

技術的媒介における行為性の再考

柿 崎 慎 也

(山形大学人文社会科学部)

はじめに

非人間（人工物等）と人間の関係において、いかなる存在が「行為する (to act)」のか、という問いは、長らく技術哲学での根本的な関心の一つであり続けてきた¹。

この問いは、技術が単なる人間の意図を実現するための受動的な道具であるという古典的な見方を乗り越え、技術そのものが持つ能動的な役割をいかに捉えるかという課題に直結する。本稿が扱う「行為性 (agency)」とは、この問題系における中核的な概念であり、人間以外の存在が「行為している」と見なされる条件を問うものである。とりわけ、科学哲学者であるブリュノ・ラトゥールによるアクターネットワーク理論 (ANT) は、この問いに対して「人間と非人間の区別を設けず、行為者 (actor) とは他の存在に差異を生じさせるものすべてである」(Latour, 1988, p.159) として、行為者の概念を大きく広げる画期的な転換を示した²。

スピードバンプが物理的に速度を落とさせ、ドアクローザーが自動的に扉を閉めるように³、非人間が単なる人間の道具ではなく、社会的秩序や知

識の構築において能動的な役割を担いうることを明らかにし、「非人間の行為性」、すなわち非人間が持つ「行為をもたらす力」が技術哲学における一つの論点として定位されたのである。

ラトゥールの影響を受けつつも、より現象学的な視点に立脚するピーター・ポール・フェルベークは、非人間が人間の知覚や行為、倫理的判断にどのように媒介的に関与するかを論じた (Verbeek, 2005)⁴。

彼は超音波診断装置が医師の診断行為を媒介し、新たな倫理的判断の場面を構成するように (Verbeek, 2005, pp.128-130)⁵、非人間が人間の行為を方向づけ、委任を求める「逆方向の代理」という概念を提起した (Verbeek, 2005, pp. 165-170)⁶。これにより、技術と人間の関係が単なる一方向の機能的構造ではなく、相互に行為性を構成し合う場であることが示された。フェルベークの議論は、技術の設計そのものが倫理を物質化する過程であることの重要性を浮き彫りにした (Verbeek, 2006)⁷。これらの理論的進展は、行

¹ Heidegger, M. (1977). ハイデガーが技術について世界を開示する一つの様態 (Ge-stell) として論じて以来、技術哲学の中心的課題の一つであり続けている。

² Latour, B. (1988). *The Pasteurization of France*, p. 159. ラトゥールは、行為者 (アクター) とは、それが人間であるか非人間であるかを問わず、他の存在に何らかの差異や変化をもたらすすべての存在を指すとした。この定義により、微生物やホタテ、スピードバンプなども社会を形成するアクターとして分析の対象となる。

³ Latour, B. (1992). "Where Are the Missing Masses?". In W. Bijker & J. Law (Eds.), *Shaping Technology/Building Society* (pp. 226-229). スピードバンプの例は、物理的な非人間 (人工物等) がいかにして人間の行動 (自動車の運転速度) に影響を与え、規範 (速度制限) を強制するかを示す、ラトゥールの代表的な分析対象である。

⁴ Verbeek, P.-P. (2005). *What Things Do: Philosophical Reflections on Technology, Agency, and Human-Technology Relations*. フェルベークは本書で、現象学と ANT を統合し、技術が人間の経験と行為をいかにして媒介し、形成するかを論じた。

⁵ Verbeek, P.-P. (2005). *What Things Do*, pp. 128-130. 超音波診断装置は、胎児を映像として可視化することで、親に「胎児に対する親としての判断」という新たな倫理的責任を課す。この技術は単に情報を提供するだけでなく、親のあり方そのものを変容させる道徳的媒介者として機能する。

⁶ Verbeek, P.-P. (2005). *What Things Do*, pp. 165-170. 「逆方向の代理」とは、人間が技術に行為を委任するだけでなく、逆に技術が人間に特定の行為 (判断や責任) を委任し、その行動を方向づけることを指す。

⁷ Verbeek, P.-P. (2006). "Materializing Morality". *Science, Technology & Human Values*, 31 (3), 361-380. フェルベークは、倫理は抽象的な原則だけでなく、技術的な非人間 (人工物等) の設計の中に具体的に「物質化」されると主張し、設計段階からの倫理的配慮の重要性を

為性を人間固有の属性と見なす伝統的な考え方に再考を迫るものであった。

こうした流れの中で、近年の研究は、人間と非人間の相互作用の構造をより精密に分析する方向へと進んでいる。著者は、ラトゥールの「代理（delegation）」概念を再構成し、そのメカニズムを理論的に補完する試みを行い、代理のプロセスを「翻訳」「分節化」「指図・規範」という三つの操作と、「表象」「代替」「変容」「転位」「行為」という五つの構成要素に基づいて分析した（柿崎, 2024）⁸。この議論の核心は、非人間が人間の行為を単に代理するだけでなく、その相互作用の過程を通じて、それまで存在しなかった新たな行為の構造そのものを生み出すという点にある⁹。この媒介性に着目することで行為性を本質的・固定的な属性としてではなく、関係的構成の中で流動的に形成されるものとして捉える視座が開かれるが、それは同時に、行為性の可変性と文脈依存性に基づく理論的再定位に向けた出発点ともなるのである。

しかし、ラトゥール、フェルベーク、そして著者の議論をもってしても、なお解明されずに残る根源的な問いが存在する。それは、行為そのものが「どこから始まるのか」という、いわば行為性の起源論的問題である。先行研究は、行為が「いかにして」媒介され、変容し、特定の形態をとるに至るかを精緻に描写してきたが、その出発点については、多くの場合、個別のアクター（人間的または非人間的的存在）のいずれかに暗黙のうちに起源を帰属させる傾向が残されている¹⁰。

強調した。

⁸ 柿崎慎也 (2024). 「B. Latour における人間及び非人間への代理 (delegation) 概念の検討」. 『山形大学人文社会科学部研究年報』第二十一号, p. 121.

⁹ 柿崎慎也. (2025). 「代理による科学的対象構築プロセスの解明」. 『東北哲学会年報』, pp. 1-20.

¹⁰ ラトゥールやフェルベークは起源論的な問いを直接立てず、ネットワークの不確定性や相互構築的な媒介性に着目している (Latour, 2005; Verbeek, 2005)。しかし本稿では、彼らの理論においても行為性が「既存のアクター」から発し、媒介を通じて変容するという図式が暗黙のうちに残されていると論じる。具体的には、ラトゥールは行為の出発点をアクターに還元する傾向があり、フェル

こうした先行研究の到達点と残された課題を踏まえ、本稿は、「媒介をめぐる問いは、もはや行為主体の所在に帰着するのではなく、諸要素の関係から生起する力の構成様態に向けられるべきである」という立場を基軸に据える。その上で、フェルベークおよび筆者による先行の議論を理論的に展開し、行為性をあらかじめ個別のアクターに内在する性質としてではなく、人間、非人間的的存在（人工物等）、データ、環境などの異質な要素が交差・接続する関係的構成のなかで同時に立ち現れる出来事的過程として再定位する視点を提示する。

本概念が対象とするのは、行為性、すなわち行為を生じさせる因子が、あらかじめ特定のアクターに内在する属性として与えられているのではなく、複数の存在者のあいだに形成される関係的構造の中から生成されるという動的過程である。この立場は、行為の担い手を特定することに重点を置くのではなく、行為性そのものがいかにして構成されるかという関係的な形成条件に注目するものである。

この理論的視座の妥当性を検証するために、本稿は以下の問いを主軸に構成される。第一に、従来の「媒介」という枠組みが抱える理論的境界は何か。第二に、フェルベークと著者の議論を批判的に継承し、媒介を「関係を形づくる力」として再定義する新たな視点はいかにして導出されるか。第三に、その理論的視座は、現代の具体的な科学技術の実践においてどのように応用可能か。

本稿では、この問いに対する具体的な検討対象として、AI シミュレーションの技術的發展とその社会的含意を取り上げる。AI シミュレーションは、設計者、アルゴリズム、訓練データ、ユーザー、ならびにシミュレートされる環境といった複数の要素が相互に関連し、継続的に再構成され

ベークの「逆方向の代理」も人間的意図を起点とする非対称性を前提としている。本稿は、この暗黙の起源論的問いを克服するため、バラッドの「内在的關係性 (intra-action)」を援用し、行為性が実体に先行するのではなく、複数の存在が交錯する関係性そのものから「同時に立ち現れる出来事」として再定位されると主張する。

る領域である。

このような構造の下では、行為の主体を特定のアクターに一義的に帰属させることが困難であり、関係性の構成に注目する必要性が生じる。したがって、本稿で提案する媒介性に関する理論的枠組みは、このような複雑な構造を分析する上で一定の有効性を持つと考えられる。

