

アクターネットワーク理論の「人工主体」研究への適用
--画像生成 AI の分析を例に

Application of actor-network theory to the study of "artificial agents"
--Analyzing AI image generators as an example

池原優斗

Abstract

This research proposes Actor-Network Theory (ANT) as a methodology for organizing and analyzing problems involving "artificial agents," which are artifacts that induce humans to experience them as a certain kind of agent, such as AI and robots. ANT makes it possible to handle the boundaries of the categories of human, "artificial agent" and "other artifacts" in a dynamic and flexible way. This enables effective organization and analysis of problems involving "artificial agents." After presenting an overview and discussion of ANT, this paper will present how ANT can be used concretely to describe and analyze problems involving "artificial agents" through an analysis that applies ANT to the problem of AI image generators.

(1) 研究テーマ

現在、AI やロボットの社会実装が進んでいる。それと同時に AI やロボットという新しい存在の登場は、例えば、AI が法的責任を持つことができるかというような、人間とモノの境界を揺るがす新たな問題を生み出している。本論では、AI やロボットのような人間がそれをある種の主体として経験することを誘発するような人工物を「人工主体」と呼ぶ¹²。「人工主体」は AI やロボットに限定されるものではなく、定義を満たすのであれば、人工生命や自動運転車など様々な人工物を含めることができる。本研究では、「人工主体」および、それが関わる問題について記述・分析する方法論としてアクターネットワーク理論 (ANT) を提示し、それをを用いた「人工主体」が関わる問題の記述・分析を行う。

(2) 研究の背景・先行研究

ロボット倫理学の分野では、ロボットを主体として扱うべきか、道具として扱うべきか、という議論が行われている。例えば、久木田 [2020 : 18] は

ロボット倫理学の課題として、次のような問題を挙げている。ロボットが人に危害を与えた際の責任の帰属はどのようにすればよいのか、人間にそっくりで人間のように振る舞うロボットに人格や人権は認められるべきか、ロボットが戦場で戦うとき人を殺傷する権限を与えるべきか、ロボットが人間の仕事を奪うようになったらどのように対応すべきか、といった事柄である。こうした論点はどれもロボットという人工物を人間と同じような主体として扱うべきなのか、という問題に行き着くものである。また、呉羽 [2021: 63] もロボット倫理学におけるロボットを道具以上のものとして扱うべきかという問題を踏まえて、社会ごとの文化による差についてどのような態度で考えていくべきかという観点からの議論を展開している。

人工物の主体性に関する問題は、ロボットだけでなく、人工知能 (AI) についても同様に存在する。例えば、2022年6月、グーグルのAI開発部門の社員が、グーグルの自然言語生成AIである「LaMDA」について、意識や感覚を持っているとし、グーグルに対して実験に使用する前に同意を求めべきであると主張した出来事があり、議論を呼んだ [ヴァランス 2022]。人間がAIに対して主体性を感じるような現象は、自然言語生成AIだけではなく画像生成AIでも報告されている。Daras & Dimakis [2022] は、画像生成AIのDALI-2は生成した画像に登場した意味をなさない文字列が、AIにとって特定の対象を指示する単語である可能性を指摘している。Daras & Dimakis [2022] で報告された現象はAIが意識を持っている根拠とはなり得ないが、人間が画像生成AIのこのような振る舞いから主体性を経験することは十分にあり得るだろう。

このような、人間がそれをある種の主体として経験することを誘発するような人工物全てを含む概念が、「人工主体」である。「人工主体」という概念を用いることで、AIやロボットなどの私たちが主体性を経験する現代の技術的存在について、統一的に議論していくことができるようになる。「人工主体」の概念は、直観的であり、議論のきっかけとする上で有効なものだと言えるだろう。

しかし、「人工主体」についての分析を進めようとする、何が「人工主体」なのかを定める境界線が曖昧で、問題の整理や分析が困難になるという問題がある。「人工主体」という概念を導入した際、人間か人工物として表現されていた存在は、人間 / 「人工主体」 / 「その他の人工物」に分けられることになる。このとき、どこからが「人工主体」でどこからが「その他の人工物」であるかをはっきりと決めることは難しい。さらに、AI等の技術がより発達し社会実装が進んでいけば、人間と「人工主体」の境界線も徐々に曖昧なも

のようになっていこう。本論は、この問題の解決を試みるものである。

(3) 筆者の主張

3-1 アクターネットワーク理論

ANT はブルーノ・ラトゥール [ラトゥール 2019] やミシェル・カロン、ジョン・ロー [カロン・ロー 1999] らによって発展した理論である。ラトゥール [2019: 24-6] による、ある研究が ANT であるかの暫定的基準を基に、概要を端的に示すならば、第一に、「非人間にはっきりとした役割が与えられている」こと、第二に、社会的なものを最初から安定させたまま、社会的なものによって、ある物事を説明しないこと、第三に、脱構築ではなく、「社会的なものを組み直すことを目指している」ことが ANT の特徴と言える。また、存在物を異種混合の集団の結果として捉え、またその行為を集団的な性質であるとする点も ANT の特徴である [カロン・ロー 1999: 249-54]。

