

平成 26 年度独立行政法人大学評価・学位授与機構田中弥生教授科学研究費事業

N P O の事業評価に関わる海外の動向調査

報告書

平成 2 7 年 3 月

一般財団法人 C S O ネットワーク

代表理事

今田 克司

NPOの事業評価に関わる海外の動向調査 報告書

目次

1 はじめに：背景と目的

- 1.1. 本調査の目的
- 1.2. 背景
- 1.3. 調査の方法

2 評価の潮流概観

- 2.1. 評価に対する意識の広がり
 - 2.1.1. 評価の要請と能力形成の広がり
 - 2.1.2. 評価専門家以外の評価観の変化
 - 2.1.3. EvalPartners と国際評価年
- 2.2. 評価の二本柱「説明責任」と「学び」、第三の柱の形成
 - 2.2.1. 「学び」の復権
 - 2.2.2. 第三の柱：マネジメント・ツールとしての評価
 - 2.2.3. 変化の理論
- 2.3. インパクト評価の「王道」と対抗軸の動き
 - 2.3.1. インパクト評価
 - 2.3.2. 展開評価
- 2.4. その他の潮流
 - 2.4.1. データ革命
 - 2.4.2. 数値化による可視化という意識
 - 2.4.3. ソーシャルインパクト分析との接点
 - 2.4.4. 評価のカントリー・オーナーシップと参加

3 考察：日本のNPO評価への活用

4 結論

<謝辞>

<資料編>

- 1 海外動向調査スケジュール
- 2 ヒアリング一覧

NPOの事業評価に関わる海外の動向調査

1 はじめに：背景と目的

1.1. 本調査の目的

本調査は、NPOの事業評価に関わる海外の一特に英国と米国の動向、潮流を概観し、それを日本のNPOに還元し、日本のNPOにおける評価文化を根づかせる動きに貢献することを主たる目的として実施された。また、調査で浮き彫りになった海外の動きを日本の評価専門家集団にも還元し、今後、海外の動きとの連動・連携を深めていくことが副次的な目的である。

1.2. 背景

筆者らが中心となり、2014年度、「評価をNPOの力にする研究会」というインフォーマルなグループを立ち上げた。その名が示すように、評価を近寄りがたいものとして敬遠してしまうのではなく、評価の考え方や手法、ノウハウを自分たちのものとすることによって、評価の専門家と対等につきあい、評価からの学びを組織や市民セクターの力にしていこうという発想である。研究会は特定非営利活動法人日本NPOセンター（筆者が常務理事を務める）と一般財団法人CSOネットワーク（筆者が代表理事を務める）の共催である。

研究会の主たる目的（目的A）は、「研究会参加者のあいだで、NPO/NGOにとっての「役に立つ」評価の理解が進み、これをもとにマネジメントツールとしての評価モデルを構築する基礎ができることである」¹。また、もう一つの目的（目的B）として、「研究会参加者のあいだで、NPO/NGOが企業の(1)本業における社会・環境インパクトを評価する、(2)社会貢献事業における成果を可視化する、ための評価リテラシーが上がり、必要と場面に応じて、企業へのアドバイザー的機能を果たしたり、改善のためのアドボカシー活動をできる素地が形成される」²があげられているが、これについては初年度である2014年度はあまり際立った取り組みをしていない。

研究会では、これまで国際NGOや助成財団を招いて研究会を実施しているほか、関連のイベントとして、日本NPOセンターの市民セクター全国会議（2014年11月22日）でセッションをもつなど、いくつかの試みを展開している。

上記のような経緯を踏まえ、今回の海外調査は、研究会のメンバーでもある独立行政法人大学評価・学位授与機構の田中弥生教授との協議から、田中教授の科学研究費を一部活用することで可能となった。

¹ 「評価をNPOの力にする研究会」計画文書より。

² 同上。

海外調査では、上記の研究会の主たる目的（目的 A）に重点をおき、もう一つの目的（目的 B）については、追加の質問として必要に応じて訊くにとどめた（その結果は 2.4.3.参照）。目的 A だけでも潮流にはたくさんの流れがあり、それ以上のメニューを組み合わせることは今回の調査の設定では困難だったからである。

今回の調査報告は、2015 年 3 月 14 日の日本 NPO 学会年次大会におけるセッション「NPO 評価の最前線と日本での活用」（モデレーター：田中弥生、今田克司）で発表するほか、田中教授との協議のもと、上記の調査目的を果たすのに資するような形で報告していく予定である。

1.3. 調査の方法

2015 年 2 月から 3 月にかけて、筆者はブリュッセル（ベルギー）、ロンドン（英国）、ニューヨーク（米国）、ワシントン DC（米国）の 4 カ所をまわり、NPO の事業評価に関わる海外の動向の調査を行った。これが今回の海外調査のメインである。

同時に、2014 年 9 月以降、海外調査のヒアリング先を選定する作業の一環で、おもにインターネットでの調査を通じて、いくつかの注目すべき動きをメモし、文献調査を行った。本報告書では、ヒアリングにとどまらず、それらについても言及している。さらに、筆者がグローバルなレベルでの CSO とドナーや国際機関との関わりに触れる勤務経験をもっているため、自身の経験もある程度本報告書の題材となっている。

ここでいくつかの留意点を記しておきたい。まず、NPO と CSO(Civil Society Organization = 市民社会組織)の用語の問題。本報告書の題名は、「NPO の事業評価...」であるが、海外の事例の際には CSO の用語を採用している。これは、CSO の方が海外においてはより一般的に使用されており、海外の文脈で NPO という用語を使うのに違和感があるためである。また、本報告書は NPO/CSO が行う事業評価を調査の対象として取り上げているが、これをよりよく理解するために、他のアクター（例えば国際機関や財団）が対政府、対 CSO 等で行っている評価も考察の対象としている。

また、筆者は昨年日本評価学会が管理する評価士の資格を取得しているが、決して評価の専門家ではなく、評価の文脈理解や事実関係で誤認があるかもしれない。さらに本報告書はヒアリングをメインの情報ソースとしているため、印象レベルにとどまっている事象も多いことを断っておきたい。

最後に、資料編ヒアリング一覧には、それぞれのヒアリングから得た参考文書やリンクの情報を掲載しておいた。依って、それらをすべて本文の脚注には再掲していない。直接のレファレンスである場合のみ、本文の脚注に載せてあることを記しておく。

2 評価の潮流概観

今回の海外動向調査において、評価の潮流（トレンド）として大きく抽出することができたものとして、3 点あげておきたい。1) 評価に対する意識の広がり、2) 評価の目

的としての学びの復権とマネジメント・ツールとしての評価の台頭、3) インパクト評価の対抗軸としての展開評価への注目、である。本項では、この3点について紹介し、加えてその他の潮流についても触れる。

2.1. 評価に対する意識の広がり

2.1.1. 評価の要請と能力形成の広がり

評価がプロの仕事として確立したのはここ20年程のことだが、特に2000年代以降、公的資金投入におけるバリュー・フォー・マネー（Value for Money [VfM]、投資に対する価値の最大化）の流れが強くなり、それに呼応するように評価の要請が強化され、その結果評価にそれまで以上の資金がすぎ込まれるようになった。

開発援助の分野においては、USAID（米国国際開発庁）、DfID（英国国際開発庁）などの援助機関がVfMを打ち出し、それに呼応するように大手の国際NGOが評価を強化していった。途上国においては、国際NGOから資金提供を受ける現地NGOに評価の波が伝播する形でトリクルダウンが起これ、評価の要請が広がっていった。

評価の要請は広まっても、それがすぐに評価の能力形成につながるわけではない。そこで、途上国国内での評価の能力強化の必要性が言われるようになり、アジア、アフリカ、ラテンアメリカなどの地域ごとの評価専門家機関（VOPEs = Voluntary Organizations of Professional Evaluators）を通じて評価のプロの養成が行われるようになっていった。

2.1.2. 評価専門家以外の評価観の変化

以上のようなドナー側の動きと評価専門家の増大は、今回の調査以前の段階でもある程度把握できていたが、今回の一連のヒアリングのなかから、評価に対する注目が他にも広がりつつあることが浮き彫りになった。

評価に特に注目しているのは、他ならぬ事業実施主体だが、これは後述の「役に立つ」評価の文脈につながっている。評価者以外が評価の重要性に気付き、視点や手法を事業運営に取り入れるようになってきており、もはや評価は評価専門家の専管事項ではなくなりつつあるということだ。

しかし一方で、評価の理解者が少ないという現実も厳然と存在している。今回ヒアリングを行った欧州連合の開発援助機関 EuropeAid や国連開発計画（UNDP）で運用している資金量に対して、評価に投入されている人員や資金は決して十分とは言えない³。

³ 開発援助関連の事業に年間90億ユーロ規模の助成を行う機関 EuropeAid の本部の評価室は17人のスタッフを抱えるのみで、事業単位の評価はおもに国レベルの担当官が行うが、事業評価の弱さが指摘されているという。また、全体の複数年計画策定過程においても、評価の必要性に関する内部の理解は十分とは言えないという。詳しくは資料編ヒアリング一覧のヒアリング1を参照。