本稿は、以下の構成で展開する。第1章では、技術哲学を中心とした媒介概念の理論史の変遷を、ラトゥールとフェルベークの議論を中心に検討し、著者による代理構成に関する議論を接続点として整理する。第2章では、著者の議論を批判的に継承しつつ、物理学者でありニューマテリアリズム（New materialism）の代表的論者の一人であるカレン・バラッド（Karen Michelle Barad）の「内在的關係性（intra-action）」の概念を補助線としながら、行為性が関係から立ち現れる動態を理論的に詳述する。第3章では、AIシミュレーションの発展の歴史と、その複層的な構造的特性を明らかにする。第4章では、DeepMindやOpenAIの研究、デジタルツイン、世界モデルといった具体的な事例を取り上げ、行為性が関係性から同時に立ち現れるプロセスがAIシミュレーションにおいてどのように具現化されているかを分析する。第5章では、これまでの分析を踏まえ、本稿の視点が媒介理論をいかに再構築し、その新規性と貢献がどこにあるのかを論じる。第6章では、この関係性から生起する行為性が含意する非線形的な時間性について考察する。最後に結論として、本稿全体の意義を要約し、今後の技術哲学における媒介概念の方向性を展望する。

本稿の目的は、行為性を個々のアクターの属性としてではなく、関係性がもたらす出来事として捉え直すことにより、複雑化する現代のテクノロジー（技術）と人間の関係性を理解するための新たな理論的・実践的枠組みを構築することにある。

第1章 技術哲学における媒介概念の理論的系譜

技術哲学における媒介（mediation）という概念は、その理論的射程を大きく広げてきた。かつて媒介は、主体である人間が目的を達成するために用いる「手段」として、その中立性が前提とされていた¹¹。この古典的道具観において、技術は人間の意図に従属する客体であり、世界との関係を変える力を持つとは考えられてこなかった。このような「中立的媒介」観に対し、20世紀後半の技術哲学は決定的な批判を加え、技術を人間経験や社会関係を積極的に構成する能動的な存在として捉え直す理論的転換を遂げた。本章では、この転換を駆動したラトゥールとフェルベークの議論を整理し、さらに著者による代理概念の再構成を接続点として、次章で提示する視座への理論的準備を整える。

この転換において、最もラディカルなアプローチを提示したのが、ラトゥールである。彼はアクターネットワーク理論（ANT）において、媒介を「行為性の翻訳作用」と捉え、人間と非人間の区別を原理的に無効にする「一般化された対称性の原理」を提唱した（Latour, 1987）¹²。ラトゥールによれば、差異を生み出す存在はすべて「アクター」であり、その振る舞いを追跡することで、科学技術や社会がいかにして形成されるかが見えてくるとされる（Latour, 1988, p. 159）。彼は、ホテルのキーホルダーに取り付けられた重りが宿泊客に鍵の返却を促すように、あるいはスピードバンプがドライバーに減速を強いるように、非人間が規範を物質的に埋め込む役割を担っていると論

¹¹ Heidegger, M. (1977). このような技術を単なる「手段」と見なす中立的道具観に対し、ハイデガーは技術を存在の真理が開示される様態として捉え直し、その見方を根本から批判した。

¹² Latour, B. (1987). Science in action: How to follow scientists and engineers through society. 本書でラトゥールは、科学的知識や技術的産物が、いかにして多様な人間的・非人間的アクターからなるネットワークの構築を通じて安定化していくかを追跡する方法論を提示した。

じた (Latour, 1992, pp. 226-229)¹³。ラトゥールの理論において、この媒介の変形作用は「仲介項 (Intermediary)」と「媒介項 (Mediator)」という重要な対比によって明確にされる。前者が意図や情報を変えることなくそのまま伝える受動的な存在であるのに対し、後者は「関係のなかで意味や作用を翻訳し、変形させ、新たなものを付け加える」(Latour, 2005, p. 39) 能動的な存在である¹⁴。ラトゥールは、非人間は常に後者、すなわち「媒介項」として扱われるべきだと主張し、それによって社会的なものの構成における非人間の決定的な役割を可視化した。

フェルベークは、このラトゥールの ANT 的視点に、現象学的な洞察を接続する形で、独自の媒介理論を構築した。彼は、非人間の行為性、すなわち「行為をもたらす力」を、人間と世界との関係を構成する要因として捉える (Verbeek, 2005, pp. 123-127)。フェルベークにとって、非人間は単に機能を提供するだけでなく、人間の倫理的判断や行為の仕方に決定的な影響を与える「道徳的媒介者」である (Verbeek, 2005, p. 114)¹⁵。彼は、人間が非人間に行為を委任するだけでなく、逆に非人間が人間に特定の行為を委任し、その行動を方向づける「逆方向の代理」というべき概念に言及している。(Verbeek, 2005, pp. 165-170)。例えば、超音波診断装置は、胎児の状態を可視化することで、親に対してそれまで存在しなかった「胎児に対する親としての判断」という新たな倫理的責任を課す (Verbeek, 2005, p.33)。この装置は、医師の知覚や診断行為を媒介するだけでなく、親の存在論的なあり方をそのものを再構成するのである。こうした分析を通じて、

フェルベークは技術の倫理性を設計の段階に埋め込まれたものとして捉える「設計倫理 (design ethics)」の重要性を強調し、技術と倫理の不可分性を論じた (Verbeek, 2006)¹⁶。

このように、媒介という概念は、従来の中立的かつ受動的な手段としての理解から脱却し、人間の経験や社会的実在を構成・形成する能動的かつ構成的な力として再定位されるに至っている。この理論的展開を敷衍する形で、著者はラトゥールの「代理」概念の精緻化に取り組んでいる。これまで比較的包括的に用いられてきた「代理」という概念を、より分析的な解像度で捉え直すことを試み、ラトゥールの代理概念がどのように成立するかを明らかにするために、そのプロセスを三つの横断的操作——すなわち翻訳 (translation)」「分節化 (articulation)」「指図・規範 (prescription)」——と、五つの具体的構成要素——「表象 (representation)」「代替 (substitution)」「変容 (transformation)」「転位 (dislocation)」「行為 (action)」——に体系的に分割し、媒介の働きがいかにか形成されるかという構造を理論的に可視化した。(柿崎, 2025, pp.8-10)。「翻訳」とは、異なるアクター (人間, 非人間) 間の利害や目的を調整し、新たな関係性を創出する過程であり¹⁷、「分節化」は、複雑な現象や概念を扱いやすい単位へと分解・再編成する操作を指す。「そして「指図・規範」は、そのプロセス全体に秩序や倫理的な方向性を与える。これらの操作を通じて、代理の五つの要素が具体的に展開される。

例えば、3D プリンタは設計者の意図を物理的な形へと「表象」し、手作業を機械の動作で「代替」する。その過程で、設計データは物理的な積層物へと「変容」し、デジタル空間から物理空間

¹³ Latour, B. (1992).

¹⁴ Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*, p. 39. 「仲介項」は入力を変えずに伝達するが、「媒介項」は入力を翻訳し、変形させ、新たなものを付け加える。ラトゥールによれば、社会的なものを理解するためには、後者の「媒介項」の働きを追跡することが不可欠である。

¹⁵ Verbeek, P.-P. (2005). フェルベークは、技術を単なる機能的な道具ではなく、人間の道徳的判断や実践を形成する「能動的な媒介者」として捉えることを提唱した。

¹⁶ Verbeek, P.-P. (2006). "Materializing Morality". *Science, Technology & Human Values*, 31 (3), 361-380. 「設計倫理」は、倫理的配慮を製品が完成した後の使用段階で付け加えるのではなく、設計プロセスそのものに組み込むことを目指すアプローチである。

¹⁷ ここで言う「翻訳」の概念は、異なるアクターの利害を調整し、一つのネットワークへと登録していく過程を指す。

へと「転位」する。最終的に、これらの連鎖が一体となってオブジェクトを造形するという「行為」が達成される。著者は、代理を単なる一方向的な機能の委任としてではなく、多様な要素と操作が絡み合いながら新たな意味と行為の構造をもたらす動的な場として捉えている。（柿崎, 2025, p.10）。

以上のように、技術哲学における媒介概念は、ラトゥールの「翻訳」としての理解、フェルベークの「倫理的・関係論」としての理解として深化されてきており、これらの理論は、行為性がもはや人間だけのものではなく、人間と技術の関係性の中にこそ見出されるべきであることを示唆している。

しかし、これら先行研究での分析にもかかわらず、行為が「いかなる場所から」立ち現れるのかという、その起源論的問題は依然として未解明のままである。先行研究は主として、行為の「変容」や「形成」といったプロセスに注目し、行為性がいかに構築されるかを記述してきたが、行為性そのものが「出現」する契機——すなわち、行為が初めて現実の中で立ち上がる瞬間——を理論的に捉えるための語彙や枠組みを未だ十分には備えていないと思われる。この残された課題に応えるために、本稿は次章以降、媒介の場を「行為の関係論的生成」によって特徴づける新たな視点の提示へと進む。

