。全体の講師は、トゥエンテ大学在籍の講師が Rip 教授を入れて 2 名のみであり、この他に、前述の英国マンチェスター大学 PREST の Dr. Keenan、ドイツ FhG-ISI の Kuhlmann 教授、米国ジョージア工科大学の Shapira 教授といった外部講師を招いて構成されている。また、ケース・スタディとして EU やオランダ大学協会 VSNU からもゲスト講師を招いている。

トゥエンテ大学の教育コースは 3 日半という短い日数で行われる（表 2）。だが、各日は朝 9 時から夜 9 時までという時間的に密なスケジュールになっているため、PREST と比べてもほぼ同様の時間数であると言える。

表2 トウェンテ大学の評価研修コース

1 日目	評価の社会的文脈	イントロダクション 評価の社会的文脈 コース概要
	評価のデザイン	評価のデザイン 演習1：目的からデザインへ（ロジックチャートの演習と講評）
	方法のカタログ	事例：オランダの大学評価におけるピアレビュー
2 日目	方法のカタログ（つづき）	方法のカタログ（概要） 中間的インパクトとしてのネットワーク 「社会的質」の評価 経済的インパクトの評価 事例：EU の研究評価 演習2：評価の方法の総合的利用（演習と講評）
3 日目	政策のための評価	評価データを用いた研究政策の改善 事例：政策的介入の方法としての評価 因果関係の理解 演習3：評価レポートの作成
4 日目	政策のための評価（つづき）	演習3（講評） コースのレビュー 将来の研究評価

トゥエンテ大学の教育コースの内容は上の表に示すようなものである。コースは大きくは次のように区分されている。

- 評価の社会的文脈
- 評価のデザイン
- 方法のカタログ

・　　政策のための評価

（1）評価の社会的文脈

最初の「評価の社会的文脈」は教育コースのイントロダクションであり、PREST 同様に評価の必要性を説明することで、なぜ評価を行うかという動機付けを再確認させるものである。この講義では、1980 年以降の研究システムの変化により、研究評価がより複雑で困難な営みとなっていることが説明される。すなわち、1980 年以前には競争的資金と委託研究との区別がある程度明確に線引きされていたために、それぞれピアレビューおよびプロジェクトの契約目標達成の分析という形で評価を行うことが可能であった。しかし 1980 年以降に戦略研究という新たなプログラム概念が生じ、より基礎的な研究開発行為についてもその社会経済的効果を中・長期的なスパンで分析することへの要求が生じてきたために、評価がより複雑なものとなってきている。

（2）評価のデザイン

次に評価のデザインとしては、PREST の Dr. Keenan を講師として、上述のロジックチャートの説明がなされる。実習の対象としては、2001 年の場合にはやはり EU の EUREKA が用いられている。

（3）方法のカタログ

評価の方法論としては社会的効果、経済的効果の分析手法が紹介される。科学技術的効果の分析手法の代表であるビブリオメトリクスについては数年前までは独立の時間割をくんでいたが、受講者の多くがその内容を既に知っている場合が多いため、独立の時間割ではなく他の説明の中に入れ込む形で簡単なものにしたという。一方、オランダでは社会的効果への注目が他国よりは強いという指摘がなされているが（van der Muelen and Rip 2000）、彼らの最近の研究成果を踏まえた分析手法の紹介がなされている。そこでは特に、新たなアクターを評価に取り込む必要が指摘され、そのようなシステム形成や、実際の環境改善や疾病治療への効果やその他の QOL への効果、規制の変化、国際的ポジショニングの変化などが評価の具体的内容として挙げられる。また、評価の示し方も社会的効果以外も含めた多様な項目をポートフォリオとして示す方法などが紹介される。
また、経済的効果についてはやはり PREST と同様に、スピンアウトの概念や BETA 法が中心的に説明される。

トゥエンテ大学ではこの「方法のカタログ」の中で「中間的インパクトとしてのネットワーク」というタイトルの授業が開かれている点が PREST とは若干異なる点である。ここでは、プログラムの中間的なアウトプットとしてのネットワークへ焦点を置くことの必要性とネットワークをトレースする方法が説明される。これは、研究開発やイノベーションが、大学研究者や民間企業研究者などの共同研究や情報交流、さらには、それらに加え

てユーザーとの交流をも通じて展開するものであるため、そのネットワークへ焦点を当てることが研究実施のプロセスを考える上でも、顧客が誰でありいかに相互作用すべきかを考える上でも重要であることを認識するために必要な授業と考えられる。また、企業間や産学間、あるいは国家間での共同研究を誘因することを目的とするプログラムではネットワーク形成はより明示的に評価でも取り上げられるべきものである。この内容は、部分的には PREST でもスピンアウト効果や行為のアディショナリティの説明、および、研究実施とその効果の関係の多様性の説明においてなされているものであるが、ネットワーク形成自体を特別な時間枠をとって教えるものではなかった。これは前述のように、PREST では別途イノベーション研究の授業が用意されている事も影響している可能性がある。
また、ピアレビューについてはオランダの VSNU（大学協会）の Steijn 氏を招いて、大学の研究評価におけるピアビューの事例を紹介する。大学評価を事例として紹介している点では PREST と同様であるが、オランダの大学評価は英国 RAE とは目的が異なり研究資金配分を行うものではないため､その方法は異なるものであることには注意が必要である。

２日目の最後にはこれら方法論を踏まえて、英国 Centre for Urban and Regional Development Studies などを対象とした実習を行う。そこでは評価目的の設定と、それに適切な評価方法の選択が学習される。

（４）政策のための評価

３日目には Kuhlman 教授により、ISI が行った臨床研究センターの評価などを事例にしながら評価と政策との関係が説明される。評価は公的資金の使用法のチェックや目標達成の確認だけを目的とするのであれば政策的意義が薄いものとなる。政策やプログラムがどのように適切に展開され、今後もされうるかを評価で問うことが重要となる。ここでは評価対象となっている資金提供プログラムと実際に評価過程で同定された効果との因果関係の問題なども議論される。

また、これらを踏まえた上でいかに評価レポートを作成すべきかが、実習を含めて教えられる。ここでは、初日に扱った EUREKA が事例として再び用いられる。

最終日には、今後の評価システムで焦点となる問題についての Rip 教授の講義でコースは締めくくられる。

欧州の研究者の間では研究評価のシステムや手法の現時点でのスタンダードについて、ある程度のコンセンサスが得られていることが伺える。ただし、社会的効果の分析やネットワーク形成の分析、サービスセクターの評価などは今後の研究評価の焦点となるものであるため、実際の評価や評価研究が進むにつれて内容は改善されていくものと思われる。

４. ジョージア工科大学

http://www.spp.gatech.edu/index.php

ジョージア工科大学（Georgia Institute of Technology）の School of Public Policy[9]では、修士・博士課程における Areas of Concentration（集中分野）の１つとして科学技術政策が設定されており、この一環として研究評価に関連する講義を行っている。

修士・博士課程のカリキュラムは、コアカリキュラムと選択科目からなっており、選択科目で科学技術政策を選ぶことが出来る。科学技術政策関連の選択科目は表３の通りである。これらの他にも、「Special Topics」として、研究評価関連の講義が行われる場合もある。

表３　GIT の修士・博士課程の科学技術関連の選択科目（PUBP は、科目コード番号）

選択科目（科学技術政策）	
PUBP 640	科学、技術及び公共政策：科学技術の支援、コントロール政策を含む科学技術と政府との関連の検討
PUBP 6402	研究政策とマネジメント：研究政策およびマネジメントの課題の検討。公、民間および非営利組織の研究活動は、戦略的計画、資源配分、技術移転および研究評価の実践を検討する際に対比させられる。
PUBP 6414	技術的革新と政府の政策革新を促す連邦政府と州の政策。革新の源と刺激。大学と企業コンソーシアムの役割。革新政策の比較。技術政策の評価。
PUBP 6415	地方、技術および政策：概念、問題、および地方開発、経済開発、産業の変化および技術政策と関係する政策の調査。
PUBP 6417	科学技術における重要観点：このコースは、学生の科学技術に関するクリティカルな思考、および市場と政治、社会との関係を刺激するよう努める。議論は、科学技術コミュニティーの社会構成、科学技術における経済と政治力の役割などのトピックを含む。
PUBP 6421	大規模な社会技術システムの開発：このコースは、科学技術の社会、政治、文化的側面、及び国の規制政策や研究開発にどのような影響を与えるかを検討する。
PUBP 6608	技術のマネジメント：外部環境
PUBP 6753	科学技術政策の比較
PUBP 6777	新興技術の分析：このコースは、技術のモニタリング、予見、評価手法の利用技術を開発する。さらに、新興技術領域の現状や見通しを検討する。

また、専門教育に関する週日の短期コースとして、以下の２つの R&D 評価がある。

①R&D のインパクトアセスメント

このコースでは、公的セクターの研究開発プログラムのインパクト測定に重点が置かれている。公的 R&D プログラムのロジックモデリングにおける基本的な概念を含み、国際的にも活用されているインパクトの概念や測定法を紹介し、受講生に対して

[9] http://www.spp.gatech.edu/

自分が所属する機関の文脈に即した測定にトライできる機会を与えている。
②R&D 評価－研究及び技術から結果を評価する
　本コースでは、科学技術プログラムのためのパフォーマンス測定と評価のデザインと実施における最新の知識と経験を提供する。講義は R&D 評価のエキスパートによってなされ、小グループによる参加型のケーススタディや各機関・国の経験について意見交換を行う。
　受講生の対象としては、
・研究を実施あるいはサポートする連邦・州・民間機関からの企画及び評価スタッフ
・大学あるいは政府機関の R&D マネージャー
　・R&D 評価を取り入れたい、あるいはその領域における理解やスキルを改善したい専門家を想定している。

5. RAND Graduate School

http://www.prgs.edu/

The Frederick S. Pardee RAND Graduate School（以下 RAND-GS）は、公共政策における初期の 8 つの大学院プログラムの 1 つとして 1970 年に設立された。当時は RAND-GS が唯一の PhD に特化したプログラムであった。また、多くの公共政策分析ツールを開発してきたシンクタンク RAND Corporation に基盤を置いていることも特筆すべき点である。RAND-GS の学生は、RAND の学際的な研究チームのメンバーとしてパートタイムで業務をこなし、フェローシップを通して経験を積んでいく。このようなアドバンスコース・ワークと OJT の組み合わせがユニークなところである。卒業生は、表 4 のようなセクターへ就職している。

表4　卒業生の就職先

セクター	比率
研究機関(Research)	59%
専門研究機関(Dedicated Research Institutions)	38%
大学(Universities (as faculty))	21%
公共部門(Public Service)	20%
政府機関(Government)(軍関係を除く)	11%
NGO (Non-Government Organization)	6%
軍関係(Military)	3%
民間部門(Private Sector)	21%

カリキュラムに関しては、数学や経済学、社会科学等を活用した研究方法に重点を置いた分野横断的な構成となっており、学生は、健康、教育、エネルギー、環境問題、公共の安全性、労働、人口統計などを含む幅広い政策イシューについて学んでいく。政策分析や経験科学や社会科学の手法、ミクロ経済学やオペレーションズ・リサーチ（OR）など 11 の必修コースがある。また、選択コースは毎年 20 以上提供され、特定の政策領域に関するワークショップや上級の方法論コースなどが含まれている。

RAND-GS のコア・コースは、Ph.D.レベルの応用政策研究の基礎を提供するものであり、大きく以下の 4 つから構成されている。

①実証的分析（3 コース）

確率論及び統計学、回帰分析及び計量経済学から選択

②ミクロ経済学（3 コース）

メイントピックとしては、消費者理論、企業理論、部分均衡分析、ゲーム理論、市

場力学、プリンシパル - エージェント(principal-agent)分析、意思決定、メカニズム・ザイン、市場均衡と市場の失敗などを含む。

③分析方法（1、1.5 コース）

不確実な状況下での選択という複雑な問題に対し、意思　決定者を支援するために用いられる数量的なツールの紹介

④政策研究（4 コース）

社会科学をベースにした研究方法の習得や RAND 研究からのケース・スタディなどを行う。

コアとなるコースは表５のような構成になっている。

表５　RAND Graduate School におけるコア・コース

年次	秋期	冬期	春期
1 年次	政策分析Ⅰ:入門 実証分析Ⅰ: 蓋然性及び統計学	政策分析Ⅱ: 社会科学方法論 ミクロ経済学Ⅰ 実証分析Ⅱ: 回帰分析	ミクロ経済学Ⅱ 実証分析Ⅲ: エコノメトリックス 費用便益分析 （1/2 コース）
2 年次	オペレーションズ・リサーチⅠ ミクロ経済学Ⅲ	政策分析Ⅲ:成功の教訓	
3 年次	政策分析Ⅳ:学位論文のための方法論		

また、選択コースとして、社会及び行動科学があり（各 10 週間)、構成は以下のようになっている

・Sociology and Policy Research
・Sociology and Policy Research
・Governance in Three Flavors
・Topics in History and Public Policy
・Collapses of Past Societies and Their Lessons for Our Own Future
・Topics in Advanced Behavioral Science: Applications of Psychological Theory to Policy Research
・Multi-Level Modeling
・Qualitative Research

また、以下のようなワークショップがこれまで開催されている。

・Long-term Policy Analysis
・Health Economics
・Science and Technology

・Welfare Reform
・Racial and Ethnic Disparities in Health
・Gender, Race and Policy
・Quality of Care
・Workshop on Quantitative Methods and Education

6. ジョージ・ワシントン大学

http://www.gwu.edu/~cistp/

ジョージ・ワシントン大学のエリオット国際関係学部国際科学技術政策センターでは、大学院教育及び研究に関して、以下のような５つの領域を網羅している。

・科学、技術及びイノベーション政策
・研究開発プログラムの評価
・テクノロジー及び産業のダイナミクス
・宇宙政策
・情報技術と国際関係

修士課程においては、国際的科学技術政策プログラムが設置されている。当該プログラムでは、分野横断的なカリキュラムが特色であり、必修科目は以下のように５コース（15 単位）となっている。

・IAFF 220　科学、技術及び国際的問題
・IAFF 229　科学、技術及び世界的諸問題の分野横断的なセミナー
・IAFF 298　自主性のある学習及び研究

以下から少なくとも２つのコースを選択する。

・IAFF 221　テクノロジーの創造と拡散
・IAFF 222　技術協力：インセンティブと戦略的提携のための政策
・IAFF 223　米国の宇宙政策
・IAFF 224　米国の宇宙政策におけるイシュー
・IAFF 225　環境政策
・IAFF 290　特別なトピックス（例：技術と国際競争、科学技術と複雑性、宇宙と国家セキュリティなど）
・ECON 255　技術的変化の経済学

7. ミネソタ大学

http://www.hhh.umn.edu/

ミネソタ大学の Hubert H. Humphrey Institute of Public Affairs では、修士課程のプログラムとして、公共政策に加えて、Master of Science(MS)がある。MS では、自然科学や工学はバックグラウンドに持つ学生に対して、経済や食料生産・健康、エネルギーや環境、セキュリティ・ポリシー、教育等における科学・技術について学生をトレーニングする。内容としては、国家間の政策及び経済関係における科学・技術のインパクトを含んでいる。MS の学生は、科学・技術の適切な促進・規制のための政策の分析やデザインについて地域的、国家的、国際的な観点から教育を受ける。必修コースは以下の通りである。

・PA 5012　公共問題の政策
・PA 5021　政策分析及びプランニングのための経済学
・PA 5701　科学と国家
・PA 5711　科学技術政策
・PA 5721　エネルギー及び環境政策
・PA 5722　環境及び資源経済政策

また、以下から２つを選択する。

・PA 5032　中間回帰分析
・PA 5033　多変量解析法
・PA 5035　調査研究とデータ収集
・PA 5036　地域経済分析
・PA 5037　地域人口統計分析

8. プリンストン大学

http://www.wws.princeton.edu/step/

プリンストン大学のウッドローウィルソン公共・国際関係学部科学工学環境政策プログラムには、科学・技術及び環境政策（STEP）プログラムがある。政策分析の分野への系統的な導入を提供することに加えて、STEP プログラムの目標として、

・科学的・技術的課題及び機会の性質
・科学及び技術的イシューを分析するために用いられる専門化された方法
・国家的・国際的な機関に関連する科学・技術のダイナミクス

のより深い理解を学生に提供することを掲げている。2005－06 年の大学院コースは、

以下のような構成になっている。

・WWS 556d　国際関係におけるトピックス：大量破壊兵器に対する防御

・WWS 582b　環境及び天然資源の経済学

・WWS 584　環境政策における科学の活用

・WWS 585b　STEP におけるトピックス：温室効果における暮らし：技術と政策

・WWS 586a/MOL 586　科学・技術及び環境政策におけるトピックス：バイオテクノロジー政策

・WWS 586b　絶滅寸前の種の保存と生態系

・WWS 586e　STEP におけるトピックス：リスク政策と規制

・WWS 594m　メンタル・ヘルス

・WWS 594b　政策分析：北朝鮮の核と米国政策の応答

・WWS 594n　政策分析：グローバル化と感染症

その他、米国において、研究開発や科学技術の分野に特に重点を置いてはいないが、公共政策の枠組みでプログラムを持っている大学としては、ジョージ・メイスン大学(George Mason University)、ロチェスター工科大学(Rochester Institute of Technology)、ラトガーズ大学(Rutgers, The State University of New Jersey)、カーネギーメロン大学公共政策学部(Carnegie Mellon University -H. John Heinz III School of Public Policy and Management)などがある。

また、技術経営管理など、自然科学・工学的アプローチが背景となっているプログラムを持つ大学としては、マサチューセッツ工科大学工学部工学システム科（技術・政策プログラム)、カーネギーメロン大学工学部工学・公共政策学科、メリーランド大学などがある。

9. 政策研究大学院大学

http://www.grips-ip.jp/index.html

政策研究大学院大学では、2004 年度から「科学技術・学術政策博士プログラム」(博士後期課程のみ）を開講している。これは、科学技術政策の課題について、歴史的、計量的、国際的、学際的に研究するとともに、高度な専門知識と深い洞察力に裏付けられた政策立案・遂行能力を持ち、国際的に活躍できる行政官や政策形成の理論と実践に通じた人材を養成することを目的するものである。

講義科目は、科学技術政策特論（必修)、特別セミナー（必修)、特別専門科目（必

修)、専門科目（選択）から構成される。科学技術政策特論は、科学技術政策全体の理解を深め、体系的に把握することを目指すものであり、2004 年度には表 5 にまとめた 14 テーマで実施されている。特別セミナーは秋学期から開講されるもので、学生の発表を中心に行い、研究テーマへの知見を深めることを目指すものである。特別専門科目は、指導教員による学生の専門に応じた議論を中心とした講義であり、専門科目は、各教員が開催するセミナー等である。学生には、指導教官と相談して、これらの科目から最低 8 単位履修することが求められる。

表6　科学技術政策特論の構成（2004 年度）

テーマ	講師
科学技術政策概論	中島邦雄（教授）
企業育成論	橋本久義（教授）
知的財産論	隅蔵康一（助教授）
経済学から見た科学技術政策	後藤晃（客員教授）
科学技術と社会	菱山豊（教授）
研究組織論	丸山瑛一（教授）
技術経営論	亀岡秋男（客員教授）
科学と政治　－科学技術政策過程論－	角南篤（助教授）
現代科学技術政策の課題	有本建男（客員教授）
国際機関における科学技術政策	井上正幸（客員教授）
技術分析	丹羽冨士雄（教授）
米国の科学技術政策	平野千博（連携教授）
科学技術政策史	中島邦雄（教授）
科学技術政策の展望	丹羽冨士雄（教授）

コースの特徴としては、個々の学生の専攻に応じて、複数の指導教員からなる「指導教員委員会（Advisors' Committee)」が指導を行う点にある。指導教員委員会は、学生の研究計画、履修状況等に応じて授業科目の履修について指導すると同時に、行政官の職にある者、専門職である者、研究者を指向する者など、学生の現職や将来の希望に応じた指導を実施している。

10. その他（MOT 大学院）

我が国においては、関連する領域として、以下のような MOT の専門大学院も開設されている。

- 芝浦工業大学大学院（工学マネジメント研究科工学マネジメント専攻）
  http://www.shibaura-it.ac.jp/shibaura-ma/index.html

- 東京理科大学大学院（総合科学技術経営研究科総合科学技術経営専攻）
  http://www.tus.ac.jp/grad/mot/
- 日本工業大学専門職大学院（技術経営研究科技術経営専攻）
  http://www.nit.ac.jp/senmon/index.html
- 日本大学大学院（グローバル・ビジネス研究科テクノロジー・マネジメント・コース）
  http://www.gsb.nihon-u.ac.jp/
- 北陸先端科学技術大学院大学（知識科学研究科技術経営（MOT）コース）
  http://www.jaist.ac.jp/ks/mot/index.html
- 立命館大学大学院（テクノロジー・マネジメント研究科）
  http://www.ritsumei.ac.jp/acd/gr/mot/index.html
- 早稲田大学大学院（アジア太平洋研究科国際経営学専攻 MOT プログラムテクノロジー・マネジメント専修）
  http://www.wiaps.waseda.ac.jp/

３－２. 学協会等が提供する研修プログラム

１. AAAS が提供するフェローシップ・プログラム

http://fellowships.aaas.org/

科学技術政策フェローシップは、全米各地の大学で科学技術分野の博士号または同等の学位を最近取得した若手の科学者が、ワシントン DC の科学技術政策の立案や意思決定を補助する目的で本格的には 1973 年に開始された。これまで 1,200 人の科学者や技術者がこのフェローシップを経験している。

志願者はアメリカ市民でなければならない。それぞれの科学技術情報や専門的知見を連邦議員や連邦政府に提供しつつ、自身も、政府がどのような仕組みで動くのかを学びつつ、科学技術の社会への影響を学ぶことができる。

フェローは、連邦議会、全米科学財団（NSF）、国立衛生研究所（NIH）、国務省（Department of State）、国防総省（DOD）、米国国際開発庁(USAID)、環境保護庁（EPA）、農務省（USDA）、食品・医薬局（FDA）、司法省（DOJ）、エネルギー省（DOE）などのオフィスに配属される。

このフェローシップには現在では毎年、約 90 人のフェローが選抜される。フェローシップは 9 月 1 日から始まり、それぞれの省庁に 1 年間派遣される前に、フェロー全員は 8 日間のオリエンテーションに参加する。また、年間を通じて科学技術と公共政策に関するセミナー、勉強会、パーティなど様々なイベントが頻繁に開催され、フェローの交流・情報交換の場となっている。

これらのフェローシップ・プログラムに加え、毎年 2 月半ばに開催される AAAS 年次大会では、フェローシップ卒業生の知的交流の場も設定されている。

AAAS フェローシップの種類と業務内容

＜科学技術政策フェローシップ＞

◆Congressional

フェローは、科学的・技術的インプットが必要な立法・政策分野における特別補佐としてキャピトル・ヒルで議会や議会委員会のメンバーとともに 1 年間過ごす。本フェローシップでは、類のない公共政策の学習経験を提供し、学術界と政府の相互作用の重要性を明示し、議会における意思決定過程に技術的バックグラウンドや外部の知見をもたらすようにデザインされている。募集人員は二人、年俸は 6 万 4 千ドル。

◆Diplomacy

フェローは、国務省の外交政策部門、もしくは国際開発庁（U.S. Agency for International Development）の国際開発部門、農務省の海外農業サービス部門、国立保健研究所フォガーティ国際センター（Fogarty International Center of National Institute of Health）の国際保健部門のいずれかで、科学・技術分野における国際問題に関して 1 年間働く。国務省は 10 名から 15 名、国際開発庁は 5 名から 10 名、農務省のプログラムは 1 名か 2 名、国立保健研究所は一人を採用予定。年俸は経験等のより 6 万 4 千ドルから 8 万 4 千ドルとなっている。

◆Defense & Global Security

フェローは、防衛政策や国家及び国際的なセキュリティ技術、システム分析に関する専門知識を外部から考察し、プロジェクトやプログラムの管理を支援する。フェローの配属先としては、国防総省（6 部局）、エネルギー省、国土安全保障省、Army Corps of Engineers' Institute for Water Resources がある。

◆Health, Education, & Human Services

フェローは、健康や環境保全、生物学的脅威、食料の安全性、教育・研究において主導するために、プログラムや政策、計画やリスク分析の改善をサポートすることが求められている。また、規制などの監視の展開・実施にも役立つことが期待されている。フェローの配属先としては、農務省、食品医薬品局、国立衛生研究所、国立科学財団がある。

◆Energy, Environment, & Natural Resources

フェローは、環境や健康を保護し、エネルギー問題に取り組み、空気や水、土壌、天然資源を守るために、プロジェクト・プログラム・政策やアウトリーチ・イニシャティブに従事する。フェローの配属先としては、農務省、エネルギー省、環境保護局、森林サービス、国立科学財団、Army Corps of Engineers' Institute for Water Resources がある。

◆Roger Revelle Fellows Program in Global Stewardship

1997 年に設置された本フェローシップでは、1 年間、国内ないし国際的な環境関連（主に生態系におけるヒトの相互作用）の分野で、議会や関連する行政府機関もしくはワシントンにある他の政策コミュニティ内の組織で働く。応募者は生物科学、物理科学、社会科学、各エンジニアリングや関連する学際領域のいずれかの分野の Ph.D を持ち、かつ最低 3 年間の学位取得後の専門的経験のあるアメリカ市民でなければならない。

＜科学技術政策以外のフェローシップ＞

◆Mass Media Science & Engineering Fellows Program

科学技術への公共の理解を高めることが AAAS の主な目標の一つであるが、本プログラムでは、20 人から 25 人のフェローに対し、夏に 10 週間にわたり、国内のラジオやテレビ、新聞、雑誌などのマスメディア機関にレポーターやエディター、制作アシスタントとして従事させる。フェローは週に 450 ドルの支給を受ける。

2. The National Academies が提供するフェローシップ・プログラム

http://www7.nationalacademies.org/policyfellows/

米国の National Academies が提供する Christine Mirzayan Science & Technology Policy Graduate Fellowship Program は、大学院の理学、工学、医学、獣医学、経営学及び法学の学生に対して、科学技術政策の創造を促すような政策研究に従事させ、科学技術と政策実務の相互作用を促進させることを意図している。プログラムを通じて学生は学術界で得られるものとは異なる重要な実務的能力を習得し、大学院生から実務的研究者（プロフェッショナル）への遷移をはかることになる。つまり、本プログラムは、AAAS のフェローシップ・プログラムが中堅研究者のキャリアー転換を促すのに対して、科学技術政策関係研究者の養成導入プログラムであるとともに、科学技術関係大学院在籍者のための副専攻研修という意味合いも持つ。

これまでに約 250 人の卒業生を送り出している。研修期間は 10 週間であり、2008 年には 2 回の研修期間が予定されている。また、配属先は図 2 のようになっている。

- 冬季：1 月 7 日から 3 月 14 日（10 週）
- 秋季：9 月 8 日から 11 月 14 日（10 週）

図2　Science & Technology Policy Graduate Fellowship Program における配属先

この 10 週のプログラムでは 4,800 ドルがフェローに対して支給される（ワシントン郊外に住んでいる場合にはさらに 500 ドルが追加される）。

フェローは配属先のシニアリサーチャーの下につき、最初の１週間目にガイダンスを受けることになる（オリエンテーション・ウィーク）。シニアリサーチャーはフェローのメンター（指導者）の役割を担い、フェローに対して NSF の役割や米国の科学技術政策の基礎を教示する。

プログラムでの次の活動は一連のセミナーをフェロー自身が企画・設計・開発することである。フェローは科学技術政策におけるホットなトピックスを３つほど設定し、セミナー企画案を作成する。どのようなテーマ類型か、講師として誰がふさわしいか等々、具体的な活動プランを作成する。そして、プログラムの終了１ヶ月前までにプランを実行する。この活動の目的は、科学技術コミュニティーにおける最先端の研究に対して理解を深めるとともに、将来、自らがプロデュースする研究活動に対する National Academy の委員会でのヒアリングやパネル討論の演習も兼ねることになる。

また、フェローにはプログラム以外の活動に対しても便宜が図られる。National Academy の理事との昼食会や現地見学、議会での公聴会やシンクタンクが開催するセミナーへの出席等、さまざま活動を通じて科学技術政策を理解する機会が与えられる。

３. PRIME が提供する人材育成プログラム

PRIME では、以下の３タイプの人材育成プログラムを提供している。

① PhD training
② Masters programme
③ Professional short courses

① PhD training

PhD training は、以下の６つの活動からなる。

- PhD Conferences
  2004 年及び 2005 年にサセックス大学 SPRU にて開催。2006 年には、コペンハーゲン・ビジネススクールにて開催。
- Summer Schools & Workshops
  2004 年にマドリードにて、2005 年にはブダペスト、2006 年には Ravenstein にて開催。
- Project Participation
  PRIME におけるすべての研究プロジェクトに PhD 学生が関与している。
- PhD Circulation

別の機関で 3 ヶ月から 9 ヶ月の間、研究する機会を提供し、相補的な専門知識と相互指導を学生にもたらす。

- Poster competitions
  最終学年の学生に対し、国際会議にて自分の業績を公開する場を提供し、欧州の PhD 市場での存在を高めさせる。2005 年にマンチェスターで、2006 年にパリ、同年にアトランタにて US-EU 会議を開催。
- PhD Days
  PRIME の博士課程の研究を紹介するために、大規模な国際会議にて PhD のプレゼンテーションを組織する。

2004-2005 年においては、25 カ国から約 120 人の PhD がこれらのイベントに参加した。

② Masters programme

2005 年には Inter-university professional Masters programme がスタートした。第 1 セメスターでは、学生は基本的なカリキュラム（科学技術の研究や政策、イノベーション分析など）を学ぶ。第 2 セメスターでは、別の大学に移って専門分野別のコースを学び、所属機関とホスト役機関の共同管理のもとで論文を書く。学生は、所属大学から Master の学位と一緒に PRIME の証明書も受け取る。15 の高等教育機関が初年度のパイロットプログラムに参加した。

③ Professional short courses

これまでに設置された Professional short courses としては、オランダ・トュウェンテ大学の Evaluation と UK・マンチェスター大学の Foresight が特筆される。PRIME では、2006 年における抱負として、現在提供しているものを拡張させることを想定している。具体的には、2008 年までに異なるトピックスに関して 5 つの年間コースを立ち上げることである。

以上、海外の学協会が提供する人材育成プログラムについてみてきたが、我が国おいては、次のような取り組みがある。

4. 政策評価相互研修会

http://www.ips.or.jp

「政策評価相互研修会」は、研究・技術計画学会の科学技術政策分科会及び研究評価分科会の両主査の呼びかけではじまったものであり、シンクタンクを含め研究者・実務家など関心のある主体に広く無料で開放し、協力して、我が国の新たなエキスパート層の人材育成・交流プラットフォーム機能を果たすモデルを構築しようと企画されたものであり、開催当初から財団法人政策科学研究所がそのノードとして継続的に企画・運営等の支援を行っている。2003 年度の途上からは、文部科学省との共催で文部科学省や関連機関の職員研修をかねるかたちで実施されている。行政、学界、シンクタンク業界などから、2003 度は 3 シリーズ計 12 回で延べ 620 名、2004 年度は計 6 回で延べ 440 名、2005 年度は計 3 回で延べ 214 名の参加者を得ている。

相互研修会でとりあげられたテーマは次の通りである。

【平成 15 年度】 ※ [ ]は、後述の文献一覧に対応

第 1 シリーズ：施策・プログラムの途上・追跡評価

第 1 回： 9 月 4 日（木）	1．イントロダクション ･･･平澤冷（東京大学名誉教授） [1]『研究開発プロジェクト等の評価手法に関する調査』 [2]『技術評価に係る評価人材の育成等に関する調査』 2．“Handbook of Research Impact Assessment”の概要紹介と検討 ･･･平澤冷 [3] Kostoff (Office of Naval Research) 3．“Evaluating Technology Programs: Tools and Methods”の概要紹介と検討 ･･･平澤冷 [8] Luke Georghiou and David Roessner
第 2 回： 9 月 10 日（水）	1．ビブリオメトリックス ･･･大野博教（電力中央研究所名誉研究顧問） [4] Barry Bozeman and Julia Melkers(ed.), ・Chapter3: pp.43-61 [6] Michel Callon, Philippe Laredo, and Philippe Mustar, ・Chapter10: pp.165-219 [7] OTA, ・Chapter3: pp.29-44 2．テクノメトリックス ･･･飯塚安伸（川鉄テクノリサーチ） [5] Gerhard Becker and Stefan Kuhlmann, ・PartⅢ: pp.285-308 3．ATP の評価 [8] Luke Georghiou and David Roessner, ･･･平澤冷 ・4.5: pp.669-671 [11] National Research Council ･･･鈴木潤（未来工学研究所）
第 3 回：	1．エコノメトリックス

9 月 17 日（水）	[10] Toolbox ･･･進藤秀夫（NEDO） ・Chapter3.3, 3.4, 3.5, 3.7: pp.82-130, 142-162 [4] Barry Bozeman and Julia Melkers(ed.), ･･･平澤冷 ・Chapter1: pp.1-16 [1]『研究開発プロジェクト等の評価手法に関する調査』 ･･･平澤冷 ・5.3: pp.71-74 ([9] PREST, et al., Chapter4.1 からの引用あり)
第 4 回: 9 月 27 日（土）	1．指標法 ･･･服部健一（ﾛｰﾗﾝﾄﾞ･ﾍﾞﾙｶﾞｰ･ｱﾝﾄﾞ･ﾊﾟｰﾄﾅｰ･ｼﾞｬﾊﾟﾝ） [10] Toolbox ・Chapter3.2: pp.71-81 [5] Gerhard Becker and Stefan Kuhlmann, ・PartⅢ: pp.243-284 2．コントロールグループ・アプローチ ･･･馬場敏幸（東大） [10] Toolbox ・Chapter3.6: pp.131-141 [6] Michel Callon, Philippe Laredo, and Philippe Mustar, ・Chapter15: pp.311-322 3．ベンチマーク ･･･新野聡一郎（三井情報開発） [10] Toolbox ・Chapter3.12: pp.201-209 4．NSF Engineering Program ･･･大熊和彦（政策科学研究所） →第 4 回（9 月 27 日）に延期 [12] Roessner, et al. [13] Roessner, et al.
	4．ロジック評価法 ･･･進藤秀夫 [10] Toolbox ・Chapter3.9: pp.173-182 [9] PREST, et al., ・Chapter4.3: pp.294-313 [4] Barry Bozeman and Julia Melkers(ed.), ・Chapter2: pp.17-42 5．システム評価法 ･･･平澤冷 [10] Toolbox ・Chapter3.10: pp.183-191 [6] Michel Callon, Philippe Laredo, and Philippe Mustar, ・Chapter19: pp.361-383 ・Chapter20: pp.385-429 [4] Barry Bozeman and Julia Melkers(ed.), ・Chapter7: pp.123-153 6．EU の Framework Program ･･･平澤冷 ◇ソシオメトリックス ◇ポートフォリオ [9]PREST, et al., ・Chapter3: pp:146-225

第 2 シリーズ：プロジェクトの事前評価

第 1 回: 11 月 26 日（水）	「制度設計」 1．イントロダクション ･･･平澤冷 ・第 1 シリーズの総括及び第 2 シリーズの概要とポイント 2．NSF における制度設計 ･･･進藤秀夫 [21] National Academy of Public Administration(2001) 3．EU における制度設計 ･･･大熊和彦 [22] Manual of Proposal Evaluation Procedures(2002.3) 4．利益相反とプロジェクト事前評価等の制度設計 ･･･伊地知寛博（NISTEP）

第 2 回: 12 月 4 日(木)	「ピアレビューとエキスパートレビューの質的改善」 1. イントロダクション ・・・平澤泠 2. Bozeman(1993)の概要紹介と検討 ・・・大熊和彦 [4] Barry Bozeman and Julia Melkers(ed.), ・Chapter5: pp.79-98 3. Toolbox(2002.8)の概要紹介と検討 ・・・加藤知彦 (NEDO) [10] Toolbox ・Chapter3.8: pp.163-172 4. Guston(2003)の概要紹介と検討 ・・・鈴木潤 [20] Guston ・Chapter6: pp.81-97 5. ディスカッション－質的改善に向けてのチェックリスト
第 3 回: 12 月 12 日(金)	「独立行政法人等の機関評価」 1. イントロダクション／我が国の特殊課題(独法) ・・・平澤泠 2. 大学の研究評価の動向 ・・・林隆之(大学評価・学位授与機構) ・ 英国RAEレビュー報告／他内外の事例 [23] Review of research assessment 3. フランスにおける独法評価 ・・・宮崎久美子(東工大) [24]Comite National d'Evaluation(2002)
第 4 回: 12 月 20 日(土)	「大型プロジェクトの事前評価－コストと契約管理－」 ・・・江崎通彦(朝日大学) 1. 大型プロジェクトの事前評価における知識を知恵にかえる方法 2. 我が国防衛庁などの事例 3. 日米における契約管理からの示唆 4. 全体討議

第 3 シリーズ:政策・施策レベルの評価

第 1 回: 3 月 19 日(金)	「政策評価の概要と課題」 1. 政策レベルの評価の概要 …平澤泠 2. 海外における政策評価の現状 ・ 欧州での取り組み－UK、オランダおよびEUにおける概況とモデルの重要性 …野呂高樹(政策科学研究所) ・ 政策決定のための事前評価的な分析—NOAA における気候変動研究の費用対効果分析 …磯野哲郎(海洋科学技術センター横浜研究所) ・ 米国での取り組み－GPRA から PART へ …遠藤悟(日本学術振興会) ・ 米国ワシントン研究評価ネットワーク会合(2003 年 12 月)におけるPARTに係る議論の概要について …進藤秀夫
第 2 回: 3 月 25 日(木)	「研究及びイノベーション政策の評価とそのインパクト」 …Luke Georghiou (Executive Director, PREST, University of Manchester, UK)
第 3 回: 3 月 26 日(金)	「イノベーション政策及びイノベーション・システムの評価」 …Svend Otto Remoe (Senior Researcher, STEP, Norway) 1. イノベーション・システムのアセスメント 2. イノベーション・ポリシーの評価事例と課題
第 4 回: 3 月 27 日(土)	「PART法(Program Assessment Rating Tool)を用いた研究開発プログラムの評価 …David Trinkle (Program Examiner, Science and Space Programs Branch, Office of Management and Budget, USA)

※ 研修会でとりあげた文献一覧

[1] 平成 13 年度経済産業省委託調査『研究開発プロジェクト等の評価手法に関する調査』(財)政策科学研究所(2002.3)

[2] 平成 13 年度経済産業省委託調査『技術評価に係る評価人材の育成等に関する調査』(財)政策科学

研究所(2002.3)
[3] Kostoff (Office of Naval Research), “Handbook of Research Impact Assessment,” Edition7., NTIS, (1997)
http://www.onr.navy.mil/sci_tech/special/354/technowatch/docs/handweb.doc
[4] Bozeman,Barry (ed.), Melkers,Julia (ed.), “Evaluating R&D Impacts: Methods and Practice”, Kluwer Academic Publishers (1993)
[5] Gerhard Becker and Stefan Kuhlmann, “Evaluation of Technology Policy Programmes in Germany,” Kluwer Academic Publishers (1995)
[6] Callon,Michel, Laredo,Philippe, Mustar,Philippe, “The strategic management of research and technology”, Editions Economica (1997)
[7] OTA, “Research Funding as an Investment: Can We Measure the Returns? -A Technical Memorandum”, NTIS (1986.4)
http://www.wws.princeton.edu/~ota/disk2/1986/8622_n.html
[8] Georghiou,Luke, Roessner,David, “Evaluating technology programs: tools and methods”, *Research Policy*, Vol.29, 657-678 (2000)
http://ideas.repec.org/a/eee/respol/v29y2000i4-5p657-678.html
[9] PREST, AUEB, BETA, ISI, Joanneum Research, IE HAS, Wise Guys, “Assessing the Socio-economic Impacts of the Framework Programme”, (2002.6)
http://les.man.ac.uk/PREST/Download/ASIF_report.pdf
[10] Joint Research Centre-IPTS and Joanneum Research, “RTD Evaluation Toolbox,” European Commission, (2002.8)
http://epub.jrc.es/docs/EUR-20382-EN.pdf
[11] National Research Council, “The Advanced Technology Program: Assessing Outcomes”, National Academy Press (2001)
[12] David Roessner, et al., “The Role of NSF's Support of Engineering in Enabling Technological Innovation,” SRI International, (1997)
http://www.sri.com/policy/csted/reports/techin/contents.html
[13] David Roessner, et al., “The Role of NSF's Support of Engineering in Enabling Technological Innovation, PhaseⅡ,” SRI International, (1998)
http://www.sri.com/policy/csted/reports/sandt/techin2/contents.html
[14] Ken Guy, “Strategic Options for the Evaluation of the R&D Programmes of the European Union”, Technopolis Ltd. (1998.11)
http://www.europarl.eu.int/stoa/publi/167406/default_en.htm
[15] ロバート K. イン『ケース・スタディの方法－第 2 版』千倉書房(1996)
[16] 林知己夫編『社会調査ハンドブック』朝倉書店(2002)
[17] 田尾雅夫・若林直樹『組織調査ガイドブック』有斐閣(2001)
[18] 小池和男『聞きとりの作法』東洋経済新報社(2000)
[19] 大谷伸介・木下栄二他『社会調査へのアプローチ』ミネルヴァ書房(1999)
[20] David H. Guston, “The expanding role of peer review processes in the United States,” in Philip Shapira and Stefan Kuhlmann (ed.), *Learning from Science and Technology Policy Evaluation –Experiences from the United States and Europe*, pp.81-97, Edward Elgar, 2003
[21] National Academy of Public Administration, “A Study of the National Science Foundation’s Criteria for Project Selection,” 2001.2
http://209.183.198.6/NAPA/NAPAPubs.nsf/0/ca5e12fc2518f12a85256a45004c5e8f/$FILE/nsfreport.pdf
[22] THE FIFTH FRAMEWORK PROGRAMME, “Manual of Proposal Evaluation Procedures,” 2002.3.4
http://www.anst.uu.se/andejons/em_en_200001.pdf
[23] Sir Gareth Roberts, “Review of research assessment –Reported by Sir Gareth Roberts to the UK funding bodies-,” 2003.5
http://www.ra-review.ac.uk/reports/roberts.asp
[24] Comité National d’évaluation, “Repères pour l’évaluation Rapport au Président de la République,” 2002
http://www.cne-evaluation.fr/WCNE_pdf/RapPdt2002.pdf

【平成 16 年度】

「新たに取り組むべき評価カテゴリーをめぐって」

日程	プログラム
第 1 回 2 月 2 日（水）	1.「新たに取組むべき評価カテゴリーをめぐって」 ・・・平澤泠東京大学名誉教授 2.「政策・施策レベルの評価の枠組みと日米欧における進展状況」 ・・・平澤泠 ・ 日 /平澤,「我が国の公共部門における研究開発評価の課題」, 研究 技術計画, Vol. 17, No. 3/4, pp.128-141 (2002). ・ 米 / Susan E. Cozzens, “Frameworks for evaluating S&T policy in the United States,” in Shapira et al. eds. (2003) ・ 欧 /Luke Georghiou, “Evaluation of research and innovation policy in Europe,” in Shapira et al. eds. (2003) 3.「EU のフレームワーク・プログラムの評価から」・・・大熊和彦（政策科学研究所） ・ Terttu Luukkonen, “ Challenges for the evaluation of complex research programmes,” in Shapira et al. eds. (2003) ・ Ken Guy, “Assessing RTD program portfolios in the European Union,” in Shapira et al. eds. (2003) 4. 質疑
第 2 回 3 月 8 日（火）	テーマ「欧米における政策・施策レベルの評価事例」 講師：遠藤悟（日本学術振興会）、鈴木潤（未来工学研究所）、馬場敏幸（法政大学）、林隆之（大学評価・学位授与機構） 紹介文献： （鈴木氏担当） ・ Louis G. Tornatzky, “Benchmarking university-industry relationships: a user-centered evaluation approach,” in Shapira et al. eds. (2003) ・ Laurent Bach, et al, “Evaluation of the BRITE / EURAM program,” in Shapira et al. eds. (2003) （馬場氏担当） ・ OECD (2002), “Benchmarking Industry-Science Relationships.” （林氏担当） ・ PREST (2002), ”Comparative Analysis of Public, Semi-Public and Recently Privatised Research Centres.” （遠藤氏担当） ・ James S. Dietz, “Factors affecting technology transfer in industry-US federal laboratory partnerships,” in Shapira et al. eds. (2003)
第 3 回 3 月 11 日（金）	テーマ「NSF における研究評価の枠組みと新手法」 講師：James S. Dietz （Program Director, Division of Research, Evaluation and Communication, NSF） - Research Evaluation at the National Science Foundation - Research Evaluation: Capacity approaches and new models for evaluation and organizational learning
第 4 回 3 月 18 日（金）	テーマ「EU におけるプログラムのメタ評価」 講師：Ken Guy （Director, Wise Guys Ltd.）

第 5 回 3 月 28 日（月）	テーマ「我が国における独立プロジェクト／プログラムレベルの評価事例をめぐる検討」 講師：佐野浩（NEDO）、鈴木隆（JST）、角田英之（海洋研究開発機構）、 柳澤和章（日本原子力研究所）
第 6 回 3 月 30 日（水）	テーマ「RAND における政策評価の概要と最近の事例」 講師：Anny Wong （Associate Political Scientist, RAND Corporation）

【平成 17 年度】

「実績の把握を中心にして」

日程	プログラム
第 1 回 3 月 6 日（月）	1. 導入「アウトカム概念の整理とその構造的把握」 ・・・平澤 泠（東京大学名誉教授） 2. 話題提供と意見交換「アウトカム指標の事例」 ・・・田原　敬一郎（（財）政策科学研究所） 3. 話題提供と意見交換「アウトカム類似概念としてのアディショナリティ分析事例」 ・・・川島啓（（財）政策科学研究所）
第 2 回 3 月 16 日（木）	1.「ヨーロッパにおける RTD 政策の評価」 ・・・Dr. Erik Arnold (Managing Director, Technopolis ltd., UK) 2. 質疑　・・・司会：平澤　泠（東京大学名誉教授）
第 3 回 3 月 28 日（火）	1.「研究開発法人におけるアウトカムの捉え方」 ・・・内田理之（理化学研究所　経営企画部　評価推進課） 2.「産業技術総合研究所の研究評価制度」 ・・・澤田　美智子（産業技術総合研究所　評価部） 3.「NEDOにおける追跡調査・評価と成果の把握の検討」 ・・・弓取修二（新エネルギー・産業技術総合開発機構研究評価部） 4.「JST戦略的創造研究推進事業における評価について」 ・・・佐々正（科学技術振興機構戦略的創造事業本部基礎研究制度評価ﾀｽｸﾌｫｰｽ） 5. 全体討議　・・・司会：平澤　泠（東京大学名誉教授）

【平成 18 年度】

日程	プログラム
第 1 回 10 月 10 日（火）	1. 課題別調査研究シリーズ（第 1 回）：評価人材 (1)評価人材―海外主要国における養成・集積状況と事例的取り組み ・・・平澤　泠（東京大学名誉教授） (2)海外主要国における評価人材養成のための基盤整備 ・・・田原　敬一郎（政策科学研究所研究員）

	(3)我が国の評価人材養成・集積上の課題と対応策 ･･･平澤　泠(東京大学名誉教授) 2. 海外国際会議報告(第 1 回) (1)2006 年 G8 研究開発評価ワーキンググループ会合について ･･･後藤　裕(文部科学省評価推進室室長) (2)Blue Sky II 2006 報告 ･･･富澤　宏之(科学技術政策研究所科学技術基盤調査研究室長)
第 2 回 10 月 26 日(木)	1. フロンティア紹介シリーズ(第 1 回) Frontiers of Evaluation: Context and 2 Cases for Application: Options (US), Networks (EU) ･･･Nicholas S. Vonortas (Professor, George Washington University, USA) 2. 全体討議　･･･司会:平澤　泠(東京大学名誉教授)
第 3 回 11 月 29 日(木)	1. 実績検討シリーズ(第 1 回) (1)科学技術振興機構における追跡調査・評価の取り組み－イノベーション創出に向けた目的基礎研究から応用・実用化研究への橋渡しに関するケーススタディ ･･･吉田　秀紀(科学技術振興機構戦略的創造事業本部基礎研究制度評価タスクフォース) (2)全体討議　･･･司会:平澤　泠(東京大学名誉教授) 2. 海外国際会議報告(第 2 回) (1)アメリカ評価学会報告　･･･中村　修(経済産業省技術評価調査課) (2)OECD 会合報告　･･･小林　直人(産業技術総合研究所評価部)
第 4 回 12 月 22 日(金)	1. 実績検討シリーズ(第 2 回):追跡調査・評価の取り組み(経済産業省の事例) (1)事例報告:光関係(レーザー加工・計測)研究開発プロジェクトの追跡評価 ･･･小池　勲(三菱総合研究所産業・市場戦略研究本部産業戦略グループ主任研究員) (2)事例報告:マイクロマシン研究開発プロジェクトの追跡評価 ･･･田村　信一(日鉄技術情報センター調査研究事業部長・主席研究員) (3)経済産業省における追跡調査・評価を巡る論点 ･･･菊池純一(青山学院大学法学部教授) (4)全体討議　･･･司会:平澤　泠(東京大学名誉教授)
第 5 回 1 月 25 日(木)	1. 実績検討シリーズ(第 3 回):レビューパネルの質的向上 (1)レビューパネルの質的向上を巡る論点　･･･平澤　泠(東京大学名誉教授) (2)物質・材料研究機構におけるサイエンスベース研究評価の現状について ･･･野田　哲二(物質・材料研究機構理事) (3)日本学術振興会事業におけるプログラムオフィサー制度の果たす役割等について－科研費を中心とした審査・評価システムの構築－ ･･･山本　一彦(日本学術振興会学術システム研究センター主任研究員／東京大学大学院医学系研究科教授)

	(4)全体討議　　・・・司会：平澤　泠（東京大学名誉教授）
第 6 回 2 月 26 日（月）	1．フロンティア紹介シリーズ（第 2 回） (1)Research and Education in Science, technology and Innovation Management and Policy at the University of Manchester (2)Lead markets as an instrument of innovation policy ・・・Luke Georghiou（Executive Director, PREST, University of Manchester, UK） (3)全体討議　　・・・司会：平澤　泠（東京大学名誉教授）
第 7 回 3 月 14 日（水）	1．フロンティア紹介シリーズ（第 3 回） (1)SciSIP's Relationship to PART ・・・Kaye G. Husbands（Science Advisor for Science of Science Policy, National Science Foundation (NSF), USA） (2)Evaluation in the Congressional Deliberation Process of the U.S. Science & Technology Budget ・・・Gerald J. Hane（Managing Director, Q-Paradigm, USA） (3)全体討議　　・・・司会：平澤　泠（東京大学名誉教授）

参考情報4. 研究開発評価関連学協会誌

研究開発評価関連学協会誌は、評価に関する専門的知識を流通する上で重要な機能を担っている。これらのジャーナルを通じて、新しい研究開発評価やマネジメントの手法について、経験的な研究、理論的研究が進められ、それらの知識を必要としている人に情報が提供されている。

1. Research Evaluation

Research Evaluation は、査読付き研究評価専門国際ジャーナルである。個別の研究プロジェクトから、研究パフォーマンスの地域間比較まで扱う。研究プロジェクト、研究者、研究センター、研究成果のタイプなどすべてカバーしている。公共・私的セクター、自然・社会科学なども含んでいる。「evaluation」という言葉は、申請書のランク付けから、現在進行中のプロジェクトやプログラムのモニタリング、研究成果の活用まですべてのフェーズを取り扱っている。Research Evaluation は、世界中の大学、政府、リサーチカウンシル、コンサルタントに広く普及している。研究者、資金配分機関、研究成果の利用者などの研究・評価に関するさまざまな主体を対象としている。アプローチ、思想、質的、量的などの制限はない。

発行：毎年 4 月、8 月、12 月の年 3 回発行
ISSN: 0958-2029(Print), ISSN: 1471-5449 (Online)
編者：Anthony van Raan 教授（蘭ライデン大学)、Susan Cozzens (米 GIT)
評編者：Michael Keenan (英マンチェスター大学)
出版社：Beech Tree Publishing

最近の掲載論文タイトル：
＜Volume 14, Number 3, December 2005 ＞
・Advantages and dangers of `remote' peer evaluation
・Handcrafted by 16 men: The impact of single and multiple authorship in collaborative research networks
・R&D evaluation in Italy: more needs to be done
・Ex ante impact assessment for research on natural resources management: methods and application to aquatic resource systems
・The dark side of mobility: negative experiences of doing a postdoc period abroad
・Taxonomy for science and engineering indicators: a reassessment

＜Volume 14, Number 2, August 2005 Special issue on the Eighth International Conference on Science and Technology Indicators, part 2 ＞

・Analysis of Spanish scientific output following the Joint Action Program (Acciones Integradas) of the Ministry of Science and Technology (MCYT)
・Key labs and open labs in the Chinese scientific research system: qualitative and quantitative evaluation indicators
・Bibliometric indicators at the micro-level: some results in the area of natural resources at the Spanish CSIC
・Differences over a decade: high tech capabilities and competitive performance of 28 nations
・Comparing and evaluating public research organisations: a unique, participatory mechanism in place in France
・Regionalisation of science and technology data in Spain
・The social accountability reporting project at Elettra
・German medical faculties in the 1990s: on-line bibliometric analysis
・Cross-disciplinary research: co-evaluation and co-publication practices of the CNRS laboratories
・Developing indicators to measure technology institutes' performance

＜Volume 14, Number 1, April 2005 Special issue on the Eighth International Conference on Science and Technology Indicators, part 1 ＞

・Tracking knowledge diffusion through citations
・Committee peer review at an international research foundation: predictive validity and fairness of selection decisions on post-graduate fellowship applications
・S&T culture: a blooming dimension
・Is external research funding a valid indicator for research performance?
・Impact of socio-economic factors on higher education in Russia
・Challenges in developing gender-sensitive indicators for Finnish researcher training
・Quantitative method and model for forecasting R&D expenditures in China
・The employment of PhDs in firms: trajectories, mobility and innovation
・Why do academic scientists engage in interdisciplinary research?
・Job advertisements as an indicator for mobility of researchers: Naturejobs as a case study

2. Scientometrics

Scientometrics は、科学の科学、科学コミュニケーション、科学政策に関する計量の視点からの国際ジャーナルであり、開発と科学のメカニズムに関する計量方法を対象とするものである。

原著論文の出版、短信、速報、レビュー論文、編者への手紙、書評が掲載され、計量的に見た科学の特徴に関する研究成果を取り扱う。特に、科学の開発とメカニズムについての統計的、数学的手法による調査に重点が置かれている。また、この分野と関連分野の国際会議やイベントに関する最新の重要な情報を読者に提供している。目録については別刷りで提供される。本誌は学際的な特徴を持つため、読者は、中央の科学機関、省庁、研究機関、研究所の図書館員等、多岐にわたる。

ISSN: 0138-9130 (Print), ISSN: 1588-2861 (Online)
出版社：Springer Netherlands
Published in cooperation with Akadémiai Kiadó
共同出版社：Akadémiai Kiadó（ハンガリー、ブダペスト）
編集長：T. Braun, Hungary
編集者：A. Schubert, Hungary
共編者：W. Glänzel, Hungary
名誉編者：M. T. Beck, Hungary; E. Garfield, USA; M. Orbán, Hungary

最近の掲載論文タイトル：

＜Issue: Volume 67, Number 1 April 2006 ＞

- Key Labs and Open Labs in the Chinese scientific research system: Their role in the national and international scientific arena
- Analyzing the association between referees' recommendations and editors' decisions
- Organizational vs. personal social capital in scientists' performance: A multi-level network study of elite French cancer researchers (1996-1998)
- Measuring internationality: Reflections and perspectives on academic journals
- Science in Brazil. Part 1: A macro-level comparative study
- Science in Brazil. Part 2: Sectoral and institutional research profiles
- Some practical aspects of fitting and testing the Zipf-Mandelbrot model: A short essay
- Creative knowledge environments for research groups in biotechnology.
- The influence of leadership and organizational support in universities and business companies

・A dense network sub-grouping algorithm for co-citation analysis and its implementation in the software tool Sitkis

＜Issue: Volume 66, Number 3 Date:　February 2006 ＞

・Law of cumulative advantages in the evolution of scientific field
・Indicators of failed information epidemics in the scientific journal literature: A publication analysis of Polywater and Cold Nuclear Fusion
・Publications resulting from Spanish radiology meeting abstracts: Which, Where and Who
・Issues in measuring the degree of technological specialisation with patent data
・Assessing the foreign control of production of technology: The case of a small open economy
・The impact of impact factor on small specialties: A case study of family medicine in Taiwan
・Scientific evaluations of citation quality of international research articles in the SCI database: Thailand case study
・Proof of a conjecture of Moed and Garfield on authoritative references and extension to non-authoritative references
・Science and technology policy in Turkey. National strategies for innovation and change during the 1983-2003 period and beyond
・Bibliometric analysis - A new business area for information professionals in libraries?: Support for scientific research by perception and trend analysis
・Towards a European economics of economics: Monitoring a decade of top research and providing some explanation
・Another ISI idiosyncrasy

3. Impact Assessment and Project Appraisal（IAPA）

Impact Assessment and Project Appraisal は、International Association for Impact Assessment (IAIA)の国際的な査読付きのジャーナルである。環境・社会・健康その他のインパクトアセスメントや費用便益分析、技術アセスメントなどをカバーしている。読者は大学、政府、公的機関、コンサルティング会社、NGO などであり、100 カ国を超える。editorials, main articles, book reviews, and a professional practice section から構成されている。

ISSN: 1461-5517 (Print), ISSN: 1471-5465 (Online)
出版社：Beech Tree Publishing
発行：毎年 3 月、6 月、9 月、12 月の年 4 回
編集者：Angus Morrison-Saunders（Murdoch University, Australia）、Maria Partidário（New University of Lisbon, Portugal）
副編集者：Technology assessment: Sharon Jones （Lafayette College, USA）、Law: Simon Marsden（Environment Agency, UK）、Health impact: Balsam Ahmad（Liverpool School of Tropical Medicine, UK）、Social impact assessment: Michael D Smith（Humboldt State University, USA）

最近の掲載論文タイトル：

＜Volume 24, Number 1, 1 March 2006＞

- Round table: Common sense in environmental impact assessment: it is not as common as it should be
- Evidence-based policy-making in Europe: an evaluation of European Commission integrated impact assessments
- Strategic environmental assessments for genetically modified organisms
- Uncertainty in environmental impact assessment predictions: the need for better communication and more transparency
- Environmental impact assessment in sub-Saharan Africa: the Gambian experience
- Characterizing environmental impact statements for road projects in North Carolina, USA

＜Volume 23, Number 4, 1 December 2005＞

- Impact mitigation in environmental impact assessment: paper promises or the basis of consent conditions?
- In search of arenas for democratic deliberation: a Habermasian review of environmental assessment
- Applying sustainability assessment models
- Mission impossible: does environmental impact assessment in Denmark secure a holistic approach to the environment?
- The place of strategic environmental assessment in the privatised electricity industry

4. RESEARCH POLICY

Research Policy は、研究政策、研究管理、研究計画を扱う学際的ジャーナルであり、主な対象領域は次のように非常に多岐にわたる。

能力と権限、起業家、進化経済学、シュンペーター経済学、産業クラスター、イノベーションのマネジメント・政策・戦略、知識（創造、移転、開発）、システムイノベーション（国内、地域、セクター）、組織としての学習、実験法、問題解決法、製品・プロセス開発、研究開発マネジメント、研究開発、研究政策、科学政策、技術マネジメント・政策・戦略.

