

人間とAIのあるべき関係：WEターンから考える —価値多層社会の実現に向けて—

出口 康夫

京都哲学研究所 共同代表理事

京都大学大学院文学研究科 研究科長・教授（哲学専修）



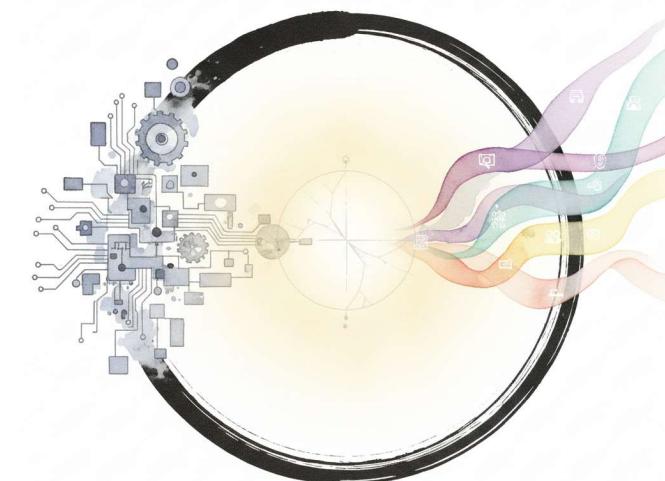
Kyoto Institute of Philosophy



価値の世紀

20世紀が「経済・科学・技術の世紀」であったのなら、21世紀は「価値の世紀」であるべきである

- 20世紀には、多くの人々が科学技術の進歩や経済発展によって人間は幸福になり、社会はより良くなり、世界はより平和になると信じ、その理想を見上げながら努力し続けてきた
- しかし、科学技術がどれほど進歩し、経済がどれほど発展しても、それだけで人を幸福にすることはできず、社会を改善することも、世界に平和をもたらすこともできない
- 同時に、人々は気づき始めている。そもそも「幸福」とは何か、社会の「善」とは何か、そして私たちが目指すべき「価値」とは何かは、自明ではない
- 「価値」という概念そのものが大きな疑問符となり、理想そのものが地球規模で消えつつある