第2章 行為性の起源を問う

前章で概観したように、技術的媒介の理論はラトゥール、フェルベークらを経て、非人間を単なる客体ではなく、人間と共に行為の場を構成する能動的な存在として認識する方向へと大きく進展してきたが、これらの理論的達成においてもなお、媒介の構造は「既存のアクター間で、行為がどのように成立し、変容するか」を記述する水準に留まっている。

行為そのものが「どこから始まるのか」、すなわち、行為性の起源論的な問題には、依然として十分な光が当てられていない。この理論的間隙を

埋めるために、本章では、行為性が特定のアクターに事前に帰属するのではなく、「複数の存在が交錯する関係の場で同時的に生起する」という視点を提起する。

この視点の中核をなすのは、行為者性、すなわち「行為をもたらす力」を、個々のアクターが所有する「発現する力」としてではなく、「関係の交錯によって引き起こされる出来事的な出現」として捉える根本的な転換である。したがって、ここで論じるのは、人間と非人間が相互に影響を与え合うという逐次的なプロセスではない。それは、人間と非人間を含む複数の存在が、ある特定の関係性の配置につくことで、それまで存在しなかった「行為の可能性と、行為者として振る舞う能力そのものを同時に獲得する」、その瞬間の出来事を指す。ここにおいて媒介性とは、もはや単なる機能的な接続手段や情報伝達の経路ではなく、新たな存在論的な条件を形づくる「形成の場」そのものとして理解される。

この理論的枠組みを深化させる上で、物理学者でありニューマテリアリズム (New materialism) の代表的論者の一人であるカレン・バラッドの「内在的關係性」の概念が、決定的な鍵を握る (Barad, 2007)¹⁸。バラッドは、量子物理学の洞察に基づき、従来の「相互作用 (interaction)」という概念を批判する。相互作用は、あらかじめ独立して存在する実体 (アクター) が、後から互いに影響を及ぼし合うという想定に立っている。これに対しバラッドが提唱する「内在的關係性」は、「実体は関係に先行して存在するのではなく、関係性そのものの内側から、特定の『現象』として『立ち現れる』」 (Barad, 2007, p. 33) と主張する¹⁹。彼女によれば、アク

¹⁸ Barad, K. (2007). Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning. 本書でバラッドは、ニールス・ボーアの量子力学解釈とフェミニスト理論を融合させ、「行為的実在論 (Agential Realism)」という独自の存在論を提唱した。

¹⁹ Barad, K. (2007). Meeting the Universe Halfway, p. 33. 「内在的關係性 (intra-action)」は、あらかじめ存在する独立した実体間の相互作用 (inter-action) とは対照的に、実体が関係性の中から現れるという根源的な生成のプロ

ター（あるいは世界の構成要素）は事前に確定された本質を持つ存在ではなく、「行為を通じて形成される」のであり²⁰、この形成のプロセスは関係性のうちに深く埋め込まれている。この視座において、世界の根源的な構成要素は「モノ」ではなく「関係性」であり、物質と意味、存在と知は分かちがたく絡み合っている（ethico-ontology）（Barad, 2007, p. 185）²¹。

バラッドは、従来の認識論中心の枠組みに対して、存在のあり方そのもの、すなわち事物や現象の構成の仕方に理論的関心を移行させようと試みている。それは世界を前提的に存在する実体の集積としてではなく、関係性や実践の中で生成されるものとして捉える視座の転換である。

この観点に立つバラッドの理論は、行為性を個別のアクターの内在的属性、たとえば主観的意図や外在的因果力に還元することを拒否する。代わりに、行為性は、アクター間の内在的な関係性や差異生成の過程そのものに「分布」しているとされる。このような分布的な行為性の理解は、媒介という概念の再定義を要求する。

すなわち、媒介とはもはや、既存のアクター同士が外在的に相互作用するための「手段的構造」ではなく、むしろそうしたアクター性そのものを可能にする「生成的構造」である。媒介は、アクターを構成する諸力の共振・交差点として機能する。言い換えれば、媒介は「アクターが先に存在し、媒介がその橋渡しをする」という図式ではなく、「媒介の過程においてアクターが事後的に立ち現れる」という構造的な理解へと読み替えられるのである。

具体例として、カーナビゲーションシステムと

セスを指す。

²⁰ *ibid.*, p. 135. バラッドの「行為的実在論」では、世界の基本的な構成単位は独立した「モノ」ではなく、それらが区別される境界線が引かれる「内在的関係」そのものであるとされる。

²¹ *ibid.*, p. 185. バラッドはこの造語を用いて、倫理（ethico）、存在論（onto）、認識論（epistemology）が、現実の構成において分かちがたく絡み合っていることを強調する。

運転者の関係を考えてみよう。フェルベークの「逆方向の代理」の枠組みでは、カーナビゲーションシステムという非人間が、運転者という人間の判断や行為に影響を与え、それを方向づけるという構図が想定されていた（Verbeek, 2005, pp. 149-152）。ここには依然として、事前に存在する「人間」と「非人間」という二項対立が前提として残っている。しかし、バラッドの「内在的関係性」の視座を導入すると、この関係の捉え方は一変する。カーナビゲーションシステムと運転者の「内在的関係」そのものが、それまで存在しなかった新たな「知覚＝判断＝行為の構造体」を一つの現象として「生み出して」いるのである。この構造体の中では、「どこが」運転者で「どこからが」カーナビゲーションシステムなのかを明確に切り分けることはできず、むしろ両者が一体となって初めて「ナビゲートしながら運転する」という特定の行為性が「出現」する。このとき、行為性は運転者にもナビゲーション機器にも偏在するのではなく、「両者を含む関係の配置そのものから生起している」。

このような媒介の再定義は、フェルベークの議論と深く接続し、同時にそれを内側から再編成する可能性を秘めている。フェルベークは、非人間が人間の倫理的実践を構成する「道徳的媒介者」として機能することを、数々の事例を通じて説得的に論じた。

しかしながら、フェルベークの分析には、媒介関係における行為性の起点が、依然として「人間的意図の変容」として描かれる傾向が残されている。「逆方向の代理」という彼の概念が、行為の主体をまず人間に置き、それが非人間によって「逆方向に」影響を受けるという、主体－客体の非対称性を前提としている。

この点において、バラッドの「内在的関係性」の概念は、フェルベークの議論の核心を維持しつつ、その制約を克服するための理論的基盤を提供する。すなわち、「人間と非人間」という二項が媒介関係に先立って存在するのではなく、「両者

が共に行為の場で出現する」という前提に立つことで、媒介の論理は全く異なる相貌を帯びる。もはや「人間の行為が非人間に影響される」のではなく、「人間と非人間の関係性の配置それ自体が、特定の行為性とそれに伴う意味作用を同時に出現させる場」となるのである。このとき、「代理」や「指示」といった語彙は、この出来事の本元的な同時性を捉えるには不十分であり、複数のエージェントが「共に」「同時に」行為を開始するという、その関係性の動態そのものを記述する必要がある。

フェルベークはまた、非人間の倫理的媒介性を「設計倫理 (design ethics)」へと接続し、「倫理的判断が設計の段階で物質的に構成される」(Verbeek, 2006, pp. 361-380) ことを強調した。これにバラッドの視点を導入すると、この議論はさらに深化する。「設計される対象」と「設計する主体」という区別そのものが、絶対的なものではなく、特定の形成的な関係として再構成される。倫理的判断は、設計者の頭の中のみ存在するものではなく、設計ツール、素材、プロトタイプ、そして未来のユーザーといった人間と非人間の関係性のうちにこそ、その萌芽を見出すべきものとして理解されるようになる。

このように、バラッドとフェルベークの議論を接続することで、媒介の理論を「誰が行為するのか」というアクター中心の問いから、「どのように行為性が共に出現するのか」という関係性中心の問いへと転回させることが可能となる。この転回こそが、本稿が目指すものである。それは、既存の媒介概念が暗黙のうちに前提としてきた主客、因果、意図といった図式を超えて、「関係的な生成」という水準において行為性を再定義する試みである。非人間と人間は、もはや主体と客体、あるいは原因と結果ではなく、一つの出来事の中で「共に構成される関係的存在」として理解されねばならない。

第3章 AI シミュレーションの構造特性

前章では、行為性が個々のアクターに内在するのではなく、関係の場で同時的に立ち現れるという視点を理論的に構築した。本章では、この視点を具体的な科学技術の実践に適用する事例として、AI（人工知能）シミュレーションの技術的發展とその構造的な特性を明らかにする。AI シミュレーションは、アルゴリズム、ユーザー、設計者、訓練データ、そしてシミュレートされる環境といった、極めて多様な構成要素が相互に、かつ複層的に交差する場面である。それゆえに、単一のアクターに行為の起点を求めることが本質的に困難であり、本稿が提唱する関係論的な視点が最も明瞭にその妥当性を発揮する事例の一つと言える。本章の目的は、AI シミュレーションの開発史を辿りながらその構造を解き明かし、次章で展開する現象分析のための土台を築くことにある。

