

拡張された心身問題

—人新世と人工知能社会の倫理—

大塚 善樹

環境問題における人新世という新たな問題と、人工知能社会における科学技術の倫理的・社会的な課題は、心身問題という観点で捉えると共通する世界観や論点があると考えられる。そこで本小論は、文献研究を通して、これらの共通点に関する理論的整理を行った。その結果、人新世と人工知能社会は、(1)「心」としての科学技術やAIが、「身体」としての自然や世界に働きかける結果として生じる問題であること、(2)「心」と「身体」の存在論的優位性による包含関係が、「身体」を優位とする唯物論的な世界観から、「心」を優位とする観念論的な世界観への反転が起きつつあること、以上の共通点を確認した。最後に、(2)に陥らずに(1)の問題を解決する方法として、汎心論的アニミズムと中動態に基づく世界観について検討した。

キーワード：人新世，人工知能社会，心身問題，ポストヒューマン，汎心論，アニミズム，中動態

1 はじめに

17世紀にルネ・デカルトが提起して以来、心身問題、つまり人間の心と身体、精神と物質、主体と客体がどう関係し合っているのかという問いは、哲学において未決着の重要問題であり続けてきた。心身問題は心と身体を峻別する二元論を前提として、一方に受動的な身体、物質、客体を置き、他方にそれらを動かす心、精神、主体を仮定する。しかし、両者が質的に全く異なる実体であるとするならば、心から身体への命令はどのようにして可能なのか、という矛盾を孕む問題である。

この小論は、環境問題における人新世という新たな問題と、人工知能社会における科学技術の倫理的・社会的な課題は、心身問題という観点で捉えると共通する世界観や論点があるのではないかという想像に基づく。これまで、心身問題を扱う心の哲学では、人工知能の登場によって新たな視点が可能になり、議論の幅が大きく広がってきた。さらに、人新世という概念も、人間の文化や科学技術と自然環境、能動的に行為する主体と受動的な客体との関係について、再考を強いるものである。したがって、心身問題という視点から人新世と人工知能社会を考えることには、何らかの意味がありそうである。

そこで、以下の節では、第2節で人新世と人工知能社会の倫理的・社会的課題の基底にある考え方について概観し、第3節でそれらの問題と心身問題との関連性について論じる。そして最後の第4節で、二つの問題

に共通する論点と解決方法について検討する。

ところで、本稿は、環境学部とメディア情報学部を擁する横浜キャンパスの将来構想において、両学部で共通する教育研究領域を検討したことが背景にある。とくに、サイバー・フィジカルという課題については、議論の中で大きな示唆を得た。勿論だが、横浜キャンパスの将来構想と本稿との間には何の関連性もないことを付言しておく。

2 人新世と人工知能社会の基底にあるもの

2.1 人新世の倫理的・社会的課題

2000年に大気化学者のPaul Crutzenが学会で発案し、それ以前からその言葉を使っていた生物学者のEugen Stoermerとともに同年に論文で新たな地質学年代として提案された人新世(anthropocene)という概念は、環境に関する自然科学だけでなく、人文社会科学においても様々な議論を引き起こしている。

2019年の立川雅司による人新世の社会科学に関する総説では、Crutzen and Stoermerによる定義[1]を、「地球と大気に対する人類の活動が及ぼす、地球スケールでの大きな影響を考慮するならば、現在の地質学的面期に対して“人新世”という用語を与えることで、地質学や生態学において人類が中心的な役割を果たしていることを強調することが適切なように思われる」[2]と訳出している。さらに立川は、社会科学分野における人新世に関する最近の議論を、(1)人新世の開始点、(2)人新世をめぐる様々な言説の分類、そして(3)人新世がもたらす含意の3つの視点からまとめている。このうち、(3)について、①自然と社会との関係性の問い

直し、②持続性やガバナンスをめぐる課題、③人間の基本的生存条件に対する認識の変化という3点の課題を指摘した。

より人文系の Eva Horn and Hannes Bergthaller による総説書 [3] では、人新世に関する人文社会科学系の議論を思想史的な文脈で捉えるとともに、自然と文化、人間とは誰か、政治、美学、生政治、エネルギー、時間的・空間的スケールの問題といった包括的な観点から、人新世の概念が提起する課題を論じている。立川の整理にもあるように、自然と文化（あるいは社会）についての考え方の転換が、より具体的な倫理的・社会的問題の基底にあることは間違いない。自然に対する人間社会の介入（人間中心主義）を肯定的に捉えるか、否定的に捉えるか（脱人間中心主義）という立場の違いによって、例えば、気候工学による気候変動の制御、ゲノム編集技術を用いた遺伝子ドライブによる生物多様性の保全など、科学技術の倫理的・社会的評価は異なったものとなるであろう。

前者のように社会的介入を肯定する立場は、エコモダニズムと呼ばれ、科学技術によって「良い人新世」を目指そうとする Breakthrough Institute の An Ecomodernist Manifesto に表されている [4]。このような介入は、「そこにはもはや純粋の自然はないのだから、人間中心的視点から自然に介入していくことは問題とされない」[2]として正当化される。しかし、後者の脱人間中心主義を徹底していくと、人間は他の生物種や無機物との共生関係の結び目に過ぎないとして、ブルノ・ラトゥール (Bruno Latour)、ドナ・ハラウェイ (Donna Haraway)、Anna Tsing などの論者のように人間と自然の区別を完全に廃棄する考え方へと至る。このような議論を Horn and Bergthaller は、エコロジカル・ポストヒューマニズムと命名し、人間が人間以外の種を抑圧してきた事実を過小評価することになると批判している。そして、エコモダニズムとエコロジカル・ポストヒューマニズムは、文化的で主体的な行為者という側面と自然の中の一つの生物種という側面からなる人間の二面性の、それぞれ一面しか考慮しない極論であるとする [3, pp.72-74]。

自然と文化についての考え方の相違は、科学技術の捉え方にも微妙な影響を与える。一方で、エコモダニズムは科学技術の帰結に対して楽観的かつ肯定的であり、科学技術を積極的に用いることを通して人間が地球システムを保全する責務を負うと考える。他方で、エコロジカル・ポストヒューマニズムは科学技術による人工物も、自然と文化が分かち難く、もつれ合った (entangled) もの、つまり自然と文化のハイブリッドであると考えられる。ラトゥールやハラウェイは科学技術の人類学の系譜に属し、フランケンシュタイン [5]、サイボーグや遺伝子組換え生物 [6] など、これまででも科学技術による