掲載される論文は、これらの活動と、経済・社会・政治・制度との関係について考察したものであり、その多くは経験的なものだが、理論的なものも含んでいる。学術的な分析者が書いたものもあれば、研究開発とイノベーションのプロセスの実務家が書いたものもある。ジャーナルは世界を視野に入れ、学術界、産業界、政府関係者、このテーマに興味のある人々を対象にしている。社会科学のジャーナルの中での impact factor の高さが、学術的なステータスを示している。

編者：M. Bell, M. Callon, H. Grupp, F. Kodama, S. Kuhlmann, B. Martin, W.W. Powell, S. Thomke, N. Von Tunzelmann
ISSN: 0048-7333
印刷： ELSEVIER

最近の掲載論文タイトル：
＜Volume 35, Issue 2, Pages 181-342 (March 2006)＞
・Measuring the knowledge base of an economy in terms of triple-helix relations among ‘technology, organization, and territory’
・Ownership matters: Intellectual Property, privatization and innovation
・The fruit flies of innovations: A taxonomy of innovative small firms
・Does technological diversification promote innovation?: An empirical analysis for European firms
・Beyond known uncertainties: Interventions at the fuel–engine interface
・Innovation and productivity in developing countries: A study of Argentine manufacturing firms’ behavior (1992–2001)
・ Conceptualising the heterogeneity of research-based spin-offs: A multi-dimensional taxonomy
・Factors affecting university–industry R&D projects: The importance of searching, screening and signalling

・Comparing firms' triadic patent applications across countries: Is there a gap in terms of R&D effort or a gap in terms of performances?

＜Volume 35, Issue 1, Pages 1-180 (February 2006)＞
・Exploration and exploitation in innovation systems: The case of pharmaceutical biotechnology
・The role of corporate scientists in innovation
・Faculty support for the objectives of university–industry relations versus degree of R&D cooperation: The importance of regional absorptive capacity
・Internationalization of innovation systems: A survey of the literature
・Competition and innovation behaviour
・The emergence of China as a leading nation in science
・Marketing/R&D integration in the pharmaceutical industry
・Characterizing the technology firm: An exploratory study
・Socio-political factors and the failure of innovation policy in Croatia as a country in transition
・Corporate governance and innovation: The UK compared with the US and 'insider' economies

5. Research・Technology Management

Research・Technology Management(RTM)は、Industrial Research Institute が1958年から発行している査読付きの隔月刊のジャーナルである。技術イノベーションに関する研究開発から製品開発、マーケティングまでのピアレビューを受けた記事を掲載している。

発行人：Edward Bernstein
編者：Michael F. Wolff
ISSN: 0895-6308 (Print), ISSN: 1930-0166 (Online)

最近の掲載論文タイトル：
＜Volume 49, Number 2, March-April 2006＞
・Netherlands "Technopole" Takes Open Innovation To Next Stage
・Innovation Rules!
・Lessons Learned from Six Sigma in R&D
・Corporate Research and Venture Capital Can Learn from Each Other

・Planning Your Firm's R&D Investment
・Building Collaborative Innovation Capability
・Strategies for Global R&D
・Tom Tries "Rank-and-Yank" Appraisal

＜Volume 49, Number 1, January-February 2006＞
・Southeast Asia Drives for Biotech Supremacy
・Strategic Accounting for R&D
・Industrial Research Institute's R&D Trends Forecast for 2006
・Hardball Innovation
・Using Real Options Discipline for Highly Uncertain Technology Investments
・Implementing Concurrent Engineering
・Evaluating R&D with "First Bounce–Last Bounce" Framework
・Allocating Patent Rights in Collaborative Research Agreements
・Is There a Best Way to Lead Scientists and Engineers?

6. JOURNAL OF ENGINEERING AND TECHNOLOGY MANAGEMENT

The Journal of Engineering and Technology Management (JET-M)は、技術、イノベーション、工業管理に関する理論と実践に関する学術査読付きの国際ジャーナルである。このジャーナルは、工業、科学、マネジメントの分野と連携している。組織の活動目標・戦略目標を達成するための技術的能力の計画、開発、実施について取り扱っている。研究開発管理だけでなく、技術関連の組織的なマネジメント問題全般についてカバーしている。新製品開発、人材管理、イノベーションプロセスマネジメント、プロジェクトマネジメント、技術的融合、マーケティング、技術予測、戦略立案などを含んでいる。

このジャーナルは、技術と、研究開発、マーケティング、製造、管理といった企業機能とのインターフェイス機能を提供している。究極的な目的は、技術、イノベーション、工業マネジメントに関する学術研究の出版のための先導的なフォーラムを提供することにより、理論開発、研究、実践に貢献することである。

編集長：Michael K. Badawy
ISSN: 0923-4748
印刷：ELSEVIER

最近の読まれている論文タイトル：

・Future management research directions in nanotechnology: A case study
・The role of social and intellectual capital in achieving competitive advantage through enterprise resource planning (ERP) systems
・Information systems, strategic flexibility and firm performance: An empirical investigation
・Innovation strategies of Asian firms in the United States
・Task partitioning in new product development teams: A knowledge and learning perspective
・CEO compensation and firm competitive behavior: Empirical evidence from the U.S. pharmaceutical industry
・The constituents of core competencies and firm performance: evidence from high-technology firms in china
・Toward a model of the effective transfer of scientific knowledge from academicians to practitioners: qualitative evidence from the commercialization of university technologies
・Transferring R&D knowledge: the key factors affecting knowledge transfer success
・The human side of radical innovation
・Technological change and the technology intelligence process: a case study
・Knowledge transfer and R&D in pharmaceutical companies: A case study
・Sustaining growth in the modern enterprise: A case study
・Coordination in collaborative manufacturing mega-networks: A case study
・Theories of organizational structure and innovation adoption: the role of environmental change
・A research agenda to reduce risk in new product development through knowledge management: a practitioner perspective
・Out of machine age?: complexity, sociotechnical systems and actor network theory
・The role of infrastructure practices in the effectiveness of JIT practices: implications for plant competitiveness
・Virtuality, communication, and new product team creativity: a social network perspective
・Technological learning, knowledge management, firm growth and performance: an introductory essay
・When teamwork really matters: task innovativeness as a moderator of the teamwork-performance relationship in software development projects

・Innovation through initiatives-a frameworkfor building new capabilities in public sector research organizations
・A tentative framework for analyzing integration in collaborative manufacturing network settings: a case study
・The meaning of success: network position and the social construction of project outcomes in an R&D lab
・Success, failure and organisational competence: a case study of the new product development process

我が国においても、以下のような学協会誌がある。

7.「研究技術計画」

「研究技術計画（The Journal of Science Po1icy and Research Management）」は、研究・技術計画学会が季刊で発行している学会誌であり、内容は、論文、事例研究、総説、論説、書評、重要文献紹介、調査研究、調査研究レポート・カタログで構成されている。2006 年 3 月現在で、Vol.20 No.3 まで発行されている。

研究開発評価に関しては、近年では 2002 年に発刊された Vol.17 No. 3-4 において「公的資金による研究開発の評価」と題する特集が組まれている。

参考情報5. 研究開発評価関連主要文献・基本資料

評価人材にとって必読の主要文献・基本資料を、性格別にとりあげた。なお、いくつかについてはウェブサイトでダウンロード可能である。

1. ハンドブック、ツールボックス、総説

Bozeman,Barry (ed.), Melkers,Julia (ed.), "Evaluating R&D Impacts: Methods and Practice", Kluwer Academic Publishers (1993)

本書は研究評価に用いられる方法論と実際の評価事例について、米国及びカナダの著者らによって 1993 年に書かれたものである。そのため、近年の米国の動向（GPRA による影響や ATP 評価など）や欧州における評価の動向を把握するには向かない反面、ピアレビュー、事例研究、ビブリオメトリックス、企業における経済性評価、オペレーションズ・リサーチ手法などの基盤的方法が順に解説されており、入門書として適している。方法論を紹介した幾つかの章では、主にプロジェクトレベルの評価について、それぞれの方法論の内容と長所・短所を説明しており、いずれにおいても方法論には限界が存在するために他の複数の方法と組み合わせて用いることが推奨される。また、実際の評価事例を報告した複数の章では、定量的手法が実際には企業・政府において使用頻度が低いこととその理由の調査結果、公的研究所の産業との関係の評価、スピンオフ効果の評価、ポートフォリオなどが議論される。

Callon,Michel, Laredo,Philippe, Mustar,Philippe, "The strategic management of research and technology", Editions Economica (1997)

本書は EU の評価研究者のネットワークにより書かれたものであり、出版時点での欧州での研究評価の最新動向を示したものとなっている。本書では特に研究開発プログラムの評価に焦点がおかれている。最初の部分で欧州各国及び EU の研究評価システムの特徴が紹介され、続いてプログラム評価において必要となる様々な視点が議論されたのち、ビブリオメトリックス手法、経済的効果の測定手法（計量経済学的手法、BETA 法、コントロール・グループ・アプローチ、前後比較）、アンケート・インタビューの手法、ネットワーク効果の評価手法が具体的な評価事例を基に記述されている。本書に含まれている内容は多様であるが、評価システムの設計からプログラム自体の概念化、そして具体的な評価作業の設計までを包括的に把握することに最適なテキストブックである。

Kostoff (Office of Naval Research), "Handbook of Research Impact Assessment," Edition7., NTIS, (1997)

本書はONRのKostoff氏がそのホームページ上で公表している研究評価関連の悉皆的文献集をまとめた概要紹介解説書（ハンドブック）の中の１冊である。本書は米国連邦政府による研究のインパクト評価に焦点をおいて、事前・中間・事後の評価について、遡及的手法（Hindsightや TRACES など)、定性的手法（ピアレビューなど)、定量的手法（費用対効果分析、文献計量学的手法、モデリングなど）の各方法論について説明を行う。本書は 3000 以上の参考文献リストを含んでおり、今後進んで学習するための足がかりとしても使える報告書である。本書ではピアレビューは最も用いられている手法であるために、その質の維持・向上のための条件が示される一方で、ピアレビューだけでは完全でなく、その他の手法をも用いる必要が指摘される。
(http://www.onr.navy.mil/sci_tech/special/354/technowatch/docs/handweb.doc)

PREST, AUEB, BETA, ISI, Joanneum Research, IE HAS, Wise Guys, “Assessing the Socio-economic Impacts of the Framework Programme”, (2002.6)

本書は欧州の研究評価研究者らによる、EU のフレームワーク・プログラム（FP）の社会経済的インパクトの評価方法に関する報告書である。本書はプログラム評価に用いられる方法論を具体的な事例を基に包括的に比較検討したものであるだけでなく、それらの背景にある評価の理念あるいは公的研究開発プログラムの理論的根拠をも深く検討している点で優れた報告書となっている。

具体的には、第２章では FP により公的支出を行うことの経済的合理性を経済学、経営学、イノベーション研究、科学技術政策研究など様々な学問的視点から考察している。これはそもそも公的プログラムの評価においてどのような点に焦点をおくべきであるか、どのような効果が得られていればプログラムは望ましいと考えるかといった理論的根拠を与えるものである。第３章では、FP に対して行われてきた評価研究自体をメタ評価する。比較群などを設定せずに期待される影響を考える「模索的評価」と、事後的に実際の影響を測る「堅い（solid）評価」とを区別したうえで、特に後者について FP 及び EUREKA と ATP の評価事例を基にいかなる次元の効果・影響がどのようなタイムスパンで評価されているか、及びその利点と欠点を検討する。第４章ではケーススタディとして、社会科学研究の評価やサービス・セクターでの研究の評価といった評価研究で今後に課題とされる分野が示されるとともに、BETA 法・費用便益分析・オプションアプローチと言った経済効果分析手法の比較検討などが行われている。このように、本書はこれまでの評価研究における方法の比較・総括と今後に課題なる分野を明らかにした良書である。
(http://les.man.ac.uk/PREST/Download/ASIF_report.pdf)

Joint Research Centre-IPTS and Joanneum Research, “RTD Evaluation Toolbox,” (2002.8)

本書は、欧州委員会の STRATA プログラムにより支援されたプロジェクトとして、Epub thematic network が 28 ヶ月をかけて RTD 政策の社会経済的インパクトを評価

するための方法論を分析したものである。このツールボックスにより、政策立案者や科学者、実務家は主な評価の概念や方法論の概略を知り、自分たちの強みや限界を見極め、政策の文脈の中に自分たちを関連づけさせることが可能となるであろう。本書が登場する時期は、EU において欧州研究圏(European Research Area)を実現させるべく新たな政策装置(Policy Instruments)の実施(例えば Network of Excellence や Integrated Projects)が開始されるのとほぼ重なることから、将来の政策的文脈における評価手続きのインパクトと評価方法の潜在的な相乗効果も狙っていることが本書の目玉とも言えよう。
(http://epub.jrc.es/docs/EUR-20382-EN.pdf)

その他、以下のような文献もある。
Becker, Gerhard and Stefan Kuhlmann, "Evaluation of Technology Policy Programmes in Germany," Kluwer Academic Publishers (1995)

Rosalie Ruegg, Irwin Feller, "A Toolkit for Evaluating Public R&D Investment Models, Methods, and Findings from ATP's First Decade," 2003.3.

2. 政策レベルの評価

Shapira, Philip and Stefan Kuhlmann ed., "Learning from Science and Technology Policy Evaluation –Experiences from the United States and Europe," 2003.

本書は、研究開発評価、特に政策レベルの評価を扱った文献の中で、おそらく最も新しいものである。2000 年 9 月にドイツで開催された欧米の評価研究者を集めたワークショップの成果を踏まえて編纂されており、一流の研究者が著者として名を連ねている。このワークショップの資料は、以下のジョージア工科大学のウェブサイトからダウンロードできる。
http://cherry.iac.gatech.edu/e-value/

政策レベルの評価では以下のような文献もフォローしておく必要がある。

- COSEPUP, "Implementing the Government Performance and Results Act for Research: A Status Report," 2001.
- Ministry of Trade and Industry Finland, "Evaluation of the Finnish Innovation Support System," 2003.
- Industries Department, "Competitiveness and Business Environment in Finland - An International Benchmarking," 2004.

3. プログラムレベルの評価

Georghiou,Luke, Roessner,David, “Evaluating technology programs: tools and methods”, *Research Policy*, Vol.29, 657-678 (2000)

本論文は、技術開発プログラムの評価の手法について、欧米の主要なプログラムの評価事例を基にして包括的にレビューしたものである。そのため、欧州及び米国の研究評価ならびに評価研究の現状を把握するには最適の論文である。論文では大学・公的機関で実施される研究プログラム、産学・官民などの連携促進プログラム、普及・エクステンションプログラムの３種に区分して評価の現状を報告している。ここでは生産関数あるいは投資利益率や内部収益率の分析は実際にはプログラムの正当化や意思決定には有効でないことが強調され、新たな動向として知識生産者と利用者のネットワークを重視する試みが指摘される。また、より具体的にアンケート対象の設定による結果の分散の問題や比較群の設定の必要などが指摘される。
(http://ideas.repec.org/a/eee/respol/v29y2000i4-5p657-678.html)

プログラム評価の参考文献としては以下の文献も有用である。

- COSEPUP, “Evaluating Federal Research Programs- Research and the Government Performance and Results Act,” 1999
- Charles W. Wessner. ed., “SBIR Program Diversity and Assessment Challenges-Report of a Symposium,” 2004.
- Institute of Medicine of the National Academies, “NIH Extramural Center Programs：Criteria for Initiation and Evaluation,” 2004.
- Ken Guy, “Strategic Options for the Evaluation of the R&D Programmes of the European Union,” 1998.11.
- (http://www.europarl.eu.int/stoa/publi/167406/default_en.htm)
- David Roessner, et al., The Role of NSF's Support of Engineering in Enabling Technological Innovation, SRI International, 1997.
- (http://www.sri.com/policy/csted/reports/techin/contents.html)
- David Roessner, et al., The Role of NSF's Support of Engineering in Enabling Technological Innovation, PhaseⅡ, SRI International, 1998.
- (http://www.sri.com/policy/csted/reports/sandt/techin2/contents.html)

4. プロジェクトレベルの評価

プロジェクトレベルの参考文献としては、以下に挙げるピアレビューシステムの構築や評価基準の設定に関するものが有用である。

- National Research Council, “Setting Priorities for Large Research Facility

Projects Supported by the National Science Foundation," 2004.
- Kostoff, "Research Program Peer Review: Principles, Practices, Protocols," 1997.
- Fiona Godlee and Tom Jefferson ed., "Peer Review in Health Sciences," Second Edition, 2003.
- Porter,A. and Rossini,F., "Peer Review of Interdisciplinary Proposals," Science, Technology and Human Values, Vol.10, No.3, 34-42, 1985.
- National Academy of Public Administration, "A Study of the National Science Foundation's Criteria for Project Selection," 2001.2. (http://209.183.198.6/NAPA/NAPAPubs.nsf/0/ca5e12fc2518f12a85256a45004c5e8f/$FILE/nsfreport.pdf)
- THE FIFTH FRAMEWORK PROGRAMME, "Manual of Proposal Evaluation Procedures," 2002.3.4. (http://www.anst.uu.se/andejons/em_en_200001.pdf)
- EPSRC, "Peer Review -A Beginners Guide," 2002.
- NSF, "NSF Grant Proposal Guide (NSF 04-2)," 2003.

5. 分析及びメトリックス

OTA, "Research Funding as an Investment: Can We Measure the Returns? -A Technical Memorandum", NTIS (1986.4)

本報告書は米国科学技術下院委員会科学政策タスクチームが提議した「研究投資を他の経済投資と比較可能な形で測定できるのか」という問いに対して OTA（技術評価局）が 1986 年当時の研究評価の研究動向をレビューしたものであり、翌年に発刊された OECD の報告書とともに研究評価が体系化される初期段階でのレビュー報告書の一つである。報告書は主に、経済的リターンの測定、非経済的アウトプットの測定（ビブリオメトリックス、科学指標）、産業及び政府での評価の意思決定プロセスへの利用に分けて事例を交えて説明されている。既にこの報告書においても、公的研究開発の利益は国防や衛生など多様なものでありそれを経済的価値（金銭）で見積もることの困難さと不適切さが指摘されている。それでも経済的価値を測定しようとするのであれば、科学を技術やイノベーションと連結する要素に注目することが必要であり、技術移転や産業ニーズ主導の研究、予測とプラニングの必要性が述べられる。また、ビブリオメトリックス手法については NSF や NIH の分析事例を引きながら今後の利用可能性を述べるとともに、これらは指標にすぎず評価と同等ではないことも指摘される。また、産業や政府での評価手法の利用状況の調査では、経済的手法は産業においても懐疑的であり、基礎研究では主観的なスコアリングモデルが主流であり、応用研究でも質的情報に基づく点数付けが主流であることが示される。産業においても政府にお

いても、R&D の計画や予算作成では多くの部門・組織・階層が集まって反復的な情報交流や将来予測による意思決定を行うことこそが重要であり、経済的モデリングなどの評価はそれを喚起するための２次的な役割しか果たさないことが述べられている。
(http://www.wws.princeton.edu/~ota/disk2/1986/8622_n.html)

National Research Council, “The Advanced Technology Program: Assessing Outcomes”, National Academy Press (2001)

本書は NRC が行った Advanced Technology Program（ATP）の評価の報告書であり、主に NRC が主催した ATP 評価に関するシンポジウムでの議論をとりまとめたものである。ATP は民間企業による研究開発プロジェクトを競争的に選択して公的資金を提供する制度であるために、市場経済に政府が介入することの正当性への批判が特に議会から継続的になされてきた。そのため、経済学者やシンクタンク・コンサルタント、GAO などによる多種多様な評価が行われており、シンポジウムではそれら評価の取り組みが紹介されている。

ATP の評価で問題とされることは、プログラム内の研究開発プロジェクトが経済効果をもたらしたかという単純なものではなく、民間企業によるプロジェクトに公的資金を提供する正当性があるのか、さらには、正当性を認めるための社会的利益がどのようなメカニズムによって得られているのかを把握することにある。この点において ATP の評価の取り組みは、依然として研究開発が実施されその効果が得られるメカニズムをブラックボックスとし、インプットとアウトプットのみからの費用対効果を求めようとする日本に多くの示唆を与えるものとなっている。

例えば前者の正当化については、GAO が 1997 年に行った評価では、ATP からの資金が得られなくても申請者が他の助成を受ける可能性があったか、助成を受けなかった場合に国民経済上の著しい損失となる可能性があったかを検討した。これは事後・追跡評価で行われるものであるが、その結果はプロジェクト選定（事前評価）のクライテリアを提供するものである。また、ATP では民間企業が行いにくいハイリスク・ハイリターンのプロジェクトを選定するためにプロジェクトの失敗割合が高くなることを許容する必要があることと、民間企業からもほぼ同額の資金拠出を求めることにより実現可能性がある程度存在するプロジェクトが申請されることのバランスがとられているというプログラム・マネジメントの特徴も評価されている。その他にも、私的利益と社会的利益の関係、ベンチャーキャピタルと ATP との関係、ATP における大企業の役割と必要性、プロジェクトの審査方法の信頼性が議論されている。

また、社会的利益のメカニズムについては、Jaffe（1996）はスピルオーバーのメカニズムを「知識スピルオーバー」「市場スピルオーバー」「ネットワークスピルオーバー」に区分し、いかなる特徴を持つプロジェクトが波及効果が大きいかを分析している。これは、プロジェクト選択・形成の指針を提供するものでありプログラムの中間評価の主要な役割と言える。

(http://www.nap.edu/catalog/10145.html)

６. 日本語による文献等

研究開発評価を主題にした日本語の体系的なテキストとしては、次のものがある。

・ 平成 13 年度経済産業省委託調査『研究開発プロジェクト等の評価手法に関する調査』（財）政策科学研究所(2002.3)

また、社会経済的側面の評価を行う際に必要となる分析的手法に関するテキストは多数存在するが、以下、代表的なものを列挙した。

・ 林知己夫編『社会調査ハンドブック』朝倉書店(2002)

・ 田尾雅夫・若林直樹『組織調査ガイドブック』有斐閣(2001)

・ 小池和男『聞きとりの作法』東洋経済新報社(2000)

・ 大谷伸介・木下栄二他『社会調査へのアプローチ』ミネルヴァ書房(1999)

資料　ワークショップ配付資料

研究開発評価研修ワークショップ
＜基礎編＞

（財）政策科学研究所

2007年9月20日(木)
於：丸ビル コンファレンス・スクエア Room 3

財団法人 政策科学研究所
Institute for Policy Sciences

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

本日の内容

- 研究開発評価研修プログラム開発について
- プログラムの獲得目標
- テキスト案の全体像
- 基礎編の演習と解説
- 質疑応答＆全体討論
- 事務局連絡

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

研究開発評価研修プログラム開発について（1）

【委託事業の目的】

- 文部科学省が平成19年度以降展開する研究開発事業等の評価を担う人材を育成するための研修プログラムのカリキュラム案と教材の開発を目的とし、それに資する調査を実施する。

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

研究開発評価研修プログラム開発について（2）

【カリキュラム案・教材案の作成の体制】

- （財）政策科学研究所内に専門委員会を設置，教材案の監修を担当
- これまでに6回の専門委員会を開催，教材案の検討を実施

専門委員会　委員名簿

主査：平澤　冷　（東京大学　名誉教授）
委員：伊地知　寛博（一橋大学　助教授）
勝本　雅和（京都工芸繊維大学　助教授）
菊池　純一（青山学院大学　教授）
調　麻佐志（東京農工大　助教授）
嶌田　敏行（茨城大学　助手）
隅蔵　康一（政策大学院大学　助教授）
林　隆之（大学評価・学位授与機構　助教授）
藤森　英俊（（独）海洋研究開発機構　チーフ）
柳澤　和章（（独）日本原子力研究開発機構　研究主幹）
弓取　修二（（独）NEDO技術開発機構　主任研究員）
吉田　秀紀（（独）科学技術振興機構　副調査役）　※委員50音順

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

研究開発評価研修プログラム開発について（3）

【基礎編の獲得目標】

- 行政機関，研究機関等における「評価の理解者」を育てる

【応用編の獲得目標】

- 評価マネジメントの実務に通じた人材（プラクティショナー）の入り口を示す。

機関・部署		養成・確保すべき人材	
府省	研究開発推進担当課	担当業務に対する自己評価等の一般的評価業務を遂行可能な人材	評価の理解者
	評価担当部署	評価マネジメントの実務に通じた人材	プラクティショナー
資金配分機関	配分業務担当部署		
	評価担当部署	評価対象の実態を把握するための分析的手法を使いこなせるプロフェッショナル	アナリスト
研究実施機関	評価担当部署		
大学等	評価担当部署		

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

テキスト案の全体像

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

研究開発評価研修プログラム

基礎編：演習と解説

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

0．基礎編テキスト案の構成 （1/3）

- 基礎編では研究開発評価のイントロダクションとして「1．評価の意義」、「2．我が国の評価制度」、「3．評価の枠組み」までを範囲とする。
- 「演習」と「解説」を基本構成とし、必要に応じて参考資料、関連トピックスをコラム等で解説

項目	内容	事例・資料等	演習の形式
1　評価の意義	【テーマ】「評価の意義」を考えるための導入部分		
1-1　なぜ評価が必要なのか？	【トピック】 公共経営における評価の必要性を理解します。		
1-1-1　環境の変化と組織経営	・　行政管理から公共経営へ ・　公共経営における評価の位置づけ ・　組織のミッションである		今日的な評価の意義を考える
1-1-2　PDCA サイクルと評価	・　PDCA サイクル ・　マネジメントの改善に役立てる		
1-1-3　コンプライアンスとしての評価	・　コンプライアンス ・　やらなければならない業務		
1-2　評価が有用であるためには	【トピック】 評価を意義あるものにするために、基本的な考え方を解説します。		
1-2-1　誤った枠組みを適用しない	・　評価の理念 （supportive, inclusive & interactive, evincive）	【コラム】 査定的な評価	評価が適切に実施されるための必要条件について理解する。
1-2-2　評価対象をよく理解する	・　評価は一品料理		
1-2-3　形式的な評価にしないためのエ夫をする	・　マニュアル主義からの脱却 ・　考える評価へ		
1-3　評価の留意点	【トピック】 評価現場の混乱や、課題となっている事柄の原因を考えるきかっけとします。		
1-3-1　評価の重複	・　機関評価と研究開発評価 ・　研究開発評価と研究者評価		評価業務の課題と原因について理解を深める。
1-3-2　評価の費用	・　目に見える費用と見えない費用 ・　十分な評価業務体制の必要性		
1-3-3　数量的評価のデメリット	・　数量的評価はどこまで可能か？		
1-3-4　評価に対するインセンティブ	・　モチベーションを考慮した評価結果の活用	【コラム】 インセンティブの誤った適用	

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

0. 基礎編テキスト案の構成 （2/3）

- 我が国の評価制度の概要を踏まえ、自らの評価業務の位置づけを理解する。

2 我が国の評価制度	【テーマ】 我が国の評価制度を体系的に理解します		
2-1 我が国の4種の評価制度と歴史的展開	【トピック】 我が国において実施されている4種の評価制度（政策評価、研究開発評価、独立行政法人評価、大学評価）について概観し、理解を深めます。		
2-1-1 政策評価法	・ 政策評価法の説明 ・ 評価と予算要求のマネジメントサイクル	【資料】 法体系のツリー図と評価主体	4種の評価制度の関連法令，主要な評価主体について理解しているか
2-1-2 研究開発評価	・ 大綱的指針の概要 ・ 総合科学技術会議の役割		
2-1-3 独立行政法人評価	・ 評価業務の集中 ・ 行政改革と独立行政法人評価		
2-1-2 大学評価	・ 認定評価 ・ 国立大学法人評価		
2-2 評価制度の歴史的展開	【トピック】 4種の評価制度の歴史的展開と適用範囲について理解します。		
2-2-1 「研究評価の基本的考え方」と「研究評価のための指針」（1986）	・ 初めての研究開発評価指針 ・ 適用の限界	【資料】 関連法令等の年表	評価制度がどのような発展を経て今日に至っているかについて理解しているか
2-2-2 大綱的指針（1997）	・ 科学技術基本法と大綱的指針 ・ 指針の特徴		
2-2-3 大綱的指針（2001）	・ 政策評価法と大綱的指針 ・ 指針の特徴		
2-2-4 大綱的指針（2005）	・ フォローアップ後の大綱的指針の特徴		
2-2-5 文部科学省指針（2005）	・ 指針の特徴		
2-2-6 適用範囲の区分	・ 各法令、指針類の適用範囲		
2-3 異なる評価枠組みにおける諸問題	【トピック】 評価における混乱を避けるために、評価制度の適用局面を踏まえ、問題の本質を理解します。		
2-3-1 研究開発評価と個人評価	・ 研究開発評価と個人評価の査定を結びつけて良いか	【事例】 A大学における教員評価の仕組み	
2-3-2 研究開発評価とSABC評価	・ 総合科学技術会議のSABC評価とはどのようなものか		
2-3-3 研究開発評価と独立行政法人評価	・ 2種類の制度の網 ・ なぜ独立行政法人評価は難しいのか		

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

0. 基礎編テキスト案の構成 （3/3）

- 評価シートの作成の仕方について理解を深める。

3 評価の枠組み	【テーマ】「評価をするにはどうすれば良いか？」を原理的に理解します。		
3-1 評価はどのように行うか？	【トピック】 評価対象、評価項目、評価基準、評定区分、評点の設定に関する基本的知識を身近な例を用いて解説します。		
3-1-1 評価対象の構成を考える	・ 評価対象の構成要素		総合スーパーA店とB店を評価・比較するためにはどのようにすれば良いか？
3-1-2 評価項目・評価基準の枠組みを考える	・ 評価の観点の構造化		
3-1-3 評定区分を設定する	・ 評定区分の設定 ・ ラジアルチャートの作成		
3-1-4 総合評価をするためには？	・ 総合評価の仕組み ・ 重み付けの意味		
3-1-5 ここまでのポイント	・ 論理の構造化の重要性	【コラム】 相対評価と絶対評価	
3-2 評価シートの作成	【トピック】 評価の枠組み・基本方針をまとめた「評価シート」の作り方について、研究開発事業等を事例に解説します。		
3-2-1 評価シートの基本構成	・ 評価シートの構造 ・ 総合評価の位置づけ		評価シートの基本構成を理解しているか
3-2-2 評価対象系	・ 内部の機能（メカニズム）に着目した評価対象の展開 ・ 機関評価の場合の展開例 ・ 研究開発事業の場合の展開例		評価対象系の展開はどのように行えば良いか
3-2-3 評価体系	・ 評価体系の展開の作法 ・ 研究開発事業の場合の評価体系		大綱的指針の評価項目をどのように展開すべきか
3-2-4 対応関係の考察	・ 評価対象系と評価体系の対応関係		評価対象系と評価項目とはどのような対応関係にあるか
3-2-5 評定区分の設定と評点	・ 評定区分のカスタマイズと評点の付け方 ・ コメントの重要性	【コラム】 パネルの間で評点が割れた時	評定区分の設定と留意点について理解しているか

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

1．評価の意義 （1/6）

- これから評価業務を担当する方へのイントロダクション

【演習1-1】
評価の今日的意義について考察しなさい。

【考察のポイント】
- なぜ、評価が必要なのか？
- 評価は何のために実施するのか？
- やらなければいけない業務だとしても、その意味はどこにあるのか？

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

1．評価の意義 （2/6）

【解説1-1】（テキスト p.1～3）

「なぜ必要なのか？」
- 行政機関や研究機関、大学等を取り巻く環境変化に対応するため、（公共）経営のためのツールとして必要性が増しているから。

「何のために実施するのか？」
- 事業や施策のマネジメントを改善するため。
  ※PDCAサイクルと評価，スパイラル・アップ

「どのような意味があるのか？」
- 今や、評価業務は行政機関等におけるコンプライアンスの確立に不可欠。

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

1．評価の意義 （3/6）

【演習1-2】

評価が有用であるためにはどのような点に配慮すべきか？

【考察のポイント】

- 評価を実施する際の理念（姿勢・態度）はどうあるべきか？
- 評価のための評価になっていないか？
- 評価のマニュアル化は可能か？

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

1．評価の意義 （4/6）

【解説1-2.1】 評価を実施する際の理念 （テキスト p.4）

継続的に評価を行うには評価の理念を確立することが必要

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

1．評価の意義 （5/6）

【解説1-2.2】（テキスト p.6）

「評価のための評価になっていないか？」

- 評価対象をよく理解し、都度、適切な評価の枠組みを考えなければならない。
- 評価は一品料理

「マニュアル化できるか？」

- 評価業務の手続きや観点をマニュアル化することはできても、評価基準などは評価対象の特性に応じてカスタマイズするしかない。
- 悉皆的な評価項目・評価基準を設定すると、評価対象によってはオーバースペックとなる（負担が増す）。

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

1．評価の意義 （6/6）

【解説1-3】 その他の評価の留意点 （テキスト p.7～9）

- これから評価業務を担当する方が理解しておかなければならない点等

評価の重複の問題

→ 「評価疲れ」を形成するひとつの要因。評価が適切に実施されていないことに起因。

評価の費用の問題

→ 一定の管理コストが掛かっている点を理解。

評価結果とインセンティブの問題

→ 誤った適用をしないことが重要。

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

2. 我が国の評価制度 （1/8）

- 我が国の評価制度の歴史的展開と適用範囲を理解し、自らの評価業務の位置づけができるようにする。

【演習2-1】

我が国の科学技術政策に関連した評価には現在、大きく分けて4種の評価制度が展開されています。政策評価（政策評価法）、研究開発評価（大綱的指針）、独立行政法人評価（独立行政法人通則法）、大学評価の4種類の評価制度について概要を理解しましょう。

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

2. 我が国の評価制度 （1/8）

【解説2-1.1】 政策評価の枠組み

- 特定3分野（研究開発事業、公共事業、政府開発援助（ODA））に属する10億円以上の重点的施策について事前評価を実施することが政令で義務づけられている。
- 特定3分野では事業評価方式 と呼ばれる、個別事業の必要性、有効性、効率性を評価する方式が要請されてい

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

2. 我が国の評価制度 （2/8）

【解説2-1.2】 研究開発評価の枠組み

•科学技術基本法（平成7年制定）に基づいた科学技術基本計画（第1期：平成8～12年度，第2期：平成13～17年度，第3期：平成18～22年度）の策定を受けて、「国が実施する研究開発評価の大綱的指針」が定められたことで整備が進められてきた。

•大綱的指針の下で、研究開発事業を展開する各省の研究開発評価指針が整備され、これに基づき、各省所轄の研究開発事業の中間評価、事後評価、追跡調査等が実施されている。

•各省の研究開発評価とは別に、総合科学技術会議では国家的に重要な研究開発事業の評価や優先順位づけ（SABC評価）を実施して いる。

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

2. 我が国の評価制度 （3/8）

【解説2-1.3】 独立行政法人評価の枠組み

- 各法人が中期目標・中期計画 に定めた事業の実施状況について、所轄府省内に設置された独立行政法人評価委員会（民間人によって構成）によって評価される。
- 評価結果を総務省の政策評価・独立行政法人評価委員会が行政改革の観点から再度評価し、これを行政改革推進本部に意見する。
  ※行政改革の意味づけが強い。

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

2. 我が国の評価制度 （4/8）

【解説2-1.4】 大学評価の枠組み

- 国公立・私立を含む認定評価と国立大学法人評価の二つの枠組み
- （独）大学評価・学位授与機構（NIAD）による教育研究活動評価

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

2. 我が国の評価制度 （5/8）

【演習2-2】

- 評価制度の歴史的展開について理解しているか？
- 現在の評価指針類は過去からの研究開発評価の考え方と比べてどのような点で改善されているかについて理解しているか？

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

2. 我が国の評価制度 （6/8）

【解説2-1.1】 我が国の評価制度の歴史的展開

65 農林省研究所レビュー
66 通産省大型プロジェクト評価
71 科学技術会議5号答申（ソフトサイエンスの振興）
84 科学技術会議11号答申：研究評価の充実　　85 臨時行政改革推進審議会答申（科学技術政策大綱提起）
85 科学技術会議政策委員会・研究評価指針策定委員会の設置
86 「研究評価に関する基本的考え方」、「研究評価のための指針」（研究評価技術策定委員会）
87 科学技術会議政策委員会：大規模プロジェクト評価の検討の進め方
87 科学技術会議13号答申：国研問題
89,90 厚生省評価マニュアル　　92 科学技術会議19号答申（ソフト系科学技術の研究開発基本計画）

95 科学技術基本法公布・施行
96 （第1期）科学技術基本計画（閣議決定）
97 国の研究開発全般に共通する評価の実施方法の在り方についての大綱的指針（内閣総理大臣決定）
97 郵政省・農水省・通産省・科学技術庁・文部省などの評価指針
98 厚生省・環境庁・運輸省・防衛庁など評価の指針
98 中央省庁等改革基本法（施行）
98 科学技術会議による評価実施状況のフォローアップ（以降毎年実施）
99 各省庁設置法等からなる中央省庁改革関連法成立
99 独立行政法人通則法（公布）、01.04（施行）
99 経済産業省における政策立案・評価指針など
00.04 大学評価・学位授与機構設置法（施行）

01.01 新府省体制発足
01.03 （第2期）科学技術基本計画（閣議決定）　01.01 内閣府・総合科学技術会議設置法（施行）
01.04 独立行政法人発足
01.06 行政機関が行う政策の評価に関する法律（政策評価法）成立（02.04施行）
01.11 国の研究開発評価に関する大綱的指針（内閣総理大臣決定）
02.03 総務省情報通信研究評価実施指針　　02.03 文部科学省政策評価基本計画
02.03 防衛庁研究開発評価指針　　02.04 経済産業省政策評価基本計画
02.04 経済産業省技術評価指針
02.04 総合科学技術会議評価専門調査会による重要研究開発課題の評価
02.06 文部科学省における研究及び開発に関する評価指針（各省も同様）

05.03 国の研究開発評価に関する大綱的指針（内閣総理大臣決定）
05.08 文部科学省における研究及び開発に関する評価指針（各省も同様）

「研究評価のための指針」（1986）
評価項目・評価基準・評定区分の区別がなく、現在のような評価シートの作成が困難な内容

「大綱的指針」（1997）
評価手法の解説に「数値的指標の活用」や「定量的評価手法の活用」が明記。

「大綱的指針」（2001）
政策評価法との対応。評価マネジメントに具体的に言及。

「大綱的指針」（2005）
支援的評価の必要性に言及。評価結果の政策策定へのフィードバックを重視

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

2. 我が国の評価制度 （7/8）

【解説2-1.2】 評価の全体的枠組みと関連法令の適用範囲

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

2. 我が国の評価制度 （8/8）

【解説2-2】 異なる評価枠組みにおける諸問題

- 研究開発評価と個人評価
  研究開発評価の結果を研究者の査定に利用できるか？
  → 研究者に対する評価の枠組みは、機関のミッション等に照らし合わせて独自に構築すべき。
- 研究開発評価とSABC評価
  総合科学技術会議（CSTP）の実施するSABC評価は各省庁が実施する評価結果とどのように整合がつくのか？
  → イノベーション等の異なる視点からの評価と、重点的予算配分に関する調整プロセスと捉えるべき。
  各省庁が実施する評価結果をエビデンスとして利用しているため、SABC評価の結果にかかわらず、真摯な評価が不可欠。
- 研究開発評価と独立行政法人評価
  独法評価に研究開発評価の結果を適用するのが困難。評価の二度手間ではないか？
  → 評価枠組みの構築と、それに必要なデータ整備に本来であれば十分な時間が必要。現在実施している独法評価は過渡的な措置として捉え、さらなる研鑽が必要。

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

3. 評価の枠組み （1/ 12 ）

【演習3-1】 評価とはどのように行うのか？

例題：総合スーパーA店とB店を評価して下さい。

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

3. 評価の枠組み （2/ 12 ）

【解説3-1.1】 評価対象の構成要素を展開する。
総合スーパーの「何」を評価するのかを明らかにする。

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

3. 評価の枠組み （3/ 12 ）

【解説3-1.2】 評価項目を展開する

- 総合スーパーをどのような観点から評価するかを明確にする。

- 評価対象の機能に併せて評価項目をカスタマイズする（評価基準を設定する）。

		価格面	品質	品揃え	利便性
食料品売り場	生鮮食品	・価格が安いか？	・新鮮か？ ・産地が表示されているか？	・品揃えは豊富か？	・レイアウトが適切か？
	加工食品	・価格が安いか？	・総菜等は作り置きしていないか？ ・味は良いか？	・品揃えは豊富か？ ・オリジナルな商品が多いか？	・レイアウトが適切か？
その他の売り場	衣料品	・価格が安いか？	・縫製はしっかりしているか？ ・デザインは良いか？	・ブランド物を扱っているか？ ・サイズは十分用意されているか？	
	日用品	・価格が安いか？	・耐久性はあるか？ ・安全か？	・定番商品が揃っているか？ ・詰め替え用が用意されているか？	
その他の施設	駐車場	・無料の条件は何か？			・エレベーターが設置されているか？ ・カーゴを使用できるか？
	飲食店	・価格が安いか？	・味は良いか？	・店舗の種類は多いか？	・待ち時間は短いか？ ・子供連れでも大丈夫か？
サービス機能	ポイントカード	・ポイント還元率は高いか？			・有効期限は十分か？
	電子マネー				・ポイントとの交換が可能か？ ・チャージャーは完備されているか？

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

3. 評価の枠組み （4/ 12 ）

【解説3-1.3】 評定区分を設定する。

- 設定した評価基準を一定の尺度で客観的に評価できるようにする。

評価基準	評定区分	評点
価格が安いか？	同地域の他店と比べてかなり安い	4
	同地域の中では比較的安い	3
	同地域の他店とあまり差はない	2
	同地域の中では比較的高い	1
新鮮か？	その日に仕入れたものだけを扱っているのでかなり新鮮である。	4
	適切な鮮度管理を行っている。	3
	新鮮とは言えないが、痛んでいるようなものは置いていない。	2
	痛んでいるものがある。	1

食料品（生鮮食品）の比較

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

3. 評価の枠組み （5/ 12 ）

【解説3-1.4】 総合評価を試みる。

- どのような評価項目・評価基準を重視するか？
- 総合スーパーのどのような機能を重視するか？

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

3. 評価の枠組み （6/ 12 ）

【演習3-2】 評価シートの作成手続きを学ぶ

- 外部評価者に評価方針を説明するための評価シートを作成しましょう。

【考え方のポイント】

- 研究開発事業の何を評価するのかを明らかにする。
- 研究開発事業をどのような観点で評価するのかを明らかにする。
- それぞれがどのような対応関係になっているかについて考察する。
- 評価基準の評定区分がなるべく客観的になるにはどのように設定するべきかについて考察する。

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

3. 評価の枠組み （7/ 12 ）

【解説3-2.1】 評価シートのレイアウト案

※評価マトリックスを表形式に展開

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

3. 評価の枠組み （8/ 12 ）

【解説3-2.2】 評価対象の展開例

機関評価の場合　　研究開発事業の場合

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

3. 評価の枠組み （9/11）

【解説3-2.3】 評価体系の展開

- 研究開発事業の評価体系の展開例

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

3. 評価の枠組み （10/11）

【解説3-2.4】 評価対象系と評価体系の対応関係と評価基準の設定

- 対応関係があるところに○印を付ける。
- ○印の内容に関わる評価基準を設定する。

評価体系 ＼ 評価対象系			必要性				有効性		効率性			
			理由		位置づけ		目的	成果	手段	体制		コスト
			上位計画	戦略	施策体系	施策展開				制度	人・組織	
実績	成果	アウトプット					○	○				○
		アウトカム	○		○	○	○	○				○
		インパクト	○	○		○	○					
	過程	アクター		○					○		○	
		マネジメント							○	○	○	○
		システム	○	○						○	○	
リソース		インプット										○
		ストック										○

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

3. 評価の枠組み （11/ 12 ）

【解説3-2.5】 評価基準の評定区分を設定する。

- なるべく客観的に判断できるよう、評定区分の説明文章を工夫する。
- 最初から相対的な価値評価（4．大変良い，3．良い，2．どちらともいえない，1．不十分）は望ましくない。

【例】
評価対象系：実績－成果－アウトプット－○×技術の開発
評価体系：成果－（評価基準）期待される成果が得られているか？
評定区分：
Ｓ．研究開発目標と比して期待以上の成果が得られている
Ａ．期待（当初の目標）通りの成果が得られている
Ｂ．期待していた（当初の目標）よりはやや不足である
Ｃ．当初の目標を全く達成していない

- 評点を付ける際の理由としてコメントを必ず書いてもらう。
  ※コメントがもっとも重要

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

3. 評価の枠組み （12/ 12 ）

【コラム】 評価パネルの間で評点が割れた時にはどうするか？

- 話し合いの機会を最低1回は設定し、評点を付けた理由をお互いに確認する。（※新たな知見が得られる機会）
- 評点が収束しない場合には、事前に委員の了承を取り付けた上で、最終的に評価パネル委員長の権限で総合的な評点を付ける。（時間とコストの兼ね合い）

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

研究開発評価研修プログラム

質疑応答＆全体討論

Copyright © 2007 by the Institute for Policy Sciences, JAPAN

第Ⅱ部　成果報告書

１．研修プログラムの概要

１.１　研究開発評価研修プログラムのターゲット

本調査を通じて開発を行う研修プログラムや教材のターゲットを明確にするとともに、どのような能力の習得を目指すのかを提示する。

（1）研究開発評価を必要としている局面

まず、我が国において、研究開発評価を必要としている局面を機関・部署別に整理すると、以下のようにまとめることができる。

表１　研究開発評価を必要としている局面

機関・部署		研究開発評価を必要としている局面
府省	研究開発推進担当課	施策・事業の自己評価、研究開発関連制度の設計・運用（大型プロジェクト、資金配分制度（プログラム、プロジェクト）、事前、中間、事後、追跡評価（今後）
	評価担当部署	上記のチェック
資金配分機関	配分業務担当部署	資金配分制度（プログラム）の設計・運用・評価・見直し、プロジェクトの評価（事前―追跡）
	評価担当部署	上記のチェック
研究実施機関	評価担当部署	内部評価の実施、とりまとめ、分析
大学等	評価担当部署	内部評価の実施、とりまとめ、分析

このように、部署によって研究開発評価との関わり方が異なってくるが、研究開発評価の質的向上を図るためには、それらの部署等における評価業務に求められる専門性の程度に応じて、人材を育成・集積しておく必要がある。

（２）研究開発評価研修プログラムの育成人材像

研究開発評価の局面に応じた人材の分類としては、次のような３類型を想定している。

- **評価の理解者：府省研究開発推進担当課**
- **プラクティショナー：府省評価担当部署、資金配分機関配分業務担当部署**
- **アナリスト：資金配分機関評価担当部署、研究実施機関評価担当部署、大学等評価担当部署、以上の他に、府省の政策評価支援部署（組織的には設置されていなくても、この種の機能が必要）**

まず、府省の研究開発推進担当課では、担当業務に対する自己評価程度の一般的評価業務が必要であり、このような軽度な評価業務に携わる多数の一般行政官や職員に対しては、プラクティショナーへの入門的な研修等を通じて、「評価理解者」へと転換させることが重要である。

また、府省の評価担当部署や資金配分機関の配分業務担当部署では、評価マネジメント

の実務に通じた「プラクティショナー」の養成が必要である。しかしながら、プラクティショナーは実務的経験なしには養成し難いため、研修プログラムのような Off-JT だけでは養成手段としては不十分である。これらの部署の人材に対しては、研修プログラムを通じて、プラクティショナーの入り口を示唆するにとどまることになる。

プラクティショナー等が実施した評価結果を、より専門的な観点からチェックしたり分析を深めたりするなど、業務内容としてより高度な専門性が要求される資金配分機関や研究実施機関の評価担当部署では、シンクタンク等の外部評価支援機関を活用するだけではなく、内部にも「アナリスト」と呼ばれるような、評価対象の実態を把握するための分析的手法を使いこなせるプロフェッショナル人材を確保しておくことが望ましい。アナリストは一定の専門的バックグラウンドを有した人材が研究開発評価に継続的に携わることで本来の機能を発揮する存在であるため、これも研修プログラムのみによって専門性を蓄積できるわけではない。

アナリスト向けの研修内容は方法論に関する高度な専門知識を整理することになるので、別途の枠組みが必要と考え、本委託調査のターゲットとしてはこれを含まないこととする。

図　研究開発評価研修プログラムのターゲットとプログラム内容

以上のような人材類型を下に、本調査において想定する研究開発評価研修プログラムのターゲットとプログラム内容の関係は上の図のように示すことができる。

まず、これまで研究開発評価に携わってきた経験のない「評価の初心者」に対して、研修プログラム（基礎編）を受講してもらうことにより、担当する評価業務についての一定の理解を得た「評価の理解者」を育成することが第一段階となる。

次に、ある程度の研究開発評価の実務経験がある者、あるいは評価担当部署に配属され

た人材に対して、研究プログラム（応用編）を受講してもらうことにより、目指すべきスキルの目安やキャリアパスの方向性に対する理解を得た人材を育成することが目標となる。研修プログラム（応用編）の受講経験者はその後の研究開発評価の実務経験を経ることにより、「プラクティショナー」と呼ぶに相応しい知識と専門的能力を身に付けることになる。

１.２　研究開発評価研修プログラムの実施方法

研究開発評価の内容は、段階的な理解を経るよりも、個別トピックスを本質的な側面から理解し、他の問題と照らし合わせて考えるという並列的な思考過程を要するものと考えられる。つまり、研修講座を簡単な命題・トピックスから出発しても、その背後にある考え方や理念など、価値観に関わる問題が階層的に集積しているため、すぐに高度な理解を必要とする局面へと議論が移る傾向がある。

したがって、座学的な研修内容では受講生の理解を促進することには結びつかず、受講者の問題認識や疑問点をその場で解決していくプロセスが必要となる。つまり、演習と解説を相互に繰り返す解説演習形式が望ましい。

このような場合、

「解説中心演習付設形式」と、「演習中心解説集約形式」があり、その典型的な方式として「プロワークショップ」と呼ばれる参加型の研修方式がある。

いずれにしても、実態としては研修課題と対象者のレベルによりこれらの形式を使い分けていくことになり、講師は教材を使って一通りの講義を行うにしても、その合間、合間で受講生同士に討論をさせ、問題認識を深めていくプロセスを加えることが望ましい。

（１）基礎編の実施方法

前述の「基礎編」の場合、多くの受講者は自己流の評価活動を多少手がけた程度であるとすれば、少なくとも前半は「解説中心」に体系的な知識を示し、自己の経験を評価論の体系の中で位置づけ、評価論の広がりと深さを実感するところから始めるのが適切であろう。一方で、基礎編の受講者の問題意識は多岐に渡り分散していることが想定される。原局原課の所属であるならば「政策評価法」への対応が義務づけられているが、担当する課題は研究開発に限定されていない。「大綱的指針」への対応が必要である部署の場合でも、比較的大型の個別プロジェクト（独立型プロジェクト）だけではなく、制度の下で展開するプロジェクト（従属型プロジェクト）の管理運営であったり制度の見直しや新制度の設計であったりする。さらには個別の独立行政法人を担当し「独法評価」に関係している場合もある。このような実態を想定すると、共通する基礎的な部分は限られ、評価論の枠組み程度となり、個別に担当する実務に話題を絞ることが望まれる。

したがって「基礎編」であってもその後半では、受講者の分布に合わせて複数の課題を設定し、「演習中心」に進めることがのぞましい。基礎編のゴールである「評価の理解者」

とは、具体的には担当している実務まわりの評価が適切に処理できる程度であると想定しておくべきであろう。

（２）応用編の実施方法

「応用編」では、課題毎に理解を深め、実務に活かせるスキルの収得を目指すべきであろう。その場合は、導入として基礎編の関連事項の確認から始めることになり、そのための短い解説と演習課題の提示の後、演習を中心に進めることが望ましい。ポイントは演習課題の選定にあり、受講者の習熟度に応じ、理解を深めるための定型的な演習から、受講者が直面している問題を題材にした応用的な課題までを傾斜的にとりまぜて課すことになる。この場合、講師の負担大きく、応用編で扱うべき項目を体系的に踏まえておき、演習課題の展開に合わせて演習の解説として適宜教示すると共に、応用的課題に対しても適切な解決策を提示する必要がある。

（３）プロワークショップ形式の実施事例

図　プロワークショップのイメージ

プロワークショップでは通常、４～５人程度を一つのグループとして、３～４つ程度のグループを配置する。講師はトピックスに関する一通りの講義を行った後、仮想的な事例

に基づいてグループに対してワークを要求する。グループはメンバー間で討論した結果をアイデアボードに書き込み、グループの代表者がそれについて説明をする。

このとき、グループの討論が活発になるよう、ファシリテーション機能が要求される。ファシリテータの役割はグループ討論の進行、タイムキーパー、質問・意見等の集約や記録等であり、優れた講師であれば、記録以外の役割を一人で担うことが可能である。

このような講座形式を採用した場合、一つの講義に許容される人数と時間がどの程度かという問題がある。座学形式であれば一度に100人程度を受講対象とすることができるが、プロワークショップでは、20 人がせいぜいとなる。また講義時間については、座学形式であれば大学の講義のように 90 分１コマで１日１コマのような運用が可能だが、プロワークショップではある程度集中して実施した方がより効果が上がることが知られている。通常は、２日～３日かけて、午前・午後で２時間１コマを１日に３コマ程度実施する。

（4）プロワークショップ形式の制約

研究開発評価研修プログラムの対象は、府省や独立行政法人等における事業推進部署の担当者や評価部署の担当者であるが、新人研修以外にまとまった研修時間を通常業務の中から割くことが困難であることが想定される。

プロワークショップ形式の講座を開催するためには、集中的な日程を用意する必要があり、受講者の属する担当部署がどれだけ研修プログラムの主旨に賛同し、受講者を送り出すことができるかという問題が生じる。科学技術振興機構の戦略的創造事業本部が平成 18 年度に実施した PO 研修コースでは、各部署から研究開発プログラムの担当者が 20 人程度参加したが、部署によっては研修コース参加を通常業務の範囲と認めず、有給休暇を取って参加したメンバーが少なからず存在していた。

したがって、研修プログラムを用意してターゲットとしている受講生が集まるかどうかは、研修プログラムの開催主体が府省の原局・原課、独立行政法人の推進部署等に受講生参加の意義を十分に通達し、受講生の所属上長の許可を取ることが最低限必要である。

また、プロワークショップ形式では、長期的には講師の育成も必要な措置と考えられる。プロワークショップ形式は 1 回あたりの受講生の人数に制限があるため、講師の不足が予想される。初期にはシンクタンク等の外部支援機関や評価研究者が講師やファシリテータの役割を果たすべきではあるが、中長期的には、評価実務経験者が研修プログラムの内容を理解し、必要であればトピックスを加えるなどして、組織的に後進の教育にあたる必要がある。

１.３　研修プログラムの内容

本調査で検討された研修プログラムの構成案を以下にまとめる。

（1）研究開発評価研修プログラム（基礎編）

日程	トピックス	獲得目標
講義と演習（1） （90 分）	1．評価の位置づけと役立て方	・マネジメントにおける評価の位置づけを理解する ・評価の役立て方を学ぶ
講義と演習（2） （90 分）	2．わが国の研究開発評価制度とその特色および変遷	・我が国の研究開発評価制度の枠組みと、業務の位置づけを行う ・我が国の評価制度の特色を海外との比較で学ぶ ・その変遷と塾度を知り、我が国の研究開発評価に課されている課題を認識する
講義と演習（3）－（5） （各 90 分）	3．評価の枠組み（1）－（3）	・評価の基本的な考え方（評価対象・評価体系）を理解する ・評価対象の階層性と評価フェーズの区分を理解する ・実績の多様な局面と区分概念を学ぶ ・評価方法の体系を学ぶ ・評価項目・評価基準・評定区分を理解する ・ピア方式とエキスパート方式の違いを学ぶ ・担当事業を事例とする評価演習を行う

（2）研究開発評価研修プログラム（応用編）

日程	トピックス	獲得目標
1 日目		
午前の部 （90 分）	1．研究開発評価の意義と演習	・研究開発評価の意義を理解する ・研究開発評価の理念を理解する
午後の部（1） （90 分）	2．研究開発評価の局面と演習－1 ～評価対象の区分・実績の区分～	・評価対象の特徴、階層構造について理解する ・成果の区分（アウトプット、アウトカム、インパクト）を理解する ・成果以外の実績概念を理解する
午後の部（2） （90 分）	3．研究開発評価の局面と演習－2 ～研究開発評価の実施時期と重視	・研究開発評価の時期を理解する ・評価時期毎の重視すべき評価項目について理解する

	すべき内容～	
２日目		
午前の部 （90 分）	１．研究開発評価の方法論と演習	・研究開発評価の方法論についての概要を理解する
午後の部（１） (90 分)	２．従属型プロジェクトの評価と演習	・研究開発課題等の採択評価やレビュー評価プロセス、評価方法について理解する
午後の部（２） (90 分)	３．プログラムの評価と演習	・研究開発施策や制度等のプログラムの評価について理解する ・プログラムの目的・目標 ・プログラムのマネジメント ・ロジックモデル
３日目		
午前の部 （90 分）	１．独立型プロジェクトの評価と演習	・重点的資金等による課題の評価、マネジメントについて理解する
午後の部（１） (90 分)	２．機関評価と政策評価	・機関評価のあり方、研究開発評価と機関評価の橋渡しの仕方等について理解する ・政策評価（施策、事業レベル）のあり方について理解する
午後の部（２） (90 分)	３．総合計画・総合政策の評価	・戦略や計画の策定方法、諸外国の事例等を通じ、今後の政策評価の参考にする

（３）Web サイトを活用した補習プログラム

オンサイトによる集中講座だけではなく、受講生が研究開発評価の現場において直面する課題や疑問点を Web サイトを活用してオンラインで補習できるシステムを構築する必要がある。このシステムを通じて、研修プログラムの内容や教材の継続的な見直しを図り、研修プログラムそのものを発展的に運営することが望ましい。

図　Web サイトのイメージ

Web サイトでは、次のような機能を備えるものとする。

- 研修用テキストの全文掲載（html，PDF）
- 講師や他の受講生との掲示板による電子会議機能
- 研究開発評価に関連する情報源情報（リンク）
- 研究開発評価関係のイベント情報に関するリンク

２．研究開発評価研修プログラム（基礎編）

教材案

基礎編　テキスト案

研究開発評価研修プログラム（基礎編）

目次

1　評価の意義 1
1.1　なぜ評価が必要なのか？ 1
1.1.1　環境の変化と組織経営 1
1.1.2　PDCA サイクルと評価 2
1.1.3　コンプライアンスとしての評価 4
1.2　評価が有用であるためには？ 5
1.2.1　評価の理念を確立する 5
1.2.2　評価対象を良く理解する 7
1.2.3　形式的な評価にしないための工夫をする 7
1.3　評価の留意点 8
1.3.1　評価の重複 8
1.3.2　評価の費用 9
1.3.3　数量的評価のデメリット 9
1.3.4　評価に対するインセンティブ 10
2　我が国の研究開発評価制度 11
2.1　我が国の評価制度 11
2.1.1　政策評価 11
2.1.2　研究開発評価 12
2.1.3　独立行政法人評価 13
2.1.4　大学評価 14
2.2　評価制度の歴史的展開 16
2.2.1　「研究評価の基本的考え方」と「研究評価のための指針」（1986） 17
2.2.2　大綱的指針（1997） 17
2.2.3　大綱的指針（2001） 18
2.2.4　大綱的指針（2005） 18
2.2.5　文部科学省指針（2005） 19
2.2.6　適用範囲の区分 19
2.3　異なる評価枠組みにおける諸問題 21
2.3.1　研究開発評価と個人評価 21
2.3.2　研究開発評価と SABC 評価 22
2.3.3　研究開発評価と独立行政法人評価 23
3　評価の枠組み 25

3.1　評価はどのように行うか？ 25
3.1.1　評価対象の構成を考える 25
3.1.2　評価項目・評価基準の枠組みを考える 26
3.1.3　評定区分を設定する 27
3.1.4　総合評価をするためには？ 28
3.1.5　ここまでのポイント 31
3.2　評価シートの作成 34
3.2.1　評価シートの基本構成 34
3.2.2　評価対象系 36
3.2.3　評価体系 38
3.2.4　研究開発事業の評価体系の考え方 40
3.2.5　対応関係の考察 41
3.2.6　評定区分の設定と評点 43
3.2.7　ここまでのまとめ 45

1　評価の意義

1.1　なぜ評価が必要なのか？

1.1.1　環境の変化と組織経営

現在我が国が置かれている状況は、かつての「追いつけ追い越せ」というキャッチアップ体制の下で高度成長を遂げた状況とは全く異なる枠組みの中にあります。経済成長率でみると、我が国は 73 年の第一次オイルショックの直後から低成長の時代に移り、また 80 年代の中頃には一人当たり GDP が世界のベストテン入りし、名実ともに経済的なフロントグループの一員にもなりました。それに加えて先進国全体の成熟社会への移行、社会全体の傾向としての知識社会化、少子高齢化の進展、90 年代に入ってからは本格的なグローバル経済体制への移行など、さまざまな社会経済的環境が変化してきました。

