

## 生成AI時代における哲学・倫理

小泉 英明  
日立製作所 名誉フェロー

## はじめに

現代は気候変動による自然環境の変化や人口の高齢化に伴う社会経済システムへの圧力といった前例のない社会問題に直面している。人類はこれらの課題を解決し、同時に持続可能な成長の道を探求しなければならない状況にある。このような状況下で誕生した生成 AI を含む新しい技術は、Society 5.0 といった未来社会のビジョンを具現化し、これらの課題への革新的な解決策を提供する可能性を秘めている。この重要な時期に、科学技術および社会科学の研究の進め方やその方向性について洞察を得るため、脳科学から新・人間学など新しい学術分野の発展に寄与してきた日立製作所の小泉英明名誉フェローにインタビューを行った（聞き手：日立総合計画研究所会長 鈴木教洋）。本稿は、小泉名誉フェローのお話を論文形式で編集したものである（編集：同研究第三部部长 鹿野健一）。

## 1. 哲学・倫理が重要になる背景

科学技術の進展はデカルトによって定式化された還元論に立脚している。この還元論は、全てをより基本的な要素へと分解し理解するという考え方で、その起源はデカルト以前のギリシャ時代の原子論にまでさかのぼることができる。しかし、この方法だけで全てを解明できるわけではない。デカルト自身も、要素還元の後には、降りて来た階段を今度は昇っていくように、統合の過程が必須であると述べている。コンピュータなどの2進体系を基調とする情報処理や情報伝達は確かにデジタル的なオンオフの機能（ビット）を基にしているが、脳神経の働きや人間の肉体的な体験を深く掘り下げてみると、その複雑さは単純にデジタルだとかアナログだとかに分類できないレベルに達する。人間はアナログといわれ

るが、神経系の情報処理は、閾値（しきいち）によってオンオフするシナプス（神経細胞間接続部）のデジタル情報伝達が基本である。これは、人間もデジタルな存在として還元してしまえば、AI と変わらなくなってしまうという点を浮き彫りにする。

さらに新しい発見やイノベーションは、物事を分解するのではなく、異なる分野の知識や技術を組み合わせること、すなわち俯瞰（ふかん）統合によって生まれる。これは、シュンペーターの経済発展理論における創造的破壊の概念とも通じる。このため、イノベーションを起こすために重要なのは、ロードマップ的な思考ではなくて非連続な遷移である。人間は、意識上で行っている情報処理よりはるかに大量の並列分散処理を意識下で行っている。それが意識に上った瞬間に、イノベーションの遷移が起こる。なぜなら神経による情報伝達速度はたかだか 200m/s であって、光速に近い電子による情報伝達とは比較にならないほど遅い。神経系が超並列分散処理を必須とするゆえんである。したがって、脳内の情報処理の最終段階である逐次処理に入って初めて意識に上るのである。意識に上るのは脳内情報処理の上澄みである。

生成 AI を含む最先端の科学技術に対する理解と応用のあり方を、哲学的な思考を通じて深く考え直さねばならない。科学技術の進展が人間や福祉にどのように貢献できるかを考える際、還元論を超えた、哲学に根差した新しいアプローチが不可欠である。

科学(Science)、工学(Engineering)、技術(Technology)という三つの領域があるが、これらの概念は互いに補完し合い、新たな発見やアプローチを生み出す土壌となる。特に工学の領域では、科学的知識を応用し人間によって自然にないものを創造する過程が含まれるが、この過程で生じる人工物や技術が自然界に与える影響については、倫理的な考慮が必要とされる。例えば、遺伝子組み換え技術のような科学的

発見に基づく技術の実現は、その応用において多くの倫理的問題を引き起こす。これらの技術が社会に及ぼす影響を考える際、科学の探求だけでなく、それを応用する工学的側面における倫理的な課題への注意が求められる。

