

政策評価・行政評価のための 評価計画策定ワークブック

—ロジックモデルに基づく執行評価・アウトカム評価—

Evaluation Plan Workbook

Innovation Network, Inc. 著

茂木康俊 訳

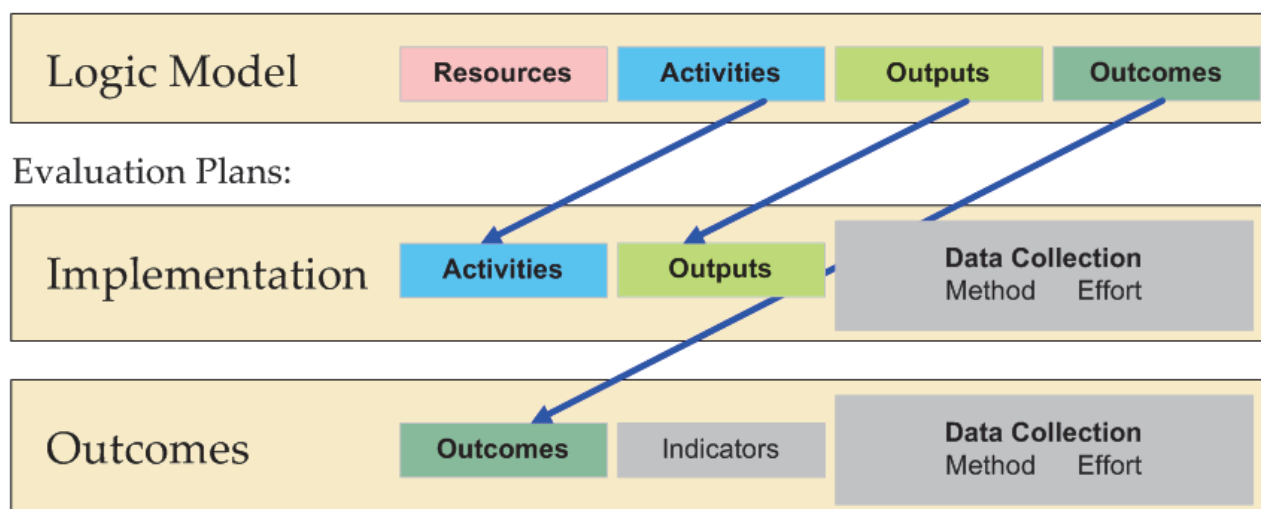
Translated by Yasutoshi Moteki



広島大学地域経済システム研究センター 発行

政策評価・行政評価のための
評価計画策定ワークブック
—ロジックモデルに基づく執行評価・アウトカム評価—
Evaluation Plan Workbook

Innovation Network, Inc. 著
茂木康俊 訳
Translated by Yasutoshi Moteki



日本語版序文

本書は、米国ワシントン D.C. に本部がある非営利組織である **Innovation Network, Inc.** が英語版を発行するロジックモデルを活用した評価計画策定のための実践的な手引き書の日本語版である。**Innovation Network, Inc.** は 合衆国内国歳入法 501 (c)

(3) の適用認定を受ける非営利組織である。非営利組織のプログラム計画と評価に関して、他の非営利組織への情報提供やオンライン上でのツールの提供、コンサルティング、研修を行うことで支援している。なお本書は 2014 年に日本語版を発行した『ロジックモデル・ワークブック』とシリーズをなしており、同書と組み合わせてロジックモデルの作成、評価計画の策定、実際の評価の実施までを簡潔に示す内容になっている。ご関心のある読者は、同書もあわせてご参照頂きたい。

米国では、1993 年に連邦政府で制定された政府業績成果法 (**Government Performance and Results Act: GPRA**) が行政分野における評価についての近年のブームを引き起こす契機となっており、その後同法の改正法である政府業績成果近代化法 (**GPRA Modernization Act of 2010: GPRAMA**) が施行され、米国連邦政府において組織のミッションを記述する戦略計画を策定し、具体的な目標を立て、指標を基に評価し改善するというサイクルが各省庁に導入され、現在も運用が続いている。米国の **GPRA** の取り組みを主に参考にし、我が国の地方自治体で事務事業評価制度の導入が広がった。同制度は 1996 年度に他の地方自治体に先駆けて三重県で導入され、他の自治体も同様の政策評価制度を導入していった。その後、2001 年に成立した「行政機関が行う政策の評価に関する法律」(政策評価法) に基づいて国レベルで政策評価制度が導入され、各省庁で政策評価の実施と結果の公表が続けられている。

国での政策評価制度導入後、全国の多くの地方自治体において事務事業評価制度を中心とした政策評価制度が、多くは行政評価という名の下に導入が進んでいるが、いわゆる「評価疲れ」の問題など、政策評価制度自体のコストや職員のモチベーションへの悪影響が指摘されるようになってきている。そのような中、注目を集めつつあるのがロジックモデルの手法である。ロジックモデルとは上位の目標 (ミッション) から事業 (プログラム) の活動レベルまでの論理的つながりを視覚的に把握するためのツールである。現在、我が国の地方自治体や行政において政策評価・行政評価の改善のためにロジックモデルの活用が広がりつつあり、平成 25 年度の総務省の政策評価に関する統一研修では「政策の「意図」の明確化とその検証手法」というタイトルで各管区においてロジックモデルに関する演習が行われ、訳者はその一部について講師を務めた。その後、平成 26 年度および 27 年度と同研修の中央研修でもロジックモデルがテーマとして取り上げられている。

本書『政策評価・行政評価のための評価計画策定ワークブック』は具体的な例に基づき、前書『ロジックモデル・ワークブック』に基づいたロジックモデルを前提に、執行評価およびアウトカム評価を具体的に実行するための知識や方法を理解し、ワークブック形式で実際に作成することを助ける手引き書である。本書の英語版の発行元である **Innovation Network, Inc.** のウェブサイトでは本書のオリジナル版やロジックモデルおよび評価計画をオンラインで作成するためのツール (**Logic Model Builder** および **Evaluation Plan Builder**) を提供しており、本書と合わせ活用することが可能である。

訳者は国際評価年にあたる 2015 年に米国シカゴ市で開催されたアメリカ評価学会 (**American Evaluation Association**) に参加し、ロジックモデルをテーマにしたいくつかのセッションを聴講し、行政を含めた公共分野においてロジックモデルを活用し