このような環境の変化は、行政においてもかつての「行政管理」の視点からではなく、「行政経営」という新しい枠組みに適応した立場からの政策展開が必要になったことを意味します。5ヶ年計画を策定し、計画期間中は粛々と事業を推進するだけの時代から、事後チェックを実施し、政策の有効性を検証しなければならない時代に移ったと言えます。

また、低成長の時代では限られた物的・人的資源を有効に活用するために、資源配分において「選択と集中」が必要となり、その選択と集中のための論理（ロジック）を生み出すことが重要になります。実効的な意思決定ができるよう、現場の裁量権を拡大し、より機動的な経営ができるようにする必要があります。

フロントグループの一員としては、固有の問題を解決するために独自の戦略をつくる必要があり、その戦略形成機能が新たに必要となります。例えば大学のような教育機関においても、18 歳以下人口の減少に直面し、社会人教育等を含めた新たな経営戦略が必要となるでしょう。

このように組織をとりまく環境が変化したことにより、適切な経営戦略の策定、意思決定を行うために、評価を実施して情報やデータを集約・分析するということが必須の作業となります。

評価を実施することの第一義は、さまざまな情報を収集し、組織の経営・マネジメントに役立てるという点にあります。

1.1.2　PDCA サイクルと評価

組織経営を実効的に行うためには目標管理システムが機能しなければなりません。製造プロセスを有するような企業では、目標を設定して活動を行い、それが効率的に行われているかをチェックするいわゆる PDCA サイクルが、ごく日常的に機能しています。PDCA サイクルとは、

１．計画(Plan)を策定する

２．計画にそって業務を実施する（Do）

３．業務の実施が計画にそっているかどうかを確認（Check）する

４．実施と計画の双方を改善する案を検討し、その結果を次の計画に活かす（Act）

というプロセスのことです。この４段階を順次行って一周したら、最後の Act を次の PDCA サイクルにつなげ、螺旋を描くように一周ごとにサイクルを向上させて、継続的な業務改善をしていきます。この業務改善のしくみをスパイラルアップ（spiral up）と呼び、ISO9000（品質マネジメントシステム）や ISO14000（環境マネジメントシステム）シリーズにも適用されている概念となっています。企業に限らず、あらゆる組織において PDCA サイクルは実施事業を効果的に展開するために重要かつ基本的なマネジメント・モデルと言えます。

図 1-1　PDCA サイクル

PDCA サイクルが上手く機能するためには、Check と Act の部分が大変重要です。なぜならば、分析的観点から実施事業を客観的に判断することで、スパイラルアップがなされるかどうかが決定されるからです。（※この二つの段階を統合し、PDS（Plan-Do-See）サイクルと呼ぶこともあります。）

評価は、評価対象である事業等の PDCA サイクルのうち、Check に相当します。評価が PDCA サイクルにおいて重要なのは、実施事業を継続的に改善する、あるいは次の事業に改善提案をするという Act のために評価が位置づけられていることによります。したがって、評価を実施する際には、評価対象全体のマネジメントが強く認識されなければなりません。

コラム 1-1：評価を実施するとどのような効果があるのか？

評価はマネジメントの改善のために実施すると解説しましたが、現実問題として、果たして本当に評価のおかげでマネジメントが改善されたかどうかを客観的なデータで示すことは難しいでしょう。しかしながら評価を導入して「有効であったかどうか」を主観的に判断することは可能です。

（独）産業技術総合研究所では、研究テーマを関連づけて「研究ユニット」と呼ばれる研究組織を構築しており、平成 14 年度からさまざまな研究ユニットに対して「ユニット評価」を導入しています。「ユニット評価」では外部レビューアから構成される評価委員会（レビューボード）が研究ユニットの実績を評価します。

ユニット長に対するアンケート調査では、「研究ユニット評価は有益だったか」という質問に対して、平成 16 年度においても 74%が「有益だった」と回答しており、平成 17 年度ではさらにその比率が 85%に拡大しています。また、「研究ユニット評価はマネジメントに役だったか？」という質問に対しては、65%前後で「役だった」と回答しており、被評価者自信が評価の有用性を認識しています。組織運営において構成員の意識が前向き変わることは非常に大切なことです。評価の効果はこのようなかたちでまず立ち現れるものと考えられます。

（出所）独立行政法人産業技術総合研究所 HP＜http://unit.aist.go.jp/eval/ci/report.html＞

1.1.3　コンプライアンスとしての評価

現代においては、評価には組織におけるコンプライアンス（法令遵守）の側面があります。「コンプライアンス」とは、「法令や社会的モラルを守ること」ということになります。これは法治国家では当然のことなのですが、コンプライアンス違反の事例は民間企業や行政も含め、後を絶ちません。

コンプライアンス違反は社会的利益を無視して組織的利益を優先するために生じると言われています。企業であれば、「利益を出すため仕方なく」と言うでしょう。「同業他社でも同じことをやっているから」という理由で反社会的行為を行う企業、「売上を伸ばすため」と言う理由で法令を破る営業マン、「会社を守るため」と言う理由で粉飾決算や総会屋対策を行う経営者などが該当します。行政や公的機関においても、組織防衛的な行動が予算の不正支出・流用等に結びつかないとは言い切れません。

したがって、どのような組織であれ、「法令を遵守し、モラルを保持し、社会に対しての説明責任を果たす」という側面がついてまわります。行政や公的機関であれば、納税者に対して、税金を使って事業をきちんと実施していることについて説明の義務があります。説明責任（アカウンタビリティ）を超えて、積極的な情報公開と説明義務を担わなければなりません。

現在では評価業務を実施することは、行政機関が行う政策の評価に関する法律（政策評価法）や独立行政法人通則法、大綱的指針や各省の評価指針等で明記されているところです。今や、評価を実施することがコンプライアンスを確立し、社会的責任を果たすことに結びついていると認識しなければなりません。

1.2　評価が有用であるためには？

1.2.1　評価の理念を確立する

評価を実施する際に常に心がけておかなければならない点は「評価する側と評価される側との間に一定の信頼関係が構築されていなければならない」ということです。そうでなければ、評価結果を組織的なマネジメントの改善に活かすことは難しいでしょう。

したがって、評価の枠組みを考える際に、査定的な態度が先に出てしまってはお互いの信頼関係の構築には至りません。どうすれば現状をより改善できるかという点を共通の目的に据えた上で、被評価者はデータや情報、意見等を提示し、評価者側はルールに則って評価するという枠組みを基本としなければなりません。

評価においてもっとも重要なことは、マネジメントの改善に役立てるために適切な評価の枠組みを構築し、関係者の理解と協力を得ることです。そのためには次のような評価の理念を確立することが有用でしょう。

図 1-2　評価の理念

「支援的であること（supportive）」とは、評価を実施することがマネジメント側と被評価者の双方にとって、業務改善等のきっかけになることを意味します。例えば、評価結果に改善提案のコメントを付す、評価結果を次の事業の企画に反映させるというような行為がこの考え方に基づいています。このことは査定的な評価とは異なり、単に「切る」ための評価ではないという考え方です（コラム 1-2 参照）。

「非排除・双方向的であること(inclusive and interactive)」とは、マネジメント側が評価のプロセスにおいて被評価者を排除するようなことはせず、双方が協力して評価を実施することを意味します。例えば評価パネルの場で被評価者に意見を述べさせる、あるいは、ある評価結果に至った理由を被評価者にきちんと説明する、というような状況が考えられます。「評価を実施するので評価シートに必要事項を記入せよ」ということだけでは、被評

価者を排除しているばかりか、一方向的な評価となってしまいます。

「明示的であること(evincive)」とは、評価の各プロセスにおいて誰が見ても分かるように文書で規定が明らかになっており、証左となる情報（エビデンス類）に基づいて評価が実施されることを意味しております。また、評価結果についても公開を原則とし、国民の目からも評価のプロセスが明らかであることも含まれています。この「明示性」が担保されずに評価が「支援的」かつ「非排除・双方向的」であると、単なる「馴れ合い」になってしまい、評価結果に関する信頼性を確保することが難しくなります。

このような理念が基本にあれば、現行の評価業務に問題があった場合、何が考え方として誤りであって、どうすれば改善できるかの判断基準となることでしょう。評価業務を定着させるためにはここで挙げた理念が必須要素となります。

コラム 1-2：査定的な評価と支援的な評価

評価においては「切る」ための評価もあり得ると考える方も多いと思われます。例えば提案公募型事業における課題採択の評価は現実問題として応募者全員を採択するわけにはいかないので、「切る」評価を実施せざるを得ないではないかとの批判もあるでしょう。

応募者に１回しかチャンスがないのであれば、その場合の評価は単純に「切る」ことを目的としていると考えても良いかもしれませんが、多くの場合は複数回公募を実施するでしょうし、同様の研究開発プログラムをその後も展開するのであれば、応募者は何回でもチャレンジできることが考えられます。

ここで重要なことは、提案公募型事業の推進者が応募者という潜在的な研究開発実施者を常にカスタマーとして認識しているかどうかという点です。事業がより良い展開を果たすためには、良質な課題が提案されることが条件であり、そのためには研究開発実施者が一定の水準でストックされていることが前提となります。事業が継続的に展開されていくためには、採択評価で「切る」にしても、研究開発実施者により良いリピーターとなってもらうための支援的な措置が必要です。例えば、不採択の明確な理由書、提案書に対する改善点の指示、他の資金配分プログラムの斡旋・紹介等が考えられるでしょう。

評価が「支援的であること」とは、評価を実施する側と被評価者とが継続的な関係を構築するための条件なのです。

1.2.2　評価対象を良く理解する

マネジメントの改善に資する評価を行うためには、評価対象である事業をより良く理解しなければなりません。研究開発事業であれば、当該事業の研究分野における特徴、研究開発のリードタイム、失敗のリスクの大きさ等を踏まえ、適切な時期に適切な評価を行うことが重要になります。

これに対して、マネジメント側の都合であらゆる研究開発事業に対して一律の評価項目・評価基準を設定したり、異なる目的の研究開発に対して評価結果のみを比較可能な数量的評価を適用したりすることは、評価そのものが事業のマネジメントと切り離されて実施されていると考えられます。

評価は基本的に一品料理であり、素材（評価対象）をいかに理解しているかが良い料理（＝良い評価）の決め手となります。また、和食の一品と洋食の一品とを単純に比較できないように、評価結果の単純な比較もまたできません。

1.2.3　形式的な評価にしないための工夫をする

我が国の行政機関では評価のプロセスをマネジメント・サイクルに明示的に導入するよりも先に、各種評価制度が整えられてしまったことにより、評価業務がマネジメントと独立に存在するかのような誤解が生じています。

行政機関の現場の声としてよく発せられるのが、業務のマニュアル化の必要性です。評価においてもマニュアル化を徹底し、要求される評価報告書をとりまとめ、現業の負担を少しでも軽くしようという考え方は当然の要請なのですが、マニュアルに従って評価を実施すれば良い評価ができるということにはなりません。

マニュアルでは最低限踏まえておかなければならない組織内の手続きを示すことは出来ても、目的も背景も異なる様々な評価対象を適切に評価するための方法論を示すことはできません。せいぜい評価の観点を示すにとどまり、それを具体的に展開するのはマネジメント側（推進部署）の責任であると言えます。

また、マニュアルではあらゆる事態を想定して、一通りの業務が遂行されるよう構成されるので、実際に評価を担当する場合には当該事業にとって過度な手続きが要求されることも考えられます。評価を形式的に規定しようとすればするほど、多様な価値観を評価しなければならない研究開発のような事業ではかえって負担が増えることにつながりかねません。

評価はマネジメントの一部であるということを認識し、当該事業にふさわしい評価の在り方を事業の企画の段階で想定しておく必要があります。

1.3　評価の留意点

1.3.1　評価の重複

現行の評価制度は評価を実施する側、受ける側から見ると必ずしも整合的であるとは言えません。研究開発評価は個別の研究開発事業等を対象としているのに対し、独立行政法人評価や国立大学法人評価などは、機関や組織（それを構成する研究グループ、研究者）を対象としています。

前者は事業等のマネジメントの改善に資することが主な評価目的となっていますが、後者は組織や個人に関する達成度評価の形式で評価されるか、場合によっては、他の組織・個人との比較（ランキング）まで行われることがあり、評価結果が査定等に利用される可能性があります。また、評価実施者もそれぞれ異なる場合が多く、前者の評価は現場に近い部署が担当し、後者の評価は組織の本部機能が担っているのが一般的です。

同じ研究活動を評価の対象としながら、評価の目的が異なるために、評価情報や評価結果を上手く活用できず、評価の二度手間や重複が発生しているのが現実です。

図 1-3　研究開発に対するさまざまな評価の局面

しかしながら、こうした問題点は評価の基本的な枠組みよりも先に、評価制度が整えられてしまったことに起因していると言えます。本テキストでは、このあたりの問題点を踏まえ、どのようにすれば評価の重複が起こらないように異なる評価制度に対応できるかについて考えてみたいと思います。

これから評価業務を担当する方々は、ある評価対象の評価に業務がとどまらず、他の組織や業務との連携が非常に重要である点を認識してください。

1.3.2　評価の費用

評価業務には一定の費用が掛かります。中間評価・事後評価を例に取ると、直接的には外部パネルを招集するための経費（謝金、会議費等)、評価業務支援のための請負費などが相当します。また、評価業務に係わる職員の人件費もこれに上乗せされます。

一方で、目に見えないかたちで存在するのが、評価を実施するために係わる内外の人件費です。事業推進部署の職員が評価業務のために資料を作成したり、研究開発実施者に対してヒアリングをするなどの費用や、研究開発実施者側においても必要書類を作成し、エビデンス類を確保するための費用が発生します。こうした業務を外注すれば委託費として一部は明らかになる場合もありますが、全てではないでしょう。

一般的に「評価に費用が掛かりすぎる」という認識があるとしたら、直接的な費用よりもむしろ目に見えない部分での負担が大きいという意味合いで語られることが多いかもしれません。

こうした費用に対する認識は、評価業務そのものの効率化によってある程度改善することも可能ですが、評価業務を本来のマネジメントの一環としてどこまで捉えるかということにも依存します。評価が文化として定着している組織においては、評価業務が負担であるという認識にはならず、やらなければならない仕事の一つとして内部化されています。むしろ、評価業務の量に見合った人員・体制・機能が確保できているかということが問題となるでしょう。

1.3.3　数量的評価のデメリット

評価結果を比較することが容易にできるため、数量的評価への期待は大きいものがあります。しかしながら、「どのような質をどのような数値で計測するか？」ということは非常に重要な問題であり、数量的評価を安易に適用するのは危険です。

例えば、研究者のアウトプットを評価する場合、海外査読付き論文数で評価したとします。このような数字が研究開発の予算配分や研究者個人の査定にまで利用されるような状況にあったとしたらどうなるでしょうか。

容易に想像できることですが、例えば 70～80 頁ほどあるフルペーパーを専門誌に投稿せず、非常に短い分量のアカデミック・レターのような雑誌に投稿することで投稿論文の数を稼ぎ、それが実績評価値として採用され、さらにそのサイクルを助長することなどが考えられます。結果、研究の質を低下させることにもつながりかねません。こうした傾向は、任期内に成果を出さなければならない期限付き雇用の研究員の立場を考慮するすると、さらに拍車がかかるものとなります。

では、論文の質が比較的反映されているとされるインパクト・ファクター（論文被引用度数）を使えば良いでしょうか？論文が引用されるためにはある程度の時期を必要としますので、これも使用できる局面が限られてくるでしょう。

数量的評価は、評価しようとする対象の一面を形式的に評価しているに過ぎません。それ以外の面を捨象して評価することがないように、評価の枠組みを十分に検討する必要があります。評価結果が何かしらの査定に利用される場合には、特に注意が必要です。

むしろ、ミッションも活動の内容も異なる様々な機関・組織、あるいは研究領域の特徴や研究開発のリードタイムも異なるような研究者の評価を実施しようとする際に、比較可能な指標が容易に設定できるかどうかという点をよくよく考えることが重要です。

1.3.4　評価に対するインセンティブ

評価業務は評価実施者側と被評価者との協働で行われます。評価実施者側は多くの場合、それを業務として実施する必要性を持っていますが、被評価者は本業である研究開発に投じる時間数が削られるという認識を持つことが一般的です。

そこで、評価に対するインセンティブを十分に考慮しなければなりません。評価が望ましい形で実施されることが前提ですが、基本的には良い評価結果を受けたら研究開発予算等の配分に反映させるというように、インセンティブを考慮することが重要です。

逆に、誤ったインセンティブの枠組みを適用することによって、研究者のモチベーションが深刻化する怖れがあります。

コラム 1-3　インセンティブの誤った適用

研究開発型独立行政法人のＡ研究所に在籍するＢ主任研究員のチームは、競争的研究資金制度の公募に応募して見事採択されました。Ｂ主任研究員のチームは向こう３年間の研究資金を外部から獲得しました。ところが、研究所側は「Ｂさんのところでは潤沢な研究資金を獲得できたので、交付金の中から配分される予算を削減し、他の研究チームに回したい」との要望を伝えました。

これでは、Ｂ主任研究員のチームがせっかく努力して外部資金を獲得したのに、そうではない研究チームに予算が配分されてしまい、努力の結果が報われないことになってしまいます。

実質的な予算の利用状況に応じて最終的な配分調整のプロセスが必要になることは仕方のないことですが、研究機関にとって外部資金の獲得は研究者個人もしくは研究チームの実績を評価する上でもっとも重視しなければならないもののはずです。したがって、予算とは別のかたちでＢ主任研究員のチームが何らかのかたちで「報われる」仕組みを作ることが必要でしょう。

2　我が国の研究開発評価制度

2.1　我が国の評価制度

我が国の研究開発に関わる評価制度には以下のようなものが存在します。

1) 政策評価（行政機関が行う政策の評価）
2) 研究開発評価
3) 独立行政法人評価
4) 大学評価

これらの評価制度がどのような法令等の下で展開されているかについて詳しく見てみましょう。

2.1.1　政策評価

「従来、わが国の行政においては、法律の制定や予算の獲得等に重点が置かれ、その効果やその後の社会経済情勢の変化に基づき政策を積極的に見直すといった評価機能は軽視されがちであった」との認識の下に、政策評価制度の導入が行政改革会議で提言されました（平成 9 年 12 月 3 日）。これを受けて、平成 13 年 1 月に中央省庁等改革の 1 つの柱として、政策評価制度がスタートし、同年 6 月に『行政機関が行う政策の評価に関する法律（政策評価法）』が制定されました（平成 14 年 4 月施行）。

また、政策評価法に基づき、『政策評価に関する基本方針』が平成 13 年 12 月に閣議決定され、この方針に従って、各府省庁が 3 年を目処に「政策評価基本計画」策定し、さらに年度毎の事後評価に関わる「政策評価実施計画」を策定して政策評価を行うこととなっています。

図 2-1　政策評価の枠組み

また、政策評価法では特定３分野（研究開発事業、公共事業、政府開発援助（ODA））に属する 10 億円以上の重点的施策について事前評価を実施することが政令で義務づけられています（行政機関が行う政策の評価に関する法律施行令，平成 13 年政令第 323 号）。これらの特定３分野では事業評価方式[1]と呼ばれる、個別事業の必要性、有効性、効率性を評価する方式が要請されています。

この他、政策評価法では総務省に他の省庁とは異なる立場からの評価機能を持たせており、総務省は複数の省にまたがる政策の評価、各府省庁の政策評価の点検活動を実施しています。点検の後、予算要求への評価結果の反映状況をとりまとめ、財務省がこれを受けて予算編成に活用します。

出所：総務省「政策評価Ｑ＆Ａ」

図 2-2　政策評価と予算要求のマネジメント・サイクル

2.1.2　研究開発評価

我が国の研究開発評価制度は、科学技術基本法（平成 7 年制定）に基づいた科学技術基本計画（第１期：平成 8～12 年度，第２期：平成 13～17 年度，第３期：平成 18～22 年度）の策定を受けて、「国が実施する研究開発評価の大綱的指針」が定められたことで整備が進められてきました。

大綱的指針の下で、研究開発事業を展開する各省の研究開発評価指針が整備され、これに基づき、各省所轄の研究開発事業の中間評価、事後評価、追跡調査等が実施されています。

[1] 個々の事業や施策の実施を目的とする政策を決定する前に、その採否、選択等に資する見地から、当該事業又は施策を対象として、あらかじめ期待される政策効果やそれらに要する費用等を推計・測定し、政策の目的が国民や社会のニーズ又は上位の目的に照らして妥当か、行政関与の在り方からみて行政が担う必要があるか、政策の実施により費用に見合った政策効果が得られるかなどの観点から評価するとともに、必要に応じ事後の時点で事前の時点に行った評価内容を踏まえ検証する方式.（出所：総務省「政策評価に関するガイドライン」平成 17 年 12 月）

図 2-3　研究開発評価制度の枠組み

また、各省の研究開発評価とは別に我が国の科学技術政策全体を俯瞰する立場から、総合科学技術会議では国家的に重要な研究開発事業の評価や優先順位づけ（SABC 評価）を実施しており、ここでの審議結果が近年では予算配分案に反映されるようになっています。

2.1.3　独立行政法人評価

国立研究所や研究資金配分機関が独立行政法人に移行し、これらの研究開発に関わる機関も独立行政法人通則法（平成 11 年制定）に基づいて独立行政法人評価の枠組みの下に置かれました。

独立行政法人評価は各法人が中期目標・中期計画[2]に定めた事業の実施状況について、所轄府省内に設置された独立行政法人評価委員会（民間人によって構成）によって評価されます。その評価結果を総務省の政策評価・独立行政法人評価委員会が行政改革の観点などから再度評価し、これを行政改革推進本部に意見するという枠組みになっています。また、行政改革推進本部の下に行政減量・効率化有識者会議（旧：独立行政法人に関する有識者会議）が設置され、ここにおいても中期目標期間を終了した独立行政法人の見直しを検討し（評価はせず）、意見を述べる機能が設定されています。

[2] 中期目標：3～5 年間の達成目標として主務大臣が策定，独法に指示．中期計画：中期目標に基づき独立行政法人が作成，主務大臣が認可．

図 2-4　独立行政法人評価の枠組み

先に評価の重複について解説しましたが、特に研究開発型独立行政法人の場合は、大綱的指針に基づく研究開発評価（主に競争的研究資金）や、総合科学技術会議による評価、さらには独立行政法人評価と様々な評価制度の枠組みの下に組み込まれており、評価業務そのものの負荷も大きいばかりでなく、それぞれの評価結果が必ずしも連携されていないという複雑な状況にあると言えます。

2.1.4　大学評価

大学評価制度は平成 15 年の教育基本法の改定に伴い、わが国の全大学に対し、文部科学大臣の認証する評価機関による定期的（7 年に 1 回）な評価を義務化したことを契機として整備されました。認証機関としては、（独）大学評価・学位授与機構や（財）大学基準協会などが挙げられます。

また、平成 15 年に国立大学法人法が制定され、これに基づいて国立大学法人評価委員会（学識経験者・有識者により構成）が文部科学省内に設置されました。国立大学法人評価委員会では、国立大学法人及び大学共同利用機関法人が定めた中期目標（6 年間）の検討や各年度終了時の評価を実施しています。

図 2-5　大学評価の枠組み

2.2　評価制度の歴史的展開

我が国における研究開発評価制度は、政策評価や省庁改革の流れと平行して展開されてきました。表 2-1 では我が国の研究開発制度の経緯を左側に、政策評価や省庁改革の経緯を右側に示しています。もっとも古くは 1965 年に当時の農林省が農業総合研究所の機関評価を実施したことがはじまりとなっています。次いで、当時の通産省が 1966 年に大型プロジェクトの評価を実施しています。

以下では、研究開発評価の指針がどのような発展を遂げてきたについて見てみることにしましょう。

表 2-1　我が国の研究開発評価制度の展開

65 農林省研究所レビュー	
66 通産省大型プロジェクト評価	
71 科学技術会議5号答申（ソフトサイエンスの振興）	
84 科学技術会議11号答申：研究評価の充実	**85** 臨時行政改革推進審議会答申（科学技術政策大綱提起）
85 科学技術会議政策委員会・研究評価指針策定委員会の設置	
86 「研究評価に関する基本的考え方」、「研究評価のための指針」（研究評価技術策定委員会）	
87 科学技術会議政策委員会：大規模プロジェクト評価の検討の進め方	
87 科学技術会議13号答申：国研問題	
89,90 厚生省評価マニュアル	**92** 科学技術会議19号答申（ソフト系科学技術の研究開発基本計画）
95 科学技術基本法公布・施行	
96 （第1期）科学技術基本計画（閣議決定）	
97 国の研究開発全般に共通する評価の実施方法の在り方についての大綱的指針（内閣総理大臣決定）	
97 郵政省・農水省・通産省・科学技術庁・文部省などの評価指針	
98 厚生省・環境庁・運輸省・防衛庁など評価の指針	
	98 中央省庁等改革基本法（施行）
98 科学技術会議による評価実施状況のフォローアップ（以降毎年実施）	
	99 各省庁設置法等からなる中央省庁改革関連法成立
	99 独立行政法人通則法（公布）、01.04（施行）
	99 経済産業省における政策立案・評価指針など
	00.04 大学評価・学位授与機構設置法（施行）
	01.01 新府省体制発足
01.03 （第2期）科学技術基本計画（閣議決定）	**01.01** 内閣府・総合科学技術会議設置法（施行）
	01.04 独立行政法人発足
	01.06 行政機関が行う政策の評価に関する法律（政策評価法）成立（02.04施行）
01.11 国の研究開発評価に関する大綱的指針（内閣総理大臣決定）	
02.03 総務省情報通信研究評価実施指針	**02.03** 文部科学省政策評価基本計画
02.03 防衛庁研究開発評価指針	**02.04** 経済産業省政策評価基本計画
02.04 経済産業省技術評価指針	
02.04 総合科学技術会議評価専門調査会による重要研究開発課題の評価	
02.06 文部科学省における研究及び開発に関する評価指針（各省も同様）	
05.03 国の研究開発評価に関する大綱的指針（内閣総理大臣決定）	
05.08 文部科学省における研究及び開発に関する評価指針（各省も同様）	

出所：（財）政策科学研究所（2002），平澤（2004）.

2.2.1　「研究評価の基本的考え方」と「研究評価のための指針」（1986）

国が実施する研究評価の在り方に関して政府によってはじめて検討されたのは、当時の首相の諮問機関である科学技術会議政策委員会の研究評価小委員会（事務局は科学技術庁）が 1986 年６月に定めた「研究評価の基本的考え方」と，９月に定めた「研究評価のための指針」でした。「研究評価の基本的考え方」では、研究評価の概念の内容や研究開発の属性等によって多様に変化し、あらゆる研究評価に共通して適用しうる研究評価システムはありえないと当時、既に指摘されています。研究開発の属性（研究の性格、予算管理区分、領域、規模、期間等）によって、研究評価のあり方を検討する必要があることを踏まえた上で、基本的には「実効性」、「継続性」、「柔軟性」、「透明性」の４点から対象を評価することが指摘されています。

続く「研究評価のための指針」では、この４点をどのように評価すべきかについて述べています。「実効性」については「基礎的研究と開発的研究を峻別して研究評価システムを検討すること，（研究者と評価者の）すりあわせを重視した評価形態とすること，（当該研究開発の意義や研究水準等の）位置同定を重視すること，研究開発の目標を明確にすることの４つが指摘され，「継続性」については，文書化されたフォーマットを使用すること，研究者および評価者の負担の少ない評価方法を用いることの２つが指摘され，「柔軟性」については，モニタリング結果等に応じ，評価項目等の見直しを適時行うこと，中間評価を実施することの２つが指摘され，「透明性」については，評価方法，評価項目，評価者等を明らかにすることが指摘されています。

しかしながら、この指針では評価論に関する初期の混乱が多く見られました。例えば、評価項目・評価基準・評定区分の区別がなく、現在のような評価シートの作成が困難な内容でした。当時、この指針の普及状況をフォローアップしたところ、指針の考え方を導入して評価を実施している国立研究機関はほとんどありませんでした。

2.2.2　大綱的指針（1997）

次の大きな転換点としては、1996 年の科学技術基本法の公布を受ける形で整備された、1997 年の「国の研究開発全般に共通する評価の実施方法の在り方についての大綱的指針」（内閣総理大臣決定）です[3]。97 年の大綱的指針では、研究開発評価の基本方針として次の４点が明記されました。

・評価基準・過程が外部からも分かる透明性のある明確な評価の実施方法の確立
・第三者を評価者とする外部評価の導入
・国民に評価結果等を積極的に公開するなど開かれた評価の実施
・研究開発資源の配分への反映等評価結果の適切な活用

我が国の評価制度の枠組みはこの 97 年の大綱的指針によってほぼ方向性が定まったとい

[3] http://www.mext.go.jp/b_menu/shingi/kagaku/hyoka/index.htm

えます。この時期に各省庁から研究開発評価に関する指針類の第１バージョンが作成されました。

この指針では、評価手法の解説に「数値的指標の活用」や「定量的評価手法の活用」が明記されていますが、研究開発評価の本質的な議論からすれば、不十分な点もありました。

2.2.3　大綱的指針（2001）

2001 年に新府省体制への移行が済むと、同年 6 月に「行政機関が行う政策の評価に関する法律（政策評価法）」が成立します（施行 2002 年 4 月）[4]。政策評価法は 10 億円以上の予算規模がある事業全てが評価の対象となり、評価の観点として「必要性」、「効率性」、「有効性」、「公平性」、「優先性」の５つの観点が盛り込まれました（第６条第２項に定められた「三．政策評価の観点に関する事項」による）。

政策評価法との整合を図るかたちで、同年 11 月に「国の研究開発評価に関する大綱的指針」が定められました[5]。01 年の大綱的指針では、政策評価法の観点である「必要性」、「有効性」、「効率性」を研究開発評価の観点とするようになりました。01 年の大綱的指針の大きな特徴は、評価対象や評価結果の取り扱い、評価実施体制について具体的に言及してある点です。

また、この指針から総合科学技術会議が実施する研究開発評価や、各省の独立行政法人評価委員会の実施する独立行政法人評価、独立行政法人大学評価・学位授与機構の実施する国立大学等の教育研究活動の評価の位置づけがなされています。

この大綱的指針を受け、各省庁から現在まで利用されている評価指針が整備されました。

2.2.4　大綱的指針（2005）

2005 年にフォローアップの結果を受けて、「国の研究開発評価に関する大綱的指針」は改定されました[6]。05 年の大綱的指針では、評価の理念として、「研究者を励ます」という文言が挿入され、研究開発評価が研究者にとって支援的なものであることを明言しました。

また、評価対象に関する調査・分析の必要性、研究開発制度等の時系列運用と評価のフィードバックが強調されており、より上位の制度、施策、総合計画等の評価の必要性が明記されています。現在の我が国の研究開発評価の枠組みは、この 2005 年の大綱的指針に基づいて展開されています。

---

[4] http://www.soumu.go.jp/hyouka/houritu.htm
[5] http://www8.cao.go.jp/cstp/hyoukasisi.pdf
[6] http://www8.cao.go.jp/cstp/taikou050329.pdf

2.2.5　文部科学省指針（2005）

2005 年の大綱的指針の改定を受け、「文部科学省における研究及び開発に関する評価指針（文部科学省指針）」が同年 9 月にいち早く改定されました[7]。文部科学省指針では、同省が展開する教育・研究の特徴に併せた評価項目・評価基準がより多様に展開され、基礎研究や学術研究において評価すべき観点が明確に述べられているとともに、研究者を奨励し、評価の文化を我が国の研究開発現場に根付かせようという、積極的な意図にあふれた内容となっています。

2.2.6　適用範囲の区分

我が国の研究開発評価の指針と関連法令の適用範囲の推移は、そのまま我が国の研究開発評価制度の進展状況と重なります。

各指針、法令の適用範囲を図示したものが図 2-6 です。ここで横軸は評価時期を示しており、縦軸は評価対象を示しています。評価時期に「途上」とあるのは、我が国で一般的な中間評価、事後（直後）評価の双方を含んでいます。この分類は、研究開発評価が盛んな欧州や米国では一般的なのですが、研究開発の成果が一定の期間を置かなければ観測できないために、研究開発事業が終了した直後における評価についてもモニタリングの意味合いで行われることが多いことによります。

評価対象で「独立型プロジェクト」、「従属型プロジェクト」とあるのは、各省庁によって事業の分類の仕方が異なるために便宜的に呼称している事業の類型です。「独立型プロジェクト」とは施策や制度の枠組みとは独立に展開される事業を指し、多くは予算規模の大きいプロジェクトが相当します。これに対し、「従属型プロジェクト」とは何かしらの施策、制度の枠組みの下で展開される事業一般を示しており、例えば研究開発制度における研究開発課題などが相当します。

1997 年の大綱的指針では、従属型プロジェクトの事前評価、途上評価、研究開発機関の評価が主な適用範囲でした。

2001 年に成立された政策評価法では、実施中もしくは終了直後の独立型・従属型プロジェクト、研究開発制度、研究開発施策が主な適用範囲であったと言えます。各省庁が実施する事業の事前評価書を総務省総務大臣に提出することが義務づけられていますが、ここでの事前評価とは省内の原局・原課が実施する自己評価であり、概算要求のための計画的意味合いが大きいために、研究開発評価という枠組みにおいては若干異色のものと言えます。現在に至るまで、我が国では政策の効果に関する分析（インパクト・アセスメント）を実施した本格的な事前評価は実施されていないのが実情です。

2001 年の大綱的指針では研究開発プログラム、研究開発制度、独立型プロジェクトの途上評価にまで適用範囲が拡大されました。また、一部のプロジェクトや制度に関しては追

[7] http://www.mext.go.jp/a_menu/kagaku/hyouka/06081122.htm

跡調査の実施の必要性を訴えています。この指針では独立行政法人評価や大学評価・学位授与機構が行う大学評価との評価の連携についても指摘されています。

2005 年の大綱的指針では、より上位の政策・施策までが評価対象となり、評価時期についても追跡調査の積極的な展開の必要性が盛り込まれるようになりました。

また、機関評価については、2001 年独立行政法人通則法が施行され、各独立行政法人は中期目標・中期計画期間を定め、計画期間終了時にその計画達成度についての評価を実施することが定められました。大学評価については、2003 年における学校教育法の改定、国立大学法人法の制定により、認証評価と国立大学法人評価が実施されることになりました。認証評価については文部科学省の認証する機関が一定期間をおいて（7 年に 1 回）大学基準評価を実施することになっています。また、国立大学法人評価では、中期目標期間を大学側が定め、これを大学評価・学位授与機構の調査・分析を基に国立大学法人評価委員会が評価することになっています。

出所：平澤（2006）を修正

図 2-6　評価の全体的枠組みと関連法令

2.3　異なる評価枠組みにおける諸問題

2.3.1　研究開発評価と個人評価

研究開発評価では研究開発に関わる実績を評価し、マネジメント改善に結びつく点を抽出することが目的となります。一方で、研究者個人に対する評価は組織人事の問題とも結びついて、何かしらの査定が前提となっています。研究開発評価の結果が個人の業績に反映されるとしたら、その妥当性はどのように担保されるべきでしょうか？

評価結果が査定と直結すると、被評価者は「評価の論理」で行動するようになることが予想されます。研究者個人は評価結果が良くなるように、評価項目・評価基準に照らし合わせた活動を戦略的に取るようになるでしょう。この仕組みが上手く機能すれば良いのですが、研究開発の質の向上につながるような仕組みは評価項目・評価基準だけで最適化されることはあり得ず、日々のマネジメントの工夫によって達成されるものです。

したがって、多くの場合は研究開発評価の結果が個人の業績査定と直結すると、望ましくない影響をもたらします。組織構成員である研究者のさまざまな活動、チームとしての共同作業、ネットワーク形成機能、地域社会への貢献等などを捨象し、ある特定の研究成果の面だけを査定の基準として取り上げると、評価を導入する前よりも組織文化の形成や研究者のモチベーションに悪影響を与えかねません。

研究者個人の評価については、研究開発評価とは切り離し、組織のミッションに従った公正な人事評価を行うのが原則であると言えます。しかも、その考え方は過度に信賞必罰を反映した査定的なものではなく、真面目な研究者を奨励し、より良い成果を生み出すために支援的なものでなければなりません。

そのためには、組織のミッションに結びつく研究者個人の活動を列挙し、その枠組みの中での活動にどれだけ貢献（努力）したかを問うような多様な評価軸を構築するべきでしょう。何ら該当する活動がない、あるいは貢献（努力）の程度が低い人間がマイナス査定になるのであれば、正直者が報われることを担保することができます。

◆事例　Ａ大学での研究者評価

Ａ大学では、教員評価を次のように実施しています。

１．教員は組織が認める活動内容（評価室が設定する評価シート）について、今後２年間の努力目標を提示する。

２．目標期間が経過した後に、自己評価を行い、採点する。

３．教員の所属する学部の長が教員の自己評価結果に対して認定評価を行い、評価項目毎に大きな開きがあるものについては、教員と学部長、ならびに大学評価室の３者で話し合い、最終的な評価結果を得る。

４．最終的な評価結果を賞与の決定や研究開発予算の配分に反映させる。

この評価方式の優れたところは、教員と大学側との間に双方向的な意見調整のプ

ロセスがあることです。大学評価室の立場は評価の枠組みを設定し、第三者的な立場で意見調整を図っています。被評価者である教員はその機会において自らの活動と努力水準に対する申し開きが可能となっています。

おそらく、A大学では評価を繰り返し実施していくことで、活動の努力水準に対する評価基準というものは評価実施者側と被評価者側とで共通に認識されるようになるでしょう。このような取り組みがその大学における組織文化や人材育成方針の形成に必ずや役立つと考えられます。

2.3.2　研究開発評価と SABC 評価

現状では各省庁は大綱的指針に則って個別研究開発事業の評価を実施しています。一方で、総合科学技術会議が重点的施策として認定している国家的研究開発事業については、総合科学技術会議において異なる評価基準で評価を実施し、最終的に SABC という優先順位付けを行います。

この評価結果は予算配分に反映され、継続案件であれば次年度の予算の増減に直結します。評価結果と予算査定の関係は、S が予算増、A が現状維持、B が再審査、C が予算減というかたちになっています。

しかしながら、行政機関が行う政策の評価に関する法律（政策評価法）に則って事前評価を実施し、さらに大綱的指針に則って中間評価を実施している研究開発事業が、継続を問題とする際に他の研究開発事業との比較において再評価・再査定されるという現状に対して、各省の事業担当原課・評価部署に混乱が生じています。

ひとつの考え方として、総合科学技術会議における重点的科学技術領域ならびに資金量の大きい競争的資金の評価は、イノベーション政策という独自の観点から評価を実施しているため、省庁をまたぐ予算配分調整の最終的な諮問機関として機能しているのだと認識して、そこでの評価が正当に行われるよう、各省における研究開発評価はより正確な評価を実施し、事業の正当性をアピールするための作業であると考えることができます。ただ、このように考えたとしても、基礎研究や学術研究に対する研究開発評価の結果が、社会経済的効果を重視するイノベーション政策の評価の観点とどのように整合性を保持できるかという点で、まだまだ総合科学技術会議の評価を見直す余地はあると考えられます。

SABC 評価には一時期、相対評価の考え方が導入されたこともあり、その評価結果に対しては相当な反発がありました。予算配分と直結する仕組みに移行したのも近年になってからの傾向です。現在では総合科学技術会議の評価結果の影響力は非常に大きいと言えます。

2.3.3　研究開発評価と独立行政法人評価

独立行政法人評価は独立行政法人通則法に則って実施される評価制度です。形式的には所轄省庁の大臣が定めた中期目標期間における中期計画を達成しているかどうかを毎年評価するもので、独立行政法人が行う事業の進捗状況（改善状況）とその成果を評価します。

研究開発評価と独立行政法人評価が関係するのは、主に研究開発型独立行政法人と研究資金配分型独立行政法人の評価においてです。大綱的指針の定義に従えば、独立行政法人が実施している競争的研究資金、基盤的研究資金、重点的研究資金に関する評価が関係してきます。それぞれの資金分類における研究開発評価は事業評価として実施されていますが、独立行政法人評価は機関全体の観点から評価が行われます。このように、研究開発を対象とする事業を実施している独立行政法人は、大綱的指針に基づく研究開発評価と独立行政法人通則法に従った評価の２種類の制度の網に掛かっています。

問題は、研究開発事業の評価結果を独立行政法人評価に活用することができるかどうかという点です。実施している研究開発事業の評価情報を積み上げ、独立行政法人評価の枠組みの中で整理・分析することができれば、評価作業の二度手間にならず、効率的に評価することができます。

しかしながら、比較的早い時期に独立行政法人に移行した機構においても、独立行政法人評価の本格的な枠組みを構築しているところはまだ少なく、現状では多くの機構において、試行錯誤の段階にあると言えます。その原因のひとつには、事業を対象とした評価を組織の事業部やその下の部署の活動を対象とした評価に組み替えることが簡単にはいかない点にあります。

通常、一つの研究開発事業を展開するだけでも、複数の部署がさまざまなかたちで関与します。そして一つの部署はさまざまなかたちで研究開発以外の事業も含む複数の事業に関与することになります。事業と部署との関係はこのように「多対多関係」になっています。つまりどちらか一方から見た関係だけで他方を定義できるのではなく、両方から参照可能な第３のカテゴリーが必要になります。この第３のカテゴリーを具体的な形で定義することが難しいために、２つの評価制度の橋渡しが上手くいかないのです。

本テキストの「応用編」では、研究開発評価と独立行政法人評価とをどのように橋渡しすべきかについて解説します。

3　評価の枠組み

3.1　評価はどのように行うか？

近所にある総合スーパーのＡ店とＢ店を評価するといった場合、誰かに説明できるような仕方で評価するにはどのようにしたら良いでしょうか？

おそらくは「なぜ、そのような評価結果になるのか？」を具体的に解説することになるでしょう。では、総合スーパーＡ店とＢ店を評価するための具体的な枠組みを考えてみましょう。

3.1.1　評価対象の構成を考える

まずはそれぞれのスーパーで比較可能な要素、つまり評価する対象の構成要素を洗い出してみましょう。仮に、図 3-1 のような構成要素を選んだとします。

図 3-1　評価対象の構成要素の選定

これだけでは、何を評価したいのかまだ説明しづらいので、さらに詳細に要素を展開したとします（図 3-2)。

図 3-2　評価対象の構成要素の展開

これで各スーパーのどのような要素を評価したいのかという全体像を示すことができました。

3.1.2　評価項目・評価基準の枠組みを考える

次に、こうした要素に関してどのような観点で評価をするのかについて説明する必要があります（図 3-3）。

図 3-3　評価項目の設定

上に示したように、スーパーの各要素を「価格」、「品質」、「品揃え」、「利便性」の４つの項目で評価するとします。しかしながら、これらの評価項目はそのままでは具体的な要素を評価する際に使えません。個別の要素について具体的に適用する際の説明が必要になります。このようなより詳細な評価項目の設定を「評価基準」を定めると言います。

表 3-1 は、スーパーの評価に際してどのような評価基準が設定できるかを例示したものです。

表 3-1　評価基準設定の例

		価格面	品質	品揃え	利便性
食料品売り場	生鮮食品	・価格が安いか？	・新鮮か？ ・産地が表示されているか？	・品揃えは豊富か？	・レイアウトが適切か？
	加工食品	・価格が安いか？	・総菜等は作り置きしていないか？ ・味は良いか？	・品揃えは豊富か？ ・オリジナルな商品が多いか？	・レイアウトが適切か？
その他の売り場	衣料品	・価格が安いか？	・縫製はしっかりしているか？ ・デザインは良いか？	・ブランド物を扱っているか？ ・サイズは十分用意されているか？	
	日用品	・価格が安いか？	・耐久性はあるか？ ・安全か？	・定番商品が揃っているか？ ・詰め替え用が用意されているか？	
その他の施設	駐車場	・無料の条件は何か？			・エレベーターが設置されているか？ ・カーゴを使用できるか？
	飲食店	・価格が安いか？	・味は良いか？	・店舗の種類は多いか？	・待ち時間は短いか？ ・子供連れでも大丈夫か？
サービス機能	ポイントカード	・ポイント還元率は高いか？			・有効期限は十分か？
	電子マネー				・ポイントとの交換が可能か？ ・チャージャーは完備されているか？

3.1.3　評定区分を設定する

さて、ここまで明らかにすれば評価を実施することができるでしょうか？答えは否です。各評価基準がどのような尺度で評価されるかを明らかにしないと第３者には分かりません。その手続きを「評定区分」を定めると言います。

表 3-2 は評定区分の設定の例です。

表 3-2　評定区分の設定の例

評価基準	評定区分	評点
価格が安いか？	同地域の他店と比べてかなり安い	4
	同地域の中では比較的安い	3
	同地域の他店とあまり差はない	2
	同地域の中では比較的高い	1
新鮮か？	その日に仕入れたものだけを扱っているのでかなり新鮮である。	4
	適切な鮮度管理を行っている。	3
	新鮮とは言えないが、痛んでいるようなものは置いていない。	2
	痛んでいるものがある。	1

表 3-2 では評定区分に対応した評点が定義されています。この数値の尺度については特に意味がなく、S、A、B、C でも構いません。評価基準がどの程度の水準で評価されているかを区別しているだけです。

このようにして、すべての評価基準に対して評定区分を設定し、評点を付けて、はじめてスーパーＡ店とＢ店の各要素についての客観的な比較が可能になります（図 3-4)。

図 3-4　個別要素に関する評価結果の表示の例

3.1.4　総合評価をするためには？

図 3-4 はスーパーＡとスーパーＢの比較をさまざまな観点から行い、個別の評価基準毎の優劣がはっきりと示されています。評価で大切なのはここまでの論理展開と評点付けに至るデータや情報が得られるかということです。多くの場合は、ここまでの情報が開示されれば、後は消費者が自分で判断することになるでしょう。

他方、このままでは全体としてスーパーＡとスーパーＢとではどちらが良いのかという疑問に答えることはできません。このように、個別要素に関する評価基準毎の比較ではなく、評価対象全体を評価することを「総合評価」と呼びます。

総合評価のためには、個別の評価結果に対する「重み付け」が必要となります。重み付けとは、すべての評価結果を同じ重みで考えるのではなく、相対的に重要と思われる評点を他の評点とは区別して「重く」見ることです。つまり、評価において、評価する側が持っているある種の価値観が反映される作業となります。

総合評価のためには２つのプロセスが必要です。ひとつは、評価項目・評価基準間の重み付けです。もうひとつは、評価対象の構成要素間の重み付けです。それぞれ詳しく見ていきましょう。

まず、評価項目・評価基準間の重み付けの考え方です。下の図は、先ほどの生鮮食品に関する評価項目・評価基準のツリー構造を表しています。

図 3-5　評価項目・評価基準間の重み付けの例

今、「新鮮か？」という評価基準に対してスーパーＡに４点の評点がついていたとします。重み付けされた評点は次のように計算されます。

スーパーＡの「新鮮か？」に関する重み付けされた評点　＝　30%×80%×４　＝　0.96

同様に、他の評価基準に対する重み付けされた評点を全て足し合わせたものが、スーパーＡの生鮮食品に関する全体的な評点となります。イメージとしては、図 3-4 のレーダーチャートで示されていた評点が、重み付けをされて棒グラフに集計されたと理解して下さい（図 3-6）。

図 3-6　評価項目・評価基準間の重み付けによる評点の集計

このようにして得られた評点を評価対象別に比較すると図 3-7 のようなグラフが描けます。

図 3-7　評価対象の構成要素別に集計化された評点の比較

一方で、評価対象の構成要素についての重み付けも存在します。これは評価する側がスーパーのどのような機能を重視するかについて数値的に示したものです。

図 3-8　評価対象の構成要素に関する重み付けの例

このような重みを得られたら、図 3-7 のグラフを集計して総合的な評点を得ることができます（図 3-9）。

図 3-9　総合評点の集計

ある評価対象を総合的に評価するというプロセスはここで説明したように、必ず評価する側の価値観が反映されます。評価項目・評価基準間の重み付けは、評価行為に関する価値観を表しているといって良いでしょう。これはどのような観点を重視するかという評価の目的に依存します。また、評価対象の構成要素に関する重み付けは、評価対象のどのような用途・機能を重視するかといった利用頻度などの現実的なニーズに依存しています。例えば経営側の観点からは、部門毎の来客数や売上高といった客観的なデータを基に重み付けを行うことが想定されます。

3.1.5　ここまでのポイント

評価の枠組みは、ここまで説明したような整理の仕方が基本となります。
まず注意しなければならないのは、評価対象の構成要素の展開と、評価項目・評価基準の展開を論理的に区別することです。この区分の仕方が曖昧だと、何を評価しているのか分からなくなってしまいます。基本的には次のような理解が有益でしょう。

・評価対象の構成要素は「評価される内容」を表しています。
・評価項目・評価基準は「評価の仕方」を表しています。

また、

・評定区分は「評価の物差し」を表しています。

多くの場合は、評価対象の構成要素別に評点が得られれば、レーダーチャートなどに図式化することで視覚的に評価結果を示すことができます。

しかしながら、評価対象の構成要素が多岐にわたったり、評価項目・評価基準の展開が詳細に行われたりすると、ある程度集計化した評価情報を示す必要が出てきます。評点そのものは単位を持たないただの主観的尺度なので、本来は足し合わせることができません

が、評点を付けた一人の人間の中での基準となる価値尺度が一定の範囲に収まるものとして、便宜的に集計を行います。

集計する際に重要となるのは、それぞれの評価項目・評価基準の間の重み付け、さらには評価対象の構成要素間の重み付けです。

・評価対象の構成要素間の重み付けは「評価される内容の重要度」に依存します。

・評価項目・評価基準間の重み付けは「評価の仕方に関する価値観」に依存します。

重み付けを行って評点を集計したものが総合評価の結果になります。

こうした手続きを一覧表にして整理したものが「評価シート」と呼ばれる評価のフォーマットです。一般的に、研究開発事業の評価では、外部有識者をパネリストとして招いて評価結果を導く「外部評価」が行われます。パネリストにどのような内容をどのような観点で評価してもらうかを説明するのに用いられるのが評価シートになります。実のある評価を実施してもらうには、よく練られた評価シートが必要不可欠です。

次節からは、研究開発事業を想定して、評価シートを作成する手続きと留意点について詳しく解説します。

◆コラム 3-1：相対評価と絶対評価

ある目的を持った評価の枠組みの中で、対象を評価する際に、相対評価と絶対評価という言葉が使われることがあります。それぞれの厳密な意味をここで理解しておきましょう。

我が国における教育評価の変遷を事例に説明することが理解の助けとなるでしょう。我が国の教育現場では 2000 年くらいまで相対評価が主流でした。相対評価の考え方は「集団の絶対数が多くなればなるほど、その成績の分布はおよそ正規分布に近づく」という統計学の理論を基本としています。評価する側（教師）は成績資料を精査した後、生徒を成績順に並べ、5 段階評定の場合、5…7%、4…24%、3…38%、2…24%、1…7%を目安として一定の割合で評定をつけます。したがって、必ず評定 5 の生徒、1 の生徒が存在することになります。

しかしながら、成績の分布が正規分布である保証はなく、また、評価結果そのものはサンプルの範囲の中でしか意味を持ちません。教師が自分の担任のクラスの生徒のみを対象に評点を付けているとすれば、そのクラスの中でしか適用できない評点であり、他のクラスとの比較はできません。ましてや、他の学校との比較はもっと不適切です。これは評点を受ける側（生徒）からすればクラスの中でのランキングというだけで、当人の能力を正当に評価していることにはなりません。少子化の問題もあり、ますますサンプルが少なくなる傾向にある教育現場において、相対的評価の持つ意味が薄れたとの指摘が各方面から寄せられました。

そこで、2001 年の指導要領改定を機に絶対評価（観点別学習状況評価）が取り入れられました。絶対評価はあらかじめ設定した到達目標に対して達成度を評価基準にして評価する方法です。例えば体育の逆上がりを例に取ると、「逆上がりができるようになる」が到達目標で、評価基準は「１．補助板なしで逆上がりができた」、「２．補助板を使って逆上がりができた」、「３．補助板を使っても逆上がりができなかった」となります。したがって、クラスの全員が逆上がりをできるようになれば、皆高い評点を得ることがあり得ます。

絶対評価では到達目標の水準をどのように設定するか、評価基準をどのように設定するかが重要となります。

ちなみにランキングは評点結果を高い順から並べるという行為を意味しますので、相対評価でも絶対評価でもランキングは存在します。

3.2　評価シートの作成

我が国の文書作成の文化において特徴的かつ有用な資産は「罫線情報の活用」です。表形式に文章を配置することで、文章情報の構造化を容易にし、位置づけまでを明らかにすることができます。

外部評価者に対して評価を実施する際の考え方を文書において示そうとした場合、前節で見たように情報の構造化が不可欠になります。したがって、罫線で整理された評価シートという形で基本方針を説明するのがもっとも有効な情報提供の方法となるでしょう。

本節では、研究開発事業の評価を想定し、評価シートを作成するまでの基本的な考え方、留意点を解説します。

3.2.1　評価シートの基本構成

評価シートの基本的な構成例を図 3-10 に示します。評価シートの左側は評価対象の構成要素を列挙します。図のように再分類を示して階層化しても構いません。次に、評価対象の構成要素に該当する評価項目を列挙します。その評価項目を具体的に説明した評価基準が次の列に入ります。右側の列には評価基準に対応した評定区分の説明が入ります。個別要素の評価については、パネリストは評価基準と評定区分の説明を中心に判断基準として評点を定めることになります。

図 3-10　評価シートの構成例

また、この評価シートでは最下段に総合評価の欄が設けてあります。総合評価は本来であれば個別の評点を重み付けして集計化することで得られるわけですが、総合評価そのものの評定区分を設定して、改めて評点を付けるというプロセスを経ることがあります。

これは、評価対象の構成要素や、評価項目・評価基準間の重み付けが多くの場合、明示的に設定することができないためです。特に外部パネルを活用した評価の場合、評価者の重み付けはそれぞれ異なるため、個別の評点から総合的な評点を導くことが難しくなります。また、評価対象全体の評価結果が定まらなければ評価が難しい評価項目も存在します。例えば費用対効果などが典型的な評価項目ですが、事業が生み出した成果等に対する評価結果が先に決定されていなければ、コストパフォーマンスを論じることはできません。したがって、総合評価の欄が設定されていることと個別評価が展開されていることは独立しているのではなく、補完的な関係にあると理解して下さい。

図 3-10 の評価シートは評価対象の構成要素から成る「評価対象系」と評価目的から展開される評価項目・評価基準から成る「評価体系」のマトリックス構造を有している点に注意して下さい。マトリックス構造とは、行と列との対応関係が要素によって関連づけられている表形式のことです。

図 3-11　評価マトリックス

評価マトリックスは行方向に「評価対象系」を、列方向に「評価体系」を展開し、それぞれの対応は評定区分に基づく「評点」によって関連づけられています。この評価マトリックスを評価対象系の構成要素別に評価項目・評価基準を繰り返し適用して並び替えたものが評価シートになっています。

評価を担当する方にとっては、この評価シートを適切に作成することがもっとも重要な仕事になります。評価シートを作成することで初めて、評価対象の何をどのように評価するのかを外部有識者（パネリスト）に説明することが可能になるからです。

以下では、評価シートの構成に従って、どのように評価対象を評価するかについての基本的な考え方を学びます。

3.2.2　評価対象系

評価対象系とは評価対象の構成要素の体系のことです。評価対象がどのような構造を持っているかについて把握することは、評価対象を正しく理解するための第一歩です。そのために、評価対象を展開し、評価すべき要素を明らかにしなければなりません。

評価対象を展開するための有効な方法は、評価対象がどのような機能（メカニズム）を内部に持っているか、それらの機能の集合が評価対象をおおよそ説明しているものかどうかを明らかにすることです。

人体を例に取ると、人間が健康的に活動するためには内部のさまざまなメカニズムが正常に作用していなければなりません。人体を骨格系、筋系、消化器系、循環器系、神経系などからなるシステムとして捉え、それぞれの要素がどのように働いているかの理解がなければ、医師は人を診察（評価）することはできないでしょう。また、機械であれば、制御系、駆動系、伝達系等の仕組みを理解していなければ、全体としての性能を把握することは困難です。

同様に、研究開発評価においても評価対象である機関、事業、課題がどのような機能を持つ要素から構成されているかについて理解することが、評価対象を展開するための方法論となります。

a)機関評価

評価対象系の展開方法として、最初に機関評価を例に考えてみましょう。機関評価では、機関全体を評価するために個別の内部組織、内部組織が行っている業務に着目する必要があります。一般に、組織における内部組織は役割・機能に応じて細分化されますので、まずは当該する組織がどのような組織構造を持っているかが問題となります。

図 3-12　機関評価における評価対象の展開（組織図）

図 3-12 のような構造は多くの独立行政法人のひな形となっているものでしょう。機関評価では事業部や他の部署の「活動」を評価することになります。図では事業部等における単独ないしは複数の事業（または活動）の評価が事業部等の評価につながり、機関全体の評価に統合される様子を表しています。

機関評価の場合、評価対象の展開に必要なことは、個別事業や活動が個々の事業部等に対応関係として把握されていることです。現在、独立行政法人は中期目標や中期計画において目標期間中にどのような事業を展開するかについて事業計画を記述しなければなりませんが、機関評価において重要なことは、それらの事業をどの部署が責任を持って実施するのかというアサイン（割り当て）が行われているかどうかです。これらの対応関係を基にして、機関の何を評価するのかが決定されます。

b) 研究開発事業

研究開発事業では、その事業が何を達成しようとしているのか、そのためにどのような機能を内部に持っているのかが評価対象系の展開の重要な点になります。図 3-13 は研究開発事業の評価対象系を展開した例です。

図 3-13　研究開発事業における評価対象の展開

研究開発事業等の評価では、大きく分けて二つの要素を評価します。それは事業の「実績（performance）」と「リソース（resources）」です。実績は事業が何を成したのか（あ

るいは成そうとしているのか）を表しており、事業内容に係わる要素から構成されます。リソースは事業にどれだけの人・モノ・金が投入されたか等を表しています。

実績はさらに「成果（products)」と「過程（process)」に分かれます。成果は事業が直接間接に生み出したモノや影響（アウトプット，アウトカム，インパクト）[8]といった要素から構成されます。過程は、事業が成果を生み出す過程における方法（研究開発システム）や研究者、研究グループの貢献（アクター)、マネジメントなどといった要素から構成されます。

リソースは当該事業に投入された人・モノ・金（インプット）の他に、事業実施者が持つ技術力、設備、ノウハウなどのストックの事業への寄与分から構成されます。

このように、研究開発事業を成り立たせている機能・要素を構造化し、当該事業に適した具体的な要素まで落とし込む作業が必要になります。

3.2.3　評価体系

評価体系とはある評価目的を達成するために論理的に展開された評価の観点をいいます。例えば、事前評価であれば事業の必要性を評価するための論理展開がなされているか、中間評価であれば研究開発目標の達成度と改善点等を評価するための論理展開がなされているかどうかが重要となるでしょう。評価体系は評価目的のために大きな評価の観点として評価項目が設定され、その下に詳細な観点からの評価基準が設定されます（図 3-14)。

図 3-14　評価体系

評価体系の構築には厳密な論理的思考（ロジカル・シンキング）が要請されます。その理由は、仮に上の図で評価項目Ａと評価項目Ｂとが依存した関係にあったとしたら（評価項目Ａの観点が評価項目Ｂの観点と意味的に区別がつかないような場合)、評価者（パネリ

[8] アウトプット，アウトカム，インパクトの概念定義については、応用編「４．研究開発評価の体系」にて詳しく説明します。

スト）が評価を実施する際に判断に困ることになるからです。ロジカル・シンキングの作法に従えば、図 3-14 は次のような論理関係にあります。

図 3-15　評価体系の論理関係

ロジカル・シンキングでは、論理を展開する際に二つのルールを使います。一つはある要素の構成要素が相互に重なりがなく、全体として漏れがないこと（MECE：Mutually Exclusive Collectively Exhaustive）。もう一つがある要素とその構成要素との関係が結論と根拠（So what? / So why?）の関係を保っていることです。