AI とシミュレーションは、それぞれが独自の発展経路をたどりながらも、その黎明期から深い関わりを持っていた。人工知能の学問的探求は、アラン・チューリング (Alan Turing) が論文「計算する機械と知能」で知的な機械の可能性を問い、「模倣ゲーム」を提案した1950年に遡る (Turing, 1950)²²。そして1956年のダートマス会議で「人工知能 (Artificial Intelligence)」という言葉が使用され、正式な学術領域として発足したとされる。初期の AI 研究は、人間の論理的推論や記号処理による知的活動の模倣にその焦点を当てていた²³。この文脈において、シミュレーションはチェスのようなゲームの盤面や、特定の課題解決の環境をコンピュータ内に再現する「知能の再現環境」として不可欠な役割を果たした (Newell & Simon,

²² Turing, A. M. (1950). "Computing Machinery and Intelligence". *Mind*, 49, 433-460. この論文で提案された「模倣ゲーム」(後のチューリング・テスト)は、機械が知性を持つかどうかを判断するための思考実験であり、AI 研究の哲学的基礎となった。

²³ この初期の AI は、記号処理と論理推論を中心としていたため、「古き良き AI (Good Old-Fashioned AI, GOF AI)」とも呼ばれる。

1976)²⁴。

1980年代以降、シミュレーション技術は複雑系科学の潮流と結びつき、既知現象の再現のみならず、未知現象の創発過程を探索する手段へとその役割を変化させていくこととなり（Epstein & Axtell, 1996）²⁵、2000年代に入ると、爆発的なデータ量の増大と計算機能力の向上を背景に、AIの主流は統計的学習理論、特に「深層学習（Deep Learning）」へと劇的に移行した（LeCun, Bengio, & Hinton, 2015）²⁶。深層学習の台頭により、AIとシミュレーションの連携は単なる環境の再現に留まらず、未来を予測し最適な戦略を発見する「予測的实践」へと質的な転換を遂げている。AIは、現象の背後にある物理法則や因果構造を人間のように「理解」するのではなく、膨大な観測データやシミュレーションデータの中から統計的な相関関係を学習し、それに基づいて最も確からしい「傾向を提示する」装置となったのである。このとき、シミュレーションはAIを訓練するためのデータ生成装置であると同時に、AIの学習プロセスそのものを方向づけ、AIの能力を形成する能動的な媒介者として機能するようになる。

この時期のAI研究、特にシミュレーションとの関連で画期的な発展を牽引したのが、DeepMindやOpenAIといった研究開発企業の登場である。DeepMindは、深層学習と「強化学習（Reinforcement Learning）」²⁷を組み合わせた深

層強化学習の分野で顕著な成果を上げた。2016年には、囲碁AI「AlphaGo」が世界トップ棋士である李世石に勝利し、世界に衝撃を与えた（Silver et al., 2016）²⁸。

これらのシステムでは、AIがシミュレートされた環境内で膨大な回数の試行錯誤を繰り返す。ここで重要なのは、AIが単に人間から与えられたルールに従うだけでなく、「環境との相互作用を通じて、人間が思いもよらなかった新たな振る舞いや戦略を同時に形成する点」である。これは、ラトゥールが言うところの、単なる「仲介項」ではない、新たなものを付け加える「媒介項」としてAIが機能していることを示す。

一方、OpenAIは大規模言語モデル（LLM）の開発を世界的に牽引しており、GPT-3などに代表されるこれらのモデルは、膨大なテキストデータを学習することで人間が生成したかのような自然な文章を生み出す能力を持つ（Brown et al., 2020）。このLLMがシミュレーションと結びつくことで、ANTにおける「行為者」としての人間のエージェント²⁹の、より現実的で複雑な振る舞いを形成することが可能になる。ここでいうエージェントとは、単に指示を実行する存在ではなく、環境との相互作用を通じて自律的に行動し、結果を生み出す能力を持つ行為性を帯びた存在を指す。AIとシミュレーションの融合は、特に「デジタルツイン（Digital Twin）」³⁰の概念と「世界モデル（World Model）」³¹の研究において、その

²⁴ Newell, A., & Simon, H. A. (1976). "Computer Science as Empirical Inquiry: Symbols and Search". *Communications of the ACM*, 19 (3), 113-126. ニューウェルとサイモンは、人間の問題解決プロセスを情報処理の観点からモデル化し、AI研究の基礎を築いた。

²⁵ エージェントベース・モデリング（ABM）を用いて、個々のエージェントの単純なルールから、いかにして複雑な社会現象が創発するかを示した画期的な研究。

²⁶ LeCun, Y., Bengio, Y., & Hinton, G. (2015). "Deep learning". *Nature*, 521 (7553), 436-444. 深層学習の三人のパイオニアによる論文。人間の脳の神経回路網に着想を得たニューラルネットワークを多層に重ねることで、データから複雑な特徴を自動で学習する能力を実現した。

²⁷ Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. 強化学習の標準的な教科書。エージェントが環境内で試行錯誤し、与えられる「報酬」を最大化するように自律的に行動方針を学習する手法を体系的に解説している。

²⁸ Silver, D., et al. (2016). "Mastering the game of Go with deep neural networks and tree search". *Nature*, 529, 484-489. DeepMindが開発したAlphaGoが、囲碁の世界チャンピオンである李世石を破った歴史的な対局について報告した論文。

²⁹ 「人間のエージェント」は、AIがその行動や意思決定を学習・模倣・予測する対象としての人間を指す。

³⁰ Grieves, M. (2016). *Digital Twin: Manufacturing in the Age of Digitalization*. デジタルツインの概念を提唱したマイケル・グリーブスによる解説。物理的な製品やシステムのライフサイクル全体を通じて、仮想空間上のモデルと連携させる構想。

³¹ Ha, D., & Schmidhuber, J. (2018). "World Models". *arXiv preprint arXiv:1803.10122*. AIエージェントが、環境のダイナミクスを予測するための内部モデル（世界モデル）を学習し、それを用いて仮想的なシミュレーショ

重要性を一層増している。

デジタルツインとは、物理的な実体やシステムの仮想的なレプリカをコンピュータ上に構築し、リアルタイムデータとシミュレーションモデルを統合する技術である。デジタルツイン上でのシミュレーション結果は物理世界への操作や改善にフィードバックされ、その結果生じた物理世界の新たな状態が再びデジタルツインに反映される。この循環的なプロセスにおいて、物理的な「モノ」と仮想的な「情報」はAIを介して関係的に結びつき、新たな行為の可能性をその関係性の中から発揮していると解釈できる。

世界モデルとは、AI エージェント³²が環境のダイナミクスを、自らの内部に学習によって獲得したモデルである。この内部モデルを持つことで、エージェントは行動の結果を事前に予測し、より効率的に計画を立てることが可能になる³³。世界モデルの研究は、AI が環境との相互作用を通じて、自律的に「知識」を構築し、それに基づいて新たな行動を形成する可能性を示唆しており、著者が論じる代理の「形成機能」、すなわち、単なる代替ではない新たな構造を生み出すプロセスと深く呼応する。

しかし、こうした目覚ましい進展には多くの課題も伴う。AI シミュレーションは、訓練データに内在するバイアスの影響を強く受け、既存の偏見を再生産・増幅させてしまう危険性がある（Crawford, 2021）³⁴。

ン内で行動計画を立てるというアプローチを提案した。

³² 特定の目標を達成するために、環境を感知し、推論し、行動する能力を持つ自律的なシステムを指す。AI エージェントと人間エージェントは、行為性の起源が異なると言える。人間は意識や身体を通じた環境との相互作用から行為性を発揮するが、AI はプログラムや世界モデルに基づく内部表現によって自律的に行動する。

³³ Hafner, D., et al. (2020). "Mastering Atari with Discrete World Models". International Conference on Learning Representations (ICLR). Dreamer と名付けられたエージェントが、観測された高次元の画像データから世界の重要な情報だけを抽出した低次元の潜在空間モデルを学習し、その中で「想像」することで効率的にタスクを遂行することを示した。

³⁴ Crawford, K. (2021). Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence. AI が中立的

また、深層学習モデルの内部的な意思決定プロセスが人間には解釈困難である「ブラックボックス問題」も深刻である（Mittelstadt et al., 2016）³⁵。AI が提示する出力を、その根拠を理解しないまま判断を委ねてしまうことで、責任の所在が曖昧になる。これは、フェルベークが警鐘を鳴らす、技術による倫理的判断の変容が、必ずしも望ましい形で発現するとは限らないことを示している。

