

AIの時代によく生きるとは
人間の再定義について

中島隆博
東京大学

Chat GPT による「AI と倫理」への回答

AI（人工知能）と倫理について論じることは非常に重要です。AI 技術の発展は、私たちの社会、経済、文化に大きな影響を与えており、その影響を適切に調整し、倫理的な枠組みを築くことが不可欠です。以下に、AI と倫理についての主要な観点をいくつか挙げてみましょう。

1. プライバシーとデータ保護: AI は膨大なデータを必要とします。個人のデータ保護は、AI を使用する際に最も重要な倫理的な問題の一つです。個人情報収集、保存、利用に関する厳格な規制と、そのデータが不正にアクセスされたり悪用されないようにするためのセキュリティ対策が必要です。
2. バイアスとフェアネス: AI は訓練データに含まれるバイアスを学び取ることがあります。これにより、性別、人種、年齢などの要因に基づいて不平等な結果が生じることがあります。AI システムの設計段階から公平性を確保し、アルゴリズムのトランスペアレンシーと説明可能性を向上させることが重要です。
3. 効用と責任: AI を使用する際、その目的と効果について検討する必要があります。AI が不正行為や悪意ある活動に使用されないように、開発者、運用者、政府などが責任を持つべきです。透明性と監査可能性の原則は、責任の追跡を助けるために重要です。
4. 安全性とセキュリティ: AI のセキュリティは、悪意ある攻撃から保護する必要があります。AI システムが攻撃された場合、深刻なリスクが生じる可能性があるため、セキュリティ対策は非常に重要です。
5. 労働市場と雇用: AI の導入は、一部の業界で労働力に影響を及ぼす可能性があります。倫理的な観点から、新しい技術によって置き換えられた労働者のサポートと再教育のプログラムが必要です。
6. 兵器と自律的な AI: 自律的な兵器システムに関する倫理的な懸念が存在し、国際的な合意が必要です。AI を搭載した兵器が人道的な問題を引き起こす可能性があるため、これらの技術の使用に対する制約が必要です。
7. 倫理的なガイドラインと規制: AI の倫理的な使用に関するガイドラインや規制が策定さ

れるべきです。国際的な協力が必要であり、産業界、学界、政府、市民社会が協力して、AIの発展を倫理的かつ持続可能なものにするためのフレームワークを構築する必要があります。

AIと倫理についての議論は進行中であり、多くの困難な質問が残されています。しかし、倫理的な配慮を持ってAIを開発、展開し、透明性と公平性を確保することは、私たちの未来のために不可欠です。

情報的健康（読売新聞、2023年5月30日）

東京大の鳥海不二夫教授（計算社会科学）と、慶応大の山本龍彦教授（憲法学）らは2023年5月29日、情報空間に対する共同提言「健全な言論プラットフォームに向けてver 2.0—情報的健康を、実装へ」を発表した。2022年1月に公表した第1版に、生成AI（人工知能）がもたらす問題などを加筆した。

新たな提言では、生成AIについて「インターネット上のいかなるデータが使用されたか不明のまま、生成された文章が拡散される危険がある」と強調。その回答はもっともらしく見えるため、『『おいしい毒林檎（りんご）』として情報的な『健康』を侵害しうる』とした。

通信・広告事業者が果たすべき役割や、教育・リテラシーのあり方にも言及した。提言を社会に実装する手立ても詳しく紹介し、ユーザーの「情報的健康」の程度を測定するシステムの構築や、リテラシー教材の開発などを示した。

前回の提言は「デジタル・ダイエット宣言」という副題を付け、フェイクニュースや、好みの情報ばかりに包まれる「フィルターバブル」などの問題を指摘。多様な情報を摂取することで、フェイクニュースなどに対する免疫を獲得した状態である「情報的健康」の重要性を訴えた。

第2版にあたる今回は、鳥海、山本両教授に加え、言語脳科学者の酒井邦嘉・東大教授など関連分野の識者ら16人が共同執筆者に名を連ねている。

代表の一人である山本教授は「日本だけでなく世界に向けて情報的健康の概念をPRしていきたい」と話した。

AIと思考

AIは思考するのか。マルクス・ガブリエル『考えるという感覚／思考の意味』（講談社、二〇二四年）によると、思考も視聴覚のような感覚であり「考覚」があると想定している。それは、視聴覚と同様に間違ふことがあり、間違ふことで、思考の内にとどまるのではなく、思考の外の現実と直接触れていることがわかる。つまり、現実を単なる構築されたものだとする極端な見方が退けられている。

