

〔特集論文〕

評価の国際的潮流と市民社会組織の役割

今田 克司¹⁾・田中 弥生²⁾

¹⁾CSO ネットワーク・²⁾大学改革支援・学位授与機構

International Trends in Evaluation and the Role of Civil Society Organizations

Katsuji IMATA¹⁾ and Yayoi TANAKA²⁾

¹⁾CSO Network Japan

²⁾National Institution for Academic Degrees and Quality Enhancement of Higher Education

The authors are working on promoting a culture of evaluation among nonprofits in Japan with the message that evaluation can be a powerful tool for civil society organizations (CSOs) to increase learning and effectiveness in solving social issues. From this perspective, there are two global trends that are currently worth exploring. One is the five-year Global Evaluation Agenda, which was adopted in relation to the Sustainable Development Goals (SDGs), and in which there is a call to action to strengthen evaluation capacities at the country level including Voluntary Organization of Professional Evaluators (VOPEs) and CSOs. The other is the so-called “the fifth wave” of evaluation, which stems from a diversification of evaluation stakeholders that includes social investors as well as dynamic and emergent operating contexts where change is the only constant and evaluation needs to be value-driven. With these trends in mind, the authors highlight the growing attention to “developmental evaluation” in the evaluation community as a useful approach to monitor and support social innovation in complex environments.

Key words: evaluation, civil society organizations (CSOs), sustainable development goals (SDGs), complexity theory, developmental evaluation

1. 本論の目的

本論は、開発分野を中心とした国際社会の評価の動向について紹介した上で、昨今の評価論やアプローチの潮流について概説し、市民社会組織（Civil Society Organizations 以下、CSO）の役割と課題について考察することを目的としている。

2015年は、国際的な開発目標であるミレニアム開発目標（Millennium Development Goals, MDGs）の最終年であり、同時にあらたな開発目標として持続可能な開発目標（Sustainable Development Goals, SDGs）を検討するという転換の年であった¹⁾。これを機に、国際機関は2015年を国際評価年と位置づけ、途上国を中心とした国レベルの評価能力を強化することを主眼とした5カ年計画を策定した。そこでは、評価専門家とともにCSOの役割が重視されている。

評価専門家の中には、SDGsで示された評価目標や社会変化の動向などに鑑み、既存の評価手法の限界を指摘する者もおり、新たな評価アプローチを構築する試みが行われている。こうした変化のなかで、評価においてCSOに求められる姿勢や役割について考察することが必要になっている。

2. 国際評価年と EvalAgenda2020

2.1. 新たな国際開発目標と国際評価年

2015年11月下旬、ネパールで開催されたグローバル評価週間において、2014年9月から一年以上にわたり行われていた世界規模での意見聴取（コンサルテーション）を終結させ、2016-2020年の5年間のグローバル評価アジェンダが採択された²⁾。

そこでは、(1)各国において質の高い政策評価が行

¹⁾ SDGsに関しては、<http://www.un.org/sustainabledevelopment/sustainable-development-goals/> (2015/12/31) 参照。

²⁾ Global Evaluation Agenda, http://www.mymande.org/evalyear/global_evaluation_agenda_2016-2020 (2015/12/31)。

われるための環境整備、(2) 評価専門機関や市民社会の組織基盤強化、(3) 評価者の能力向上、(4) 評価機関や評価者の能力向上のためのネットワークの構築が掲げられている。これらより、途上国政府の政策立案能力を向上するためには、評価能力の向上が不可欠であること、そのためには政府のみならずCSOなどの能力強化が必要であると考えられていることがわかる。

とはいえ、CSOを巻き込んだ評価力向上の必要性については以前より指摘され、既に取り組みが進められていた。UNICEF(国連児童基金)とIOCE(評価協力国際組織)³は、2012年に国際評価パートナーシップ・イニシアチブ(EvalPartners)を設立している。これは、各国のCSOの評価能力と政策・世論形成能力の強化を図ることを目的に設立されたもので、最終的な目的はCSOを通じて国家レベルの政策がエビデンスに基づき、公平性や政策効果に重きを置いて形成されるようになることである⁴。特に途上国内のCSOの評価の能力強化を図るため、世界各国の評価専門機関のネットワーク⁵を通じて、評価者の養成を実施していくことが構想されている。

また、EvalPartnersが推進母体となって、2015年が国際評価年と定められたが、これはMDGsからSDGsへと交代するタイミングを見据えたもので、その数年前から議論が行われていた。すなわち、2013年9-10月に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ける決議⁶、さらに2014年12月の国単位での評価能力構築を推進する国連総会決議があり、そこで国際評価年についての言及がなされている⁷。そして、新たな開発目標が定められたタイミングで国際評価年を定めたことは、国際開発の分野において評価の重要性を意識する関係者が多いことと無縁ではなかった。

2.2. 国際開発目標の構造と評価

このように評価力の構築が注目される背景には、開発分野におけるアカウンタビリティの意識の強さ

に加え、MDGsやSDGsの構造と関係があると考えられる。

MDGsやSDGsは、複数の開発目標から構成されるが、さらにその下位概念としてターゲット、さらに下位概念に指標を設けており、3層構造を有している。

SDGsを例にとって具体的に説明しよう。2015年9月の国連持続可能な開発サミットにおいて採択されたSDGsの成果文書、「我々の世界を変革する：持続可能な開発のための2030アジェンダ」(外務省仮訳)(国際連合2015)を見てみたい。