このように、AI とシミュレーションが統合された領域は、単なる技術的課題を超えて、「意味と判断がどこから生まれるのか」という根源的な哲学的問題に直結している。その複雑な構造を捉えるには、個々のアクターの性能を分析するだけでは不十分であり、多様な要素が織りなす媒介性に基づく関係論的視座、すなわち、「行為が関係性からいかにして生起するか」を分析する視点が不可欠となる。

第4章 AI シミュレーションにおける行為性の関係論的分析

前章では、AI シミュレーションの技術的發展と、それがもたらす複層的な関係構造を明らかにした。本章では、第2章で提示した関係論的な視座、すなわち行為性が多様な要素の「内在の関係性」から同時的に立ち現れるという枠組みを用い、AI シミュレーションの具体的な実践場面を分析する。特に、① AI の訓練プロセス、② デジタルツインの運用、③ AI による世界モデルの構築、という三つの側面から、行為性が生成される関係性の動態を詳細に解き明かす。ここでの目的は、AI の行為性が単一のアクターに由来するのではなく、具体的な現象を通して、その関係論的な生成の様相を示すことにある。

な技術ではなく、その開発と運用が権力、政治、そして地球環境に深く根ざしたものであることを、データ収集からアルゴリズム、インフラに至るまで多角的に論じている。

³⁵ Mittelstadt, B. D., et al. (2016). "The ethics of algorithms: Mapping the debate". Big Data & Society, 3 (2), 1-21. アルゴリズムがもたらす倫理的問題を体系的に整理し、特にその不透明性（ブラックボックス性）やバイアス、説明責任の欠如といった課題を指摘している。

4. 1. AI 訓練過程における行為性の生成

AI, 特に深層学習モデルの訓練は, 人間（設計者）, データ, アルゴリズム, そして計算資源が複雑に絡み合うプロセスである。このプロセスは, 各要素が単に相互作用（interaction）するだけでなく, バラドの言う「内在的関係性」を通じて, 行為性を同時的に生み出す様相を呈している（Barad, 2007）。

AI モデルの訓練において, データはアルゴリズムを「学習」させるための基盤となる。しかし, データは単なる静的な情報源ではなく, アルゴリズムとの関係性の中で初めて意味を持つ能動的な存在である。画像認識 AI, 特に深層畳み込みニューラルネットワーク（DCNN）の訓練を考えてみよう³⁶。ImageNet のような巨大なデータセットには数百万枚のラベル付けされた画像が含まれているが, これらのデータが個別に概念の意味を内包しているわけではない。

例えば, データセット中の特定の画像に「猫」というラベルが付与されていたとしても, その画像自体が「猫」という普遍的な概念を直接的に表現しているわけではない。アルゴリズムが, 画像データの中から特定の特徴（例えば, 耳の形状, 目の配置, 毛並みなど）を抽出し, それらの特徴間の関係性を統計的に学習することで, 初めてモデル内部に「猫らしさ」といった認識の構造が「形成される」。この「猫らしさ」は, 個々のデータに内在するのではなく, データとアルゴリズムの相互作用によって創発的に構築される概念である。したがって, 猫の具体的なイメージは, AI がデータを学習し, 分類能力を獲得する過程で生じる認識の一例として挙げられる。このプロセスは, データがアルゴリズムを一方向的に駆動する線形的なものではない。むしろ, データセットに内在するパ

³⁶ Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). "ImageNet Classification with Deep Convolutional Neural Networks." *Advances in Neural Information Processing Systems*, 25. AlexNet と呼ばれる DCNN が, 画像認識コンペティション ILSVRC で圧勝し, 深層学習ブームの火付け役となった歴史的な論文。

ターンや偏りと, アルゴリズムが持つ特定の構造的バイアスが「内在的」に関係し合うことで, 特定の「認識能力」という現象が同時に出現するのである。

これは, データとアルゴリズムが相互に作用し合い, 認識の構造そのものを「共に」構成する関係性の現れである。意味やバイアスは, データかアルゴリズムのどちらか一方に存在するのではなく, 両者の関係性から「もたらされる」のである。

また, AI の設計者は, モデルのアーキテクチャ選択やハイパーパラメータ調整などを通じて, 訓練プロセスに決定的な影響を与える³⁷。しかし, 一度訓練が始まると, 特に強化学習の文脈では, AI は設計者の直接的な制御を超えて, 自律的に学習を進める。

DeepMind が開発した AlphaGo やその後継の AlphaZero は, この点を象徴的に示している。設計者は, 囲碁のルールと報酬関数を設定するだけである。その後, AI エージェントは自己対局というシミュレーションを繰り返す中で, 最適な戦略を自ら探索していく。その過程で, 人間の定石からは考えられないような, しかし極めて効果的な「新手」を「もたらす」ことがある³⁸。

この現象は, 設計者の意図が AI の行動を完全に規定するというトップダウン的な構図でも, AI が完全に自由に振る舞うというボトムアップ的な構図でも説明できない。

むしろ, 設計者による学習環境の「設定」という介入と, AI によるその環境内での自律的な「探索」というプロセスが, 共起的に作用し合うこと

³⁷ Zoph, B., & Le, Q. V. (2016). "Neural architecture search with reinforcement learning". arXiv preprint arXiv:1611.01578 強化学習を用いてニューラルネットワークのアーキテクチャ自体を自動探索する手法を提案した。これは, 人間が手作業で行っていた設計プロセスを自動化し, より高性能なモデルを効率的に発見できる可能性を示した点で画期的である。

³⁸ Metz, C. (2016, March 12). "How Google's AI Viewed the Move No Human Could Understand". *Wired*. AlphaGo が李世乭との第2局で見せた37手目は, 人間のトップ棋士の誰にも理解できず, 当初はバグかと思われたが, 結果的に勝利の鍵となった。これは AI が人間の知識を超えた創造性を示した例として広く知られている。

で、予期せぬ、しかし創造的な行為の形式が出現するのである。設計者の存在がAIの学習空間を限定し、方向づけつつも、その限定された空間の中でこそ、AIは自己組織的に新たな行為の様式を「発揮する」。ここでの行為性は、設計者の意図にも、AIのアルゴリズムにも還元できず、両者の関係性が作り出す動的な場そのものに帰属するのである。これは、著者が論じる、多様な要素が絡み合いながら新たな行為構造をもたらす「代理」の形成プロセスとも重なる。

4. 2. デジタルツインの運用における 行為性の生成

デジタルツインの運用は、物理的な実体と仮想的なモデルがリアルタイムで連携し、相互に影響を及ぼし合う、極めて動的なシステムである。このシステムの中では、行為の関係論的生成が特に明確な形で現れる。

スマートファクトリーにおける製造ラインのデジタルツインを例に考えよう。物理的な製造ラインから収集されたリアルタイムデータは、即座にデジタル空間上の仮想モデルにフィードバックされ、デジタルツインは物理世界の現状を極めて高い忠実度で反映し続ける³⁹。この情報伝達は一方通行ではない。デジタルツイン上ではAIを用いたシミュレーションが実行され、予知保全や生産計画の最適化といった、未来に関する予測や計画が生成される。そして、このデジタルツインが「もたらした」予測や計画は、物理的な製造ラインへの具体的な指示としてフィードバックされ、その「行動」を現実に変容させる。

この反復的で循環的なフィードバックループにおいて、物理的な「モノ」と仮想的な「情報」は、AIを介して「同時に行為を形成し、互いの状態を絶えず変え続ける」。物理世界とデジタルツインは、もはや独立したアクターとして存在するの

³⁹ Tao, F., & Qi, Q. (2019). "Make More Digital Twins". *Nature Machine Intelligence*, 1 (1), 1-2. デジタルツインの応用が製造業だけでなく、より広範な領域に拡大していくことの重要性を説いている。

ではなく、一つの「内在的關係性」の中で行為を「共に生み出す」のである。デジタルツインは、単なる物理世界の鏡像 (representation) ではなく、物理世界そのものと相互に作用し、共に新たな現実を「形成する」生成的な媒介装置として機能する。人間のオペレーターは、デジタルツインが提示するシミュレーション結果に基づいて判断を下す。しかし、この判断は、人間の純粋な意図だけで決定されるものではない。デジタルツインがどのような情報を、どのような形式で提示するかは、人間の知覚の枠組みを構造化し、思考の方向性を強力に「方向づける」。これはフェルベークの言う、技術による「道徳的媒介」の一例である (Verbeek, 2011)。

この関係性において、人間はデジタルツインの出力を単に受動的に受け入れるだけではない。その情報に触発されて新たな問いを立て、シミュレーションのパラメータを変更して別のシナリオを試す。つまり、人間とデジタルツインは、互いの行為を誘発し合いながら意思決定の場を共に作り出すのである。このとき、行為性は人間の判断力にも、デジタルツインの計算能力にも還元できない。それは、人間とAI搭載デジタルツインとの間の「内在的關係性」から出現する。デジタルツインは、人間の意図を単に代理するのではなく、人間の意思決定プロセスそのものに新たな意味と行為性を「出現させる」強力な媒介者として機能するのである。