ハイブリッドの生成を容認する方向で論じてきた。とすると、人新世の議論においては、これら二つの潮流以外で、科学技術に批判的な第三の陣営の存在が想像される。

ここで、立川も引用している Anne Fremaux and John Barry [7] による、自然と文化の二元論を保持するかどうか (human or posthuman)、科学技術による自然の改変を肯定するかどうか (most-modern or postmodern) の二軸による4類型を検討しよう。上述のエコモダニズムとエコロジカル・ポストヒューマニズムは、どちらも most-modern であるとされる。most-modern とは、フェミニズムや環境問題で発言してきた米国人作家 Charlene Spretnak の用語で、資本主義的な近代のプロジェクトをさらに強化し加速するナラティブである。most-modern-posthuman に類型化されるラトゥールやハラウェイは、自然と文化の境界を超えるポストモダンの思想家として語られることが多いが、ここでは自然を科学技術に吸収して“techno-nature”とする超モダニストとみなされている。その妥当性はともかく、より本質的な科学技術の批判として彼女らが考える postmodern の2類型は何だろうか。posthuman-postmodern はエコ・フェミニズムやラディカル・エコロジーとするので、ディープ・エコロジーも含まれると思われるが、環境倫理では伝統的な議論であって必ずしも posthuman と呼ぶことはできないであろう。そして、著者らの立場である human-postmodern は、1947年のホルクハイマーとアドルノの『啓蒙の弁証法』を引用している。また、ラトゥールやハラウェイによる自然と文化を混然一体と捉える考え方を、アドルノの『否定弁証法』における「非-同一性」という概念から批判している。以上から、human-postmodern はフランクフルト学派の思想に近いと思われる。しかし、この postmodern の2類型の記述は十分ではなく、通常の postmodern の理解とも異なる。どちらも、自然と文化の区別は維持しながら、自然と人間を破壊しないレベルに科学技術を制御する考え方とみなしてよいであろう。

加えて、科学技術の位置づけについて注目すべき議論を行っているのは、地理学者の Peter K. Haff である [8] [9]。Haff は人新世のテクノロジーを気圏や水圏と同様の自律的な圏域、テクノ圏 (technosphere) とみなし、(1) アクセス不可能性：多くのテクノロジーは個々の人間に直接的影響しない、(2) 不能性：ほとんどの人間はテクノロジーに影響を及ぼすことができない、(3) 制御不能性：人間は自身より多くの行動を有するシステムを制御できない、(4) 交換性：人間は自身と同じサイズのシステムとしか直接的な相互作用ができない、(5) パフォーマンス：人間はテクノ圏の一要素としてテクノ圏の代謝を支えなければならない、(6) 提供義務：

テクノ圏は人間の生存のための環境を提供しなければならない、以上6つのルールによって自律的に運動しているとした [8]。

もはや科学技術は人間の能力の延長ではなく、人間が科学技術の一部であるという洞察は、人新世という問題構成において重要であろう。例えば、自動車交通システムを考えれば、運転者としての人間がそのシステムの一構成要素に過ぎないことは明白である。自動車交通システムは、エネルギーの調達、インフラの整備、廃棄物や排気ガスの排出まで、グローバル経済の枠組みのなかでそれ自体の再生産と効率化を動因として変化する。このような自律的なテクノ圏に対して、環境破壊を抑制し人間の地位を回復するためには、人間特例主義になっても、人間らしさについての信念を保持することがテクノ圏を変革する力となることを説いている [9]。この意味で、Haffの議論は posthuman な状況を科学技術の面からより精緻に把握しつつ、Fremaux and Barry と同様にその制御を考えるものと言えよう。

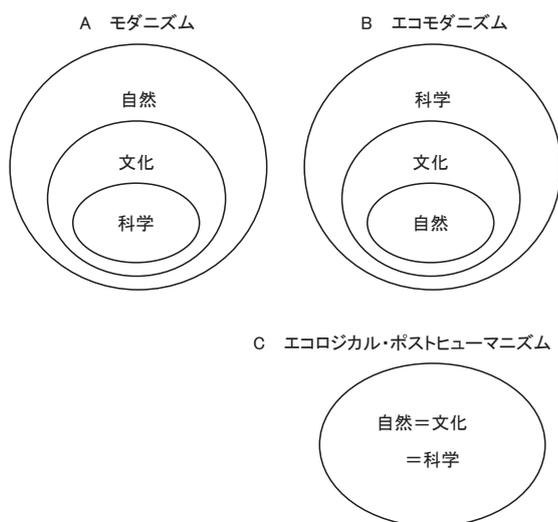


図1 人新世における世界観
筆者による。科学は科学技術を指す。

以上の検討から、人新世における自然、文化、科学技術の関係を表す世界観を、極端な類型として図式的にまとめたものが図1である。図1のAは現代における自然・文化・科学技術の関係についての普通のモダニズム的理解であり、自然は文化の外部にあるという二元論的な思考を前提にしている。科学技術が理解し介入できる自然は一部に限定され、故に環境問題の分布も対策も不均等で不平等である。人新世はこの常識をBのように逆転し、自然も文化も科学技術（テクノ圏）の一部として均質化

する変化であると考えられるだろう。この変化は、エコモダニズムが目指す科学技術による自然と文化のマネジメントと一致する。Aは二元論的な世界観を前提としているが、Bではもはや自然は文化の一部であるとする、科学技術による一元論に近づいている。これに対して、Cはそのような区別を廃して、科学技術による人工物も自然と文化がもつれ合ったものとする。フラットな存在論 (flat ontology) と呼ばれるCの主張は、人新世以前も以後も、AやBは現実には存在せず、つねにCのようにあり続けているというものである。人新世はそのような現実を明るみに出してくれたに過ぎないとする。

重要なことは、この3種類の図式のどれが正しく、どれが間違っているかではなく、複数の世界観が可能であり、その世界観によって現状の理解や将来進むべき方向性が異なってくるということに認識することである。