開発目標は、17分野(貧困、飢餓、健康と福祉、教育、ジェンダー、水と衛生、エネルギー、雇用と経済成長、産業・技術革新・社会基盤、格差、都市とコミュニティ、消費と生産、気候変動、海洋資源、陸上資源、平和・正義・制度、パートナーシップ)から構成されている。さらに、各分野ごとに、ターゲット、指標が設定されている(図1)。

例えば、第1の目標である貧困は次のように構成されている。

目標1:「あらゆる場所のあらゆる形態の貧困を終わらせる。」

ターゲット:目標1について7つのターゲットが設定されている。例えば、ターゲット1.1は、「2030年までに、現在1日1.25ドル未満で生活する人々と定義されている極度の貧困をあらゆる場所で終わらせる」と示されている。

指標:ターゲットに対応してひとつまたは複数の指標が設けられる。例えば、ターゲット1.1に関する指標は、「性別、年齢、雇用状態で細分化された国際貧困ライン以下の人口の割合」が示されている⁸。

このように、MDGsやSDGsとして掲げられた目標は、比較的抽象的な表現になっているが、その下位に位置するターゲットには複数のより具体的な目標が設定されており、さらにその下位には指標が設定されていることから、目標をより具体的かつ計測可能なもので測定していこうという意図があることがわかる。

一般に、このような国際目標は汎用性が求められるために抽象度が高くなる傾向にあるが、目標の設定が曖昧であると達成状況の検証が困難になる。しかし、MDGsやSDGsにおいては、目標に加えてターゲットと指標を設けることによって、目的達成状況の測定や

³ 英語名は International Organization for Cooperation in Evaluation.

⁴ The International Evaluation Partnership Initiative, <http://mymande.org/evalpartners/the-international-evaluation-partnership-initiatives> (2015/12/31).

⁵ これを通称、VOPEs (Voluntary Organizations of Professional Evaluators) と称する。 <http://aea365.org/blog/susan-kistler-on-vopes-and-the-evalpartners-innovation-challenge/> (2015/12/31) 等参照。

⁶ Third International Conference on National Evaluation Capacities 2013, <http://www.nec2013.org> (2015/12/31).

⁷ News Center: Adoption of a New United Nations General Assembly Resolution Highlighting Evaluation Capacity Building for the Achievement of Development Results at Country Level, <http://www.unevaluation.org/mediacenter/newscenter/newsdetail/105> (2015/12/31).

⁸ 本稿執筆段階において指標は合意に至っておらず、ドラフト段階。2016年3月に合意に至る予定。2015年11月2日時点でのドラフト一覧は、<http://unstats.un.org/sdgs/files/meetings/iaeg-sdgs-meeting-02/Outcomes/Agenda%20Item%204%20-%20Review%20of%20proposed%20indicators%20-%202015.pdf> (2015/12/31)。

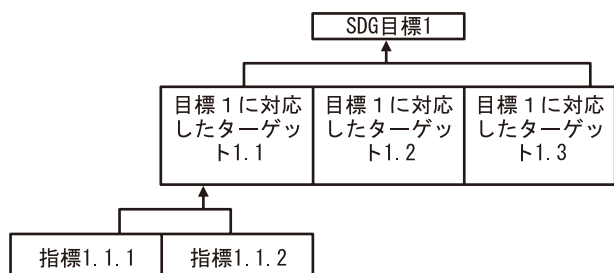


図1 SDGsの3層構造
出所：筆者作成

検証を可能にしている⁹。換言すれば、そこで設定された指標に照らして目標の達成状況を検証することが求められており、そこから評価を重視する姿勢が読み取れる。実際、MDGsに関しては2005年以降毎年MDGs達成の進捗状況に関する年次報告書が発表されており¹⁰、SDGsにおいても同様の検証作業が行われることが予測できる。また、すでにSDGs策定過程において、策定後のモニタリングや検証の重要性は複数のステークホルダーから指摘されており、進捗状況を的確に検証していく各方面からの圧力はMDGsのときよりも強くなっていくことは想像に難くない。

このような背景もあり、成果文書の「フォローアップとレビュー」の原則の一覧の部分（国際連合2015: 11-12および33、一部翻訳を改訂）に以下のような進捗測定や評価に関する言及が見られる。

48.（本件アジェンダを達成するための）指標は、こうした（フォローアップ）活動を支援するために整備される。誰一人も取り残さないよう進捗を測定するためには、質が高く、アクセスが容易で、時宜を得た細分化されたデータが必要である。このようなデータは、政策決定の鍵となる。現存する報告メカニズムからのデータと情報は、可能な限り活用されるべきである。アフリカ諸国、後開発途上国、内陸開発途上国、小島嶼開発途上国、中所得国をはじめとする開発途上国における、統計能力の強化のための努力を強化することに我々は合意する。我々は進捗を測定するために、GDP指標を補完する、より包括的な手法を開発することにコミットする。

74g. これら（フォローアップとレビュー）は、評価は各国主導で行われるべきであるが、その際、正

⁹ SDGs指標に関しては、<http://unstats.un.org/SDGs/>（2015/12/31）参照。

¹⁰ Millennium Development Goals Reports, <http://www.un.org/millenniumgoals/reports.shtml>（2015/12/31）。