ある評価体系を考えるときに評価の観点が３つあったとしたら、

- 評価項目それぞれが独立の観点であり、その３つを明らかにすればほぼ評価目的を達成することができること。
- ３つの評価項目を明らかにすることが評価の結論と根拠を述べていることになっていること。

を満たしていれば論理的に展開されていることになります。同様に、ある評価項目に関して、評価基準が３つあったとすれば、

- 評価基準それぞれが独立の観点であり、その３つを明らかにすれば該当する評価項目の内容をほぼ表現することができること。
- ３つの評価基準を明らかにすることが評価項目の結論と根拠を述べていることになっていること。

を満たしていれば、評価基準が論理的に展開されていることになります。

では、研究開発事業の場合ではどのように評価体系を構築すべきでしょうか？

3.2.4　研究開発事業の評価体系の考え方

我が国の研究開発事業の評価体系を展開する際には、大綱的指針や政策評価法の枠組みにしたがって考察することが有用です。大綱的指針では研究開発評価の観点として、「必要性」、「有効性」、「効率性」の３つを挙げています。これらを、評価項目を括るカテゴリーとして考えると図 3-16 のような論理展開が可能です。

図 3-16　研究開発事業の評価体系の展開

必要性とは、研究開発事業を実施することの正当性を述べることです。その説明要素としては、事業を実施する政策上の理由と位置づけが考えられます。理由については、上位計画（総合計画）や総合戦略と事業内容とが整合的かどうかという点がチェックされます。位置づけについては、施策体系の中での位置づけと、歴史的展開の中での位置づけ（政策の連続性）が適切かどうかチェックされます。

有効性とは、研究開発事業が達成する中身についての妥当性を述べることです。これに関しては事業の目的とそれに対応する成果の両面から説明することができます。目的については、事業がどのような効果を意図しているのか、また、それに資するための研究開発目標は何であるかがはっきりと定められているかどうか、目標水準は適切なものかどうかがチェックされます。一方、成果については、意図した研究開発成果が得られているかどうか、目標達成度はどのようになっているかがチェックされます。

効率性とは、研究開発事業を実施する上でどの程度効果的な運用が図られたかどうかに

ついて述べることです。これを評価するためには研究開発の手段（研究開発方法とそのマネジメント）や体制（人・組織）、コストの面から精査することが必要です。

このように評価体系の論理構造を明確にすることで、「有効だから必要だ、効率的だから有効だ」といった評価における初歩的な間違いを避けることができ、それぞれを説明するための評価基準を明確に設定することが可能になります。

評価基準については、評価対象となる個別の研究開発事業の特徴にあわせてさらに詳細に展開することが必要になるでしょう。例えば、原子力や宇宙開発のような巨大技術に係わる研究開発事業であれば個別の要素技術の開発に関する評価だけではなく、システム全体についての評価基準があって然るべきです。研究拠点（COE）や技術移転機関（TLO）のような拠点形成を目的とした事業であれば、産学連携の進捗状況や地域交流の活性化等に関する評価基準が考えられるでしょう。また、研究者だけでなく、インテグレーターやコーディネーターといった役割を担う人材に対する評価の観点も必要になるでしょう。

このように、当該事業の目的に応じて、評価体系で重視すべき評価項目・評価基準をカスタマイズすることが重要です。

3.2.5　対応関係の考察

評価対象系と評価体系が定まったら、次はそれぞれの対応関係をもっとも詳細なレベルで考察する作業に移ります。評価対象系では評価すべき個別の要素まで展開されたものと、評価体系では細分類のレベルの評価項目とで対応関係を表にしてまとめます。対応関係があれば、星取り表のイメージで○印を付け、後から○印の中身を具体的に検討しましょう。

表 3-3　評価対象系と評価体系の対応関係

評価対象系＼評価体系			必要性				有効性		効率性			
			理由		位置づけ		目的	成果	手段	体制		コスト
			上位計画	戦略	施策体系	施策展開				制度	人・組織	
実績	成果	アウトプット					○	○				○
		アウトカム	○		○	○	○	○				○
		インパクト	○	○		○	○					
	過程	アクター		○					○		○	
		マネジメント							○	○	○	○
		システム	○	○						○	○	
リソース		インプット										○
		ストック										○

表 3-3 はこれまで見てきた評価対象系と評価体系との対応関係をプロットした例です。もちろん、これらの対応は研究開発事業の目的や特徴に応じて変化します。この対応関係表を基に、評価対象の構成要素別に評価基準をカスタマイズして、具体的な記述を盛り込みます。

表 3-3 で示された対応関係毎に評価基準の具体的な記述と評定区分の定義に関する記述が完成したら、表 3-4 のように列項目（評価体系）を行項目（評価対象の構成要素）毎に対応させて配置します。これが評価シートの完成型となります。

表 3-4　評価シートの完成型

評価対象系			評価体系			評価基準	評定区分
実績	成果	アウトプット	有効性	目的		具体的な記述	具体的な記述
				成果			
			効率性	コスト			
		アウトカム	必要性	理由	上位計画		
				位置づけ	施策体系		
					施策展開		
			有効性	目的			
				成果			
			効率性	コスト			
		インパクト	必要性	理由	上位計画		
					戦略		
				位置づけ	施策展開		
			有効性	目的			
	過程	アクター	必要性	理由	戦略		
			効率性		手段		
				体制	人・組織		
		マネジメント	効率性	手段			
				体制	制度		
					人・組織		
				コスト			
		システム	必要性	理由	上位計画		
					戦略		
			効率性	体制	制度		
					人・組織		
リソース		インプット	効率性	コスト			
		ストック	効率性	ストック			

評価シートを作成する上で大切なことは、評価対象の構成要素については、繰り返しはなく、評価項目に関しては異なる評価基準で繰り返し質問されるということです。例えば、アウトプットに対するコストの評価基準は、「他の研究開発方法に比べて低コストで実験データが得られたか」などの記述が想定できる一方で、マネジメントに対するコストの評価基準は、「研究開発フェーズ毎の研究開発期間を短縮することができたか」などの記述が想定できます。

評価対象系と評価体系とを論理的に区別し、評価対象の構成要素毎に評価基準を作成することができれば、評価の枠組みはほぼ整ったと言えます。

3.2.6　評定区分の設定と評点

評価基準に対する回答の方法としては、事実関係を文章で記述する場合と何かしらの価値観が反映される場合とがあります。前者の場合はきちんと記述されているかどうかという判断基準しか適用することができませんが、後者の場合は評価基準に対する評価者の思考の枠組みをある程度設定しないと適切な評価結果を得ることができません。

評価基準のうち、評価者の判断に左右されるようなものについては、評価の尺度を定める必要があります。この尺度を評定区分といいます。またどの評定区分に該当するかを形式的にあらわしたものが評点です。

【例】
評価対象系：実績－成果－アウトプット－○×技術の開発
評価体系：成果－（評価基準）期待される成果が得られているか？
評定区分：
　Ｓ．研究開発目標と比して期待以上の成果が得られている
　Ａ．期待（当初の目標）通りの成果が得られている
　Ｂ．期待していた（当初の目標）よりはやや不足である
　Ｃ．当初の目標を全く達成していない

評定区分はできるだけ客観的に判断できるよう、具体的に設定する必要があります。ここでは研究開発計画において達成しようとしていた目標水準を基準にして評定区分を設定しています。

しかしながら、上の例では、当初の研究開発目標（ミッション・ステートメント）の水準が極めて挑戦的であった場合、ＳとＡの解釈が変わってくることも想定できるでしょう。したがって評価者の主観や価値観を完全に排除した評点付けというものはそもそも存在しません。

ここで重要なのが、「評点付けを行う場合には必ず理由をコメントのかたちで付してもらう」ということです。なぜＡではなくてＳを付けたのか？その理由を評価者から収集することが外部評価を実施することの主要な目的となります。

◆コラム 3-2：評価者（パネリスト）の間で評点が割れた時

ある研究開発事業の評価において評価パネルを活用して評点を付けたとします。この時、パネリストの付けた評点が次のような分布になったとします。

ケース１

ケース２

ケース１は評点が完全に割れた状態です。パネリストの間で当該事業に対する評価結果が正反対に出ているケースです。また、ケース２は少数のパネリストがまったく違う見解を示しています。このような場合、他のパネリストにはない独自の観点から評価意見を述べていることが想定できます。

いずれの場合においても、パネリストの間で意見調整を行う必要があるでしょう。新しい知見が得られる機会としてむしろ積極的に意見交換を図るべきです。なぜならば、一意の評価結果を決めることよりも、評価結果に至る理由を明らかにすることがより重要だからです。

時間的制約があるために、最終的にはパネル責任者（評価委員長）の権限で評価結果を一任するにしても、評点の集計には少なくとも２回のプロセスが設定されている必要があります。最終的にパネル間で合意に至らずとも、反対意見や少数意見はコメントのかたちで留意事項や課題として活かせることになります。

3.2.7　ここまでのまとめ

第１章では、評価を実施する上での心構えを解説しました。特に評価の理念については、後の章でも繰り返し登場することになります。第２章では我が国の研究開発評価制度の成り立ちと概要を解説しました。評価業務を担当した場合、自らの業務がどのような評価制度の下に展開されているか、それを常に意識しておく必要があります。

第３章では、評価の基本的な仕組みをスーパーの比較という例題に沿って学びました。ここでは、評価する対象と評価の論理を区別することが重要になります。また、総合的な評価が決まる仕組みを考えることは、本来であればいかに前提がしっかりしていなければならないか、ということも学ぶべき点として重要です。システマティックな評価が現実にはそれほど多く取り入れられていないことの理由は、主観的にしか評価ができない対象が多いのではなく、全体として評価するためのロジックを厳密に構築することが困難だからという理由によります。

続く研究開発評価に係わる評価シートの作成の部分では、評価対象の区分の仕方や評価項目の展開の仕方を学びました。ここまでの内容を理解された方は、「研究開発事業の成果は妥当であるか？」とか「体制・マネジメントは妥当であるか？」といった乱暴な評価基準を作成することはないでしょう。担当する事業の評価において、少なくとも外部評価者がさまざまな「妥当性」の根拠を客観的に判断できるように評価の枠組みを評価対象に応じて真剣に考えなくてはなりません。

評価対象をより深く理解しなければ、具体的な評価基準を示すことはできません。したがって、具体的な評価基準を設定するのは、事業を実施している担当部署の人間がもっとも適任となります。研究開発事業を企画し、マネジメントを担当し、研究開発評価の具体的な評価基準を設計できる人材をプラクティショナーと呼びますが、プラクティショナーは実務経験を積まなければその役割を十分に果たすことができません。少なくとも、関連する研究開発領域の「畑」を歩んで来られた方が事業推進部署にいないと、当該研究開発事業の評価シートを正しく設計することは難しいでしょう。

研究開発評価を形式的なものにしないためにも、優れたプラクティショナーの育成・蓄積が急務となっています。応用編テキストではより研究開発評価に特化した内容を体系立てて議論しますが、そこで展開される議論を踏まえてようやくプラクティショナーの入り口に立つことができるとイメージして下さい。

本テキストは、プラクティショナーとなられる皆さんが、評価の現場で困難な事象に遭遇したときに立ち戻るための考え方を示しているに過ぎませんが、これまで評価に係わってきたプロジェクト・マネージャーや評価パネリストの意見をかなり反映したものとなっています。評価の現場では新しい問題点が日々生まれていますが、まずは研究開発評価に関するこれまでの議論の中から比較的整理され、共有されている経験・考え方を学んでいきましょう。

３．研究開発評価研修プログラム（応用編）

教材案

研究開発評価研修プログラム（応用編）教材案

研究開発評価研修プログラム（応用編）

目次

1　研究開発評価の体系 1
1.1 研究開発評価の意義 1
1.1.1 大綱的指針が示す研究開発評価の意義 1
1.1.2 研究者が抱く評価への懐疑的な思い 2
1.1.3 評価と説明責任 3
1.1.4 評価と意思決定の関係 3
1.2 研究開発評価の理念 6
1.2.1 研究開発評価の理念とは何か 6
1.2.2 研究開発評価の理念はどうあるべきか 7
1.2.2.1 支援的であること 7
1.2.2.2 非排除・双方向的であること 9
1.2.2.3 明示的であること 9
1.3 研究開発評価の局面 11
1.3.1 評価対象の区分 11
1.3.2 研究開発の実績の区分 15
1.3.2.1 成果の区分 16
1.3.2.1.1 アウトカム 16
1.3.2.1.2 アウトプット 19
1.3.2.1.3 インパクト 21
1.3.2.2 成果以外の実績 23
1.3.2.2.1 アクター 24
1.3.2.2.2 マネジメント 25
1.3.2.2.3 システム 27
1.3.3 研究開発評価の実施時期と重視すべき内容 28
1.3.3.1 評価時期 29
1.3.3.1.1 研究開発の計画＝実施＝普及の各段階と評価時期 29
1.3.3.1.2 政策評価法における事前評価と事後評価 29
1.3.3.1.3 評価時期毎の重要性 32
1.3.3.2 評価時期毎の重視すべき評価項目 34
1.3.3.2.1 事前評価で重視すべき評価項目 34
1.3.3.2.2 中間評価で重視すべき評価項目 38
1.3.3.2.3 事後評価で重視すべき評価項目 40

1.3.3.2.4 追跡評価で重視すべき評価項目 42
1.4 研究開発評価に関わる専門人材 43
1.4.1 レビューア 43
1.4.2 プラクティショナー 44
1.4.3 アナリスト 49
1.5 研究開発評価の方法論 50
1.5.1 方法論の一般論 50
1.5.1.1 調査＝分析＝評価のための方法論 50
1.5.1.2 定性的評価と定量的評価 53
1.5.2 定量的評価法の基本 54
1.5.2.1 比較考量の基本的な考え方 54
1.5.2.2 数値データはどのように扱うべきか 58
1.5.3 研究開発の「質」に関わる評価法 63
1.5.3.1 評価対象に応じたレビューの仕組み 63
1.5.3.2 パネル運営 64
2 従属型プロジェクトの評価 69
2.1 従属型プロジェクトのマネジメント 69
2.1.1 公募実施のための準備 71
2.2 レビュー評価 74
2.2.1 レビューア 74
2.2.1.1 レビューアの確保 74
2.2.1.2 レビューアの選定 75
2.2.2 レビュー・システム 75
2.2.3 レビュー・マネジメント 79
2.2.3.1 レビューアの所属機関に対する配慮 79
2.2.3.2 評価パネルの運営 80
2.2.3.3 レビューアの能力と客観性 80
2.2.3.4 パネル間、分野間の正規化 80
2.2.3.5 評価項目としての総合評価 81
2.2.3.6 評価の匿名性 81
2.2.3.7 ピアレビューのコスト 81
2.2.3.8 評価の倫理基準 82
2.2.3.9 その他の留意事項 82
2.3 レビュー評価の手続き 85
2.3.1 レビュー・システムの設計 85
2.3.2 レビューア・データベースの作成 86

2.3.3 レビュー評価指針の作成 88
2.3.4 レビューアへの委嘱手続き 90
2.3.5 評価パネルの設計 90
2.3.6 採択の判定 91
3 プログラム（施策・制度等）の評価 93
3.1 プログラムを評価するとはどういうことか？ 94
3.2 ＲＯＡＭＥＦの設定 96
3.3 プログラム評価のための準備 100
3.3.1 プログラムの再設計 100
3.3.1.1 必要性の再設定 100
3.3.1.2 有効性の再設定 101
3.3.1.3 効率性の再設定 103
3.3.2 ロジックモデルの作成 106
3.3.2.1 プログラムとロジックモデル 106
3.3.2.2 ロジックモデルの作成の仕方 108
3.3.2.3 ロジックモデルのチェック 111
3.3.3 パフォーマンス指標の設定 112
3.4 プログラムの事前評価 116
3.4.1 プログラムの目標 116
3.4.1.1 成果に関わるプログラムの目標 116
3.4.1.2 マネジメントに関わるプログラムの目標 117
3.4.2 評価システムの運用 118
3.4.2.1 評価の体制 118
3.4.2.2 評価の仕組み 119
3.4.2.3 評価結果の活用の方法 119
3.5 プログラムの中間・事後評価 120
3.5.1 中間・事後評価の意義 120
3.5.2 プロジェクトの実績のデータ 120
3.5.2.1 インプットデータの類型化 120
3.5.2.2 アウトプットデータの類型化 122
3.5.2.3 アウトカムのデータ 123
3.5.3 プログラム固有のデータ 124
4 独立型プロジェクトの評価 127
4.1 独立型プロジェクトに要求されること 128
4.1.1 プロジェクトと創造のプロセス 128
4.1.2 プロジェクトの経済性 131

4.2 独立型プロジェクトの評価とは. 135
4.2.1 事前評価の段階 135
4.2.1.1 プロジェクトの事前評価項目 135
4.2.1.1.1 プロジェクトの目的・目標 136
4.2.1.1.2 プロセスの設計 137
4.2.1.1.3 体制・マネジメントの設計 138
4.2.1.2 プロジェクトの構造化とロードマップ 139
4.2.1.3 実施体制・責任の明記 141
4.2.1.4 費用分析の実施 141
4.2.2 独立型プロジェクトの中間評価 143
4.2.2.1 各サブ・プロジェクトの進捗状況 143
4.2.2.2 予算執行状況 143
4.2.2.3 意思決定・R&D マネジメント・体制 143
4.2.3 独立型プロジェクトの事後評価 144
4.2.3.1 目的・目標達成状況 144
4.2.3.2 プロジェクトのマネジメント 145
4.2.3.3 副次的成果 146
4.2.4 独立型プロジェクトの追跡評価 147
4.2.4.1 追跡評価の実施時期 147
4.2.4.2 追跡評価の方法論 147
5 機関の評価 149
5.1 機関評価はなぜ難しいのか. 150
5.2 機関評価の枠組み. 152
5.2.1 機関の何を評価しなければならないのか？ 152
5.2.1.1 ミッションから活動まで 152
5.2.1.2 機関のアウトカム・インパクトとは？ 154
5.2.1.3 活動と組織構造 156
5.3 機関評価の方法論. 159
5.3.1 業務分析の実施 159
5.3.2 活動へのコストの割り当て 161
5.3.2.1 活動基準会計（ABC） 162
5.3.3 マネジメントの評価 163
5.3.4 スタッフ組織の業務の位置づけ 163
5.3.4.1 マトリックス型組織とクロス・ファンクショナル・チーム 163
5.4 機関評価の方法論を踏まえた中期目標・中期計画の策定. 166
5.4.1 独立行政法人評価の枠組み 166

5.4.2 中期目標・中期計画の策定 168
5.4.2.1 組織のプログラム化 169
5.4.2.2 業務運営の効率化 169
5.4.3 年度計画の策定 169

研究開発評価研修プログラム（応用編）教材案

1　研究開発評価の体系

1.1　研究開発評価の意義

科学技術基本法の規定に則り、第Ⅰ期科学技術基本計画が閣議決定されてから 12 年、最初の大綱的指針（1997 年）が閣議決定されてから 10 年、独立行政法人通則法が定められてから８年経過し、さらに、行政機関が行う政策の評価に関する法律（政策評価法）が公布されて 5 年が経過しました。この間、我が国の研究開発評価はかなりの部分まで制度化されてきたと言えます。

しかしながら、この間、研究開発評価を巡る環境の中で問題視されたことは、研究開発事業等を実施している側の評価業務に関わる負担感からの「評価疲れ」や、被評価者（＝研究者）が評価結果を意識しすぎて、本来の研究開発の質が担保されていないのではないかということでした。

それでは研究開発評価は何のために実施されているのでしょうか。この素朴な疑問に対する答えは色々な観点から導かれますが、大切なことは「評価の行為と結果をどのように活かしていくべきか」という点です。

評価そのものも一定のコストと時間を費やして行われる業務ですので、それに関わる人間に上手く働いてもらうためには目的意識が必要です。また、目的のより上位にはそれを支える理念が必要でしょう。

ここでは、大綱的指針に記述されている研究開発評価の意義を具体的に解説します。

1.1.1　大綱的指針が示す研究開発評価の意義

最新の「国の研究開発評価に関する大綱的指針」では、冒頭に評価の意義として以下のような内容が取り上げられています。

- 評価を適切かつ公正に行うことにより、研究者の創造性が十分に発揮されるような、柔軟かつ競争的で開かれた研究開発環境の創出を実現することができる。
- 評価を支援的に行うことにより、研究開発の前進や質の向上、独創的で有望な優れた研究開発や研究者の発掘、研究者の意欲向上、より良い政策・施策の形成等の効果が得られる。
- 評価結果を積極的に公表し、優れた研究開発を社会に周知させることにより、研究開発に国費を投入していくことに関し、国民に対する説明責任を果たし、広く国民の理解と支持が得られる。
- 評価結果を適切に予算、人材等の資源配分に反映することにより、研究開発を重点的・効率的に行うことができる。

出所：「国の研究開発評価に関する大綱的指針」（平成 17 年 3 月内閣総理大臣決定）p.3

ここで例えば、「評価を適切かつ公正に行う」と、なぜ「研究者の創造性が十分に発揮されるような、柔軟かつ競争的で開かれた研究開発環境の創出を実現することができる」のか、疑問に思われないでしょうか。

公募を前提とした競争的研究資金制度のような仕組みを例に挙げると分かりやすいと思いますが、「柔軟かつ競争的で開かれた」制度を創設して実施した場合、制度運営の改善のために評価を行う必要が出てきます。公募段階であれば採択評価、課題終了後であれば事後評価が実施されるわけですが、これらの評価が「適切かつ公正に」実施されることで、制度の枠組みが改善されるとともに、「研究者の創造性が発揮され」、良質な課題を創出することが期待されています。

つまり、研究開発評価を実施することの意義として、まず第一に研究開発制度や研究開発事業全体のマネジメントに資するという点が強調されて良いでしょう。

1.1.2　研究者が抱く評価への懐疑的な思い

また、「評価を支援的に行うこと」で、なぜ「研究開発の前進や質の向上、独創的で有望な優れた研究開発や研究者の発掘、研究者の意欲向上、より良い政策・施策の形成等の効果が得られる」のでしょうか。

研究者は一般的に評価に対して懐疑的です。研究開発評価に対するアレルギーには次のようなものが考えられるでしょう。

研究の意図や動機は研究者の個人的なものであり、他と比較して評価できないのではないか？

自分の研究分野を評価できるほど深い知識を持った研究者がいるとは思えないので、妥当な評価ができないのではないか？

研究成果をきちんと出しておけば、評価を行う必要はないのではないか？

社会経済的に役立つ研究へ傾斜していく中で、基礎研究や学術研究がまともに評価されるとは思えない。

評価結果が個人査定に用いられると分かっているので、不都合な情報は出したくない。

評価を実施するとさまざまな書類を作成しなければならず、相当な負担が掛かり、研究に没頭する時間が減る。

出所：財団法人政策科学研究所「研究開発プロジェクト等の評価手法に関する調査」報告書，2002.

「負担が掛かる」を除くと、基本的に研究者と評価実施側との信頼関係に関わる問題であると理解できます。研究開発評価は必ずしも成果の側面だけを取り上げるわけではないので、研究者には評価の目的をきちんと説明することが大切でしょう。そして何より、評価業務に対して両者の間に継続的な信頼関係を構築するためには、「評価を支援的に行う」ことが条件となります。

評価実施側が最初から査定的な態度で評価に臨むと、信頼関係の構築には決して至りません。事前や中間評価段階では、どのようにすれば研究開発事業や研究開発課題がより良いものになるかについて双方が協議し、必要な措置を講じることができます。事後評価であれば次の段階への発展的課題を抽出するなどの措置を取ることができます。

評価において何かしら査定的な側面があったとしても、査定理由が明確で両者の間で合意できるものであれば、研究者は次に何を成すべきか、どうすれば良いかにつなげることができます。評価実施側からサジェスチョンを与えることもできるでしょう。

評価業務に対する負担感についても、評価の意義を十分理解できれば必要な作業を実施するという認識が生まれ、負担感の軽減につながることが予想できます。

つまり、研究開発評価では、評価業務を研究開発者側にとって「支援的に行う」ことがもっとも重要であり、その結果として「研究開発の前進や質の向上、独創的で有望な優れた研究開発や研究者の発掘、研究者の意欲向上、より良い政策・施策の形成等の効果が得られる」ことにつながることが期待されるのです。

1.1.3　評価と説明責任

「評価結果を積極的に公開」が国民に対して説明責任を果たすことに結びついている点については広く了解が得られるところでしょう。

しかしながら、単に評価結果を開示するのではなく、なぜそのような評価結果に至ったのかを示すことができなければ意味がありません。つまり、評価が説明責任を果たすことに役立つためには、評価のプロセスそのものが明確なエビデンスに基づいて実施されなければ、評価結果の妥当性を判断することはできません。

1.1.4　評価と意思決定の関係

「評価結果を適切に予算、人材等の資源配分に反映することにより、研究開発を重点的・効率的に行うことができる。」に関しては、先の支援的評価の考え方と若干矛盾するように思われるかもしれません。個別事業の範囲の中での話であれば、マネジメントの改善により、「重点的・効率的に研究開発を実施する」ことに結びつくでしょう。しかしながら、異なる研究開発事業間の比較において、予算配分や資源配分に評価結果が反映されるとしたら、それは評価とは別の次元の話、すなわち「意思決定」の問題になります。

出所：財団法人政策科学研究所作成資料

図 1-1　評価と意思決定の関係

「意思決定」と「評価」は不可分ではないのか？という疑問を持たれる方もいるでしょう。たしかに評価は何かしらの意思決定に用いられるために存在するといえます。

しかしながら、意思決定は評価結果以外の要因も考慮した総合的な判断の上に行われます。例えば、研究開発プログラムの当該年度の予算が縮減になった、競争的環境を醸成するために『足きり』の基準が厳しくなった。同一プログラム内での類似課題を統合する必要が出てきた、等々が考えられます。

つまり、評価結果とは別の要因（それらは多分に事業を取り巻く環境、時代認識、上位政策等の変化によって移り変わるもの）によって判断の基準が規定されるため、意思決定をより合理的に行うためにも、評価行為それ自体は意思決定と独立して行われる必要があります。

ここで、評価の機能・役割とは、次のように心得ていただければよいでしょう。

評価の機能・役割とは、評価対象を「つまびらかに」することである。

つまり、大綱的指針で述べられている「評価結果を適切に予算、人材等の資源配分に反映させる」については、「予算、人材等の資源配分」に関わる意思決定を行うために研究開発評価の結果を活用することを断っているに過ぎず、意思決定そのもののプロセスは別にあることを理解しなければなりません。

例えば、内閣府の総合科学技術会議が行う SABC 評価を考えてみましょう。総合科学技術会議では各省において評価された研究開発事業を別の観点で審査して SABC の総合評点

を付けます。財務省がその評点結果を参考に、予算配分を実施します。

総合科学技術会議では、評価結果に対して、イノベーション政策の観点、他の重点的事業との比較の観点が反映されて SABC が決定されます。つまり、仮に研究開発事業の中間評価の結果が S であったとしても、総合的な判断から総合科学技術会議の場で S になるとは限りません。しかし一方で、研究開発事業の中間評価がしっかり行われていなければ総合科学技術会議において再評価することはできません。つまり、評価結果が査定的な意味合いの意思決定に活用されることがあっても、評価そのものが果たすべき役割はなんら変わることがないのです。

我が国の研究開発評価の現場において意思決定と評価が厳密に区別されていると認識できる場面は少ないかもしれません。しかし、評価は意思決定のための「お墨付き」ではありません。意思決定側が評価結果をどのように使うかについての明確な判断基準と、どの段階で実質的な意思決定がなされるかについての責任と権限が明らかになってはじめて、評価が「活かされ」るのです。

1.2　研究開発評価の理念

前述の大綱的指針における研究開発評価の意義に関する記述には、その文言を審議する過程で、この国の研究開発評価の理念なるものが形成されてきたと考えることができます。

研究開発評価の理念とは、評価行為を成り立たせている根本的な原則のことと考えて下さい。ここではなぜ研究開発評価の理念が必要で、それを理解することが大切なことかを解説します。

1.2.1　研究開発評価の理念とは何か

研究開発評価は新しい問題に常に直面し、定型的な業務としてこなすことのできない実務的能力が試される仕事です。したがって、マニュアルだけでは対応できない状況において、基本に立ち戻って考える原理原則が必要になります。研究開発評価の理念を理解しておくことで、新しい問題に対してより有効な対処方法を考えることができるでしょう。

研究開発評価の理念とは次のように言うこともできるでしょう。「評価の理念とは、評価の現場において個別具体的な問題に直面したときに、判断基準となる価値観、感性、美意識である。」

既存の理論に従ってロジックを積み上げ、それを現実に投射し、アクションを起こす、という合理性に基づく計画的思考は、さまざまな問題に直面する現場ではほとんど機能しません。より実務的な問題に対処するためには、行動の規範基準である理念がどのような仕事にも必要であり、関係者の中で共有されていることが大切です。

本テキストの基礎編の冒頭でも登場しましたが、研究開発評価の理念は、「支援的であること（Supportive）」、「非排除・双方向的であること（Inclusive and Interactive）」、「明示的であること（Evincive）」の３つが基本です。

出所：財団法人政策科学研究所作成資料

図 1-2　研究開発評価の理念

この理念は、我が国の研究開発評価の現場の経験から導かれたものです。海外では異なる評価の理念が、独自に形成されています。その国の研究開発マネジメントや科学技術政策の仕組みに適切な評価の在り方が存在して当然ですし、文化的背景、歴史認識、政治手続によっても評価の理念は変わってくるでしょう。

日本のように、過度に競争的でない、かつ長期的展望の下に研究開発を実施する国では、評価の理念においてもその特徴を活かすような考え方が望ましいと言えます。

1.2.2　研究開発評価の理念はどうあるべきか

1.2.2.1　支援的であること

研究開発評価において支援的であることとは、研究開発実施者を「励まし、育む」評価であることを意味します。

「励まし、育む」ためには最初から「切る」という姿勢では評価が成り立ちません。また、評価結果が査定と直接的に結びついているような仕組みも望ましくありません。

なぜならば、評価を組織的な業務として定着させ、意義のあるものにするためには信頼関係の構築が最も重要になるからです。被評価者が評価実施側に対して信頼感を持つためには、評価を実施することが被評価者にとって有用であることが条件です。また、現実問題として評価を実施する場合に詳細な情報は被評価者側が持っているため、信頼関係の構築なくして正確な情報に基づく評価は無理であると言えましょう。

では、被評価者が評価を有用であると認識するためにはどのような措置が考えられるでしょうか。表 1-1 は評価対象である被評価者別に、評価時期毎に想定される支援的措置をとりまとめたものです。

課題評価、つまり研究者が提案した課題の事前（採択）評価において、評価実施側が応募課題の一定数を切らざるを得ない状況を想定してみましょう。採択された場合であれば、提案課題のより円滑な実施のために必要な措置を支持したり提案することが可能でしょう。しかしながら、不採択の場合においても、応募した研究者がよりよい課題を今後提案してくれるよう、「励まし、育む」必要があります。そのためには、なぜ不採択になったのか、どうすれば今後同様の課題に応募する時に採択されるようになるかについての情報を提示することが重要です。ファンドマネージャーが提案課題により相応しい他機関のファンドを紹介するということも必要な措置です[1]。

中間評価の場合であれば、課題や事業の途中段階でのコメントをレビュー評価から得ることができます。研究開発の最終的目標の確認、軌道修正や研究開発方法に対するコメントなど、研究開発実施側が気づかない点をフォローすることができます。

実際のところ、我が国の研究開発事業では中間評価時に研究開発の方向性や具体的目標が明確になるというケースが少なくありません。レビューアからの意見を参考にして、共

[1]米国 NSF ではプログラムマネージャーが応募課題に対してより相応しいプログラムの紹介を実施しています。

通の目標を認識したり、研究グループ間の相互作用が生まれたりするという効能が中間評価には認められます。

事後評価であれば「次にどうつなげるか？」が問題になります。研究開発テーマの発展性、継続性などに対してレビューアからコメントをもらい、支援的措置の可能性を研究者側に提案することができるでしょう。企業等が実施者である実用化研究の場合、事後評価は研究開発段階を脱して次の段階に移行できるかどうかを問う重要な評価になります。この場合も、研究開発以外の事業展開を含め、的確な支援措置を講じることが必要となります。

表 1-1　評価における支援的措置

対象	事前（採択）評価	中間評価	事後評価
研究者	【採択の場合】 ・事業目的からみた研究内容への提案 ・研究計画への改善提案 ・他の研究グループとの協力等、研究体制への提案 【不採択の場合】 ・不採択理由の明記（どこが足りなかったのか？） ・改善提案（どうすれば採択に結びつくか？） ・研究開発テーマに相応しい他のファンドの紹介	・事業目的からみた研究内容への改善提案（研究方法等） ・事業目的からみた研究活動への提案 ・事業目的からみた成果利用に関する提案	・研究内容の発展の仕方に対する提案（次の段階のテーマは何か？） ・継続的な支援措置の可能性に対する提案（他のファンドの紹介等） ・事業目的からみた成果利用に関する提案 ・優れた研究に対する表彰制
企業・大学等（研究開発実施者）	【採択の場合】 ・研究開発計画への改善提案 ・コンソーシアムや産学連携等の研究開発体制に関する提案 ・責任／管理体制への提案 【不採択の場合】 ・不採択理由の明記 ・改善提案 ・研究開発提案に相応しい他のファンドの紹介	・事業目的からみた研究開発内容への改善提案（事業アプローチ等） ・事業目的からみた成果利用に関する提案 ・知財化・標準化に向けての提案	・研究開発の発展方向に関する提案 ・継続的な支援措置の検討（研究開発以外の事業への展開） ・成果の普及に関する提案 ・優れた研究開発に対する表彰制度
機関（独法等）	【中期目標・中期計画】【年度評価】 【中期目標期間終了年次の評価】 ・具体的な目標設定の方法に関する提案　事業展開の方向性に関する提案　次期中期目標期間における事業展開の方針に関する提案 目標の測り方に関する協議・提案　事業目標の見直し等の提案（上方・下方修正）　次期中期目標期間の目標設定に関する提案 数値目標に関する共通理解	【年度評価】 ・事業展開の方向性に関する提案 ・事業目標の見直し等の提案（上方・下方修正）	【中期目標期間終了年次の評価】 ・次期中期目標期間における事業展開の方針に関する提案 ・次期中期目標期間の目標設定に関する提案

出所：財団法人政策科学研究所作成資料

機関評価の場合は、事前評価の段階では中期目標・中期計画の策定に関して省庁のアドバイザリー・ボード（独立行政法人評価委員会等）から適切な提案が出せるかが問題となるでしょう。機関のミッションを具体的に展開し、事業目標とどのように位置づけるか、どの水準に目標を定めるかについての具体的な方針を協議する必要があります。これは年

度評価や中期目標期間終了後の評価の枠組みについて議論することにもつながりますので、機関評価の枠組みを支援的に構築する作業であると言っても良いでしょう。

年度評価（中間評価）では、詳細なデータに基づいた事業目標の達成見込みを踏まえながら、事業展開の方針に関する提案や支援的措置の必要性を講じることが重要になります。中期目標期間終了年次の評価（事後評価）では、次期目標期間における事業展開の方針についての前向きな提案が求められるでしょう。

研究者にしろ、組織にしろ、被評価者側にとって評価が有用な機会でなければ、評価に必要な詳細なデータや情報を被評価者側が提出するインセンティブはありません。ましてや、最初から査定的な態度（つまり、「切る」評価）で評価を実施すると、お互いの信頼関係を構築することは難しく、継続的な評価の実施は不可能になります。

1.2.2.2　非排除・双方向的であること

研究開発評価において非排除的であるということは、被評価者が意見を述べる機会を排除しないということです。

研究開発評価では、一般的に外部評価者による評価パネルを組織し、ここでの検討が評価結果に反映されます。通常、評価は資料・データに基づくドキュメントを基に行われますが、評価者が被評価者の意見を直接聞くというプロセスも重視されます。

この時に、被評価者である研究開発実施者は、ドキュメントには盛り込むことのできない研究開発への熱意、専門家としての展望、より広い研究の意味づけ等を評価者側に伝えることができます。

このようなやり取りが手続きとして原則的に確保されているということが、研究開発評価の枠組みの中で大切なことになります。

1.2.2.3　明示的であること

研究開発評価が明示的であることとは、評価のプロセスが明確な文書類に基づいて実施され、かつ評価結果が国民に容易にアクセスしやすいかたちで公開されることを意味します。

評価が実施されるためにはその前提としてさまざまな文書規定が存在します。例えば、提案公募型研究開発事業の場合では、事業の説明、応募要領、採択審査の基準が応募者に対して示されていなければなりません。応募者は応募要領に基づいて申請書と研究計画書を提出します。採択審査は事務手続きレベルでは申請書の内容を確認し、研究内容については研究計画書を基にメールレビュー等の手続きを踏まえて審査されます。

採択審査の結果については、Web 等を通じて広く公開し、どの研究開発課題が採択されたかを明らかにします。また、不採択の課題については、申請者に対して不採択の理由書が送付されます。

採択課題の提案者に対する契約においても、さまざまな取り決めが行われます。支出等

の規定、評価時における情報開示義務、成果利用や知財化、上市の状況等の報告義務なども盛り込まれます。

研究開発評価実施時においても、成果報告書、ヒアリング記録、アンケート、それらをとりまとめた評価報告書等の各種文書が作成されます。

このように、評価に限らず研究開発マネジメントで作成される各々の文書はお互いの記述内容を参照するといった一連のドキュメント・システムを構築しています。これは、ある文書が必ず他の文書の理由付けや根拠を述べているためです。最終的に評価結果を国民に公表するにしても、その結果に至った経緯が後から検証することができなければ意味がありません。

評価を含む研究開発マネジメントにおいて、明示的であることが要請されるのは、後から第三者が検証可能であることを担保するためです。

この明示性が確保されることで、支援的、非排除的であっても、馴れ合いではなく、緊張感のある関係を事業推進側と研究開発実施側とで保つことができるのです。

1.3　研究開発評価の局面

本節では、研究開発評価の What と When について解説します。What とは、研究開発評価の対象についての理解や、研究開発事業等のいかなる実績を評価すべきかについて学びます。

また、研究開発評価は適切な時期（When）に実施されなければなりません。評価時期においてどのような内容が重視されるべきかについて解説します。

1.3.1　評価対象の区分

大綱的指針では、研究開発評価の評価対象を次のように区分しています。

表 1-2　大綱的指針における評価対象の区分

大綱的指針の定める評価対象		定義
研究開発施策		国及び府省の政策目標や機関等の設置目的を達成するために策定した研究開発政策、戦略、制度、プログラム等
研究開発課題		研究者等が具体的に研究開発を行う個別のテーマ
	競争的研究資金による課題	「研究者の自由な発想に基づく基礎研究」と特定の政策目的を実現するための「研究目的を指定された研究」
	重点的資金による課題	国が定めた明確な目的や目標に沿って重点的に推進される課題
	基盤的資金による課題	研究開発機関の長の責任において、機関の設置目的等に照らして行われる研究開発課題
研究開発機関		国立研究所、研究開発型独立行政法人、研究開発資金配分型独立行政法人等
研究者		研究開発機関、大学等の研究者、研究チーム

出所：財団法人政策科学研究所作成資料

各府省の実施している研究開発事業等の制度的枠組みが異なるため、大綱的指針では上のような区分を適用し、それぞれの評価対象についての留意点をとりまとめています。大綱的指針の区分と各府省が実施している研究開発事業との対応を考えるときに混乱しやすいのは、大綱的指針の区分は原理的であり、実際の研究開発事業は事業形態に応じて区分されていることです。

大綱的指針の原理的区分とは、研究開発施策を政策ツールとしての枠、つまりプログラムとして捉え、研究開発課題をプログラムの下に位置づけられるプロジェクトとして捉えていることです。つまり、政策の階層構造を意識して区分されています。

一方、実際の各府省が実施している研究開発事業は助成事業か、委託事業か、補助事業かという契約形態による区分、または、テーマ公募型事業か、委託先公募型事業か、もしくは特定の独立行政法人への委託事業かという区分にしたがって展開されており、基本的に事業を政策単位もしくは予算単位として運用しています。

研究開発評価において重要なことは、事業形態がどのようなものであれ、評価の目的に応じて評価対象を原理的に区別することです。本テキストでは、大綱的指針の区分をさらに発展させ、より一般的な評価ができるように評価対象を次のように区分します。

表 1-3 本テキストにおける評価対象の区分

本テキストの評価対象の区分	定義	大綱的指針との対応	政策評価法との対応
総合政策・計画	研究開発を含む総合的な政策や計画。	研究開発政策・戦略	対象
プログラム	総合政策・計画の特定の目的を達成するための政策パッケージもしくは政策手段。プログラムの下で走るプロジェクトを階層構造として持つ。	制度・プログラム	対象
従属型プロジェクト	プログラムの下で走る個別事業、課題。	競争的研究資金による課題 重点的資金による課題	対象（10億円以上の費用を要するものについて事前評価の義務付けあり）
独立型プロジェクト	総合政策・計画の特定の目的を達成するための独立したプロジェクト。 大型プロジェクト。	重点的資金による課題 基盤的資金による課題	対象（10億円以上の費用を要するものについて事前評価の義務付けあり）
機関	研究開発機関、資金配分機関、大学等	研究開発機関等	対象外
研究者	研究者個人、研究チーム	研究者等	対象外

出所：財団法人政策科学研究所作成資料

「総合政策・計画」とは、府省が作成する政策体系とスケジュールのことです。ここでは、府省が実施する研究開発と研究開発以外の政策の機能的連携を考慮し、各事業の政策的位置づけを行うとともに、政策全体としてのスケジュールを定めます。総合政策・計画の評価では、政策全体が国民生活等にどのように役立ったかの観点から評価されます。

「プログラム」とは、総合政策・計画におけるある特定の政策目的を果たすための政策手段のパッケージです。研究開発を含む事業や、提案型研究開発公募事業などの研究開発制度、競争的研究資金による研究開発制度が相当します。プログラムの評価では、プログラムの下で走る課題やプロジェクトの個別の評価ではなく、プログラム全体としての実績を見ることになります。

「従属型プロジェクト」とはプログラムなどのある枠組みの下に位置づけられる個別の研究開発事業、課題のことを指します。従属型プロジェクトでは、採択評価、中間評価、事後評価が行われます。

「独立型プロジェクト」とは国の定める戦略領域や国家基幹技術などの研究開発で、比較的大型のプロジェクトを指します。例えば、宇宙開発や原子力、スーパーコンピュータ開発などのナショナル・プロジェクトが相当します。独立型プロジェクトでは、その規模ゆえに、研究開発実施者側の体制・マネジメントが重視されます。

「機関」とは国から交付金を得ている研究開発機関、資金配分機関、国立大学法人等のことを指します。多くは独立行政法人評価、国立大学法人評価の枠組みの中で評価されます。

「研究者」とは研究機関、大学等に属する研究者または研究チームのことであり、任期付きかどうかを問わず、公的研究資金で研究開発を実施しているすべての人材が相当します。大綱的指針では研究者の業績評価のあり方を示しています。

このうち、特に重要な区分が総合政策・計画、プログラム、従属型プロジェクト、独立型プロジェクトの区分です。これらは評価対象として階層構造を有し、上位の施策目的に対して下位の施策が手段として位置づけられることになります。このような原理的区分を行うことで、評価の重複を避け、評価を効率的に実施することが可能になります。

出所：財団法人政策科学研究所作成資料

図 1-3 研究開発施策の階層構造

機関における階層構造とは機関の組織構造を意味しますが、評価対象として機関の階層構造をどのように考えるかについては、第６章「機関の評価」で詳しく解説します。

コラム：政策体系

評価対象がどのような階層構造を持っているかを具体的に図示することは大変重要です。図 1-4 は経済産業省が実施する研究開発に関わる施策体系を図示したものです。経済産業省では、研究開発事業を「研究開発プロジェクト」と呼び、研究開発プロジェクトだけからなる施策を「研究開発プログラム」と呼んでいます。それ以外にも、施策レベルでは研究開発を一部含むもの、競争的研究資金制度を含むものが存在します。また、競争的資金制度（＝提案公募型事業）の中で採択されるプロジェクトを「課題」と呼びます。

出所：財団法人政策科学研究所作成資料

図 1-4 経済産業省における研究開発評価対象の階層構造

ある研究開発プロジェクトを評価しようとした場合、そのプロジェクトの意義や政策的位置づけ等は上位施策の目的に依拠することになります。施策の評価を実施しようとした場合、より上位の政策、あるいは省が掲げるミッションに依拠してその重要性なり意義が評価されることになります。

このように、評価対象が階層構造を持つということは、上位政策の目的に対して下位の

施策または事業が「手段」として位置づけられている、ということを意味しています。下位の事業が「手段」である以上、上位の政策に対して当該事業にはどのような貢献が期待されているかが明確になります。

プロジェクトを束ねる枠組みが単なる科学技術領域であったり、産業分野であったりする政策体系は、プロジェクトの重複を生み、資金の効率的・効果的な運用にはつながりません。政策体系が目的＝手段関係から構成されるとき、研究開発プロジェクトもより効果的に運用され、研究開発評価もより有効に活用されることになります。

1.3.2　研究開発の実績の区分

研究開発事業や制度で評価される項目のうち、もっとも重要なものは「実績」に関わる項目です。実績には通常、「何を産み出したか？」に係わる「成果」の部分と、新しい取り組みとして「どのように産み出したか？」に係わる「過程」の部分が着目されます。

出所：財団法人政策科学研究所作成資料

図 1-5　実績概念の区分

1.3.2.1　成果の区分

成果を区分する概念としては、これまで「意図的／副次的」、「直接的／間接的」、「形式的／本質的」といったものが適用されてきました(表 1-4)。しかしながら、物事の多くは二元論で割り切れるものではなく、それらの区分概念が同時に組み合わされて表現されるものと考えられます。

表 1-4　成果の区分

区分概念	定義
意図的 (intended)	目的・目標に沿った成果
副次的 (unintended)	目的・目標に含まれない成果
直接的 (direct)	事業等によって直接的に産み出された成果
間接的 (indirect)	事業等の成果の二次利用によって産み出された成果
形式的 (formative)	形式的(量的)に把握が可能な成果
本質的 (essential)	本質的な価値を有する成果、成果の質的側面

出所：財団法人政策科学研究所作成資料

こうした区分を適用すると、評価対象が上手く当てはまらない場合にどのように評価するべきかという問題に直面します。例えば、「副次的ではあるが、直接的かつ本質的な成果」、「意図的ではあるが、間接的な成果」という場合には、こうした類型化は成果の特徴を述べる分には有益かもしれませんが、評価のためのカテゴリーとしてはあまり意味を持ちません。

近年、「アウトカム評価」という言葉がよく聞かれるようになってきました。「そもそもアウトカムとは何か？」という議論よりも先に、なんとなく社会経済的価値に相当するものという了解が先に出来てしまい、当該事業のアウトカム探しをしなければならいという混乱が生じているように思われます。

本節では、成果の区分概念として、基礎研究から実用化研究の場合においても適用可能な類型化概念を説明し、なぜ「アウトカム評価」が必要なのかについての理解を深めます。

1.3.2.1.1　アウトカム

最初に、難解なアウトカムの概念について学びましょう。

大綱的指針には「事前評価や追跡評価における効果（アウトカム）、波及効果（インパクト）等に係わる評価手法について改良する」とあるだけで、アウトカムが何であるかの具体的な記述は見られません。

具体的な記述の例として、独立行政法人産業技術総合研究所編「産総研の研究開発評価のあり方（中間まとめ）」にアウトカムに関する記述があります。

”アウトカムとは、研究開発成果の本質的ないし内容的側面。研究開発のミッションが達成された結果。例えば、学術論文については、科学技術コミュニティで評価を得た内容。社会経済的な効果を目的とした研究の場合は、社会経済的な製品やサービスの価値的な内容。”

前半部分の「研究開発成果の本質的ないし内容的側面」という定義の仕方は、アウトカムという現象を捉えようとする試みに他なりません。現象論でアウトカムを理解しようとすると、その現象に当てはまらないケースに適用することが困難になります。

例えば、戦略的拠点形成プログラムのような事業において、アウトカムを示せと言われたときに、価値ある学術論文の本数やどのような経済効果があったかを評価しても、戦略的拠点形成プログラムの本質を評価したことにはならないでしょう。

ここで後半部分「研究開発のミッションが達成された結果」という定義が重要な意味を持つことになります。

アウトカムとはもともと政策評価において使われている言葉です。施策や事業がどのような効果を上げたかを検証するための概念です。つまり、研究開発に限らず、評価の対象となる施策や事業の内容によって、アウトカムの定義、範囲が異なることがまず基本的な理解として必要です。

アウトカムが政策評価に用いられる概念であるということは、その定義と範囲が施策や事業の目的・目標（ミッションをより具体的に展開したもの）に依存するという意味を持ちます。ここで重要なことは、ある施策や事業が目的・目標としてどのような範囲（時間的・空間的）で有効な手立てを考えているのか？ということです。つまり、施策や事業の責任範囲を明らかにすることが大切です。

（1）目的・目標設定

ここで、施策や事業の目的・目標とはどのようなものか整理しておきましょう。

目的（Objectives）とは願望、期待などが反映された、現状とは異なる新たなステージのイメージであると考えて下さい。例えば、「CO2 排出量を削減する」ことがある技術開発の主目的だとすれば、目的表現は「CO2 削減に資する○○技術を開発する」ということになります。

出所：財団法人政策科学研究所作成資料

図 1-6　目的と目標

目標（Targets）の設定は、目的表現を具体的に観測可能な達成水準として示すことです。目標を定めるという行為は、茫漠とした目的の範囲をある程度限定し、もっとも実現すべき事項がどこにあるかという「境界」を定めることです。例えば、「CO2 排出量を削減する」という目的に対して、「総量としてどれだけ削減するか？」、「いつまでに削減するか？」、「どの産業の排出量を削減するか？」、「削減コストはいくらまで許容できるか？」といった具体的な目標が必要になります。

つまり、目標が設定されれば、その目標を達成することで施策や事業の目的の全部ではないにしろ、主要な部分が達成できることを意味します。

（2）目的・目標とアウトカムの関係

目的・目標が明らかになったら、アウトカムをどのように考えるべきでしょうか。

施策や事業は、それらを措置する何らかの理由に基づいて実現すべき内容が構想されます。これを施策や事業の「目的」であるとすると、通常、「目的」（objectives）は施策企画者の様々な願望を含んでいます。しかし実施に移される施策や事業はそれらが実現すべき内容を厳選し、確実に実現を目指す内容に絞られなければなりません。この厳選された内容を「アウトカム」とよび、施策実施者が社会に対してその実現を約束する、いわば契約内容を表しています。

したがって、アウトカムの対象・範囲は施策として「意図した結果」（intended objectives）に相当することになります。

では、目的・目標とアウトカムの関係はどのようになっているでしょうか。下の図 1-7 は目的・目標とアウトカムの範囲を概念的に示したものです。「目的」はその内容を具体的

に示す「目標」にブレークダウンされますが、その際にアウトカムの内容（意図した結果）を代表するような目標を選びます。その目標の達成に向けて施策を展開することが、アウトカムをもたらすことになります。

出所：財団法人政策科学研究所作成資料

図 1-7　目的・目標とアウトカムの関係

ここで重要なことは、アウトカムが具体的に何であるか、ということよりも、「目標の達成に向けて施策を展開すること」に注意を払わなければならないことです。施策や事業が遠大な目的・目標を掲げつつ、実態は単なる試作品の開発や要素技術の開発だけに終わっている、というような研究開発の例は枚挙に暇がありません。施策目標として掲げた以上、それを実現するために施策としてどのような手だてを打っているのか、が重要であり、その結果がどのような形で現れているかは別に検証すれば良いことなのです。

つまり、アウトカムが先にあるのではなく、施策や事業の目的・目標とそれを実現するための具体的な措置があって、その結果（アウトカム）が何であるのかが問われる本質となります。

したがって、ある事業やプログラムの評価制度を変えただけでは、アウトカムは議論することはできません。事業やプログラムの企画・設計段階における目的・目標設定をきちんと展開するための、マネジメントツールが必要になってくるのです。

1.3.2.1.2　アウトプット

アウトカムが「意図した結果」であることを理解した上で、では、アウトプットとは何なのでしょうか。

一般的にはアウトプットとは入力に対して、あるメカニズムが産み出した出力結果を言います。施策や事業の場合、当該施策や事業の実施期間中に実施者が産み出したものがア

ウトプットということになります。しかしながら、「何を産み出したか？」という問いかけだけでは、アウトプットを列挙することはできません。アウトプットとは「目に見えて（tangible）」、「数えられるもの（countable）」であることが条件です。

下の図 1-8 は目的・目標、アウトカムとアウトプットの関係を示したものです。

平澤冷(2006)，内閣府総合科学技術会議発表資料

図 1-8 アウトカムからみたアウトプットの位置づけ

アウトカムが目的をブレークダウンした目標の実現によって規定される「意図した結果」であるのに対し、アウトプットは目標の実現に向けた「活動の水準（level of activities）」であると言えます。すべてのアウトプットは何らかの意思の下に実施された行為の結果として把握されます。その意思とは目標達成に向けた努力のことであり、その行為の結果とは、さまざまな活動の水準として把握されます。それらの水準を、目に見えるかたちに、数えられるものとして把握したものがアウトプットということになります。

つまり、アウトプットとは目標実現に向けた努力の証であり、できるだけ悉皆的に把握列挙する必要があります。

ここで注意しなければならないことは、アウトプットの数が多ければ多いほど良い、という論理にはならないことです。努力や活動の証ではあっても、価値判断はアウトカムの評価に委ねられます。アウトプットから何が産み出されたのかが問題であって、形式的に把握される数量に価値が見いだされることはありません。

例えば、論文の数が多ければ多いほど良いという議論にはならないはずです。数が多ければ良いものが確率的に産み出されるという論法は、マネジメントを改善していくという考え方と相反するものです。

アウトプットは単に活動の水準を表しているに過ぎないことを心に留めて下さい。

「優れたアウトプットの創出を目指す」という記述が公的研究機関の評価指針などによく見受けられますが、本テキストの枠組みの理解からすれば、それはアウトカムを意味していることになります。

1.3.2.1.3　インパクト

インパクトとは、施策の実施者が意図した結果以外の成果をもたらしたものを言います。下の図は先のアウトプット、アウトカムの関係図と施策の関係者を対置したものです。「施策の企画者」とは政策立案者を指し、「施策の実施者」とは研究開発であれば、研究開発機関や研究者ということになります。

今、何かしらの施策を展開したとしましょう。例えばある疾病の治療薬の開発を例に取ると、この研究開発施策が最終的に想定している「施策の受益者」は当該疾病に疾患している患者ということになります。

出所：平澤冷(2006)，内閣府総合科学技術会議発表資料

図 1-9　施策の受益者

ここで、創薬を例にとると、新薬開発は前臨床段階と臨床試験の各フェーズをクリアしなければ最終的に薬が患者に提供されることはありません。

ところが、大学や研究機関は有用物質の発見や化合物の生成について研究開発を実施し、治験の段階はもっぱら臨床を担当する大学病院等に委ねられることになります。有用物質の発見や前臨床段階までの研究開発の成果は、薬理効果の可能性を示しているだけで、人

体に投与して効果があり、かつ安全であるという保証を得るためには臨床段階を経なければなりません。

図 1-10　新薬開発のフロー

したがって、創薬に関する研究開発の場合、研究開発フェーズに応じたカスタマーが誰なのか？という区別をすることと、研究開発施策そのものの責任範囲がどこまでなのか、という「線引き」が重要になります。

前臨床段階までが施策の責任範囲であれば、特定の受益者（製薬会社や治験における対象患者）が施策の意図した最終受益者になるでしょう。臨床段階までを含み、医薬品市場における競争力の向上までを図るのであれば、特定の受益者からさらにすすみ、新薬の対象患者全体が最終的な受益者になるでしょう。

平澤冷(2006)，内閣府総合科学技術会議発表資料

図 1-11　施策の責任の範囲とインパクト

アウトカムとインパクトを分かつのは、「施策としてその効果が発現する手だてを用意しているかどうか？」という判断基準が妥当です。

創薬の場合、有用物質の発見が成果（アウトプット）であれば、薬としての本質的な価値（対象患者への薬理効果）がある程度確認された（前臨床試験が済んだ）時点でアウトカムが達成されていると判断できます。この場合のインパクトは、「有用物質から別の症状に効く新たな薬理効果が見いだされた」、「その発見を契機に有用物質の同族体がさまざまな研究開発機関で研究され、発見された」、あるいは、「それらの同族体の方が先に上市された」という場合も含み、施策の責任の範囲外で実施された行為によって得られた効果と判断することができます。

このことは、インパクトの大きさをもって研究開発の質を評価することはできても、インパクトを施策の実績そのものとして考えることはできない、ということを意味しています。なぜならば施策の誇大広告になってしまうからです。

1.3.2.2　成果以外の実績

研究開発評価では、成果以外の実績もきちんと評価しなければなりません。科学技術の振興のためには、優れた成果が創出されるように、人材育成や研究開発ネットワークの構築、研究開発環境の整備や地域クラスターの形成など、新しく効果的な仕組みを産み出したかどうかという、研究開発の「過程」についても評価することが重要です。

米国の財務省（OMB）が行っている政策プログラム評価のためのツール「PART（Program Assessment Rating Tool）」では、プログラムがもたらした成果以外の実績、過程（Process）についても評価されます。過程では、主にアクター（実施体制）、マネジメント、システム（拠点形成などの仕組み）が評価されます。

基礎研究や戦略的な研究領域においては、評価実施時期に成果がかたちとなって現れるとは限りません。むしろ、政策の効果としては人材育成や研究開発拠点の形成、ネットワーク形成等において一定の効果が得られていることが重視されなければならないでしょう。

これらの過程における実績を定量的に示すことは、高度な分析ツールを必要としますが、事業やプログラムの意図した効果がきちんと現れているかどうかを検証する際には非常に有効です。

ここでは、分析ツールについての詳しい解説を省きますが、過程を評価する際の観点を概説します。

1.3.2.2.1　アクター

研究開発プログラムや事業を展開する場合、さまざまなアクターが介在します。研究開発プログラムや事業の推進者（多くは官庁、資金配分機関）、研究開発の実施者（企業、研究開発機関や大学等の研究者）、成果の一次利用者、二次利用者などが想定できるでしょう。

それぞれのアクターにとって、研究開発プログラムや事業がどのような影響を及ぼしたかについて、明らかにすることが実績・実態把握の第一歩になります。

表 1-5　アクターに関する評価の視点

アクターの種類	明らかにすべき内容
研究開発実施者	・人材育成面 -学位取得者数 -受賞・表彰の有無 -若手研究者・ポスドク等の参加度 -男女共同参画の度合い ・ネットワークの形成 -共著論文・共同発表の数 -学術団体の分科会形成等 -特許の共同申請状況 ・人材交流の実績 -国際共同研究の実績.