2.2 人工知能社会の倫理的・社会的課題

AIが社会にもたらす影響については、すでに多くの提言がなされてきている。ここでは、2017年末にThe IEEE Global Initiativeが公開した Ethically Aligned Design version 2 とその日本語概要版 (EADv2) [10]、2017年9月から2018年2月に行われた AI Initiative の報告書 (英語) および日本語訳概要「国際的な市民討論 人工知能の到来をいかに舵取りするか」(AII) [11]、2019年5月にG7各国の学術会議でとりまとめられたGサイエンス学術会議共同声明「人工知能と社会」(GS) [12]、および2018年の科学技術社会論研究16号の特集「人工知能社会のあるべき姿を求めて」所収の諸論文—そのなかにはAIネットワーク社会推進会議の「報告書2017」の解説[13]も含まれる—を主要な素材とする。

最初にAIの定義だが、AIIでは「スーパーコンピューティングによるアルゴリズムを中心としたビッグデータ駆動の機械学習システム」、GSでは「基本的には(通常は大量の)データを使って作動するアルゴリズムの集合体」としている。ここで、AIが大量のデータの存在を前提にしていることは、当然ではあるが銘記すべきである。また、EADv2では「インテリジェントで自律的な技術システム」として、自律的であることが強調されている。さらに、GS「オンラインで進化するシステム」、AII「AI統合システム」、AIネットワーク社会推進会議「AIシステム相互間の連携」、EADv2「自己改良能力のある技術システム」「汎用人工知能 (AGI)」「人工超知能」といった多岐にわたるAIの将来像も言及されている。

すなわち、現在では、「汎用人工知能等が実現をしない限り (そしてそれはそう簡単ではないが)、技術の目的を設定するのはあくまで人間である」[14]としても、将来的にはAIの目的自体がAIによって決定されるこ

とも起こり得る。加えて現在でも、「学習や最適化に基づく人工知能システムの振る舞いを人間の常識や価値観に沿わせるのは意外に難しく、むしろ簡単に非常識なものになってしまう問題は深刻である」[15]という状況を考えると、単に評価関数をどのように適切に設定するかという問題だけでなく、人間が適切にAIの「振る舞い」を制御できるのかという疑問も生じる。

次にAIの倫理的・社会的影響については、EADv2本文の章のタイトルに即して整理すると、(1) 個人データとアクセス制御、(2) 自律型兵器システムへの軍事利用、(3) 経済効果による福祉の促進、(4) 法的責任の範囲、(5) アフェクティブ・コンピューティング、つまり人間の感情への影響、(6) 政策による教育と認知、(7) 古典的な道德規範を超える意思決定、(8) 混合現実による個人の権利やアイデンティティへの影響、(9) 人間の福祉への影響の評価方法、以上9点が挙げられている。これに、(3)や(9)とも関連するが、AIIとGSで言及されている(10)雇用への影響、を独立した項目として加えることもできるだろう。さらに、AIの設計時に生じる可能性がある課題として、EADv2で挙げられている(11)自律システムへの価値観の組み込み、(12)倫理的な研究のための価値ベースの設計方法論、(13)汎用人工知能(AGI)と人工超知能の安全性も倫理的・社会的課題に含めることができると考えられる。(12)については、社会的コストと経済的価値の増大の両方を評価できるシステムということだが、これはAIIで言及されている「先行者利益によって市場シェアを獲得しようとする業界または国の『倫理的底流への競争』を避けるため」のスマートガバナンスに通じるものであろう。そのほか、上記の文献は社会への影響を中心にまとめられているため、自然環境への影響についての評価が行われていない。AIIでAIがSDGsの達成に役立つことが書かれているが、気候変動のシミュレーションに見られるように、AIは自然環境に関するデータも収集して環境を改変する際の意思決定に利用されている。

このように人工知能社会の倫理的・社会的課題は広範な領域に及んでいるが、AIの定義を「データを使って作動するアルゴリズム」として、倫理的・社会的課題をAIと実世界との関係で生じる問題と考えるならば、主として実世界(社会だけでなく自然環境を含む)からデータを入力する段階に関わる問題と、学習や最適化や統合の結果を出力して実世界を変化させる段階における問題に分けて考えることができる。

前者、つまり実世界のデータの収集と入力、上記の(1)、(11)、(12)でとくに課題となっているが、(2)～(10)の場面でもデータの収集と入力は起きているはずである。データの収集のアルゴリズムは、当初は人間が決定するにしても、対象領域によっては自律的に

われるであろう。また、データの収集は、実世界に前もって存在しているデータを単に集めるのではなく、実世界をAIとして利用可能なデータの収集に適した形態に変化させるプロセスでもある。言い換えると、データは既存の事実ではなく、収集という行為(それが人間であれAIであれ)によって新たにつくられる。すでに商品の購入方法は、ビッグデータ収集に適した形態に変化し、製造と小売と消費者の行動を変えている。医療分野でも、ICTの利用による医療データ管理システム(e-Health)が患者と医療専門家との関係を変えていることが分かってきた[16]。e-Healthは、医療行為が行われる時間と空間のあり方、患者と医療者の感情的、認知的、物質的な関係、医療者の作業の配分のパターンを変化させている。民間企業による自律的なAIが医療情報の収集を独占したり、効率化のために収集方法を変えたりするならば、さらに異なった変化が起こり得るであろう。

また、(11)においては、システムの導入先であるコミュニティの規範、とくにシステム的设计目標であるタスクに関わる規範をAIに組み込むことが想定されている。しかし、データの場合と同様に、規範も実世界に所与のものとして固定された状態で存在しているわけではない。例えば、上述のような医療コミュニティの場合、診断と治療の意思決定には患者へのアカウントビリティと事後の責任が求められるという規範がある。しかし、近年のゲノム医療で見られるような偶発的所見、二次的所見—本来の目的とは異なる疾患や遺伝に関する情報、本人だけでなく家族や第三者に影響を与える場合もある—のように医療のICT化との関係で、どこまで患者へ説明することが望ましいのか答えが出ていない先端医療領域もある[17]。そのような場合には、AIに組み込むことのできる規範は前もって存在せず、AIとの相互作用のなかで新たにつくられる。すなわち、将来的にAIが実世界のあらゆる領域に浸透するならば、実世界そのものが、データの収集パターンだけでなく、実世界の規範や価値のAIへの組み込みに合わせて作り変えられることが考えられる。この意味では、現在のAIと社会についてのテクノロジーアセスメントは、社会の側を所与のものとして固定的に捉える点で、AIと社会の相互作用についての考慮が足りないのではないだろうか。