確な根拠データをもって実施されることが求められる。そして、そのデータの要件としては、質が高いこと、アクセスが容易なこと、時宜を得たものであることに加え、収入、性別、年齢、人種、民族的属性、移住者の法律上の地位、障害、地理的属性、及びその他各々の国内での状況に応じた属性による細分化が可能なデータが生成されることがあげられる。

2.3. グローバル評価アジェンダ

一方、ネパールのグローバル評価週間で採択された2016–2020年のグローバル評価アジェンダは、EvalAgenda2020と呼ばれ、2020年までの成果目標として4つの主要な要素が掲げられている。それらは以下の通りである（EvalPartners 2015: 7–8）。

(1) 評価のための環境整備

- ・社会のあらゆるセクターが評価の価値を理解し重要視する。
- ・評価が国の政策や法体系のなかで義務づけられる（あるいは奨励される）。
- ・評価のために必要な資源配分がなされる。
- ・評価結果を保存しアクセスを確保するデータシステムが存在する。
- ・利害関係者が評価情報を活用したいと思う。
- ・評価従事者の専門家集団としての地位が向上する。

(2) 組織基盤

- ・評価や評価者を発展させ支援するVOPEsやCSOs、研究者、その他の機関が十分に存在し、必要とされるデータを生成し共有している。
- ・これらの諸機関が、質の高い評価とは何かを判別、推進することができる。
- ・これらの諸機関が、必要に応じて他の関連機関と連携することができる。
- ・これらの諸機関が、データ生成や評価の質を保証するのに必要な資源を充当し、評価へのアクセスを容易にし、評価の結論や提言を適切にフォローすることができる。
- ・これらの諸機関が、評価の分野の進展に応じて成長・発展することができる。
- ・学術機関は、評価のコースを適切に運営することができる。

(3) 個人の能力強化

- ・評価者、評価の依頼者、評価の使い手の評価能力が向上する。
- ・評価の依頼者や使い手が、評価の価値や質の高い

評価のやり方に関する理解を高め、評価の提言をより積極的に活用する。

- 特定の地域や分野において十分な数の評価者が存在し、質の高い評価を実践している。また、これらの評価者が総体として多様な専門分野の知識を保持している。
- これらの評価者が評価原則、理論、方法、アプローチに精通している。
- これらの評価者が評価の価値をよく理解し、文化の違いに敏感である。
- これらの評価者が常に学び、改善に努めている。

(4) 上記3側面の統合

- 政府、議員、VOPEs、国連、財団、企業などの諸機関が評価の実践、評価分野の革新、評価の能力強化のために共同で資源を投入する。
- 評価についての議論ができる共通言語や専門用語群が存在する。
- 評価に関わる複数のパートナーが定期的に国レベル、国際レベルの学習機会を活用する。
- SDGsの「誰一人取り残さない」原則が前述の評価システムの3つの要素に通底する基本価値として組み込まれる。¹¹

以上、国際評価年やEvalAgenda2020にかかる一連の動向をレビューしてきたわけだが、それらから次の点を読み取ることができる。すなわち、i) 評価に関する正しい理解が、評価関係者のコミュニティの外では未だ欠落しているという認識があること、ii) 評価の担い手の育成が急務であり、それには政府、国際機関などのマルチステークホルダーによる取り組みが肝要であること、さらに、iii) 途上国の評価能力の強化のためには、評価専門家だけでなく、CSOの役割が肝要であることである。

評価においてCSOが期待される背景には、開発事業の現場における、評価の使い手としてのCSOの役割がある。評価は、専門家が実施してその報告書が評価依頼者（多くの場合は資金提供者）に読まれるだけではよく活用されたとはいえない。事業の実施主体は多くの場合CSOであるから、CSOが評価を上手に活用して、事業実施から学びを得、事業改善へとつなげていかなければ評価の意味は半減してしまう。また、評価が活用できると気づいたCSOは、評価のプロセス、例えばデータの収集や評価者との対話に積極的に

参加するようになり、結果としてこれが評価の質を上げることに貢献していく。

3. 評価の潮流にみるステークホルダーの変化

3.1. 評価の潮流～「第5の波」の到来

前節では、開発分野と評価にかかる国際的動向において、CSOへの期待が高まっていることを概観した。では、評価論そのものに関しては、これまでいかなる変遷がみられているのだろうか。ここでは、Picciotto (2015)の「第5の波」論に沿って概説してみたい。

Picciottoは、現在、評価の「第5の波」が到来していると述べる。「第5の波」という表現は、Vedung (2010)の評価の4つの波の議論を参照し、これに新たな潮流を加えたところから来ている。ここでは、ベースとなったVedungの議論を説明した上で、「第5の波」について説明する。

まず、Vedungの評価の4つの波とは、「第1の波：評価の科学化」、「第2の波：評価の民主化」、「第3の波：評価における市場主義の導入」、「第4の波：評価の第2の科学化」である。これは一般的な評価論の潮流（田中 2005: 139-146）と合致するものであることからそれらもふまえて、以下に解説する。

「第1の波」評価の科学化（1965-75頃）：政策形成過程に科学的知見を兼ね備えた合理主義が導入された時代。

米国でケネディ大統領が「政策に科学を」と述べたことが契機になり、政策形成過程に科学的な視点を入れることの重要性が説かれ試みが行われた。特に、貧困層や有色人種への公的サービス提供が開始され、国家予算に占める社会保障費が急増した時代であり、投じた国家予算に対する効果の検証が求められた。そして政策効果検証を契機に政策評価方法の開発、施行が進められた。