本稿の主題である生成 AI に対する哲学と倫理の関心は、人間が自然界に本来存在しない技術を創出し、その技術がさらに人間のように振る舞い、新たなものを生み出すという事実由来する。この複雑な関係を適切に理解し、利用していくためには、技術背景とその社会的影響の両方を考慮するための哲学と倫理への研究が必要である。地球上に生命が誕生したのがおよそ 38 億年前であることは大切な事実である。やがて大量発生したシアノバクテリア（藍藻（らんそう））が酸素を含む大気を生んだ。進化の長い歴史の中で、現生人類（唯一のヒト属現存種である亜種ホモ・サピエンス）はまだ 20 万年程度のわずかな歴史しかなく進化の途上にあるが、言語や抽象的思考を獲得して、高度な文明を異常なスピードで築いた。従って、最も近い種のチンパンジーと比較して、前頭前野以外の基本的な脳構造は同じであり、長い進化の歴史をそのまま宿している。現生人類の脳は、理性が野性をかろうじて抑えているのが実態と解釈すべきであり、繰り返される戦争の理由はそこにある。暴力行為によって、いったん、野性へのスイッチが入ってしまうと野生動物と同じレベルになって、人間の倫理は吹き飛ぶ場合が多い。脳神経科学の視座からは、いったん始まった戦争を止めるのは極めて難しく、戦争が始まらないことに最大限の努力を集中すべきことが見えてくる。今後は脳神経科学に依拠した合意形成の科学が必要となる<sup>1</sup>。

また、脳は進化途上であって、現実と仮想現実を区別する術（すべ）を持たない。記憶の連鎖をたぐって夢か現（うつつ）かを判断するしかない。しかも

<sup>1</sup> 例外的なケースとして、戦争を一時的に止められた 1993 年のオスロ合意がある。紛争の中でも最も解決が難しいとされるイスラエル/パレスチナ問題に、この合意が実現した背景には、北欧・英国を中心とした学術組織（Common Security Forum）の現地調査があった。まさに「現場」の「一次情報」を、歴史を含めて再収集したことが契機となった（オスロ合意の直後に開催された国際会議で、交渉に関わった人々から筆者は直接話を聴く機会があった）。

生存確率を上げるために存在する「快・不快」という評価指標は、人間もそれ以外の動物と基本は変わらない。生存確率を上げる「快」の方向へと、生存に有利な行動を本能的に繰り返す。この「快」を追求する本来の姿が、人工的な世界でゆがんでくると、広範な依存症（薬物、ギャンブル、ビデオゲーム）の問題が深刻さを増してくる。経済ですら欲望によって倫理限界を超えた格差が生まれて、効果的に改善できる術を持たなくなる。

## 2. 脳科学における言語と生成 AI

生成 AI には、言語処理だけでなく、画像処理なども扱うマルチモーダルモデルが出現しているが、やはり現状では大量言語モデル（LLM: Large Language Model）に代表される言語処理に特筆すべき点がある。そこで、人間の脳における言語と生成 AI が作り出す言語の違いについて考察してみる。

ここで重要なのは、言語が単に文法的に正確な構造を持つだけでなく、意味論的にナンセンスな表現も可能にすることである。この点について、現代言語学の父といわれる Noam Chomsky 博士は、かつて「Colorless green ideas sleep furiously」（色のない緑の考えが猛烈に眠る）という例を挙げ、文法的に正しくとも意味をなさない文章を示した。この言語の特性は、人間が事実と離れた抽象的な概念も扱える能力を示唆している。

また、ネアンデルタールがホモ・サピエンスと混血していることを明らかにした（2022 年度ノーベル生理学・医学賞）化石人類の DNA を研究する遺伝学者の Svante Pääbo 博士は、ホモ・サピエンスが言語に関わる特定の遺伝子を持っており、これがネアンデルタールとの主な違いの一つであることも明らかにしつつある。現生人類が初めて抽象的な概念を扱う言語というツールを持ったということが、これらの研究によって示唆されている。

言語によって、人類は未来という抽象的な思考ができるようになった。例えば、「三カ月後」という時間の概念を、マイム（身ぶりや手ぶり）で表現することは極めて困難であるが、言語を用いれば瞬時に表現することができ、言語や記号が時間概念や未来

を扱う上で優位であることは明らかである。さらに言語を用いれば、誰一人として経験していない死後の世界や、極楽・地獄などの抽象概念を構築することも可能である。

また、こうした未来を概念的に考えることにより、人類は10年後に得られる報酬を今得ているかの如く快感に浸ることもできるようになった。このような言語による未来の概念化は人間特有の現象であり、現代科学においても完全には解明されていない。生成AIの言語モデルは、人間の言語を用いた複雑な思考や創造的な表現をある程度まで模倣する能力を持つが、未来に関する抽象的な概念化や、それに伴う精神的な快感を完全に再現することは、現在のところ人間にしかできない。しかし、このあたりもこれからの研究次第かもしれない。