研究開発プログラムや事業の推進者	・プログラムマネージャー（PM）の資質・配置状況 -専任の PM の人数 -PM のバックグラウンド ・マネジメント能力 -PM に対する研究実施者からの評価（アンケート等） -PM に対するプログラム・ディレクター（PD）の評価
成果等の利用者	・利用者の属性 -属性毎の利用者数 -用途等

出所：財団法人政策科学研究所作成資料

1.3.2.2.2　マネジメント

研究開発評価において、マネジメントについて評価するといった場合、二つの意味が考えられます。

ひとつは、研究開発実施者側における R&D マネジメントの質についての意味。もうひとつは、研究開発施策や事業の推進者側のマネジメントの質についての意味です。

研究開発実施者側の R&D マネジメントについては、巨額の研究開発プロジェクトや研究開発機関等の評価以外で、改善課題として問題にされることはあまりないでしょう。あまりにずさんな管理をしていた場合には、実施者側の管理能力が問われることは想像に難くないですが、このような例は多くの場合、契約事項や倫理規定への違反というかたちで対処ができます。

むしろ、公的研究開発資金を効率的に運用したかどうかの説明責任は、ファンディングを実施している側にあると言えます。すなわち、事業推進主体である省庁の原局・原課、資金配分機関の推進部署におけるマネジメントの在り方が問われることになります。

また、マネジメント評価が対象とするのは、組織としての運用能力ですので、プロジェクトレベルよりも、プログラムレベル、機関レベルでの評価においてより重要になります。

ここでは研究開発の評価項目としてマネジメントを議論する場合には、研究開発実施者（研究者、研究チーム、企業等）のマネジメントではなく、ファンディング実施側のマネジメントと公的研究開発機関（独法、大学等）における組織的なマネジメントを中心的に扱います。

どのようなマネジメントが優れていたかの判断基準を明らかにすることは難しいでしょう。また、マネジメントが改善されたから優れた成果に結びついたということも証明することはなかなかできません。では、どのようにマネジメントは評価されるべきなのでしょうか。民間企業のマネジメント評価のツールを参考にこの問題を考えてみましょう。

（1）バランス・スコア・カード（BSC）によるマネジメントの評価

民間企業では、マネジメントの評価にはバランス・スコア・カード（BSC）がよく利用されています[2]。BSC では、「財務の視点（過去）」「顧客の視点（外部）」「内部業務プロセスの視点（内部）」「イノベーションと学習の視点（将来）」の“4 つの視点”からマネジメントの改善度を評価します。具体的にはそれぞれの視点を代表する指標群を設定し、改善目標をたて、それをモニタリングすることで業務改善の進捗度を把握します。

出所：財団法人政策科学研究所作成資料

図 1-12 BSC によるマネジメントの評価軸

財務の視点では、「財務的業績の向上のために、株主に対してどのように行動すべきか」ということを明らかにします。公的機関の場合であれば、「財務当局に対して、ひいては納税者に対して、どのような財務的改善の努力をするべきか」というように読み替えることができます。

顧客の視点では、「経営戦略を達成するために、顧客に対してどのように行動すべきか」ということを明らかにします。これも、「研究開発事業が想定している顧客（カスタマー）に対して、どのようにサービスの改善を図るべきか」というように読み替えることができるでしょう。

内部業務プロセスの視点は、「顧客と株主をより満足させるために、どのような業務改善を図るべきか」という点を検討します。つまり、業務の効率化指標を考えるということになります。

イノベーションと学習の視点とは、「経営戦略を達成するために、どのようにして変化と改善のできる能力や環境を維持するか」ということを検討します。これも、「優れた研究開発課題を発掘・支援するために、どのようにして変化と改善のできる人材能力と環境を維持・発展すべきか」というように読み替えることができるでしょう。

---

[2] Robert S. Kaplan と David P. Norton らが提唱したマネジメント評価ツール。

これらの視点を具体的な指標に落としたのが表 1-6 です。

表 1-6 BSC によるマネジメント指標の例

視点	評価指標（例）	単位	事業	機関	目標	結果
財務の視点	事業予算に対する人員数	円／人	○	△		
	事業予算に対する事業期間	円／日	○	△		
	事業予算に占めるグラントもしくは委託契約費の比率	%	○	△		
	事業費に占める人件費の比率	%	○	△		
	事業費に占める当期取得固定資産額の比率	%	○	△		
顧客の視点	普及・啓発費（全額）	円		○		
	新規応募数の伸び（対前年比）	%	○			
	クレーム発生件数	件	○	○		
	顧客からみた満足度（アンケート等による指標化）	－	○	○		
	Web閲覧数の伸び（対前年比）	%		○		
内部業務プロセス	内部業務日数の短縮	日	○	○		
	意思決定プロセスの明確化	－		○		
イノベーションと学習	担当人材一人あたりの研修日数	日		○		
	担当人材一人あたりのOFF-JT機会	回		○		
	担当人材一人あたりの研修費用	円／人		○		
	担当人材の満足度	－		○		
	担当人材の業務に関連した保有資格	件		○		

対前年比（%）

出所：財団法人政策科学研究所作成資料

BSC を利用したマネジメントの評価では、PDCA サイクルを有効に機能させるマネジメント上の検討項目について、目標値を定め、それをクリアしたかどうかをチェックします。目標値が挑戦的であるかどうか、どの視点、検討項目を重視するかは経営戦略に従います。公的研究開発資金のマネジメントでは、経営戦略は事業の目的、あるいは機関のミッションに従うことになります。

マネジメントの評価は組織の自己改革に対する評価であるため、努力目標を自ら設定し、達成したかどうかを問うという、実務的なプロセスの繰り返しになります。マネジメントが功を奏したかどうかは、全体的な財務状況の改善や、組織に対する顧客の評価によって中長期的な指標として確認するしかありません。

1.3.2.2.3 システム

研究開発事業やプログラムによっては、何かしらの直接的な成果を意図してではなく、優れた研究開発成果を生み出すための仕組みや、専門人材を育成するための制度を構築するという目的を持つものが少なくありません。

このような事業やプログラムにおいて、実績として評価すべき項目が「システム」です。

例えば文部科学省が展開している戦略的研究拠点形成プログラムでは、「組織の長の優れた構想とリーダーシップにより、研究開発機関の組織運営改革や国際的に魅力ある研究拠点の創出を図る」とあります。

このような新しい仕組みを目的としたプログラムでは、その仕組み（システム）自体を

評価しなければなりません。表 1-7 は戦略的研究拠点形成プログラムの中間評価における評価項目・評価基準を抜粋したものです。

表 1-7 システムの評価

Ⅰ．進捗状況（目標達成度）（ミッションステートメントに対して）		
		○ 計画に沿って順調に進捗しているか
		○ 客観的な情勢の変化等（国立大学の法人化等）に対して適切に対処できているか
		○ 当初の計画どおりに進捗していない場合、当初目的を達成する見込みはどれほどか
Ⅱ．組織運営の妥当性		
		○ 調整費と調整費以外の外部資金、内部資金が、それぞれ組織運営構想に則って適切に充当されているか
		○ 総括責任者は組織運営構想実現のための必要な権限と責任を有しているか
		○ 単なる研究の実施だけになっていないなか
		○ 機関として組織運営構想の実現に向けて必要な取組・支援（人事制度・給与制度の改革等）を行っているか
Ⅲ．組織改革の成果		
	（研究拠点としての波及効果）	
		○ 組織改革により形成される研究拠点は、国際的な水準から見てトップレベルのものとなっているか
		○ 組織運営構想の実施により、研究拠点としてのポテンシャルの増加が期待されるか
	（組織運営構想の波及効果）	
		○ 研究開発システム改革のモデルとなることが期待されるか
		○ 他の研究機関にも波及し得る積極的な取組がなされているか
	（情報発信）	
		○ 広報は一般向けを含め十分に行われているか
		○ 組織改革の取組について、関係機関への情報提供・情報交換により周知を図っているか
Ⅳ．実施機関終了後における取組の継続性・発展性の見通し		
		○ 実施機関の長による支援が十分行われてきたか、また、実施機関終了後も、それらの支援が継続され、機関本来の取組としての発展が期待できるか
		○ 育成機関終了後においても、研究拠点となる体制を整備するための計画を有しているか
		○ 研究拠点として機関終了後の発展性が期待できるか
		○ 育成機関終了後の運営に必要な財政的・人員的措置が計画されているか

出所：文部科学省

システムの評価で重要な観点は、「事業等によって作り上げられた新しい仕組みが持続的なものであるかどうか」という点です。先の表の評価項目Ⅳにあるように、事業が終了した後も継続して研究拠点が存続し得るかどうかが問われています。初期の拠点形成のための投資が途絶えてしまったために研究拠点が存続できないとなれば、それは投資だったのではなく、単なる予算配分だったということになるからです。

1.3.3 研究開発評価の実施時期と重視すべき内容

本節では、いつ研究開発評価を実施すべきか（When?）という評価時期の問題と、適切な評価時期にどのような内容を評価すべきかという問題について解説します。

1.3.3.1　評価時期

1.3.3.1.1　研究開発の計画＝実施＝普及の各段階と評価時期

研究開発評価の時期は評価対象である研究開発の時間的経過に則して決まります。評価時期は一般的に、「事前評価（ex-ante）」、「中間評価（mid-term）」、「事後評価（ex- post）」の３つに区分されます。「事後評価」はさらに、「直後評価」と「追跡評価」に区分できます。

我が国の研究開発評価において「事後評価」と定義しているものの多くは、事業終了後１年以内に実施されていることから、「直後評価」に相当することになります。また、中間評価については、３年以上の研究開発期間を有する事業の中間年度に実施されることが一般的です。この中間評価以外に、毎年主要データを収集し、進捗状況の把握を行うことを「途上評価（monitoring）」といいます。

出所：財団法人政策科学研究所作成資料

図 1-13　研究開発段階と評価実施時期

1.3.3.1.2　政策評価法における事前評価と事後評価

先の図 1-13 は、研究開発評価一般の評価時期の定義です。これとは別に、我が国において展開されている政策評価の制度の中での研究開発評価の時期に関する位置づけは、それぞれ異なっています。この関係性を整理しておかないと、研究開発評価の業務と政策評価の業務との対応が難しくなります。

政策評価法では、各府省に対して、３年から５年程度の中期的な「基本計画」と毎年度の「実施計画」を策定することを規定しています。

特に、平成 17 年 12 月に改定された「政策評価に関する基本方針」では、各府省は評価と予算との対応関係を明確にするように求められたため、評価対象となる政策体系の見直し、評価単位の大括り化が図られています。

したがって、基本計画策定の時点において研究開発を含むさまざまな施策の基本的な方向性についてもこの段階で事前評価が行われています。

さらに、政策評価の年間スケジュールを見てみましょう。多くの府省においては、各年度当初の業務のスタートに合わせて、年度末頃に毎年度の政策評価の実施計画が策定され、この実施計画に基づき評価書の作成作業が行われます。一般政策についての評価書は、予算要求や政策の企画立案に反映するため、８月末の概算要求期限までに作成・公表されます。政策評価の結果は、予算査定等に活用され、年末には翌年度の政府予算案が決定されます。

出所：総務省「政策評価Ｑ＆Ａ」より抜粋

図 1-14　政策評価の年間スケジュール

実施計画策定の段階で、次年度の実施事業に対する事前評価が行われ、その評価結果が「評価書」として取りまとめられます。各府省が企画・立案した事業計画案に対して、外部有識者からなる委員会等でコメントを集約する機会も設けられてはいますが、基本的には事業の事前評価は府省の自己評価で行われます。　文部科学省であれば、事前評価は「事業評価書（新規・拡充事業等）」に取りまとめられます（参考：各省の政策評価関連文書）。

表 1-8 各省の政策評価関連文書

省名	分類		名称	URL
文部科学省	基本計画		文部科学省政策評価基本計画	http://www.mext.go.jp/a_menu/hyouka/seido/001.htm
	実施計画		文部科学省政策評価実施計画	http://www.mext.go.jp/a_menu/hyouka/seido/002.htm
	政策評価		実績評価書	http://www.mext.go.jp/a_menu/hyouka/main_a11.htm
	事業評価	事前評価	事業評価書（新規・拡充事業等）	
		事後評価	※実績評価書に含まれる、もしくは個別事業の事後報告書	
経済産業省	基本計画		経済産業省政策評価基本計画	http://www.meti.go.jp/policy/policy_management/top-files/plan-top.htm
	実施計画		事後評価実施計画	
	政策評価		経済産業省年次報告	http://www.meti.go.jp/policy/policy_management/top-files/Accountability-Report.htm
	事業評価	事前評価	概算要求等に係る事前評価書	http://www.meti.go.jp/policy/policy_management/top-files/jizen-top.htm
		事後評価	事後評価書	http://www.meti.go.jp/policy/policy_management/top-files/jigo-top.htm
厚生労働省	基本計画		厚生労働省における政策評価に関する基本計画	http://www.mhlw.go.jp/wp/seisaku/hyouka/keikaku-kekka.html
	実施計画		厚生労働省における事後評価の実施に関する実施計画	
	政策評価		実績評価書	
	事業評価	事前評価	新規事業に関する事業評価書	
		事後評価	※実績評価書に含まれる、もしくは個別事業の事後報告書	
国土交通省	基本計画		国土交通省政策評価基本計画	http://www.mlit.go.jp/hyouka/
	実施計画		国土交通省事後評価実施計画	
	政策評価		政策チェックアップ結果評価書	
	事業評価	事前評価	予算概算要求に係る個別研究開発課題評価書	
		事後評価	個別研究開発課題の評価書	
環境省	基本計画		環境省政策評価基本計画	http://www.env.go.jp/guide/seisaku/h18/kihon.html
	実施計画		環境省政策評価実施計画	http://www.env.go.jp/guide/seisaku/h18/keikaku.html
	政策評価		環境省政策評価書	http://www.env.go.jp/guide/seisaku/h17_jigo/index.html
	事業評価	事前評価	地球環境研究総合推進費における戦略的研究開発領域の新規プロジェクトに係る事前評価書	http://www.env.go.jp/guide/seisaku/kenkyu/index.html
		事後評価	※政策評価書に含まれる、もしくは個別事業の事後報告書	http://www.env.go.jp/guide/seisaku/h17_jigo/index.html

出所：財団法人政策科学研究所作成資料

政策評価法で定めるところの事後評価についても、基本計画が事後評価の対象となるため、基本計画の終了時点が評価時期になります。一方、実施計画で当該年度に行う事後評価は、基本計画に位置づけられた事業の事後評価となります。それぞれの事後評価の扱いについては階層が異なることに留意しなければなりません。

現状では、政策評価法の定める事後評価については、実施事業の事後評価を以て報告とする場合と、基本計画の達成状況や進捗状況に関する評価を以て報告とする場合というように、府省によって政策評価における事後評価の位置づけが異なっています。今後は、政策体系の見直しとともに、総合政策・計画や研究開発戦略の事後評価もしくは途上評価の在り方が問われることになるでしょう。

1.3.3.1.3　評価時期毎の重要性

研究開発評価の現場において、多く寄せられる疑問に「事後評価をなぜしなければならないのか？」という声があります。その理由は、中間評価であれば評価者のコメントを下に研究開発の軌道修正なり、改善提案が可能であるが、事後評価の段階では研究開発が終了しているので、前向きなコメントを得ることができない、あるいは評価結果を活かすことができないという意見です。また、さまざまな研究開発事業の終了年度が重なれば、事後評価を実施するだけで大変な負荷になり、その有用性もよく分からないというジレンマを評価現場では抱えています。

後者の負担感については評価の実施方法とも関連するので、ここでは取り上げませんが、事前評価、中間評価、事後評価、追跡評価が何のために実施され、どのような効用があるかについて、ここに取りまとめたいと思います。

(1)事前評価の意義

政策評価法で義務づけられている以上に、事前評価の意義は大きいものがあります。なぜならば、後に続く中間評価、事後評価において、事前評価の内容が必ず参照されるからです。

研究開発事業が複数年度に渡って実施される場合には、事前評価の時期の担当者と中間・事後評価の時期の担当者が同一人物であるとは限りません。事業の実施理由と目的・目標が事前評価の段階で明らかにされていなければ、後任の担当者が困ることになるだけでなく、中間評価や事後評価が事業に関する「そもそも論」で終始してしまう場合も想定されます。

事前評価において、事業の意義、目的・目標、期待される効果、事業推進者の所掌範囲が明確にされているからこそ、後に続く中間評価や事後評価において達成度評価が可能になるのです。研究開発評価において当面のわが国の課題は、事前評価をきっちりと実施するということです。

(2)中間評価の意義

おそらく研究開発評価を実施する上で、関係者の間で中間評価ほど有意義に捉えられているものはないでしょう。なぜならば、提案公募型事業や委託先公募事業では、事業の事前評価の段階で目的・目標に対する明確なコミットメントが出来ないために、中間評価の段階ではじめて関係者間における事業の方向性や成果の目標水準等の確認ができるという場合が少なくないからです。

加えて、被評価者である研究開発実施者は、関連領域の専門家から有益なコメントを得

ることができる機会として積極的に捉えることができます。

研究開発の仕切り直しや事業推進側と研究開発実施者側が共通認識を得る場としても、中間評価は有用視されています。

(3)事後評価の意義

事後評価については、事業終了時点における目的・目標の達成状況についての確認もさることながら、次の施策に活かすための知見を得ることが重要です。

また、事後評価は評価対象の品質保証（QA：Quality Assurance）を兼ねています。これは、研究開発に従事した研究開発実施者の実績にもなりますので、別の形での評価（組織内での実績評価、顧客の企業に対する評価等）に利用されることになります。事後評価で S を獲得したという品質保証が、研究開発実施者の次の研究開発への資金供与や融資等の機会に結びつくのです。

むしろ、事後評価の評価結果ほど本来戦略的に活用できるものはないと言っても過言ではないでしょう。評価が適切に実施されていることが大前提ですが、展開次第では、新しい研究開発案件の品質保証システムを構築することも可能だからです。

加えて、資金配分機関では、自らのマネジメントも評価項目として評価されます。事後評価において、資金配分機関のマネジメントが適切であったかどうかの評価は、機関評価のエビデンスとして活かされることになるでしょう。

研究開発実施者の立場から考えれば、事後評価が存在すること自体がある種のプレッシャーとなって、節度ある公的資金の使用とスケジュール管理を促す効果もあるといえます。

(4)追跡評価の意義

追跡評価については、フォローアップの観点が重要になります。政策の効果が果たして本当にあったのかどうかを確認するための作業といえるでしょう。したがって追跡評価については評価目的とする政策効果が十分に出現する時期を見越して、実施時期を設定しなければなりません。

追跡評価で確認する効果、特に社会経済的効果が、ある特定の事業の結果としてもたらされているとは考えられません。同様の社会経済的効果を目的とした事業を政策レベルで統合し、その政策としての効果が得られているかどうかを確認するというスタンスが評価の作業面からも有効となります。

1.3.3.2　評価時期毎の重視すべき評価項目

政策評価法や大綱的指針で定めるところの「必要性」、「有効性」、「効率性」の観点も、評価時期によっては、どのような評価基準で評価すべきかは本来異なります。このような点を踏まえ、評価時期毎に重視すべき評価項目を考えてみましょう。

1.3.3.2.1　事前評価で重視すべき評価項目

事前評価の段階では、研究開発はまだスタートしていません。ここでは研究開発の目的・目標設定や実施体制（責任体制）、研究開発計画、研究開発方法、コストなどについて、明示化された文章やプレゼンテーションをもとに評価します。

提案公募型の事業における個別課題に関わる採択評価を除いて、一般に国が実施する研究開発事業の事前評価は政策評価当局（総務省）への提出資料である評価書に、他の事業と併せてとりまとめられます。これは、政策評価法の第十条に規定されている業務であり、各省庁とも政策評価担当部署（多くは官房）が推進部署からの提出書類を基に作成します。

事業評価書では、「必要性、効率性又は有効性の観点その他当該政策の特性に応じて必要な観点」を評価項目として記載する必要があります。この中で、事前評価の段階ではもっとも重要な評価項目は、「必要性」と「有効性」に関する記述です。

必要性とは、「なぜ事業を実施するのか（Why？）」が問われています。また、有効性とは、「その事業によって何を成すのか（What？）」が問われています。この二つについて詳しく見てみましょう。さらに効率性についても、事前評価の段階で記述すべき点を指摘しておきたいと思います。

（1）必要性

事業の必要性を述べるためにはどのように考えるべきでしょうか。直接的な答えは、「事業の正当性（rationale）」を論理的に展開することです。ある事業が必要であるという根拠として、形式的な側面と本質的な側面の両方から述べることが必要でしょう。

形式的な側面とは、事業が実施主体の総合的な計画や政策の中でどのように位置づけられているかということです。これは、研究開発に限らず、政府の実施する事業は予算枠組み（一般会計か特別会計か）、政策体系の中で形式的に位置づけられ、国会等の承認を得てオーソライズされている必要があります。

本質的側面とは、形式的側面とは別に当該事業の意義を明らかにすることです。事業を行う前提条件としての問題認識・背景等が明らかにされ、対策の必要性を述べることになります。また、国として実施する必要性を別途記述しなければなりません。

図 1-15 は必要性に関するロジックの展開例です。「事業の正当性」は、「政策の中での位置づけ」、「事業の意義」に大別され、それぞれがさらに論理展開されています。「政策の中での位置づけ」は、「政策体系の中での位置づけ」と「歴史的経緯の中での位置づけ」に分けられています。

出所：財団法人政策科学研究所作成資料

図 1-15　必要性のロジック展開（事前評価）

「政策体系の中での位置づけ」は当該事業が、1)総合政策・計画に包摂されているか？、2)（省内の政策の中で）重複はないか？、3)（他省との政策と比べて）差別化されているか？という点を述べることで網羅できます。

「歴史的背景の中での位置づけ」は継続・拡充事業において特に明記しなければならない点です。1)前身の事業は何か？、2)これまでの事業が一定の効果を上げているか？、3)（これまでの事業と比較してどのように）高度化されているのか？という点を明らかにします。

一方、「事業の意義」については、「問題認識・背景等」と「公的関与の必要性」に分けられます。「問題認識・背景」では、1)（事業が前提としている）環境や状況の説明、2)施策主体としての問題認識、3)政策手段としての事業の役割が明記される必要があります。

「公的関与の必要性」では、1)民間ではなく、国が実施する理由、2)地方ではなく、国が実施する理由、3)官民の役割分担等が明記されます。

「事業の正当性」を述べることは、事業の目的を裏付けるための記述に相当します。事業がフォーカスすべき目的がどこにあり、どのような役割が期待され、政策体系の中でど

のように位置づけられているかを明記することで、実施事業の根拠が形成されます。

事前評価の段階で、必要性が明確に記述されていれば、中間評価や事後評価の段階で必要性を再度評価する必要がなくなります。もし、中間評価の時点で事業の前提条件に大きな変化があれば、改めて事業の意義を考える必要はありますが、研究開発に限らず、国の実施する事業は中長期的視野の下で展開されていることが一般的ですので、そうそう見直されるようであれば、政策評価のプロセスにおいて中止案件として議論の俎上に上がることになるでしょう。

（2）有効性

事業の有効性を示すには、事業の目的と目標を設定し、そのために何を実施するかについて明らかにすることです。

事前評価の段階で本来明らかにしておかなければならない点は、実施事業によって何が産み出され、どのような効果を国民生活にもたらすか、ということです。

この段階で、中間評価、事後評価、しいては追跡評価の際に検討されるアウトカムをどのように捉えているかが中心的な問題になります。

図 1-16 は有効性のロジック展開を目的と目標、アウトカムとアウトプットの観点から見ています。

「事業によって何を成すのか？」という問いかけに直接的に答えるものは、「事業の目的」です。「事業の目的」では、事業の成果によってもたらされるであろう、国民生活や社会の新しいイメージ（ビジョン）が記述されます。これは、実現を確約するものではなく、政策担当者（部署）の願望・期待を込めた目的表現です。また、ここでの表現が、当然のことながら上位の施策や総合政策・計画等とリンクしていなければなりません。

願望や期待だけでは当然のことながら困ります。事業実施主体として、何を実現させ、実現させるために何をするのかを明確にしなければなりません。ここで「事業の目標」を設定することになります。「事業の目標」とは、「事業の目的」で描いたビジョンを具体的に検討し、「これくらいの成果が得られれば事業目的に対して十分寄与した」と考えられる目標水準のことです。「事業の目標」は同時に、目標達成のための政策措置が及ぶ範囲を定める行為に他なりません。すなわち、事業の責任範囲であり、その結果はアウトカムに相当します。

「事業の目標」が設定されたら、その目標達成のために必要な活動を列挙します。事前評価の段階ではまだアウトプットが出ていませんが、事業目標の達成に資するアウトプットを想定し、そのための出力装置として事業の活動を設計することになります。

出所：財団法人政策科学研究所作成資料

図 1-16　有効性のロジック展開（事前評価）

（3）効率性

事前評価の段階ではアウトプット、アウトカム等が得られていませんので、この評価の観点を厳密に記述することは不可能です。しかしながら、事業の個々の活動において、どのような体制・仕組み、予算等で実施するかについての記述は可能です。ただし、これらは事業を実施する段階において常に見直され、結果的に当初の体制や仕組み、予算配分で行われないことが想定されますが、中間評価や事後評価の段階で、そのようなマネジメントの変更が有効であったかどうかを検証することができます。

少なくとも、事業全体の管理機能、執行体制や責任体制については、事前評価の段階で明らかにしておくべき事項です。

さらに、事前評価の段階では、事業全体のコストパフォーマンスについて、ある程度の比較考量を行う必要があります。他の研究開発手段と比較して、当該事業が目的達成のために優れた手段であることを裏付ける根拠を示すことで、当該事業の目的に対する効率性について述べることができます。

効率性記述の良い事例:

文部科学省平成 19 年度事業評価書（新規・拡充事業）より

「X線自由電子レーザー装置の利用開発」　研究振興局基礎基盤研究課

【効率性の記述部分】

本装置は、大型放射光施設 Spring-8 に隣接して整備することで、世界で唯一、放射光施設と X 線自由電子レーザーが共存する施設となり、２つの光を同時に用いる実験がここでのみ可能となるほか、運用面でも効率化が実現される。

また、日米欧で熾烈な国際競争力が繰り広げられる中、我が国は高加速勾配加速管や真空封止型アンジェレーターなど独自に開発した技術により、欧米に比べ全長４分の１以下、総コスト２分の１程度と、コンパクト化と低コスト化を実現可能としている。

さらに、本事業のために開発した装置架台は、40 年前に開発された送電線用絶縁体を利用することにより、従来用いられてきた架台の 20 倍という高い安定性を示しており、古い技術に新しい応用をもたらすなど、効率的な技術開発が進められている。

1.3.3.2.2　中間評価で重視すべき評価項目

中間評価の段階では、事業の成果の一部が出されていることが想定されます。特に、提案公募型事業の場合、研究開発制度の下で採択された研究開発課題の一部は終了していることでしょう。

成果がある程度出現していることを前提に、中間評価において重視されるべき評価項目とは何でしょうか。再び、「必要性」、「有効性」、「効率性」の観点から、評価項目を検討してみましょう。

（1）必要性

事業の正当性は事前評価の段階で確保されていると考えるのが妥当です。仮に、事前評価の時点と中間評価の時点とで、事業が依って立つ問題認識に大幅な修正を加える必要がある場合、改めて事業の意義を考えることになります。

事前評価の段階で、事業が置かれている環境や問題の重要性等に関するデータを示しておき、データのフォローアップを行っていれば、中間評価時において事業の問題認識が有効であるかどうかの確認になります。

(2)有効性

中間評価の段階では、アウトプットがある程度出現しているため、「事業において何が行われたか？」を確認することができます。ふたたび、有効性のロジック展開図 1-17 を見てみましょう。中間評価で行わなければならないことは、事業の目標の達成状況を確認することです。既出のアウトプットが設定している事業目標に貢献しているかどうかがチェックされ、さらに事業全体で意図している結果（アウトカム）がもたらされそうかどうかを評価することになります。

事業の目標達成状況が期待通り、もしくは期待を上回るものであれば、中間評価の評価パネルの評点は高いものになります。一方、このままでは事業目標が達成できない、あるいは、アウトカム実現の観点からより良い研究開発方法、実施体制へのサジェスチョンが得られたら、個々の事業の活動を見直すことになります。この場合、評点は低いものになるでしょう。

ただし、中間評価はあくまでも途中段階におけるチェックであり、仮に評点が低くても、中間評価時の評価パネルの意見を受け入れて、事業が改善されたのであれば、事後評価においては当該事業は高く評価されなければなりません。この意味でも、研究開発実施者や事業実施者にとって、中間評価は重要な意見交換の場であることが分かります。

出所：財団法人政策科学研究所作成資料

図 1-17　有効性のロジック展開（中間評価）

(3)効率性

中間評価の段階では、アウトプットの発生状況が確認できます。これらのアウトプットを得る過程で、効率性の観点からチェックすべき点は、主に研究開発の実施体制の妥当性、研究開発実施側ならびに事業実施側のマネジメントの妥当性、資金ショートの場合の措置、研究開発方法等の見直しといった、主に事業運営に関わる項目となります。

効率性に関わる観点は、事業の目標達成状況と密接な関係があります。いくら研究開発体制やマネジメントが優れていると評価されても、肝心の成果が思わしくないのであれば評価した意味がありません。単純に、「マネジメントは妥当だったか？」と問いかけられても、目標の達成状況が計画どおりであれば、「妥当である」としか答えようがありません。

したがって、中間評価の段階で効率性の観点から評価する場合には、どのようにすれば事業運営をより良く改善できるかという観点から実施されるべきです。

例えば、提案公募型事業の場合、提案書申請に関わる手続きの効率化、採択審査プロセスの改良等について制度実施期間中にどのような試みを行った結果、課題の応募数が増加している等の裏付けが得られれば、マネジメントの改善を明確に示すことができます。

1.3.3.2.3　事後評価で重視すべき評価項目

事後評価の段階では、事業終了に伴い、すべてのアウトプットが出揃っている状況にあります。事後評価の段階では、研究開発実施者に対して研究開発の軌道修正や改善提案を行う必要性がないために、その効用を疑問視する声もあります。

評価時期毎の意義の部分で解説したように、事後評価の一つの目的は当該事業の品質保証を行うことにあります。

公的資金を投入した研究開発事業が、当初の目標を達成し、かつ、質的に優れていたかどうかを検証する場が事後評価と言えます。

もう一つの目的は、事業実施主体のマネジメントに関わる問題です。当初の目標設定が目的に対して妥当であったか、研究開発体制として不備はなかったか、資金量は適切であったかどうかに関わる事後チェックとして機能し、後継の事業や、新規の事業展開のための教訓を得る場となります。

ここでも、「必要性」、「有効性」、「効率性」の三つの観点から見てみましょう。

(1)必要性

事後評価において事業の必要性を評価するということは、事業の目的とするものが有意義であったかどうかを判断することです。

事後評価において事業の意義を再確認するという行為は、仮に有効性、効率性の観点から当該事業が低い評価を受けたとしても、「事業目的に対して目標設定や研究開発方法、マネジメントなどが適切でなかった」という問題点の特定化のために必要なことと言えます。

また、事前評価の段階での状況と、事業が終了した段階での状況とを比較して、依然として事業の意義が失われていないのであれば、実施機関として次の施策展開を考える必要があります。

(2)有効性

事後評価においては、有効性の評価がもっとも重要になります。事後評価の段階で有効性を評価するということは、事業の目標の達成度を評価することです。再び、有効性のロジック展開図を見てみましょう。

まず、当該事業におけるアウトプットを悉皆的にリストアップすることから始めます。このとき、事業として意図していない副次的な成果も含めます。

次に、それらのアウトプットが事業の目標に対して寄与しているかどうかを精査します。ここで、事業の目標の達成状況が確認されます。

事業の目標の達成状況が確認されたら、事業の目的、すなわちアウトカム（意図した結果）が実現しているか、あるいは実現しそうかについての評価を行います。

出所：財団法人政策科学研究所作成資料

図 1-18　有効性のロジック展開（事後評価）

(3)効率性

事後評価の段階ではすべてのアウトプットが出揃い、事業の目標達成度が明らかになります。この段階において、はじめて社会経済的な効果について検討することが可能になります。事業の目標がそのような効果に対して設定されていたとすれば、この段階において効果の確認を必要とします。

しかしながら、市場導入や普及を伴うような効果の計測については、その可能性に対する精査が必要になります。市場関係者の意見聴取、開発された技術や試作品の引き合い等の実績がなければ、多くの場合は絵に描いた餅となりかねません。

基盤技術、観測技術、計算科学技術等における開発の成果については、その後の利用計画を明確にすることで、社会経済的効果を見積もることができます。

マネジメント等の評価項目に関しては、中間評価時の評価パネルからの提案を受け入れ、改善努力をしている場合には、高く評価する必要があります。

1.3.3.2.4　追跡評価で重視すべき評価項目

追跡評価では、事業終了から一定期間経過後、事業の目的が達成されているかどうかをフォローアップすることがもっとも重要になります。ここで、事業の目的とはビジョンや願望を含めた、「社会の新しいステージ」を意図したものです。

一般に、アウトカム評価では、ターゲットとする社会経済的な効果を指標のかたちで限定し、その変化を見ることで、政策の効果を勘案します。厳密な因果関係はともかくとして、結果がよければ、施策として展開した各事業が意味を持っていたと考える検証の仕方です。

例えば、ある技術開発プログラムの目的が、「技術開発を通じた新市場の創出と、企業競争力の強化」といったものであった場合、プログラムに参画した企業の製品開発状況、市場化の状況、収益の状況等の情報を収集し、プログラムに参画していない同業他社との数値と比較するなどします。この結果によって、一定の効果が認められるようであれば、プログラムは意味のあるものであったと解釈することができます。

科学技術の場合であれば、科学技術領域別の論文被引用度等のさまざまな指標に関する国際ランキングをモニタリングし、わが国の科学技術政策の効果を論じることが可能です。

追跡評価で得られるデータはどのようなものであれ、今後の政策研究に必要なデータとなります。

いかなる政策も単純に失敗だったということは断定できないでしょう。国が展開する政策は何かしらの影響を持って、現実の社会を構成しているからです。しかしながら、どの部分は効果的であり、どの部分では無効だったという点を整理し、次の政策につなげるよう、データを収集することは喫緊の課題です。「思いつき」のレベルの政策が「失敗だった」と言われないよう、政策自体の高度化のために追跡評価は必要であると言えます。

1.4　研究開発評価に関わる専門人材

本節では研究開発評価に関わる専門人材（Who?）について学びます。

研究開発評価は単に大学の先生を呼んできて、評価パネルを構成し、審議してもらうだけでは成り立ちません。研究開発評価には、その役割に応じて専門的能力を発揮する高度な人材が必要となります。

研究開発の評価活動に関わる人材は、その人材が担うべき評価機能により大別すると、レビューア（reviewer）、プラクティショナー（practitioner）、アナリスト（analyst）の 3 種類に区分できます。

表 1-9 評価人材の分類

分　類	レビューア reviewer (evaluator)	プラクティショナー practitioner	アナリスト analyst
定　義	評価パネルを構成し、評価対象の質的側面を専門的観点から明確にする。評価対象領域の専門的人材であり、エバリュエータとも呼ばれる。 ディシプリン内部の評価に携わるピアレビューアと、学際的ないし実務的内容に関する評価に携わるエキスパートレビューアとがある。	行政関連機関内部で評価の実務や運営に携わり、評価運営の実務的専門性を有する人材。 行政一般を担ういわゆる「ジェネラリスト」が、評価に係る組織内でのOJTや外部での教育・研修等の機会を経て評価の実務的専門性を獲得し、「プラクティショナー」と呼ぶに相応しい実務的評価人材に成長する。	評価対象を分析するための高度な手法を活かし、評価対象の実態を深く把握し、評価作業を専門的見地から遂行する人材。 評価に係る「スペシャリスト」であり、深い評価活動や経験等の研鑽を経て、「プロフェッショナル」と呼ぶに相応しい高度な手法を駆使できるようになる。

出所：財団法人政策科学研究所作成資料

本節では、特に欧州で通例的に使用されている人材分類の考え方に従って、研究開発評価人材の役割とキャリア・パスについて解説します。

1.4.1　レビューア

最初に、評価パネルを構成するレビューアについて解説します。

レビューアとは研究開発の質的側面に対して評価を行うことができる専門家です。多くは、研究開発の成果が属する専門領域における専門家であり、大学等の研究者、研究機関の研究者がその任を担います。

レビューアは採択評価の場合であれば、研究プロポーザルに目を通し、当該研究の新規性、独創性等を評価します。中間・事後評価では、産み出された成果に対して専門家としての評価を行います。

レビューアは組織を代表する者であってはなりません。あくまでも個人が主体です。同じ研究領域の発展、人材の育成のために社会的貢献活動の一端として、その役割を担う存在といえます。したがって、レビューとはボランタリーベースで行われるものであるという認識がレビューアと評価実施側双方に必要です。

また、レビューア自体の育成・蓄積については、何かしらの方法論があるわけではない、ということも認識すべきでしょう。彼らは自分たちが属する専門領域でのプロフェッショナルであって、レビュー・システムとはそうしたプロの観点からみた評価を重視するという仕組みです。評価を行うための能力・見識はレビューアとなる前の段階、すなわち研究者としてのキャリアによって育まれるものです。評価実施側ができることは、そのようなプロに対して、どのようなレビューを行って欲しいかの具体的な「依頼」をすることです。

1.4.2　プラクティショナー

（1）プラクティショナーとは

プラクティショナーとは行政関連機関（省庁、独立行政法人を含む）で研究開発評価の実務や運営に携わり、評価運営の実務的専門性を有する人材であるのことです。典型的な職種としては、プログラムの運営一般に携わる「プログラムオフィサー（PO）」 がこれに該当します。（参照　コラム：プログラムオフィサー）

しかしながら、PO のすべてがプラクティショナーというわけではありません。行政関連機関の内部職員は多くの場合、ゼネラリストであり、研究開発評価のための専門的人材と呼ばれるためには知識と経験を必要とします。

表 1-10 は米国の州立研究機関のプログラムマネージャーの中途採用に関する募集要件です。プラクティショナーと呼ばれる人材の必要要件がどのようなものか参考にしてみましょう。

表 1-10 プラクティショナーの応募要件（評価マネージャーの中途採用例）

職位　評価マネージャー（Evaluation Manager）

内容　研究開発評価ディレクターの下、機関の一つ以上の研究開発プログラムの評価を担当します。内部スタッフや外部協力者、研究開発実施者とともに、評価プランやモニタリングシステムの開発・改善を行います。

必須条件
・社会科学、公衆の健康もしくは関連する領域における修士の学位を持つ者。
・プログラム評価の経験が3年以上あること。
・質的データかつ（もしくは）量的データの評価・分析手法に関する知識を持つ者
・質的データや量的データの分析経験を有する者。

【基本的な業務内容（責任分野）】
1. プログラム評価全般のマネジメント
a. ロジックモデルや評価プランの開発
b. データ収集ツールの開発
c. データ収集作業の調整
d. データ入力やデータ統合の管理
e. 量的・質的データの分析
f. プログラムのモニタリングシステムの設計
g. 評価結果のフィードバックメカニズムやレポートシステムの開発
h. 評価手続きや評価方法論の一貫した適用

2. 内部スタッフや外部協力者とともに、評価能力形成を支援する
a. 評価戦略やデータ収集、パフォーマンス計測に関する内部スタッフへの指導
b. 研究開発実施者に対する評価の技術的な支援を提供
c. 評価やデータ活用における能力開発の管理
d. 内部スタッフに対し、政策決定や意思決定に評価データを活用するための指導
e. 外部協力者との効果的な共同作業

3. 評価スタッフと研究開発実施者のマネジメント
a. 評価スタッフの統括
b. 外部発注のマネジメント
c. 研究開発実施者との契約とモニタリング

4. 監査と成果の普及啓発
a. 監査報告書の作成
b. 内部スタッフとの協働による出版物等の普及啓発の促進
c. 他の評価スタッフによる報告書の統括

5. 評価委員会メンバーとしての役割
a. 評価委員会への貢献
b. 機関ミッションに関するワークグループや戦略企画会議への参加
c. 必要に応じて評価委員会と評価プランやマネジメントについて協議
d. プロポーザル開発や資金獲得業務への支援

6. 望ましいスキル、知識、態度
1. 口頭ならびに文書によるコミュニケーション・スキル
2. チームワークと良好な対人関係
3. プロジェクト・マネジメントのスキルとグループの調整スキル
4. 細心の注意力
5. 前向きな思考と自主的な働き
6. 多種多様な業務をこなすことの出来る組織的スキル
7. サーベイやデータ収集ツールの開発経験
8. SPSS や SAS などの主要な統計ソフトウェアを用いた分析経験
9. オフィスソフトを十分使いこなせるスキル

給与
マネージャー職：5000～6500ドル／年
シニア・マネージャー職：6000～7500ドル／年

出所：<http://www.monster.com/>より財団法人政策科学研究所作成

上の例は、公衆の健康を改善するための研究開発プログラムを実施している州立研究所のプログラムマネージャーに関する応募です。必須の要件として、公衆健康に関する修士以上の学位を要求しているのはそのためです。一方で、公衆健康のような社会的インパクトを評価するためには、社会科学の知識が不可欠です。したがって、社会科学の修士以上の学位がまず第一に要求されています。

また、データ分析に関する知識と経験も要求されています。この州立研究所では、内部スタッフ以外に外部の協力者と連携して評価業務を遂行していますが、データ分析を自前でやるにせよ、外注するにせよ、データ収集、データ分析の方法論についての知識が必要不可欠であることを示しています。これらの知識を教育・指導することも業務のうちに含まれています。

当然のことながら、組織的業務を営むためのさまざまなコミュニケーションスキルも重視されます。したがって、ある程度の経験を積んだ人材でなければ、プログラムマネージャーは務まりません。必須条件にプログラム評価に３年以上従事した経験が盛り込まれているのは、最低でもそのくらいの習熟期間を必要としている業務内容であるためです。

(2)プラクティショナーのキャリア・パス

科学技術政策の先進諸国において、一般的にプラクティショナーと呼ばれる評価専門人材はさまざまなバックグラウンドを持ちつつも、評価論に関する学習の機会を獲得し、評価マネジメントに携わっています。学習の機会がどのようなものであるかについては、研究開発マネジメントの発達に併せて変遷してきました。例えば、スウェーデンの研究開発資金配分機関である VINNOVA（スウェーデン・イノベーションシステム庁）の評価部門（実質的な戦略形成部門）は分析能力を備えた高度人材が集積していますが、彼らのキャリア・パスは世代によって異なっています（図 1-19 参照）。

VINNOVA のシニアクラスのマネージャー（もしくはディレクタークラス）は研究開発評価の創世記から歩んできた世代と言えます。技術者・研究者としてのキャリアから、オンザジョブトレーニング（OJT）等によって、試行錯誤を繰り返しながら評価専門人材としてのキャリアへ転換を図りました。この世代の蓄積が、後に高等教育機関の専門コースの土台となっています。

ミドルクラスのマネージャーは、ある時期に職場を離れ、高等教育機関の専門コースで再学習し、評価専門人材としてのキャリア転換を図っています。専門コースでは、研究開発マネジメント、経営学、経済学などの主に社会科学関連の修士以上の知識を習得しています。

もっとも若手のマネージャーもしくはスタッフは、最初から社会科学関連の大学院を修了し、専門職として採用されています。彼らの課題は、科学技術に関するバックグラウンドがないために関連する分野に関する知識を OJT で学ばなければならないことですが、研究開発評価やマネジメントについての専門性は最初から身に付けています。

出所：財団法人政策科学研究所作成資料

図 1-19　VINNOVA におけるプラクティショナーのキャリア形成パターン

また、学習の機会として現在でも非常に有益とされているのは、各国のプラクティショナーが集う研究開発評価や科学技術政策（あるいはイノベーション政策）に関する国際会議への出席と、そうした場における情報交換です。欧州やアメリカを中心に、このような国際会議が年複数回開催されているため、人々は積極的に参加し、研究開発評価のネットワーク作りを行っています。

我が国においても、総合科学技術会議で評価専門人材の育成と蓄積が急務であることが議論されました(「平成 16 年度の科学技術に関する予算、人材等の資源配分の基本的方針」)。研究開発評価を効果的を推進するためには評価を含めた研究開発マネジメントを担当する専門人材の育成こそが重視されなければなりません。マネジメントに関する知識や経験は外部から人を調達すればよいというものではなく、組織内部で自前で蓄積しなければ定着しません。このような観点からすると、大学の研究者を PO として一定期間出向させるような評価制度では、研究開発評価がおぼつかないだけでなく、地に足の着いた戦略形成もままならない状況が今後も続くと言わざるを得ません。

本テキストはプラクティショナーの入り口を示すための手引きとお考え下さい。

コラム：プログラムオフィサー

プログラムオフィサー（PO）とはプログラムの運用に携わる者という意味です。PO の職階が、プログラム・ディレクター（PD）とプログラムマネージャー（PM）です。

出所：財団法人政策科学研究所作成資料

図 1-20 プログラムオフィサー

PD は当該プログラムの責任者です。科学技術政策の先進諸国では通常、プログラム内部に関わる予算執行等の決裁権限や PM の人事権を有しています。

PM は実務を遂行する専門職です。プログラム内で遂行されるプロジェクトの企画、公募、採択、モニタリング、評価等の運用に携わります。

この他に業務を補佐するスタッフを含め、プログラム運営の人材を PO と総称します。

「競争的研究資金制度改革について」(平成 15 年 4 月 21 日総合科学技術会議)を踏まえて、2004 年度に導入された我が国での PD・PO の呼称には、定義上の混乱があったと言えます。

また、我が国では政策のプログラム化が進んでいないこともあって、PD、PO の名称を外部からプログラム運営のために招聘された者のみに当初用いられていました。その後、PM に内部職員を充てるケースも出てきていますが、人事ローテーションのサイクルが早いため、評価の実務的経験を組織的なナレッジとして蓄積することの困難性が指摘されています。

例えば米国 NSF（National Science Foundation）では常時 600 人の PM が在籍していますが、４割は１～３年の期限付きの大学研究者等からなる出向職員、６割がパーマネントの職員で構成されています。持続的、発展的なプログラム・マネジメントが実施されるためには、内部職員からなる PM の集積が不可欠であると言えます。

1.4.3　アナリスト

アナリストとは、研究開発評価の対象となるさまざまな事象を、専門的なツールを用いて分析することができる人材を言います。ある手法や技能についての専門性を有するだけでしたら、大学や研究機関の研究者もアナリストと呼べるかもしれませんが、研究開発評価の戦略、目的、枠組みの中で評価手法を位置づけ、全体を俯瞰しつつ分析もできるという点で、より職業的能力を発揮する人材がイメージとなります。具体的には、研究開発評価を専門とする高等教育機関の研究者や公的機関の戦略形成部門の専門家、シンクタンクの専門家などが相当します。

アナリストはプログラム評価やナショナル・プロジェクトの評価に貢献します。研究開発のインパクトを分析の対象とすることが多いため、社会科学関連のバックグラウンド（経済学、経営学、社会学等の PhD）を持った人材が主流です。彼らの駆使する分析手法は高度にシステマティックであり、データ収集、データ分析、システム分析などに関する知識が要求されます。また、評価の方法論だけではなく、政策措置（policy instrument）や科学技術の知識についても造詣がなくては研究開発評価を分析面から担うことはできません。

アナリストは、一定の専門性を身に付けた上で研究開発評価の支援を行うため、その育成方法は OJT によるものが主流となっています。

経験を積んだアナリストは「プロフェッショナル」と呼ばれます。

1.5　研究開発評価の方法論

本節では研究開発評価の方法論（How）についての概要を学びます。はじめに研究開発評価の方法論の概要を把握し、次に効果計測の考え方や指標の見方について解説します。さらに、研究開発の質に関わる評価方法として一般的なレビュー法の基本について解説します。

1.5.1　方法論の一般論

1.5.1.1　調査＝分析＝評価のための方法論

研究開発評価を実施するためには、評価対象に関する情報やデータをまず収集し、これらを分析した上で、評価の俎上に載せる必要があります。正しい分析のためには綿密な調査が必要であり、正確な評価のためには正しい分析が必要であるという意味で、これらのプロセスは互いに密接に関係しています。しかしながら、このプロセスは機能としては独立であり、研究開発評価の業務プロセスを考えた場合には区別しなければなりません。

出所：財団法人政策科学研究所作成資料

図 1-21　研究開発評価のプロセス

特に、研究開発に伴って得られる情報やデータのみに依拠し、評価の目的を構造化して専門的な分析することをせず、レビューアに諮るような評価の仕方を「単純評価」と呼びます。我が国の例で言えば、科学技術研究費補助金（科研費）の個別課題に対する評価がこの方法であると言えます。

一方、単純評価で使用される情報やデータ以外に幅広く情報を収集し、評価の専門家（アナリスト）がその専門性を発揮して分析を加え、その結果を下にレビューアに諮るような評価の仕方を「複合評価」と呼びます。評価対象である研究開発プロジェクトやプログラムが実施機関側の何かしらの政策目的やミッションに基づいて行われる場合には、効果計測・分析のプロセスが必要になりますので、複合評価でなければ評価したことにはなりません。

表 1-11 は調査＝分析＝評価のための方法論を体系的に示したものです。

表 1-11　研究科発評価の方法論

調査法		分析法	評価手法	
情報収集	データ収集		手法区分	
インタビュー（関係者） ヒアリング（専門家） アンケート調査 事例調査 文献調査 顧客調査 社会調査	既存統計 アンケート調査 申請・管理データ データマイニング テキストマイニング ： ： ：	ケース分析 コンテント分析 計量文献分析（ビブリオメトリックス） 　引用分析、共引用分析、共語分析 　謝辞分析 計量技術分析（テクノメトリックス） 　特許分析、サイエンスリンケージ分析、 　技術仕様分析 計量経済分析（エコノメトリックス） 　費用分析、費用便益分析、費用効果分析 　生産関数分析、消費者余剰分析 　財務指標分析 計量社会分析（ソシオメトリックス） 　表明選好法、各種指標分析 多変量解析 システム分析	定性的評価	レビュー法 　メールレビュー，ブラインドレビュー パネル法 　ピアパネル、エキスパートパネル 　ボードパネル
			半定量的評価	評点法
			定量的評価	指標法 　単一指標、複合指標 比較評価法 　対比較年度比、対計画比 　コントロールグループ比、前後比較 　非実施仮説比 ランキング ベンチマーキング ポートフォリオ

出所：財団法人政策科学研究所作成資料

調査法には評価対象に関連する情報の質的側面に関する情報収集と、量的側面に関するデータ収集とがあります。情報収集として一般的なものが関係者に対するインタビュー調査や、評価対象の関連分野における専門家へのヒアリング調査、アンケート調査などがあります。また、同様の研究開発に関する国内外の事例調査、文献調査なども含まれます。評価対象の目的が社会的インパクトや経済効果にある場合には、顧客調査（市場調査）や社会調査が必要になるでしょう。

データ収集では、評価対象に関連する既往の公式統計や調査報告等に掲載されているデータにあたることが必要になります。データが存在しない場合にはデータ化するための作業が必要になります。関係者に対するアンケート調査、マネジメント側が保持している申請・管理のデータ、文献情報からのテキストマイニング、データマイニングなどを実施し、望ましいデータを探索することになります。

データ収集に関して、特に対象の質的側面に対して何らかの測定基準を設け、その基準に照らし合わせて測定し、データ化する手法を「メジャメント（measurement）」と呼びます。

分析法の多くは、評価対象についての仮説検証を行うためのツールです。特に、既に測定されている統計データや数量データを加工し、評価対象の評価項目の内容を表すように分析結果を示したり、評価項目との関係性を裏付けるようなデータを得るような定量的分析手法のことを総称して「メトリクス（metrics）」といいます。メトリクスの多くは統計的手法を用いるため、データサンプルが多くなければならないので、一般的には経年的、もしくはプログラムのように複数のプロジェクトに横断的な分析において使用されます。

表中のケース分析は事例比較から傾向を抽出する定性的な分析法です。コンテント分析

はテキストマイニングの結果から統計的推論を用いて構成概念の枠組み作りを行います。計量文献分析（ビブリオメトリクス）は学術論文などを対象に成果の利用度指標を作成するのに用いられます。計量技術分析（テクノメトリクス）は特許データを対象に分析を行います。計量経済分析（エコノメトリクス）はコストデータや売上、社会費用などの経済データを分析の対象にする諸手法の総称です。計量社会分析（ソシオメトリクス）は社会的影響に関する指標化のツールです。これらの複合的な利用と発展的な分析方法に、各種多変量解析やシステム分析を組み合わせることになります。

メジャメントやメトリクスを活用するためには高度な専門知識を必要とするため、個別手法に精通したアナリストが分析の任を担います。

評価法にもさまざまな手法がありますが、これらのどれを採用するかという問題は、最終的な評価をどのようなかたちで行うか、という問題に置き換えることができます。一般的には、評価パネルを構成し、そこでさまざまな情報、データ、分析結果を勘案し、評価対象に総合的な評点をつけることになるでしょう。その意味では、表に挙げた定量的評価の手法群は分析法に含まれるかもしれません。

研究開発評価においてもっとも重要な評価法はレビュー法、パネル法です。科学技術の質的に深い側面を捉えるためには、ピアもしくはエキスパートの専門的知見による方法しかありません。ただし、その運営は非常にデリケートであり、高度なマネジメントが要求されます（第２章参考）。

評点法（scoring）は本来定性的な評価内容を評定区分に照らし合わせて採点する方法です。ここで注意しなければならないことは、評点（score）それ自体は絶対的な意味を持つ数値ではないということです。これは、あるパネルの評点と別のパネルの評点を単純に足して２で割ることに意味がないことを考えれば自明でしょう。評点法の導入は、評価結果を明晰に示すことができるという点で望ましいものではありますが、評点そのものが重要なのではなく、なぜそのような評点がついたのかという理由（コメント）を明らかにすることが重要であると理解して下さい。

定量的評価のうち、指標法は評価対象に関するさまざまな定量的表現を正規化し、指標（indicator）を作成して他の評価対象と比較する方法です。単一のデータから指標を作成する場合を単一指標、複数のデータを組み合わせて指標を作成する場合を複合指標といいます。

比較評価法は評価対象のパフォーマンスを相対的に評価する方法です。代替案がない場合には、評価対象の対前年比や前後比較、あるいは対計画年比をとります。プロジェクトやプログラムの関与が及ばなかったグループと比較して、評価対象のパフォーマンスを検証するコントロールグループ・アプローチや、プロジェクトやプログラムがなかった場合を想定して、それをベースラインとした時と比べてパフォーマンスがあったかどうかを実証的に検証する非実施仮説比があります。

ランキングはある指標に関して、評価対象群を上位から並べて比較する方法です。ベン

チマーキングはひとことで言うと「それまでのベスト」と比較してどうかを見るという評価方法です。実験結果に関するチャンピオンデータなどのベストプラクティスを参照し、評価対象の優秀性を判断するのに用いられます。ポートフォリオは評価対象について、２つ以上の指標を用いて評価座標軸を作成し、評価対象群の偏りを見るための手法です。

1.5.1.2　定性的評価と定量的評価

先の表に示した調査法、分析法、評価法においては定性的な情報を扱う手法と定量的な情報を扱う手法とがあります。

評価に関わる誤解の一つに、「定量的評価＝客観的」、「定性的評価＝主観的」という考え方が根強くあります。定量的手法は一般的に結果の明晰性が高く、情報伝達の際の任意性が低いという面がありますが、反面、評価対象のある局面のみを限定的に切り取って数量化しているという特徴があります。

定性的手法は結果の表現を主に文章による叙述に頼ることになりますが、事象の本質を捉えることに関してはむしろ適切な場合の方が多いと言えます。

例えば、ワインの解説をする際に、ワインの甘辛度、酸味度、タンニン含有量、アルコール度等を指標化したところで、本当にそのワインが美味しいかどうかの判断材料にはなりません。むしろ、「ブラックベリーやプラム、カシスといった熟したフルーツの風味が・・・。」というような表現にこそ、飲む人の嗜好を判断させるための本質的な内容が含まれていると言えるでしょう。

通常、研究開発評価の俎上に上がるものは、

・先行的事例に乏しく、比較不可能な事象を扱うケースが多い
・一般化が困難で、マネジメントなどの経験的・学習的要素が多い
・因果関係が明らかでなく、先行指標を得ることが難しい（妥当でない）

というような問題を含む対象であることがほとんどです。

特に研究開発評価では、成果の質に関わる評価をピアレビューアやエキスパートレビューアに委ねることになります。レビューアは彼らの知識や経験を総動員してコメントを作成します。

したがって、「定性的評価だから主観的であり、客観性に欠ける（＝信頼性が低い）」ということにはなりません。

むしろ定性的評価において重要なことは、「どのような文脈で語られているか？」ということです。歴史認識、比較対象、課題の喫緊性などが背景にあった上で、本質的な意味づけが行われることになります。

1.5.2　定量的評価法の基本

1.5.2.1　比較考量の基本的な考え方

ある研究開発のパフォーマンスを定量的に示すためには、比較の視点が絶対に必要になります。代替案（alternatives）が始めから存在していれば、比較は簡単でしょう。しかしながら、研究開発はプロセスの革新であれ、改良であれ、基本的に「これまでにない」モノを作り上げる行為に他なりません。そのため、比較考量をどこに求めるかが大変重要になります。

研究開発評価においてパフォーマンスを比較する場合の基本的な考え方を以下にまとめました。

（1）チャンピオンデータ

研究開発のアウトプットの性能等を比較する際に、もっとも直感的で分かりやすい方法は、当該分野のチャンピオンデータと比較するというものです。通常は、これまで世界一だったスコアを大きく更新したという場合にこの方法が用いられます。

図 1-21 は地球シミュレータの演算性能を比較した図です。地球シミュレータがそれまで世界最速であった米国のスーパーコンピュータ ASCI-White を実効性能で５倍も引き離した事実をもっとも雄弁に物語っています。

出所：財団法人政策科学研究所作成資料

図 1-22　チャンピオンデータとの比較の例（地球シミュレータの演算性能）

また、世界一の更新でなくても、性能に関わる数値を費用対効果の指標として示すことで、コストパフォーマンスをチャンピオンデータと比較することもできます。

他に比較可能な対象がある場合には、まずは性能指標等に関して直接比較を行うことが原則であると言えます。

（2）前後比較（"Before-After" comparison）

ある研究開発を実施する前と後とで、ターゲットにしている指標がどのように変化したかを見る比較の方法です。例えば、ある技術開発プログラムに参加した企業の製品技術が、プログラムに参加する前とした後で比較して、大幅に向上しているような場合に、それを当該プログラムのパフォーマンスとして示す考え方です。

出所：財団法人政策科学研究所作成資料

図 1-23　前後比較の考え方

しかし、前後比較は当該技術分野に実績を有する企業がプログラムに参加している場合などには、プログラムのパフォーマンスとしての因果関係を保証する比較の仕方ではありません。その企業が元々行っていた研究開発の蓄積分や継続努力が結びついて性能を上昇させている部分と、当該プログラムによって研究開発が加速された部分とを切り分けて考えないと、本来のプログラムのパフォーマンスを示したことにはならないからです。

(3)非実施仮説("With-Without" comparison)

前後比較が観測される性能指標のある一定期間内の向上分を示しているのに対して、非実施仮説による比較考量は、当該プログラムの「真の貢献」部分を示します。

先ほどと同じ例で考えてみましょう。技術開発プログラムに参加している企業は、当該技術分野における研究開発の蓄積があり、ある時期から国が実施しているプログラムによって研究開発資金を助成してもらい、研究開発を実施したとします。

ここで、当該技術開発プログラムの寄与度はどのように考えたらよいでしょうか。もし、その企業が技術開発プログラムに参加せず、自前で研究開発を継続したとしたら、製品の性能指標はどのように推移していたでしょうか。これを推計するための基本的な考え方が「非実施仮説」です。図 1-24 を用いて非実施仮説によるパフォーマンスの計測方法を解説します。

出所：財団法人政策科学研究所作成資料

図 1-24 非実施仮説の考え方

仮にそれまでの性能指標のトレンドがデータから得られたとしましょう。この時、企業が自前で研究開発した場合の達成したであろう性能指標の値は P*2 になります。一方、観測された性能指標の値は P2 です。P*2 と P2 の差は、技術開発プログラムに参加したことで、当該企業の研究開発が加速された部分と解釈できます。つまり、この差を技術開発プログラムのパフォーマンスとして示すことができます。

非実施仮説による比較考量は、継続性のある事象の評価では不可欠です。「もし、研究開発プログラムに参加していなかったら（Without)」という状況を、BAU （Business As

Usual)ケースと呼びます。BAU で推計される指標と、「プログラムに参加した場合（With）」の現実の指標との差が、非実施仮説における比較考量となります。

(4)コントロールグループ・アプローチ

非実施仮説では、文字通り実施しなかった場合を想定することで比較考量としていますが、実施しない場合を現実のサンプルに求めることもできます。この比較考量の求め方をコントロールグループ・アプローチと呼びます。このもっとも代表的な例が、新薬の治験（臨床試験）で行われる「比較試験」です。

新薬の治験では、プラセボ（偽薬）を用いて薬理効果を検証します。 新薬を投与する患者のグループと、全く治療効果が無い「プラセボ」を投与する患者のグループ（プラセボ群）に分けて、それ以外は全く同じ条件で治験を実施し、それによる臨床データを測定し、両者のデータを比較して治療効果を確かめます。プラセボは新薬と見た目も味までも同じに製剤されているので、患者はもちろんのこと、投与する側の医師ですら区別することができません。偽薬の投与によっても症状が改善してしまうというプラシーボ効果すらもデータのバイアスとして考慮した非常に厳密な比較試験です。