「第2の波」評価の民主化（1975年頃より）：利害関係者による対話や討議を重視。

政策評価が導入された当初は、評価者は、評価対象を第三者として客観的に観察する主体として捉えられていた。この作業過程においては、評価者の独立性が強調され、評価対象となる人々とコミュニケーションは取られなかった。しかし、評価の目的が評価対象となる事業の改善や被評価者の資質の向上に資することであるのならば、評価目的や方法について評価の依頼者や評価対象者・団体との対話や議論が必要であるという考え方が徐々に普及していった。日本語版が2001

¹¹ 2015年11月下旬のネパール会議後に集約されている最終文書ドラフトより。本稿執筆段階において最終文書は仕上がっていないが、最終文書においても、ここに言及した4つの基本要素に変更はない予定である。

年に出版されたパットン (2001) の『実用重視の事業評価入門』の英語での初版は1978年に出版されているが、これはこうした考え方を表す代表的な著書である。

「第3の波」評価における市場主義の導入 (1980頃より) : NPM (ニュー・パブリック・マネジメント) の波に乗った効率化やRBM (成果重視マネジメント)。

1980年代、行政改革の基礎をなす新思想として、新自由主義が登場する。そして、行政改革の新たなアプローチとしてニュー・パブリック・マネジメント (以下NPM) が、各国政府によって導入されてゆく (大住 2003: 20-22)。NPMは、競争原理の導入などの方針で構成されるが、そのうちのひとつが業績、成果にかかる評価の導入である。また、その関連で、事業の結果や成果により焦点をあてた「成果重視マネジメント」が論じられ、成果検証の視点から評価の必要性が認識された。

「第4の波」評価の第2の科学化 (1995頃より) : エビデンス主義の導入と、RCT (ランダム化比較試験) などに代表される精緻なインパクト評価の潮流。

評価手法において科学的で精密な視点を導入し、評価結果の信頼性を向上しようとする動向。実施した政策や事業の純効果 (ネット・アウトカム) を明らかにし、政策や事業と効果の直接的因果関係を明らかにしようとした。そのためには、ランダム化比較試験などもともと医療分野で開発された方法論を採用した¹²。こうしたアプローチによる評価は一般にインパクト評価と言われる (Rossi et al. 1999: 235-275)。

以上のように、評価の大きな潮流としては、1980年代以降、成果重視や特定の介入による成果の抽出の流れがあるが、多くは業績測定と呼ばれる手法、すなわち、特定の成果指標をモニターし、その進捗を確認することに留め、事業と成果の因果関係については深く追求しないアプローチが採用されていた。しかし、特に国際開発の分野では、ランダム化比較試験によるインパクト評価が「ゴールド・スタンダード」と呼ばれ、数ある評価のアプローチのなかでも最上位に位置するものという考え方が普及するようになった。

これに関連して、公的資金投入におけるバリュー・フォー・マネー、すなわち投資に対する価値の最大化が言われるようになり、それは国際開発の分野にも波及して、欧米の援助機関を中心にこの基準で事業の選定をするようになっていった。このように投資効果を求めるようになると事業と効果の因果関係をより明確にする必要が出てくる。それが、インパクト評価が求められるようになったひとつの所以といえる。

これらの評価の潮流に、Picciottoは「第5の波」を加えた。「第5の波」においては、「ゴールド・スタンダード」はもはや存在せず、社会的イノベーション分野における評価は、「成果重視であるよりも価値中心となる」(Picciotto 2015: 8) と論じている。その背景には、上述したSDGsの影響がある。Picciottoは、「SDGsの到来もあり、社会的インパクトの測定は長期的、二次的、間接的で、予測していなかった結果を測定しようとするのが大事になる。そこでは、個々の介入に成果を帰属させようとする、大概は不毛に終わる試みを超えていかなければならない」(Picciotto 2015: 7) と述べている。

また、Picciottoは、社会的インパクト評価の波の到来により、評価に携わるアクターの多様化が起こっており、特に企業セクターやNGOの役割が大きくなっていると指摘する。確かに、昨今の状況は、倫理的投資やソーシャル・インパクト・ボンド (SIB) などの投資や債券市場への社会性評価導入などの新しい評価インストゥルメントの誕生が、評価にかかわるステークホルダーの関係図を書き換えている様相を示している。

3.2. 潮流変化とステークホルダーの多様化

また、評価専門家の役割に対する認識も変化している。Vedungの第1の波では、社会学者など評価に関連した技術や知識を有する者が評価を担うことが想定されていた。第2の波では、利害関係者による対話や討議を重視する方向へ関心が向いたが、これは、専門家中心からより多くの人々を評価に参加させ、理解を広げてゆこうというものだ。第3の波と第4の波においては、評価作業に求められる専門性が高まる一方で、評価者コミュニティの外の理解はさほど進まなかったと捉えることができる。この観点から Picciotto の議論を敷衍すれば、現在、「第5の波」においては、評価の対象が多様になり、事業価値により注目が集まり、評価結果の活用も投資目的などより幅広くなっている。そのような状況下、第3、第4の波のような専門家中心の評価から、より多様な関係者の理解を広げる動きが加速しつつあるといえる。

¹² ランダム化比較試験は、政策評価に科学的視点が導入された1960年代、すなわち「第1の波」でも用いられている。しかし、現実の社会生活においてコントロールグループを作ることが難しく、思うような結果が出なかったこと、またこうした手法にはかなりのコストがかかることから次第に用いられなくなっていった。