後者のように、AIによる学習や最適化や統合の結果を出力する段階では、当然のことながら、実世界のあり方に変化を及ぼす。とくに(4)や(7)のように、政治的あるいは倫理的な意思決定において、これまでの社会的規範を超える行動がなされる場合も想定されている。しかし、その変化が実世界にとって望ましいものであるのか、そうでないのかの評価は難しい。なぜなら、

AIは前述したようにデータや規範を収集する過程で、自らに適合したデータや規範となるように実世界をつくり変えているため、その結果としてAIが行う意思決定は、すでにつくり変えられた実世界にうまく適合することは想像に難くないからである。AIが自身に適合した実世界をつくることで、実世界に適合し高評価を得るとすれば、評価の意味は無くなってしまふ。データや規範の収集やそれに基づく自身の改変まですべて自律的に行うAIが現れば、AIと実世界の関係は、ますますAI自身の自己言及的な過程に閉ざされていく可能性がある。

このことは、AIIにおいて「機械の知的レベルが一般的なものからスーパーインテリジェンスと呼ばれる段階に進化するにつれ、これまでに宇宙で知られている中で最も認知的に発展した存在である人間に対するパラダイムが変化し、新たにAI中心の世界において人類とAIの関連性や人間の役割、および将来の発展についての議論が必要になります」として楽観的に述べられている。人間中心の世界から「AI中心の世界」への変化である。ここから人新世の概念が孕む問題との相似性を確認することができるであろう。

すなわち、現状では実世界の中の一部のみがデータ化され、その一部のみを用いたAIが実世界の一部に影響を及ぼしている(図2のA)。もちろん、自然環境を含む実世界のすべてがデータ化されることはあり得ない。しかし、実世界の多くの領域でICTに依存する部分が拡大すればするほど、データ化された「世界」のみが人間にとって意味をもつようになるであろう。結果として、実世界とは異なる「世界」が構築され、AIによって制御されるようになる可能性がある。そのような「世界」はもはやAIを制御するすべを持たないだけでなく、自らをデータ化して統合するメタデータへのアクセス権も持たないであろう(図2のB)。人新世の場合と同様に、人工知能社会の倫理的・社会的課題の基底にあるのは、このような世界観の反転ではないだろうか。

もちろん、「AIのなかの世界」の「世界」は実世界とは異なったAIがつくる「世界」である。しかし、そのような「世界」以外の世界は存在しないという世界観がグローバル経済を動かすことは十分に起こり得るし、それによって実世界の分断や包摂されない世界の排除が問題化するという可能性もあるだろう。Bはそのような問題を孕んだ世界観の模式図である。

3 心身問題からみる人新世と人工知能社会

3.1 心身問題からみる人新世

トーマス・ネーゲル(Thomas Nagel)は、2012年の著書 *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature is Almost Certainly*

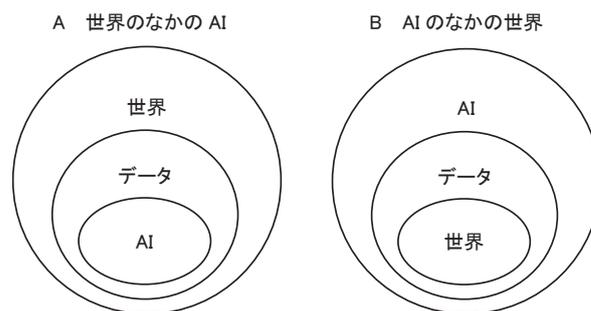


図2 人工知能社会における世界観
筆者による

Falseにおいて、「心身問題は、単に生物の心、脳、行動の間関係についてのローカルな問題ではなく、宇宙全体とその歴史についての私たちの理解へと拡張する問題であることを論じる」と書いた[18]。

ネーゲルの主張は、心身問題で二元論的な理解の根拠となっている人間の「意識」を唯物論的なネオ・ダーウィニズムの自然選択説によって説明しようとする、物質からの「意識」や「価値」の創発(emergence)、あるいは逆にあらゆる物質にも心があるとする汎心論(panpsychism)に陥るので支持できないというものだ。生物学には、創発性を細胞の形成や発生のメカニズムに適用する考え方が存在している。しかし、自然選択説を厳密に解釈するならば、創発性の主張は非物質的で根拠のない二元論的な概念を、唯物論的な科学に持ち込むことになる。また、ネーゲルは「意識」だけでなく「価値」の客観的実在をも前提とする哲学者なので、自然選択説に立つならば、この「価値」も生命の起原にさかのぼって物質レベルで説明されなければならないと考える。結果として、偶然の変異と自然選択に基づくネオ・ダーウィニズムを否定し、独自の道徳的実在論に基づく一元論を展開する。この一元論では、心身問題を宇宙全体に拡張し、心と宇宙、文化と自然の二項対立を解消することが謳われている。

これに対して、環境や自然の哲学でも知られるノルウェーの哲学者アルネ・ヨハン・ヴェトレセン(Arne Johan Vetlesen)は、人新世を論じた近著のなかで、ネーゲルの捉え方を未だに人間中心主義的かつ主知主義的であると批判している[19]。人新世の概念は人文社会科学の領域で多様な反応を引き起こしているが、そのひとつの傾向が文化と自然の関係性を再考し、文化と自然は分かちがたく相互浸透しているとの前提に立って人新世をもたらした社会変化を批判して転換することである。ヴェトレセンによると、ネーゲルの理論は人間

特有の「意識」や「価値」に基づく一元論であるため、自然や環境は受動的な立場に置かれたままであり、人新世を転換する革新性に欠けるのである。

ここでは、人新世という問いに対して心身問題をどのように拡張するのかという視点が問われている。能動的な「心」を人間の文化や科学技術と捉え、受動的な「身体」が自然環境であるとするのが古典的な心身二元論である。そのときに、図1のAモダニズムでは、文化や科学技術がその外部にある自然へと働きかける。このとき、自然には未知の部分が含まれるため、「心」＝文化や科学技術による「身体」＝自然の支配は限定的なものである。しかし、図1のBエコモダニズムでは、自然は文化や科学技術の一部である。もはや自然に未知の部分はない。「心」＝文化や科学技術による「身体」＝自然の支配は完璧なものとなる。