た評価がアメリカ評価学界や 50 カ国以上にも及ぶ大会の参加者の各出身国において現在も重要視され、関心を集めていることを確認した。それに関連して、2016 年 2 月発行の『評価クォーター』誌に掲載した論文で同学会の動向について議論を行ったが、評価における「ポジティブ・アプローチ」というキーワードに訳者は強く関心を持っている。2016 年 11 月 12 日木曜に開催された会長企画の「評価への模範的ポジティブ・アプローチ—世界からの視点」(Exemplary Positive Approaches to Evaluation: Views from Around the World)のセッションやその他の報告がテーマとしている。我が国における地方自治体における政策評価は公共経営や行政改革などの概念に結びつけて論じられることが多く、所管課で内部評価に関わる職員のみならず、評価取りまとめ部署の職員にも大きな負担をかけるにも関わらず、評価が職員やステークホルダーに有益な効果をもたらしているのか明確でなく、公共組織への市民の批判的な視線や効率性への強い要望と相まって、むしろネガティブな結果をもたらす、職員にとって気が重い作業になっているように思われる。NPO などの非営利組織も含めた公共組織全般が、ロジックモデルを活用し、執行評価・アウトカム評価を計画に基づいて適切に、かつ自ら積極的に行うことが重要であり、各ステークホルダーに図示やデータでの表示も含めわかりやすく自らの成果を表現し、それが各ステークホルダーに伝わり成果が認められることで、組織の構成員が前向きに業務に取り組むことにもつながる。そのような本来のプログラム評価の理念が我が国の政策評価・行政評価に浸透することに本書の翻訳が寄与できれば望外の幸せである。

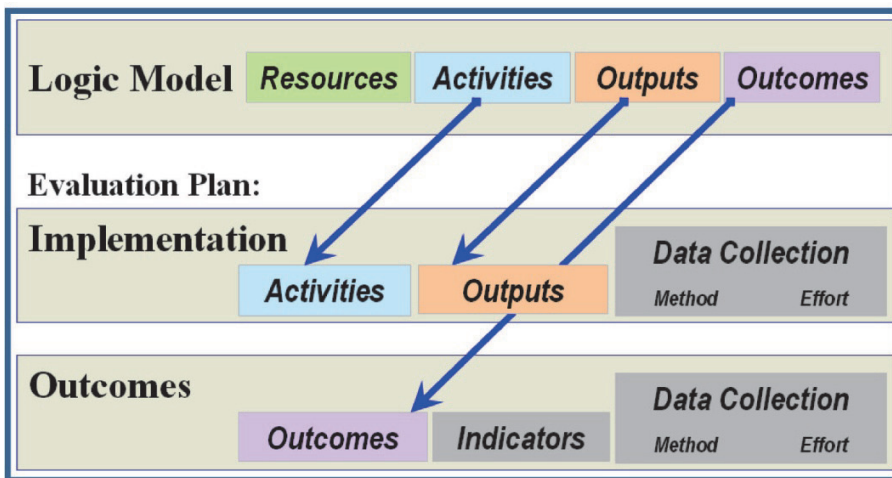
最後に、本書の翻訳出版の発行をお引き受け頂いた広島大学大学院社会科学研究科附属地域経済システム研究センターの伊藤敏安教授、本翻訳出版に関しご協力とご理解を頂いた Innovation Network, Inc. の Director である Johanna Morariu 氏および Veena Pankaj 氏、そして、本書の翻訳出版の準備にあたり陰に陽に訳者を支援して下さった多くの方々に、厚く御礼申し上げます。

2016 年 10 月

広島大学大学院社会科学研究科 准教授
茂木 康俊

* 本書は、平成 27 年度の地域経済研究協議会の研究助成「地域医療構想実現に資する医療における情報・評価に関する研究—個人情報情報の利活用とロジックモデルに着目して—」の研究成果の一部である。

Evaluation Plan Workbook



政策評価・行政評価のための評価計画策定ワークブック

目次

	ページ
イントロダクション.....	7
本ワークブックの利用法.....	7
評価とは何か?.....	8
証明 vs. 改善—評価の歴史の概要.....	8
我々のアプローチ.....	8
評価の諸原則.....	9
なぜ評価を行うのか?.....	10
評価計画の策定.....	11
執行評価とアウトカム評価: 片方だけの実施は可能か?.....	12
執行評価.....	13
何を行ったか: 活動およびアウトプット.....	13
どの程度うまく行えたか: 追加的評価項目.....	14
アウトカム評価.....	16
どのような違いを生じさせたか?.....	16
指標.....	16
堅固な指標ステートメントの要素.....	18
指標の例.....	19
指標の目標設定.....	20
複数の指標.....	20
直接的指標 vs. 間接的指標.....	22
データ収集の概要.....	23
データ収集の方法.....	23
データ収集の容易さ.....	25
計画の見直し.....	26
付録: 評価テンプレート	
執行評価	
アウトカム評価	

イントロダクション

本ワークブックの利用法

イノベーションネットワークの評価計画ワークブックへようこそ。本書はプログラム評価[program evaluation]の計画策定に関する諸概念および各プロセスを紹介するものである。このワークブックに沿うことで、エンパワーメントのためのツールとして評価を理解できるようになることが望まれる。そして、組織がより効果的になるのに評価がどのように役立つのか学び、組織のプログラムの実施とアウトカムの両方を評価するための計画を策定できるようになるだろう。

本書は2番目のワークブックである。読者がこのワークブックを使用する前に、ロジックモデル・ワークブックを利用することをおすすめしたい。

読者は次に示す例のうちから最も合う方法で本書を利用されたい。

- プログラムの評価計画の作成に単独で役立つ手引き書として
- オンラインの Evaluation Plan Builder (Point K Learning Center www.innonet.org) の利用者への補足的資料として
- 評価計画策定についての対面研修[in-person training]への補足資料として



このチェックリストのアイコンは、情報の収集、行動の開始、テンプレートへの記入、あるいはオンラインの Logic Model Builder への入力など、ワークブックの中であなたが何か行動を行うべきか、または何かを記録すべきポイントで現れる。

あなたは、Innovation Network の Point K Learning Center (すなわち www.innonet.org における一連の、オンラインでの計画策定、評価ツール、および関連資料) において Evaluation Plan Builder を使用しオンラインで評価計画を作成できる。このオンラインツールは評価計画の作成プロセスをあなたに手引きし、既存のロジックモデルをあなたの評価計画のための枠組として使用することができ、またあなたの作業を保存し後で作業に戻れたり、同僚が見直しや検討をしたりできるように作業を共有できるようになっている。

紙で作業を行うことを好む場合、評価計画のテンプレートがこのワークブックの終わりに付されている。あなたの評価計画に対して、修正および更新の余地を残すために、このテンプレートを数部コピーするとよい。