4. 3. AIによる世界モデルの構築における 行為性の生成

AIによる世界モデルの構築は、環境のダイナミクスをAIが内部的に表現するプロセスであり、ここにおいても行為の関係論的生成が深く関連する。

AIが世界モデルを構築する際、まず大量の環境データが入力される。AIはこれらの時系列データから、環境がどのように変化するかのパターンを学習し、将来を予測するための内部モデルを形

成する (Hafner et al., 2020)⁴⁰。このプロセスは、データが一方向的にモデルを「形成する」というものではない。AIモデルのアーキテクチャ（例えばTransformer）や学習アルゴリズムが、入力データストリームのどの側面を「重要」と見なし、どのように抽象化するかを決定する (Vaswani et al., 2017)⁴¹。つまり、入力される環境データとAIモデルの構造的バイアスが相互に作用し、AI内部における「現実」の表現を同時に形成するのである。

このAIによって構築された「現実」の表現は、客観的な現実の完全な写しではない。しかし、この「ずれ」を持つ内部モデルこそが、AIの行為の源泉となる。AIは現実世界との相互作用を通じて、自らの世界モデルを継続的に更新し、新たな状況に適応するための行動を同時に生み出す。これは、行為の生成が静的なものではなく、環境との持続的な関わりの中で動的に展開するプロセスであることを示している。

AIが世界モデルを構築すると、それは将来の予測や行動計画の策定に用いられ、AIは、物理的な世界で実際に行動する前に、内部世界モデルの中で仮想的なシミュレーション（「想像」とも表現される）を実行し、様々な行動の結果を事前に評価する (Hafner et al., 2019)⁴²。

自動運転車の世界モデルを例にとってみよう。そのモデルは、センサーデータに基づき、周囲の交通状況の未来を予測し、安全な走行経路を計画する (Waymo, 2023)⁴³。このAIによる交通計画は、その内部モデルに基づいて生成されるが、現実の

交通状況という外部環境によってミリ秒単位で絶えず修正・再構築される。自動運転車が予測に基づき行動を起こすと、それが周囲環境に影響を与え、その影響はセンサーデータとしてAIにフィードバックされる。

この連続的なループの中で、AIの内部的な行為（予測と計画）と環境の応答は不可分に結びつき、同時に発生することで、交通システムにおける行為の全体像を形成する。

このようにAIの行為性は、内部モデルに閉じることなく、環境との絶え間ない内在的関係性の中で継続的に形成され、同時に発現する。したがってAIは、環境を単に理解する客体ではなく、環境そのものを変容させ、新たな行為の場を形成する存在となる。

第5章 関係性の動態がもたらす 媒介理論の再構築

これまでの章では、行為性が関係性の力学から立ち現れるという視点を理論的に構築し（第2章）、それがAIシミュレーションという具体的な技術実践においていかに現象的に現れるかを分析してきた（第4章）。本章では、これらの議論を総括し、この関係論的な視点が、従来の技術哲学における媒介理論に対して、いかなる新たな視点を導入し、それをいかに再構成できるかを検討する。目的は、本稿が提示する理論的枠組みの独自性と、技術哲学分野への理論的貢献を明示することである。具体的には、①行為性の起源、②媒介の構成的機能、③関係性の非還元性、という三つの観点から媒介理論の再構築を試みる。

5. 1. アクターから「関係の場の出現」へ

従来の技術哲学における媒介概念は、ラトゥールによるアクターの拡張 (Latour, 2005)、フェルベークによる「逆方向の代理」 (Verbeek, 2005) などを通じて、行為が人間と非人間の相互作用の中で変容し、分散することを明らかにした。しかし、これらの理論は依然として、行為が「既

⁴⁰ Hafner, D., et al. (2020). "Mastering Atari with Discrete World Models". International Conference on Learning Representations (ICLR).

⁴¹ Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). "Attention Is All You Need". Advances in Neural Information Processing Systems 30 (NIPS 2017).

⁴² Hafner, D., Lillicrap, T., Ba, J., Norouzi, M., & Tucker, G. (2019). "Dream to Control: Learning Behaviors by Latent Imagination". arXiv.

⁴³ Waymo. (2023). "Measuring surprise in the wild". arXiv preprint arXiv:2305.07733.

存のアクター」から発せられ、それが媒介を通じて変容するという図式を、完全には乗り越えておらず、行為の「起源」は、最終的には個々のアクターの能力や性質に還元される傾向が残っていた。

これに対し、本稿が提示する視点は、行為性の捉え方そのものに再考を促す。それは、行為性が特定のアクターに事前に内在する性質なのではなく、多様な要素が特定の関係的な配置につくことで、その関係の場そのものから、同時に立ち現れる出来事として行為性を捉えるということである。第4章で分析したAIの訓練プロセスは、この考え方を明確に裏付けている。そこでは、データセット、アルゴリズム、設計者の意図が「内在的」に関係し合うことで、初めて「認識能力」という特定の行為性が「出現」した。この行為性は、いずれかの要素に事前に存在していたわけではない。それは、三者の関係性が整った瞬間に「共に生成された」のである。

このように、本稿の視点は、行為の起源を単一の主体に還元せず、関係性の出現そのものを行為の始まりと捉える。この観点から、媒介理論は「誰が行為するか」という主体の問いから、「どのような関係の配置が、どのような行為性を生み出すのか」という、関係性の動態を問う方向へと再構築される。これは、行為の主体を特定する分析から、行為性が発現する条件を考察する研究への移行を意味する⁴⁴。

5. 2. 媒介の形成能力

本稿は、媒介を既存のアクター間の単なる仲介としてではなく、新たな知覚、判断、および現実自体を構成する動的な場として再定義する。この構成能力の強調は、従来の媒介概念では捉えきれ

⁴⁴ 哲学者ルチアーノ・フロリディ (Luciano Floridi) は「情報行為者性 (informational agency)」の概念を提唱した。彼によれば、情報エージェントとは、環境との相互作用を通じて状態を変化させる存在であり、意図や意識をもたずとも行為をなすものと見なされる (Floridi, 2013, p. 145)。このとき「行為性」は、意識や自由意志の有無に還元されず、「ふるまいに帰属可能な結果」を基準として定義される (Floridi, 2013, p. 148)。

なかった関係性の創造的側面を前景化させる。フェルベークは技術が「道徳的現実を共同構成する」と論じ (Verbeek, 2011) たが、本稿の視点はこの洞察を発展させ、媒介が全く新しい存在論的条件、すなわち「新しい種類の現実」そのものを構築する可能性を提示する。

第4章のデジタルツインの事例は、この点を明確に示している。物理的な製造ラインと仮想的なデジタルツインが形成するフィードバックループは、単に物理世界を効率化するだけではない。それは、物理世界とデジタル世界が分かちがたく結びついた、新たな「現実」を「構築」しているのである。

この新たな現実の中では、何が「実体」で何が「情報」という区別は、もはや自明ではなくなる。

AIによる世界モデルの構築もまた、AI自身の行為の基盤となる独自の機能的環境を生成する。この生成の側面は、媒介が予期せぬ、しかし有意義な新たな事態を誘発する可能性を秘めていることを示す。これは、バラッドの「内在的関係性 (intra-action)」が、関係性そのものから「現象」を切り出す (agential cut) という考え方と共鳴する (Barad, 2007, p. 178)⁴⁵。行為の関係論的生成とは、まさにこの「切り出し」の瞬間であり、新たな現実が関係性の中から形成される動的なプロセスなのである。

5. 3. 関係性の非還元性

最後に、本稿の視点は、人間と非人間の関係性を、主体と客体、原因と結果といった、還元的な二元論的枠組みで捉えることの限界を決定的に示す。AIシミュレーションにおいて観察される人間とAIの複雑な協働関係は、その好例である。第4章で論じたように、デジタルツインを用いた意思決定において、人間の判断とAIの提示する

⁴⁵ Barad, K. (2007). Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning. Duke University Press.