一方、UNDPでは国別、テーマ別、地域／グローバル単位で事業評価を行っているが、独立評価室（IEO）は、スタッフ20人（評価専門家10人）の小さな事務所。国ごとおよび地域（Regional Bureau）ごとにM&E（モニタリングと評価）の担当官をおいているが、専門性にバラツキがあり、全体の底上げが課題となっているという。詳しくは資料編ヒアリング一覧のヒアリング5を参照。

米国の民間財団における評価の意識を捉えたものとして、Center for Evaluation Innovation (CEI) が 2013 年に発行した調査報告書がある。これを見ると、調査に参加した 31 財団のうちの 50%が、助成金に対する評価への資金投入が増加したと回答しており、評価の必要性を意識している財団が増加傾向にあることがうかがえる⁴。

2.1.3. EvalPartners と国際評価年

途上国を中心に VOPEs を通じた評価者養成の機運が高まるのと同時に、上記の流れに乗って、CSO の評価能力強化の必要性が言われるようになり、UNICEF (国連児童基金) と IOCE (International Organization for Cooperation in Evaluation) が共同で、2012 年に EvalPartners (国際評価パートナーシップ・イニシアチブ) を設立した⁵。EvalPartners の目的は、世界各国において CSO の評価能力と政策・世論形成能力の強化を図り、その結果として国レベルの政策が客観的な事実に基づいて形成されるようになることとある。EvalPartners では特に公平性 (equity) とジェンダーの平等、社会正義の実現に力を入ると、設立の際に採択されたチェンマイ宣言に記されている⁶。

EvalPartners が母体となって、国連ミレニアム開発目標 (MDGs) から持続可能な開発目標 (SDGs) へと交代する年であり、新たな開発目標へのモニタリングの機運の高まりが見られる 2015 年を国際評価年と定めた。これは、2013 年 9-10 月、ブラジルで開催された第 3 回国家評価能力国際会議において決定したもので⁷、これを受けて 2014 年 12 月には、国単位での評価能力構築を推進する国連総会決議が承認されている⁸。

このような流れのなかで、評価のための環境整備を加速させるためのツールキットも作成され、CSO による評価を充実させていこうという動きがグローバルなレベルで結実しつつある⁹。

2.2. 評価の二本柱「説明責任」と「学び」、第三の柱の形成

なぜ評価をするのか。この質問に対する答えは、伝統的に「説明責任を果たすため」と「学んで改善に役立てるため」の 2 つであったといえる。これを評価の二本柱と呼んでみよう。通常、前者のために行うのが総括評価 (summative evaluation) であり、後者のために行うのが形成評価 (formative evaluation) である。

2000 年代以降、この二本柱における力点は、「説明責任」の方に大きくシフトしていった。評価に熱心な先進国政府において、おもに国に台所事情から、さきの VFM のか

⁴ <http://www.evaluationinnovation.org/sites/default/files/Benchmarking%20Evaluation%20in%20Foundations.pdf>

⁵ <http://mymande.org/evalpartners/the-international-evaluation-partnership-initiatives>

⁶ <http://www.ioce.net/en/PDFs/Chiang%20Mai%20Declaration%202012.pdf>

⁷ <http://www.nec2013.org>

⁸ <http://www.unevaluation.org/mediacenter/newscenter/newsdetail/105>

⁹ http://www.mymande.org/sites/default/files/toolkit/UNICEF%20NY_Advocating%20for%20Evaluation_Web_1.pdf

け声が高まっていったためである。今回ヒアリングをした開発援助の国際機関、欧州連合の援助機関（EuropeAid）と国連開発計画（UNDP）はともに、ドナーからの VfM の圧力には答えなければならないと感じていた。また、今回のヒアリングには含まれていないが、文献調査や筆者の経験からいっても、開発に携わる国際 NGO に対する VfM の圧力も同様に強くなっている。開発援助のように、先進国の国内に支持基盤をもたない分野に対しては、いかに効果的に国民の税金が使われたかを証明せよとの圧力が強くなるのは至極当然のことと言えよう。

2.2.1. 「学び」の復権

ところが今回の一連のヒアリングで見えてきたのは、二本柱のもうひとつである「学び」の復権とでも言える事態だ。これはなぜだろうか。「説明責任」を果たすためにも、資金は効率良く使われなければならない。しかし、評価という作業から事業実施者が学ぶことがなければ、改善は望むべくもないというのもあたり前のことだ。EuropeAid の事業評価においても、説明責任一辺倒から、最近、活用特に学びや改善のための活用のための評価という位置づけが見直されつつあるという。その一環で、特に助成相手が CSO の場合、EuropeAid 側が評価するのではなく、CSO 自身が適切に評価できるような力量を身につけてもらう方に力点をシフトしてきているという。DfID においても、同様の動きが見られるようだ。

2.2.2. 第三の柱：マネジメント・ツールとしての評価

しかし今回の一連のヒアリングで「学び」の復権よりも強く、多くの面接者が指摘していたのが、評価とマネジメントの接近とでも言える事態だ。

総括評価、形成評価に加え、「実用重視」の評価という評価目的を打ち立てたのはパットンだが¹⁰、その流れを汲んで、「役に立つ評価」という視点から評価をマネジメントに直結させようという動きが加速している。CSO の世界にパフォーマンス主義、成果主義を浸透させようという考えが、評価のもつ実用性をフル稼働させようとしていると見立ててよいだろう。

この流れは、後述するデータ革命、特に今やリアルタイムで事業の現場からフィードバックを得られるような技術があることに着目している。KPI(Key Performance Indicators)をしっかりと作り、フィードバックから即学びや改善に結びつけることで、評価が生きてくる。中間評価、事後評価といったある特定のタイミングにとどまらず、評価と改善が事業運営の不断のサイクルに組み込まれるのだ。

そうであれば、より活用すべきは外部評価ではなくむしろ内部評価だ。そして、それは事業実施者の評価能力が問われるということを意味する。評価の専門家は外部にいて、評価のユーザとしての事業実施者が事業の始まりから終わりまで外部の専門家のアドバイスを受けながら事業を二人三脚的に実施するという姿は、まさにパットンが実用重

¹⁰ マイケル・クイン パットン（2001）『実用重視の事業評価入門』（清水弘文堂書房）

視の評価で説いたことであり、これをうまく回すことによって、評価は使われる評価となる。

このように、評価の知見をマネジメントに取り入れようという動きは特に英米の開発 NGO のあいだで顕著に見られる。英国の NGO のアンブレラ組織 BOND や、米国の InterAction は、そういった動きの牽引役にもなっている。BOND には、Effectiveness チームという NGO の事業の有効性や妥当性を考えるチームがあり¹¹、InterAction には Evaluation and Program Effectiveness Working Group (EPEWG) というワーキンググループがある¹²。どちらにおいても、効果 (Effectiveness) という単語がキーワードになっているが、評価の有効活用によって事業効果の向上を企図している。

この関連で、上述の VfM についても、BOND は、DfID などからの圧力の強いことを念頭に、英国 NGO としてどう振る舞うべきかの指針を出している¹³。基本的なスタンスとして、NGO は、政府から一方的に VfM の考え方を押し付けられるのではなく、その概念を正しく運用し、事業ごとの VfM を比較することで事業の改善につなげ、さらに評価をうまく活用することで VfM の推進者として自らを位置づけられるようにしなければいけないと説いている。評価の活用が、VfM のかけ声で一方的に守勢に立たされてしまうことを防ぎ、自ら VfM の推進者として実践を示していくという攻めの姿勢に転じることを可能にするのだ。

2.2.3. 変化の理論

評価がマネジメント・ツールなのであれば、事業マネジメントの核である事業の形成過程に評価の知見が活用されなければならない。評価の世界で評価可能性のアセスメントと呼ばれる場面だが、今回のヒアリングで、複数の面談者がこの重要性を指摘していた。

さらに、今回の 11 件のヒアリングを通じて、もっとも多く聞かれたことばのひとつが変化の理論 (Theory of Change) であろう。変化の理論自体は、変化の理論センターというウェブサイトもあり¹⁴、決して新しい概念ではないが、DfID や UNDP などのドナー・機関も含め、その有効性に注目しているようだ。

変化の理論センターは、変化の理論について、1) 長期的目的を達成するための条件、事業の結果、アウトカムなどの構成物を抽出し、2) 構成物間の関連を矢印でつなぎ、3) 変化の道として図式化したものと規定している¹⁵。インプット→活動→アウトプット→アウトカム→インパクトという典型的なロジックモデル¹⁶も変化の理論の一種と言えるが、変化の理論の特徴は、直線的因果関係をベースに論理構築をするロジックモデルよりも柔軟に事業のアウトプットやアウトカムを設定できる点だ。したがって、矢印の方

¹¹ <http://www.bond.org.uk/effectiveness>

¹² <http://www.interaction.org/work/monitoring-evaluation>

¹³ <http://www.bond.org.uk/effectiveness/value-for-money>

¹⁴ <http://www.theoryofchange.org>

¹⁵ <http://www.theoryofchange.org/what-is-theory-of-change/>

¹⁶ プログラムセオリー、ログフレームなどの呼び名で呼ばれることも多い。

向は一様ではないし（逆進もあり）、複数の構成物から矢印が同じ目的や事象に向かって出ていたりもする。また通常、事象に関わるステークホルダーの存在や関わりの程度を明示化することで、事業実施過程において、どのステークホルダーといついかにどう関わるべきなのかを示唆するツールにもなる。