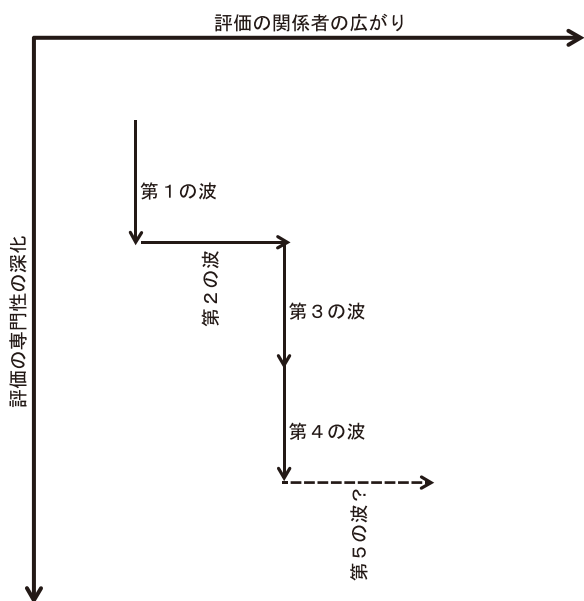


図2 Vedung と Picciotto による「評価の波」における
評価の専門性と広がり
出所：筆者作成

図2は、縦軸に評価の専門性の深化、横軸に評価関係者の広がりを据え、第1から第5の評価の波の動向を示したものである。専門性と関係者の広がりの間にはある種の背反関係があるのだが、評価の波は両者の間で行き来していることがわかる。

この状況認識は、筆者が昨年行った評価に関する海外調査でも浮き彫りになっていた点である（今田2014）。従来、評価に注目しているのは直接の関係者、すなわちドナー等の評価依頼者、事業実施主体と評価者自身だったが、現在では政策決定者、投資関係者、事業の利害関係者を含む間接的な関係者が評価の重要性に気づいている。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。

Picciotto が論じるように、社会的インパクト評価の波の到来により、評価に携わるアクターの多様化が起こっている。社会的投資を出発点とした動きには上述の倫理的投資やSIBに加えて、社会的リターン投資(SROI)、ESG投資¹³などの動きが含まれるが、評価と投資を直接連動させることによって、評価の活用が加速することが予測されている。また、こういった動きに連動して、成果の可視化や事業単位の成果の把握

を超えて地域や分野での複数の事業の成果やインパクトを測るコレクティブ・インパクトの動きが進みつつある¹⁴。さらに、評価の対象となる事業やイニシアチブのフィードバックを提供する層を広げることを企図するフィードバック・ラボなどの動きもあり¹⁵、多様なアクターが評価に関わるなかでCSOの役割にも注目度が増している。さまざまな新しい評価方法が誕生するなかで、評価のステークホルダー関係図が次々と書き換えられているのが現状といえよう。

4. 新たな評価アプローチへ～複雑な世界

4.1. 複雑系理論の評価への応用

評価理論の基本には、インプット→活動→アウトプット→アウトカム→インパクトというロジックモデルに代表されるように、単線系モデルがある。すなわち、物事には原因と結果があり、ある事象を生起させる原因をつきとめることができるとする論理実証主義の考え方である。Vedungの第1の波の時代は、社会科学でいう「科学化」の時代に対応している。すなわち「科学化」とは、自然科学の伝統にのっとり、記述し、説明し、将来の事象を予測できるようにすることである（Fay 1975）。この時代に評価の基礎づけがなされたことを考えると、その背景に、論理実証主義的な社会の捉え方があることは自然なことといえる。

しかし、もはやランダム化比較試験は「ゴールド・スタンダード」ではないとする考え方の背景には、伝統的な評価理論では社会現象を正確に捉えることはできないという主張が存在する。

ここで、評価の専門家たちが評価に応用できる理論として注目してきたのがいわゆる複雑系の理論およびシステム理論である。ただし、これらを評価に応用するのは容易ではない。Ramalingam et al. (2008: 8) は、システム理論も含めた複雑系理論の主要な概念として、10の要素を抽出しているが、そのうち、複雑系とシステムに関するもの、および複雑系と変化に関するものは以下の7つである。

複雑系とシステム

- ・複雑な繋がりや構成要素の相互依存性
- ・変化を生起させるフィードバックのメカニズム
- ・エマージェンスの考え方（要素の相互作用により

¹³ 例えば日本の年金積立金管理運用独立行政法人（GPIF）は2015年9月28日、資産運用においてESG（環境・社会・ガバナンス）の取り組みに優れた企業へ投資を行うESG投資の推進の一環として、同月16日に国連責任投資原則（PRI）に署名したことを発表した。http://sustainablejapan.jp/quickestg/2015/09/28/gpif-pri/18941（2016/1/5）。

¹⁴ 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。https://www.collectiveimpactforum.org（2016/1/5）。