出所：財団法人政策科学研究所作成資料

図 1-25 コントロールグループ・アプローチの考え方（例：新薬の比較試験）

このとき、新薬を投与されたグループは対象群（コントロールグループ）と呼ばれ、プラセボを投与されたグループは非対象群と呼ばれます。対象群と非対象群との間に統計的な明確な差異が認められたとき、新薬の薬理効果が実証されるのです。

研究開発評価においてコントロールグループ・アプローチを使用する場面はサンプルを多く必要とするプログラムレベルでの評価に限定されると考えられます。例えば、プログラムに参加した企業と参加していない企業との間にどのようなパフォーマンスの差異があったかを追跡調査等で明らかにするような場合が相当します。

1.5.2.2　数値データはどのように扱うべきか

私たちは普段気にすることなく、数値化されたものをデータと呼んでいます。しかし、データには種類があり、それらを理解しなければ誤った活用をしてしまうことになります。まずはデータの種類を学びましょう。

データには大きく分けて質的データと量的データとがあります。質的データは数値であっても四則演算が適用できないデータのことです。それに対して量的データは四則演算が可能なものといえます。

出所：財団法人政策科学研究所作成資料

図 1-26　データの種類

質的データはさらに名義尺度（nominal scale）と順序尺度（ordinal scale）に分かれます。名義尺度とは、カテゴリーの分類に使用されるデータのことで、その数値の順序、大きさには全く意味のない情報のことです。データベースのコード番号なども名義尺度の一つです。順序尺度とは大小関係や順位の相対的な比較に用いられる判定条件に対応したデータです。研究開発評価における評点は順序尺度になります。順序尺度は数値の値そのものに意味があるわけではないので、平均値を取ることができません。しかしながら、最頻値（mode）や中央値（median）を取ることで代表性を表現することが可能です。

量的データも間隔尺度（interval scale）と比例尺度（ratio scale）に分かれます。間隔尺度は距離尺度とも呼ばれますが、数値の差にのみ意味があるデータです。もっとも一般的なのは例えば 0～1 の範囲で正規化された指標です。つまり、パーセンテージ（%）で示されるデータです。間隔尺度では数値の相対的な距離を定めているだけですので、比を取ることができません。「＋4 ポイント」という表現はできても、「3%上昇」という表現はできません。

量的データのうち、比例尺度になってはじめて、四則演算の全てが適用できます。比例尺度のデータとは絶対的な値を持つデータのことです。つまり、比が定義できる絶対零点を持つデータです。

表 1-12 データの4分類

分類	説明	例	対応するデータ
名義尺度 Nominal scale	カテゴリーの分類に使用される。平均値を求めても意味がない。	血液型:A→1，B→2，O→3，AB→4 性別:男→0，女→1	質的データ
順序尺度 Ordinal scale	大小関係や順位の相対的な比較に使用される。数値の絶対的な大小に意味がないので、平均値を取ることはできない。中央値であれば可。	判定:とても良い→5，良い→4，普通→3，悪い→2，とても悪い→1	質的データ ※順序や大小関係に等間隔を仮定すれば、間隔尺度として計算できる。
間隔尺度 Interval scale	距離尺度とも呼ぶ。数値の差にのみ意味があるだけなので、「%上昇」といった表現はできない。	温度(摂氏)、テストの得点	量的データ
比例尺度 Ratio scale	比尺度とも呼ぶ。数値の差とともに数値の比にも意味がある。比が定義できることは絶対零点を持つことを意味する。	温度(華氏)、距離、時間、重量	量的データ

出所：財団法人政策科学研究所作成資料

さて、データの種類について理解が深まったところで、研究開発評価における数値データ活用の例を見てみましょう。

(1)評点の場合

研究開発評価の評点について考えてみましょう。ここでは、新エネルギー・産業技術総合開発機構における評点法の適用例を見てみましょう。この場合の評点と評定区分の定義は以下のようになっています。

表 1-13　新エネルギー・産業技術総合開発機構の評定区分

判定基準	
1. 事業の位置づけ・必要性について	
・非常に重要　→A ・重要　→B ・概ね妥当　→C ・妥当性がない、又は失われた　→D	考慮事項（参考） (1) NEDO、産総研の事業としての妥当性 ［a b c d］ (2) 事業目的の妥当性 ［a b c d］
2. 研究開発マネジメントについて	
・非常によい　→A ・よい　→B ・概ね適切　→C ・適切とはいえない　→D	考慮事項（参考） (1) 研究開発目標の妥当性 ［a b c d］ (2) 研究開発計画の妥当性 ［a b c d］ (3) 研究開発実施者の事業体制の妥当性 ［a b c d］ (4) 情勢変化への対応等 ［a b c d］
3. 研究開発成果について	
・非常によい　→A ・よい　→B ・概ね妥当　→C ・妥当とはいえない　→D	考慮事項（参考） (1) 目標の達成度 ［a b c d］ (2) 成果の意義 ［a b c d］ (3) 特許の取得 ［a b c d］ (4) 論文発表・成果の普及 ［a b c d］ (5) 標準化（産総研担当分） ［a b c d］
4. 実用化・事業化の見通しについて	
・明確に実現可能なプランあり　→A ・実現可能なプランあり　→B ・概ね実現可能なプランあり　→C ・見通しが不明　→D	考慮事項（参考） (1) 成果の実用可能性 ［a b c d］ (2) 波及効果 ［a b c d］ (3) 事業化までのシナリオ ［a b c d］

出所：新エネルギー・産業技術総合開発機構作成資料

上の評定区分に従って、評価委員会の各委員が評点を付けます。その結果が下のような図で開示されています。

評価項目	平均値	素点（注）						
1．事業の位置付け・必要性について	2.0	B	B	C	B	C	C	B
2．研究開発マネジメントについて	1.2	B	B	A	B	A	A	D
3．研究開発成果について	2.5	C	B	C	B	D	D	C
4．実用化・事業化の見通しについて	1.5	D	B	A	C	B	A	B

（注）A=3，B=2，C=1，D=0 として事務局が数値に換算し、平均値を算出。

出所：新エネルギー・産業技術総合開発機構作成資料

図 1-27　新エネルギー・産業技術総合開発機構の評定結果の例

このような結果を得るためには、前提条件として、順序尺度である評点に対して間隔尺度を適用しているという点が指摘できます。つまり、評点 A と B の差は C と D の差に等しいという解釈を前提としています。また、個人の評点の差についても、個人間で互いに等しいという解釈を前提としています。

評価結果を明示的に表すにはこのような表現方法もやむを得ないと思われます。しかしながら、より重要なことは、このような評価結果を得るに至る評価委員会でのマネジメントの在り方となります。本来、個人の順序尺度である評点を、評価委員会共通の間隔尺度として用いるためには、評点を導くための評定区分に対する理解が委員会で共有される必要があります。評定区分に対して解釈の個人差が大きく生じないように、評価シートが設計されていなければなりません。

また、こうした評点を示すとしても、各委員の評点（素点）がどのような分布になっているかを補助的に示すことが大切になります。平均値だけを見た場合、分布が平均値の周辺に集まっている場合と、極端に評点が割れている場合との違いが判別できないからです。

出所：財団法人政策科学研究所作成資料

図 1-28　平均値が同じでも分布が異なるパターンの例

表 1-14　経済産業省の技術評価結果（3点満点）の例

評点結果

評　価　項　目	平均点	標準偏差
1．事業の目的・政策的位置付け	2.83	0.41
2．研究開発目標の妥当性	2.33	0.52
3．成果、目標の達成度の妥当性	1.83	0.41
4．波及効果、事業化についての妥当性	2.00	0.63
5．研究開発マネジメント・体制・資金・費用対効果等の妥当性	1.83	0.41
6．総合評価	2.50	0.55

出所：経済産業省平成 18 年度技術評価報告書より抜粋

したがって、NEDO 技術開発機構の例のように、各評価委員の素点を示すか、経済産業省の実施している研究開発プロジェクトの技術評価のように、平均点の他に標準偏差（値が大きいほど評点が割れている目安になります）を示すなどの工夫が必要となります。

(2)費用便益比の場合

研究開発評価において、費用便益比を計算することは極めて難しいと言えます。事業終了直後では研究開発投資の効果が出現していることはほとんど希だからです。また、さまざまな効果を金額換算する方法も用意されてはいますが、一律に適用できるような換算値は存在しません。しかし、何らかの予測を立て、かつ高度な金額換算手法を用いて便益が計算できたとすれば、研究開発投資費用に対する便益比が計算できます。

問題は、このような比率の持つ数値的意味です。費用便益比は当該プロジェクトの経済性効率（パフォーマンス）を見ているだけであって、利益の大きさを見るものではありません。また、分子と分母がそれぞれ独立に計算されているものを対比した複合指標であるため、他のプロジェクトとの比較において絶対基準を持ちません。したがって、非常に高度な計算を行い、定量的分析を厳密に行っているように見えても、比較の観点からは、費用便益比は間隔尺度と同じと言えます。

つまり、プロジェクトが経済効率的（この言葉の意味も本来は深く考える必要がありますが）に実施されたかどうかについて、他のプロジェクトとの相対的な比較の観点なしに「大変効率的に実施された」、「ある程度効率的に実施された」、「効率的に実施されたとは言い難い」というような評定区分を客観的に数値化したということと同じ意味合いしか持ちません。

費用便益比を比較可能な数値として意味づけるためには、「投資の機会費用」という考え方が不可欠です。つまり、投資の一部を金融機関からの借り入れで賄う場合、かつ、事業に財務的収入が事業の想定している計画年限に見込める場合においてのみ意味を持ちます。なぜならば、この場合、費用便益比（あるいは内部収益率）が大きいほど、投資の回収期間が短い、つまり収益を生みやすいという説明になるからです。ですから、民間の研究開発投資のために試算することには一定の意味があります。

翻って、国が行う研究開発投資においてすべからく費用便益比を計算する意味がどこまであるでしょうか。個別の研究開発プロジェクトの目的にもよりますが、おそらくは計算する労力が大きいだけで、実務的な意味合いからは全くないと言っても過言ではないでしょう。あるとすれば、事業管理主体の財務当局へのエクスキューズか、財務当局自身の国民へのエクスキューズという意味合いしか残らないでしょう。そうした数値が保証するものは、「お金を無駄に使ったわけではない」という意味であって、「どれほど役に立ったか」を明らかにしていないからです。

心ある経済学者は費用便益分析よりも費用効果分析を推奨しています。費用効果分析では、効果に相当する数値を物理量であらわすため、単純に何に役に立ったかを明らかにし

やすいからです。例えば、温暖化対策のための研究開発投資であれば、成果がもたらす CO2 削減量 1 トンに対する投資額が明確になります。3,000 円／t-CO2 という数値であれば、他の研究開発プロジェクトとその効果を比較することができるからです。この場合の数値は比例尺度である言えます。

1.5.3　研究開発の「質」に関わる評価法

一般的に、研究開発の「質」を評価するためのもっとも有効な方法はピア、エキスパートによるレビューです。我が国の研究開発評価の制度は、早くから外部の専門家、有識者等による評価を導入し、評価委員会や評価分科会などの評価パネルが研究開発評価の主軸を担ってきました。

しかしながら、レビュー評価法は研究開発の、とりわけ成果の「質」については有効な評価方法ですが、その他の評価項目に関して万能に機能するわけではありません。また、その「質」に関しても、ある単独のディシプリン内の評価で済むものから、学際的な評価が必要な場合、社会経済的影響が大きいものについては、関係者の意見が反映されなければならないものまで、さまざまです。

つまり、学識経験者や有識者と呼ばれる方々を単純に配置すれば良いのではなく、評価対象の特性に応じてレビューの仕方やレビューアの構成を考えなければなりません。また、評価パネルを設置し、委員長を任命したら自動的に審議と評価が行われるのではなく、パネルの運営も重要な仕事になります。

ここでは、原則論としての評価パネルの在り方を考えてみましょう。

1.5.3.1　評価対象に応じたレビューの仕組み

評価対象である研究開発事業の目的に応じて、構成されるレビューアのタイプが異なります。

一般に、レビューアのタイプにはピア（Peer）とエキスパート（Expert）があります。ピアの原義は、「仲間内、同業者」という意味です。研究開発評価においてのピアレビューの意味は、あるディシプリン（学問の一分野）内における専門的知識を持った研究者仲間内での評価ということになります。

一方、エキスパートは「有識者」と言えば良いでしょうか。単一のディシプリンだけでなく、複数の学問分野、あるいは技術領域、法制度、経営、マーケティング等に関する幅広い知識を持ち、研究開発の成果に対して複眼的なものの見方ができる人材が相当します。これを担うのは研究者だけでなく、実業家、専門家、行政職等の人材であり、これらの人々が研究開発事業等の目的に応じて配置されます。

表 1-15 ピアとエキスパートの適用区分

	科学技術的価値の側面を重視	社会経済的価値の側面を重視
単一ディシプリン	・ピアレビュー ・ピアパネル	
複合的ディシプリン	・エキスパートレビュー ・エキスパートパネル	・エキスパートレビュー ・エキスパートパネル ・エキスパートボード

※ディシプリン（Discipline）の原義は「しつけ」。学問領域毎に専門的知識を見つける「しつけ」や「訓練方法」が異なるため、それらが機能する「特定の研究領域」の意味に用いられるようになった。

出所：財団法人政策科学研究所作成資料

上の表 1-15 はピアとエキスパートの適用区分についてまとめたものです。研究開発成果の科学技術的価値を重視した評価を行うことを前提に考えてみましょう。単一ディシプリンの場合でしたら、研究開発の成果は当該分野の研究者内でのレビューをもとに評価を行うことしかできません。一方、複合的ディシプリンの場合、さまざまな領域の専門家が集い、お互いの評価について話し合うことが重要になります。学際的領域や新領域については、異なるディシプリンの専門家が集い、各々評価を行うにせよ、パネル・リーダーはそれらの言い分に理解のある、幅広い見識を持った人材（エキスパート）でなければなりません。

また、社会経済的価値の側面を重視した評価を行う時には、単一ディシプリンのピアレビューは原則として意味を持ちません。社会経済的価値そのものが、さまざまな学問分野の関心領域だからです。むしろ、ここで問題となるのは、より評価の難しい、社会的インパクトが大きいと考えられる事業やプログラムの評価において、専門家の判断だけによらない評価の枠組みが新たに必要となる点です。表中、エキスパートボードとあるのは、評価に最終的な責任を負うことの出来る、意思決定者により近い存在の有識者の集まりです。

エキスパートによるレビューでは、パネル構成やレビューアの選出方法がマネジメントの重要な点になります。これについては、次の章で詳細に見ていくことにします。

1.5.3.2 パネル運営

レビューアを一堂に会し、レビューアからなる委員会（パネル）を組織し、民主的なプロセスに従って評価を行う方法をパネル評価と言います。

評価パネル方式では、異なる専門家の意見を聞くことが前提ですので、もっぱら学際的研究か、新領域の研究、もしくはミッション型の研究開発が評価対象となります。つまり、評価パネル方式では少なくともエキスパートが参画することが前提となります。

ここでは評価パネルの運営の原則を学びましょう。以下の図 1-29 の流れでは、次のようなアクターが登場します。

評価委員会は、評価パネル（パネルリーダーを含む）と、評価事務局、そして評価され

る側である、被評価者（研究開発実施者ならびに事業推進者）から構成されます。

評価パネルは評価事務局から示される評価項目・評価基準に従ってプロジェクトの評価を行います。そのために必要な情報やデータ、場合によってはヒアリングや現地見学等から得られた情報を下に評価を行います。評価事務局は評価パネルの選定、委員会招集のロジスティクス、評価項目・評価基準の作成を行います。また、被評価者に対しては評価方針の事前説明や、評価結果の通知を行います。被評価者は、評価に先立ち、事業の成果報告書を提出し、さらに評価が公正に行われるように評価委員会に対して評価に必要なデータの提供を行ったり、評価理由に異議がある場合の申し立てを行うことができます。

出所：財団法人政策科学研究所作成資料

図 1-29　評価委員会のアクター

(1)レビューアの選定

評価パネルのレビューアを選定します。評価対象の科学技術領域や社会経済領域にしたがって、学識経験者、民間の技術者、その他のステークホルダーを選定します。

その他のステークホルダーとしては、より広い見知から意見のできる科学ジャーナリストや技術開発であれば関連技術領域の企業人等が相当します。

この中で、専門領域を別にして、評価経験の有無からおおよそのレビューアの構成を考えます。例えば、７人のレビューアを選定するとしたら、５人程度は評価経験がある人材、２人程度は初めて評価を担当する人材というように配置します。

これには、レビューア自身も評価経験を積むことで、より有効なレビューを行うことができるという理由と、レビューアのパイを増やして、研究開発評価のレビュー人材の厚みを増す必要があるからです。

新人のレビューアは、過去の評価対象から優秀な評価を受けた研究者等や、他の有識者からの紹介というかたちで発掘することができます。

(2)パネルリーダーの選出

評価事務局としては、選定したレビューアの中から、評価経験の豊富で、かつ、幅広い知見を有した人材（エキスパート）に対して予めパネルリーダーの任を担ってもらうよう、依頼をする必要があります。そして、第１回目の評価委員会において、他のレビューア全員から承認を受けるかたちでパネルリーダーを正式に選出することが望ましいでしょう。

パネルリーダーはさまざまな意見に耳を傾け、議論の方向性が評価の観点とずれているようであればこれを修正するなどのファシリテーターの役割があります。また、最終的な評価がレビューアの中で割れてしまった場合には、パネルリーダーの権限で評点を決定するなど、他のレビューアに対して、評価パネルをコントロールする立場にもあります。

したがって、民主的な手続きを踏まえて選出されることが原則として必要です。

(3)評価方針等の説明

第１回目の評価委員会において評価事務局からレビューアに対し、どのような評価を行って欲しいのかについてのレクチャーをする必要があります。ここで、評価事務局として評価方針の説明、評価対象に関する評価項目・評価基準を記した評価シートの説明を行います。

評価シートに対する委員からの疑問、修正要望等があれば、この段階で修正し、評価委員会において共通の理解を得る必要があります。

(4)評価対象の説明と情報・データの確認

第１回目以降の評価委員会で、評価対象の説明を被評価者が行います。研究開発事業であれば、事業推進者が施策担当者の立場から説明をし、研究開発実施者が活動の内容、成果の内容についてデータを踏まえ説明します。

評価パネルは評価項目や評価基準に照らして、判断に必要なデータに不足があれば、被評価者にデータや情報の提供を求めます。また、アウトカムに関する指標や分析結果が必要であれば、評価事務局に対して調査・分析の要請を行います。

研究開発事業によってはオンサイトの評価を実施する必要があります。拠点形成や実験設備、大型の研究開発基盤に関わる評価であれば、現地の見学やヒアリングが不可欠でしょう。

なお、評価パネルは評価委員会で得た情報・データ等を他者に教えてはならないなどのルールが必要な場合があります。特に、特許などの知的財産に絡むような成果を扱う場合には、あらかじめ、そのようなルールを設定し、レビューアに誓約してもらうことが重要です。

(5)審議

評価委員会を何回か経て、評価パネルによる最終的な審議を行います。審議の段階では、被評価者は退席し、委員の間で意見交換をした後に評点付けなどの作業に移ります。パネルリーダーが主導的な役割を果たすのはこの段階です。何回かの意見交換の後、やはり評点が付けられないようであれば、時間的な制約もあるため、パネルリーダーに一任するなどの措置が採られる場合もあります。

(6)評価報告書の作成

評価事務局は評価結果である評点の集計と、コメントの集約作業を行います。通常、評価報告書案を作成した後、被評価者と評価パネル双方に対して報告書案を開示します。被評価者は評価報告書に記載された事実関係に間違いがある場合、その内容によって評価結果に影響をもたらしていると考えられる場合には、評価パネルに対して異議申し立ての機会を与えられて然るべきでしょう。評価パネルも評価報告書に記載されたコメントが誤って解釈されていないかどうかをチェックし、双方合意の上で承認され、評価報告書が完成します。

(7)評価報告書の公表

評価報告書案が承認された後、Web 等で評価報告書として公開します。評価報告書には、評価パネルの氏名や肩書きが明記されているので、評価事務局のみならず、評価パネルも評価結果に対して責任を負う立場にあります。

2　従属型プロジェクトの評価

本章からは評価対象毎に評価論のポイントを解説していきます。各評価対象に共通の問題は前章で既に解説しましたので、評価対象に固有の問題に焦点を当てることになります。まずは、評価対象としてもっとも案件の多い、従属型プロジェクトの評価について学んでいきましょう。

2.1　従属型プロジェクトのマネジメント

従属型プロジェクトとは「ある一定の枠組み（＝プログラム）の下で展開される個別の研究開発課題のことです。具体的には、競争的研究資金による研究開発制度の個別課題、府省・資金配分機関等が実施する提案公募型事業の個別プロジェクトが相当します。

従属型プロジェクトの特徴は、プロジェクトの研究開発目的、研究開発領域・分野等の基本的な方向性が、上位の枠組みである制度によって規定されていることです。したがって、プロジェクトの評価の観点（必要性、有効性、効率性）は制度の目的に合致するかどうかで判断されます。

評価の枠組みが上位の施策によって規定されている場合、個別プロジェクトの評価では研究開発方法、研究開発成果の質などの内容的側面をいかに充実させるかが中心的課題になります。そのためには評価だけではなく、プロジェクトの採択から最終的な評価までを一連のマネジメントとして捉え、より良い成果が生み出されるように工夫をすることが重要になります。

出所：財団法人政策科学研究所作成資料

図 2-1　従属型プロジェクトのライフサイクル

図 2-1 は従属型プロジェクトのライフサイクルを描いたものです。

通常、制度全体の応募規定や公募要領に従い、公募を開始します。研究開発実施者は申請書と研究計画書を作成し、公募期間内に提出します。

制度実施側は申請書・計画書を受理した後、審査に移ります。ここで、採択評価のためのマネジメントが発生します。レビュー方式、レビューアの選定が行われ、レビューが行われます。一方、制度実施側は資格審査や倫理規定に違反していないかどうかの判断をし、レビュー結果と付き合わせて最終的な採択の判断を行います。

採択・助成が開始されたら、各種契約手続きが始まります。近年では、競争的研究資金の受給に伴い、研究者の所属機関に一定のオーバーヘッドを認めていますので、契約の対象は所属機関と結ぶことが一般的となってきました。不正使用問題などの教訓から、今後は所属機関側の管理体制がより一層、問われることになるでしょう。

研究開発がスタートしたら、モニタリングを実施する必要があります。中間報告会などでアドバイザリーボードとの意見交換をするなどの機会を設けても良いでしょう。

研究開発が終了したら、研究開発実施者は成果報告書を提出し、これに基づいて課題の終了評価が行われます。採択評価と同様、レビューアを配置し、評価結果を導くためのマネジメントが必要となります。

このような従属型プロジェクトのライフサイクルを踏まえ、以下では、各段階でのマネジメントの工夫、レビュー評価に関する理解を深めることとします。

2.1.1　公募実施のための準備

従属型プロジェクトの公募に際して、制度実施側が最低限行わなければならないことは、次のような点です。

- 公募要領の作成
- 採択評価の評価項目・評価基準の作成
- 公募方法・時期等の検討
- 採択評価の方法
- 採択結果の通知

このうち、公募要領の作成については、大筋のものは制度（プログラム）設計の段階で行われています。プログラムにおいて研究開発領域や課題設定が行われている場合には、それに対応した詳細な公募要領を作成する必要があるでしょう。

採択評価の評価項目・評価基準は、研究開発領域毎、設定されているテーマ領域毎に作成される必要があります。これらの評価項目・評価基準については、公募要領等に記載して、申請者にも十分に通知しなければなりません。

公募方法・時期等については、もっとも工夫の余地がある部分です。現状では、一部の資金配分機関において年複数回公募などの取り組みも行われていますが、さらなる工夫が考えられます。例えば、欧米の資金配分機関で行われているような２段階公募方式（Ａ４２枚程度のプレプロポーザルを提出・審査した後に、審査に通った申請者のみがフルプロポーザルを提出できる仕組み）や、通年で申請を受け付ける仕組み等があります。研究者の申請作業と制度実施側、ならびにレビューアの負荷を軽減しつつ、より良い研究課題を募集する仕組みについては、海外の資金配分機関に学ぶべき点が多いように思われます。

採択評価の方法はレビュー評価法が採られます。レビュー評価では適切な仕組みとマネジメントの構築が不可欠です。従属型プロジェクトにおいてレビュー・システム、レビュー・マネジメントは中心的な要素になりますので、節を改めて詳細に見ていくことにします。

採択結果の通知については、採択が決まった申請者に対してよりも、不採択となった申請者に対する配慮が必要です。我が国では、科学研究費補助金の採択結果の通知システムが先進的な事例と言えます。

コラム：科学研究費補助金（基盤研究等）の審査

科学研究費補助金（基盤研究等）の審査では、ピアレビューによる２段階審査が実施されています。審査期間中は審査委員が誰であるかを公表しないブラインドレビュー方式を採っています。

第１段審査の評定要素は、①研究課題の学術的重要性・妥当性、②研究計画・方法の妥当性、③研究課題の独創性及び革新性、④研究課題の波及効果及び普遍性、⑤研究遂行能力及び研究環境の適切性の 5 つが設定されています。これにその他の評価項目を加え、総合評点を付けるようになっています。各要素における評定基準は以下の通りで、それぞれに評点区分が設定されており、4 段階で評点を付します。

①研究課題の学術的重要性・妥当性

・学術的に見て、推進すべき重要な研究課題であるか。

・研究構想や研究目的が具体的かつ明確に示されているか。

・応募額の規模に見合った研究上の意義が認められるか。

②研究計画・方法の妥当性

・研究目的を達成するため、研究計画は十分練られた物になっているか。

・研究計画は遂行する上で、予期される問題点に付する配慮、問題が生じた時の対応などが検討されているか。

・研究期間や経費配分は妥当なものか。

・公募の対象としていないような研究計画に該当しないか。

③研究課題の独創性及び革新性

・研究対象、研究手法やもたらされる研究成果等について、独創性や革新性が認められるか。

④研究課題の波及効果及び普遍性

・当該研究分野もしくは関連分野の進展に対する大きな貢献、新しい学問分野の開拓等、学術的な波及効果が期待できるか。

・科学技術、産業、文化など、幅広い意味で社会に与えるインパクト・貢献が期待できるか。

⑤研究遂行能力及び研究環境の適切性

・従来受けた研究費での研究経過・研究成果を評価するとともに、これまでの研究業績等から見て、研究課題に対する高い遂行能力を有していると判断できるか。

・複数の研究者で研究組織を構成する研究課題にあっては、組織全体としての研究遂行能力は十分高いか、また各研究分担者は十分大きな役割を果たすと期待されるか。

・研究課題の遂行に必要な研究施設・設備等、研究環境は整っているか。

第 2 段審査では、第 1 段審査委員の審査結果をもとに、「総合評点」に重点を置きつつ、各要素の評定結果およびコメント等も勘案してより広い立場から総合的な審査が行われます。「総合評点」および上記の 5 つのの評定要素については、各第 1 段審査委員の評点を研究種目（審査区分）ごとに平均点と標準偏差をそろえるように電算処理が行われ（T スコア化）、各審査委員の素点と T スコア化した評点が合議審査（第 2 段審査）の資料として提出されます。また、その他の評価項目については、各第 1 段審査委員の評定結果がそのまま合議審査（第 2 段審査）の資料として提出されます。合議審査を経て、採択・不採択が決定されます。

応募者が応募時に審査結果の開示を希望すれば、応募した細目における不採択課題のうちのおおよその順位や評定要素毎の平均点等が通知されます。この開示情報により順位やどの評価項目が低かったかについて事後的にチェックができますので、次の応募時の反省材料にすることができます。

なお、審査終了後に、審査委員の氏名等は独立行政法人日本学術振興会（JSPS）のホームページ及び印刷物で公開されます。

2.2　レビュー評価

本節ではレビュー評価の全体像を概観します。最適なレビュー評価というものはなく、実施するプログラムの目的・性格によって、さまざまな仕組みが考案されています。しかしながら、単に仕組みだけ作れば良いというものではなく、実務者のマネジメント努力が発揮されてはじめて有効なメカニズムとして機能します。

ここでは、レビューアをどのように選定するか、レビュー・システムをどのように構築するか、レビュー・マネジメントとはどのようなものかについて解説します。

2.2.1　レビューア

レビューアは専門家の立場からプロジェクトの内容について審査する評価人材です。ある特定のディシプリンの内部で評価する場合にはピアレビュー、複数の領域に跨って、あるいは新領域を評価する場合にはエキスパートレビュー方式が採用されます。

レビューアは組織や国から独立した存在であることが要請されます。つまり、個人の能力・資格によって評価に対して責任を全うするものであって、組織や国の代理人であってはなりません。非常にパブリックな存在であると言えます。

2.2.1.1　レビューアの確保

レビューアは基本的にその国の研究者のストックに依存しますが、基礎研究や応用研究等の研究プログラムであれば、国外の研究者をレビューアとしてストックすることも可能です。例えば、EU が実施しているイノベーション促進研究開発プログラム（Framework Programme：FP）では欧州域内 15 ヶ国の研究者をレビューアの対象としています。

レビューアの確保は、海外の資金配分機関でも悩みの種となっています。資金配分機関のプログラムマネージャーは自分の担当する科学技術領域の学会誌に目を通し、国際学会や関連学会に頻繁に出席し、研究者の中からレビューアとして有望な人材をリストアップすることに余念がありません。例えば、米国国立科学財団（National Science Foundation：NSF）は毎年約 42,000 件の研究提案を受け付け、そのうち、審査によって高い評価を受けた約 11,000 件に対し研究資金を提供していますが、このような大量の申請を捌くためには常時 50,000 人の評価者を動員しています（１つの申請提案に対して、最低３人のレビューアが審査に関わります）。このうち、毎年 10,000 人程度が申請審査を初めて経験するよう、リストの更新が継続的に図られています。

レビューアをストックする機能を資金配分機関の外に配置している事例もあります。ドイツ学術振興会（DFG）では、ピアレビューアの選定は DFG のプログラムマネージャーが行いますが、ピアレビューアのリスト化や能力保証についてはドイツ学術アカデミー連合などのサイエンスコミュニティが担っています。

研究者の立場からレビューアを引き受けるとはどういうことでしょうか。先ほどの NSF の例ですと、多い人で年間 30 本程度のプレプロポーザルに目を通し、20 本近いフルプロポーザルを読み、かつパネル評価にも参加します。フルプロポーザルのレベルになると、学術論文を読むのと変わらないほどの労力を要します。

しかしながら、多くの研究者は NSF のレビュー依頼に対して、よほどの場合がない限り NO とは言いません。科学技術の振興、特に若手研究者育成のためという社会貢献の一環として評価への参画が社会的に認知されているからです。また、レビューができるということは研究者として一流であることの証でもあります。したがって必要な経費と若干の謝礼が支払われるとしても、多くの研究者は基本的にはボランタリーでレビューアを引き受けています。

我が国の場合ですと、予算規模がもっとも大きいプログラムである科学研究費補助金（基盤研究等）の審査委員はデータベースに登録されている人材としては 40,000 人ですが、第 1 段審査委員は 4,100 人です。一方、科研費の応募件数は約 10 万件であり、研究種目毎に１課題あたり 3 人又は 6 人が審査を行い、1 人当たり平均 100 件の審査を担当していることになります（多い人で 200 件を超えることもあります)。

この数字は研究者のボランタリーのレベルを超える負荷として、関係者の間で認識されています。科研費制度を持続的に発展させるためにも、ピアレビューの質と量の確保は急務であると言えます。

2.2.1.2　レビューアの選定

日本の科研費や NSF の研究･教育プログラムでは審査委員が誰であるか分からないよう、匿名性を保持したブラインドレビューと呼ばれる審査方式が採られます。審査委員は研究領域毎にリスト化され、一定の申請数を割り当てられるように決めていきます。

一方で、申請者が自ら審査してもらいたい研究者を１名選択できるようなシステムもあります。英国工学・自然科学研究会議（EPSRC）の産学共同研究プログラムでは、申請者が利害関係のない範囲で審査委員（レフリーと呼ばれます）を１名指名することが義務づけられています（その他に、EPSRC のプログラムマネージャーが大学、研究機関等から３名のレフリーを選出します)。これは、先端的な研究領域という特徴もあり、申請者が自分の研究内容について、最先端の研究者に興味を持ってもらうための措置でもあります。

レビューアは通常、１課題につき３～４名程度が担当になります。論文の査読審査と同じ体制と考えてよいでしょう。各レビューアはプログラム実施側が用意する評価項目・評価基準に基づき、採択評価を行います。

2.2.2　レビュー・システム

レビュー・システムとは、プロポーザルに対するレビュー評価を採択審査の中でどのよ

うに位置づけ、結果を活用するかという仕組みのことです。

ここでは、欧州委員会（EC）の実施しているイノベーション促進プログラムであるフレームワークプログラム（FP6)を例に、レビュー・システムを解説します。

図 2-2 はプロジェクトの提案（プロポーザル）の受理から採択（融資の決定）に至るまでの手続きをフローチャートにしたものです。

プロポーザルが受理されると、まず資格審査に移ります。ここで FP6 の資格要件を満たしていないものは自動的に不採択になります。

資格審査を通りますと、サブプログラム（研究開発領域）毎のエキスパートが個人的に審査をします。審査はダブル・ブラインドレビュー方式（採択者・申請者双方が匿名）で行われます。FP6 のエキスパートは欧州域内 15 ヶ国に点在しているため、メールレビューが基本となります。FP6 では 2006 年の公募でおよそ 3,400 名のエキスパートが審査に関わっています。

出所：<http://www.cordis.lu/fp6/>より財団法人政策科学研究所作成

図 2-2　欧州フレームワークプログラム（FP6）のレビュー・システム

表 2-1　FP6 のレビューアの数

プログラム分野	エクスパートの人数
1. Life sciences, genomics and biotechnology for health	603
2. Information society technologies	349
3. Nanotechnologies and nanosciences, knowledge-based multifunctional materials and new production processes and devices	31
4. Aeronautics and space	
5. Food quality and safety	191
6. Sustainable development, global change and ecosystems	400
7. Citizens and governance in a knowledge-based society	66
8. Policy support and anticipating scientific and technological needs	326
9. Horizontal research activities involving SMEs, research and innovation	※1
10. Specific measures in support of international cooperation	77
11. Support for the coordination of activities	※1
12. Support for the coherent development of policies	※2
13. Human resources and mobility, research infrastructures	1,149
14. Science and society	158
15. Euratom	33
Total	**3,383**

※1 2006年の応募実績なし
※2 ″7. Citizens and governance in knowledge-based society″と共通

出所：<http://cordis.europa.eu/fp6/experts2006.htm>より財団法人政策科学研究所作成

個別審査が終了すると、同一課題に対して選出されているレビューア間でのパネル審査に移ります。ここで、一次採択としての評価結果をプログラムの実行委員会に提出します。

次に、オプション的な扱いになりますが、申請プロジェクトによってはヒアリング（オンサイトもしくはオフサイト）が実施され、一次審査のレビューアの他に数名のエキスパートを加えた評価パネルが構成されます。評価パネルはヒアリングの結果も踏まえ、最終的な評価点をプログラム実行委員会に告知します。

プログラム実行委員会は評価パネルが導いた各プロジェクトの評点をランキングし、採択候補を絞り込みます。その後、契約手続きに移る段階で、申請者との細かい交渉が始まります。

また、FP6 ではプログラムにおいて各種倫理規定が設けられています。プロジェクトの内容が倫理規定に抵触しないかどうかは採択評価のさまざまな段階で確認されます。

交渉の結果、プログラム実行委員会は申請者に対して、研究開発体制やプロジェクト管理体制、予算、倫理規定や支払規定に関するコンサルテーションを実施し、合意が得られ

れば、融資が決定されます。すなわち、申請は最終的に採択されることになります。

このように、採択評価の段階で、幾重にもエキスパートがレビューを行い、場合によってはヒアリングや現地見学等を実施するなどの工夫を重ねて、最終的な評価が行われます。ここまでの労力を採択評価に費やしているからこそ、その後の評価のプロセスがより合理的に実施されることにつながっています。

例えば、FP6 ではプロジェクトに関して中間評価のような制度化された評価は存在しません。FP6 のプロジェクトは３～５年の研究開発期間が与えられますが、各プロジェクトの実施者は、オープンな年次報告会等を開催し、広く意見聴取を行うなど、独自の活動によって自らの研究開発の進捗状況や研究開発の方向性を確認する機会を積極的に作っています。最終年度にはファイナルレポートを提出しますが、直後の事後評価も制度化されておらず、本格的な評価はプログラムレベルでの事後評価（終了後２年以内に実施）か、追跡評価（終了後５年以上の期間を置く）に委ねられます（ただし、財務状況の確認等、定型書式による年次報告は義務づけられています)。

2.2.3　レビュー・マネジメント

レビュー・システムを円滑に機能させるためには、さまざまな運営上の工夫が必要です。この工夫をレビュー・マネジメントと呼ぶことにしましょう。

レビュー・マネジメントはレビュー・システムの各段階で必要となります。ここでは、Kostoff (1997)を参考に、ピアレビューの質的向上のための８項目とそれに関連した留意事項を紹介したいと思います。

2.2.3.1　レビューアの所属機関に対する配慮

レビューアは大学や研究開発機関に所属する研究者であることがほとんどです。レビュー作業は、研究者に対して一定の労力を強いるため、プログラムマネージャーがレビューアの所属する機関の上司に対して明確なコミットメントをしなければなりません。

コミットメントの内容は、レビューアに評価を依頼するプロポーザルの数、評価内容、委嘱期間、パネルとしての参加、パネルの開催回数・時間、交通費等の支給の有無など、通常の委員委嘱と同様の手続きが明記されるべきでしょう。

また、レビュー作業が公益に資するものとしての認識をもってもらい、所属組織における研究者の評価を高めるよう、配慮しなければなりません。大学や独立行政法人のように交付金で運営されている機関に対しては、所属の研究者が国の実施する研究開発においてレビュー機能を提供しているという事実は、機関評価の重要な側面として認知されるべき点でもあります。

Kostoff はレビュー・マネジメントのもっとも重要な手続きとして、この項目を一番に挙げています。

2.2.3.2　評価パネルの運営

評価パネルを設定した時に、評価実施側はパネルを上手にファシリテーションする必要があります。ピアパネルのレベルでは、１課題につき３人程度ですので、パネルリーダーを選出し、運営を任せるよりも、プログラムマネージャーが直接ファシリテーションした方が効率的であるとも言えます。

これに先立ち、プログラムマネージャーはレビューアを選出する段階で、パネル評価の結果が偏らないように恣意的でない方法でレビューアを選ばなくてはなりません。

また、事後評価における評価委員会のような場においても、委員の質疑応答を先導し、彼らのコメントを要約し、評価レポートを作成するのは、ひとえにプログラムマネージャーの力量に掛かっていると言っても過言ではないでしょう。

レビュー・システムが上手く機能するかどうかは、裏方であるプログラムマネージャーのモチベーションと能力に掛かっていると Kostoff は指摘しています。

2.2.3.3　レビューアの能力と客観性

第３に重要な要素として、Kostoff はレビューアの能力と客観性を挙げています。ピアレビューアといえども、担当分野の専門的能力だけではなく、研究対象の複合的な側面、例えば隣接研究分野、研究により潜在的に影響を受ける技術・システム、プログラムのミッション等に対する理解がなければ、評価の役に立つことは出来ません。

一方で、プログラムマネージャーもレビューアに対して、評価内容を専門分野に限定するのではなく、プログラムの目的や評価対象となるプロジェクトの最終的な目的にまで拡張して評価してもらうように配慮しなければなりません。このことによって、レビューアは評価経験を積み、革新的な提案についても公正な評価を行えるだけの潜在的能力を身に付けることができます。

2.2.3.4　パネル間、分野間の正規化

レビュー・システムの採択判断に近い段階では、複数のテーマやプロジェクトを比較検討し、ランキングによって採択評価を行うなどの措置が採られます。そのため、評価パネル間、もしくは分野間の正規化を図るようにレビューアの構成を設計する必要があります。

ある程度類似性を持つ分野については、幅広いバックグラウンドを持つレビューアを共通に配置することで評価の正規化が図れます。極めて異なる分野については、プログラム実施機関の当該分野に関する戦略的価値を反映し、採点の厳しさや偏向性の補正が必要になります。その場合においても、レビューアのうち１名は幅広いバックグラウンドを持つエキスパートを配置し、評価内容の正規化を図るように努めるべきと、Kostoff は指摘しています。

2.2.3.5　評価項目としての総合評価

レビュー評価を行う場合、設定された各評価項目の他に総合評価の項目を設けているケースが多く見られます。なぜ各評価項目の評点の和や平均点を用いず、別途、総合評価の評点を付けなければならないのでしょうか。

第１に、各評価項目に対するウェイトがレビューアによって異なるという点が指摘できます。これについては、より適切な評定区分の表現を工夫することで、ある程度の問題の解決にはなります。

第２に、より重要なことですが、プロポーザルを評価するための評価項目以外に、レビューアが重要と考える要素が必ず存在するからです。例えば、当該プロポーザルが国の研究開発よりも産業内の研究開発課題として相応しいと評価者によって判断された場合には、技術的な評価項目の評点が高くても、総合評価の評点が低くなる可能性があります。

このような可能性がある限り、個別の評価項目・評価基準以外に総合評価の欄を設けることは有用と考えられます。

2.2.3.6　評価の匿名性

レビュー評価を行う際に重要な要素として、評価の匿名性が挙げられます。通常、レビューアについては採択審査の期間中は匿名性が保持されます。一方、申請者のプロポーザルを匿名とすべきかどうかについては、プログラムの特徴によって判断されるべきでしょう。

例えば、学術研究のために広く助成されるようなプログラムでは、ダブル・ブラインドレビュー（申請者も匿名）が推奨されます。性別、国籍、年齢、所属機関名などに対する偏見を排除し、純粋にプロポーザルの中身について審査する方が学術研究の機会を提供するという面では大きなメリットになります。

一方で、応用研究や実用化研究などのプログラムの場合には、レビューアが評価項目として重視するのは、研究開発の実施体制や実績です。このようなプログラムでは、申請者を匿名とするメリットはあまりないでしょう。

2.2.3.7　ピアレビューのコスト

レビュー・システムは研究者がボランタリーベースで参加する仕組みとはいえ、多大な時間コストの上に成り立っているシステムです。研究者がレビューに費やす時間は、１申請あたり１時間だとしても、割り当てられるプロポーザルが多ければ、とてつもない負荷になります。

また、申請する研究者もプロポーザルを作成するには、その何十倍もの時間を掛けていることでしょう。我が国の申請メカニズムのように、最初から各種申請書類と同時にフルプロポーザルを提出しなければならない場合には、研究者は１回の公募に全身全霊を掛けて望まなければなりません。たとえ他に申請できるプログラムの機会があったとしても、

競争的研究資金制度においては重複申請が禁止されているために、審査結果が出るまでは申請することができません。完全なアイドル状態に陥るわけです。

そのため、公募型の研究開発プログラムが発達している諸外国では、２段階公募方式を採用している資金配分機関が少なくありません。１段階目では各種申請書類と簡単なプレプロポーザルを提出し、レビューアはプレプロポーザルの中から有望な研究開発課題をスクリーニングします。プレプロポーザルはせいぜい、A4 で１～２枚程度に収まっていますので、割り当てられたプロポーザルの件数が多くても、時間コストを削減することができます。

一方、申請者の側も、プレプロポーザルの作成にはそれほど時間を掛けずともよく、さらに一次審査の結果が下りるまでは他のプログラムへの申請も同時並行的に可能です。フルプロポーザルを書くにあたっても、一次審査のレビューアからのコメントで、どこを詳しく書いて欲しいかの指示を受けていますので、より具体的なプロポーザルの作成が可能となっています。

こうしたレビュー・システムが成立するためには、十分な人員とマネジメントが不可欠といえます。

2.2.3.8　評価の倫理基準

より良く設計されたレビュー・システム、配慮の行き届いたレビュー・マネジメントがあったとしても、最後に問題となるのはレビューアの倫理観です。

人が人を評価するということは、ある種の特権が与えられるということです。評価のプロセスにおいてはさまざまな利害の衝突が想定されます。またレビューアがその気になれば、秘密情報の流出、違法行為などの可能性がないとはいいきれません。

どのような文書にレビューアが署名したところで、最終的な拠り所になるのは、真の評価者たらんとする研究者の美意識・モラルです。それを支えているのは研究者コミュニティの自己規制しかありません。

レビューアが審査の段階ではなはだしい倫理規定の逸脱に及んだ場合、レビューアとしての資格取り消しや当該プログラムの申請停止処分といった対処だけではなく、学協会における会員資格の剥奪や職場での減棒等、厳しい対応が必要と考えられます。

また、それを行うに判断の拠り所となる倫理規定と罰則をあらかじめ作成しておくことが重要になります。

2.2.3.9　その他の留意事項

（1）パネリストの入れ替え

評価パネルの委員候補が固定化してしまうと、既得権益が生まれたり、評価組織が硬直化するなどの弊害が予想されます。一方で、すべての委員を毎回入れ替えてしまうと、評価の継続性がなくなり、レビューアの学習効果が発揮されなくなってしまいます。したがっ

て、一般的にはパネリストの一部を毎年更新することが望ましいマネジメントになります。例えば、評価委員の任期が４年であれば、毎年４分の１を入れ替えるなどの工夫が考えられます。

(2)利益相反(利害関係者の排除)

評価者がプロポーザルに対して利害関係を持つ場合、プロポーザルの審査から外れるよう配慮しなければなりません。しかしながら何をもって利害関係と認めるべきでしょうか。

科学研究費補助金（基盤研究等）では次のように利害関係を定義し、審査委員自らが該当すると判断した場合には当該研究課題の評価に加わらないこととしています。

表 2-2 科研費の利益相反の取扱

審査に関する利害関係の排除（利益相反）については、下記のとおり取扱うこととする。

(1) 審査委員自身が、研究課題の研究代表者又は研究分担者である場合は、審査に加わらないこととする。

(2) 審査委員が、研究課題の研究代表者又は研究分担者との関係において、次に掲げるものに該当すると自ら判断する場合は、審査に加わらないこととする。

① 親族関係もしくはそれと同等の親密な個人的関係

② 緊密な共同研究を行う関係
（例えば、共同プロジェクトの遂行、共著研究論文の執筆、同一目的の研究会メンバーにおいて、緊密な関係にある者）

③ 同一研究単位での所属関係（同一講座の研究者等）

④ 密接な師弟関係もしくは直接的な雇用関係

⑤ 研究課題の採否が審査委員の直接的な利益につながると見なされるおそれのある対立的な関係もしくは競争関係

（「科学研究費補助金（基盤研究等）の審査方針」Ⅰ．５）

出所：独立行政法人日本学術振興会「平成 18 年度科学研究費補助金第１段審査の手引」より抜粋

重要なのは（２）の規定です。審査員がこの部分の規定に抵触するかしないかということが直接的な問題なのではなく、審査員の倫理規範上、審査に関わるのが望ましくないと判断するかしないかが問われています。例えば①について言えば、親族関係と同等の親密な個人関係とは、具体的にどのような関係なのかは人によって様々な解釈があり得るでしょう。しかしながら、親密な個人関係にあるから甘い評価をするということでないかぎり、本来は問題ないことです。

Kostoff が評価は倫理の問題であるといったのは、このように評価のメカニズムが個人の価値判断に依って立つ部分が大きいからです。

(3)守秘義務

利益相反とならんで、評価者に課せられる倫理規定の一つが守秘義務です。採択評価の進捗状況や、成果物に関する情報等の漏洩は重大な守秘義務違反の例です。

例えば、アウトプットに関する守秘義務が守られないと、研究開発実施者側はその後評価に必要なデータをパネルに提出しなくなることが予想されます。企業の研究開発が絡む場合には、情報漏洩は重大な経済的損失を生む恐れがあるため、単にレビューアが罰則を適用されたというだけでは済まない問題となるでしょう。

形式的には守秘義務に関する誓約書（罰則等を含む）にサインをしてもらうという手続きが必要ですが、重要なことは評価者の倫理規定として、科学技術コミュニティに広く認知してもらうということに尽きます。

(4)オブザーバー

評価パネルにおいて、評価者、被評価者でもなく、事務局の側に立つ者でもない外部的な存在として機能する人材にオブザーバーがいます。オブザーバーの役割は、評価セッションの進行の仕方、公正性、公平性、手続きを改善する方法や、評価項目・評価基準に対して独立した意見を述べることにあります。オブザーバーはプロポーザルや成果報告書に対する意見を述べるのではなく、評価という手続きが公正かどうかを見る者です。

評価者も被評価者も、また評価事務局もオブザーバーの存在を意識することで、評価セッションを客観的に運営できることになります。ある程度規模の大きい事業については、評価委員会を公開し、オブザーバーを配置するなどの工夫が必要となります。

2.3　レビュー評価の手続き

2.3.1　レビュー・システムの設計

最初に、自らが実施する従属型プロジェクトのレビュー・システムを設計する必要があります。プロジェクトのプロポーザル提出から採択までのフローの中で、どのようにピアレビュー等を位置づけ、意思決定に役立てるかについて明らかにする必要があります。

前節の FP6 の採択評価のフロー図 2-2 を参考に、レビュー・システムのフロー図 2-3 を作成しましょう。基本的な枠組みを先に作成し、オリジナルな試みについて加えていくことが頭の整理に役立ちます。

出所：財団法人政策科学研究所作成資料

図 2-3　採択評価のフローの基本形

上の図 2-3 にさらに必要な情報は、誰がどのような役割を果たすのかという責任分担に関する情報です。採択評価の責任分担を考えながら、実施体制を記載しましょう。

出所：財団法人政策科学研究所作成資料

図 2-4 採択評価の手続きと役割分担

こうして作成したレビュー・システムを公募要領等に掲載し、申請者やレビューア等への採択審査のプロセスに対する理解を促す説明資料として役立てます。

2.3.2 レビューア・データベースの作成

評価に協力してくれるレビューアをどれだけストックしているかが、資金配分機関の資産といっても過言ではないでしょう。レビューアに関する情報はデータベースの形式で情報を蓄積し、資金配分機関内で共有することが一般的です。

データベースはフィールド（データ項目）の設計に情報収集の仕方、情報の活用の仕方がすべて反映されます。プログラムマネージャーはレビューアに関する必要な情報がすぐに取り出せるよう、メンバー間で十分に協議した上でデータベースを設計しなければなりません。

下の表 2-3 では、専門分野等に関するデータとして、メニュー形式での入力を設定しています。資金配分機関がどのような専門分野のレビューアを欲しているかという観点から、活用のしやすい分野・領域の区分を設定し、レビューア候補者を分類していくことになります。

表 2-3 レビューア・データベースのフィールドの設計例

フィールド名	データ種別
氏名	テキスト
生年月日	日付データ
性別	2値データ
所属	テキスト
現住所	テキスト
Tel & Fax	リンク
e-mail	リンク
略歴	テキスト
主たる専門分野	メニュー形式
扱うことのできる隣接分野	メニュー形式
当機構における評価経験	テキスト
当機構のプログラム申請経験	テキスト
委嘱状等送付先住所・宛名等	テキスト
委員委嘱中	2値データ
備考欄	テキスト
謝金ランク	メニュー形式
謝金等先支払口座	テキスト
源泉徴収書等送付先住所	テキスト
支払履歴	日付データ
最終更新日	日付データ
最終更新者	テキスト

出所：財団法人政策科学研究所作成資料

作成したデータベースからピアレビューやレビューパネル等の審査員候補として抽出します。データベースはメンテナンスが重要になりますが、特に審査経験等のデータを反映することで、レビュー人材の拡充を資金配分機関が定型的な業務の一環として図っていくことが可能になります。

例えば、毎年レビュー審査に関わる人材の 1 割を、レビュー経験のない新しい人材でまかなうという基本方針を固めることで、どの分野のレビュー人材がどれだけ必要となるかという具体的な情報収集戦略が立てられます。こうすることで特定のレビューアに審査業務が集中しないよう、マネジメントができるようになるだけでなく、レビュー人材の層を厚くすることにも貢献します。

出所：財団法人政策科学研究所作成資料

図 2-5 レビューア・データベースの活用と更新

2.3.3 レビュー評価指針の作成

レビュー人材に審査を依頼する際に重要なことは、どのような審査をして欲しいかということを具体的に示した文書を作成することです。

レビュー評価指針はおおよそ次のような構成を持ちます。

1) プロジェクトが属するプログラム，サブプログラムの目的

プログラムの目的に則したプロジェクトの採択についての理解を促します。

2) 審査スケジュール

レビュー評価の期日を定めます。

3) 利害関係者規定の説明と該当した場合の対処

プロポーザルの申請者と利害関係にあると判断できた場合、審査を辞退してもらうことになります。

4) レビュー審査の評価項目・評価基準・評定区分

プロポーザルをどのような観点，基準から評価するべきかを明らかにします。

5) 総合評価点の評定区分

総合的な評価点を付けてもらいます。また、その評価点の理由をコメントのかたちで記載してもらうことが必要です。

6) 倫理規定に関わる方針[3]

プログラムに関係する科学倫理規定の適用を考慮して、倫理規定に抵触する怖れがあると判断される場合にはその旨を記載してもらいます（最終的な判断はプログラムの実行委員会等で決定されます。）

これらの中で我が国のレビュー評価システムの質的向上を図るためには、評定区分の定義をしっかりと示すことがもっとも重要です。

評定区分の悪い例

（１）社会的・経営的ニーズ

５：基礎研究の技術は社会的・経営的ニーズとピッタリ合致している

４：基礎研究の技術は社会的・経営的ニーズと合致している

３：基礎研究の技術は社会的・経営的ニーズとある程度合致している

２：基礎研究の技術は社会的・経営的ニーズから若干外れている

１：基礎研究の技術は社会的・経営的ニーズから外れている

出所：長広仁蔵（1995）「評点法による研究開発の進め方と評価」日刊工業社.