本論では ANT の議論のうち、特に次の二つを取り上げたい。一つは、人間だけでなく、生物やモノである非人間もアクターとして捉えるという議論。つまり、ANT は人間と非人間を対称的に扱うのである。もう一つは、エージェンシーについての議論である。ANT においてエージェンシーは、人間および非人間であるアクターの中にあるというよりも、アクターが織りなす布置から生まれてくる [青山 2008: 177]。したがって、ANT は分析の対象となる事柄について、それを人間と非人間が含まれるアクターが関係し合っているネットワークとして扱い、そして、エージェンシーはそのネットワークの布置連関から生まれてくると捉えるのである。

ANT は事象を記述し、分析することに長けたフレームワークであるが、ある対象に対して批判的態度を持ち続けるにはどのようにすればよいのか、といった規範的な議論への応用も検討されている。以下で取り上げる土橋 [2015] は、ケータイ的な情報環境における「批判」を ANT によって考察したものである。土橋 [2015] は、ケータイという可動的で可変的なメディアと人間はどのように関係を取り結び、人間はそうした技術に対して批判的な姿勢をどのようにすれば持つことができるのか、という問題に対して、ANT の視点から分析している。

土橋 [2015: 19-22] は、可動的かつ可変的なケータイが常にそばにある情報環境を、デスクトップパソコンのような固定的な空間に設置されるメディアによる情報環境と対置する。そして、アドホックなニーズに柔軟に応じるため利用者の生活スタイルや価値観と摩擦を起こさずに利用者の文脈に適合するケータイを、人間と対峙する固定的メディアと異なり、人間と協働する

メディアであるとする [土橋 2015: 21-2]。ANT を用いて人間とケータイの協働を捉えれば、人間とケータイはハイブリッド・コレクティブであり、エージェントは人間とケータイと空間の関係的な布置連関から、関係的な効果として立ち上がってくるのだとする [土橋 2015: 22-4]。

ただ、土橋はケータイとの関係が協働であったとしても、ケータイ的な情報環境への批判的な構えを持ちづける必要があるとする [土橋 2015: 31]。しかし、ユビキタス化とアプリケーション化によって特徴付けられるケータイ的な情報環境においては、技術に対する批判的構えを持つための距離が存在せず、また、その都度異なるエージェントが立ち現れてくるため、一貫した論理によって批判するための根拠が存在しない [土橋 2015: 30-1]。土橋 [2015: 32-3] は批判のための「距離」と「根拠」の不在に対して、外部ではなく、内部から目の前にあるケータイをめぐる布置連関に思考や実践を「付加」し、その布置連関を組み替えていくような批判のあり方を提案している。

土橋 [2015] が扱ったケータイ的な情報環境において起きている問題は、正にケータイ（スマートフォン）を含む移動するコンピュータのアプリケーション上に現れ、ユビキタス的に存在する「人工主体」にも共通するものだろう。更に言えば、土橋 [2015] において問題とされているのは、情報技術が否応なく実装され浸透される世界で、どのように批判的態度を持ち、批判的な受容が可能であるかということであり、より広範な問題に対して、ネットワークの内部から思考や実践を付加するという批判のあり方は有効であるかもしれない。いずれにせよ、内部から布置連関を組み替えることによる批判に着目することは、ANT の記述的分析を「人工主体」の具体的問題に応用し、実際性のある議論を展開するための一つの有効な方法となりうる。

本論では、人間と非人間を対称的に扱う ANT を、「人工主体」が関わる問題の整理・分析のための方法論として提案する。つまり、本論の主張は、ANT を用いることで、人間／「人工主体」／「その他の人工物」という区分そのものを問い直す形で、境界線の曖昧さという問題を乗り越え「人工主体」が関わる問題を整理し分析することが可能となるということである。

次の節では、ANT を用いて画像生成 AI について記述・分析する。この画像生成 AI についての考察では、主に「人工主体」と「その他の人工物」の境界の曖昧さに焦点を当てる。また、土橋 [2015] が扱った、情報技術が否応なく実装され浸透する世界で、どのように批判的態度と受容が可能であるかという観点からも ANT を用いて記述・分析する。