変化の理論は、後述する「発達評価」との関連性も大きい。従来のロジックモデルの応用が効かないような分野、すなわち直線的因果関係が描きにくく、環境の変化や不確定性の要素を勘案しなければならない複雑な事象でより力を発揮する。ガバナンス、能力強化、組織の変化、政策形成などがそういった分野に当たり、こういった分野での事業が特に開発援助の場面で増えてきたことも、変化の理論の台頭と無縁ではない。ただし、一部に言わせれば、社会的事象に直線的因果関係が描けるものの方がむしろ例外で、ようやく評価の意識がそのことに気づいただけ、ということかもしれない。

2.3. インパクト評価の「王道」と対抗軸の動き

筆者は、2009年にエジプトのカイロで開かれたインパクト評価の国際会議に出席したことがある。その場でもっとも印象に残ったのが、数値化と統計的処理を駆使したインパクト評価を評価の「王道」と位置づける“主流派”とそれを必ずしも万能とは見ない一派との覇権争いであった。

今回の一連のヒアリングでも、この対立は色濃く評価の世界を彩っている印象を残した。

2.3.1. インパクト評価

念のため記しておくが、ここで言うインパクト評価は、DAC 5項目のインパクトの評価ではなく、ネット・アウトカム、すなわち、事業の介入以外の影響を排した純アウトカムを測定する一連の評価手法のことである。RCT（ランダム化比較試験）を筆頭に、さまざまな実験的手法が開発されており、「科学的」アプローチとして評価の「王道」を歩んでいる。

インパクト評価という手法自体は、以前より存在していたが、2006年に Center for Global Development が「いつ我々は学ぶのか？」というペーパーを発表し、社会開発の分野でたくさんの資金が投じられているにもかかわらず、科学的知見がほとんど得られていないことに疑問を呈した¹⁷。インパクト評価によってそれまでの開発のやり方を変えようと提唱する文書で、これを契機にインパクト評価に向かう流れが加速していった。

以来、世界銀行、DfID、ゲイツ財団など、開発資金を多くもつ影響力のある機関が推進していることもあり、評価の世界における主流派の立場から動いていない。インパクト

¹⁷ “When Will We Learn - Improving Lives through Impact Evaluation,”
<http://www.cgdev.org/publication/when-will-we-ever-learn-improving-lives-through-impact-evaluation>

ト評価を推進するネットワークとして、3ie（インパクト評価国際イニシアチブ）が存在する¹⁸。

政策決定者の客観指標、量的指標志向は強くなっており、今後もその方向に変化の兆しは見られない。また、インパクト評価は因果関係を証明する手段だから、「なにをしようになった」とプラスまたはマイナスの原因を特定できることが、政策の説明責任達成にとって大きな力となる。

このようなアプローチに対して「社会は実験室ではない」と反発する向きがあるのは当然だろう。また、インパクト評価はほかの評価手法に比べてコストが高くなりがちという理解があり、果たしてその投資に見合う評価結果が出せるのかを疑問視する声も根強い。さらに、インパクト評価を推進する立場からも、実験的・統計的手法による厳密な因果関係の確立から貢献度の測定にまで間口を広げようという、いわばインパクト評価再考を促した文書が出ている。これは、英国評価学会会長を務めるなど評価の世界では重鎮のエリオット・スターンが2012年にDfIDの委託で書いた「インパクト評価のデザインと方法の範囲を広げる」というペーパーで、インパクト評価にまつわる論争に一石を投じたものとして知られている¹⁹。

とはいえ、評価の精緻化は因果関係の確立によってなされると考える思想は評価の世界に強く根を張っており、評価手法の革新も多く、「王道」の地位はしばらくゆるがないだろう。

2.3.2. 展開評価

インパクト評価が評価のひとつの世界観（世界は操作可能、予測可能で物事の原因は還元可能）を表しているとするれば、その対極にある世界観（世界は操作困難、予測困難、原因は特定困難）を体現している評価の体系がある。展開評価（Developmental Evaluation）の名の下にこの世界観が多くの評価専門家の信任を勝ち得ていることが、今回の一連のヒアリングで明らかになった²⁰。これはそもそもパットンが1994年から提唱している評価の考え方だが、2011年に一冊の書物にまとめてことで、体系化の証となったものだ²¹。

端的に言って、展開評価とは、パットンが「実用重視」の流れで開発した評価の類型で、特に複雑な事象において目標設定や介入の仕方が単純には決められないような事業

¹⁸ <http://www.3ieimpact.org/>

¹⁹ “Broadening the range of designs and methods for impact evaluations,” <http://r4d.dfid.gov.uk/Output/189575/>

²⁰ Developmental Evaluation という用語を耳にして、すでに日本語への訳語があるか少しチェックしてみたが、見つからなかった。途上国で実施される事業についての評価ではないので、開発評価では誤解を生むし、発達心理学系統の評価でもないので、発達評価としても混乱のもとになるだろう。あとは「発展」か「展開」あたりが考えられるが、状況の変化に応じて事業を展開させていくのと同様に、評価の展開させていくというおおよその含意であるため、展開評価としてみたい。

²¹ Michael Q. Patton (2011), Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use, New York: The Guilford Press.

<http://www.guilford.com/books/Developmental-Evaluation/Michael-Quinn-Patton/9781606238721> 参照。

の評価用に開発された。単なる一手法ではなく、より高レベルの評価の枠組み、類型である。複雑な事象に対し、ロジックモデルなどの単純な因果関係の評価図式では評価が難しいとの意識は従来からあり、それが変化の理論の主流化へとつながっているのだが、変化の理論等のツールも含め、評価作業の全体に一貫したアプローチを採用したものでいえよう。

展開評価の主唱者たちは、いわゆる複雑系の理論（complexity theory）に展開評価の理論的支柱をおいている。また、前述の総括評価や形成評価との比較で語ることが多い。これは、以下のように図式化できる。

図1：事業の発展段階と適した評価

事業の骨格ややり方が	事業の活動、実施者、背景は、	どんな評価が適しているか
<ul style="list-style-type: none"> すでに確立している 結果もある程度予測可能 	<ul style="list-style-type: none"> 事業の活動はすでに確立している。あまり変化は予定されていない。 実施者には「どうやればうまくいくか」の経験値がある 事業の成果や価値を体系的に捉える用意ができています 	総括評価
<ul style="list-style-type: none"> まだ改善の余地あり 基準づくりの過程 	<ul style="list-style-type: none"> 事業の主要な要素はまだ形成途上 実施者は方法や活動を改訂中 成果はだんだん予測可能になりつつある 事業が行われている背景もだいたいわかりつつある 	形成評価
<ul style="list-style-type: none"> 発展中 生成中 	<ul style="list-style-type: none"> 実施者はまだ異なる方法や活動の実験中 「どうやればうまくいくか」についても不確定要素が大きい 事業を進めるうえでの新たな疑問や課題、機会等が出てきている 	展開評価

Hallie Preskill & Tanya Beer (2012), Evaluation Social Innovation (FSG & Center for Evaluation Innovation) より一部改訂

展開評価が脚光を浴びてきたもうひとつの流れに、米国民間財団を中心に 2000 年代後半からブームになった戦略的フィランソロピーの動きとそれに対する反発がある。この頃から米国の大手財団は、それまでより財団としての優先順位に基づいたアウトカムやインパクトを前面に打ち出した選択的な助成活動を行うようになった。戦略的な投入をすれば、それに見合った社会的リターンが見込めるはずで、それまでフィランソロピーがそのような成果を出せていなかったのは、戦略的な発想に欠けていたからであると

いう考え方で、フィランソロキャピタリズム (philanthrocapitalism) の名前で知られた流れになった。

しかし、結局のところ、社会は「機械」のように動かず、そもそもの目標設定や介入と結果の組み立てに対して予想外の事態が多く持ち上がった。展開評価は、そういった反省をもとに、複雑な社会に見合ったプログラム構築や評価のためには、違った手法が必要ということで注目されているのだ。上記の図で示したように、生成中 (emergent) な事態に対応すべく開発された評価の考え方といってよい。

展開評価に関しては、実際にどう評価を行い、評価者の役割は何かなど、まだ未知数の部分も多い。しかし、「世界は操作困難、予測困難」の世界観に立つ評価の専門家にとって、今後大きな思想的バックボーンを提供していくことはまず間違いないと考える。

展開評価の枠組みのなかに位置づけられる手法のひとつとして、Outcome Mapping 手法から派生した Outcome Harvesting が編み出されている。Outcome Harvesting は、あらかじめ特定された目的をもとに事業のアウトカム達成度を見るのではなく、「起きたこと」を綿密に拾い上げていくことで、その「起きたこと」と事業の関連を見ていこうとするものである²²。

2.4. その他の潮流

2.4.1. データ革命

今回のヒアリングで、複数からデータ革命の話題が出た。従来、評価における悩みの種の一つはいかにデータを取得するかだったが、ICT 革新によって、データは容易に手に入るようになってきている。特に、1) 事業の参加者や受益者がデータの自主的な発信源となることをソーシャルメディアが可能にしていることと、2) データを時間が経ってではなく、リアルタイムで発信・集積・分析できるようになったことは、特にマネジメント・ツールとしての評価という位置づけや今後の展開評価の進化を予測するうえで、大きな意味合いをもつ。