上の評定区分の設定の仕方には２つの問題点があります。まず一つは、「社会的ニーズ」と「経営的ニーズ」という異なる評価の観点を同じ評価基準に同居させて評定区分を定めていることです。二つめには、「ピッタリ合致している」、「合致している」、「ある程度合致している」、「若干外れている」、「外れている」といった表現に対する判断基準が用意されていないことです。

このような評定区分の設定の仕方では、レビューアは「社会的ニーズはあるものの、経営的ニーズを満たしていない」場合や、合致している度合いについて何を基準にすれば良いか分からない場合には評点を付けることができません。

評定区分の設定には次の２つの原則を心掛けるようにして下さい。

[3] 文部科学省であれば、科学技術･学術審議会生命倫理・安全部会の指針「ライフサイエンス分野における生命倫理に関する取組」＜http://www.lifescience.mext.go.jp/bioethics/seimei_rinri.html＞，厚生労働省であれば、「医学研究に関する指針」＜http://www.mhlw.go.jp/general/seido/kousei/i-kenkyu/index.html＞等が最新の倫理規定となる。

- **異なる評価基準を同じ文章に持ち込まない。**
- **優劣に対する判断基準を必ず文章中に記載する。**

2.3.4　レビューアへの委嘱手続き

レビューア候補者は大学や公的研究機関、民間企業の研究所等、組織に所属している研究者がほとんどです。彼らに対して一定以上のエフォートを要求する以上、所属機関に対して正式な委嘱の手続きを踏まえる必要があります。

委嘱状は次のような内容を備えている必要があります。

1) 当該プログラムの審査員として委嘱する研究者への評価

レビューアとして相応しい経験･見識を有した候補者である旨を所属機関の上長に伝えます。

2) 当該プログラムの審査に関わることの公益性に関する説明

所属機関側が研究者をレビューアとして参加させることの社会的意義を説明します。

3) レビューアの負荷に関する情報

レビューアが何本のプロポーザルの審査を､いつまでに実施しなければならないかについての情報をコミットします。

4) その他、謝礼・経費支払い等に関する情報

審査員に対して支払われる謝礼ならびに、必要経費についての情報を明確にします。

レビュー・システムは科学技術政策を振興する上で必要不可欠な仕組みです。審査員としてさまざまな組織の研究者が参加しやすいよう、研究者の所属機関に対して最大限の心配りをする必要があります。

2.3.5　評価パネルの設計

メールレビューなどの評価方法だけでなく、パネル評価を実施する場合には、評価パネルの人数、構成、権限等に関する設計を行う必要があります。

従属型プロジェクトの評価パネルの平均的な人数は 5～7 人程度です。評価パネルの人数は奇数で構成する方が、パネルリーダーの判断の責任を明確にすることにつながります。例えば、採択評価においてメンバー間で票が割れた場合、通常はパネル間での意見調整を行うことで票を集約させられることが考えられますが、結論が出ずにパネルリーダーに一任というかたちで最終的な判断を出す場合も想定できます。奇数で評価パネルを構成しておけば、最後の 1 票がパネルリーダーの責任として明確になります。

評価パネルの構成はプログラムの目的にも依存します。社会経済的なインパクトを目的

に含むプログラムであれば、プロジェクトを評価するレビューアにはエキスパートが含まれていなければなりません。少なくとも、パネルリーダーは視野の広い、複合領域について扱えるエキスパートである必要があります。

また、評価パネルにまったく評価経験のない人材を最低１名は毎回組み入れる必要があります。これは、パネル評価を通じて評価人材を育成・蓄積するためにも必要な措置です。

2.3.6　採択の判定

プロジェクトの採択結果がまとまったら、プログラム運営主体は採択結果を速やかに申請者に知らせる必要があります。採択／不採択の結果は単純に総合評点（スコア）を知らせれば良いというわけではありません。なぜ、採択／不採択となったかの理由を文書にして明らかにするとともに、評価項目・評価基準毎の評点も明らかにすることが望ましい措置となります。

従属型プロジェクトはプログラムが存続する限り、何度でも申請の機会が設けられています。また、一定数の申請者が存在しなければ、本来、競争的資金制度のような研究開発施策は成り立たない制度であると言えます。申請者にプログラムのリピーターとなってもらうためにも、とりわけ不採択の理由・評価結果の詳細情報を明らかにし、より良いプロポーザルを次の機会に申請してもらうことが、結果的にはプログラムの質を高めることになります。

3　プログラム（施策・制度等）の評価

政策評価やマネジメントが対象としているプログラムの定義を Web 等で調べると、二つの意味で用いられていることが分かります。

- A collection of projects that are directed toward a common goal
- A broad framework of goals to be achieved, serving as a basis to define and plan specific projects

前者はシステムとして機能しているプログラムです。共通の目的のために、さまざまな政策ツール（instruments を組み合わせ、それぞれが補完的に機能しているようなイメージです。後者は、幅広い枠組みとしてのプログラムです。計画や特定のプロジェクトを定義するために、大きな括りとしての政策領域です。

政策評価にとって重要なのは前者の意味でのプログラムです。評価対象がシステムとして把握できなければプログラムではありません。

我が国の政策の基本単位は事業ですが、プログラムとして評価すべき対象か、プロジェクトとして評価すべき対象かは事業の目的と構成によります。

- プログラムとしての事業：事業の目的が組織のミッションに資するものであり、かつ、中長期的には社会的改善をもたらすもの。また、全体が複数のテーマ、プロジェクトから構成されるもの。（例：研究開発助成制度、拠点形成、人材育成、技術移転等）
- プロジェクトとしての事業：より具体的な研究開発目的があり、短期的に達成されなければならないもの。（委託研究開発事業等）

プロジェクトとしての事業に関しても、本来であれば大きな政策的枠組みの中で手段として位置づけられている必要があります。これを政策のプログラム化と呼びます。米国では政策の基本単位（予算単位）がプログラムとなっていますので、プログラムを評価することが財務当局との取り決めになっています。そのために、PART（Program Assessment Rating Tool）と呼ばれるプログラム評価方法に従って、評価が行われます。

政策のプログラム化は我が国にとっても重要な課題のひとつですが、本テキストでは研究開発プログラムの評価に限定して、特に競争的研究資金による研究開発制度をイメージして解説します。

3.1　プログラムを評価するとはどういうことか？

プログラムの例として競争的資金による研究開発制度を考えてみましょう。プログラムの政策目的にもよりますが、プログラムの存在意義とは基本的には良質な研究開発課題を集約し、プログラム自体も改善を重ね、持続的に発展し、多くの優れた研究開発成果を生み出す仕組みとして機能することと考えられます。

プログラムはこの意味で、ある一定期間内に成果を出さなければならない従属型プロジェクトや独立型プロジェクトとは性格を異にするものです（もちろん、プロジェクト型のプログラムも多く存在します)。

しかしながら、仕組みとしてのプログラムはよほど上手く設計され、マネジメントされていなければ機能しません。また、課題設定についても時事の環境の変化を受け入れ、研究開発のトレンドを踏まえたものでなければ、応募枠の申請数を一定数維持することは難しいでしょう。採択評価の手続きも申請者の立場に立ったものでなければ、プログラムのリピーターが減ることにもなります。

ここで具体的なデータを見てみましょう。図は独立行政法人科学技術振興機構の実施していた「独創的シーズ展開事業」の申請数と採択数の推移に関するものです。

出所：文部科学省データより財団法人政策科学研究所作成

図 3-1　独創的シーズ展開事業の申請数と採択数の推移

上の図 3-1 ではプログラムの応募枠毎の採択率が描かれています。募集年度によって、採択率にかなりの変化が見られます。新規課題や継続課題が一緒になっている、あるいは課

題の質も年によってバラツキがあることを考慮すると、こうした数字の推移だけで何かを説明することは難しいですが、ここで確認しておくべきことは、プログラムは必ずしも計画通りに実施することができない、という点です。

例えば上の図 3-1 にあるように、年度によっては申請数が半減しているというような状況をプログラムの計画時に見通すことできるでしょうか。このような状況を想定せず、相対評価による採択率の目安をプロジェクトの採択に関して設定しているとすれば、年によって採択された案件の質が大きくバラつくことになり、プログラム全体のパフォーマンスにも影響を及ぼすことになるでしょう。

このように、プログラムは本質的に初期最適の発想では設計できない、という特徴を強く持っています。初期最適化ができないということは、プログラム・マネジメントが必須であり、プログラムを評価するということは、プログラムのマネジメント・サイクルを評価するということになります。

3.2　ＲＯＡＭＥＦの設定

1990 年代後半に、イギリス貿易省（DTI）であるプログラム評価が実施されました。1983～88 年に実施された情報通信技術開発に関する産学連携プログラム、"Alvey Programme" の追跡評価です。Alvey Programme に対する評価結果は「失敗」でした。失敗のもっとも大きな原因は、目的・目標設定が妥当でなかったことと、目標を達成するための措置を政策として上手く設計できていないことが挙げられています（企業の参加が期待していたよりも少なく、効果が十分に挙げられなかったためです）[4]。

しかしながら、この追跡評価はその後の英国のプログラム設計・運用方法に重要な知見をもたらしました。いわゆる"ROAMEF"と呼ばれるプログラム・マネジメント・サイクルの原則です。ROAMEF とは、Rationale（政策的位置づけ）、Objectives（施策の目的、目標）、Appraisal（事前における規定類・基準類の策定）、Monitoring（実施段階における情報収集）、Evaluation（施策の事後評価・追跡評価）、Feedback（政策のノウハウベースの獲得、次の施策策定への知見）からなる一連のマネジメント・サイクルの頭文字をとったものです。

図 3-2　ROAMEF サイクル

ROAMEF は、評価のみを対象としているのではなく、プログラムのライフサイクルの各段階においてどのような観点を持ってマネジメントにあたるか、ということを重視しています。わが国の研究開発評価の３つの観点（必要性、有効性、効率性）と比較すると、ROAMEF の位置づけは次のようになります。

[4] <http://www.bopcris.ac.uk/bopall/ref22284.html>

表 3-1 ROAMEF と評価体系の論理構造

	評価項目	小項目	評価基準	
Why （必要性）	理由	上位の計画	計画に包接されているか	Rationale
		戦略	全体的な視点から妥当か 先見的な視点から妥当か	
	位置づけ	施策体系	重複はないか バランスはとれているか	
		施策展開	高度化が図られているか	
What （有効性）	目的		明確に定められているか 目標は低すぎないか 目的は妥当であったか	Objective
	成果		期待される成果はコストに比し大きいか 得られた成果は計画を上回っているか 最終的に得られた成果は他の場合に比し大きいか	
How （効率性）	手段		評価方法が具体的に定められているか 評価方法は目的に適合しているか 関連する活動主体にインセンティブを与えているか	Appraisal
	体制	制度	ロードマップは定められているか 見直し制度を内包しているか	Monitoring
		人・組織	責任体制は的確か 外部とのネットワークに配慮されているか	Evaluation
	コスト		他と比較して優位であるか 管理費用は低く抑えられているか	Feedback
その他	対象固有の項目他			

出所：平澤冷,「平成１７年度研究開発評価研修」配付資料

Rationale はプログラムの必要性を政策的位置づけによって説明するものです。この段階では、政策体系の中で当該プログラムがきちんと位置づけられているかを明らかにします。したがって、政策体系が規定されているのであれば、プログラムの根拠は妥当であると判断されます。わが国の政策評価の枠組みで言えば、上位施策との関連を述べることで、研究開発制度の必要性を上位施策が決定された段階において既に担保している、と考えることに相当します。仮に当該制度が上位施策との関連性を直接的には表現できないような場合においても、当該制度の持つ「意義」を上位施策の目的と関連して説明すれば、Rational であると説明することができます。この「Rational である」という説明が、「なぜプログラムを実施するのか？（Why）」に対する回答となり、すなわち、当該制度の「必要性」を明確にすることと同義となります。

Objectives はプログラムの中身を説明するものです。プログラムの目的・目標を明確にすること、それはすなわち、社会的・経済的インパクトをも事前の段階で想定しておくことに他なりません。わが国の政策評価の枠組みで言えば、事前評価の段階であれば研究開発制度の目的・目標を明記すること、中間・事後評価の段階であれば得られた成果や目標達成度を明記することに相当します。そして、パフォーマンスの評価とは別に、当該制度がどのような社会的・経済的インパクトをもたらすかについて、定性的に記述することが

重要になります。この objectives を明確に記述することが、「プログラムで何を行うのか？（What）」に対する回答となり、それは当該制度の「有効性」を明確にすることにつながります。

Appraisal から Feedback まではプログラムの運営に係わるチェック項目と言って良いでしょう。ここで Evaluation が一連の手続きに位置づけられているのは、Evaluation そのものは、プログラムにおいてあらかじめ計画されていた内容が真に達成されていたかどうか、について分析等を通じて理解する行為だからです。したがって、わが国で使用される「評価」という言葉の範囲と、英語の「Evaluation」という言葉の範囲は必ずしも一致せず、むしろ「評価」が「Evaluation」を内包していると理解した方がわかりやすいかもしれません。Appraisal から Feedback までの内容を明確にすることは、「どのようにプログラムを運営するのか？（How）」を明らかにすることです。このことはわが国の政策評価の枠組みで言えば、「効率性」の観点から研究開発制度を評価することになります。

ROAMEF サイクルが優れている点は評価体系として各要素が独立の構造となっている点ばかりでなく、プログラムの実施に係わる一連の手続きとして、目的＝手段関係が個々の要素に関して成立している点です。このことは ROAMEF を逆に読むとよく分かります。すなわち、Feedback を得るためには Evaluation が必要であり、Evaluation のためには Monitoring による情報収集が必要であり、Monitoring する項目を決めるためには Appraisal が必要であり、Appraisal で規定すべき内容は Objectives が明確にされていなければならず、Objectives を政策として展開するためには Rationale である必要があるからです。つまり、プログラムの改善（Feedback）という最終的な目的のために各手続きが位置づけられていることになります。

図 3-3　ROAMEF サイクルと制度の運用

出所：財団法人政策科学研究所作成資料

プログラムのマネジメント・サイクルを評価するためには、ROAMEF サイクルがもっとも参考になります。現在、ROAMEF サイクルは英国のみならず欧州各国の科学技術政策、イノベーション政策の重要な指針として、機能しています。

以下では、ROAMEF サイクルを原則としつつ、プログラムの評価時期においてどのような評価項目を重視すべきか、評価のために必要な手続き等について解説します。

3.3　プログラム評価のための準備

わが国において制度評価が本格的に実施されるようになったのは、ここ数年のことです。多くの場合には現在実施している、あるいは、過去において実施した研究開発制度をどのように評価したらよいかということが中心的な問題になります。

また、府省に関しては政策評価法に基づき毎年実施事業の事前評価書を提出する必要があります。継続事業や拡充事業というかたちで、研究開発制度を見直す際にも、再設計の手続きは必要になります。

評価時期毎に留意すべき点は後に詳細に見ていくとして、アウトカムを産み出すためのシステムとしてプログラムを再設計するための考え方を挙げることにします。

3.3.1　プログラムの再設計

3.3.1.1　必要性の再設定

ROAMEF に従えば、まず最初に行わなければならない点は対象となるプログラムの Rationale を定義することです。これは必要性を明確にすることです。

必要性の再設定を行うためには以下の作業が必要となります。

- 政策体系の整理
- 制度の意義を裏付けるデータ（仮説の根拠）

政策体系の整理については、当該制度に関する上位施策からの位置づけと、類似制度との差別化が必要になります。制度の意義を裏付けるデータについては、制度が関心を持つ事象の特徴や推移を示すデータを用意し、制度が中長期的な目標としてどのような効果を実現するためのものかについての説明を行います。

図 3-4　プログラムの意義の示し方

3.3.1.2　有効性の再設定

ROAMEF の Objectives を明らかにします。必要性のところで述べた意義、つまり中長期的な目標と制度実施期間内に達成すべき短期的目標との因果関係や影響の大きさを考慮して設計されなければなりません。

ここで、陥りやすいのが次のようなロジック展開です。

図 3-5　陥りやすいプログラムの設定の仕方

中長期的な目標、つまり問題となっている状況の改善の方向を示したとして、プログラムによる総合的な取り組みの必要性を訴えるまでは十分な根拠と問題認識を示すことができます。しかしながら、なぜ個別の取り組みが必要かについての考察がプログラム起案者の経験的な知識から設定されているだけであって、具体的な根拠がありません。

プログラムの再設計のために本来なすべきことは、課題となっている状況の分析です。上の事例では、なぜ諸外国と比較して日本のランクが下がったのかに対する分析が必要とされています。諸外国の科学技術政策の動向、予算の重点化のトレンド、人材開発のための仕組み・工夫、良質な研究開発課題の採択の仕方など、参考とすべき点を踏まえているかどうかがまず抑えられていなければなりません。また、国内固有の条件として良質な成果を生み出すためのリソースが充足しているかどうかの検討も必要です。関連研究領域の研究者の数、学会の動向、ネットワークの形成状況、インフラ・設備面の充足状況など、どこにボトルネックがあるかについての分析も必要です。つまり、政策研究が不可欠であるということです。

このような分析を踏まえなければ、具体的な取り組みとして掲げた項目の達成目標が議

論できません。つまり、中長期的な目標に対するプログラムによって改善されるべき個別の施策目標が議論できないだけでなく、いつまでにどの程度の改善を図るべきかの目標水準の議論ができなくなってしまいます。

図 3-6 プログラムの有効性の再設定の仕方

政策研究についてはプログラム実施機関が自ら調査・分析しても、適切な他の分析事例を参照しても構いません。審議会等の議論だけでなく、国や独立行政法人が実施している数々の調査分析の報告書がこのようなロジック展開においてレファレンスとして活用されることで、政策がオープンに議論され、政策研究自体の質も向上すると考えられます。

さて、上のようなロジック展開を行うと何が見えてくるでしょうか。中長期的な目標、つまりプログラムの目的は「●●分野における日本のランキングの向上」です。これは中長期的なアウトカムといってよいでしょう。次に、目標水準の項はプログラムの実施期間中であればアウトプットとして把握できます。原因やボトルネックの部分は、それぞれプログラムが達成すべきアウトカムといって良いでしょう。プログラムの所掌範囲はこのアウトプットとアウトカムであり、その結果を通じて果たして日本のランキングが上がるかどうかはプログラムのコントロールの範囲外ということになります。しかし、範囲内のアウトプット、アウトカムについては確実に実現しなければなりません。

このようにプログラムの所掌範囲を明確にし、かつ実現すべきアウトカム（プログラムの目標）を論理的に定めることが、ROAMEF の Objectives を明らかにするという行為なのです。

こうしたロジック展開の他に、政策の有効性を示す方法としてはインパクトアセスメントと呼ばれる方法があります。インパクトアセスメントはある程度需要予測がつく政策領域に関して適用される方法論です。例えば、規制の影響、公共事業の経済性評価、教育プログラムや保健衛生プログラムなどの影響が相当します。インパクトアセスメントでは、プログラムの恩恵を受けるユーザーがどれだけいるか、ユーザーがどのような恩恵を受けるかを推計し、その影響（インパクト）の大きさでプログラムの有効性を明らかにします。

研究開発プログラムの場合でも、市場化や普及啓発までを視野に入れたプログラムにおいては、有効性のロジック展開をインパクトアセスメントの手法にしたがって行う必要があります。

3.3.1.3　効率性の再設定

プログラムの意義、目的・目標が定まったとして、それをどのように展開するかは政策手段としてのプログラムの中身を明らかにしなければなりません。今度はプログラムの設計図を作る作業に移ります。これがなければ、個別課題の採択（Appraisal）や情報収集（Monitoring）、中間･事後評価（Evaluation）ができません。

プログラムの設計図に必要なのは次の要素です。

- 時間軸
- リソース
- アクター
- メカニズム

時間軸については説明の必要がないでしょう。研究開発制度の多くは予算執行上、事業の形態をとるために３年ないしは５年といった一定予算年度の枠が最初から決まっています。より詳細なタイムテーブルはプログラムの詳細が設計された後に実現可能なスケジュールを検討することで作成されます。

リソースについてはさまざまなものが考えられます。設計しようとしているプログラムの事業予算だけではなく、研究開発に携わろうとしている多様な参加主体のリソースが投入されてはじめてプログラムが実施可能となります。例えば地域における研究開発拠点の形成を目的としたプログラムでは、大学、企業、公設試験機関などがどのようなリソースを持ち寄り、持続可能な拠点形成メカニズムを提案するかが問われることになります。ここでのリソースとはポテンシャル（潜在的能力）といっても良いでしょう。

アクターとはプログラムの枠の中で活動する組織や個人を指します。プログラムがどのような参加主体による活動を経て、最終的な裨益者（beneficiary）にどのような効果をもたらすのかを特定しなければなりません。

メカニズムとは、いかなるアクターのどのようなリソースを用いてどのような活動を経てアウトプットやアウトカムをもたらすかという、プログラムの仕組みのことです。

以上の要素を整理して、次のような大まかなプログラムのロジックモデルを考えてみましょう。

出所：財団法人政策科学研究所作成資料

図 3-7　プログラムの時間軸・リソース・アクター・メカニズム

図 3-7 はある特定の技術開発分野に関する拠点形成プログラムをイメージしたものです。参加するアクターは大学、研究開発独法、コーディネート機関（TLO 等）、地域内の企業、都道府県等の公設試験機関です。さらに研究開発の成果のユーザーとしては、域内・域外のプログラム非参加の企業、最終的な消費者が想定されます。プログラムではこのイメージに適合的な地域オリジナルの仕組みの提案を受け、採択した場合には中核機関に対してプロジェクト予算を配分するものとします。

メカニズムとしては、TLO 等のコーディネート機関が地域内企業のニーズと大学・研究開発機関等のシーズをマッチングし、課題を特定した後に研究開発を促進させ、得られた研究開発成果の権利化を行います。同時に、公設試験機関等を利用し、実用化研究を同時に進め、企業に試作品の開発までを請け負わせます。権利化・実用化された技術は、地域内外の企業からの引き合いを受ける段階で付加価値化され、市場での普及を通じて一定の社会経済的効果をもたらします。これらの効果も重要ですが、拠点形成により、こうした仕組みを持続的に発展させていくことがこのプログラムの主たる目的になっているものと

します。

ここで、プログラムの所掌範囲は、研究開発拠点から生み出された技術等が権利化され、ロイヤリティを生むまで、あるいは試作品・実用化技術等が企業の引き合いを受けるまでとなっています。

以上のような状況を想定しつつ、効率性の再設定を行ってみましょう。

まず、プロジェクト採択（Appraisal）についてです。プロジェクトの採択に関わる評価では、プログラムの目的・目標を達成するに十分なポテンシャルを有しているかどうかということが最大のポイントです。ここで、リソースの賦存状況に関して精査が行われなければなりません。図の例であれば、リソースの賦存状況として確認すべき点は次のようなポイントが挙げられるでしょう。

- 研究開発能力（実績等）
- 地域内企業の特性，産業基盤等
- コーディネート人材の適切性
- コーディネート機関の業務執行能力
- 利用可能な地域内のインフラ・設備等
- 中核機関におけるプロジェクト管理執行能力

次にプロジェクトの進捗状況の管理（Monitoring）について考えてみましょう。採択評価の折に、中核機関におけるプロジェクトの管理執行能力を条件として挙げているのは、必要な時に必要な情報としてプロジェクトのデータが得られることを期待してのことです。プログラム評価はプロジェクト評価よりも広い観点から評価されます。例えば、競争的研究資金の枠組みであれば、ポストドクターの採用状況などがプログラムレベルで集計できなければなりません。中核機関の管理能力が低ければ、プロジェクト予算でどこの研究室がポストドクターを何人採用したかどうかすら把握していないことでしょう。大学教員等のエフォート管理も重要です。

管理能力がしっかりしているということは、必要なデータがすぐに提出できるということと同義です。データの蓄積方法については何かしらの指導が必要だとしても、プロジェクト業務の遂行中にデータを蓄積可能な人員・体制を確保していなければ、管理能力に問題があるということになります。

モニタリングとは、資金配分機関とプロジェクト実施機関との間にあらかじめ定められた項目についての情報供出に関わる契約があって、業務遂行中にデータを蓄積し、報告すれば良いだけの評価手続きです。モニタリングデータの推移に資金配分機関側として問題を見出すようであれば、プログラムのアドバイザリーボード等の意見を聴取し、必要に応じてプロジェクトの責任者に対して指導を行うことができます。

中間・事後評価（Evaluation）の段階ではどのような再設定が必要でしょうか。アウト

プットについてはモニタリングで情報を蓄積していれば、特に追加的にデータの提出を要請する必要はありません。上の事例であれば、アウトカムについてもデータ化することが可能でしょう（ライセンシング件数、技術のオファー数等)。プロジェクトから得られるこれらのデータを集計するために、プログラムレベルで集約可能な類型化が必要となります。

また、拠点形成プログラムのような取り組みでは、その仕組み（システム）が持続的かどうかという点が重視されなければなりません。拠点が形成されたことにより、マッチングファンドによる企業との共同開発が進んだ、優れた研究開発が他の競争的資金の発展課題として採択された、というような事例があれば、それらを類型化して他のプロジェクトと共通に実績を評価する必要があります。地域におけるネットワークの活性化を評価する場合でも、活性化を示す指標を他のプロジェクトと共通で指定する必要があるでしょう。

このように、プログラムの評価では個別プロジェクトから得られるデータをプログラムレベルで集計するために、またポートフォリオのようなかたちで分析するために、データの類型化（カテゴライズ）が必要になります。この作業は、本来、プロジェクトの成果報告書を作成する際に共通の仕様を課していれば済む問題ですが、プログラム評価よりも前にプロジェクト評価を先に実施している場合には、プロジェクトにより提供されているデータの量、範囲等が異なっている場合がほとんどです。こうした場合には、プロジェクト評価書のメタ評価（共通の枠組みで比較すること）を前提に、データを収集しなおさなければなりません。

ROAMEF サイクルではプログラムレベルでの評価のためにプロジェクトの評価が位置づけられており、Monitoring や Evaluation に活用されるデータをあらかじめ設定しておくことがプログラム・マネジメントであることを示しているのです。

3.3.2 ロジックモデルの作成

3.3.2.1 プログラムとロジックモデル

プログラムとは中長期的な目的の達成のために設計される政策手段のパッケージです。研究開発に限らず、プログラム化された政策は特定の問題解決のために、さまざまな政策手段を組み合わせ、その複合的な効果を問題解決のために集約させるために設計されています。

ロジックモデルはさまざまな政策手段の意図した効果をどのように最終的な問題解決に結びつけるかを描くためのツールです。

下の図 3-8 は一般的な問題解決のためのプログラムに関して、想定される因果関係を包括的に示したロジックモデルの例です。

出所：T. J. Chapel (2005) "Integrating Program Planning and Evaluation Using Logic Models"より抜粋

図 3-8　包括的なロジックモデルの例

ロジックモデルでは、通常、「意図した結果（アウトカム）を導くために何が必要か？」という問いかけを繰り返し、プログラムのすべての活動に至るまで因果関係を展開していきます。

ロジックモデルの作成で重要なことは、プログラムの所掌範囲がどこまでかを明確に定義することです。上の図では、効果（Effects）で括られている部分の最初の列のボックスが短期的アウトカム（Short- term outcome s ）として挙げられています。通常、プログラムによって実現される効果はこの部分までで、後は社会がプログラムの影響を受けながら変化していき、最終的な目的（意義に相当）である長期的アウトカム（Long-term outcomes）を実現するであろうという期待を示しています。

しかしながら、長期的アウトカムの実現のためには、中期的アウトカムが先に実現される必要があり、さらにそれに先だって、短期的なアウトカムが実現されなければならないという、因果関係を示している点で、ロジックモデルは大変有用なコミュニケーションツールとなります。

3.3.2.2　ロジックモデルの作成の仕方

プログラムの実施者側は、評価の専門家とロジックモデルの開発について一緒に仕事をすることを考えなければなりません。ロジックモデルに関する経験を持つ評価の専門家とプログラムマネージャーの対象分野に関する専門知識を組み合わせることにより、ロジックモデルをより効果的に構築することができます。

ロジックモデルを構築するためには、次の構成要素を明らかにする必要があります。

- 活動（activities）：そのプログラムの下で、関係するアクターが従事する主要な活動は何か。すなわち、アウトカムの達成に貢献する主要な活動は何か。
- アウトプット（outputs）：主要な活動のアウトプットは何か。すなわち、活動が着手されたことを示すものは何か。アウトプットはアウトカム実現のための活動の水準であり、目に見えて（tangible）、数えられる（countable）ものとして把握される。
- 短期的アウトカム（short-term outcomes）：活動やアウトプットから生じる短期のアウトカムである。ロジックモデルにおいてアウトカムは、概して「増大した」とか「改善した」というようなアクションワードを持っており、活動やアウトプットの結果を示している。プログラムの所掌範囲として最低限責任をもつべき結果として把握される。
- 中期的アウトカム（mid-term outcomes）：短期的アウトカムの連鎖における次のつながりは何か。逆に、プログラムの意義でもある長期的なアウトカムの実現に必要な条件は何か。これらのアウトカムは、中期的アウトカムであると解釈される。
- 長期的アウトカム（log-term outcomes）：　政策の最終的なアウトカムは何か。どうしてこれらの活動が行われているか。これらは、一般に実感されるより長い時間がかかり、政策自身を超えた影響を仮定し、より戦略的なレベルにあることがらである。

しかしながら、ロジックモデルにも記述しづらいプログラムの構成要素もあります。例えば、

- どのようにプログラムが実施されるかについての、特定のマネジメント上の詳細事項
- 組織経営、設備運用、業務のプロセスに焦点をあてたもので、スタッフ雇用、装置の購入、ドキュメントのやり取りのような活動。

これらは、重要なプログラムの活動ですが、通常、ロジックモデルに入れることはできません。

ロジックモデルを開発するためにはプログラムの主要な構成要素を特定する必要があります。一つの効果的なアプローチは、評価の専門家のサポートを受けながらグループディスカッションによってモデルを開発することです。プログラムに対してさまざまな視点を持つ個人は、ロジックモデルの開発に貢献できます。

まず最初に議論すべき事は、下の表 3-2 のように、活動、アウトプット、アウトカム（短期的～長期的）に該当する要素をリストアップすることです。

表 3-2　ロジックモデル作成の第1段階

	活動 （activities）	アウトプット （outputs）	短期的アウトカム （short-term outcomes）	中期的アウトカム （mid-term outcomes）	長期的アウトカム （long-term outcomes）
構成要素	• 既存施設の利用開放 • インフラ・設備の高度化のための技術開発 • 科学技術コーディネータによるシーズとニーズのマッチング • マッチング課題の発展的研究開発の実施 • 各種普及啓発活動 • 海外共同研究者の参画 • 海外評価者によるパネル評価	• 施設の利用者増加、利用者属性の多様化 • インフラ・設備の高度化達成水準 • 科学技術コーディネータによるマッチング件数 • マッチング課題の研究開発成果 • 普及啓発の活動水準 • 若手研究者の参加数 • 海外共同研究者の参加数 • 海外評価者の参加件数	• 既存施設のアクセスの多様化 • インフラ・設備の高度化・拡充 • 産業界との共同研究の拡充 • 海外からの成果引き合いの増加	• 良質な課題に対する十分な研究資金の助成 • 優れた実験設備等へのアクセスの向上 • 若手研究者の参入 • ナレッジの増加 • オープンなコミュニティの形成	• ●●分野における日本のランキングの向上
	プログラムによる対策の活動内容	プログラムによる対策の活動水準	プログラムによる対策の結果	現在の状況の原因となっている個別要因の改善	改善したい状況（プログラムの意義）

出所：財団法人政策科学研究所作成資料

長期的アウトカムの欄については、プログラムの意義の部分で説明された、改善したい状況がそのまま入ります。中長期的アウトカムの欄については、関心となっている状況の分析から得られた、個別要因の改善目標が入ります。短期的アウトカムの欄については、それらにつながるプログラムが直接的に及ぼし得る対策の結果が入ります。ここまでの議論では、プログラムの有効性を明らかにした際の問題認識がそのまま適用できます。

プログラムのアウトプットについては、短期的アウトカムに寄与するであろう活動の水準が列挙されます。アウトプットは形式的には数値データとして把握可能なものとなります。

活動の欄にはアウトプットを生み出す仕組みとして何を実施するのかの説明がリストアップされます。

次に、上の表 3-2 をフローチャートとして表現し、因果関係を考察していきます。

出所：財団法人政策科学研究所作成資料

図 3-9　ロジックモデル作成の第2段階

表 3-2 でリストアップした項目をボックスで配置し、それぞれの因果関係を矢印で結びます。原因となるボックスから結果となるボックスに対して→の矢印を引きます。上の図では、「ナレッジの増加」と「オープンなコミュニティの形成」が双方向の矢印で結ばれていますが、このような関係は共益関係です。お互いがお互いの実現に対して影響し合うような場合には双方向の矢印で結びます。

ここまでの図 3-9 をプログラムマネージャー間で議論して作成したら、評価の専門家に見てもらうことが重要です。

3.3.2.3　ロジックモデルのチェック

評価の専門家は因果関係に誤りがないか、必要な要素が抜けていないかをチェックします。特に、アウトカムのロジック展開に関するフローに着目してチェックします。

アウトカムのロジック展開で重要な点は次の２点です。

- あるアウトカムは意図するアウトカムの直接的な因果関係を持つか？
- アウトカム同士の関係はいつでも成り立つか？

前者の質問に対する検討内容は、「アウトカムの実現に対して媒介的な機能（mediator）が抜け落ちていないか？」ということです。例えば上の図では、「産業界との共同研究の拡充」と「海外からの成果引き合いの増加」が「良質な課題に対する研究資金の増加」という中期的アウトカムをもたらしていますが、このような因果関係が直接的に成立するでしょうか。

出所：財団法人政策科学研究所作成資料

図 3-10　ロジックモデルにおける媒介的機能（mediator）の例

直接的に、つまり放っておいても因果関係が成り立ちそうであれば、ロジック展開としては問題がありませんが、何かしらの媒介的機能がなければロジック展開に無理があると考えられる場合には、具体的な媒介的機能を検討することになります。上の例では、プログラムの次の措置として新たなファンディング・システムの導入の必要性を検討しています。

また、後者の質問に関する検討内容は、「アウトカムの実現に対して、社会経済的環境の変化が決定的に作用しないかどうか？」ということです。もし、アウトカムの実現に対して環境の変化が大きいようであれば、ロジックチャートにおいて、外的要因を調整的機能（moderator）として考慮する必要が出てきます。例えば作成したロジックモデルの図 3-11 で、「施設利用者の増加、利用者属性の多様化」と「既存施設のアクセスの多様化」の因果関係において、施設利用料金（コスト）が施設運営機関の事情で高騰した場合には、限られた研究開発実施者しか利用できないということになり、多様化が促進されるとは限りません。

出所：財団法人政策科学研究所作成資料

図 3-11　ロジックモデルにおける調整的機能（moderator）の例

アウトカムが、環境要因が変化しないという前提で設定されている場合、政策担当者としてどこまでのリスク要因をプログラムの運営上考慮しているかという点がチェックの対象となります。

このような修正を経て、ロジックモデルは適宜見直される必要があります。

3.3.3　パフォーマンス指標の設定

ロジックモデルでアウトカムを示した後に重要なことは、アウトカムが達成されているかどうかを判断するための評価指標を特定することです。

ここでは、ロジックモデルにも転用が可能なインディケーターの例として、米国行政管理予算局（OMB：Office of Management and Budget）が採用している PART（Program Assessment Rating Tool）の業績評価指標を紹介します。米国では OMB が５年毎に省庁の行政プログラムの評価を行っており、PART と呼ばれるプログラム評価ツールを採用しています。PART とは、次のような特徴を持っています。

- 行政プログラムの業績（パフォーマンス）を測定し、業績を予算審議に反映させる
- 以下の4つのセクション（合計25個）の共通質問事項から構成される：
  - １．プログラムの目的とデザイン
  - ２．戦略的プランニング
  - ３．プログラム管理
  - ４．プログラム結果／説明責任
- 省庁側は単に質問に答えるだけではなく、その答えの根拠となる証拠書類を提示することが求められる。

PART が導入された結果、省庁側は自らのプログラムの業績を定量的に測定するために、プログラムのアウトプットやアウトカムに対してインディケーターを設定し、パフォーマ

ンスの計測を行うようになっています 。

ここでは、米国エネルギー省（DOE）の風力エネルギープログラムのインディケーターの設定の仕方について紹介しましょう。

事例：米国エネルギー省（DOE）風力エネルギープログラムのインディケーター

【プログラムの目的】

風力エネルギープログラム（Wind Energy Program）は DOE のエネルギー効率・再生可能エネルギー局（EERE : Office of Energy Efficiency and Renewable Energy）により運用されるプログラムであり、風力エネルギーの実現可能性およびその利用レベルの向上を目的としています。このプログラムは、より大きな枠組みである風力・水力発電技術プログラム（Wind and Hydropower Technologies Program）の構成要素の 1 つであり、次のような構造を持っています。

出所：EERE Web サイトより財団法人政策科学研究所作成

図 3-12 米国エネルギー省（DOE）の風力エネルギープログラム

【プログラムの目標】

プログラムの目標には次の４つが掲げられています。

・2012 年までに大規模な風力システムを用いた発電コストをクラス４の風で３¢／kWh（陸

上）またはクラス３の風で５¢／kWh（海上）まで下げる。

・2007 年までに分散型風力システム（distributed wind system）を用いて現在クラス 5 の風力で得られている 3¢／kWh のコストをクラス 3 の風力で達成する。

・風力が国のエネルギー需要に支障なく貢献できるよう、2012 年までに電力市場ルールや相互接続影響、運営戦略、系統計画等の整備をプログラムの活動の一環として完了させる。

・2010 年までに少なくとも 30 の州で最低 100MW クラスの風力エネルギーの利用が行われるよう普及啓発事業を展開する。

【プログラムのインディケーター】

風力エネルギープログラムのアウトプット・インディケーターには次のような発電単価が目標値ならびに実績値として示されています。

出所：EERE データより財団法人政策科学研究所作成

図 3-13　米国エネルギー省（DOE）風力エネルギープログラムのアウトプット・インディケーター

また、アウトカムのインディケーターとしては、次のような指標が目標値ならびに実績値として示されています。

出所：EERE データより財団法人政策科学研究所作成

図 3-14 米国エネルギー省（DOE）風力エネルギープログラムのアウトカム・インディケータ

ここで注意しなければならないことは、PART では、改善しようとしているアウトカム指標に対して、どれだけ当該プログラムが具体的に貢献したかどうかの寄与率等は示されていないということです。また、アウトカム指標である風力発電設備の導入州の数については、本来であれば州の特性（人口、風況、既往電源の構成、送電網等の整備状況）の差などを考慮して、単純に集計できそうもありませんが、プログラムの目標として大型風力発電設備の導入目標（2010 年までに 30 州）を掲げているため（また、そのための普及啓発事業もプログラムの活動の範囲となっています）、このようなかたちでプログラムのパフォーマンスを示しています。

この理由は、PART ではプログラムが対象とする特定の社会経済的状況が、プログラム実施期間中もしくは終了後一定期間中に結果として改善されていれば良いという立場に立っていて、インディケーターの設定の焦点はどれくらいパフォーマンスが向上しているかということを目標値と実績値とで比較して定量的に把握することに当てられているためです。

PART のアプローチはある程度の合理性を持っていると考えられます。政策展開において大切なことは、国民に結果を目に見えるかたちで示すことであり、結果が良ければプログラムの妥当性は担保されるというものです。また、プログラムの内部に結果を導くための政策的措置が組み込まれていることで、正確な寄与率は判断できなくても、一定の寄与があったとプログラム実施者側は主張することができます。PART の厳密な点は、こうしたパフォーマンス指標をプログラムの予算折衝の段階で実施側から提供することを要請している点です。つまり、後から「いいとこ取り」ができないよう、最初から政策の効果を「約束」させているのです。

3.4　プログラムの事前評価

わが国の場合、プログラムの事前評価に相当するのは、概算要求時に府省が作成する毎年の事前評価書において記述される研究開発制度等の自己評価です。

研究開発制度等のプログラムは複数年度展開されることが一般的で、かつ、発展的に継続される場合が少なくありません。つまり、プログラムの事前評価では、プログラムの循環的改善に関する記述が中心となります。

循環的改善とは、目標を明確に定め、モニタリングによる達成度評価の結果を毎年の事前評価の指標に用いることです。

ここでは、プログラムの事前評価書で特に記述すべき内容について解説します。

3.4.1　プログラムの目標

プログラムの意義、アウトカムの設定の仕方についてはすでに見てきました。ここでは、プログラムの目標について考察してみましょう。

プログラムの目標とは何でしょうか。事前評価書に記載すべき内容としては、成果に関わる内容と実施機関側のマネジメントに関わる内容が抑えられている必要があります。

3.4.1.1　成果に関わるプログラムの目標

ミッション型のプログラムにおいては、アウトカム（意図した結果）が得られているかどうかを示すことが目標達成状況を確認する手段となります。学術研究や基礎研究、応用研究等への助成を目的としたプログラムでは、優れた成果の輩出が目的であるため、成果の評価軸におけるポートフォリオから期待した成果群が出現しているかどうかを示すことが目標となります。

例えば、応用研究への研究開発助成プログラムの評価結果をこのような図 3-15 にして示すことができるとします。評価軸は当該技術の実用化に関する評価項目と、市場性に関する評価項目です。それぞれの評価項目に関する評点を軸にして、研究開発課題の評価結果をプロットすると、ポートフォリオを描くことができます。

実施しているプログラムがポートフォリオのどの領域における研究開発課題の発掘を行うべきかの戦略的意図（＝目標）があれば、課題の終了評価の結果を毎年プロットしていくことで、戦略的目標の達成度を視覚的に示すことができます。

出所：財団法人政策科学研究所作成資料

図 3-15　成果の評価軸による課題のポートフォリオ

3.4.1.2　マネジメントに関わるプログラムの目標

提案公募型研究開発事業においては、プロジェクトの採択審査から終了評価まで一連のマネジメントが不可欠になります。マネジメントの改善目標をバランススコアカードのような方法で示すことで、毎年の事前評価書におけるプログラムの目標を記述することができます。

表 3-3　マネジメントに関わるプログラムの目標

視点	内容	単位	目標	H17	H16
財務の視点	事業費の推移	金額／年			
	事業費全体に占める管理費用	%			
	事業担当者の人数	人			
顧客の視点	応募申請数の伸び	申請数			
	顧客満足度	%			
	Web閲覧数の伸び（プログラムの認知度）	閲覧数			
	評価報告書のダウンロード数	DL数			
内部業務プロセスの視点	十分な審査体制（1課題あたりの審査人数）	人／課題			
	採択審査の決定から助成契約までの期間	日			
イノベーションと学習の視点	担当者の研修日数	日／人			
	担当者のOff-JT機会	回／人			
	中間評価におけるプログラムマネジメントの評点	評点			

実績と目標を
記載

出所：財団法人政策科学研究所作成資料

プログラムの事前評価では、モニタリング等で得られるプログラムに関わるデータを活用し、循環的にプログラムを改善していくことで目標設定を行うことになります。

3.4.2　評価システムの運用

プログラム・マネジメントを循環的に改善していくことの目的は、より良質な研究開発課題を見いだし、研究開発を促進することで優れた成果を生み出すことです。良質な課題や優れた成果は、プログラム自身が評価システムを内包し、評価システムを運用することで見いだされるものです。

したがって、事前評価の段階でプログラムの評価システムをきちんと明記しておく必要があります。評価システムとは、評価の体制、評価の仕組み、評価結果の活用方法に関する一連の取り決めのことです。

3.4.2.1　評価の体制

評価の体制とは評価から意思決定に至る実施体制のことです。採択評価であれば、次のような実施体制と役割が描けるでしょう。

出所：財団法人政策科学研究所作成資料

図 3-16　評価の体制と評価のフロー

プログラムマネージャーは一連の審査のサポートを行います。最初にプロポーザルに対し、形式的な資格審査（formal examination）を行います。次に、ピアレビューを実施するため、レビューアの選出を行います。選出されたレビューアはレビュー評価を実施します。ピアレビューの結果から一次採択案が提出されると、それらについて専門家や有識者からなる評価委員会がランキングを行い、採択評価を行います。プログラムマネージャーはパネル運営に関わり、パネル評価が円滑に行われるようコントロールします。評価委員

会は一時採択案のランキングを行い、最終的な採択結果の案をとりまとめます。最後に、プログラムの責任者であるプログラム・ディレクターが承認して採択結果が決定されます。

このように、評価における実施体制と役割、あるいは責任体制をあらかじめ明確にしておくことが重要です。

3.4.2.2　評価の仕組み

評価の仕組みとは、具体的にどのような評価項目・評価基準でプロポーザルが評価されるかを明らかにすることです。また、必要に応じて、ヒアリングや現地見学などのオプションを実施するなどの措置についても明確にしておきます。

3.4.2.3　評価結果の活用の方法

中間評価ならびに事後評価の結果について、どのように活用するかについて評価システムを設計する段階で明らかにしておきます。例えば次のような活用の位置づけが考えられます。

- 中間評価では、プログラム・マネジメントの妥当性について評価し、事後評価段階でのプログラム実施側のマネジメント改善の判断材料とする。
- 中間評価における成果の出現状況を踏まえ、毎年の採択評価のあり方を検討する際の材料とする。
- 事後評価における成果の出現状況から、後続プログラムの企画立案、マネジメントのあり方の判断材料とする。

3.5　プログラムの中間・事後評価

研究開発プログラムの中間・事後評価についてはどのように考えるべきでしょうか？プロジェクトの場合と異なり、研究開発実施者への助言を得ることはありません。

むしろプログラム実施側のマネジメントを評価することが中心的な問題になります。ここでは、プログラムレベルでの目的・目標とは何か、プロジェクトのデータをプログラムとして評価するための考え方などを学びます。

3.5.1　中間・事後評価の意義

プログラムが５年程度の中期にわたる実施期間を定めている場合、中間年において中間評価を実施し、プログラム・マネジメントの適切性を評価することが重要になります。

プログラム・マネジメントが上手く機能しているかどうかは成果の出現状況と、マネジメント固有の目標設定に対して達成されているかどうかという二つの面から判断されます。

3.5.2　プロジェクトの実績のデータ

プログラムにおける成果とは、プロジェクトの実績をプログラムレベルで集計可能なデータとしてとりまとめ、類型化したり、ポートフォリオを組むことによって示されます。

つまり、プロジェクトのデータを要素としてプログラムの全体像を描く作業が必要になります。以下では、プログラムのロジックモデルとも適合的になるように、インプット、アウトプット、アウトカムのデータの整理の仕方を考察します。

3.5.2.1　インプットデータの類型化

プロジェクトのインプットに関わるデータを集約します。インプットには、プログラムによる投資額・助成額の他に、マッチングファンドであれば民間企業の出資額、投入要素としての研究者数、プロジェクト従事者等の人材、利用設備等の情報が必要になります。

表 3-4 プロジェクトのインプットデータの例

インプット			データ
事業費	公的資金		円
	民間資金		円
事業費内訳	人件費		円
	直接経費	設備費	円
		消耗品費	円
		データ･ソフトウェア等購入費	円
		旅費	円
		その他	円
	間接経費		円
人材	研究者	男性	人
		女性	人
	プロジェクト従事者	男性	人
		女性	人

出所：財団法人政策科学研究所作成資料

これらのデータから、プログラムレベルでの集計を行います。事業費のような連続的なデータは、階級区分を行うことで度数分布にして示すことができます。また、多くのデータはプロジェクト間でデータラベルを統一（詳細なものは統合し、大まかなものは内訳を調べ）して円グラフにすることで、全体的な傾向を示すことができます。

出所：財団法人政策科学研究所作成資料

図 3-17 プログラムレベルでのインプットデータの集計例

3.5.2.2 アウトプットデータの類型化

アウトプットについてもプロジェクト間で集約可能なように類型化を行う必要があります。

表 3-5 プロジェクトのアウトプットデータの例

アウトプット		データ
論文数	海外査読付き	本
	国内査読付き	本
	DP	本
学会報告	招待講演	回
	学会発表	回
国際学会・シンポジウム等	招待講演	回
	学会発表	回
技術報告書		本
特許	IPC分類別出願数	件
ワークショップ開催数		回
報告会開催数		回

出所：財団法人政策科学研究所作成資料

アウトプットデータについては、把握可能な活動の水準を悉皆的に収集しなければなりません。アウトプットの数値自体に意味があるのではなく、どのような活動を行ったのかという類型に意味があります。プロジェクト報告書では、個別具体的な活動が列挙されていますが、それらを類型化し、プロジェクト間で比較可能なデータに落とし込む作業が必要になります。

また、特許データなどは、単に件数を集めても意味がありません。当該特許の IPC（International Patent Code）４桁分類程度の情報を併せて収集すれば、プログラムによるパテントマトリクスなどが描けるため、後々の分析に有用です。

3.5.2.3　アウトカムのデータ

アウトカムのデータについては、プログラムの目的・目標に依存するために、一概に述べることはできません。一部の社会経済的効果（市場化、商品化、起業化等）をプログラムの主たる目的に据えている場合には、そのようなデータをプロジェクトの追跡調査から収集しなければなりませんが、中間報告や事後報告の段階で満足に得られるかどうかは不確定です。

中間評価や事後評価の段階では、プログラムレベルでの集約化されたアウトカムを評価するためには、アウトカムのポテンシャルについてデータ分析するしかありません。

ここでは、高度な分析手法を用いずとも、アウトカムのポテンシャルを評価する方法を例示します。

プログラムの目的・目標は、プロジェクトの評価項目・評価基準にも反映されていなければなりません。プロジェクトの評価項目・評価基準はパネル評価等によって、スコア化されるケースが一般的です。例えば以下の評定区分によって、アウトカムのポテンシャルが評点付けされているとしましょう。

表 3-6　アウトカムに関連する評価項目・評価基準・評定区分の例

評価項目	評価基準	評点	評定区分
研究開発成果の経済的効果は認められるか？			
	成果は実用化の水準に達しているか？	A	実証試験の結果も良好で、十分に実用化の水準に達している
		B	実用化のためには、まだ実証試験の継続が必要である
		C	実証試験段階まで達していないので実用化を判断できない
	成果の市場化の可能性はあるか？	A	革新的な性能向上と製造プロセスの両立が可能であり、十分に市場化の可能性がある
		B	・革新的な性能向上は認められるが、製造プロセスの大幅な更新が必要であり、量産効果が得られなければ市場化は難しい ・製造プロセスの流用が可能だが、性能向上の幅がまだ大きくなく、引き続き開発の必要性を認める
		C	一定の性能向上は認められるが、既存の製造プロセスを踏襲できないため、コストが高く市場性が低い

出所：財団法人政策科学研究所作成資料

上の表 3-6 の評定区分にしたがって、各プロジェクトが評点付けされているとしましょう。ここで、評点Ａを３点、Ｂを２点、Ｃを１点と定義してスコア化します。すると、研究開発成果の経済的効果というアウトカムに対して、「実用化」と「市場化」の二つの評価軸でプロジェクトのポテンシャルを評価することができます。

出所：財団法人政策科学研究所作成資料

図 3-18　ポートフォリオによるアウトカムのポテンシャルの示し方

このようなポートフォリオを作成し、アウトカムのポテンシャルを視覚的に示すためには、アウトカムをパネルに評価してもらうための評価項目・評価基準、特に重要なものが評定区分の設定の仕方になります。プログラムの意図した結果であるアウトカムが、どのような条件をクリアすれば実現可能かについて考察し、プロジェクト評価の際に、その可能性を検討してもらうことが必要です。評価パネルが評点を付けるのに混乱するような評定区分の設定の仕方は厳に慎むべきです。

3.5.3　プログラム固有のデータ

プログラム固有のデータとは、プロジェクトの活動や成果とは独立したデータのことです。例えば、申請数などは、プログラム全体のデータとしてしか把握されません。

プログラム固有のデータをできるかぎり悉皆的に把握し、蓄積することでマネジメントの改善に資する要因を検討することができます。以下はプログラム固有のデータの例です。

表 3-7 プログラム固有のデータの例

プログラム固有のデータ	単位
プログラム全体の予算	円／年
公募説明会実施回数	回
公募説明会実施会場数	箇所
公募説明会参加人数	人
年間公募回数	回
電子申請の導入の有無	ー
公募申請数	件
公募に際して必要な書類の数	種類
レビュー審査数	件
1課題あたりの審査員数	人
採択案件数	件
採択率	%
プロジェクト中止案件数	件
研究開発期間の延長申請数	件
プログラム事業費に占める資金配分機関の管理費用	%
プログラムマネジャーの人数	人
評価に関わるコスト （プロジェクトの終了評価、プログラムの中間・事後評価）	円／年

出所：財団法人政策科学研究所作成資料

プログラム固有のデータをいくら眺めても、すぐれた研究開発課題やプロジェクトが生み出されるわけではありませんが、プログラム全体のマネジメントを考える上での参考情報として整備しておく必要があります。

4　独立型プロジェクトの評価

独立型プロジェクトとは、総合政策・計画の特定の目的を達成するための独立した研究開発プロジェクトのことです。大綱的指針の分類では、予算規模が比較的大きく、重点的資金や基盤的資金による研究開発課題が相当します。

特に、総額が約 300 億円を超える研究開発投資については、総合科学技術会議が必要に応じて専門家・有識者を活用し、府省における評価結果も参考として調査・検討を行い、その結果を受けて評価を行い、その結果を公開するとともに、評価結果を推進体制の改善や予算配分に反映させることとしています。

本章では、国が実施する研究開発のうち、比較的大型の独立型プロジェクトの評価について、重視すべきポイントを解説します。

4.1　独立型プロジェクトに要求されること

4.1.1　プロジェクトと創造のプロセス

大型の研究開発、基盤開発等でもっとも要求されることは、プロジェクトの目的・目標を計画年度内に確実に達成し、成果の運用を図っていくことです。そのためには、確実な実施体制や実効的なマネジメントが不可欠になります。また、全体を統括する強いリーダーシップが必要となります。

ここで、技術開発プロジェクトにおける「創造のプロセス」を図式化したものをご紹介しましょう。「創造」は今日的には「イノベーション」と置き換えて構わないでしょう。

出所：宮田秀明(2004)「理系の経営学」を基に財団法人政策科学研究所作成

図 4-1　技術開発プロジェクトにおける創造のプロセス

この図 4-1 は民間企業の製品開発を想定すると非常に分かりやすくなります。例えば、Apple 社の携帯音楽プレイヤー「iPod」を当てはめてみましょう。「ビジョン」には製品企画段階における究極的な目的が入ります。初期の iPod の開発ビジョンは「ユーザーの CD コレクションを丸ごと持ち運べる携帯音楽プレイヤーを開発する」というものでした。

続く「コンセプト」では製品のスペックを決定します。要求される開発目標水準を具体的なスペックに落としていく作業が発生します。このプロセスには、デザイン性、質感等の感覚的要求水準も含まれます。

「コンセプト」が固まったら、どのように開発するか、商品展開するかビジネスモデル

として成り立たせるか、という具体的な「モデル」の設計を行います。iPod の場合、開発当初から iTunes との連携を前提に設計されていましたが、2005 年からはネット音楽配信ビジネスと組み合わせることで新しいビジネスモデルを創出しました。

「モデル」が決定したら後は「ソリューション」を選択することになります。ここで言う「ソリューション」とは、与えられたビジネスやサービスのモデルを実行するためのやり方（実行手段）のことです。iPod や iPhone では「ソリューション」からプロダクツの発想を得るのではなく、先にプロダクツのコンセプト、モデルがあって、必要な技術を外部調達するというアプローチが採られています。音楽プレイヤーを作ったことのない企業だから採用できた開発戦略であると言えます。

大型プロジェクトが失敗する事由のひとつとして、「ビジョン」からいきなり「ソリューション」への発想のパスが挙げられます。いわゆる「シーズ型」の発想でプロジェクトを企画しても、結果としてイノベーションに至らないという事例は、「コンセプト」や「モデル」の検討が十分になされていないことが多いと言えます。

ここまでの開発戦略が立てられていたら、後はマネジメントをしっかりと行うことです。この段階で非常に重要なことは、「意思決定」と「実行」に関する権限が研究開発実施者側に委ねられているかどうかということです。開発の段階では予期しないことが発生します。状況に応じて適切な判断を全体的な観点から行える人材がいて、それを実行に移すリーダーシップが要求されます。

国が実施する大型の研究開発プロジェクトも、最終的なプロダクツが商品か、公共的基盤を有するものかという違いはあっても、創造のプロセスという観点から見れば民間企業のそれと変わりはありません。例えば、上の創造のプロセスに地球シミュレータ開発プロジェクトを当てはめて考えてみましょう。

地球シミュレータは高度科学演算へのニーズが高まってきたことと、バブル後遺症に悩む我が国において高性能計算機の技術維持・発展を図る観点から、1997 年に当時の科学技術庁主導で開始されました。プロジェクト計画期間は 5 カ年、総額 600 億円の巨費を投じられています。

写真出所：地球シミュレーションセンター提供資料

図 4-2　地球シミュレータの創造プロセス

地球シミュレータ開発プロジェクトでは、「世界最速のコンピュータを開発する」というハード上の開発目的と、「仮想地球を実現し、気候変動、地殻変動等のシミュレーションを行う」という利用目的が設定されていました。こうしたビジョンを実現するための開発基本コンセプトとしては、「気候シミュレーション等の科学技術計算に特化した性能要件を満たす」というものでした。

地球シミュレータの開発モデルは当時、世界の高性能計算機の主流であったスカラー型プロセッサによるグリッドシステムではなく、ベクトル型並列処理方式でネットワークを組むという選択をしています。この基本的方針を受け、ソリューションとしては、開発を請け負った NEC のメインフレーム SX-5 を 1 チップ化し、集積度を高めるという研究開発が実施されました。

ここまでの開発方針をとりまとめ、精力的に研究開発マネジメントを展開し、リーダーシップを発揮したのが故三好甫博士（地球シミュレーション研究開発センター長）でした。三好博士は地球シミュレータの完成を見ることなく他界されましたが、佐藤哲也博士（地球シミュレータセンター長）がシステムアップと基幹ソフトの改良を引き継ぎ、完成させました。

地球シミュレータは 2002 年 4 月に LINPACK ベンチマークで実効性能 35.86TFLOPS を記録しました。この数値は当時最速であった米国のスーパーコンピュータ ASCII-White に実行速度で 5 倍以上も上回るものでした。

地球シミュレータの成功要因は、2 位以下のスカラプロセッサ超並列コンピュータと比較し、ベクトル計算機特有の高速高バンド幅のメモリシステムおよび単段クロスバーネットワーク接続による低レイテンシ（データ転送の遅延速度）によるものと分析されています。つまり、創造のプロセスにおける「コンセプト」と「モデル」とが正しかったことの証明です。

独立型プロジェクトでは、このように研究開発評価の前提として、プロジェクト自体の設計がきちんと行われているかどうかという点が大変重要になります。公的資金を巨額に費やす独立型プロジェクトでは、何よりもプロジェクトの目的・目標が意図した通りに実現しているかが最も重要であり、それを担保するためのコンセプトやモデル、正しいソリューションが選択されているかという点がチェックポイントになります。

逆に、プロジェクトの意義を事後的に解釈するのは厳に慎まなければなりません。プロジェクトが失敗だったかどうかの判断を分けるのは、プロジェクトの当初の目的が果たされたかどうかの 1 点です。むしろ、なぜ失敗したのかの原因解明を評価のプロセスによって明らかにし、新しいプロジェクトへの知見の集積をもたらすようにすることが研究開発評価の本来のあり方です。

4.1.2　プロジェクトの経済性

予算規模の大きい独立型プロジェクトでは、公的資金を投入することに対する説明として、プロジェクトの経済効果がどれだけあるか、またプロジェクトが経済効率的であるかが財務当局ならびに国民の重要な関心事項となります。

したがって、プロジェクトの全体としての目的がいかなる社会経済効果をもたらすか、コストパフォーマンスがどれくらいなのかについての客観的予測が評価の一方の課題となります。

社会経済的効果については基本的に次のような価値を貨幣換算することで示すことができます。

- 財務的便益：プロジェクトの成果が商品化され、市場において普及した際に得られる収益
- 経済的便益：従来の製品や技術と比較して、時間短縮、省エネ、省資源、である場合のコスト削減便益。もしくはコスト一定で比較した場合の大幅な性能向上による便益。
- 社会的便益：プロジェクトの成果が普及することによってもたらされる国民への便益。排出物削減、事故リスク削減、健康影響の低減等に関わる便益
- オプション価値：現状の事業環境下では財務的便益が得られなくても、為替レートや原油価格が大幅に変化した場合には価格競争力が得られる場合の価値。

図 4-3 は大型実験施設を例にした場合のプロジェクトの経済性評価の枠組みを示したものです。

出所：財団法人政策科学研究所作成資料

図 4-3 大型実験設備開発プロジェクトの経済性評価の枠組み

プロジェクトの経済性評価を行うには、まずプロジェクトによって何がもたらされるのかを定量的に評価する必要があります。

もっとも基本となるのはコスト要因です。大型実験設備の開発プロジェクトではプロジェクト全体の事業費はコストにはなりません。実験設備のライフサイクルに伴う運用費、バックエンドコストが考慮されていなければ、作った後のことは「知らぬ存ぜぬ」という無責任なことになってしまいます。また、民間企業の参画を条件としている場合には、民

間企業の持ち出し部分も費用として考慮しなければなりません。

次に、実験設備のような研究基盤の開発であれば、建設しようとしているインフラの仕様が定まっていることが前提条件となりますので、設備の性能がある程度把握可能です。つまり、設備の技術的特性を性能指標によってできるだけ客観的に評価する必要があります。性能指標を効率よく、システマティックに行う一般的な方法論は存在しません。技術領域における専門的知識と経験による勘といった「目利き」の要素も大変重要です。しかしながら、当該技術をどのように評価するかという観点がある程度共通の認識としてあれば、その代理指標を専門家が選ぶことは可能です。例えば「高性能」といった場合には「出力」や「応答速度」などが考えられるでしょう。

また、設備の用途についても評価する必要があります。利用計画に基づいた個々の研究プロジェクトの成果利用の面から、あるいは産業利用による民間の研究開発促進の効果、実験設備が所在する地域への効果（雇用、財政等）を示すことも重要です。

最終的には、そのような個々の研究開発プロジェクトがもたらす社会的便益、経済的便益まで評価することができれば経済性評価として申し分ないわけですが、これについては、設備を利用する研究開発プロジェクト群が具体的に定まってからでないと評価できませんので、評価時期に依存すると言えます。

こうした枠組みと経済性評価手法との関係について述べておきましょう。

まず、開発目標と仕様が定まっていればどの段階でも確実に適用できるのが費用分析（CA: Cost Analysis）です。上の図で言えば、実験設備の性能指標に対していくらの建設費用が掛かるかという指標が作成できます。こうした指標が他の類似設備の指標と比較したとき、どれだけのコストパフォーマンスを有しているかで一定の経済性評価が可能となります。

次に、実験設備を利用した研究開発プロジェクトが利用計画として具体的に選定されていれば、これらの研究開発プロジェクトが目標としている成果から社会経済的影響が推計されます。この時、社会経済的影響が物理量で示される場合には、費用効果分析（CEA: Cost-Effect Analysis）が適用されます。費用効果分析では、ある効果（例えば CO2 排出削減量）に対していくら費用が掛かるのかを明らかにします。費用はこの場合、当該設備のライフサイクル・コストが相当しますが、ある特定の研究開発プロジェクトが当該設備を全て利用する訳ではないので、何かしらの配賦基準（利用時間数等）で割り引く必要があります。また、個別の研究開発プロジェクト自体の費用も考慮しなければなりません。計算される指標は、○○円／効果（例：2000 円／t- CO2）という限界費用の形式で表されます。限界費用は他のプロジェクトや施策による効果と比較するための指標です。CO2 排出削減効果であれば、予算制約に応じて限界費用の低い対策から順次実施するという戦略が展開できます。

社会経済的影響が何かしらの価値評価を反映して貨幣換算価値で示される場合には、費用便益分析（CBA: Cost-Benefit Analysis）を適用します。実務的には、費用便益分析はもっ

とも労力を要する経済性評価になります。貨幣換算係数は非常に高度な推計手法を要します。また、その不確実性も大きいため、費用便益比も本来は一意には決まらないことに留意する必要があります。費用便益比は、効果（円）／費用（円）の形で表現されます。通常、費用便益比が１を超えていれば、社会的便益が費用を上回るため、プロジェクトの実施が社会的に望ましいという結論になります。

4.2　独立型プロジェクトの評価とは

独立型プロジェクトは予算規模が大きく、複数年度に渡って展開されることが一般的です。また、プロジェクトの全体が複数の研究開発ステージに分けられていたり、平行したサブ・プロジェクトを内包し、全体での研究開発目標の達成のために位置づけられています。

このような特徴を持つ独立型プロジェクトを評価する場合、プロジェクト全体をごろりと評価することはできません。プロジェクトをいくつかの要素（サブ・プロジェクト）に分け、その要素についての評価を積み上げることが重要です。つまり、独立型プロジェクトの階層化を行い、階層化された個々の要素について評価することが重要になります。評価対象をシステムとして捉えることで、当初の研究開発目標が達成されたかどうかについての評価が行われます。

また、独立型プロジェクトには多くの人材、組織が参画します。これらの統括的な実施体制や研究開発マネジメントがプロジェクトの成否を決定的に左右します。したがって、他の評価対象と異なり、研究開発実施者側の体制・マネジメントがもっとも重要な評価項目として検討されなければなりません。

これ以外にも、独立型プロジェクトは先端的な領域、フロンティア領域において研究開発が実施される場合が一般的であるため、新しい知の発見、ナレッジの拡大、ネットワークの形成、スピルオーバー等の副次的な効果が期待されます。こうした項目についての情報を整理することが非常に大切になってきます。

以下では、独立型プロジェクトの評価で特に重要な点を評価時期毎に指摘します。

4.2.1　事前評価の段階

4.2.1.1　プロジェクトの事前評価項目

独立型プロジェクトの事業計画はどのように策定されるでしょうか。多くのプロジェクトは、定型的もしくは非定型的な情報を下に、事業推進部署が原案を作成することになります。プロジェクトの原案に対して、有識者等による外部評価を実施し、プロジェクト原案を修正するプロセスが入る場合もあります。ここでの有識者とは多くの場合、プロジェクトにおいて展開される研究開発領域の専門家が内容をチェックするために配置されます。

出所：財団法人政策科学研究所作成資料

図 4-4　政策形成プロセスと事前評価プロセス

修正されたプロジェクト案が大型案件（総額約 300 億円以上の重点的課題）であれば、総合科学技術会議においてさらに評価が実施されます。ここでの議論で必要性について評価されることは稀でしょう。上の図 4-1 にある通り、プロジェクトのビジョンそのものはさまざまな情報を勘案し、プロジェクトの意義について十分に検討されているからです。