### 3. 生成AI時代における一次情報の重要性

AI技術、特に生成AIの進展は今日、人間が書いたものと見分けがつかないほどの文章を生成する能力に到達し、人間の能力を上回る場面もある。しかし、この進歩にはフェイクニュースの拡散やハルシネーション（誤情報：原義は「幻覚」）といった新たな課題も伴う。これらの問題は社会における情報の信頼性を損ねるだけでなく、社会インフラなどに深刻な影響を及ぼすことが懸念されている。

言語を通じて伝えられる情報が複雑に合成されたものである場合、人間は、論理的に正しいかどうかでその真実性を見極めてきた。しかし、人間の言語を模倣した生成AIの高度な情報生成により、人間が論理的にその真偽をひもとくことが不可能になる。このため複合される前の一次情報にアクセスすることの重要性を強調したい。一次情報は情報が生成された最初の点であり、最も信頼性が高い。それは歴史的にも組織中枢からの一次情報が最も貴重とされてきたゆえんでもある。

生成AIの進化は医療分野においても顕著な成果を示している。例えば、疾患の早期発見や治療法の選定において、大量のデータからパターンを識別し、医師の診断をサポートする役割を担っている。これ

により、より正確で迅速な医療サービスの提供が可能になっている。しかしこのような有用性を認めつつも、患者の健康と安全を確保するために、医療情報における一次情報の重要性は依然として高く、診断や治療計画の根拠となるデータの正確性と信頼性が求められる。

ここで日立製作所を例にとると、創業以来、一次情報の重要性和それに基づいた倫理的な情報取り扱いが企業文化の核となっている。創業者である小平浪平や久原房之助は、技術開発と事業展開において、現場で得られる直接的なデータや情報を重視した。この「現場主義」は、正確な一次情報に基づく意思決定を促し、企業の信頼性と持続可能性の確保に寄与してきた。一次情報への深い理解とそれを尊重する文化の構築により、長期的な視点での企業価値を高めてきたことを示している。

今後の生成AI時代を見据えたとき、一次情報へのアクセスとその活用は、情報の信頼性を確保するために一層の重要性を帯びてくる。AIによる情報生成が進む中で、真実と誤情報を見分け、信頼できる知識の構築には、原点に立ち返ることが必要である。一次情報への深い理解と適切な活用が、私たちを誤情報から守り、健全な情報社会の基盤を築く鍵となると考える。

### 4. 人間社会的倫理観から自然的倫理観へ

倫理に関するこれまでの学術界の議論は、アリストテレスやプラトンといった古典哲学にその起源を持ち、人間の習俗や慣習に深く根差した倫理観を探究してきた。しかし、プラネタリーバウンダリ（人類が地球上で持続的に生存していくために超えてはならない環境限界）という考えが定着してきたことで、現代の倫理学は自然科学の知見を踏まえた、より広範な視野からのアプローチを必要としている。

地球上の生命は、非常に薄い大気圏内のバランスに依存しており、このデリケートなバランスが崩れれば、全ての生命に深刻な影響を及ぼす可能性がある。人間だけでなく地球上の全生命が相互依存しているという現実を基に、地球上の全ての生命体とそ

の生存環境を尊重する普遍的な倫理観の形成が必要となる。

デジタル時代、特に生成 AI の時代に突入し、人類の技術が未曾有の速度で進化している現在、人類が地球環境に与える影響は、これまで以上に大きくなる。これまでの人間中心の倫理から、自然界とのより調和した関係を基盤とする科学的知見に基づいた、より包括的な倫理観が必要とされている。

#### 執筆者紹介



小泉 英明（こいずみ ひであき）  
日立製作所名誉フェロー、日本工学アカデミー顧問（前上級副会長）  
1971年東京大学教養学部基礎科学科卒業、同年日立製作所計測器事業部入社。1976年理学部に論文を提出し、東京大学理学博士。2000年基礎研究所所長、2003年技師長、2004年フェローを経て、2017年より名誉フェロー

「心と脳の科学」という新たな transdisciplinary 分野を提起し、道を開いた研究者として世界中に知られ、偏光ゼーマン原子吸光法の創出・実用化による環境計測をはじめに、f-MRI・光トポグラフィーによる脳機能計測技術を通じて脳科学から新・人間学など新しい学術分野の発展に寄与。東京大学先端科学技術研究センター フェロー・ボードメンバー、中国工程院外国籍院士・東南大学名誉教授。国際工学アカデミー連合（CAETS）理事、米国・欧州・豪州などの各種研究機関や財団のボードを歴任。著書に『アインシュタインの逆オメガ：脳の進化から教育を考える（Evolutionary Pedagogy）』（パピルス賞受賞作品、文藝春秋社刊）