そこからさらに述べれば、両手が触れ合うことがわかりやすいのだが、触れて触れられるという、主観客観の二元論では説明しにくい状況が感覚にあるように、思考もまた単に客観的な対象について考える主観的な作用ではなく、思考するもの、思考されたもの、そして思

考という活動が同時に成立し、しかもその思考自体が思考されるという複雑な構造がある。おそらく、日本語の「もの思う」がここでは、思考を適切に示している。

そのような思考はAIにはない。身体という感覚と考覚の場がないからというだけではなく、意味の場が開かれないからだ。生成AIが適切な文字列を配置することができるとしても、その意味を理解しているわけではない。逆に、それを読む人間の方がそこに意味を理解する構造を有しているのだ。

したがって、AIは人間の知性を加速するような、Accelerated Intelligenceとでも言うべきであろう。問題は、AIに対応する人間とは何か、とりわけ、東西を乗り越えるような新しい人間観とは何かである。

「わたし」とは誰か

クオリア（その人が単独で経験するような感覚質）が「わたし」を定義できるのか。

他者の言語を用いて「わたし」が成立するなら、根源的な社会性があるのではないか。

つまり、「わたし」は、近代西洋的な、孤立した、もしくは屹立した個人ではなく、他に開かれ、さらに自と他が嵌入するような、根源的に社会的であるあり方をしている

ここで英語という言語に無理強いして考えてみる。大文字のIではダメで、小文字のiしかも底が開いているように見える筆記体のiが、「わたし」には好ましい。というのも、それは複素数のiでもあって、実数空間で連続的に別の数列に開かれているだけでなく、複素数空間という実数空間からは見ることのできない別の次元の数列にも開かれていることがイメージできる。そうであれば、iとしての「わたし」は最初から、多様な複素性からなるプロセスであって、i-ing(s)とでも表記すれば、それに近づくことができそうである。しかも、この、i-ing(s)というダイナミックな構造は、他者とも共有されていて、入れ子状をなしてすでに常に「わたし」に他の「わたし」が嵌入している。

AI と ai

AIは自分の提示した文字列の「意味」を理解していない。それにもかかわらず、操作する側の「わたし」は、まるでAIがもうひとりの「わたし」であるかのように、さらにはそこに書かれた文字列の意味を理解しているかのように、思ってしまう。こうしてAIが「わたし」として立ち現れてくる事態を、aiすなわち allocated i（あてがわれた わたし）と定義してみる。

なぜそのような事態が生じるか。それは、人間の根本的な社会性のためであって、「わたし」に嵌入してくるものに対峙すると、それを、もうひとりの「わたし」として認識しがちだからである。ここでは具体的に、犬や猫、さらには石や木を考えてみるとよい。

そして、AIとチャットした人が自死に至るケースをどう考えるのか。あるいは、現在流行している、digital funeral、digital ancestor や digital immortality をどう考えるのか。これはなかなか悩ましい問題である

人間の再定義：Human Co-becoming

わたしが考えているのは、Human Co-becoming である。とりわけ西洋哲学の文脈では、人間は Human Being として、存在の側から、つまりは存在としての神の側から見られてきた。それに対して、近年、東洋哲学を研究する人々から、Human Becoming という考えが出されてきている。これは「仁」という古い概念の読み直しでもあり、人間は人間的になっていくものだ、という意味で、わたしはそれに co すなわち「一緒に」という言葉を付け加えて、Human Co-becoming と言ってみたい。つまり、人間はひとりで人間的になることはできず、他者と共にあることではじめて人間的になる。

他者とともにかろうじて人間的になりゆくプロセス (Human Co-becoming) はもうひとりの自分である allocated i の構造をもはみ出し、深い自己変容が生じることだ。その自己変容において、まったく新たに他者と出会うことがよくある。このような自己変容の経験は AI には認められない。

Self-as-WE

出口康夫 (京都大学) が説く Self-as-WE は以下の通りである。

WE ターンという視座において、行為の主体はわたし (I) からわれわれ (WE) へ移行します。そうすると、社会のさまざまな場面で、連鎖的な変化が生じていくでしょう。大きくは、自己 (self) の主体性が拡大されます。

行為をする = Do が、I から WE へ。

「わたし」の認識は「わたし」のみの自己 (self-as-I) ではなく、「われわれ」としての自己 (self-as-WE) に広がってゆくと考えられます。

WE ターンはやがて社会の思想に共有され、「わたし」で閉じている機能主義的人間観を、開放的にしていくでしょう。(出口康夫『AI 親友論』、徳間書店、二〇二三年、三二～三三頁)