情報は、互いを「形作り」、互いを「修正する」。この相互作用は、どちらか一方に分析的・存在論的な優位性があるわけではない。むしろ、両者が「共に作用し合い」、一つの「意思決定の場」という現象を「生み出す」、不可分な関係にある。

この関係性は、単純な因果の連鎖では記述できない。それは、両者が互いの「条件」となり、互いの行為の可能性を「同時に起動させる」という、循環的で非線形的な関係性である。この「非還元的な関係性」こそが、本稿が捉えようとする中核的な洞察である。

この視座は、AI倫理における責任の所在の問題に対しても新たな光を当てる。自動運転車の事故の責任は誰にあるのか。還元的な思考は、責任を個々のアクターに押し付けようとする。しかし、本稿の視点に立てば、事故という「出来事」は、人間、AI、車両、環境といった多様な要素が織りなす関係性のネットワーク全体から「出現した」ものとして捉えられる。これは責任を曖昧にするための議論ではない。むしろ、責任を個人の帰責の問題から、そのような望ましくない出来事を生み出す「関係性の設計」の問題へと転換させることを促すのである⁴⁶。

結論として、行為性の関係論的生成という概念は、媒介理論を行為の「主体」から「関係の場」へ、媒介の機能を「伝達」から「形成」へ、そして分析の枠組みを「二元論」から「非還元的な関係性」へと、三つのレベルで再構築する。これにより、複雑かつ自律的な現代のテクノロジー（技術）と人間との関係性を適切に捉えるための理論的枠組みを提示することが可能となると考えられる。

⁴⁶ Elish, M. C. (2016). "The Moral Crumple Zone: Cautionary Tales in Human-Robot Interaction". *Engaging Science, Technology, and Society*, 2, 40-60. 事故やエラーが発生した際、責任が複雑なシステム全体ではなく、最も近くにいる人間のオペレーターに不当に押し付けられる傾向を「道徳的クランプルゾーン」と呼び、警鐘を鳴らした。

第6章 関係性の動態が明らかにする時間性に関する考察

本稿で提示してきた、行為性が関係性の力学から生起するという視点は、人間と技術の関係性を空間的な配置や構造の観点から捉え直すだけでなく、その関係性が展開する「時間」という概念そのものに対しても、新たな解釈を要請する。従来の技術哲学における時間概念は、多くの場合、過去の意図が現在の行為を規定し、現在の行為が未来の結果を引き起こすという、線形的な因果律に基づいた一方向の流れとして、暗黙のうちに想定されてきた⁴⁷。しかし、本稿が論じる行為の同時的で非還元的な出現は、この単純な時間認識を根本から問い直し、行為が複数の時間軸の複雑な交錯から立ち現れるより動的な時間像を明らかにする。

本章では、この関係性の動態が含意する時間性について、①時間性の非線形性、②過程としての持続性、③未来の開かれ方、という三つの側面から哲学的考察を深める。

6. 1. 時間性の非線形性と「出来事的現在」の出現

線形的な時間観においては、時間は過去から現在、未来へと不可逆的に流れる。これに対し、本稿が提示する視点に立てば、時間はもはや一方向の流れではなく、複数の、そして異質な時間的要素がダイナミックに絡み合い、予測不可能な形で新たな「現在」をもたらす非線形的な様相を呈する。

この時間性の非線形的な特性を理解するためには、バラッドが提唱する「内在的關係性」の概念が重要な手がかりとなる。バラッドによれば、「出来事 (phenomenon)」とは、事前に固定された

⁴⁷ Prigogine, I., & Stengers, I. (1984). *Order out of Chaos: Man's New Dialogue with Nature*. ニュートン力学に代表される古典科学の世界観では、時間は可逆的で決定論的なものとして扱われたが、散逸構造論は不可逆的で創造的な時間の側面を明らかにした。

個体（アクター）同士の相互作用の結果としてではなく、関係性そのものの内部においてアクターと出来事が「同時に」生成されるプロセス、すなわち「行為的分節（agential cut）」によって成立する（Barad, 2007, p. 140）。この形成の瞬間は、過去の出来事の直線的連続によって機械的に導かれるものではなく、複数の時間的ベクトルが交差し、質的に新しい構成が創出される点にこそ本質がある。すなわち、ここで生じる出来事は、単なる時間の経過や因果的展開とは異なり、非線形的かつ特異な時間構造を有する現象的契機として理解されるべきである。

AI シミュレーションの実践において、この「出来事的現在」の出現は顕著に観察される。深層学習モデルの訓練プロセスでは、訓練データという「過去」の記録の集積と、アルゴリズムという「現在」において作動する学習ルール、そして設計者の意図という「未来」への方向づけが相互に作用する。この相互作用を通じて、モデルはしばしば「予測不可能な形で立ち現れる」新たな知識を生み出す。この「立ち現れる」瞬間こそが、「出来事的現在」に他ならない。そこでは、異なる時間軸を持つ要素が共鳴し合い、新たな行為の可能性を切り開くのである。

6. 2. 時間的厚みと過程としての媒介

本稿は、行為の生成を単一の瞬間で完結する離散的な出来事としてではなく、継続的な過程として捉える動態として理解することを提唱する。AIによる世界モデルの構築は、一度学習すれば完了する静的な過程ではない。AI エージェントは、環境との絶え間ない相互作用を通じて、自らが持つ世界モデルを継続的に更新する。自動運転車の例では、AIは過去の走行経験に基づいて現在の交通状況を予測し、その予測に基づいて行動する。その行動の結果は新たな経験としてフィードバックされ、未来の予測と行動を方向づける。このプロセスは、過去、現在、未来が相互に浸透し合う「時間的な厚み」を持った持続的な過程である。

デジタルツインにおいても同様に、物理世界とデジタル世界との「連続的なフィードバックループ」が存在する。このループは、単なる瞬間的な反応の連鎖ではない。それは、物理システム、デジタルシステム、そして人間が一体となって、「時間の中で共に進化していく」持続的なプロセスなのである。媒介性とは、ある特定の時点における機能的な接続に限定される概念ではない。それは、関係性が時間的経過の中で連続的に自己を再構築し、新たな行為可能性を創出していく生成的な動態として理解される。

6. 3. 未来の開かれ方と不確定性

本稿が提示する理論的枠組みは、未来が過去の延長として決定論的に存在するわけではないことを示している。むしろ未来は、関係性の非線形な相互作用から現れる本質的に予測不可能な領域として認識される。

AI シミュレーションの分野では、この未来の「開かれ方」が特に顕著である。強化学習の文脈でAIが人間には予測できなかった新たな戦略を「もたらす」ことがある（Silver et al., 2017）⁴⁸。この「出現」は、過去の条件だけからは因果律的に説明できない。それは、システムが自らの力で「未規定なものの実現」を果たした瞬間であり、未来が常に創造的に「開かれている」ことを力強く示唆する。

デジタルツインによる未来予測は、確定的な未来像を示すものではない。これは、現在のデータに基づき構築される蓋然性の高い状態であり、AIのアルゴリズムとモデルによって常に調整される。この調整可能性は、未来が固定された終着点ではなく、関係性の中で共に構成される開かれた領域であることを意味する。AIの予測は、未来を確定するのではなく、未来に対する新たな問いや選択肢を提示し、システムの最適化と効率化

⁴⁸ Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Sifre, A., van den Driessche, G., Chen, Y., Lillicrap, L., Leach, M., Kavukcuoglu, K.,

のための調整を促すために存在する。

この不確定性は、単に私たちの知識が不足していることに起因する認識論的なものではない。多様な要素が非線形かつ持続的に交錯することで、規定できない行為が出現する現象は、未来が単なる過去の反復ではなく、予測を超えた現象によって構成される領域であることを示唆する。したがって、技術と人間の関係は、予測可能な因果律の連鎖を超え、常に未定義の可能性を生成する動的なプロセスとして捉えられると考えられる。

結 論

本稿は、「非人間と人間の関係において、いかなる存在が『行為する (to act)』のか」という技術哲学における基本的な問題設定を起点とする。この問いに対して、従来の媒介概念に内在する理論的限界を明らかにし、それを乗り越える構成として、行為性を単一の存在に帰属させるのではなく、複数の構成要素が交差する関係の構造の中で立ち現れるものとして捉える立場を提示した。

この観点から、媒介理論の再構成を理論的検討と事例分析の両面から試みた。議論の目的は、現代におけるテクノロジー（技術）、とりわけAIとの相互関係における行為性を理解するための分析枠組みを構築することにある。

第一に、媒介概念の再検討にあたり、ラトゥールおよびフェルバークの理論的貢献を参照しつつ、既存の議論では行為の成立条件が構造的に問われていない点に着目した。これに対して、バラッドによる「内在的関係性」の概念を導入し、個別のアクターが事前に与えられるのではなく、関係そのものの内部から構成されることを理論的立場として採用した。これにより、分析の焦点は主体の特定から、行為が生起するための構造的条件の検討へと移行する。

第二に、この立場を具体的な技術的対象に適用するため、AIシミュレーションに関する複数の事例を取り上げた。例えば、機械学習における訓練過程では、データとアルゴリズムが固定された

順序で接続されるのではなく、両者の調整過程において相互に影響しながら構成される。また、AlphaGoに見られる意図しない打手の出現は、設計者の意図とAIの振る舞いが固定的な関係がないことを示している。さらに、デジタルツインにおける物理空間とサイバー空間の対応、環境データとモデル生成との相互依存関係も、行為が単独の系に還元されない構造を例証している。これらの検討を通じて、AIの振る舞いは事前に定義可能な出力ではなく、多様な構成要素の相互作用として記述されるべき対象であることが明らかとなった。