こういった観点から見て注目すべき動きは多い。Feedback loop といったキーワードで検索するとたくさん事例が出てくるが、それらはほとんどがここ数年のものだ²³。また、Feedback Lab²⁴ や Accountability Lab²⁵ といった、ともに Lab という命名からわかるような「実験室」的試みもある。前者は、事業の受益者のフィードバックをリアル

²² Outcome Harvesting については、以下など参照。

http://www.outcomemapping.ca/resource/outcome-harvesting#inline_en

<http://betterevaluation.org/resource/overview/OutcomeHarvesting>

<http://aea365.org/blog/ricardo-wilson-grau-on-outcome-harvesting/>

²³ 例えば Reboot のナイジェリアでの活動など。

<http://reboot.org/2015/02/13/how-to-listen-so-that-people-talk/>

²⁴ <http://feedbacklabs.org>

²⁵ <http://accountabilitylab.org/learn-AboutUs.html>

タイムで事業実施者等に還元する仕組みを模索しており、後者は、同様の試み（受益者に限らず市民一般からのフィードバック）を説明責任のテーマでしている。

ただし、統計的手法もないままやみくもにデータを集める向きもあり、信頼できるデータを取るには専門性が必要だということが熟慮されていないという声も聞かれており、これは今後の課題だろう。

2.4.2. 数値化による可視化という意識

評価の量的手法と質的手法に関していえば、教科書的な答えは、どちらが優位ということではなく、評価の目的や分野、種類によって使うべき手法が変わってくる。できれば、両方を生かす混成手法が好ましいというのが正解だろう。

ここ数年の Dfid の動きを見ると、上述したようなインパクト評価に対する一定の見直しとともに、一時の数値主義一辺倒ではなく、量的と質的の混成手法を推奨する傾向があるという声が今回のヒアリングで聞かれた。とはいえ、政策決定者の量的指標志向は全般的にこれからも弱まることはないだろうから、市民は、自分たちが大切だと思う問題を数量化して示す必要がある。

そこでひとつ注目したいのが New Economics Foundation (NEF) による、市民が関心ある領域で問題を可視化する手段としての測定・数量化の役割に注目・実践している一連の活動だ。評価に関連する分野として、NEF では、2006 年頃より、National Accounts of Well-being（一種の幸福度指標）を開発している。社会の発展の度合いを経済指標以外の指標で見る試みとして、イギリスではある程度浸透しており、これを使った測定を行う自治体も増えている²⁶。SROI（社会的投資リターン）もそうだが、数量化（あるいは貨幣価値換算）の際に、多くのステークホルダーが関与し、その過程と結果にオーナーシップを持つことで、普及が加速している。

2.4.3. ソーシャルインパクト分析との接点

日本においては、企業やベンチャー・フィランソロピスト、投資家の間で関心が高まっているソーシャルインパクト関連の動き²⁷に敏感に反応し、グローバルな動きと同じ速度で情報共有がされている感がある。そこで、今回のヒアリングでもこういった動きとより伝統的な評価の世界の接点について質問してみたが、反応はまちまちであった。一部は、こういった民間セクター主導の動きには評価の専門家は鈍感だと言う。これは、伝統的な評価の世界がおもに政府系の資金で成り立っていることに起因するのではないかという見解だ。一方で、インパクト投資等の流れは、いかに測定を効果的にやるかという点においてこれまでの評価の知見に注目し、少なくともここ数年活用してきているという声も聞かれた。また、昨年ぐらいより、企業の社会活動評価の流れに評価の世界が注目するようになってきており、昨年の米国評価学会（AEA）の大会では、ソーシ

²⁶ <http://www.neweconomics.org/publications/entry/national-accounts-of-well-being>

²⁷ 例えば Social Impact Analyst Association (<http://www.siaassociation.org>) といった情報交換の仕組みも立ち上がっている。

ヤルインパクトや SROI に関するセッションがいくつか見られたが、これは新しいことだ、という感想もあった。

こういったなかで、評価の経験と民間セクターの動きをうまくハイブリッドさせて注目されているものに、FSG と Aspen Institute による Collective Impact Forum(CIF)がある²⁸。コンサル会社である FSG が企業の戦略づくりで培った知見と、システムの変化がいかにか起こるかの研究などを題材に構想が練られたもので。大きな規模でシステムの変化を起こそうと思ったら、組織単体ではなく、より大きな単位でビジョンの共有、測定指標の共有などが起こらないといけないという発想から成り立っている。

CIF で体現しているような組織という点でなく地域コミュニティのような面での変化を起こしそれを測定するという発想は、じつは評価の世界の特に展開評価の発想に近い。実際、複雑な社会変化を起こすために面的な発想と、ビジョンや指標の共有を前提として複数の事業を同時に進めようという試みは、地域研究などの分野においては伝統的に行われており、これを評価という土俵で実践することは、これまで多く議論されてきたテーマを評価という新しい皮袋に入れる試みと捉えられるのかもしれない。

2.4.4. 評価のカントリー・オーナーシップと参加

2005 年の援助効果パリ宣言以降、開発援助の世界では援助を受ける側のカントリー・オーナーシップが言われている。そして昨今は「カントリー」は政府だけでなく CSO や市民が含まれることが当然視されている。その流れで、評価についてもローカル・オーナーシップが言われ始めている²⁹。データ革命がリアルタイムのフィードバックを可能にしていることとあいまって、「現場の声」を直接事業に反映させることを規範としようとする機運が高まっている。

これは評価の国単位での能力強化とも呼応する動きで、それ自体は好ましい。しかし一方、評価への参加には時間と手間がかかり、「現場」の声としては、「生活していくのがやっとなのだから、参加はいいからどうすればうまくいくのか教えてくれ」といった声も聞かれているようだ。参加には常にこういった問題が内在するため、事業実施者および評価者には状況に応じた判断が求められる。

3 考察：日本の NPO 評価への活用

以上、今回の海外動向調査から浮き彫りになった事業評価の潮流について概観してみた。さて、このような海外の状況から、日本の NPO が学ぶべきことはなんだろうか。これについては、「評価を NPO の力にする研究会」等の議論を通じてより深めて論じるべき点を明らかにしていきたいが、ここでは以下の 4 点について言及しておきたい。

²⁸ <https://www.collectiveimpactforum.org>

²⁹ 例えば、InterAction は、今年（2015 年）に入って、Local Ownership in Evaluation を発行している (<http://www.interaction.org/sites/default/files/Local%20Ownership%20in%20Evaluation%20-%20FINAL.pdf>)

3.1. マネジメント・ツールとしての評価

「役に立つ評価」という視点から評価をマネジメントに直結させようという動きが加速しているのは日本の NPO にとっても示唆に富む。リアルタイムのフィードバックによる改善、KPI 重視、内部評価重視といった一連の視点の応用によって評価と改善が事業運営の不断のサイクルに組み込まれる姿というのは、日本の NPO も目指すべき姿ではないだろうか。

その際、評価専門家が多くの NPO にとって遠い存在である現状を鑑みるに、評価の専門家と評価のユーザとしての事業実施者が事業の始まりから終わりまで二人三脚で協力するというのは、どのような形で構想できるのだろうか。また、そもそも外部評価の役割をいかに捉え直せばよいのだろうか。事業実施者の評価リテラシーを上げることは必要かつ有用だが、外部者として評価者が関わる重要性も大きい。評価のすべての機能を団体が内在化してしまうのではなく、評価に関しては内外の役割分担をうまく図ることが成功の鍵になるのではないだろうか。

3.2. 変化の理論

変化の理論の有用性に関しては、すでに研究会でも意識してきたが、事業の入口の段階で単線系 (linear) でない変化の道を描くことの重要性は大いに納得できる。ただし、日本の NPO においてはそもそもロジックモデルを構築したことがない場合も多く、いきなり複雑な変化の道を描くことは難しいかもしれない。一定のトレーニング・モジュールを作成して、能力強化を体系化していくことを構想すべきではないだろうか。

3.3. 展開評価

上述したように、展開評価についてはその運用面でまだ未知数も多い。しかし、多くの日本の NPO は今後特に発展中、生成中の社会課題と向き合って事業を行っていくことが多くなることが予測され、この評価体系のポテンシャルは大きいと考える。加えて、日本の評価専門家集団を巻き込みながら、この評価体系の普及を図って行くべきだろう。

3.4. 数値化による可視化

データ革命の活用は言わずもがなだが、市民は、自分たちが大切だと思う問題を数値化することによって世の中に示していく必要があるという見解には賛同する。数値化は効果的なコミュニケーションの手段であり、的確な数値の提示は NPO の主張に説得力を増すことができる。また、多くのステークホルダーが関与し、その過程と結果にオーナーシップを持たせることで、普及を加速させるという手法は日本の NPO も採用すべ