評価にはしばしば否定的な含意が結びつけられる。その1つは、それが人々に対し行われるものであるということである。つまり、ある者が評価されるということである。それとは対照的に、参加型評価[Participatory evaluation]は、プログラムあるいはコミュニティの中の人々によってコントロールされるプロセスである。それは彼らが自身の成長とエンパワーメントのための定式的で、内省的なプロセスとして行うものである。

M. パットン『定性的評価の方法』¹

評価とは何か？

評価とは、ステークホルダーがプログラムをより良く理解し、その有効性を改善し、および/または、今後のプログラム策定に関する決定をすることを可能にするプログラムに関する情報の体系的な収集のことである。

証明 vs. 改善—評価の歴史の概要

評価は、プログラムに関わる人々が、プログラムをより良く理解し、改善することを手助けするツールとしてこれまで常に見られていたわけではない（現在も、常にそのように見られているわけではない）。

歴史的には、プログラムがより効果的に機能するように改善することについてよりも、むしろプログラムが機能したかどうか証明することに、評価は焦点を置いていた。このように証明に焦点を置いていたことは、「客観的な」外部の評価者が評価を行っていたことを意味していた。リサーチデザインには、因果関係を評価するために対照群または比較群を使用して、厳しい科学的基準を用いた。評価は、プロジェクトの終わりに出てきて、プログラムが成功したか失敗したかのみ焦点をあてていた。そして、成功することに貢献したのは何か、成功することを妨げたものは何かを学ぼうとすることはなかった。結局は、このタイプの評価はしばしば、評価プロセスからプログラムスタッフおよびその他の職員を切り離していた。つまり、これらのステークホルダーは、めったにプログラムに関する自身の疑問に対する答えを学ばず、また、プログラムを改善するのに役立つ情報を彼らが得ることはめったになかった。

我々のアプローチ

我々は、評価がエンパワーメントの方法になりうると信じている。参加型評価[Participatory evaluation]は、組織が自身の成功を定義し、自らの評価の質問を提示し、ステークホルダーおよび地域住民をプロセスに関与させることに力を与える。外部から押しつけられるよりむしろ、評価は、プログラムのステークホルダーがプログラムによって達成されると予想されること（および、その時期）を確認することに役立つ。その結果、プログラムへの各自の期待が調整されることが確実になる。プログラムに何が投入され、プログラムが何を行い、何を生産しているか、およびプログラムが何を達成しているかを体系的に見ることによって、この評価アプローチは、プログラムのステークホルダーが成果[results]について責任を持ち、プログラムの改善方法について学ぶことの両方を可能とする。

¹ 2nd edition, 1990; p. 129.

評価の諸原則²

我々は、以下の場合に、評価が最も効果的になると確信している：

- **プログラムの計画策定および実施[delivery]へのリンクがなされる。** 評価は計画と執行に情報を与えるものである。あなたにいくらか時間に余裕があるか、またはあなたがそれをするを求められたときにだけ、評価を行うべきではない。むしろ、評価はプログラムの有効性に不可欠のプロセスである。
- **利害関係者の参加を確保する。** 評価の結果で影響を受ける者は当該プロセスに関わる権利を有する。参加は彼らが評価の目的を理解し、評価に情報をもたらすのに役立つだろう。参加はまたステークホルダーが評価結果を出すことに貢献し、そしてそれを受容することを促進する。これはプログラムの改善のための評価結果の利用の見込みを増加させる。
- **組織の学習し自省する能力を支援する。** 評価は本来それ自体が最終目的ではない。継続的学習に貢献できるよう、組織の中核的マネジメントプロセスの一部となるべきである。
- **プログラムが奉仕するコミュニティを尊重する。** 評価は、地域住民を尊重し、彼らに対する質問について思慮深くなければならない。評価はプログラム参加者およびプログラムにより影響を受けるか、関連を持つその他の者に対して「行われる」ものであるべきではない。むしろ、評価は、プログラムを改善するのを助け、コミュニティの需要をより満たすための情報を生み出すために、彼らの知識と経験を引き出すべきである。
- **最少の努力で最も多くの情報の収集を可能にする。** あなたはすべてを評価することはできないし、そうする必要もない。あなたが知る必要があることに焦点を合わせなさい。プログラムについて説明責任を果たし、プログラムの改善を行うためにあなたとそのステークホルダーが必要とする重要な情報はどのようなものであるか？

² この節の情報の一部は以下の文献を参考にしている: Earl, Sarah et al. *Outcome Mapping: Building Learning and Reflection into Development Programs*. International Development Research Centre (Canada), 2002.

なぜ評価を行うのか？

良く設計され、実施される評価を行うことは、**あなたの利益**になるだろう。それは以下の点であなたに役立つだろう：

- **プログラムを理解し、改善する。** 最善に運営されたプログラムでさえもいつも完全な成功をもたらすわけではない。あらゆるプログラムが改善できる。評価で集められた情報は、プログラム改善のための手引きを提供できる。あなた方の進行中の仕事に評価を組み入れると、有益な情報を獲得することになり、（その結果として）絶えず情報を収集し、変化し、向上する「学習組織」が形成されるであろう。
- **プログラムの基礎となるセオリーを確かめる。** プログラムの実施と同様に、プログラムの短期、中期、長期の実績に関して系統的にデータ収集することは、あなたが、プログラムの基礎となる仮説が正確であるかどうか（そして、その条件）、または、それらが修正される必要があるかどうかを理解することに役立つ。
- **プログラムのストーリーを話す。** 評価を通じて集められたデータは、あなたがあなたのプログラムが行っていることや、達成していることについて説明するのに役立つ説得力のある情報を提供しうる。評価結果はステークホルダー、資金提供者、および政策立案者に対しあなたのプログラムを説得する際に強い枠組みを提供する。
- **アカウントビリティを果たす。** 評価は、あなたが資金の責任ある管理者であることを示すことに役立つ。
- **関係者に情報提供する。** プログラムを評価し、改善した非営利団体は、より広い非営利のコミュニティと信頼し得る結果[results]を共有できる。結果を共有できる共同体は、より効果的となりうる。
- **資金集めの活動をサポートする。** あなたのプログラムの明確な理解（あなたが成功裏に行ったこと、および正確にどのようにアウトカムを達成したか）は、あなたが業務を続け、活動を拡大あるいは活動を重ねるための追加の資金を集めることに役立つ。