3-2 画像生成 AI について ANT を用いて記述・分析する

2022 年は複数の影響力がある画像生成 AI が公開された。2022 年 4 月 6 日に DALL-E-2 が発表された [OpenAI n.d.] のを皮切りに、7 月 12 日に Midjourney のオープンベータが開始された [@midjourney 2022]。そして、8 月 22 日に Stable Diffusion が一般に公開された [Stability AI 2022]。画像生成 AI は、従来は人間にしかできないと思われていた創造的な絵を描くことができる。そのような画像生成 AI の登場は作家や、創造性を人間が犯されることのない領域だと考えていた人々に衝撃を与え、イラストレーターの職が AI によって奪われるのではないかという不安を抱かせた。Stable Diffusion を開発した Stability AI の CEO がキャリアを心配する若いアーティストに対して、「イラストやデザインの仕事は、とても退屈な仕事だ。芸術的かどうかではなく、あなたは道具なのだ」、「画像生成 AI は今後大きく発展していく。お金を稼ぎたいなら、新しい技術からチャンスを見つける方がずっと楽しくなるはずだ」というメッセージを発して物議を醸した [松浦 2022]。また、日本においては、画像生成 AI サービスの mimic について、一部の作家から悪意のある第三者が勝手に画像を使用することを止められないという反発があり、サービスを一時停止したという出来事があった [高木 2022]。

ANT の枠組みでこの状況を考えるならば、2022 年現在は、画像生成 AI という新しいアクターの登場によって、従来のイラストを書くことに関するネットワークが組み替えられようとしている最中であると言える。また、自由にパソコンやスマートフォンから使える画像生成 AI は、アプリケーション的かつユビキタス的であり、上に述べた「人工主体」および土橋 [2015] で論じられている「ケータイ」の特徴と同様である。したがって、画像生成 AI も批判のための「距離」と「根拠」が不在であるという問題を抱えていると言えよう。

画像生成 AI の登場は、イラストを描くことに関する既に安定したネットワークに存在していたイラストレーターにとって、不安を抱かせ、著作権や職種そのものに対する脅威として存在するようになった。このとき、画像生成 AI は正にイラストレーターに脅威をもたらす「人工主体」として存在していたと言えるだろう。

しかし、一部のアクターは画像生成 AI と「対峙」して、拒絶するのではなく、別の関係のとり結び方を模索し始めた。例えば、Midjourney を用いたイラスト背景の作成方法を複数の作家が研究していたことが記録されている [@golden_haniwa 2022]。また、画像生成 AI を用いて漫画を書く実践は池田・柳澤 [2022] の例など複数行われている。そうした漫画の制作者が、熟

練した作家が絵を描くようには自由に描けないという苦勞を語っているケースもあり、このような実践を通して、徐々に画像生成 AI が差し迫った脅威にはならないことが明らかになっていった。

このような「協働」は、めまぐるしく変化する画像生成 AI が含まれるイラストを描くことに関するネットワークのただ中で行われた、「批判的実践」である。そしてこの、「批判的実践」の「付加」は画像生成 AI に関する布置連関を組み替え、画像生成 AI という存在をイラストレーターの職に関する差し迫った脅威から、創作活動での活用の可能性があるツールのような存在に変化させた。これは土橋 [2015] が提案した、ネットワークの内部から批判を付加する実践の一つであると言えるだろう。

さらに、この画像生成 AI という存在の変化は、「人工主体」と「その他の人工物」の区別を固定したものとして捉えない ANT の視座によって描きやすくなった変化である。画像生成 AI は、当初はイラストレーターにとって、「すべてのイラストレーターの仕事を奪っていく万能なイラストレーター」というような「人工主体」としての存在であった。しかし、それが徐々にデジタルイラストソフトのツールのような「その他の人工物」に近い存在へと変化していった。

3-3 流動的な「人工主体」の存在のしかたを捉える

上の例では、画像生成 AI が「人工主体」として存在するかは流動的であり、それらが存在するネットワークによって存在のしかたが変化した。これは人間／「人工主体」／「その他の人工物」という区分で対象を固定的に捉えていては記述することのできないものである。この「人工主体」の流動的な存在のしかたを捉えられることこそが ANT を方法論として導入する利点であると考えられる。

(4) 今後の展望

本研究の課題について二つ述べておきたい。まず、本論では ANT の中でも人間と非人間を対称的に扱うことと、エージェンシーに関する議論を中心に用いて記述・分析を行った。しかし、ANT にはより多くの理論的な要素があり、それらを十分に活用できているとは言いがたい。したがって、ANT の持つ各理論的要素を俯瞰し、総合的に「人工主体」について考察していくことが課題となる。もう一つは、人間は技術に対して批判的であるべきかどうかそれ自体を ANT によって説明する必要があるかを検討し、そうであるならば、それを ANT で説明することである。本論では、土橋 [2015] で提示

された技術に対して「批判的な構え」を持つべきであるという規範を、前提として議論した。しかし、よりラディカルに考えるならば、なぜ人間というアクターが、別の特定のアクターである技術（を活用したメディア等）に対して、「批判的な構え」を持つ必要があるのかという問題が生じる。「批判的な構え」を持つ必要性自体も、アクターのネットワークの関係性の中から説明される必要があるかもしれない。この点についてより研究を進めることも本研究の持つ課題の一つである。