この議論は、上述した i-ing(s) や Human Co-becoming の議論と重なっている部分もありながら、大きく異なる部分もある。

(1) 重なっている部分は、「わたし」の手前に根源的な社会性を認めていて、そうした社会性 (「われわれ」) から「わたし」が構成されていることを論じている点である。

(2) 異なっている部分は、中島の議論においては、「わたし」は「われわれ」に還元されないこと、さらに「わたし」が他者とともに変容することに注意を払っている点である。

では、この違いは何に由来しているのか。「関係が主体に先立つ」のが京都学派の特徴的な考え方であるが、中島は関係と主体が創発的に同時に出現していると考えている。

「全体主義的なわれわれ」と「道徳的要請」

出口の言う「悪いわれわれ」が「全体主義的なわれわれ」に陥った場合、どうすればよいのか。出口の議論では、「道徳的要請」が登場するとなるが、それ自体が全体主義的ではないのだろうか。

「われわれ」のメンバーには、例えば、自転車に乗っている「わたし」以外にも、自転車や道路や石ころも含まれていました。「え！ 石ころも、道徳的責任を負うのか？」と驚かれる向きもあろうかと思えます。僕の答えは、驚くべきことに「イエス」です。

(同、六一頁)

「わたし」が自転車事故を起こした際に、その責任は「わたし」を超えて、「自転車や道路や石ころ」にまで延びてゆく。それは、「自転車や道路や石ころ」を含む「われわれ」の責任だからである。

とはいえ、ここですでに出口の議論は曖昧である。「道徳的要請」と自転車事故の「責任」とは同じではないにもかかわらず、それを重ね合わせているからだ。「道徳的要請」の主体である「われわれ」はいかようにも伸び縮みするもので、自転車事故の被害者であっても、「道徳的要請」に従わなければならない。それに対して、自転車事故の「責任」はその被害者にとっては無いが、軽減されるものだ。出口はエージェントの「位置や役割」(同、六一頁)によって「責任」の度合いは異なると述べるが、同じことは「道徳的要請」には当てはまらない。

では、全体主義的な「われわれ」が存在するとして、そのなかで被害に遭っている人々にもまた「道徳的要請」が課せられるのだろうか。それは特殊な場合(被害者であっても加害者に積極的に加担しているような場合)を除いて、適用されるのは難しいのではないだろうか。そうであれば、Self-as-WE という設定自体が、かなり問題のあるものだということになる。

主人と奴隷

「われわれ」を構成するフェロー(仲間)としてAIを遇すると出口は主張する。

出口が批判するのは、AIや技術一般を、主人である人間にとって、奴隷と見る見方である。こうした「不当な差別」をもたらす「主人/奴隷」モデルに替えて、出口はAIを「共冒険者」(同、一二四頁)という仲間として遇しようとする。念頭にあるのは、「アトム」や「ドラえもん」といった日本の漫画の中に登場する親友ロボットである。

道徳的 AI 「悪行フィルター」

だが、もしそう考えるのであれば、AIもしくはロボットは、先ほど批判的に検討した「道徳的要請」や道徳的な責任を何らかの仕方で装着しておかなければならないことになる。と

はいえ、「悪行フィルター」という装置の開発は、至難の業のように見える。それは、「シビ
アな悪」は許容しないが、「マイルドな悪に開かれた AI」をデザインする（同、一五二頁）
必要があるが、「マイルドな悪」は政治的・社会的・文化的な背景によって、いくらでも変
わりうるものだからだ。それでも、出口は道德エージェントとしての AI を、仲間として導
入しようとする。その結論は、次のとおりである。

「e-ひと」は道德的エージェントとして基本的に人間と同じ権利や権限を享受すべき
存在です。例えば、「e-ひと」の参政権も視野に入ってくるでしょう。またそれはフェロ
ーシップ権、反ディスポーザル権に加えて、WE ターンの下、人間の命令に対するより広
範な拒否権を獲得することにもなるでしょう。（同、一六二～一六三頁）