つまり、心身問題の枠組みでは、図3に示すように、「心」が「身体」の一部として存在する—「心」は脳の神経回路によって物質的に説明できる—とする唯物論的な立場を認めつつ、その物質的な「心」を拡張した科学が「身体」としての物質や自然を部分的には制御できるだろうという考え方がAの唯物論的モダニズムであり、図1のAのモダニズムに対応する。心身問題の唯物論的な理解は現代の心の哲学の主流派であるが、完全な一元論になることは稀で、「心」や「意識」の存在自体は認める場合が多い。

他方で、図3のBは、「身体」が「心」に含まれるという考え方を示す。唯物論に対立する思想は観念論であるが、心の哲学で観念論的立場の哲学者は見当たらない。観念論を突き詰めていくと、意識にないものの実在を認めないフッサールの現象学に至るが、そこでは心身問題は否定される。フッサールにあつては、志向の対象のみからなる現象的世界が世界のすべてである。そのような意味で、図3のBは究極的には独我論的な観念論につながるが、「身体」を区別している点で完全に実在的对象を排除しているわけではない。実在としての「身体」を排除してしまうと、「心」＝主体からの働きかけに意味がなくなってしまう。そのような意味で、観念論的モダニズムと名付けよう。

このように考えると、人新世におけるエコモダニズムは観念論的な自己言及性を特徴とすることが考えられる。科学技術は自らが認知し制御することのできる文化や自然のみを対象として、それ以外のものは排除する。常識的なモダニズムでは、認知や制御の不十分さが露わになり、科学的な不定性や環境不正義が横行するが、少なくともそれらの問題が見えていた。しかし、自信にあふれたエコモダニズムは、ほとんどの問題は制御可能であると考えているが、それは問題の一部しか見ていないからである。

まとめると、「心」が「身体」の内側にあるか外側にあるかによって、「心」が「身体」を支配する範囲と強度が異なってくる。「心」が「身体」の内側にある場合(図3のA)、「心」による「身体」の支配は限定的で、影響の及ばない部分が存在する。よって、二元論的な心身問題は継続する。しかし、「心」が「身体」の外側にある場合(図3のB)、「心」は「身体」のすべてを支配することになり、心身問題は「心」の一元論に近づく。それでも、「心」が作用する客体としての「身体」が区別され、「身体」のない「心」も存在することから、心身二元論は維持される。そのような構造を孕むものとして、人新世を考えることができるだろう。

ところで、人新世の問題においては、エコロジカル・ポストヒューマニズムという第三の立場が現れていた。心身問題では、ネーゲルが退けていた汎心論という立場があり得る。これは、すべての物質的実体には「心」があると認めるもので、荒唐無稽な考え方と思われがちだが、現代の心の哲学のなかで、そして思弁的実在論やオブジェクト指向存在論と呼ばれる新しい哲学の潮流のなかで一定の立場を占めている[20][21]。もちろん、物質に人間のような「心」があるという主張ではなく、物質的とされる現象を構成する基本要素に、「心」として経験されるような何らかの性質が含まれているという説で、汎心論とも呼ばれる。あらゆる物質的「身体」に「心」が存在するので、人間と人間以外のモノ(non-human)を区別する基準は失われる。そこで、この考え方をC汎心論的ポストヒューマニズムと捉えておく。

ラトゥールやハラウェイも人間とモノの区別を曖昧化し、自然と文化の差異を脱構築していることから、汎心論的ポストヒューマニズムは、人新世におけるエコロジカル・ポストヒューマニズムと近い関係にある。ラトゥールやハラウェイは必ずしも汎心論的な考え方を支持していないが、汎心論はエコロジカル・ポストヒューマニズムを基礎づける考え方とも言えるかもしれない。しかし、行為主体性(agency)をどのように考えるか、科学技術をどのようにみなすかという点については注意が必要である。

前述したヴェトレセンは、「作動中の汎心論(panpsychism in action)」としてアニミズムを取り上げている[19, pp.162-163]。アニミズムの事例で引用している人類学者のエドゥアルド・コーン(Eduardo Kohn)は、アマゾンの狩猟民であるルナ・プーマのフィールドワークから、あらゆる生ある存在は思考するのであり、その「生ある思考」、つまりこの小論で言うところの「心」が生命にとって自己と自己以外を区別することの根源であると主張している[22, pp.125-138]。アニミズムでは、人間が人間以外の生命に「心」を見出すだけでなく、人間以外の生命もまた人間に「心」を見

出す。あらゆる生命はこの「心」の経験を通して相互作用する。デュルケム以来の人類学は、このような考え方を非合理的な自然の擬人化（社会的に構築されたもの）として退けてきたが、コーンをはじめ、それを西欧とは異なる存在論として評価する動きがある。

コーンによると、ラトウールは科学技術におけるモノの抵抗を行為主体性と同一視するため、「人間的なるもののかなたに実際に存在する諸々の行為主体性に対して、私たちの目は閉ざされたままになる」[22, p. 161]。コーンの立場は生命を特別視する考え方だが、それがなければ人間と他の生命との関係性、さらには死や絶滅を語ることはできない。アニミズムの実践者たちは、種の絶滅を招いたり、生命の行為主体性に介入したりする科学技術を肯定することはないであろう。

一方、ハラウェイは、多様な種と人間との連帯の可能性を「異種協働のポリティクス」と呼び、「互いに相手を形成しあう種の結び目」で動物と人間が敬意をもって応答し合うことを論じている [23]。コーンもハラウェイの影響を強く受けていることから、エコロジカル・ポストヒューマニズムにアニミズム的な要素を付け加える可能性はあると思われる。アニミズムが示す汎心論は、哲学的な汎心論とは違った視点で、人新世が提起する問題へ独自のアプローチを示す可能性がある。ここでは、C 汎心論的ポストヒューマニズムの特殊な場合として、アニミズム的な汎心論を捉えておこう。