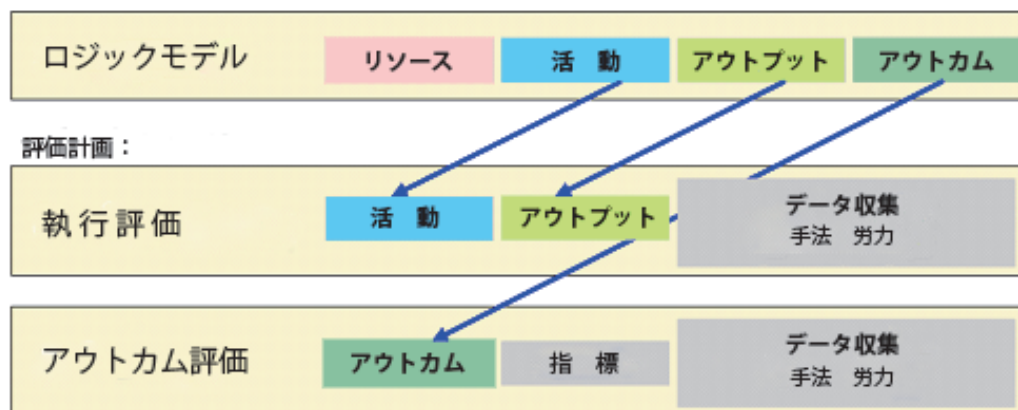
評価計画の策定

評価計画は、プログラムについて持つ質問を特定し整理し、答えを得るための道筋を描くものである。各組織が評価で検証するほとんどの質問は以下の3つのカテゴリーに分類される:

- 何を行ったか？
- どの程度うまく行えたか？
- プログラムはどのような違いを生み出したか？（どのような変化がプログラムによって生じたか？）

あなたのプログラムのロジックモデルは評価計画の基礎を形成するだろう。ロジックモデルを見ることで、プログラムに関する答える必要のある質問を見つけるだろう。評価の計画策定の目的は、これらの質問を特定し、答えを見つけるための道筋を計画することである。

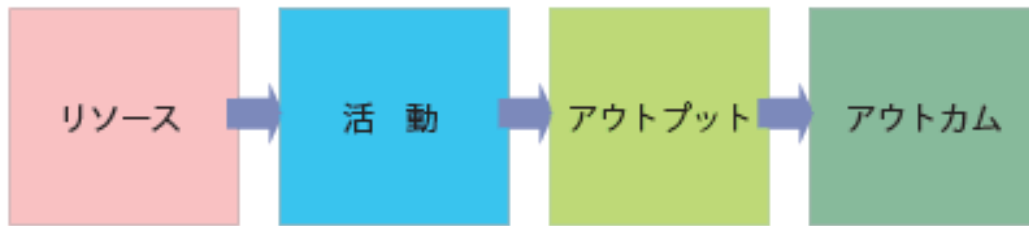
図1 ロジックモデルと評価計画の関係



主に2つの形式の評価が、これらの質問に答えるのに役立つ。

- 1) **執行評価:** あなたは計画通りサービスあるいは活動を実行しているか？ あなたは意図している対象集団に到達しているか？ 想定している参加者数に達しているか？ あなたが期待する生産物[products]につながっているか？ どのように参加者はこれらのサービスと活動を認識するか？ これらの質問は執行に関するものである。
- 2) **アウトカム評価:** 対象者はあなたが達成しようとした知識、態度、行動、または認識における変化を経験しているか？ あなたの業務のアウトカムは何であるか？ あなたの対象とする参加者の間で何が達成されているか？ これらの質問はアウトカムに関するものである。

図2 プログラムの各プロセス



執行評価とアウトカム評価: 片方だけの実施は可能か?

評価が効果的であるためには、両方のタイプの質問に答えるべきであると私たちは確信している。もう一方を行わずとも片方を実施できるが、両方を行ったほどは学べないだろう。

「昔の時代」には、非営利組織の評価は、プログラム活動に関して記録して、報告することに焦点を合わせていた。そして、非営利組織および他の組織はその際、計画通りプログラムの活動を実行した場合、望まれた結果が個人、家族、組織、またはコミュニティに生じると想定した。一般的に、非営利のコミュニティは、アウトカムを除外して、執行に関して報告することに焦点を合わせていた。近年、振り子は逆方向に揺れ動いた: 非営利団体は執行に強調を置くよりも、アウトカムを測定し報告するプレッシャーを受けている。

私たちが使用する評価のフレームワークはアウトカムおよび執行の両方を組み込んでいる。

プログラムは複雑である。望まれたアウトカムをどの程度達成したか理解する必要がある。また、プログラムのどの側面がそれらの業績に貢献したか、そして、理想的な結果を得るためにどんな障害がプログラムに存在するかを学ぶ必要がある。

このワークブックは:

- プログラムの執行を評価する計画の策定過程の全体を順に提示する。
- プログラムのアウトカムを評価する計画を作成するのに役立つ。

あなたが最初にどちらを行うかは重要でない—アウトカムから始めるのを望むならば、そのように行ってよい。

執行評価:

何を行ったか？ どの程度うまく行えたか？

活動	アウトプットおよび執行に関する質問	データ収集の方法 (どう測定するか)	データ収集の労力 (既に持っている、 低、中、高)
活動のグループ	アウトプット		
	質問		
活動のグループ	アウトプット		
	質問		

上図は執行評価の計画の各要素を概略として示している。

何を行ったか: 活動およびアウトプット

執行の評価計画策定はプログラムの**活動**の特定から始まる。活動は、プログラムが望まれるアウトカムを獲得するために実施する行動である。あなたのプログラムが多くの活動を伴う場合は、これらの活動を各活動カテゴリー（プログラムの中で密接に関連する活動のグループ）に分類することになるかもしれない。

アウトプットはプログラムの活動の具体的な成果物[products]である。アウトプットは活動に関するエビデンスでもある。執行評価においては、アウトプットはあなたの活動を評価するために実際に測定する項目である。アウトプットを計測することは以下のような質問に答えることである: 私たちは何を行ったか? これは評価で多くの場合最も簡単で、最も直接的なプロセスである。



評価計画の作成: ロジックモデルからの情報

(Point K Learning Center を www.innonet.org から閲覧し、我々の Logic Model Workbook をダウンロードするか、オンラインの Logic Model Builder を利用してみてください。)

- あなたが紙で作業している場合:
 - ロジックモデルから活動カテゴリーと個々の活動を取りだしなさい、そして、執行評価計画のテンプレートの「活動」の欄にそれらを記入しなさい。
 - ロジックモデルから各アウトプットを取り出し、評価計画の「アウトプット」の欄にそれらを記入しなさい。
 - アウトプットを見直し、あなた方の活動からさらに別の「生産物 [products]」を検証したいかどうか決めなさい。これらをアウトプットの欄に記入しなさい。
- あなたが Point K でオンラインで作業している場合は、ここでの情報はあなたの Logic Model からあなたの評価計画にあらかじめ自動入力される。