きだろう。評価の世界で編み出されているこれら量的手法の試みを参考にしながら、日本の NPO の発信力の強化が模索できるのではないかと考える。

4 結論

今回の海外調査を通して、現在の世界の評価地図をある程度把握することができた。今後は、それをいかに日本の NPO に還元し、日本の NPO における評価文化を根づかせる動きに結びつけるかを、「評価を NPO の力にする研究会」などを通して考察していくことになる。また、展開評価の動向などに関して、日本の評価専門家集団とも情報・意見交換したいと考えている。

事業評価に関しては、アウトカムやインパクトをいかに測るかに多くが関心を払ってきており、その流れは今後モ続くであろうが、近年、知識の生成、共有、管理といった評価の活用が見直されてきているという。学びとは得た知識をいかに共有、管理し、それを実践に役立てるかによってその価値が決まる。日本の NPO における評価文化の普及とは、まさにその意味での学びを流通させていく営みになっていくのであろう。

<謝辞>

今回の海外調査は、独立行政法人大学評価・学位授与機構の田中弥生教授との協議から可能になったもので、田中教授の科学研究費を活用させていただいている。このような機会を与えてくださった田中教授、および大学評価・学位授与機構の方々に深く感謝したい。

また、資料編ヒアリング一覧に示したが、今回の調査にあたり、ヒアリングの依頼を受け容れ、貴重な時間を日本からの評価の学徒に割いてくださったヒアリング対象の方々にも深く感謝したい。一様にとっても快くヒアリングを受け容れていただき、日本の NPO や市民社会の発展に向けた激励の言葉を多くいただいた。これらの方々にもここで感謝するとともに、今回の調査で得た学びを活用し、日本の市民社会の進展のために寄与していきたいと考えていることを述べておきたい。

N P O の事業評価に関わる海外の動向調査

(資料編 1)

海外動向調査スケジュール

	日時 (2015 年)	機関・団体	場所
1	2月25日(水)	Evaluation Unit, Directorate General for Development and Cooperation (DevCO)- EuropeAid	EuropeAid Rue de la Loi/Wetstraat 41, 1040 Bruxelle ベルギー、ブリュッセル
2	2月26日(木)	New Economic Foundation(NEF)	New Economic Foundation 3 Jonathan Street, London SE11 5NH 英国、ロンドン
3	2月27日(金)	Overseas Development Institute (ODI)	Overseas Development Institute (ODI) 203 Blackfriars Road, London SE1 8NJ 英国、ロンドン
4	2月27日(金)	Keystone Accountability	Keystone Accountability 222 Kensal Rd, London W10 5BN 英国、ロンドン
5	3月3日(火)	Independent Evaluation Office, UNDP (United Nations Development Programme)	UNEP/UNDP 220 East 42nd Street, Room 2046 New York, NY 10017 米国、ニューヨーク
6	3月3日(火)	UNEG (United Nations Evaluation Group)	UNEP/UNDP 220 East 42nd Street, Room 2036 New York, NY 10017 米国、ニューヨーク
7	3月5日(木)	En Compass, LLC	スカイプ面談
8	3月6日(金)	The Evaluators' Institute The George Washington	The Evaluators' Institute The George Washington

		Univ.	University 2036 H Street NW (Samson Hall 105) Washington, DC 20052 米国ワシントン DC
9	3月6日(金)	FSG	スカイプ面談
10	3月6日(金)	Aspen Planning & Evaluation The Aspen Institute	Aspen Planning & Evaluation Program 1333 New Hampshire Avenue, NW Washington DC 20036 米国ワシントン DC
11	3月6日(金)	Center for Evaluation Innovation	Center for Evaluation Innovation 1625 K Street, NW, Suite 1050 Washington, DC 20006 米国ワシントン DC

N P O の事業評価に関わる海外の動向調査
(資料編 2)
ヒアリング一覧

ヒアリング 1

日時	2015 年 2 月 25 日 (水)	
場所	EuropeAid Rue de la Loi/Wetstraat 41, 1040 Bruxelles ベルギー、ブリュッセル	
機関・団体	Evaluation Unit, Directorate General for Development and Cooperation (DevCO)- EuropeAid	
面談相手	Catherine Pravin, Chef d'unite adjoint, Juergen Lovasz,	Catherine.PRAVIN@ec.europa.eu Juergen.Lovasz@ec.europa.eu, Catherine Pravin のアシスタント Christiane.Oris@ec.europa.eu
面談のポイント	<p>1) DevCO/EuropeAid および EC の評価について</p> <ul style="list-style-type: none"> • DevCO/EuropeAid は EC のなかで開発援助関連の事業に年間 90 億ユーロ規模の助成を行う機関。拠出先には国連機関等の国際機関、途上国政府、CSO などが含まれ、事業が実施されるのは EU の拠点のある 125 の国・地域にのぼる。 • 評価に対する意識が高まったのは 2000 年以降。EU 議会が監査室を通してプレッシャーを強めてきたのが 2005 年。「説明責任を果たす」ことが評価の一大目標になった。 • 本部の評価室は 17 人のスタッフを抱えるのみで多くはできないので、事業レベルの評価やより日常的な ROM(Result-Oriented Monitoring)は国の担当官が中心で行い、本部では戦略レベル (5-7 年のサイクルで戦略プランが作られている) の評価に徹している。 • ここ数年、一番大きなレベルである戦略評価と、ROM のあいだの事業評価の弱さが指摘され、その改善のために尽力している。 • これを受けて、評価についての意識や基本的やり方等を徹底するため、指導文書を作成したり、トレーニング機会を提供したりして、評価能力の向上に努めている。昨年、Evaluation Matters と題したガイダンス文書を作成した (1)。 	

	<ul style="list-style-type: none"> • 評価では、助成事業の設定されたアウトカムやインパクトに対する貢献度を適切に測ることと事業実施のプロセス評価の両方に力を入れている。 • 一番最近では、2014-2020年の7年計画が作られた。評価室からも、おもに評価可能性のアセスメントの観点から計画作成に関わったが、評価の必要性に関する内部の理解は十分とは言えない。 <p>2) 評価の潮流について</p> <p>2.1. 説明責任と学び</p> <ul style="list-style-type: none"> • DevCO の評価では、説明責任を果たすことが主眼とされてきたが、最近、活用、特に学びや改善のための活用のための評価という位置づけが見直されつつある。 • その一環で、特に助成相手が CSO の場合、EC 側が評価するのではなく、CSO 自身が適切に評価できるような力量を身につけてもらう方に力点をシフトしてきている。 <p>2.2. 国際評価年</p> <ul style="list-style-type: none"> • 国連等が中心になって、2015年の国際評価年として数々の啓発活動を企画している。2015 年は欧州開発年（European Year of Development）でもあるので、この2つを抱き合わせた開発の分野における評価の重要性に関する啓発事業をいくつか持つ予定だ。
<p>参考文書／ リンク</p>	<p>(1) https://ec.europa.eu/europeaid/node/71167</p>

ヒアリング 2

日時	2015年2月26日(木)	
場所	New Economic Foundation 3 Jonathan Street, London SE11 5NH 英国、ロンドン	
機関・団体	New Economic Foundation(NEF)	
面談相手	Olivier Vardakoulias	Olivier.Vardakoulias@nefconsulting.com
面談のポイント	<p>1) NEF の取り組みについて</p> <ul style="list-style-type: none"> • 評価に関連する分野として、NEF では、2006 年頃より、National Accounts of Well-being を開発している。社会の発展の度合いを経済指標以外の指標で見る試みとして、英国ではある程度浸透しており、これを使った測定を行う自治体も増えている(1)。 • 関連で、地域からの資源流出を抑えることを念頭に、Plugging the Leaks というイニシアチブも始めた(2)。 • これらの試みで、NEF は、市民が関心ある領域で問題を可視化する手段としての測定・数量化の役割に注目し、実践している。 <p>2) 評価の潮流について</p> <p>2.1. 量的志向</p> <ul style="list-style-type: none"> • 政策決定者の客観指標、量的指標志向は強くなっており、今後もその方向は変わらないだろう。だからこそ、市民は、それを利用して自分たちが大切だと思う問題を数量化して示す必要がある。 • 数量化の際に、多くのステークホルダーが関わり、その過程と結果にオーナーシップを持つことが重要だ。 • 数量化は必要な流れだが、SROI のような貨幣価値化は必ずしも進むべき方向ではない。社会や環境の価値を万人が合意できるような経済的価値には転換できないからだ。 • また、社会は実験室ではないので、RCT (ランダム化比較試験) の手法にはどうしても限界がある。それにその種の評価には多大な資金がかかり、果たしてその投資に見合う結果が出せるのか疑問だ。 	

	<p>3) その他</p> <ul style="list-style-type: none"> • Social Impact Analyst Association (3) など、ソーシャルインパクト測定のコミュニティのネットワーク化も進んでいる。NEF も含め、社会課題の測定に関しては多くの人の知見を動員して大きな流れを作っていくべきだ。
<p>参 考 文 書 / リ ン ク</p>	<p>(1) http://www.neweconomics.org/publications/entry/national-accounts-of-well-being (2) http://www.neweconomics.org/publications/entry/plugging-the-leaks (3) http://www.siaassociation.org</p>