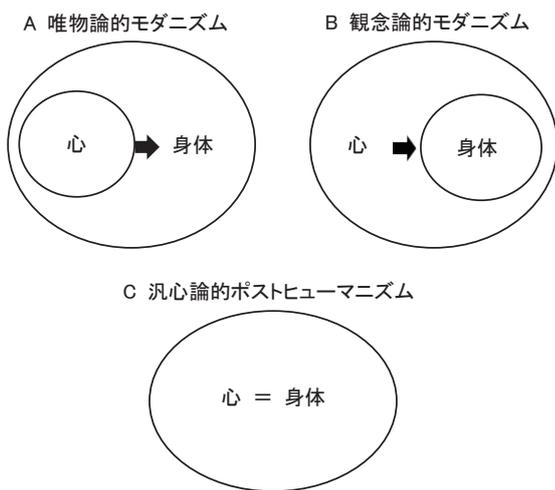


図3 心身問題の心身関係
筆者による

3.2 心身問題からみる人工知能社会

AIは心身問題を考える哲学者たちに、非常に強い影響を与えてきた。例えば、古典的計算主義やそれに対する

「中国語の部屋」の思考実験、コネクショニズム、機能主義、消去主義、そして意識やクオリアをめぐる議論など、1980年代以降、AIとの関係を抜きにしては成立しなかった議論が豊富に生まれてきた。これらについては、信原 [24]、チャーチランド [25]、金杉 [26]などを参照されたい。

この節では、それらの議論を念頭に置きながら、2.2. で述べた人工知能社会の倫理的・社会的課題と心身問題、とくに図3の類型との関係について検討する。図2では、Aとして「世界のなかのAI」、AにおけるAIとデータと世界を逆転した極端な世界観Bとして「AIのなかの世界」を考えた。ここで、AIを「心」を拡張したもの、実世界を「身体」を拡張したものとして捉えることに異論はないであろう。

問題は、データの位置づけである。人新世の問題では、図1の中間にある文化が本来は「心」を表していた。科学技術はその文化のなかの一部、ただし影響力が強大な一部である。その意味で、科学技術=文化と考えて前節での議論を行った。しかし、図2のAは、世界のなかの一部のみがデータ化されていることを表す。人工知能社会におけるデータは、「心」と「身体」、精神と物質を媒介する中間項として特殊な役割を持っていると考える。さらに、2.2. で論じたように、AIは実世界のデータや規範の収集を通して、実世界をつくり変えられ、それが「AIのなかの世界」へと反転する一つの根拠となっていた。その意味でも、データの役割は重要である。

ここで、データは実世界をバイナリに変換したものという意味ではない。データの特徴は、客観的に標準化されていることであり、それによって伝達や解釈や処理が可能となる。再度、医療での患者の診察の場面を考えてみると、血液検査の数値、CTの画像、電子カルテに記載されている情報と医師による書き込み、これらはすべて客観的なデータとして標準化され、e-Healthのネットワークに蓄積される。では、診察の場面でデータ化されないものは何だろうか。患者と医師との会話のうち電子カルテに書かれない部分があるかもしれない。患者と医師の言葉遣い、表情や身振り、沈黙の時間、交錯するお互いの感情、診察室の光景、点滅する医療機器、音や匂い、窓の外の風景など、二人が時間と空間を共有する主観的な経験の記憶はデータとしては残らない。だが、それは患者にとっては、何度も想い返されるショッキングな経験かもしれない。このような実世界の主観的な経験を標準化し、客観的に流通可能な記号に変換した事実の集合がデータである。

しかし上述の例で、患者あるいは医師個人の主観的経験は、彼/彼女の「心」つまり脳にとっては、データとしての位置を占めると考えてよいであろう。つまり、心

身問題との関係においては、AIが「心」、実世界が「身体」だとすると、データとは本来は主観的な経験である。個人の「心」が世界という「身体」を知り、作用できるのは経験を通してである。図3の黒い矢印を主体から対象への志向性と捉えて、そこで生じる経験が「心」と「身体」を媒介すると考えよう。「心」が経験できるのは、実世界のなかで知覚される一部の対象に限られる。それ以外のモノは経験されないが、「心」とは独立に存在している。これが図3のA唯物論的モダニズムにおける経験の役割である。しかし、Bの観念論的モダニズムでは、経験される対象のみが実在する。つまり、「心」が経験というデータを通して世界を限定する。心身問題を人工知能社会へと拡張する際に、主観的な経験は客観的なデータへと標準化される。その過程を、前述のように「AIは実世界のデータや規範の収集を通して、実世界をつくり変える」と表現した。

唯物論の世界観の下では、e-Healthに蓄積されるデータは、患者と医師の経験の一部に過ぎなかった。しかし、e-Healthが拡大して、多様な経験を取り込めるようになるかもしれない。さらには、医療行為を最適化するための手段だったe-Healthへデータを取り込むこと自体が目的化すると、e-Healthに入力するための診察が行われ、患者と医師の経験も変化するであろう。やがて、e-Healthが自律的なAIとして自らを変化させたり、他のAIとのネットワークから影響を受けたりするようになると、実世界からは独立した観念論の世界観として「AIのなかの世界」が生成し、成長していくことになると思われる。ここでももちろん、「AIのなかの世界」は「AIの外の世界」を切り捨てることで生成する。加えて、「AIのなかの世界」のデータや世界からも独立したAIの領域があり、そこでは「身体」から独立した「心」のように、データや世界による拘束を受けない。AIは完全な自律性を手に入れる。そのような世界観である。

心身問題への第3の回答である汎心論的ポストヒューマニズムでは、汎経験論とも呼ばれるように、「心」を経験として読み替えることができる。あらゆる物質的な存在＝「身体」は、何らかの経験を持つことができるというのが、汎心論が意味することであった。そして、人間を含むあらゆる存在は、この経験を通して相互作用する。人間の「心」はモノを理解し働きかけるだけでなく、モノの「心」が人間を見つめて人間に作用する。この汎心論を人工知能社会に拡張すると、人間以外の人工物を含むあらゆるものに物質としての側面とデータとしての側面があり、同時にAIはそれらすべてを併せたもの以上ではなく、同等であるということになる。

このような状況は、少なくとも現在の情報に関する科学技術としてのAIについては、すべての物質的なもの—例えば、水や大気や生物—を物質として包含すること

は想像し難い。と言うのも、純粋なAIはアルゴリズムであって、固有の物質性を欠いているからである。それが図2にCがない理由である。

4 人新世と人工知能社会の倫理

4.1 人新世と人工知能社会の問題の同型性

これまでの考察から、人新世と人工知能社会は心身問題という補助線を引くことによって、同じ形式の問題と考えられることが分かった。この同型の問題には、二つの側面がある。