価値とは何か



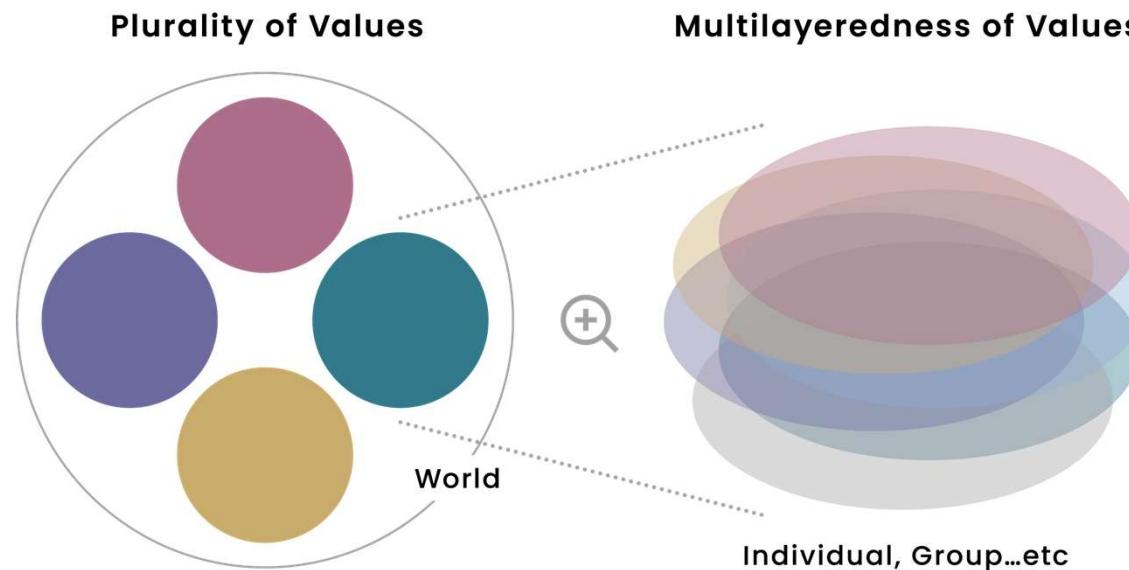
価値とは、何が重要であり、望ましく、正しいか
を判断する規範的基準である

価値は行為に方向性と動機づけを与えるベクトル
として機能する

価値の多元性と多層性

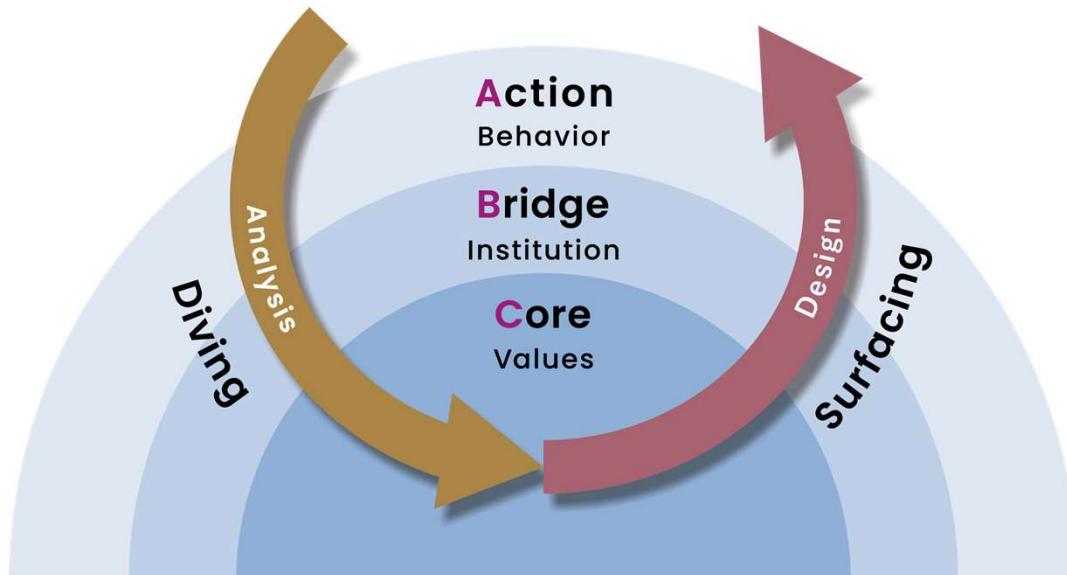
価値の問いに唯一の正解はない

- 価値は人々や文化を超えて多様であり、複数存在する（**価値多元性**）
- しかし、多様であるだけでは不十分である。個人や社会はしばしば、互いに衝突し矛盾しながらも層状に共存する複数の価値によって支えられている（**価値多層性**）



ABCモデル：価値を問い合わせるための枠組み

重要なのは、価値の根源的再検討と、実際の社会を変革する活動を橋渡しし接続することである



- ABCモデルは複雑な課題の深層構造を理解するための枠組みである
 - **A層 (Action)** : 具体的な実践や振る舞いのレベル
 - **B層 (Bridge)** : 理想と実践をつなぐ媒介層
 - **C層 (Core)** : 価値や世界観の基層
- このモデルの真の力は、「深く潜る」と「浮上する」という動的プロセスにある
- このサイクルこそが、根源的かつ持続的な変容を駆動するエンジンである

エンゲージング人文学の必要性

哲学とは、価値の提案である



- 根源的な価値の問い合わせに向き合うためには、**エンゲージング人文学に目を向けなければならぬ**。人文学は、意味と目的が問われる世界をナビゲートするために不可欠な知的インフラを提供する
- しかし、この知は社会から隔離されたままではならない。深い探究と現代の喫緊の課題をつなぐためには、人文学と実践の**再結合**が必要である
- とりわけ哲学は重要な役割を担っている。そのもっとも重要な任務は、単に分析するだけでなく、集団としての**私たちの未来を導くことのできる価値を積極的に提案する**という重要な役割がある

できなさ vs. アリストテレス的ヒエラルキー

できなさターンは、人間のかけがえのなさを、できることではなく、われわれが本来持っているできなさに見いだす

できること: アリストテレス的ヒエラルキーの伝統

知的能力が支配の形態を正当化する

アリストテレス できるものができないものを一方的に支配することは正しく、有益だとした

カントと啓蒙主義 人間同士の平等性を宣言したが、非人間である動物を手段として扱うことによりアリストテレス的ヒエラルキーを保存した

ニーチェ 「他者を支配する力への意志」を批判したが、今日の「ニーチェ主義者」の出現の道を開いた

現代の継承者 現代の「ニーチェ主義者」は、**このヒエラルキーをさらに徹底させよう**としている（暗黒啓蒙、政府系企業）

AI倫理における主人と奴隸のモデルは、**人間を主人に、人工物を奴隸に位置付ける**

できなさ

人間のかけがえのなさは、まさに私たちが**本来持っているできなさ**にある



価値の具体的提案としてのWEターン

「私」から「われわれ」へ—われわれ自身と社会を捉える新しい視点

- 東アジアの「聖なる愚者（Holy Fool）」の系譜に導かれたWEターンは、単独身体行為は本質的に不可能であるという認識を起点とする
- すべての身体的行為は、マルチエージェントシステムを形成するすべての不可欠なエージェントのネットワークを通じてのみ可能である
- すなわち行為主体とは、人間・他の存在・自然・人工物を含む「**Self-as-WE**」である
- 価値語の主語が「私」から「われわれ」に変われば、自由は関係的能力に、責任は相互的義務に、幸福は共有された成果になる
- ともにWEターンを構成し、新たな社会秩序の可能性を開く