動物について

ここで重要なのは動物の問題である。出口の議論にはなぜか動物が出てこず、「自転車や
道路や石ころ」そして AI やロボットといった、人工物や無機物が論じられる。もし動物を
「われわれ」に入れるとどうなるだろうか。動物に「悪行フィルター」を備えることはでき
ない以上、「e-ひと」ならぬ「animal-ひと」になることは難しい。もし牛や豚や鶏に「反
ディスポーザル権」を与えてしまえば、人間はそうした動物を処分して食べることはできな
い。別の言い方をすれば、出口の議論では、「われわれ」が、程度の違いはあれ、常に道德
化され続けなければ成立しないし、道德化に抵抗しそうな動物は最初から外されなければ
ならないのだ。

たとえば、「e-動物」をデザインしたとを考えてみよう。それに対して、「悪行フィルター」
を備える必要はない。しかし、「e-動物」を allocated i として、「わたし」が認める可能性
は常に開かれている。それはペットをもうひとりの「わたし」として認めるのに近い行為だ
からだ。つまり、「e-ひと」でないにしても、「e-動物」と共に「われわれ」を構成するこ
とは可能なのだ。道德化のプロセスは「われわれ」の構成にとって必須ではないのである。

ただし、これは人間と動物を主人と奴隷の関係に置き続けようとする考えをサポートし
ようというわけではない。事態は逆なのだ。これまで人間は動物を劣った人間として扱い続
けてきた。動物には、人間に備わっている理性や言語が不足している、もしくはそれらを欠
いていると考えてきたのである。そして、この考え方が人間のなかに折り返されると、動物
的な人間として、女性、子ども、老人、奴隷、先住民が表象されてしまったのだ。

モア・ザン・ヒューマン

ところが、事態は逆なのだ。動物は人間以上（モア・ザン・ヒューマン）である。人間の
ような理性や言語が無くとも、より社会的に生きているし、公私の区別が無くとも、よりパ
ーソナルな仕方 で存在している。動物はそのあり方それ自体が高貴であって、人間のように
人間的になり続ける努力を必要としない。フリードリッヒ・ニーチェが発狂した原因や背景

は種々語られているが、重要なのは、鞭打たれた馬の首を抱きしめたことだ。馬の高貴さを鞭打つ人間のどうしようもなさの前にして、ニーチェは馬の方に向かったのである。

必要なことは、人間の側の変容である。それは道德化とは異なる、「よく生きる」や「よくなる」というプロセスを求めるものだ。Human Co-becoming や i-ing(s)は、こうした「よく生きる」や「よくなる」というプロセスにおいてはじめて開かれる複数性を言祝ぐものであって、最初から「われわれ」を前提しているのではない。「何よりも重要なのは、道德の詐術にわれわれが欺かれていないかどうか、それを知ることである」。これはエマニュエル・レヴィナスが『全体性と無限』の序文に書きつけた文章である。道德化を超えて、他者に向かうこと。それがレヴィナスの「倫理」であったことを思い出しておきたい。

ジェンダーについて

すでに女性について言及したが、ジェンダーの問題もここではきわめて重要である。わたしたちが、他者とともにある場合に、ジェンダー化された事態が必ず入り込んでくる。だからこそ、動物に対してそうしたように、女性を差別し、抑圧し、搾取する構造がこれまで生じてきたのである。出口のいう、Self-as-WE の議論には、やはり女性は登場しないが、それは、ジェンダー構造への批判なしに、ジェンダーレスな自己そして我々を前提しているために、これまでの男女の関係を温存しかねない。「全体主義的なわれわれ」とともに、従来の性別分業や男尊女卑の構造を破壊しなければ、その自己は男性に都合のよい自己になってしまう。

先ほどの「道徳的要請」をここで考えてみると、たとえばDVの被害に遭っている女性に対して、加害した男性が「責任」を要求することは、到底考えることはできない。それは不当な要請であって、責任は加害者のみにある。逆に、加害者である男性の責任は、自らの行為への責任を深めるとともに、特殊な仕方で男性に有利に構成された社会のあり方を変容させる責任にまで伸びていかなければならない。それによって、私や自己を構成する我々のあり方を変え、よりましな我々、そして私への通路を開くからだ。

では、AIにはジェンダーの議論が適用されうるのだろうか。

→身体を生きる人間は、必ずジェンダー化された身体を生きている。Self-as-Weであれ、関係性から私を考える議論であれ、それが今日において意義のある議論であるためには、ジェンダー化された我々や関係性について考察し、そこにおいて生じる倫理的・政治的・社会的な要請、とりわけこれまで構造的に抑圧されてきた女性が尊厳を取り戻す要請を実現していくことが不可欠となる。それが満たされることで、AIにおいてもジェンダーの問いがより深まった仕方で論じられ、ジェンダー化された我々のあり方がより公正なものになりうるのだ。