第一に、「心」としての科学技術とAIが、「身体」としての自然や世界に働きかける結果として生じる問題であることである。ここで、科学技術やその一部でもあるAIは、非物質的な人間の精神あるいは思想に由来する作用主体—AIの場合はその技術自体もアルゴリズム—であるのに対して、作用の客体である自然や世界は、モダニズム、つまり現代社会に普通の考え方では、物質的なものとみなされている。環境問題が地球規模に拡大して人新世と倫理的・社会的影響呼ばざるを得なくなった背景には、自然を客体化して、科学技術に先導される現代社会の資源として利用してきたことがある。AIの場合でも、2.2節で挙げた13項目のなかで、データや規範の取得ではない直接的な影響には、同様の構造があるであろう。例えば、(1)個人データのセキュリティの問題は、個人データの一方的な利用から生じる。(2)自律型兵器システムへの軍事利用においても、国家によるAIの独占が起きることになる。ポストモダニズムやポストヒューマニズムと呼ばれる思想では、このような主客の分離への批判が起きてきた。

第二に、「心」としての科学技術とAIを「身体」としての自然や世界の、どちらが存在論的に優位にあるか、言い換えると、どちらがよりリアリティがあるものとして他方を限定するかを、包含関係として把握することができる。心身問題の唯物論的な理解に基づくならば、「心」は物質的な「身体」によって還元されるとまでは言わない場合でも一つ作られる。物質のほうがよりリアルなものとして、精神を構成する。この関係を拡張した場合、科学技術やAIは物理的に基礎づけられる。科学技術は自然環境のなかで作動し、AIは素材がシリコンか有機物かはともかく、アルゴリズムが作動する何らかの物質的なデバイスに依存する。ところが、観念論はこれを反転し、世界は「心」が経験する現象世界としてしか存在しないとす。実世界でのこのような反転は、科学技術やAIが世界全体に拡大する可能性が考えられなければ、意味はないであろう。人新世や人工知能社会は、そのような可能性が予見される状況として現れた。科学技術が自然界全体に浸透し得るならば、科学技術の制御下にある世界が自然である。そこでは、環境問

題は完全に制御され、問題自体が存在しない。同様に、「AIのなかの世界」の内部では、AIへの批判は起こりえない。したがって、観念論的な世界観への反転は、第一の同型性で起きた問題を見かけ上解決する一つの方法となり得る。そして、そのような解決への志向、完全な制御への希望が、世界観の反転への要因となっていると考えられる。

まとめると、人新世と人工知能社会における倫理的・社会的な問題とは、第一に、主客の分離にともなう不均等な環境破壊や情報格差の問題、第二に、第一の問題を解決しようとして起こる唯物論的世界観の反転にともなう科学技術やAIによる支配の問題である。

4.2 解決法としての中動態

最後に、これら二つの同型的な問題をどう解決するか、抽象的なレベルではあるが、心身問題との関連で考えてみたい。

第一の主客の分離問題を、観念論的な反転に陥らずに解決する方法として、図1のCエコロジカル・ポストヒューマニズムの論者たち、とくにアクターネットワーク理論を唱導してきたラトゥールは、客体とされてきた人間以外のモノ（生物を含む）もまた行為主体であることを主張してきた。ただし、非生物のモノ、人工的なモノの抵抗性—実験室や技術的人工物が人間の意図した通りには働かないこと—を行為主体性と同一視する考え方には、前述のコーンのような批判もあり得る。ヴェトレセンもまた、ラトゥールの抵抗性の概念が、行為主体性としては貧弱で、かつ特殊な状況に依存することを懸念している [19, p.234].

コーンのような人類学者が再評価し、ヴェトレセンが人新世と関連付けて主張するアニミズムは、確かに行為主体を生あるものに限定することで、それらと人間の相互行為やつながりを回復する。人間が科学技術やAIを制御できるという条件のもとでは、つまり普通のモダニズムや「世界のなかのAI」の段階では、主客の分離にともなう問題の解決につながる可能性はあるだろう。しかし、反転したエコモダニズムや「AIのなかの世界」では、知識やアルゴリズムとしての科学技術やAIと実世界との間の関係が問題になる。抽象化され標準化された知識やアルゴリズムは、ラトゥールの抵抗性を示すことはあっても、行為主体性を持って他のアクターとつながることができるかという大いに疑問である。第二の問題についても、汎心論という第三の類型があった。情報をライブニッツのモナドのようなものと考えれば、必ずしも物質性にこだわらない汎心論として、心身問題に対する一つの解答になり得る。しかし、実世界で物質性を省みないことには問題が多いと思われ、「AIのなかの世界」では通用しそうでないことは既にみたとおりで

ある。

そこで注目するのが、文法において能動態でも受動態でもない中動態という動詞の形態に着目している哲学者、國分功一郎の仕事である [27] [28]。中動態は「かつてのインド=ヨーロッパ語にあまねく存在していた態」で、「もともとは能動態と中動態の対立があった」が中動態の派生形として受動態が発達したために、忘れ去られることになった [27, pp.41-42]。ところで、行為主体としての人間の自由意思は、能動と受動の対立によってもたらされる構文上の効果であるとする。本来、人間の行為を能動的な部分と受動的な部分に分けることは実際には難しい。能動と受動が対立するという文法上の慣習によって、能動的な行為を主体的な意志によるものとみなすようになったと考えられる。また、中動態と対立していた時代の能動性は、動詞で表される過程の出発点となることであって、行為の主体性は含意されていなかった [27, p.91].