第三に、本稿では時間についての考察を理論構成における不可欠な要素として導入した。行為がある特定の瞬間において完結するものではなく、関係の展開過程の中で形成されるという立場に立つとき、時間は単なる背景条件としてではなく、行為の成立そのものに関与する構成要素となる。とりわけAIの動作過程においては、ナノ秒単位の演算、ユーザーとのリアルタイムの応答、制度的評価との接続といった異なる時間スケールが交差している。このような複数の時間的要素の重なりは、行為の構造を理解するうえで不可避の変数であり、したがって、技術哲学における媒介概念の再構築には時間論的視点の組み込みが理論的に要請される。時間は単線的な因果連鎖としてではなく、重層的で可変的な構造として取り扱われる必要があると考えられる。

以上の検討を通じて、本稿が示す媒介概念の再考は、人間と非人間の関係を、単なる主体と客体の固定的な枠組みで捉えるのではなく、構造的かつ時間的に変化しうる関係として分析する視点への転換であり、この視点によって、行為性の所在に関する従来の理解を超え、関係の構造がどのように変動するかを捉えることが可能となるのである。

参考文献

Barad, K. (2007). Meeting the Universe Halfway:

- Quantum Physics and the Entanglement of Matter and Meaning. Duke University Press. [バラッド, K. (2023). 『宇宙の途上で出会う: 量子物理学からみる物質と意味のもつれ』水田博子, 南菜緒子, 南晃訳, 人文書院]
- Batty, M. (2018). *Inventing Future Cities*. MIT Press.
- Bosch. (2021). "Digital twins in manufacturing: A guide to the key technology of Industry 4.0." Bosch Global.
- Brown, T. B., et al. (2020). "Language Models are Few-Shot Learners." *Advances in Neural Information Processing Systems*, 33, 1877-1901.
- Crawford, K. (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.
- Deng, J., et al. (2009). "ImageNet: A large-scale hierarchical image database." *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248-255.
- Elish, M. C. (2016). "The Moral Crumple Zone: Cautionary Tales in Human-Robot Interaction." *Engaging Science, Technology, and Society*, 2, 40-60.
- Epstein, J. M., & Axtell, R. (1996). *Growing Artificial Societies: Social Science from the Bottom Up*. MIT Press.
- Grieves, M. (2016). *Digital Twin: Manufacturing in the Age of Digitalization*. Gartner.
- Ha, D., & Schmidhuber, J. (2018). "World Models." *arXiv preprint arXiv:1803.10122*.
- Hafner, D., et al. (2019). "Dream to Control: Learning Behaviors by Latent Imagination." *International Conference on Learning Representations (ICLR)*.
- Hafner, D., et al. (2020). "Mastering Atari with Discrete World Models." *International Conference on Learning Representations (ICLR)*.
- Haugeland, J. (1985). *Artificial Intelligence: The Very Idea*. MIT Press.
- Heidegger, M. (1977). *The Question Concerning Technology and Other Essays*. Harper Perennial. [ハイデガー, M. (2013). 『技術への問い』 関口浩訳, 平凡社ライブラリー]
- Kitchin, R. (2014). *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. Sage Publications.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). "ImageNet Classification with Deep Convolutional Neural Networks." *Advances in Neural Information Processing Systems*, 25.
- Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Harvard University Press. [ラトゥール, B. (2001). 『科学が作られているとき——人類学的考察』 川崎勝, 高田紀代子訳, 産業図書]
- Latour, B. (1988). *The Pasteurization of France*. Harvard University Press.
- Latour, B. (1992). "Where Are the Missing Masses?" In W. Bijker & J. Law (Eds.), *Shaping Technology/Building Society* (pp. 225-258). MIT Press.
- Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford University Press. [ラトゥール, B. (2019). 『社会的なものを組み直す——アクターネットワーク理論入門』 伊藤嘉高訳, 法政大学出版社]
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). "Deep learning." *Nature*, 521 (7553), 436-444.
- McCorduck, P. (2004). *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence*. CRC Press. [マコーダック, P. (1983). 『コンピュー

- タは考える—人工知能の歴史と展望』黒川利明訳, 培風館 [原著1984年]
- Metz, C. (2016, March 12). “How Google’s AI Viewed the Move No Human Could Understand.” *Wired*.
- Mittelstadt, B. D., et al. (2016). “The ethics of algorithms: Mapping the debate.” *Big Data & Society*, 3 (2), 1-21.
- Newell, A., & Simon, H. A. (1976). “Computer Science as Empirical Inquiry: Symbols and Search.” *Communications of the ACM*, 19 (3), 113-126.
- O’Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown. [オニール, C. (2018).『あなたを支配し, 社会を破壊する, AI・ビッグデータの罠』久保尚子訳, インターシフト]
- Prigogine, I., & Stengers, I. (1984). *Order out of Chaos: Man’s New Dialogue with Nature*. Bantam Books. [プリゴジン, I., & スタンジェール, I. (1987).『混沌からの秩序』伏見康治, 伏見謙, 松枝秀明訳, みすず書房]
- Saxe, A. M., McClelland, J. L., & Ganguli, S. (2019). “A mathematical theory of semantic development in deep neural networks.” *Proceedings of the National Academy of Sciences*, 116 (23), 11537-11546.
- Silver, D., et al. (2016). “Mastering the game of Go with deep neural networks and tree search.” *Nature*, 529 (7587), 484-489.
- Silver, D., et al. (2017). “Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm.” *arXiv preprint arXiv:1712.01815*.
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. MIT Press. [サットン, R. S., & バート, A. G. (2022). 強化学習 (第2版)(奥村エルネスト純・鈴木雅大・松尾豊ほか監訳). 森北出版]
- Tao, F., & Qi, Q. (2019). “Make More Digital Twins.” *Nature Machine Intelligence*, 1 (1), 1-2.
- Turing, A. M. (1950). “Computing Machinery and Intelligence.” *Mind*, 59 (236), 433-460.
- Vaswani, A., et al. (2017). “Attention is all you need.” *Advances in Neural Information Processing Systems*, 30.
- Verbeek, P.-P. (2005). *What Things Do: Philosophical Reflections on Technology, Agency, and Human-Technology Relations*. Pennsylvania State University Press.
- Verbeek, P.-P. (2006). “Materializing Morality.” *Science, Technology & Human Values*, 31 (3), 361-380.
- Verbeek, P.-P. (2011). *Moralizing Technology: Understanding and Designing the Morality of Things*. University of Chicago Press.
- Waymo. (2021). *Waymo Safety Report*.
- Zoph, B., & Le, Q. V. (2016). “Neural architecture search with reinforcement learning.” *arXiv preprint arXiv:1611.01578*.
- 柿崎慎也 (2025). 「代理による科学的対象構築プロセスの解明——ブリュノ・ラトゥールにおける実在論的アプローチに関する考察」『東北哲学会年報』, (41), 1-20.
- 柿崎慎也 (2024). 「B. Latour における人間及び非人間への代理 (delegation) 概念の検討」『山形大学人文社会科学部研究年報』第二十一号, 119-132

Reconsidering Agency in Technological Mediation

KAKIZAKI Shinya

This paper addresses the philosophical question of agency—that is, who or what acts—in the context of technological mediation. It focuses particularly on the relationship between humans and nonhumans, such as artifacts and AI. Traditionally, technology was viewed as a passive tool subordinate to human intention. However, Actor-Network Theory (ANT) by Bruno Latour and the phenomenological approach of Peter-Paul Verbeek have expanded the concept of agency to include nonhumans as active participants that shape social order and human action.

ANT, as developed by Latour, argues that any entity capable of making a difference—human or nonhuman—should be considered an actor. Verbeek builds on this idea, showing how artifacts mediate human perception, action, and ethics. He introduces the concept of “reverse delegation,” in which artifacts influence human behavior and responsibility. Both theorists challenge the traditional view that agency is exclusive to humans.

This paper critiques these theories for implicitly attributing the origin of action to either humans or nonhumans. Drawing on Karen Barad’s concept of intra-action,” the author proposes a shift: Agency does not preexist in individuals, but rather emerges relationally as a dynamic event within networks of humans, artifacts, data, and environments.

The paper illustrates this concept by analyzing AI simulation technologies, in which agency is distributed among designers, algorithms, data, users, and environments. For instance, in AI training and digital twin systems, agency and meaning emerge from the interaction of all elements rather than from any single actor.

The study concludes that agency should be understood as an emergent property of relational configurations rather than as an attribute of isolated actors. This relational perspective provides a new framework for analyzing complex technological practices such as AI, in which meaning and responsibility are produced by networks of human and nonhuman elements.