¹⁵ Feedback Lab, http://feedbacklabs.org（2016/1/5）参照。

表1 複雑系の特徴と評価への応用

複雑系の特徴	評価への応用
複雑なシステムは常に変化の途上にあり、多くの場合、予測可能ではない。	評価のデザインおよび実践を、状況適合型で柔軟性に富み、反復可能なように設計せよ。
すべてはつながりあっている。部分はシステムの一部であり、全体に影響を与える。	構成要素やつながり具合を含め、システム全体を捉えるようにせよ。
情報は学びを促進する燃料であり、システムの繁栄を支える。	フィードバックの回路や情報へのアクセスを強化することでシステムの学習能力を高めるよう支援せよ。
背景・文脈が重要。事業の成否がそれにかかることも。	背景・文脈に注意を払い、それらの変化に柔軟に対応できるようにせよ。
個々の状況はユニークなものであり、よってベストプラクティスの適用範囲は限られている。むしろベストプリンシプル（最善の原則）を考えるべき。	既定の行為にしばられるよりも実践に役立つ効果的な原則を探すようにせよ。
場合によって、エネルギーや集約点の出所は異なる。	システム内でのエネルギーや影響が発生する場所や力の流れを観察せよ。
物事の関係性は物事自体と同等かそれより大きな重要性をもつ。	システム内の関係性、相互依存性に注目せよ。
原因と結果は単線系でも予測可能でもなく、一方向に動くものではない。より反復するものである。	非単線系で複数のベクトルをもつ関係性、予期された結果や予期されなかった結果を説明するようにせよ。
自由に動くことのできる半ば独立し多様な主体からパターンが生成される。	システムの異なるレベルに出現する、一回きりまたは再現するパターンを観察せよ。

出所：Preskill and Gopal (2014: 5)

事象が生起すること。多くの場合、予期せぬ事態の招来)

複雑系と変化

- ・非単線系の変化
- ・初期状態の微小な違いによるのちの大きな違い (例：バタフライ効果)
- ・フェーズ理論と段階的变化
- ・カオス現象 (見た目の無秩序の裏にある秩序)

これらのうち、構成要素の相互依存性、エマージェンス論、非単線系の変化などが特に評価の理論家に注目されてきた。そして、複雑系の理論を評価の体系に応用することを試みたものとして、Preskill and Gopal (2014) によるものがある。そこでは、複雑系の特徴とそれが評価にいかに応用されるかが概説されているが、その内容を表1にまとめた。

また、より実践家向けに複雑系の理論を概説し、現実の社会でいかに変化が起こるかを叙述した Westley et al. (2006: 9) は、わかりやすくいえば、複雑な問題の例は、「子どもを育てる」ことで、それは単純な問題である「ケーキを焼く」とことと対比できると述べている。説明を簡略化するためか、挙げられた事例がデフォルメされている感があるが、Westley et al. の主張を表2にまとめた。

Westley et al. がこの比喻に込めたメッセージは、社会の変化は「ケーキを焼く」ようには、あるいは「ロケットを月まで飛ばす」ようには起こらない、という

もので、現在でも評価におけるロジックモデルが、事象の変化を「ロケットを月まで飛ばす」とことと捉えている場合が多いことに警鐘を鳴らしている。

そこでロジックモデルに代わり、非単線系の世界の把握を評価に応用しようという試みが「変化の理論」¹⁶に見られる。「変化の理論」センターというウェブサイトもあり¹⁷、ロジックモデルも包含する変化のロジックをより幅広い観点から捉えるモデルとして使われている。

「変化の理論」については、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されている¹⁸。典型的なロジックモデルも「変化の理論」の一種といえ、そのロジックの展開や構造については様々なバリエーションが生まれているものの、その基本はロジックを構成する要素間を因果関係で結ぶ単線形の考え方である。他方で、「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ではなく、事業を直接構成する要因のみならず、間接的に関係する要素も視野に入れ、それをつなぎ、しかも逆進性も認めている。また、当初想定していた目標値に

¹⁶ Theory of Change の訳。

¹⁷ Center for Theory of Change, <http://www.theoryofchange.org> (2016/1/5)。

¹⁸ <http://www.theoryofchange.org/what-is-theory-of-change/> (2015/12/31) 参照。今田 (2014) で言及している。

表2 単純な問題と複雑な問題

	単純	複雑 A (complicated の訳)	複雑 B (complex の訳)
例	ケーキを焼く	ロケットを月まで飛ばす	子どもを育てる
操作手順	肝心なのはレシピ	厳密な手順と方程式が必須	厳密な手段は限られた応用性しかもたず、逆効果のこともある。
再起性	レシピに従うことによって同じケーキを焼くことができる	一度成功すれば、その手順に従うことにより、次の成功の可能性も高い	子どもを一人育てることによって経験値は上がるが、次も同じようにやってもうまくいくわけではない。

出所：Westley et al. (2006: 9) より抜粋

についても、モニタリングや観察を通じて、変更可能であり、場合によっては効果の発現に影響する主たる要因も変更することができる。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。

注意しなければならないのは、「変化の理論」は、ロジックモデルを否定するものではないということだ。事業によっては「ロケットを月まで飛ばす」のように、論理構築（あるいはそれよりも単純なもの）で事足れる場合もあり、その場合には単線系のロジックモデルが有用である。しかし、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の場合は、「変化の理論」を描いた方がより事象を正確に説明できると考えられる。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評価を必要とする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。

4.2. 発展型評価

発展型評価 (Developmental Evaluation)¹⁹ は Patton が 1990 年代なかばから提唱している評価の考えから発展させたもので、2011 年に一冊の書物にまとめられ、評価者コミュニティのなかで注目を浴びている。

発展型評価とは、単なる一手法ではなく、文献によってはアカウンタビリティの確保を主眼とする総括評価、事業改善を主眼とする形成評価と対比されることからわかるように、より高レベルの評価の枠組み、類型である。Patton (2011: viii) は、「発展型評価」について次のように説明している。