ヒアリング 3

日時	2015 年 2 月 27 日 (金)	
場所	Overseas Development Institute (ODI) 203 Blackfriars Road, London SE1 8NJ 英国、ロンドン	
機関・団体	Overseas Development Institute	
面談相手	Simon Hearn	S.Hearn@odi.org.uk
面談のポイント	<p>1) 本人および ODI について</p> <ul style="list-style-type: none"> • ODI は英国の開発シンクタンクとして知られている。特に調査研究と政策形成のつながりを強めることを志向する研究事業が多い。 • Hearn 氏は、ODI の職員として、事業計画の手法として知られている Outcome Mapping(OM) コミュニティのコーディネーターをここ 7 年務めている (1)。OM は、評価の手法としてもその有効性が認識されるようになってきている。 <p>2) 評価の潮流について</p> <p>2.1. インパクト評価について</p> <ul style="list-style-type: none"> • ODI では、インパクト評価に関するペーパーなども多く出しており、分野の進展のために積極的に発信している。RCT(ランダム化比較試験) がうまくいくかどうかは、入口の段階での評価可能性のアセスメントの役割が大きい。 • インパクト評価は、世界銀行、英国国際開発庁 (DfID)、ゲイツ財団など、開発資金を多くもつ影響力のある機関が推進しているため、相変わらず評価の主流の立場を動いていない。 • ただし、ここ数年の DfID の動きを見ると、量的手法と質的手法を両方生かす混成手法を推奨する傾向があるようだ。また、変化の理論など、評価をより精緻な論理の組み立てのもとに展開することに注意を払っていると見受けられる。 • こういった、実験的・統計的手法による厳密な因果関係の確立から貢献度の測定にまで間口を広げようという、いわばインパクト評価再考を促した文書として、Elliot Stern が 2012 年に DfID の委託で書いたペーパーが知られている (2)。 <p>2.2. 開発 NGO の動き</p> <ul style="list-style-type: none"> • 英国の開発 NGO のアンブレラ組織 BOND や、米国の InterAction の動きは注目に値する。BOND には、Effectiveness チームという NGO の事業の有効性や妥当性を考えるチームがあり (3)、InterAction には Evaluation and Program Effectiveness Working Group (EPEWG) という 	

	<p>ワーキンググループがある（４）。</p> <ul style="list-style-type: none"> • 評価を事業の改善のために活用とするこれらの試みのなかから、「現場の声」を聞き、フィードバックを事業に反映させるという機運が高まっている。 • VfM (Value for Money)についても、BOND は、Dfid などからの圧力の強いことを念頭に、英国 NGO としてどう振る舞うべきかの指針を出している（５）。基本的なスタンスとして、NGO は、政府から一方的に VfM の考え方を押し付けられるのではなく、その概念を正しく運用し、事業ごとの VfM を比較することで事業の改善につなげ、さらに評価をうまく活用することで VfM の推進者として自らを位置づけられるようにしなければいけないと説いている。 <p>2.3. 変化の理論(Theory of Change)</p> <ul style="list-style-type: none"> • Dfid が変化の理論に注目しているのは、従来のログフレーム的なアプローチが環境の変化や不確定性にうまく対応できていないことを見たものといえる。 • ログフレームも変化の理論の一種だと見ることができるが、変化の理論を活用ことでより柔軟に環境の変化に対する対応を構想することができる。 • ガバナンス、能力強化、組織の変化、政策形成など、従来のログフレーム的なアプローチが成立しにくい分野でのプログラムが増えてきたこともこの流れと関係している • こういった環境が変化するなかで事業を実行しているときのアウトカムについて考える手法のひとつとして、Outcome Harvesting が編み出されている。Outcome Harvesting は、あらかじめ特定された目的をもとに事業のアウトカム達成度を見るのではなく、「起きたこと」を綿密に拾い上げていくことで、その「起きたこと」と事業の関連を見ていこうとするもの（６）。 <p>2.4. 変化は遅い</p> <ul style="list-style-type: none"> • 評価の専門家集団が存在し、ひとつの産業として成立しているせいか、時代の流れに比べて変化は遅い。 • 例えば、企業やベンチャー・フィランソロピー、投資家の間で関心が高まっているソーシャルインパクト関連の動きに対しても、評価学会等の反応はあまり勢いが無い。これは、評価の世界がおもに政府系の資金で成り立っていることに起因するのではないか。 • 同様に、緊急人道支援の分野では ALNAP が中心になって評価の仕組みを確立しており（７）、その分野と開発を中心とした評価の世界の交わりもあまりないようだ。
--	---

	<p>2.5. EvalPartners</p> <ul style="list-style-type: none"> • そんななかで、EvalPartners が中心になって途上国の CSO の評価能力を高めようと尽力しているのは注目に値する。なかでも、2015 年を国際評価年と位置づけて様々なイベントが企画されている。
<p>参考文書 ／リンク</p>	<p>(1) http://www.outcomemapping.ca</p> <p>(2) “Broadening the range of designs and methods for impact evaluations,” http://r4d.dfid.gov.uk/Output/189575/</p> <p>(3) http://www.bond.org.uk/effectiveness</p> <p>(4) http://www.interaction.org/work/monitoring-evaluation</p> <p>(5) http://www.bond.org.uk/effectiveness/value-for-money</p> <p>(6) Outcome Harvesting については、以下など。 http://www.outcomemapping.ca/resource/outcome-harvesting#inline_en http://betterevaluation.org/resource/overview/OutcomeHarvesting http://aea365.org/blog/ricardo-wilson-grau-on-outcome-harvesting/</p> <p>(7) http://www.alnap.org</p>

ヒアリング 4

日時	2015年2月27日（金）	
場所	Keystone Accountability 222 Kensal Rd, London W10 5BN 英国、ロンドン	
機関・団体	Keystone Accountability	
面談相手	David Bonbright Chief Executive	davidb@keystoneaccountability.org
面談のポイント	<p>1) 本人について</p> <ul style="list-style-type: none"> • Bonbright氏は、Keystone Accountabilityの代表として、変化の理論、事業受益者フィードバックなどで概念の普及、ツールの開発を行っており、財団や国際機関に一定の影響力をもっている。筆者とは旧知の仲。 <p>2) 評価の潮流について</p> <p>2.1. マネジメント・ツールとしての評価</p> <ul style="list-style-type: none"> • 評価は役に立たなければ意味がない。その点で、注目すべきはCSOの世界にパフォーマンス主義、成果主義を浸透させようとしている一派だ。 • 例えば、Performance Imperative（1）。CSOのハイパフォーマンスのための「7つの条件」を提示しているが、そのなかには「学びの文化」、「内部のモニタリングと改善」、「外部者による評価」が含まれている。 • 例えば、Giving Evidence（2）。著者のCaroline Fiennesは、通常、評価は「やりすぎ」で要らない情報がたくさん出ていると指摘する。 • マネジメント・ツールとしての評価を浸透させるには、もっと内部評価をやらなければいけない。現状は、やらされている外部評価が多い。 • むしろ必要なのは、KPIをしっかりと作り、リアルタイムでフィードバックを得て、それを即学びや改善に結びつけることだ。ICT革新で、リアルタイムのフィードバックは可能になっている。 • こういったことを推奨するために、Feedback Labを立ち上げた（3）。 	
参考文書／リンク	<p>(1) http://leapofreason.org/performance-imperative/about-pi/</p> <p>(2) http://www.tinyspark.org/podcasts/poor-data-caroline-fiennes/</p> <p>(3) http://feedbacklabs.org</p>	

ヒアリング 5

日時	2015年3月3日(火)	
場所	UNEP/UNDP 220 East 42nd Street, Room 2046 New York, NY 10017 米国、ニューヨーク	
機関・団体	Independent Evaluation Office, UNDP (United Nations Development Programme)	
面談相手	Fumika Ouchi Evaluation Advisor	fumika.ouchi@undp.org
面談のポイント	<p>1) 本人について</p> <ul style="list-style-type: none"> 1998年から国連機関等で評価の仕事に携わっており、世界銀行研究所、国連麻薬犯罪事務所 (UN Office on Drug and Crimes)、国際原子力機関 (IAEA)、現在の国連開発計画 (UNDP) 独立評価事務所 (Independent Evaluation Office, IEO) と歴任している。 UNDP では国別、テーマ別、地域/グローバル単位で事業評価を行っているが、IEO は、スタッフ 20 人 (評価専門家 10 人) の小さな事務所。UNDP の指示系統に組み込まれるのではなく、直接理事会に報告する機構になっている。 <p>2) UNDP の評価について</p> <ul style="list-style-type: none"> UNDP では国別プランを 4-5 年のサイクルで作成するが、その評価に関して National Reference Group (NRG) をつくり、政府、ドナー、受益者、CSO などに入ってもらってプランの妥当性等を審議してもらう。NRG を評価の一連のプロセスに巻き込むことによって、サイクルの最初から最後までステークホルダーとの協議の場が保たれるようになった。 UNDP では国ごとおよび地域 (Regional Bureau) ごとに M&E (モニタリングと評価) の担当官をおいているが、専門性にバラツキがあり、全体の底上げが課題となっている。 <p>3) 評価の潮流について</p> <p>3.1. インパクト評価について</p> <ul style="list-style-type: none"> ドナー (DfID=英国、SIDA=スウェーデンなど) からのプレッシャーは強い。ただし、UNDP のプログラムの性格として、i) 単体ではなく他の国連機関や他の機関と共同事業をやっている場合が多く、UNDP の貢献度を抽出するのが難しい、ii) ガバナンス関連 	