どの程度うまく行えたか: 追加的評価項目

活動および関連するアウトプットを記録することで、あなたが**何**を行ったかが明らかになる。しかし、それは評価の各目的のためには十分ではない（なぜなら、あなたは、ロジックモデルで活動とアウトプットは既に特定していたからである）。執行評価の目的は、**どの程度うまく行えたか**を理解することである。

次のステップは、あなたの活動とそれらのアウトプットについてあなたが持つ他の質問を特定するためのものである。どのような情報が、あなたのプログラムの執行状況をよりよく理解するのに役立つだろうか？ 以下はあなたが考慮しうる各種の質問例である:

- **参加:** 対象の聴衆は、期待したように活動に参加したか？ なぜであるか？ あるタイプの人々が実際よりも多く含まれていたり、あるいは逆に少なく含まれていたりしていないか？ それはなぜであるか？
- **品質:** 提供したサービス/資料を対象の聴衆が価値あるものとして認識したか？ それらは適切であったか？ その分野の第三者はどのようにその品質を受け取ったか？
- **満足:** プログラムのサービスで影響を受けた者がそれらを肯定的に受け取ったか？ なぜであるか？ 誰が最も満足し、誰が最も満足しなかったか？
- **文脈:** どのような他の要素が、計画通りプログラムを実施するあなたの能力に影響を及ぼしたか？ どのような、政治的、経済的あるいはリーダーシップの問題が介在し、その結果あなたのプログラムの期待していたアウトカムに影響を及ぼしたか？

質問の選択

あなたのプログラムの執行に関して答える必要のある追加の質問をブレインストーミングしなさい。「参加/品質/満足/文脈」は目安としてのみ使用しなさい。すべての活動に対しそれぞれの質問を1つは持つという必要はないし、質問をこれらのカテゴリに制限する必要もない。あなたのリストを、プログラム改善に最も大きい影響を与える「知る必要のある」質問に重点を置きなさい。

これらの質問への回答はプログラム改善のために豊かなフィードバックを提供できる。それはアウトプットと活動の簡単な目録という意味を超えて、しばしばあなたのプログラムが対象としている人々の視点を明らかにする。それでも、それらはアウトカムよりもプログラムの執行に関連するものである。それらはあなたが何を行ったかを扱うものであり、あなたのプログラムの結果何が変化したかを扱ってはいない。



評価計画の策定: あなたの評価計画のテンプレートの質問の欄か、オンラインの Evaluation Plan Builder の「執行に関する質問」の欄に、各質問を入力しなさい。

執行評価はあなたが何をしたか、そして、あなたがどの程度うまく行えたかに関する重要な情報を提供する。あなたが学ぶ内容は、あなたの当初のプログラム計画と比較して進捗を測るベンチマークとして働きうる。

- あなたは、ことによると計画通り活動を行っていない; あるいは
- あなたがそれらの活動を行っているが、それらはあなたが意図した成果物 [products]/アウトプットにつながっていない、もしくは
- それらは意図したアウトプットにつながったが、品質あるいは満足のレベルが、あなたが望んでいたレベルではなかった。

このような情報は、あなたのプランの調整、活動の変更、あるいはあなたの理論的想定の見直しが必要の有無を決めるのに役立つ。あなたの執行状況を評価することは、アウトカムを評価する前に、あなたの取組にフィードバックループを提供しうる。

アウトカム評価: どのような違いを生み出したか?

最近では、善意、誠実な努力、または模範的プログラムの執行が、我々が働きかけている人々にとって望ましいアウトカムをもたらすと非営利組織が想定するのはもはや許容されない。アウトカムの達成を測るための計画を策定するのに時間をかけることは重要である。

あなたのロジックモデルにおいて、望まれる各アウトカム—すなわち、あなたの業務の結果として目にすると期待されるさまざまな変化—は、既に特定されている。アウトカムは知識、技能、態度、行動、モチベーション、決定、政策、および状態における変化としてしばしば表現される。それらはさまざまな個人、コミュニティ、組織、またはシステムの中に生じるものである。

指標

あなたがどの程度良く成果を達成したかを評価するために、あなたが想定したアウトカムに対し指標を決定することが必要になる。

指標は、あなたのプログラムが意図した成果を達成しているかどうかを示す証拠[evidence]あるいは情報である。指標は**測定可能**で観察可能な特徴を持つ。それらは次のような質問に答えるものである: 「生じた変化を我々はどのように知ることができるか?」

我々はアウトカムをしばしば抽象的概念または野心的希望として定義しがちである。指標はアウトカムを測るものである。指標は、そのような概念または野心的希望について具体的な情報を提供する明確な性質または行動を表すものである。しばしば、1つのアウトカムは2つ以上の指標を持つ。あなたが指標を定めるとき、以下のように問うことは役に立ちうる。: 「アウトカムはそれが生じるときにどのような形態であるか? それが生じたかどうかどのように知ることができるか? 何を見ることができるか?」

指標は次のようなものとなるべきである:

- **有意義である:** 指標はプログラムの主要なステークホルダーにとって重要な情報を提供する。プログラムの「成功」についての定義が人によって異なりうることを念頭に置きなさい。成功がどのようなものであるかに関して主要なステークホルダーの間で意見が一致することは、あなた方の評価結果が支持されることを確保するのに重要である。
- **直接的である:** 指標または各指標の組み合わせが、アウトカムを表現するために、アウトカムの不可欠な構成要素を十分に捉えていなければならない。アウトカムを十分に測るためには、いくつもの指標が必要となることもある。しかし、用いられる指標の数について基準があるわけではない。複数の指標が必要となることはしばしばあるが、3 または 4 を超える場合に、アウトカムが複雑過ぎて、よりよく定義されるべき状況になりうる。また、指標はアウトカムと同じ種類の変化を反映するものでもなければならない。例えば、アウトカムが態度または意見の変化に関するものであれば、指標は行動の変化を表すものであってはならない。
- **役に立つ:** 指標によって提供される情報は、プログラムの改善のために実用的に使われうるものである。
- **実際に収集できる:** 指標のためのデータは収集が重荷であるようなものではない。時宜を得た方法で、そして、手頃な費用で指標に関するデータを集めることができるか考慮しなさい。指標が上述した他の評価基準を満たすものの、収集のための労力[effort]による負担が重すぎるものとなることもある。本書の評価テンプレートは必要となる労力のレベルに注意する機会をもたらしてくれる。

堅固な指標ステートメントの要素

評価に役立てるため、堅固な指標ステートメントになるよう以下の 4 要素を含めるべきである:

- **どのくらい[How much]**。成功と考える達成レベルを示す対象集団における変化の量を明確にしてください。これはあなたの業務の目標[target]を定めるものである。また、これをあなたのプログラムの基本水準[baseline]および妥当な変化のレベルの理解に基づいたものとしなさい。
- **誰が[Who]** あなたが測定する対象集団[target population]を特定しなさい。
- **何を[What]** あなたが測定しようとする状態、行動、または性質について記述しなさい。
- **いつ[When]** 変化が起こる時間枠を示しなさい。

例えば、経済的エンパワーメントの研修プログラムのアウトカムは、参加者がより良く資産管理を行えるようになることかもしれない。そのアウトカムのための指標の 1 つは次のようなステートメントになる:

- 75%の参加者は、当該プログラムの開始後 6 カ月以内に無料の銀行当座預金口座を開設する。

プログラムの参加者の大部分がペイデイローン業者などの高価な金融サービスで自身のお金を管理していることを、事前プロセスで知っているとは仮定しなさい。参加者の 75%が当該プログラムの開始後 6 カ月以内に無料の銀行当座預金口座を開設するならば、あなたはそれを資産管理の方法の改善の指標の 1 つと見ることができるかもしれない。

上述の要素にステートメントを分けると、以下のようになる:

- どのくらい[How Much]=75%
- 誰が[Who]=プログラムの参加者
- 何を[What]=無料の銀行当座預金口座の開設
- いつ[When]=プログラムの開始後 6 カ月以内

以下はアウトカムおよび指標の例である。

アウトカム	指標
母親になったばかりの者が、子どもの発達に関する知識を高める。	母親になったばかりの者で、プログラムの参加者の75%が講習の最後に子どもの発達に関する簡単な問い[short survey]にきちんと答えることができる。
対象の参加者は子どもの虐待やネグレクトの兆候に関する知識を高める。	教育キャンペーンが終わった6カ月後に、50%のコミュニティ・フォーカス・グループのメンバーは、児童の虐待やネグレクトの兆候を認識できる。
住民は、近隣地域が、子どもたちにとってより安全となっていると感じる。	60%の近隣住民が、子どもたちにとって近隣地域が1年前より安全となっていると確信していると答える。
プログラムの資金源の多様化	1年間で、4種の資金源（公共、民間企業、個人の寄付者、財団）のそれぞれが、プログラムの収入の55%を超えず、かつ、少なくとも15%を占めること。
法律扶助対応弁護士の文化的問題への対処能力の向上	80%の法律扶助対応弁護士が、ワークショップの最後に文化的な問題に関する学習について自主報告する。 6カ月以内に、訓練を受けた弁護士の60%のクライアントが、弁護士が文化的な問題に関して知識豊富であったと示す。
立法者の政策オプションに関する認識の増加	6カ月以内に、調査された立法者のうち45%が、異なる政策オプションがあることについて認識していることを示す。
義務教育レベルの生徒（K-6）が、オペラに関する知識を示す。	講座の終わりまでに： -65%の子どもは演じられるオペラの物語を認識できる。 -65%がオペラにおける3人の登場人物について説明できる。 -50%が、登場人物の歌っているパートについて、声のタイプ（ソプラノ、アルト、テノール、バス）を認識できる。
若者が、長期のアルコール、たばこ、その他の薬物（ATOD）の使用/乱用をもたらす結果に関する知識を高めた。	講習の最後に -90%の参加者が、自身がATODの使用に関連するリスク/害に関して知識を得たと報告する。 -80%が、アルコールまたはその他の薬物を使用しないことが重要であると回答する。

あなたが指標を策定する際、アウトカムが不明確であるか、あるいは曖昧であると認識することがあるかもしれない。このようなプロセスは、あなたのアウトカムを再検討、あるいはさらに明確化する機会を提供する。



評価計画の策定：各アウトカムについて検討しなさい—それらは、明確であって、曖昧でないものか？

指標の目標設定

指標は、あなたのプログラムが意図するアウトカム、すなわちあなたが各個人、グループ、あるいはコミュニティにおいて目にすることを望む変化、を生み出ししていることを示す[indicates]ものである。有意義なものとなるためには、関連するアウトカムの良好な達成状況を示すと考えられる変化の量を定めるべきである。

目標を定めるのは困難が伴うと感じられるかもしれない。だれも非現実的な期待に対し責任を負いたくはない。それと同時に、私たちは皆各プログラムが本当の変化につながってほしいと願うものである。

あなたはどのように現実的な目標を選択するか？以下の目安を検討しなさい：

- あなたの最低限度の基準[baseline]は何であるか？
- あなたの経験が何をあなたに語っているか？
- 当該分野の標準[standards]はどうなっているか？
- ステークホルダーは何を期待するか？

あなたのプログラムが新規のものであったり、以前この種の情報を集めたことが全くないためベースラインの情報を持たないこともある。その場合は、あなたが行う最初の評価がベースラインの情報を提供しうる。後年には、あなたは進捗を示す現実的な目標を定めることになるだろう。