発展型評価とは、社会的イノベーションなど、目的が固定されているというよりも目的自体が変化し、時間軸も予め設定されているというよりも流動的で前進的な対象を評価するための方法である。そして、そこから得ようとするのは、外部への説明責任というよりも、イノベーションや変化から学習することにある。

この記述からもわかるように、発展型評価は、「変化の理論」等のツールも含め、複雑な事象と向き合いながら評価の作業をいかに柔軟かつ系統立って推し進めることができるか、その可能性を追求している。ただし、「変化の理論」がロジックモデルを包含するものであると同様に、「発展型評価」も従来の評価のやり方を否定するものではなく、評価の対象となる事象の種類によって、評価アプローチを使い分けるといった考え方に立っている。そこで、発展型評価を既存の評価アプローチとの比較において説明すべく、表3のように整理してみた。

「発展型評価」は、現実の事象に「複雑 B」の世界を反映するものが多いのにもかかわらず、評価の理論やツールにはそれに対応しているものが少なかったため、評価コミュニティのなかで大きな注目を浴びている。しかしながら、評価の体系としてはまだ発展途上で、2011年のPattonの著作のあともなにか「発展型評価」でなにかがそうでないのか、議論は続いている²⁰。

¹⁹ Developmental Evaluation に関しては、今田 (2014) では「展開評価」と訳出したが、本稿では「発展型評価」という訳語を使用している。この用語は、現段階で日本語への定訳が存在していない。ただし、発展型評価と訳しても展開評価と訳してもわかりにくさは同程度で、この点については筆者が 2015 年 11 月に開催された AEA (American Evaluation Association) の年次大会で Patton 氏と話し合ったところ、意味としては “adaptive evaluation” という用語も使えと言われた。これを使うとすれば「適合評価」となるが、本稿ではこれは採用していない。おおよその含意として、状況の変化に応じて事業を発展・展開・適合させていくのと同時に、評価を発展・展開・適合させていくというタイプの評価を指し示している。

²⁰ Patton (2011) より以前の段階で、発展型評価を系統立てて簡易に説明しようとする試みはあった。Gamble (2008) や Dozois et al. (2010) などを参照。

表3 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
すでに確立している 結果もある程度予測可能	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
まだ改善の余地あり 基準づくりの過程	事業の主要な要素はまだ形成途中 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
発展中 生成中	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.	発展型評価

出所: Preskill and Beer (2012: 6) より一部筆者加筆. 今田 (2014) より再掲.

評価は理論だけでなく実践例を伴って初めて理解が進むことから, Patton et al. (2016) は不確実性の高いプログラムにおける評価の12事例を集め, それらから浮かび上がってくる「発展型評価」に共通する要素を抽出して, 基本原則を見出そうとしている.

このように, 「発展型評価」をめぐる言説は, いまだ発展途上にあるが, より現実の事象や変化に即した評価手法を模索しようとする試みであることから, その基本理念や実践例も含め, 日本の評価コミュニティや非営利セクターに紹介が進むことが期待される.

5. 結論: 新たな「学び」のシステム構築に向けて

「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. 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. 特に公金が支出される事業に関しては, それが国内事業だとしても開発援助だとしても, バリュー・フォー・マネーのプレッシャーも強く, 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いえよう.

こうした社会的動向に鑑みれば, 評価技術のさらなる開発が必要であるが, 同時に, 評価論を発展させていくには, いかにより動的で常に変化する現実を的確に捉え, 評価の体系のなかに組み込むかが肝要となる. 本論で紹介した「複雑系」や「発展型評価」はその代表例である. 従来の評価のアプローチは, 過去に設定した目標の達成状況を検証することが主たるもので, いわば静的なアプローチであった. しかし, 先のような大きな環境変化においては, 目標設定の変更や, より多様なステークホルダーを想定して, 事業に

かかわる主要アクターの変更もありえることを想定した上での評価が求められている. それは, 評価対象を線形から非線形的に, 静的なものからより動的に捉えることであり, 自ずと新たな評価論・評価技術の開発が求められることになる. そうしたアプローチは, 一方で解析技術などより専門的な知見や技術, 他方でステークホルダーの適切な巻き込みや評価への参加が求められ, いずれにしても相当な時間, コストを要することを視野に入れておかねばならない.

しかし, CSOにとっては追い風といえるような現象も見出せる. 前述の評価の二本柱のもうひとつである「学び」の復権が起こっているからだ. 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は当たり前のことだが, Eval Agenda 2020で明確に述べられているように, CSOの評価の能力形成が急務とされており, 評価の実践を通じて同時に評価能力を形成すると捉えることができる. また, 特に「複雑B」の世界観に立てば, 評価とはすなわち学びの共有や生起しつつある事象への適合をいかに的確に行うかという作業を包含するものであり, ここでも学びとの接近が指摘できる.