注

1. 「人工主体」についての研究はまだ始まったばかりだが、既に複数のプロジェクトが実施されている。2020年度には戦略的創造研究推進事業のプロジェクト企画調査として「人工主体の創出に伴う倫理的諸問題を分析・討議するプラットフォームの構築に向けた企画調査」が行われた[科学技術振興機構 2022]。また、2022年度からはトヨタ財団の助成対象プロジェクト「人間と人工主体の共存のあるべき姿を学際的に問うための新たな枠組み『人工主体学』の構築に向けて」が開始されている[トヨタ財団 日付不詳]。
2. 本論の「人工主体」の定義は、科学基礎論学会 2022年度秋の研究例会ワークショップ「AIに主体性を帰属させること：学際的アプローチの中間報告」の要旨である宮原[2022]における定義を参考にした。

(5) 参考文献

- 青山征彦 2008「アクターネットワーク理論が可視/不可視にするもの--エージェンシーをめぐって」『駿河台大学論叢』35: 175-185.
- 池田佳恋・柳澤あゆみ 2022「AIが絵を描く？ 進化する画像生成 AIの最前線」『NHK』10月8日 <https://www3.nhk.or.jp/news/html/20221008/k10013851401000.html> 2022年11月4日閲覧.
- ヴァランス, C. 2022「AIに『感情がある可能性』 グーグルのエンジニアが主張」『BBC ニュース』6月14日 <https://www.bbc.com/japanese/61793898> 2022年12月8日閲覧.
- 科学技術振興機構 2022「JSTプロジェクトデータベース — 人工主体の創出に伴う倫理的諸問題を分析・討議するプラットフォームの構築に向けた企画調査」<https://projectdb.jst.go.jp/grant/JST-PROJECT-20343952/> 2022年12月11日閲覧.
- カロン, M. & J. ロー 1999 (1997)「個と社会を超えて--集団性についての

- 科学技術社会論からの視座」林隆之訳 岡田猛ほか編『科学を考える--人工知能からカルチュラル・スタディーズまで 14 の視点』北大路書房 pp. 238-257.
- 久木田水生 2020「ロボットの倫理--友達ロボットから殺人ロボットまで」『日本ロボット学会誌』38(1): 18-22.
- 呉羽真 2021「日本人とロボット--テクノアニミズム論への批判」『Contemporary and Applied Philosophy』13: 62-82.
- ゴールデンハニワ@Eve ghost enemies 発売! (@golden_haniwa) Together 投稿 2022年8月14日「キャラを自分で描ける人によって『Midjourney』にマンガやイラストの背景を描かせる研究が着々と進んでいる模様」<https://together.com/li/1930520> 2022年11月4日閲覧.
- 高木克聡 2022「イラスト作画AI、作家の敵か味方か 画風学習し新たな絵 盗用を恐れ炎上」『産経新聞』9月13日 <https://www.sankei.com/article/20220913-64SABJB6PBPLPFPPRKNKD2GJ64/> 2022年10月9日閲覧.
- 土橋臣吾 2015「移動するモノ, 設計される経験--ケータイの可動性と可変性をめぐって」『マス・コミュニケーション研究』87: 17-35.
- トヨタ財団 日付不詳「助成対象詳細 | 公益財団法人トヨタ財団」<https://toyotafound.secure.force.com/psearch/JoseiDetail?name=D21-ST-0012> 2022年12月10日閲覧.
- 松浦立樹 2022「「イラストやデザインの仕事は退屈」--Stable Diffusion 開発元の代表インタビュー記事が話題」<https://www.itmedia.co.jp/news/articles/2210/26/news183.html> 2022年11月4日閲覧.
- 宮原克典 2022「AIに主体性を帰属させること--学際的アプローチの中間報告」https://phsc.jp/dat/rsm/20221020_01.pdf 2022年12月10日閲覧.
- ラトゥール, B. 2019 (2005)『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 法政大学出版局.
- Daras, G. & A. G. Dimaki. 2022. Discovering the Hidden Vocabulary of DALLÉ-2. arXiv preprint arXiv:2206.00169.
- Midjourney(@midjourney) Twitter 投稿 2022年7月13日 午後3:41 <https://twitter.com/midjourney/status/1547108864788553729> 2022年11月4日閲覧.
- OpenAI. n.d. Timeline. <https://openai.com/timeline/> 2022年11月4日閲覧.
- Stability AI. 2022. Stable Diffusion Public Release. <https://stability.ai>

[/blog/stable-diffusion-public-release](#) 2022 年 11 月 4 日 閲覧.

(北海道大学)