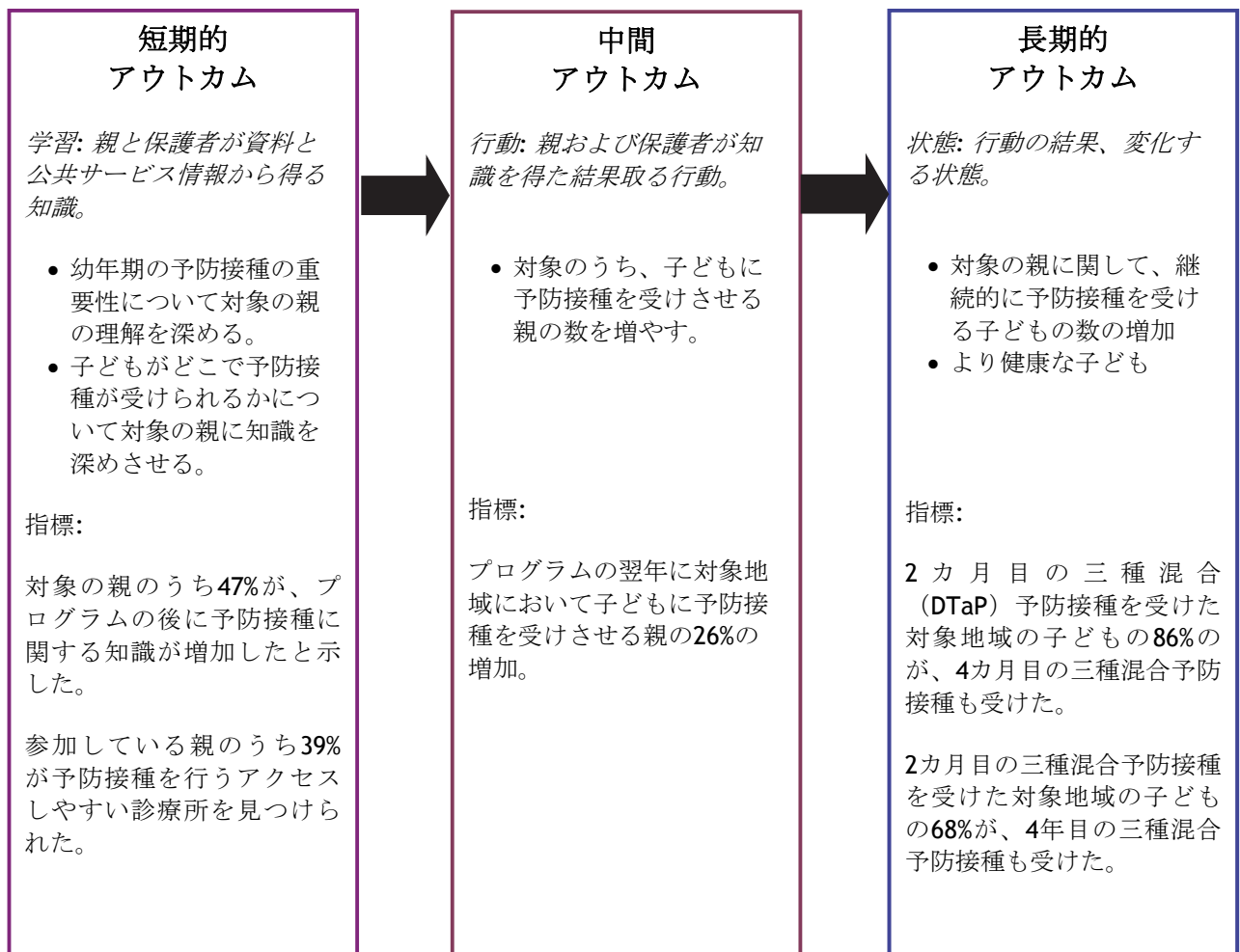
複数の指標

アウトカムが、それと密接に関連がある1つの指標のみを持つことがある。アウトカムがあなたの組織が提供する新しいサービスについてのメンバー内での認知を増加させることであるならば、あなたの唯一の指標は、あなた方のメンバーの80%が6か月以内に新しいサービスについて知っているということを報告することであるかもしれない。喫煙減少プログラムのアウトカムは、参加者が禁煙することである。この場合、指標はプログラムに参加した後に禁煙する参加者の割合になるだろう。この場合、1つの指標のみが必要となるだろう。

複雑なアウトカムには、複数の指標を定めうる：プログラムのアウトカムが参加者自身の求職スキルの改善であるならば、プログラムで強調されるスキルを反映しうる複数の指標が必要となる。指標は以下のようなものを含みうる：

- トレーニングコースの最後の模擬面接において基準を満たす参加者の割合
- コース終了後1カ月以内に質の高い履歴書を作れる参加者の割合
- トレーニングコースの最後までに、求人一覧の主要な3つの情報源を認識している参加者の割合

ロジックモデルを策定する際に、あなたはアウトカムの連鎖（チェーン）の定義を行っていた。この連鎖は変化の理論[theory of change]を明確にし、現実的な期待を設定するのに役に立つ。また、このチェーンはあなたが何を達成しているかについて知るといふ、評価の見地からも重要である。下のチェーンにおける各欄に対応する各指標がある。



直接的指標 vs. 間接的指標

指標が、あなたが評価しているアウトカムに直接関連し、同じタイプの変化についてのエビデンスとなっていることを確かめなさい。以下に、アウトカムの例題についての直接的指標と間接的指標の例を示す。

アウトカム	指標
参加している新しく子を持った母親が自身の子どもに予防接種を受けさせる（行動）。	<u>間接的</u> : 参加している新しく子を持った母親のうちで、予防接種の重要性を認識している者の数/%（認識） <u>直接的</u> : 1年以内に予防接種を受けさせている参加している母親の子ども数/%（行動）
参加している子どもが良いスポーツマンシップの原則を理解している（知識）。	<u>間接的</u> : プログラムを終了した後にチームに参加する子どもの数/%（関連しない行動） <u>直接的</u> : 参加している若者のうちで、その季節の終わりまでに5つの良いスポーツマンシップの行動を認識できる者の数/%（知識） 毎年プログラムで、運動選手である生徒の間でのけんかと言論の減少した数/%（行動。ただし、実地の知識を示す。）
十代の対象者が特定の環境に関する健康被害についての知識を高める（知識）。	<u>間接的</u> : プログラムの最初の6カ月の間にトピックに関する小冊子を受け取る学生数/%（アウトプット） <u>直接的</u> : プログラムの最初の年の終わりに、3つの健康被害について認識できる対象の学生数/%（知識）



評価計画の作成: あなたの評価計画テンプレートの指標の欄の中、または、オンラインの Evaluation Plan Builder の「指標/データ収集」の欄の中に指標を書き入れなさい。

データ収集の概要

評価のすべての要素の中で、データ収集が最も困難なものとなりうる。我々の多くが、統計の専門家や質的データ収集に従事する評価の専門家である必要があると信じているが、それは全く、正しくはない。

このワークブックは、あなたが評価計画を完成するのに役立つように、データ収集に関する基本概念を紹介する。別途、データ収集ワークブックが近いうちに Point K Learning Center で利用可能になるだろう。

データ収集の方法:

あなたが必要とする情報を集める最も良い方法は何であるか？

これまでのところ、あなたは、評価したいものは何か、そして、測定するものが何かについて確認した。執行に関する評価においては、これらは、活動および関連するアウトプット、追加の質問である。アウトカム評価では、これらは、プログラムのアウトカムおよび関連する指標である。

ここで、あなたはデータを集めるための方法を考えることになる。アウトプット、執行に関する質問、および指標は、あなたが測定する**対象**である。データ収集の方法はあなたがどのようにこれらを測定するかという**方法**である。

データ収集におけるゴールは、あなたが使用する収集のための道具の**数を最小化**して、それぞれから集める**情報量を最大化**することである！

必要とする情報を得るための最も良いデータ収集方法を選ぶ際には、以下の点を考慮しなさい:

- どの方法があなたのプログラムと対象の人々にとって最も悪影響がないものか？
- どの方法が費用を賄え、うまく実施できるものか？
- あなたの情報源から情報を得るためにどの方法が最善のものか（文化的な適切性および他の背景的問題の考慮）？

最も一般的なデータ収集の戦略は大きく以下のカテゴリーに分けられる。

1. 文書の調査

印刷された資料（プログラムの記録、研究レポート、センサス・データ、健康記録、予算が含まれる）の分析。文書の調査は執行評価のための活動とアウトプットに関するデータ収集の一般的な方法の1つである。