	<p>のプログラムが多いので、インパクト評価がしやすい種類の事業ではない、などの理由であまり普及していない。</p> <ul style="list-style-type: none"> • ただし、GEF (Global Environment Facility) においては RCT (ランダム化比較試験) を活用するようになってきている。 <p>3.2. 変化の理論 (Theory of Change) の活用が奨励されるようになってきている。</p> <ul style="list-style-type: none"> • 従来のログフレームと違い、事象が単線系 (linear) でなく複雑系 (complex) であることにより対応しているから。また、事象に関わるステークホルダーの存在や関わりの程度を明示化することで、どのステークホルダーといついかに協議すべきなのかを示唆してくれる。 • UNDP において NRG の手法が主流化していったのが 2012 年頃。関連して、世銀や国際機関は参加型評価を多用するようになってきている。 <p>3.3. 評価に関する関心が全般に高まっている。</p> <ul style="list-style-type: none"> • 評価者以外が評価の重要性に気付き、視点や手法をプログラム運営に取り入れるようになってきており、もはや評価は評価者の専管事項ではなくなりつつある。 • その意味でも、評価のトレーニングの需要は高まっており、その重要性は増している。国連機関では、IPDET のトレーニングをよく活用している (1)。
<p>参考文書／ リンク</p>	<p>(1) http://www.ipdet.org</p>

ヒアリング 6

日時	2015 年 3 月 3 日 (火)	
場所	UNEP/UNDP 220 East 42nd Street, Room 2036 New York, NY 10017 米国、ニューヨーク	
機関・団体	UNEG (United Nations Evaluation Group)	
面談相手	Jin Zhang UNEG Programme Specialist	jin.zhang@undp.org
面談のポイント	1) UNEG について <ul style="list-style-type: none"> 国連機関における評価を横断的に扱い、情報交換、規準の設定、評価能力の底上げなどを担う機関として、1984 年に前身の 'Inter-Agency Working Group on Evaluation が設立。2003 年に UNEG と改名。 現在の活動の指針は 2005 年に作成された「規範(Norms)」および「規準(Standards)」文書に規定されている (1)。 昨年 (2014 年)、重点分野として人権とジェンダーの平等に配慮した評価に関するガイダンス文書を作成した (2)。 2) 国際評価年について <ul style="list-style-type: none"> 2013 年 9-10 月、ブラジルで開催された第 3 回国家評価能力国際会議において、2015 年を国際評価年とすることが決定した (3)。これを受けて、2014 年 12 月には、国単位での評価能力構築を推進する国連総会決議が承認されている (4)。 	
参考文書 /リンク	(1) http://www.uneval.org/document/detail/21# http://www.uneval.org/document/detail/22 (2) http://www.uneval.org/document/guidance-documents (3) http://www.nec2013.org (4) http://www.unevaluation.org/mediacenter/newscenter/newsdetail/105	

ヒアリング 7

日時	2015年3月5日(木)	
場所	スカイプ電話(雪のためワシントンDCの諸機関は閉鎖)	
機関・団体	En Compass, LLC 11426 Rockville Pike, Suite 229, Rockville, MD 20852 米国メリーランド州	
面談相手	Tessie Catsambas President	tcatsambas@encompassworld.com
面談のポイント	<p>1) 本人について</p> <ul style="list-style-type: none"> 個人コンサル会社 En Compass LLC の代表。評価分野で 25 年の経験。特に質的評価や AI(Appreciative Inquiry)の専門家。米国評価学会 (AEA) に面談を打診して紹介された。 下記の EvalPartners に関して、2012-14 の 3 年間、AEA の代表として、IOCE に参画し、EvalPartners の設立に関わった。 本人の略歴は (1)。 <p>2) 評価の潮流について</p> <p>2.1. 評価のプロ化の確立と広がり</p> <ul style="list-style-type: none"> 評価がプロの仕事として確立したのはここ 20 年程のことだが、「科学的」とされる手法と組織論・経営論との関連で成長してきた。 公的資金投入における Value for Money (VfM、価値の最大化) の流れが強くなり、評価の要請も強化され、その結果評価に多くの資金がつき込まれるようになった。 国際開発の分野において、USAID (米国国際開発庁)、DfID (英国国際開発庁) などの援助機関が VfM を打ち出し、それに呼応するように大手の開発 NGO が評価を強化し、それが途上国の NGO にも伝播する形でトリクルダウンが起こり、評価の要請が広がって行った。 同時に、途上国国内での評価の能力強化の必要性が言われるようになり、アジア、アフリカ、ラテンアメリカなどの地域ごとの評価専門家機関 (VOPEs = Voluntary Organizations of Professional Evaluators)を通じて評価のプロの養成が行われるようになっていった。 <p>2.2. EvalPartners と国際評価年</p> <ul style="list-style-type: none"> 上記の流れに乗って、CSO の評価能力強化の必要性が言われるようになり、UNICEF (国連児童基金) と IOCE (International Organization for Cooperation in Evaluation) が共同で、2012 年に EvalPartners (国際評価パートナーシップ・イニシアチブ) を設立した (2)。EvalPartners の目的は、世界各国 	

	<p>において CSO の評価能力と政策・世論形成能力の強化を図り、その結果として国レベルの政策が客観的な事実に基づいて形成されるようになること。EvalPartners では特に公平性 (equity) とジェンダーの平等、社会正義の実現に力を入れると、設立の際に採択されたチェンマイ宣言に記されている (3)。</p> <ul style="list-style-type: none"> • EvalPartners が母体となって、国連ミレニアム開発目標 (MDGs) から持続可能な開発目標 (SDGs) へと交代する年であり、新たな開発目標へのモニタリングの機運の高まりが見られる 2015 年を国際評価年と定めた (4)。 • このような流れのなかで、Catsambas 氏も共編者となって、評価のための環境整備を加速させるためのツールキットを作成するなど、CSO による評価を充実させていこうという動きがグローバルなレベルで結実しつつある (5)。 <p>2.3. 変化の理論や評価可能性のアセスメントなど、事業の入口で評価の視点を入れた介入を行うことの有効性に関する認識が高まっている。</p> <p>2.4. RCT (ランダム化比較試験) など、評価の「科学」的運用を比較優位と捉える潮流は強く、そういった手法に資金も集まりやすい。</p>
<p>参考文献／リンク</p>	<p>(1) http://www.encompassworld.com/blog/authors/tessie-tzavaras-catsambas</p> <p>(2) http://mymande.org/evalpartners/the-international-evaluation-partnership-initiatives</p> <p>(3) http://www.ioce.net/en/PDFs/Chiang%20Mai%20Declaration%202012.pdf</p> <p>(4) http://mymande.org/evalyear/Declaring_2015_as_the_International_Year_of_Evaluation</p> <p>(5) http://www.mymande.org/sites/default/files/toolkit/UNICEF%20NY_Advocating%20for%20Evaluation_Web_1.pdf</p>

ヒアリング 8

日時	2015年3月6日(金)	
場所	The Evaluators' Institute The George Washington University 2036 H Street NW (Samson Hall 105) Washington, DC 20052 米国ワシントン DC	
機関・団体	The Evaluators' Institute The George Washington Univ.	
面談相手	Ann Doucette Director	doucette@email.gwu.edu
面談のポイント	<p>1) 本人および TEI(The Evaluators' Institute)について</p> <ul style="list-style-type: none"> Ann Doucette 氏は、TEI のディレクターで、評価と保健政策の専門家。医療、学校関係の評価事業でのアウトカム測定の実験が豊富 (1)。 TEI は ジョージ・ワシントン大学にあり、1996 年に始動。現在数種類の認定評価士コースを開講している。姉妹機関として、他機関と連携して評価事業を行う The Midge Smith Center for Evaluation Effectiveness がある (2)。 <p>2) 評価の潮流について</p> <p>2.1. ソーシャルメディア等を通じたデータ取得</p> <ul style="list-style-type: none"> 従来、評価における悩みの種の一つは、いかにデータを取得するかだったが、ICT 革新によって、データは容易に手に入るようになってきている。これはまさにデータ革命の名にふさわしい。 ただし、統計的手法もないままやみくもにデータを集める向きもあり、信頼できるデータを取るには専門性が必要だということをよく知らなければいけない。 <p>2.2. 量的方法重視と質的方法重視の意見の相違</p> <ul style="list-style-type: none"> 評価の世界ではこれが一番の対立点だが (イデオロギー対立と言ってもよい)、今も解消していない。 量的方法では、やはり RCT (ランダム化比較試験) やそれに準ずる実験的手法への信仰は強い。私見では、2000-08 年のブッシュ政権時代に公立学校が 	