今田 (2014) でも述べたが, 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Pattonだが, 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. その後, 30年を経て, 前述した「発展型評価」の流れにつながり, 複雑な世界における実用的評価のあり方として普及しつつある. Patton et al. (2016: 3) が明らかにした「発展型評価」の基本原則のひとつに, タイムリーなフィードバックというものがある. テクノロジーの進歩により, 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ることが技術的には可能になっている. 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留まらず、評価と改善が事業運営の不断のサイクルに組み込まれていく。

そうであるならば、発展型評価のアプローチをより効果的に活かすことができるのは外部評価ではなくむしろ内部評価である。より正確に言えば、外部者である評価者を事業実施のプロセスに巻き込んだ事業実施者主体の評価プロセスを構築・実施することであり、内部評価の新たな地平といえるだろう。実際、発展型評価の原則のひとつには、「共創」²¹の文字が見られる (Patton et al. 2016: 306–307)。そして、これは事業実施者の評価能力が問われることを意味する。評価の専門家は外部にいて、評価の使い手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 Patton が実用重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っていく。

もし原語との違いにこだわらないのであれば、「発展型評価」は「伴走型評価」と訳出するのがよいのかもしれない。評価者と事業実施者の二人三脚は「発展型評価」の理想とする姿である。「発展型評価」においては、評価者のカウンターパートは社会的イノベーションの実践者と位置づけられている。評価は社会変革のためにあるという評価理論の伝統は、発展型評価に至って、社会的イノベーターという相方を見つけ、次のステージへと歩みを進めている。

Final version accepted May 12, 2016

参考文献

Dozois, Elizabeth; Langlois, Marc and Blancet-Cohen, Natasha (2010) *DE201: A Practitioner's Guide to Developmental Evaluation*, The J.W. McConnell Family Foundation and the Inter-

- national Institute for Child Rights and Development.
EvalPartners (2015) *Global Evaluation Agenda 2016–2020: EvalAgenda2020*, Draft version 7 of 19 December 2015.
Fay, Brian (1975) *Social Theory and Political Practice*, Sydney: Allen & Unwin.
Gamble, Jamie A.A. (2008) *A Developmental Evaluation Primer*, The J.W. McConnell Family Foundation.
今田克司 (2014) 「海外の非営利組織事業評価の動向に関する情報収集」 田中弥生科学研究費調査基盤 (B) 「公共の境界領域の評価～政府、企業と非営利組織の輪郭～」の一環として行われた海外の非営利組織の事業評価に関する動向調査。
国際連合 (2015) 「我々の世界を変革する：持続可能な開発のための2030アジェンダ」 (外務省仮訳) (<http://www.mofa.go.jp/mofaj/files/000101402.pdf>) 2016/2/29.
大住莊四郎 (2003) 『NPMによる行政革命—経営改革モデルの構築と実践』日本評論社。
パットン・マイケル・クイン (2001) 『実用重視の事業評価入門』大森彌監修, 山本泰・長尾眞文編, 清水弘文堂書房。
Patton, Michael Q. (2011) *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*, New York: The Guilford Press.
Patton, Michael Q.; McKegg, Kate and Wehipeihana, Nan (2016) *Developmental Evaluation Exemplars: Principles in Practice*, New York & London: The Guilford Press.
Picciotto, Robert (2015) The 5th wave: Social impact evaluation, Rockefeller Foundation Working Paper.
Preskill, Hallie and Beer, Tanya (2012) *Evaluation Social Innovation*, FSG & Center for Evaluation Innovation.
Preskill, Hallie and Gopal, Srik (2014) *Evaluating Complexity: Propositions for Improving Practice*, FSG Report.
Ramalingam, Ben and Jones, Harry with Reba, Toussaint and Young, John (2008) Exploring the science of complexity: Ideas and implications for development and humanitarian efforts, Working Paper 265, Overseas Development Institute.
Rossi, Peter. H.; Freeman, Howard E. and Lipsey, Mark W. (1999) *Evaluation: A Systematic Approach*, Sixth Edition, Sage Publications.
田中弥生 (2005) 『NPOと社会をつなぐ—NPOを変える評価とインターメディアリ』東京大学出版会。
Vedung, Evert (2010) Four waves of evaluation diffusion, *Evaluation*, vol.16, no.3, pp.263–277.
Westley, Frances; Zimmerman, Brenda and Patton, Michael Q. (2006) *Getting to Maybe: How the World Is Changed*, Vintage Canada.

²¹ Co-Creation の訳で、事業実施者と評価者が評価を共に創りあげていく志向性を表わしている。

要 約

市民社会組織（CSO）にとって、評価が学びとなり、社会課題の解決に向けた効果的なツールとなるとの着眼点から、日本のNPO/NGOに向けた評価普及の試みに取り組んでいる。その観点から、本論は、グローバルな場で起こっている2つの流れに注目した。

ひとつは、2015年、持続可能な開発目標（SDGs）採択の年を契機として、評価の担い手の育成が急務との発想をもとに、国レベルでの質の高い評価のための環境整備や評価専門機関とCSOの基盤強化などを主眼とする今後5年間のグローバル評価アジェンダが作られたことだ。もうひとつは、評価専門家のあいだで進展しつつある、評価の「第5の波」と呼ばれる潮流である。それは、評価の対象のみならず、それを取り囲む関係者が多様になり、事業の「成果」と同様に「価値」に注目が集まり、評価結果の活用も投資目的など、急速に幅広いものになっている状況をさす。

こうした流れのなかで注目したいのが、システム理論や複雑系理論を援用した「発展型評価」と呼ばれる評価アプローチで、従来のやや静態的な評価アプローチから、より動的なアプローチそして開発されているものだ。これらは、社会的イノベーションなどをモニターし、支援するものとしての使い道が議論されている。

キーワード： 評価、市民社会組織、持続可能な開発目標、複雑系理論、発展型評価