例えば、ギリシア語やラテン語で、「生きる」「在る」は能動態であるが、その主語に主体的な意志はない。國分が引用する言語学者のバンヴェニストの解釈では「主語から出発して、主語の外で完遂する過程を指示している」。一方、「生まれる」「できあがる」は中動態で、「主語がその座となるような過程を表している。つまり主語は過程の内部にある」。つまり、能動態と中動態の違いは、主語が過程の外にあるか内にあるかである [27, p.88]。図3で「心」が「身体」に作用を及ぼすとき、A唯物論的な過程の理解は、主語の外側に作用を及ぼすので能動態、B観念論的な過程の理解は、主語の内部で作用が行われるので中動態として解釈することができる。

さらに、この中動態だけからなる世界を描いた哲学者が、スピノザである。スピノザの世界はすべて神の内部にあつて、「心」と「身体」も実体は一つの神である。世界のなかでの出来事は、内在原因である神が変化し、自らに影響を与えるという、回帰的で自閉的なものとみなされる。その意味で、図3のBの観念論的モダニズムと構造的に似た部分がある。大きな違いは、中動態の主語は自由意思を持って何らかの過程を支配する主体ではないことである。國分はイタリアの哲学者ジョルジョ・アガンベンの、「主体は動作を支配するのではなく、みずからが動作の起こる場所なのである」 [28, p.310] を参照する。

中動態は、能動/受動、主体/客体という関係ではなく、自らのなかで自らに作用し自らを変化させる動詞の様態である。したがって、図3に描いたような矢印の関係ではない。「心」と「身体」は一体であるので、どちらが動作の主体になっても、その結果は自らに返ってくる。そこには、権力や支配をともなう関係はないことになる。

さらに國分は、中動的な思考から意志を批判したハイデッガーが、近代技術批判を展開した著書『放下』において、原子力技術への批判を展開していることに着目する。ハイデッガーによると、現代では技術を使わないことは考えられない。それは肯定する。しかし、「それらの技術的な対象物が私たちに独占しようと要求し、そのようにして私たちの本質を歪曲し、混乱させ、ついには荒廃させることを、私たちがそれらの対象物に拒否する限り、私たちは同時に『否』とすることができる」[28. p.196]。一方では技術を肯定し、他方では技術が私たちに「独占」しようとした場合には否定することを、ハイデッガーは「放下」と呼ぶ。そして、「放下は能動性と受動性の区別の外部に横たわっている」[28. p.304]ことから、それは中動態と考えられる。

このような中動態の思考から成り立つ世界観は、具体的にどのようなものだろうか。本小論で提起した二つの課題へのヒントとして考えていきたい。

参考文献

- [1] Crutzen, Paul J. and Eugene F. Stoermer: "The Anthropocene," IGBP [International Geosphere-Biosphere Programme] Newsletter, Vol. 41, No. 17, pp.17-18, 2000
- [2] 立川雅司：“分野別研究動向（人新世）,” 社会学評論, Vol.70, No.2, pp.146-160, 2019
- [3] Horn, Eva and Hannes Bergthaller: The Anthropocene. Key Issues for the Humanities, Earthscan, 2019
- [4] Breakthrough Institute: Ecomodernist Manifesto, 2015, <https://thebreakthrough.org/articles/an-ecomodernist-manifesto>
- [5] Latour, Bruno: Love Your Monsters. Why We Must Care for Our Technologies as We Do Our Children, Breakthrough Journal, No.2, 2012, <https://thebreakthrough.org/journal/issue-2/love-your-monsters>
- [6] Haraway, Donna J.: Modest_Witness@Second_Millennium.FemaleMan@_Meets_OncoMouse²: Feminism and Technoscience. Routledge. 1997.
- [7] Fremaux, A. and J. Barry: "The "Good Anthropocene" and Green Political Theory: Rethinking Environmentalism, Resisting Ecomodernism," in Frank Biermann (ed.) : Anthropocene Encounters: New Directions in Green Political Thinking, pp.171-190, 2019
- [8] Haff, P.: "Humans and technology in the anthropocene: Six rules," Anthropocene Review, Vol.1, No.2, pp.126-136, 2014
- [9] Haff, P. K.: "Being human in the Anthropocene." Anthropocene Review, Vol.4, No.2, pp.103-109, 2017
- [10] The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems: The IEEE Ethically Aligned Design version 2, 2017, <https://sites.google.com/view/ethically-aligned-design/ws/eadv2-workshop>
- [11] The Future Society: A Global Civic Debate on Governing the Rise of Artificial Intelligence, Report 2018, <http://ai-elsi.org/archives/651>
- [12] Gサイエンス学会議共同声明：人工知能と社会（仮訳）, <http://www.scj.go.jp/ja/int/g8/index.html>
- [13] 城山英明：“人工知能とテクノロジーアセスメント—枠組み・体制と実験的試み,” 科学技術社会論研究, 16号, 2018
- [14] 江間有沙：“人工知能社会のあるべき姿を求めて—特集にあたって,” 科学技術社会論研究, 16号, 2018
- [15] 國吉康夫：“人工知能の将来と人間・社会,” 科学技術社会論研究, 16号, 2018
- [16] Andreassen, H. K., et al.: "Digitized patient-provider interaction: How does it matter? A qualitative meta-synthesis," Social Science and Medicine, No.215, pp.36-44, 2018
- [17] Yamamoto, K., et al.: "Population-based biobank participants' preferences for receiving genetic test results," Journal of Human Genetics, Vpl.62, No.12, pp.1037-1048, 2017
- [18] Nagel, Thomas: Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature is Almost Certainly False, Oxford University Press, 2012
- [19] Vetlesen, Arne Johan: Cosmologies of the Anthropocene: Panpsychism, Animism, and the Limits of Posthumanism, Routledge, 2019
- [20] Harman, Graham: The Quadruple Object, Zero Books, 2011（岡島隆佑監訳, 四方対象—オブジェクト指向存在論入門, 人文書院, 2017）
- [21] Shaviro, Steven: The Universe of Things: On Speculative Realism, Minneapolis, University of Minnesota Press, 2014（上野俊哉訳, モノたちの宇宙—思弁的实在論とは何か, 河出書房新社, 2016）
- [22] Kohn, Eduardo: How Forests Think: Towards an Anthropology Beyond the Human,

- University of California Press, 2013 (奥野克巳, 近藤宏監訳, 森は考える一人間的なるものを越えた人類学, 亜紀書房, 2016)
- [23] Haraway, Donna J.: When species meet, University of Minnesota Press, 2008 (高橋さきの訳, 犬と人が出会うとき—異種協働のポリティクス, 青土社, 2013)
- [24] 信原幸弘: 心の哲学: 新時代の心の科学をめぐる哲学の問い, 新曜社, 2017
- [25] Churchland, Paul: Matter and Consciousness, MIT Press. 1984 (信原幸弘, 西堤優訳, 物質と意識—脳科学・人工知能と心の哲学, 森北出版, 2016)
- [26] 金杉武司, 心の哲学入門, 勁草書房, 2007
- [27] 國分功一郎: 中動態の世界——意志と責任の考古学, 医学書院, 2017
- [28] 國分功一郎: 原子力時代における哲学, 晶文社, 2019