	<p>RCT で評価を行うことを義務化し、その流れが今も続いていると考えている。</p> <ul style="list-style-type: none"> 質的方法では、MSC(Most Significant Change)、AI(Appreciative Inquiry)、CAI (Critical Appreciate Inquiry、AIが「何が起きたか」に注目するのに対し、CAIは「何が起きなかったか」に注目する) など、手法の改善が見られている 評価の専門家が増えるにつれ、量的、質的を問わず手法を増え、手法過多とも呼べる現象が起こっているのではないか。 <p>2.3. 展開評価 (Developmental Evaluation)</p> <ul style="list-style-type: none"> 注目すべき動きに、展開評価がある。これは、パットンが「実用重視」の流れで開発した評価の種類で、特に複雑な事象において目標設定や介入の仕方が単純には決められないような事業の評価に適している (3)。 <p>2.4. ソーシャルインパクト評価の流れへの注目</p> <p>昨年ぐらいより、企業の社会活動評価の流れ (インパクト投資等) に評価の世界が注目するようになってきている。昨年の米国評価学会 (AEA) の大会では、ソーシャルインパクトや SROI に関するセッションがいくつか見られたが、これは新しいことだ。</p> <p>2.5. その他の新しい動き</p> <ul style="list-style-type: none"> 政治学の分野で活用されていた process tracing が評価の世界で注目され、評価専門家が使うようになってきている。事業の結果から政策決定の有効性をトレースする手法。 事業評価に関しては、アウトカムやインパクトをいかに測るかに多くが関心を払ってきたが、近年、知識の生成、共有、管理といった評価の活用が見直されてきていると思う。
<p>参 考 文 書 / リ ン ク</p>	<p>(1) https://tei.gwu.edu/about-tei</p> <p>(2) https://tei.gwu.edu/about-mscee</p> <p>(3) Michael Quinn Patton, 2011, Developmental Evaluation, New York: The Guilford Press (http://www.guilford.com/books/Developmental-Evaluation/Michael-Quinn-Patton/9781606238721)</p> <p>Developmental Evaluation については、ほかに以下参照 http://betterevaluation.org/plan/approach/developmental_evaluation https://www.scribd.com/doc/8233067/Michael-Quinn-Patton-Developmental-Evaluation-2006</p>

<p>http://www.fsg.org/tabid/191/ArticleId/708/Default.aspx?srpush=true http://tamarackcommunity.ca/downloads/vc/Developmental_Evaluation_Primer.pdf http://mccconnellfoundation.ca/assets/Media%20Library/Publications/DE%20201%20EN.pdf</p>
--

ヒアリング 9

日時	2015年3月6日(金)	
場所	スカイプ電話	
機関・団体	FSG	
面談相手	<p>Hallie Preskill, Managing Director (在シアトル)</p> <p>Jennifer Splansky Juster, Director, Collective Impact Forum (在サンフランシスコ)</p>	<p>hallie.preskill@fsg.org</p> <p>jennifer.juster@collectiveimpactforum.org</p> <p>コンタクト先</p> <p>Dane Smith, Managing Director (Dane.Smith@fsg.org)</p> <p>アシスタント</p> <p>Rany Chin (Rany.Chin@fsg.org)</p>
面談のポイント	<p>1) FSG と Aspen Institute による Collective Impact Forum (CIF)について</p> <ul style="list-style-type: none"> • FSG では、企業の戦略づくりで培った知見と、システムの変化がいかに起こるかの研究などを題材に、CIF 構想を練った。大きな規模でシステムの変化を起こそうと思ったら、組織単体ではなく、より大きな単位でビジョンの共有、測定指標の共有などが起こらないといけないと考えたため。 • 従って、FSG には、Preskill 氏のような評価の専門家もいるが、CIF はむしろ FSG の戦略ワークから出てきたもの。 <p>2) CIF の実際の運用について</p> <ul style="list-style-type: none"> • 現在はまだ試用段階で、いくつかのケーススタディーを集めている。 • すでに、なにがシステムの変化を促すか、重要な要因は見えてきている。それには、責任あるリーダーシップ、適切なネットワーク、緊急意識の共有、適切な資源などが含まれる。 	
参考文書／リンク	<p>CIF に関しては、FSG のホームページに各種資料が存在する。</p> <p>https://www.collectiveimpactforum.org</p>	

ヒアリング 10

日時	2015年3月6日(金)	
場所	Aspen Planning & Evaluation Program 1333 New Hampshire Avenue, NW Washington DC 20036 米国ワシントンDC	
機関・団体	Aspen Planning & Evaluation The Aspen Institute	
面談相手	David Devlin-Foltz Executive Director	david.devlin-foltz@aspeninst.org
面談のポイント	<p>1) 本人について</p> <ul style="list-style-type: none"> Aspen Institute で長く評価、特にアドボカシー評価の分野で活動しており、開発した Advocacy Progress Planner は知られている(1)。 <p>2) 評価の潮流について</p> <p>2.1. ビッグデータの影響</p> <ul style="list-style-type: none"> ICT 革新で、データを時間が経ってではなく、リアルタイムで取れるようになったことは大きい。 これを評価に関連する動きに結びつけようとしているもので注目すべきものに Feedback Lab (2) や Accountability Lab (3) がある。ともに Lab という命名からわかるように「実験室」的試みで、前者は、事業の受益者のフィードバックをリアルタイムで事業実施者等に還元する仕組みを模索しており、後者は、同様の試み(受益者に限らず市民一般からのフィードバック)を説明責任のテーマでしている。 <p>2.2. 展開評価 (Developmental Evaluation)</p> <ul style="list-style-type: none"> 事象が複雑なのにログフレームなどの単純な因果関係の評価図式が主流という評価の現状に一擲を投じた功績は大きい。 <p>2.3. 企業のソーシャルインパクト</p> <ul style="list-style-type: none"> インパクト投資等の流れは、いかに測定を効果的にやるかという点において評 	

	<p>価の知見に注目し、少なくともここ数年活用している。</p> <p>3) その他</p> <ul style="list-style-type: none"> • Collective Impact Forum について。組織単体ではなく、地域とかを面で見えてそこに共通のビジョンや評価手法を適用しようとする発想は、Aspen のなかでおこなれている Voices from the Field の一連の仕事の功績が大きい(4)。 • 2005年のパリ宣言以降、カントリー・オーナーシップが言われ、昨今は「カントリー」は政府だけでなく CSO や市民が含まれることが当然視されている。その流れで、評価についてもローカル・オーナーシップが言われており、それ自体は好ましい(5)。しかし一方、評価への参加には時間と手間がかかり、「現場」の声としては、「生活していくのがやっとなんだから、参加はいいからどうすればうまくいくのか教えてくれ」といった声も聞かれているようだ。
<p>参 考 文 書 / リ ン ク</p>	<p>(1) http://planning.continuousprogress.org</p> <p>(2) http://feedbacklabs.org</p> <p>(3) http://accountabilitylab.org/learn-AboutUs.html</p> <p>(4) http://www.aspeninstitute.org/publications/voices-field-iii-lessons-challenges-two-decades-community-change-efforts</p> <p>(5) 例えば、InterAction は、今年(2015年)に入って、Local Ownership in Evaluation を発行している (http://www.interaction.org/sites/default/files/Local%20Ownership%20in%20Evaluation%20-%20FINAL.pdf)</p>

ヒアリング 1 1

日時	2015 年 3 月 6 日 (金)	
場所	Center for Evaluation Innovation 1625 K Street, NW, Suite 1050 Washington, DC 20006 米国ワシントンDC	
機関・団体	Center for Evaluation Innovation	
面談相手	Julia Coffman Director	jcoffman@evaluationinnovation.org
面談のポイント	<p>1) 本人について</p> <ul style="list-style-type: none"> アドボカシー評価、政策評価、システム評価などの分野の専門家。特に米国民間財団とのつながりが強く、先進的な財団と共同で評価の革新について発信している (1)。 <p>2) 評価の潮流について</p> <p>2.1. 戦略的フィランソロピーの反動</p> <ul style="list-style-type: none"> 2000 年代後半ぐらいから、米国の大手財団は戦略的フィランソロピーのかけ声のもと、アウトカムやインパクトをそれまでより前面に打ち出した選択的な助成活動を行うようになった。しかし、社会は「機械」のようには動かず、いろいろな所で軋轢のもとになった。 その反動が今起こっている。複雑な社会に見合ったプログラム構築や評価のためには、違った手法が必要ということで注目されているのが展開評価 (Developmental Evaluation) の流れ。 <p>2.2. 評価の「王道」と対抗軸の関係</p> <ul style="list-style-type: none"> 評価の「王道」は依然として RCT (ランダム化比較試験) に代表されるようなインパクト評価。VfM(Value for Money) の目からは、政策決定者の構想通りの結果が効率良く出たかで事業の善し悪しが判断される。 一方で、より開かれた視点で、起きている事象から学び、それを体系化しようとする試みは評価の主流にはならないが、確かな一派を形成している。 	

	<p>3) その他</p> <p>CEI では、財団の評価に関する情報交換のための「評価ラウンドテーブル」を定期的 的に開催しており (2)、その関連で、財団の評価の動向に関して出版物を出して いる (3)。</p>
参 考 リ ン ク	<p>(1) http://www.evaluationinnovation.org/about-us/staff</p> <p>(2) http://www.evaluationroundtable.org</p> <p>(3) 例えば、 http://www.evaluationinnovation.org/sites/default/files/Benchmarking%20Evaluation%20in%20Foundations.pdf</p>