むしろ、評価専門委員会での議論では、独立型プロジェクトの予算案に対する有効性ならびに効率性に関わるポイントの明確化が事業推進部署に対して要求されます。特に、プロジェクトの目的に対する明確な目標水準の設定、ならびに研究開発方法、事業推進体制、事業推進上の権限の設定等が重視されます。

4.2.1.1.1　プロジェクトの目的・目標

独立型プロジェクトの目的を設定することはプロジェクトから得られるアウトカムを明らかにすることです。しかしながら、独立型プロジェクトでは、プロジェクトの目的について「○○構築を通じた次世代技術の習得」や「△△関連分野における科学技術人材の育成」といった派生的な部分の記述が目立ち、現実に投資の対象となる研究開発の中身や技術開発目標と掲げられたプロジェクトの目的とが対応していないような印象を受けるものが少なくありません。

プロジェクトの目的はけしてお題目ではなく、全体として何を達成すべきかを明確にコミットするものでなくてはなりません。そのためには、プロジェクトの中で実施される研究開発の目標水準や技術水準が目的達成のための最適な手段であることを担保するような記述になっていなければなりません。

独立型プロジェクトの場合には、予算年度が終了した時点で明確に目的達成の評価ができるよう、具体的なプロジェクトの目的を設定する必要があります。この目的が達成されたかどうかでプロジェクトの失敗・成功が客観的に分かるように事前評価の段階で措置しておくことが重要になります。誤解を恐れずに言えば、プロジェクトは失敗しても構わないのです。およそ研究開発を実施して失敗しないということはあり得ません。ただし、「なぜ失敗したのか」に対する教訓を得ることに対して努力がなされなければ本当に「税金の無駄遣い」で終わってしまいます。

出所：財団法人政策科学研究所作成資料

図 4-5 プロジェクトの目的･目標の設定

独立型プロジェクトの目的・目標の設定では、何を以ってプロジェクトの失敗／成功なのかをあらかじめ明らかにしておくことがプロジェクト・マネジメント上もっとも重要なことです。意義と目的を混同してはいけません。

4.2.1.1.2 プロセスの設計

研究開発目標を達成するための手段として、研究開発や事業実施の方法を明記しなければなりません。このとき、目標に対する手段としてなぜそれが最適なのかについての判断情報が必要になります。

研究開発目標が設定されたら、いきなり研究開発方法の各論に移るのではなく、プロジェクトのモデルが検討される必要があります。サブテーマの構成やサブテーマ毎の予算規模、

全体スケジュールなどを、全体最適の観点と持続可能性の観点から検討します。

プロジェクトのモデルが設定されたら、ようやくソリューションについて検討します。技術シーズとして何が採用可能か？、研究開発方法等のプロセスは何が最適か？、研究開発方法等の見直しも含め、サブテーマの目標を達成するためにどのようなスケジュールで研究開発を実施すべきか？といった内容を詰めていきます。

出所：財団法人政策科学研究所作成資料

図 4-6 プロジェクトのプロセスの設計

プロジェクトによってはコアとなる技術シーズを発展させることでビジョンを描く場合も想定されるでしょう。しかしながら、目的・目標表現ならびにプロセスの設計では、常に、「目的＝手段」関係を優先させなければなりません。そうすることで、プロジェクトの意義から個別の研究開発課題までを一貫した論理で説明することができるからです。独立型プロジェクトの事前評価では何よりも理由付けが重視されます。逆に、こうした論理展開が難しい場合には発展させようとしているシーズに問題があると考えなければなりません。

4.2.1.1.3 体制・マネジメントの設計

ソリューションのレベルまで詰めることができたら、後は具体的な研究開発の実施主体を割り当てます。また、プロジェクト全体の統括的立場の責任者（個人名が望ましい）を指定することが重要です。統括的立場の責任者は、プロジェクト執行上の全権限を持っていることが重要です。

4.2.1.2　プロジェクトの構造化とロードマップ

大型研究開発案件である独立型プロジェクトでは、プロジェクト内の各研究開発やその他のプロセスを構造化して、プログラムのように設計する必要があります。つまり、プロセス別に目標を定め、プロジェクト全体の目的・目標に対してどのような手段（寄与）関係にあるのかを明記する必要があります。

図 4-7 は米国エネルギー省のプログラム、「炭素隔離プログラム（Carbon Sequestration Program）」のうちの一つのプロジェクトである、炭素捕捉（Carbon Capture）研究開発プロジェクトの内容を構造化したものです。

出所：米国エネルギー技術研究所（NETL）"Carbon Sequestration Technology Roadmap and Program Plan 2006"より、Carbon Capture Project の概要図

図 4-7　大型プロジェクトの構造化とロードマップの例

図 4-7 では、プログラム（DOE の政策）に対して炭素捕捉研究開発プロジェクトがどのように位置づけられ、プロジェクト内の各サブ・プロジェクトが目的別に配置されていま

す。全体の流れとしては、エネルギー転換プロセスの中で取り組むべき研究開発（緑字）と、本流である炭素捕捉プロセスの研究開発が推進され、研究開発の進捗状況によって最適な組み合わせや技術の切り替えなどを検討する段階（cross cut pathway）を経て、プログラムの達成目標にコミットするという展開になっています。その他にも、上位政策であるプログラムの補助事業によって民間が取り組むべき課題（赤字)、プログラム期間中に呼び検討されている課題（青字）が補完的に配置され、当該研究開発プロジェクトだけで目標年次における CO2 排出量の削減が達成されるわけではないということが理解できます。

この事例で重要な点は、最終的な目標達成年次は定めていても、プロジェクト内の研究開発テーマについて詳細なスケジュールは設定されていないことです。それはプロジェクトの最終段階で「切り替え経路（cross cut pathway)」が用意されていることからも分かるように、研究開発課題が必ずしも期待通りに実施されるという保証がないため、予め複数の代替的手段を同時に走らせて、ある時期に取捨選択をするという意思決定プロセスを組み込んでいるからです。

つまり、プロジェクトの初期段階ではあらゆる可能性を模索しつつも、一定の時期にモノにならない研究開発課題については、積極的にコントロールし、最終的なプログラムの目標に寄与するものだけを残すという、「意思決定→実行」の創造のプロセスが機能しています。このプロセスを反映して、米国エネルギー技術研究所では、上のプロジェクト概要図を毎年更新しています。

上位のプログラムの目標年次は、プログラムのロードマップによって定義されていますが、科学技術先進国のロードマップは基本的にアウトカム・ベースで設定されます。技術開発だけでなく、産業界のインセンティブや市場トレンドを想定しつつ、相互に補完的な役割を期待してプログラムの達成目標年をプロットします。したがって、その下にぶら下がる研究開発プロジェクトでは、プログラムのロードマップに従って、プロジェクト・マネジメントが実施されことになります。

我が国でも経済産業省をはじめとして、さまざまな技術ロードマップが作成されています。独立型プロジェクトの開発目標は、省庁の戦略目標であるロードマップとリンクが取れていなければなりません。

一方で、ロードマップ自体の改善も不可欠です。アウトカム・ベースではなく、シーズ側からの発想でロードマップを作成している場合には、全体最適が図れないことと、代替的な研究開発プロセスをコントロールできないという欠点があります。つまり、そこに描かれている目標年次は絵に描いた餅になってしまいます。本来、何かしらの問題解決のための手段であるはずの研究開発が予算執行上の目的になってしまうという悪循環に陥りやすくなってしまいます。

4.2.1.3　実施体制・責任の明記

プロジェクトを構造化した後は実施体制・責任を明記します。先の DOE の炭素捕捉研究開発プロジェクトでは、複数の研究開発機関（公的研究開発機関や民間の研究所）が参加しています。下の表 4-1 では、構造化された各研究課題毎に実施機関が設定され、研究内容（description）、組織の持つ技術的優位性（Advantages）、研究開発目標（challenges）、研究開発のアウトカム（R&D pathway）が記されています。

表 4-1　大型プロジェクトの実施体制・責任の記載例

	Description	Advantages	Challenges	R&D pathways
Hydrates *Status:* Laboratory scale *R&D by:* Simtech LANL	At high pressure and low temperature (200 psi, 4 °C) CO_2 and water form solid macromolecular structures called hydrates. The solids are separated from the liquid stream and then heated to break down the hydrates and recover CO_2.	• Has the potential to produce captured CO_2 at several hundred psi of pressure.	• Requires refrigeration energy to counteract heat of hydrate formation (ΔH_F 600-1,400 Btu/lb CO_2 captured). • Cold spots in the hydrate formation reactor can cause ice formation and operational problems.	• Additives to speed up the hydrate formation reaction and enable 90% CO_2 capture. • Heat exchanger design to ensure uniform temperatures in the hydrate reactor.
Enzymatic CO_2 Sorbents *Status:* Laboratory scale *R&D by:* Carbozyme CO2Source	Enzyme-based system achieves CO_2 capture and release by mimicking mammalian respiratory mechanism.	• Fast kinetics lower system size and cost. • Resistance to SO_x and NO_x. • pH swing based operation offers potential to produce CO_2 above atmospheric pressure.	• 100 °F operating limit and exothermic CO_2 sorption reaction require cooling of flue gas. • Entrained solids in flue gas from coal boilers mayblock membrane channels. • Possible sensitivity to acid gases.	• Develop and test thermal management approaches • Demonstrate insensitivity to SO_x, NO_x, and acid gas
Amine-enriched Sorbents *Status:* Laboratory scale *R&D by:* NETL in-house	A carbon material with amine compounds fixed upon it exposed to a CO_2-rich process stream. The amine sites absorb the CO_2. The temperature of the material is raised to release the CO_2.	• High storage capacity (4 moles CO_2 per kg solid sorbent) and use of tertiary amines allows potential for lower energy required per CO_2 captured.	• Relative to a liquid, it is difficult to lower and raise the temperature of a solid material – can cause slow de-sorption rates. • Small diameter particles can cause high pressure drop across absorber.	• Improve sorption capacity, kinetics • Improve stability, resistance to poisons • System design for gas/solid contact/heat transfer • Multi-pollutant capture
Ionic liquids *Status:* Laboratory scale *R&D by:* U. of Notre Dame Sachem; Merck	Ionic liquids are a broad category of organic chemical compounds consisting of anionic and cationic components. They can dissolve gaseous CO_2 and are stable at temperatures up to several hundred degrees C.	• In a gasification application, can avoid cooling down syngas and heating it back up. • Physical solvent requires little heat for CO_2 recovery.	• Room temperature ionic liquids are not manufactured commercially and are expensive ($350 - $2,000/kg). • High viscosity may hinder usefulness.	• Develop better understanding of CO_2 dissolution mechanism • Screen materials for CO_2 solubility, viscosity, thermal stability • Conduct toxicity tests

出所：米国エネルギー技術研究所（NETL）"Carbon Sequestration Technology Roadmap and Program Plan 2006"より抜粋

4.2.1.4　費用分析の実施

事前評価段階で研究開発費の総額を見積もるのは大変難しいことです。特に予算規模の大きい独立型プロジェクトでは、挑戦的な目標性能を実現する装置等の開発にいくら掛かるを事前に明らかにすることはほとんど不可能といっても良いでしょう。

ただし、完成するであろうシステムの建設費や維持費用、総業費等は既存のデータからある程度見積もることができます。研究開発費について具体的な見積額を弾くことができなくても、成果物であるシステムのライフサイクル・コストを想定することは最低限必要な作業であると言えます。

ライフサイクル・コストを推計するためには運用のための想定シナリオが必須となります。例えばスーパーコンピュータの開発においては、完成直後の運用局面では NLS（National Leadership Super-computer）として国内外最高水準の計算機としての役割を果たすことが期待され、その後、後発のスーパーコンピュータの性能が追い越した段階で、NIS（National Infrastructure Super-computer）としての運用局面に移行します。現在稼働中の地球シミュレータはすでに最高性能計算機がもたらす計算科学の先導者としての役割を終え、科学技術インフラとして大学や民間企業を含む研究機関の共同利用施設として機能しています。段階的な運用計画も想定した上で、独立型プロジェクトは実施されなければなりません。

研究開発費の見積については、過去もしくは海外の同種のシステムの開発にどれだけコストが掛かったかについてのサーベイが得られていれば、一定の目安にすることができます。他の事例と比較して、新型のシステムを開発する際に考慮すべき要因（性能指標、立地条件、物価水準、研究開発のリードタイム、投入資材量、研究者数等）を加味することで、妥当な研究開発費用の幅を示すことができるでしょう。

コラム：我が国の科学技術予算の問題点 『冗長性と継続性』

我が国の科学技術予算には冗長性がない、と指摘されます。一番分かりやすい事例は宇宙開発です。科学技術衛星を１基打ち上げるプロジェクトが実施される場合、失敗した場合のバックアップコストを予算として認めていないことが問題視されています。つまり、予備のロケットと衛星を用意して、発射時期をずらして待機させておく、というような運用体制になっていません。バックアップがなされていれば、１基目の打上に失敗してもすぐに原因のフィードバックをかけることができ、その後の開発に活かすことができます。

また、科学技術予算の連続性も問題視されています。大型プロジェクトが一度成功すると、すぐには継続的な研究開発予算が認められません。しかしながら、技術動向に関する調査研究や得られた要素技術等の発展研究を継続的に展開しなければ、次の「玉」を出す時に最初の構想から手がけなくてはなりません。大型プロジェクトにありがちな、「初年度のシステム設計とフィージビリティスタディーに○○億円」は本来必要のないコストだったのかもしれないのです。

「ワンショット」で研究開発プロジェクトを遂行せざるを得ないため、また、機会が限られているために必要以上に高水準の技術開発に向かいやすいという傾向になります。こうした予算制約は、逆に研究開発のリスクとコストを高め、さらに失敗した場合には全て無駄になってしまうという綱渡りのような研究開発を実施者に課すことになります。

目先のコストが重視された結果、返ってコスト高になってしまうというジレンマが科学技術政策には存在します。科学技術関連予算においてはリスクを許容する仕組みの導入こそが中長期的に研究開発の失敗の可能性とコストを下げることに結びつくのです。

4.2.2　独立型プロジェクトの中間評価

独立型プロジェクトの中間評価は、実施計画を遂行する上で非常に重要な位置づけを占めます。プロジェクトの予算や研究開発計画の見直し、あるいは撤退を検討しなければならないからです。中間評価では、計画に対して進捗状況がどのようになっているか、あるいは想定したコストをオーバーしていないか、適切なマネジメントが実施されているかという点が重要になります。

4.2.2.1　各サブ・プロジェクトの進捗状況

事前評価の段階において、プロジェクトは構造化され、いくつものサブ・プロジェクトが研究開発計画に従って実施されています。それらの実施状況をモニタリングすることが最初の作業になります。

技術開発プロジェクトでは、要素技術等の開発目標が個別に設定されていることでしょう。性能指標を個別の技術開発（サブ・プロジェクト）について設定し、中間評価時点の指標が開発目標に対してどの程度達成しているかを把握することになります。

4.2.2.2　予算執行状況

中間評価時点で、進捗状況と併せて把握する必要があるものに、予算の執行状況があります。現時点での性能を得るために、これまでいくら研究開発に投資されたか、また試行的に行った調査や研究開発にいくら投じているのか、こうした情報をサブ・プロジェクトの範囲で整理することで、今後の研究開発予算がどれくらい必要になるかの目安が出来ます。

研究開発予算は許されている限り使用するというものではなく、執行上の理由を必要とします。中間評価以降の期間で毎年度どれだけの研究開発予算が必要になるかは、研究開発のフェーズにも依存しますが、これまでの予算執行状況からある程度の予測を立てることができます。

4.2.2.3　意思決定・R&D マネジメント・体制

独立型プロジェクトでは、研究開発実施者のプロジェクト遂行上の意思決定の迅速さ、R&D マネジメントの妥当性、研究開発体制の明確さが求められます。ただし、これらの評価項目については、明確な判断基準を用意できません。中間評価時にこれらの評価項目を評価するには、次のような場合が考えられます。

1) 試行的、もしくは並行的に開発していた技術方式（サブ・プロジェクト）をプロジェクト全体の観点から採択・中止の判断した場合
2) 見直す必要のあったサブ・プロジェクトの実施主体（参画機関）に対して、その後のプロジェクトにおける関与のあり方を判断した場合
3) 全体のスケジュールに調整の必要があり、何かしらの意思決定を行った場合

つまり、意思決定の必要性がなく、計画通りに進んでいるうちはトップマネジメントや実施体制のあり方に問題があるわけではなく、何かしらの軌道修正や調整が図られた時に、その判断が最善であったかどうかが問われることになります。特に、並行的な研究開発を進め、プロジェクトのある段階においていずれかの成果を採択するような「切り替え経路（cross cut pathway）」が設定されている場合には、トップマネジメントの妥当性が評価されることになります。

4.2.3　独立型プロジェクトの事後評価

独立型プロジェクトの事後評価の最大の目的は、プロジェクトが成功だったか失敗だったかを判断することです。この総合的な評価が行われた後に、各評価項目を評価する意義が明らかになります。つまり、成功要因もしくは失敗要因の洗い出しが独立型プロジェクトの事後評価の本質的な行為になります。

4.2.3.1　目的・目標達成状況

独立型プロジェクトの事後評価で問題となるのは、「研究開発目標は達成できたが、事業の目的は達成できなかった」というケースが存在することです。このような場合、研究開発実施者側が最善を尽くして優れた研究開発成果を生み出したこと自体は高く評価されるべきですが、事業目的の設定の仕方が誤っていたかどうか、または事業目的を達成するための措置がプロジェクトに組み込まれていたかどうかについては厳しく評価されなければなりません。

事後評価の段階で成功か失敗かを議論することを避けるためにも、事前評価の段階でどのような状況がプロジェクトの成功と認められるかについて明らかにしておく必要があります。事業の目的とは意図した結果（アウトカム）が得られているかということですが、独立型プロジェクトの場合、意図する者は国に相当します。研究開発目標や技術開発目標を評価することとは別に、意図した結果を得られているかどうかについては、事業実施者（国）の政策能力と説明責任が問われています。

出所：財団法人政策科学研究所作成資料

図 4-8　事後評価の目的・目標達成状況とプロジェクトの成功／失敗

4.2.3.2　プロジェクトのマネジメント

事後評価の段階では、プロジェクトの成功・失敗が明らかになっています。この判断を前提に、マネジメントが有効であったかそうでなかったかが評価されます。

マネジメントに対する評価はプロジェクトの総合的な評価とセットになっています。いくらマネジメントが高く評価されても、結果が伴わなければマネジメントが有効に機能したとは言えません。逆に、マネジメントに見るべきものがなくてもプロジェクトが上手く行った、という解釈も成り立たないでしょう。プロジェクトを遂行する上での何かしらの努力が行われていたと考えるのが一般的だからです。つまり、事後評価の段階では、マネジメントに関する評価項目は総合的な評価と入れ子の関係にあると言えます。

出所：財団法人政策科学研究所作成資料

図 4-9　事後評価におけるプロジェクト・マネジメントの評価の位置づけ

事後評価におけるプロジェクトのマネジメントに関する評価のポイントには次のようなものが考えられます。

出所：財団法人政策科学研究所作成資料

図 4-10　プロジェクト・マネジメントの評価の観点

「総合化（synthesis)」の観点とは、様々な要素（サブ・プロジェクト等）から構成されるプロジェクトの全体最適を図っているかどうかを問うものです。「事業に関係のないサブ・プロジェクトがある」、「積極的な関与のない参加主体がいる」、「プロジェクト情報が散逸している」等の状況が発生していれば、それはマネジメントにおいて総合化の観点が欠落していると言えるでしょう。総合化はプロジェクトの実施期間全般において適宜行われる必要があります。

「正しい選択（trade-off)」とは、プロジェクトの途中段階で二者択一的な意思決定を迫られる状況において、トップの責任において望ましい選択がなされたかどうかを検証することです。こうした重要な決定を実行委員会などで多数決で決めているようでは、マネジメントの自主性が保たれているとは言えないでしょう。こうした意思決定はサブ・プロジェクトの成果を待つため、プロジェクト実施期間の中盤以降で行われます。

「優先順位（priority)」はプロジェクトの時間的制約・予算的制約の下で実施するサブ・プロジェクトに対してどのように優先順位を付けたかという観点です。これはプロジェクトのスケジュールや予算配分を調整するという行為ですが、研究開発計画がまだ十分に練られていない状態でスタートしたプロジェクトでは必須のマネジメントになります。優先順位付けは通常、プロジェクトの初期段階で行われます。

「循環・連鎖（spiral)」はサブ・プロジェクト間の効果的なコラボレーションを図るための観点です。ある研究チームのミーティングに必ず関係性のある他の研究開発チームのメンバーを加えるなど、研究者間の人的交流の側面が重視されます。こうした工夫はプロジェクト実施期間に限らず、研究者ネットワークを構築していくためにも必要な措置であるといえます。

「妥協（compromise)」はプロジェクトを計画期間内に終了させるために必要なマネジメント行為です。時間的制約、予算的制約を考慮して、実現可能なシステムを構築するために研究開発目標の見直しを行います。最初から妥協をするのでは挑戦する意味がないので、妥協に関わるマネジメントはプロジェクトの終盤に検討される必要があります。

以上のようなプロジェクト・マネジメントの観点を、評価対象である独立型プロジェクトの特性に合わせてカスタマイズし、評価項目・評価基準を作成することで評価が可能となります。

4.2.3.3　副次的成果

事後評価の段階においては、プロジェクトからもたらされた副次的な成果も数多く生み出されています。目的・目標としていたシステムの構築以外に、科学技術的に価値のある成果（論文や学会発表)、産業技術へのスピルオーバー（技術的波及）の可能性、ネットワークの構築（研究者コミュニティの形成）等、評価時点で把握可能な成果に関する情報を悉皆的に収集する必要があります。

プロジェクトの活動の一環として、これらの情報を Web 上でアーカイブ（書庫）として管理するのも有効な手段です。プロジェクト終了後も研究者ネットワークの中枢として、研究開発成果に関する情報を更新・管理することができます。この良い事例には、第５世代コンピュータプロジェクトアーカイブス＜http://www.icot.or.jp/＞などが挙げられます。

4.2.4　独立型プロジェクトの追跡評価

ナショナル・プロジェクトについての最終的な評価は追跡評価によって明らかになると言っても良いでしょう。巨額を投じたプロジェクトがいかなる社会的影響をもたらしたのかについて、フォローアップすることでプロジェクトの最終的な評価が固まります。

大綱的指針では、次のように追跡評価を位置づけています。

> *研究開発においては、終了後、一定の時間を経過してから、副次的効果を含め顕著な成果が確認されることもまれではない。こうした点を踏まえ、学会等における評価や実用化の状況を適時に把握し、必要に応じて、研究開発施策、研究開発課題等について追跡評価を行い、成果の波及効果や活用状況等を把握するとともに、過去の評価の妥当性を検証し、関連する研究開発制度等の見直し等に反映する。*
>
> 大綱的指針 p.9

4.2.4.1　追跡評価の実施時期

追跡評価の時期については、大綱的指針では具体的に言及していませんが、欧州等で実施されている研究開発プログラムの本格的な事後評価（ex-post evaluation）がプログラム終了から 5 年経過を目安に実施されていることからも、終了時点から 5 年目くらいが妥当と考えられます。

追跡評価の実施時期がそれ以上長ければ、今日的意義から評価対象を眺めてしまうため、研究開発評価というよりは歴史的分析になってしまいます。5 年程度であれば、関係者の記憶も鮮明であるため、プロジェクトに携わった研究者から確度のある情報をヒアリングすることが可能です。

4.2.4.2　追跡評価の方法論

追跡評価については、評価のプロセス（情報収集＝分析＝評価）そのものが政策研究として専門的な学術領域を占めています。したがって、評価対象に対して普遍的に適用できるアプローチは存在しません。通常の学術研究と同様に、仮説検証のために最適な方法論を選択するというアプローチが採用されます。

したがって、追跡評価のためには、政策研究のための仮説を構築しなければなりません。この仮説が追跡評価の評価項目に相当すると考えて下さい。

独立型プロジェクトの追跡評価に関する問題仮説の例

1) 評価対象となる独立型プロジェクトは研究開発基盤としての役割を果たし、良質な研究開発成果をその後も生み出し続けている。
2) 評価対象となる独立型プロジェクトは高度な産業技術基盤を提供し、競争力のある工業製品を開発するのに大きく貢献した。
3) 評価対象となる独立型プロジェクトの経済的効果は大きく、イノベーション促進の観点からも公的資金を投入した意義が大きかった。
4) 評価対象となる独立型プロジェクトで開発された要素技術は、事業の目的であった○○分野だけにとどまらず、他の分野においても活用され、大きな技術的波及をもたらした。
5) 評価対象となる独立型プロジェクトでは、国内外の研究者コミュニティの形成に役立ち、そのネットワークから多くの先導的な研究者を排出することになった。
6) 評価対象となる独立型プロジェクトでは、世界の科学技術の進展に大きく寄与し、その後の研究開発の礎となった。

このような問題仮説を検証するための方法論にはさまざまな政策研究のアプローチが考えられますが、それらは高度な専門的知識やツールを必要とします。主に社会科学分野の研究者やアナリストと呼ばれる高度専門家集団の支援を受け、初めて本格的な追跡評価が実施できるでしょう。

追跡評価をシンクタンク等の外部支援機関に調査研究案件として委託する場合、上のような問題仮説を委託元である評価関連部署が設定し、それに対して有効なアプローチを提案してきたところを採択するという手続きが追跡評価に対してはもっとも有効かつ現実的なリソースの使い方となることでしょう。

5　機関の評価

大綱的指針における機関評価の記述には、「研究開発機関等の設置目的や研究目的・目標に即して、機関運営と研究開発の実施・推進の面から行う」とあります。

「機関運営と研究開発の実施・推進の面から行う」とはどういう意味でしょうか。評価の対象となる機関全体の運営状況と、そこで実施されている個別の研究開発活動（資金配分機関であれば実施事業）を同時に評価するという意味です。しかしながら、機関運営と実施している活動との間には一定の関係性があり、両者を独立に評価することはできません。機関運営と機関で実施されている諸活動とを結びつける方法論や組織モデルがあって初めて機関全体を評価することが可能になります。

本章では機関評価における課題を整理したうえで、より実効的な機関評価を実施するための方法論について学びます。

5.1　機関評価はなぜ難しいのか

ある組織を評価するといった場合、通常は定型的な枠組みに従ってある時点におけるデータを見ることになります。財務諸表が良い例でしょう。財務諸表は企業の一年間の経済活動に関する成績表です。企業は経理を中心とした経営内容を財務諸表のかたちで社会に公表し、利害関係者の監視を受けることで説明責任を果たすことができます。

財務諸表は企業の活動の結果ですから、活動に対する努力が実を結べば、自ら財務諸表上の経理指標は改善されることが期待されます。企業の利害関係者は結果としての数値を毎年見ることで、企業の全体のパフォーマンスをチェックすることができます。

ところが、財務諸表のデータは企業の経営者にとって企業の健康状態を教えてくれても、どうすれば健康体になれるかについての情報を教えてくれません。例えば、ある業務を効率的に実施し、業績を延ばすと同時にコストも下げるという目標に対して通常の企業会計データは役に立たないのです。つまり、ある業務を中心としたコストデータの管理ができなければ、業務の改善につながらないということです。

業務の改善のためには管理会計とよばれる会計手法を導入する必要があります。一般的な経理データから、特定の業務のために費やされた直接経費、間接経費を何からの基準を用いて配賦し、業務の原価を計算します。

管理会計手法を導入するためには、業務がどのような活動から構成されているかという業務分析、活動のために費やされるコストを計算するための方法論の開発、通常の経理データを必要に応じて読み換えることのできる経理システムの導入など、大変なマネジメント費用が発生します。しかし、それだけの投資を行うことで、ようやく業務の改善に対する評価が可能になります。そして、改善が結果として企業の財務諸表における経営指標の向上に結びつくのです。

翻って、公的研究開発機関、資金配分機関の評価について考えてみましょう。独立行政法人通則法で定められているところでは、「各事業年度に係る業務の実績に関する評価」（第32条）ならびに「中期目標に係る業務の実績に関する評価」（第34条）の２種類の評価を実施なければなりません。特に、後者の「中期目標に関わる業務の実績に関する評価」では、機関の設置目的（ミッション）から展開される具体的な活動領域と、実施している業務の関係性を明らかにした上で、中期計画期間の実施業務がどのようにミッションに貢献しているかを評価しなければなりません。

機関評価で問題となっているのは、実施している研究開発事業等の評価と機関全体のパフォーマンスの評価を結びつける方法論がないままに評価の重複が起こっていることです。具体的には次のような問題が挙げられるでしょう。

- 機関の設置目的（ミッション）と、それから展開される活動目的・目標に対して、実施している事業の目的・目標が直接的に結びついていない。（実施事業の評価結果を機関評価に利用できない）

- 実施事業ならびに機関全体のパフォーマンスを測定するための方法論が導入されていない。（業務分析が出来ていない）

こうした問題を考えるときに、先に紹介した管理会計的な考え方が大変役に立ちます。機関を評価するためには、その構成要素をどのように分類し、いかに再構築するかということが大変重要になります。次節からは、そのプロセスを詳しくみていくことにしましょう。

5.2　機関評価の枠組み

5.2.1　機関の何を評価しなければならないのか？

民間企業はあるビジネスモデルにしたがってキャッシュフローを確保することで経営が成り立っています。したがって民間企業を評価する場合には、財務的な指標を観ることで、実施している事業の効率性等を企業経営の総体として評価することができます。

ところが、公的研究機関や資金配分機関は公的資金を運用して事業を展開しています。機関の依って立つ理由は、独立行政法人通則法や各独立行政法人の個別法によって明らかにされているように、何らかの公共的なミッションを果たすことにあります。公的研究機関や資金配分機関はその財務状況をみて評価が可能なのではなく、機関のミッションに沿った個別具体的な事業展開について、その改善努力や事業の効果を国民の観点から評価しなければなりません。

つまり、公的機関の評価では機関の「状態（state）」を診るだけでなく、機関の「活動（activity）」についても診る必要があり、その「活動」は組織のミッションに結びついている必要があります。

5.2.1.1　ミッションから活動まで

旧国立研究機関や資金配分業務を実施していた特殊法人は、独立行政法人通則法の制定（1999 年）を機に独立行政法人として逐次移行していきました。独立行政法人は中期目標・中期計画を作成しなければなりません（独立行政法人通則法第 29 条，第 30 条）。

中期目標では、独立行政法人の設置目的（ミッション）を果たすために、目標期間中に以下の点を明確に定める必要があります。

1)　中期目標の期間（3 年～5 年）
2)　業務運営の効率化に関する事項
3)　国民に対して提供するサービスその他の業務の質の向上に関する事項
4)　財務内容の改善に関する事項
5)　その他業務運営に関する重要事項

この中で、中期目標を実現するための中期計画を策定するに当たって記述に工夫が必要な事項は 2)と 3)です。なぜならば、当該独立行政法人のミッションと国民に提供するサービスや業務を具体的に展開し、その上で個別業務の効率化について述べなければならないからです。さらに、これらの展開がきちんと実施されないと、毎年の年度計画が具体的に記述できないことになります。

独立行政法人のミッションから具体的な活動に至るまでの連鎖がしっかりと出来ていないと、中期計画・年度計画は策定できません。そして、独立行政法人評価は中期計画、年度計画の目標達成度を評価する仕組みになっているので、評価自体が成り立ちません。

ミッションから活動に至るまでの連鎖を作り上げるということは組織の理想モデルを構築するという行為に相当します。

一方で、現状の独立行政法人の組織構造がどのようなシステムを形成しているか（組織の実態はどのようになっているか？）を整理して、組織の実態モデルを用意する必要があります。組織の実態モデルは単に組織図や機構図を見れば良いというものではなく、組織内部の部署がどのような機能を（相互に）果たしているか、という観点から捉える必要があります。

出所：財団法人政策科学研究所作成資料

図 5-1　組織の理想モデルと実態モデル

上の図 5-1 は組織の目的に応じたさまざまな機能に着目して構築した組織の理想モデルと、それと対比した形で表した組織の実態モデルを表したものです。理想モデルでは、組織のミッション、目的（objectives）に応じて事業等の活動（activity）が位置づけられ、そのための機能（function）が明確にされ、各ユニットは担当する機能が明確になっています。この状態を「組織のプログラム化」と呼びます。

一方、組織の実態モデルでは、必ずしもミッションに対応した事業目的を持たず、事業そのものが自己目的化しています。また、目的を達成するための活動や機能が明確に位置づけられておらず、各ユニットも複数の業務を担当するような構造になっていて、各ユニットの業務に対する貢献度も分かりません。

こうしたモデルは組織の目的や活動、機能をどのような事業レベルで定義するかによって描かれるパスが異なりますが、理想モデルと実態モデルとを分ける最大の特徴は、上位

目的や上位の活動に対して手段となる活動やユニットが独立な要素として機能しているかどうかという点です。実態モデルでは、事業目的やユニットの業務に重複があり、効率的な運営がなされていないと考えられます。

組織の理想モデルと実態モデルはどのような組織においても通常は乖離しており、また、接近している場合にも、常に理想モデルの再構築と、実態モデルをその理想モデルに近づけていくというシステム・ダイナミクスが働いています。組織の理想モデルに実態モデルを近づける行為を本来「組織経営」と呼びます。

中期計画・年度計画で描かれる組織像や事業のイメージは「理想モデル」でなければなりません。現実の組織が上に挙げたような「実態モデル」であったとしても、理想モデルに近づける努力を行う必要があります。

5.2.1.2　機関のアウトカム・インパクトとは？

機関評価との対応で考えれば、機関のアウトカムやインパクトとは何かという問題を考慮する必要があります。先のミッションから活動の連鎖に至るモデルと、評価の観点から整理した活動からアウトカムやインパクトに至る連鎖とを比較してみましょう。

出所：財団法人政策科学研究所作成資料

図 5-2　組織のプログラム化と機関評価の観点

図 5-2 は組織のミッションから個別の活動までに至るプログラム化の方法（プログラム・デザイン）と、組織の個別の活動から影響経路を追っていく機関評価の観点とを対比して表したものです。

図 5-2 の左側から見ていきましょう。ビジョン（vision）とはプログラムや組織が究極的に達成したい事象、あるいは状況の変化を意味し、組織の意思（will）を表明したものです。

ただし、その組織が存在しなければ実現しないという類のものではなく、ビジョンとして描かれる事象や状況は他の主体やプログラムの影響を受けて成立します。したがって、中期目標・中期計画等における表現としては、「～に資する」、「～に貢献する」、「～を目指す」といった記述の仕方になります。ビジョンとはその実現のための行動を伴うものであり、行動を伴わないものは「祈り（prayer）」でしかありません。ビジョンが達成された状態を通常、「目的の地（Goal）」と呼びます。

ミッション（mission）とは、そのプログラムや組織が存在することで確実に影響を及ぼすと言うことのできる事象、あるいは状況の変化そのものを意味します。したがって、目的表現としては、「～を果たす」、「～を実現する」といった表現が用いられます。ミッションはプログラムや組織が目的として表明したことに対する結果責任を負うものであるため、この段階までが自己評価の責任の範囲内にあるものと考えることができます。

目的（objectives）はミッションの実現に貢献すると考えられる活動の結果のイメージです。組織で展開されるさまざまな活動（事業、業務）が何をもたらし、ミッションといかに結びついているのかが示されなければなりません。つまり、目的を上位ミッションに対する手段として位置づけることで、連鎖を繋げていきます。中期計画等では「○○を通じ、～を果たす」といった表現の「○○」に相当する部分となります。

目標（target）は目的実現のために達成しなければならない活動の水準を表しています。年度計画等において、「平成△△年までに○○の効率を□□%達成する」という表現で客観的に示されます。

プログラム・デザインは、目標を達成する手段である具体的な活動（activity）を位置づけることまでが最小限の作業範囲となります。組織の場合であれば活動は事業や業務に相当します。事業や業務に投入されるリソースがどれくらいか、また、どのような体制・マネジメントで実施すべきかといった部分を設計することは通常、プロジェクト・デザインと呼ばれます。プログラム・デザインとプロジェクト・デザインが結びついて初めて、「サービスの提供や業務の質の向上」と「業務運営の効率化」とを同時に議論できます。

個別の業務に関して「業務運営の効率化」が把握できれば、中期目標の検討事項である「4) 財務内容の改善に関する事項」がリアルな数値の積み上げによって検討することが可能となります。

他方、ビジョン、ミッション、目的、目標が評価の観点からどのような対象として位置づけられるかを示しているのが図 5-2 の右側になります。個別の活動が生み出したものがアウトプット（outputs）であり、これは目標と対応する形で評価されます。また、目的と対応するのが短期的なアウトカム（short-term outcomes）になります。短期的なアウトカムの把握までは事業レベルの評価（独立型プロジェクト、プログラムの評価）を活用することで得られます。

ミッションと対応するのが長期的なアウトカム（long-term outcomes）です。これは、機関が存在しなかった場合を非実施仮説にして、当該機関のサービスの提供がもたらした

効果を計測する試みになります。研究開発機関であれば、実施した研究開発の成果の総体（論文や特許）が科学技術領域においてどの程度のシェアを占め、同領域の発展に貢献したかを精査することで検証可能でしょう。資金配分機関であれば、研究開発助成や補助金の受け手（カスタマー）に対するアンケート等で、公的研究資金が果たした役割について評価してもらうなどの方法が考えられます。

ビジョンに対応するのはインパクトです。インパクトのレベルでは、我が国の科学技術水準に関する国際比較など、マクロ的な視点で評価することになりますが、当該機関のみの貢献に帰することはできません。機関評価においてインパクトまでの立証責任はなく、内外の科学技術政策に関する研究成果に委ねられることになるでしょう。むしろ、そのような研究成果を引用し、機関が提供するサービスの影響が結果としてどのように現れているかどうかを確認するにとどまると考えられます。

5.2.1.3　活動と組織構造

組織のプログラム化と評価の観点との関係を踏まえた上で、機関評価のために必要な措置は、組織の活動と組織構造に関する対応関係の把握です。

組織の理想モデルと実態モデルの説明で触れたように、現実の組織の活動はミッションから目的、目標、活動へと独立な関係性を保って展開されているわけではありません。まずは組織の中の個別のセクションがどのような業務を担当しているのかについて、業務分析を行う必要があります。

出所:平澤冷(2007),「独法評価への対応～製品評価技術基盤機構での経験から～」,日本原子力研究開発機構講演資料

図 5-3　担当部署と担当業務の把握：製品評価技術基盤機構（NITE）の例

図 5-3 は独立行政法人製品評価技術基盤機構（NITE）の業務と担当部署との対応関係を描いたものです。しかし、これらは主務とする業務を描いたものであり、それぞれの業務に何かしらのかたちで関与するセクションを業務別に展開すると、各セクション別の業務が 660 項目にも上りました（下図 5-4）。

出所：平澤冷(2007),「独法評価への対応～製品評価技術基盤機構での経験から～」，日本原子力研究開発機構講演資料

図 5-4　NITE の業務分析の結果

NITE では業務項目別に業務目標を定め、各セクションの業務に関わる費用を推計し、業務目標と対応させることで業務効率を計測するためのデータベースを 2 年間掛けて作成しました。

さらに、NITE では組織活動のプログラム化を実施しました。NITE の提供するサービスについて、アウトカムの観点から事業分野を類型化し、事業分野別のアウトカムとコストデータを対比することで業務効率の改善効果を数値に基づいて客観的に示しています。これらの改善効果を組織全体で積み上げることで、NITE 全体のコスト改善効果を示しています（次頁図 5-5・5-6）。

NITE の機関評価の経験[5]は現時点における独立行政法人評価の理想的な事例と言えます。しっかりとしたシステム設計の下、組織の業務分析、活動基準会計手法（ABC）による業務費用の推計、国民の視点（アウトカム）からの事業分野の類型化とミッションの再定義、機関全体のパフォーマンスの計測を行うことができたのです。

[5] http://www.nite.go.jp/hyouka.html

コストの把握 ― 化学物質管理分野のコスト構造

リスク評価に関する新たな情報の提供、化審法改正（届出等1．8倍）やNITEへのワンストップ化、化管法のPRTR届出の完全施行（届出1．2倍）などの業務増加に対して、業務に対応した機動的な人員配置、業務マニュアルの整備、外部委託の積極的な活用などによる効率的な業務実施により、コスト増加を抑制し、コストパフォーマンスを向上させた。

コスト改善取組のポイント

総合情報提供（CHRIP、リスク評価（NEDOプロ））

・CHRIPについて、英語版作成や構造式検索対応の追加実施による業務増加に対して、外部委託の効率的実施や、データ信頼性確保のための作業をマニュアル化するなどの効率化により、データ収集を開始した平成14年度に比べ1%のコストを削減した。

化審法関係業務

・平成16年度から 3省DBの構築開始により経費が増加しているが、期初から継続している業務については、3省の届出窓口一本化や化審法の改正に伴う業務増加（届出等平成14年比約1．8倍、相談平成15年比約3．7倍）などによる大幅なコスト増加を、熟練職員の集中的配置等の機動的な配置等により効率化し、全体で36%のコスト増加に抑制した。

化管法関係業務

・ 事務処理軽減などのためのシステム改修、行政監察指摘への対応のためのコストが増加。しかし、集計業務については、マニュアル化などによる集計作業の効率化、電子届出の推進、外部委託の効果的・効率的実施などの継続的な努力により、化管法のPRTR届出完全施行による届出対象事業者の増加（約1．2倍）にもかかわらず、集計業務を開始した14年度に比べてコストを10%削減した。（13年度はシステム開発中）

出所：平澤冷(2007),「独法評価への対応～製品評価技術基盤機構での経験から～」，日本原子力研究開発機構講演資料

図 5-5　NITE における事業分野別のアウトカムとコストデータの対比

コストの把握　―　nite全体のコスト構造

●期初からの継続業務については、5年間トータルで9．2%のコスト削減を達成。
●第一期中に開始したBRC業務では中期目標で想定していなかった海外微生物収集事業の展開など当初の想定外の大きな成果を上げ、MLAP認定業務、講習関係業務では短期間での業務立ち上げに加え業務量の大幅な変動への着実な対応を図ることにより、高いパフォーマンスを示すとともにコストの増加抑制を図った。
●期初からの業務においても、ゲノム解析業務での独自の前処理方法の開発や独自ツールの開発による効率化の追求、化審法業務での三省共通事務局化による業務量増加への効率的対応、認定業務での複数のプログラムでの横断的な内部職員の活用による業務量変動の吸収、製品安全業務での事故情報収集・検証件数の増加という環境下での社会的インパクトの高い情報発信とコスト削減の両立など、高いパフォーマンスの提供とコスト削減を両立した。

コスト改善取組のポイント

○バイオテクノロジー分野

・収集菌株の迅速な簡易同定方法の確立
・独自の前処理方法の開発や独自ツールの開発によるゲノム解析業務での効率化の追求

○化学物質管理分野

・CHRIP構築でのデータ信頼性確保のための作業のマニュアル化による効率化
・化審法業務での三省共通事務局化による業務量増加に対し、職員の機動的配置等による効率的対応

○認定業務

・複数のプログラム間での横断的な内部職員の活用による業務量変動の吸収と外部審査員の育成・活用の推進

○人間生活福祉分野

・製品安全業務において事故情報収集・検証件数が大幅に増加する中で「誤使用ハンドブック」等の社会的インパクトの高い情報発信とコスト削減を両立
・受講者変動の著しい講習業務において増加コスト抑制とサービス向上の両立を達成。

出所：平澤冷(2007),「独法評価への対応～製品評価技術基盤機構での経験から～」，日本原子力研究開発機構講演資料

図 5-6　NITE 全体のコスト改善効果

5.3　機関評価の方法論

5.3.1　業務分析の実施

(1)業務対応表の作成

機関評価のためには、機関で実施している業務を全て洗い出します。そして、業務を担当している部署（セクション）を業務毎にリストアップします。この作業では、次のような業務対応表をデータベース形式で作成することが望ましいでしょう。

表 5-1　業務対応表の例

NO.	事業分野名	事業名	業務名	業務内容	担当セクション	所属事業部
T-00001	○○推進分野	△△事業	××業務	・・・・	□□推進室	○○事業本部
T-00002	○○推進分野	△△事業	××業務	・・・・	◇◇事業所	○○事業本部
T-00003	○○推進分野	△△事業	⋮	⋮	⋮	⋮
T-00004	○○推進分野	△△事業	⋮	⋮	⋮	⋮
T-00005	○○推進分野	△△事業	⋮	⋮	⋮	⋮
T-00006	○○推進分野	△△事業	⋮	⋮	⋮	⋮
T-00007	○○推進分野	⋮	⋮	⋮	⋮	⋮
T-00008	○○推進分野	⋮	⋮	⋮	⋮	⋮
T-00009	○○推進分野	⋮	⋮	⋮	⋮	⋮
T-00010	○○推進分野	⋮	⋮	⋮	⋮	⋮
T-00011	⋮	⋮	⋮	⋮	⋮	⋮
T-00012	⋮	⋮	⋮	⋮	⋮	⋮
T-00013	⋮	⋮	⋮	⋮	⋮	⋮
T-00014						

出所：財団法人政策科学研究所作成資料

業務対応表はボトム・アップで作成すると良いでしょう。業務を統括する最小単位のセクション毎に、年間を通じて発生する業務をすべてリストアップしてもらいます。次に、その業務がどのような事業に関連したものかについて対応付けします。

さらに具体的な事業名が、機関においてどのような事業分野に位置づけられるかについて対応付けします。この時、事業分野に関して、ミッションを実現するための手段として位置づける（アウトカムの視点から位置づける）ことが出来れば、組織のプログラム化に役立ちます。

出所：財団法人政策科学研究所作成資料

図 5-7　事業分野の再設定と組織のプログラム化

業務対応表を各フィールド別に集計した場合、次のような対応関係の把握が可能になります。

出所：財団法人政策科学研究所作成資料

図 5-8　業務対応表の集計結果

担当セクション別に集計した場合、担当セクションから実施している業務の一対多関係が導かれるでしょう。逆に業務別に集計した場合には担当セクションへの多対多関係が見えてきます。業務から事業への対応も多対多関係であることが分かります。

業務分析では、組織の活動の基本単位を業務に求めた場合、事業に対してどのような重複や担当セクションの関わりがあるかという部分が組織の実態モデルとして見えてきます。この関係性を組織の理想モデルに近づけるためには２つの方法が考えられます。

1)　事業から業務への関係がなるべく一対多関係になるように、業務の重複を避ける。

2)　担当セクションに対して主たる業務を割り当てる。

(2)データベースの関連づけ

業務対応表にすべての業務を記入することができたら、インプットデータの登録と、アウトプットデータの登録を行います。インプットデータの最小単位はおそらくどのように優秀な経理システムを導入していたとしても、事業部単位の把握が限界でしょう。

表 5-2　インプットデータの登録

NO.	コスト	万円	所属事業部
C-00001	人件費	2,000	○○事業本部
C-00002	設備費	150	○○事業本部
C-00003	⋮	⋮	⋮
C-00004			
C-00005			
C-00006			

出所：財団法人政策科学研究所作成資料

また、アウトプットデータの最小単位は実施事業単位であることが想定されます。これらのデータは研究開発評価のデータをそのまま流用できます。

表 5-3　アウトプットデータの登録

NO.	事業名	アウトプット分類	件数
O-0001	△△事業	特許申請数	10
O-0002	△△事業	論文	12
O-0003	△△事業	試作機	3
O-0004	△△事業	⋮	
O-0005	⋮	⋮	
O-0006	⋮	⋮	

出所：財団法人政策科学研究所作成資料

これらのデータを先の業務対応表の集計結果と接続すると、下のようなリレーショナル・データベース（RDB）を作成することができます。

出所：財団法人政策科学研究所作成資料

図 5-9　業務分析のためのデータベースの作成

上のような RDB を作成して、ようやく、アウトプットとインプットの対応を、業務という機能を媒介にして関係づけることが可能になり、「業務運営の効率化」を評価することができます。

5.3.2　活動へのコストの割り当て

業務分析を実施した結果、インプットとアウトプットの対応関係については見えてきましたが、ある業務ないしある事業に一体いくら年間に費やされたかを明らかにするために

は管理会計手法を適用する必要があります。代表的な管理会計手法が活動基準会計（ABC: Active Based Counting）と呼ばれるものです。

5.3.2.1　活動基準会計（ABC）

活動基準会計とは、高度に発達した組織活動において管理費用や調整費用などの間接経費が膨張するに従い、無視できないコスト要因になってきたことを経営上の問題として捉える必要性から生み出された原価計算手法のことです。活動基準会計によって、活動ごとに把握された原価情報などを活用し、コストの視点から活動の管理に重点を置くことが可能になります。活動の分析を通じてプロセスのムダ（非付加価値的活動）が明らかにされるので、組織の実態モデルをリエンジニアリングする際に役立つ手法として注目されています。

活動基準会計の手続きは次の通りです。最初に、業務の原価として考慮しなければならないコストを列挙します。コストデータは会計年度において確定された財務データから抽出します。

次に、コストデータを部門別に配分します。積み上げでデータを蓄積している場合には問題ありませんが、コストデータによっては把握できない場合は、部門全体の年間の予算比率等によって配分します。同様に、担当セクション別のコスト比率も推計します。

出所：財団法人政策科学研究所作成資料

図 5-10　活動基準会計の考え方

担当セクションが業務に対して費やしている費用については、何かしらの配賦基準（コストドライバー）を基に配分する必要があります。活動基準原価計算では、この配賦基準を活動の度合いをもっとも良く表している指標を設定することで代理変数とします。例えば、労働時間（人・時）などが代表的な例です。

すべての業務についてコストが配分されたら、業務別にコストを集計します。これが、活動基準原価と呼ばれるものになります。

業務分析によって、機関の活動は「業務」－「事業」－「事業分野」への対応が図られています。業務の活動基準原価を積み上げることで、さまざまな観点からコストの集計が可能になり、アウトプットやアウトカムとの対比が可能になります。

5.3.3　マネジメントの評価

業務運営の効率化については、実施している業務それ自体の目標水準が定められていなければなりません。業務分析において業務をリストアップすると説明しましたが、それを業務別に推計するという作業は、実は業務の戦略的な類型化が前提となっています。

業務の戦略的な類型化とは、マネジメントの観点から業務がどのように位置づけられるかという問題を解くことです。こうした観点に有効な手法として、バランススコアカード（BSC）によるマネジメント評価があります。

バランススコアカードについては、本テキストの第 1 章 1.3.2.2.2 節（p.26）において詳細に解説していますので、ここでは重複して解説することを避けたいと思います。

5.3.4　スタッフ組織の業務の位置づけ

業務分析で「事業」－「業務」－「担当セクション」の対応が比較的可能なのはライン部門であると言えます。ところが、研究開発機関や資金配分機関では、企画調整部門、研究開発評価部門、戦略策定部門等、多くの事業に関わるスタッフ組織が重要な役割を果たしています。

これらの部門に業務分析を単純に適用すると必要以上に多くの業務単位を作り出してしまうことになります。このような問題を解決するために、担当セクションではなく、実際の業務を遂行するための機能を重視した組織構造を仮想的に設計し、業務分析に組み込む方法が考えられます。

5.3.4.1　マトリックス型組織とクロス・ファンクショナル・チーム

図 5-11 はフィンランド科学技術庁（TEKES）の組織構造を図示したものです。TEKES は日本でいうところの NEDO 技術開発機構のような資金配分機関です。TEKES では典型的なマトリックス型組織を構築しています。

図 5-11 の縦軸は資金配分業務（Core Processes）が採られ、横軸は事業領域（Technology and Research Area）が採られています。また、右下には戦略形成・分析部門が両方の業務軸 b にコミットするような形で組織構造を示しています。

出所：TEKES ヒアリング資料(2006.1)より抜粋

図 5-11　マトリックス型組織（TEKES の例）

マトリックス型組織は、形式的には他の組織と同様に各セクションが独立したツリー構造に位置づけられています。しかし、機能的にはスタッフ組織とライン組織が区別され、両方の組織が事業を共有するかたちで業務を遂行するようになっています。

このとき、組織の垣根を越えるようにそれぞれのセクションからプロジェクトに応じてメンバーを募り、チームを作るような実施体制をクロス・ファンクショナル・チーム（CFT）と呼びます。CFT は日産のゴーン社長がフェニックス計画で導入したマネジメント手法として一躍有名になりました。

CFT の考え方を業務分析に導入すると、次のような関係性の把握ができるようになります。

出所：財団法人政策科学研究所作成資料

図 5-12　CFT を導入した業務分析

図 5-12 では、業務を直接担当のセクションに関連づけるのではなく、セクションのメンバーから構成される CFT に関連づけしています。このようにすることで、CFT が実施すべき業務をセクションに割り当てることが間接的に可能になり、かつ、CFT が媒介となることで割り当ての組み合わせ数を大幅に減少させることができます。つまり、その後の活動基準原価計算の手間が掛からなくなります。

CFT は必要に応じて結成されるため、組織がフラットでかつ、機動的な構造になっていなければ機能しません。しかし、一度このような組織構造を作り上げることができれば、業務効率の評価の対象を CFT に絞り込むことができるため、機関評価の手続きを大幅に短縮することができます。

CFT の考え方は、組織のプログラム化を一層進めたものであると言えます。

5.4　機関評価の方法論を踏まえた中期目標・中期計画の策定

独立行政法人評価は中期目標・中期計画・年度計画に対して評価が実施されます。したがって評価の方法論よりも先の段階でこれらの計画の策定方法を厳密に行う必要があるのですが、より具体的な計画の策定方法は、予め業務分析などを通じて得られたデータから、機関評価の枠組みを踏まえた上で計画をデザインすることが望ましいと言えます。

中期目標・中期計画は独立行政法人の経営方針を規定する重要な文書です。これをマニュアルなどで形式的に作成するのではなく、良い評価結果が得られるように、むしろ戦略的に取り組むべき対象として機関のマネジメント情報を活用するべきでしょう。

5.4.1　独立行政法人評価の枠組み

独立行政法人制度は、政策実施機能に係る一定の事務・事業を担う独立の法人格を持つ法人を設置し、国の事前関与・統制を制限し、法人の運営における自主性・自律性を確保するという主旨で開始されました。一方で、主務大臣の指示する明確な達成目標の下で、その業務の実績を事後的に評価し、その結果を法人の業務運営の改善に反映させ、また、毎年毎年の長の責任や役職員の処遇等に反映させ得るという仕組みにより、業務運営の効率化と国民に対して提供するサービスの向上等、国民の求める成果の実現を図ることを目的としています。

中央省庁等改革基本法（公布：平成 10 年 6 月 12 日）ならびにこれを受けて施行された独立行政法人通則法（公布：平成 11 年 7 月 16 日）では、独立行政法人の主務省（当該独立行政法人を所管する内閣府又は各省をいう）に、その所管に係る独立行政法人に関する事務を処理させるため、独立行政法人評価委員会を置くことが明記されています（中央省庁等改革基本法第三十九条，独立行政法人通則法第十二条)。

また、独立行政法人通則法第 32 条第 5 項及び第 34 条第 3 項による基準の適用により、全政府レベルでの評価機関として、各府省評価委員会から通知された独立行政法人の業務の実績の評価結果について総務省の政策評価・独立行政法人評価委員会が二次評価を行い、必要があると認めるときは、当該評価委員会に対し、勧告することが可能とされています。

このように、事後評価に重点を置くということが制度の基本の一つであることから、独立行政法人の業務の実績の評価が、中立・公正な立場から客観的に実施されることが重視されています。このため、各府省に第三者評価機関である評価委員会が置かれて評価を行うことに加えて、総務省に全政府レベルの第三者評価機関である政策評価・独立行政法人評価委員会が置かれることにより、独立行政法人の評価の客観的かつ厳正な実施を確保する仕組みが実現することが期待されています。

次の図 5-13 は独立行政法人制度の全体像を示したものです。

出所：総務省（http://www.soumu.go.jp/hyouka/dokuritu-gyouseihoujin.htm#3）

図 5-13　独立行政法人制度の全体像

独立行政法人制度は、「中期目標」や「中期計画」による中期的な管理を行う点に制度上の特徴があります。「中期目標」は、3 年から 5 年を期間として、主務大臣から独立行政法人に示されるものであり、それには、業務運営の効率化や、国民に対して提供するサービスの質の向上、財務内容の改善などについての目標が掲げられています。独立行政法人は、この「中期目標」を達成するため、自ら「中期計画」を作成して主務大臣の認可を受け、さらに「中期計画」及び年度ごとの「年度計画」をもとにして毎年度の業務を行います。

独立行政法人の業務の実績評価には、毎年度の業務の実績について行われる年度毎の評価と、中期目標期間における業務の実績について行われる中期目標期間ごとの評価との 2 種類があります。

独立行政法人の業務実績については、専門的な知識を持つ第三者で構成される府省の独立行政法人評価委員会による評価と総務省政策評価・独立行政法人評価委員会による評価活動により、ダブルチェックされるようになっています。

独立行政法人は、まず、府省の独立行政法人評価委員会により、各年度の業務実績や中期目標期間の業務実績について、中期計画の実施状況、中期目標の達成状況等を考慮の上評価されます。府省の独立行政法人評価委員会は、評価の結果、必要があると認める場合には、法人に対して業務運営の改善などを求めることができます。

次に、総務省の独立行政法人評価委員会が、府省の独立行政法人評価委員会による評価結果の通知を受け、第三者的な立場から調査・審議を行い、必要があると認める場合に意見を述べることができます。

中期目標の期間の終了時において、主務大臣は、独立行政法人の業務を継続させる必要性、組織の在り方その他その組織や業務の全般にわたる検討を行い、その結果に基づいて、必要な措置を講ずることになります。この検討を行うに当たっては、府省の独立行政法人評価委員会の意見を聴くこととされています。

また、総務省の政策評価・独立行政法人評価委員会は、中期目標の期間の終了時において、独立行政法人の主要な事務や事業の改廃に関して、主務大臣に対して勧告できることとされています。つまり行政改革からの評価の視点（査定的な視点）が非常に強い制度となっています。したがって、独立行政法人評価では、定量的に評価可能な部分を少しでも広げて、中期目標・中期計画に盛り込むという対策が必須になります。

5.4.2　中期目標・中期計画の策定

このような特徴を持つ独立行政法人評価では、評価に先だって、評価の対象として委員の審議に耐えうるしっかりとした中期目標・中期計画を策定する必要があります。

前節の機関評価の方法論で学んだ業務分析の考え方を中心に、中期目標・中期計画の策定の仕方を考えてみましょう。

5.4.2.1　組織のプログラム化

中期目標の最初の頁には、当該独立行政法人の設置目的ミッションが定義されています。このミッションを実現していくために必要な条件として、何が状況として変化しなければならないのかについて、考察しなければなりません。すなわち、アウトカムの観点から、独立行政法人の事業領域を類型化する必要があります。このことは、事業領域での活動が活発になればなるほど、ミッションの実現に貢献しているという因果関係を暗黙のうちに前提としています。

次に、事業領域が定まれば、中期目標期間中に実施予定の事業を事業領域と対応させる手続きに移ります。広報や普及啓発などの事業についても、きちんと位置づける必要があります。

実施事業の位置づけが出来た段階で、中期目標期間中の事業計画を事業毎に記します。この時、主に担当するセクションがどこなのかについてのアサインを中期計画で明らかにしておく必要があります。

ここまでの作業で、中期目標・中期計画の段階での組織のプログラム化は一通りの作業として完結します。

5.4.2.2　業務運営の効率化

業務分析で得られた業務に対して、独立行政法人全体のマネジメントの観点から、戦略的に類型化を行います。中期計画の段階では組織全体のレベルでの記述で構いませんので、バランススコアカードを作成し、中期目標期間中に達成されるべき業務の改善目標を立てます。

改善目標に対して、具体的にどのような活動を実施するのかについて個別に記述します。

5.4.3　年度計画の策定

年度計画では、中期計画に記載された各事業別の業務運営の効率化について記述する必要があります。業務分析が実施されていれば、かなり詳細なレベルのバランススコアカードが作成できていますので、年度内に改善の見込みがある項目について目標を設定します。

また、年度計画においては、前年度のパフォーマンスを改善目標の目安とすることが可能です。業務分析データベースから前年度の対事業運営のコストデータを集計しておくことで、改善目標幅を実績から想定することができます。つまり、エビデンスに基づいた計画策定が実施可能になるのです。