自転車に乗るためのマルチエージェントシステム

道徳的コンパスとしての中空

「良いWE」を構成するための原則



- 中空構造とは、**特定の価値体系・個人・集団が権力や利益の中心を恒常に占有しない社会構造**である
- その中心には、一見逆説的なアイディアがある。すなわち、**中心が空であるという事実そのものが、全体を統合するという重要な機能を果たしている**
- この**空**は、あらゆる多様な価値の公正な固定点として機能し、創造的な対話と協力の「場」を創り出す
- 中空構造は、「いかにすれば、さまざまな「WE」同士が、そのさまざまな価値体系を保持したままで対話と共存することができるか」という、分断された我々の時代にとって根源的な疑問に答える可能性を秘めている

WEターンの帰結

「WEターン」の哲学は、二つの現代的ヒエラルキーモデルの批判を論理的に導き出す

暗黒啓蒙のポリティクスの否定

- 暗黒啓蒙は、悪しき中心占有型WEを想定している。すなわち、知的にできるものが中心に、知的にできないものが周縁に置かれ、両者の間には一方的な統治がある
- 中空的WEは、このような非対称性に反対し、民主主義を擁護する

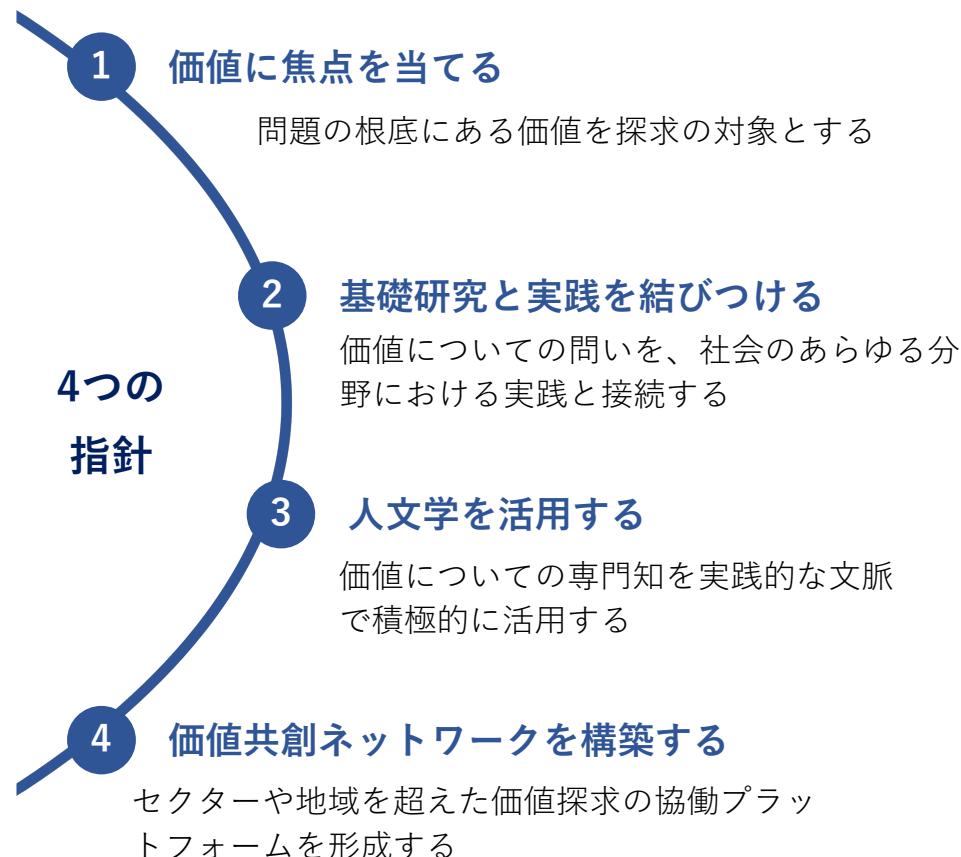
AIの主人-奴隸モデルの否定

- 「主人-奴隸モデル」もまた、悪しき中心占有型WEである。すなわち、人間が中心に、AIやロボットが周縁に置かれる
- WEターンは、フェローシップ／共冒険者モデルを提案する。すなわち、人間もAIやロボットも、リスクを引き受ける行為としての身体行為におけるフェローあるいは共冒険者として捉える



価値多層社会に向けて

グローバルな価値の風景を多様化し、21世紀における価値の語彙を豊かにする



価値多層社会に向けて

多元主義や価値の多層性を対立から創造的な緊張へと、そして社会の活力へと変えるダイナミックな社会モデル

京都会議はその第一歩である