2. 観察

通常は、観察用のチェックリストと訓練された観察者を利用して、定式化され、かつ体系的な方法で状況、行動、活動を観察する。これは実際のイベントや状況を経験することが（それらについて話を聞くことよりむしろ）評価の重要な部分である場合に使用するのが良い方法である。

3. 関係者と話をする

インタビュー（直接、もしくは電話での）または集団面接[focus groups]を通して、参加者および他のステークホルダーから言葉での回答を収集する。特定のグループの人々から、複雑な、あるいは極めて個人的な考えを聞く際には役立つ方法である。

4. 関係者から記述による回答を集める

記述による回答を、調査（対面式、電子メール、オンライン、郵送、電話での）、テスト、または日報/ログを通して集める。日報に関するケースを除いて、あなたが多くの人々からの多くの情報を必要とするか、すべての回答者から同様の情報を利用可能とすることが重要な場合に、この方法はしばしば使用される。

5. その他の方法

写真、録音テープ、CD、視覚アートワークの、画像/マルチメディア形式のデータを調査する。専門家が、その専門知識で活動または製品を評価する分野では、専門家または同僚による調査[peer reviews]を実施する。ケーススタディ、すなわち、1つのユニットのインテンシブな調査は、学習目的で用いられ、そのユニットはしばしば模範例[exemplar]または避けられるべきモデルとして扱われる。



ブレインストーミング: あなたが測定する各項目に対応するデータ収集方法を検討しなさい。

データ収集の容易さ

あなたがデータ収集の方法を決める際に、次のステップとなるのは、データを集めるのに必要な労力のレベルを確認することである。データ収集のための新しいツールを作成するのに必要な費用と時間を考慮しなさい。また、実際にデータを集めて、分析するのに関係する費用を考慮しなさい。ある方法が他のものより費用が高くなることがある。例えば、インタビューは質問紙調査より多くの時間（そして、リソース）が必要となる。多くの方法に関して、データベースソフトが必要となる。あなたが選ぶそれぞれのデータ収集の方法には、あなたが使用しうる既存のツール（インテーク調査など）が手元にあるかどうかを検討しなさい。そうでなければ、新しいデータ収集ツールを作成して、使用するのに必要な労力の量について考えなさい。それが「低」「中」または「高」のどのレベルの労力を必要とするかどうか評価しなさい。



評価計画の策定: あなたがこれまでに確認した各方法について、それらを用いた場合にデータ収集に必要となる労力のレベルを評価しなさい。以下のようなスケールを利用する: 既に持っている（あなたの組織が既に利用可能な方法について）; 低; 中; 高。

データ収集の方法とそれらに対応する労力のレベルを評価計画テンプレートもしくはオンラインの **Evaluation Plan Builder** に記入しなさい。

計画の見直し

あなたの執行テンプレートとアウトカムテンプレートを一つの評価計画にしましょう。時間をかけて、現時点での計画を価値のある評価につながるように見直さない。

あなたの計画を見直すことについて次のような点に留意されたい:

- 執行評価のために、あなたが行っている活動の質を向上させるのに役立つ追加の質問があるか？
- あなたは、質問のリストを、少数の「知る必要がある」トピックに狭めたか？
- あなたはすべてのアウトカムに対して指標を定めたか？ 3を超える指標が必要なアウトカムがあるか？ その場合、そのアウトカムは、再定義されるべきか、あるいはいくつかの意図する変化に分けられる必要があるかどうか検討しなさい。
- 各指標の記述の大部分あるいはすべてが「誰」「どのくらい」「何が」「いつ」という主要な要素を含んでいるか？
- 各指標に対し、妥当な目標[targets]を定めているか？

プログラムの計画策定、または評価に何か質問があるか、個別対応サービスに興味をお持ちの場合は、私たちのウェブサイト (www.innonet.org) を訪れるか、以下に連絡してください。

Innovation Network, Inc.
1625 K Street, NW, Suite 1050
Washington, DC 20006
(202) 728-0727
info@innonet.org
www.innonet.org

- **Copyright Policy**
Innovation Network, Inc. grants permission for noncommercial, share-alike use of these materials provided that attribution is given to Innovation Network, Inc. For more information, please see our [Creative Commons License](#).

執行評価の計画[Implementation Evaluation Plan]

活動	各アウトプットおよび執行に関する質問	データ収集の方法	データ収集の労力 (既に持っている、 低、中、高)
プログラムの要素	各アウトプット		
	質問[Questions]		
プログラムの要素	各アウトプット		
	質問[Questions]		

アウトカム評価の計画[Outcomes Evaluation Plan]

アウトカム	指標	データ収集の方法	データ収集の労力 (既に持っている、 低、中、高)

訳者略歴

茂木康俊 (もてき やすとし)

MOTEKI Yasutoshi

1978 年群馬県生まれ。2001 年九州大学法学部卒業。2007 年九州大学大学院法学府博士後期課程単位取得満期退学。2008 年に九州大学から博士（法学）の学位を取得。九州大学大学院法学研究院助教などを経て、現在は広島大学大学院社会科学研究科准教授。専門は行政学・公法学。政策評価・行政評価について研究を行っている。

メールアドレス等の問い合わせ先については、ホームページをご参照ください。

ホームページ <http://www.moteki.org/>

Copyright Policy for Japanese Edition

Center for Research on Regional Economic Systems grants permission for noncommercial use of these materials provided that attribution is given to Innovation Network, Inc, Yasutoshi Moteki, and Center for Research on Regional Economic Systems. For more information, please see the [Creative Commons License](#).

Attribution-NonCommercial-ShareAlike 3.0 United States (CC BY-NC-SA 3.0 US)

<http://creativecommons.org/licenses/by-nc-sa/3.0/us/>

About the Japanese edition

Author: Innovation Network, Inc.

Translator: Yasutoshi Moteki

Publisher: Center for Research on Regional Economic Systems, Hiroshima University

Published Date: October 15, 2016

政策評価・行政評価のための評価計画策定ワークブック
—ロジックモデルに基づく執行評価・アウトカム評価—
(原題 : Evaluation Plan Workbook 英語版 2005 年発行)

2016 年 10 月 15 日 第 1 版第 1 刷発行

著者	Innovation Network, Inc.
訳者	茂木康俊
発行	広島大学地域経済システム研究センター (大学院社会科学研究所附属) 〒730-0053 広島市中区東千田町 1 丁目 1 番 89 号 電話 082(542)6991 FAX 082(249)4991
印刷	株式会社 中本本店
ISBN	978-4-9907890-1-5



ISBN978-4-9907890-1-5

広島大学地域経済システム研究センター