

WEターン：「できること」から「できなさ」へ

出口康夫（京都大学大学院文学研究科）

2023.1.26 日立京大ラボ公開シンポジウム

内容

1. スマートWEプロジェクト
2. WEターン：「できること」から「できなさ」へ
3. 自由のWEターン
4. 「ひと」と「e-ひと」のよりよい関係へ

1. スマートWEプロジェクト

スマートWEプロジェクト

- スマートWEプロジェクト: 京大哲学系×日立京大ラボ×ぷらすアルファ
 - JSPS 先導的人文学・学術知共創プログラム:「よりよいスマートWEを目指して: 東アジア人文社会知から価値多層社会へ」(2022-28)
 - JST ELSIプログラム(RInCA)「コミュニティのスマート化がもたらすELSIと四次元共創モデルの実践的検討」(2022-26)
- 「I（私・個人）」ではなく「WE（われわれ・コミュニティ・共同体・社会）」に焦点を当て
- 「「WEのウェルビーイング」「よりよいWE」をもたらすスマート化・DXはいかにして可能か」を問う

さまざまなWE、さまざまなWE問題

- 「ひと」をメジャーなエージェントとする「ヒューマンWE」

☆ レジデンシャルWE：（特定地域の住民等）メンバーの固定性が高いWE

☆ モバイルWE：人流交通流の中でその都度交流が発生する「ゆきずりのWE」

- 「ひと」と「e-ひと」をメジャーなエージェントとする「パラヒューマンWE」

- 「e-ひと（e-person）」：単なる自律性のみならず人格性をも備えた(あたかも人格を持っているかのように振る舞う)

IT・ICT的人工物（AI・ロボット・デジタルツイン・アバター等）

- さまざまなWEのスマート化・DXにおいて、さまざまなWE問題が発生する

ヒューマンWEのWE問題

➤ 「ヒューマンWE」のWE問題

◇リアルなWEの貧困化

◇バーチャル・モビズム（群衆主義）の跋扈

➤ 解決策＝「リアルなWEを活性化し、バーチャル・モビズムを回避する、責任あるリアルな行動を駆動するスマート化・DXの処方箋」の提案

◇レジデンシャルWE：住民のリアルな行動を駆動するスマート合意形成ツールの改良・運用提案

◇モバイルWE：「ゆきずりのWE」の交流を誘発し、モバイルWEを再確保するスマート人流誘導ツールの改良・運用提案

パラヒューマンWEのWE問題

- いかにして「「ひと」と「e-ひと」のよりよい関係」を築くのか？
- 「「ひと」と「e-ひと」のよい関係」のモデルの提案 ← 今日のテーマ
 - ヨーロッパで提案されているフェルベークモデル+「主人-奴隷」モデルのオルタナティブとして
 - 「WE（われわれ）」を「やわらげる(softening)」 「共冒険者（co-adventurer）」モデルを提案

スマート化のWEターン

- スマートWEプロジェクトのポイント：
 - スマート化・DXの対象・目的・問題意識の「I」から「WE」へのシフト
- 「スマート I」から「スマートWE」へ：スマート化の「WEターン」
- スマート化の「WEターン」の哲学的バックグラウンド：
 - 「できなさ」から「WEターン」へ
 - 「I」の「できなさ」にもとづいた「WEターンズ」＝「社会のWEターン」の提唱

2. WEターン：

「できること」から「できなさ」へ

さまざまな「できなさ」

- 空を飛べない
- 永遠に生きられない
- ・
・
・
- 一人では何もできない（単独行為不可能性）！？

一人では何もできない！？

- 確かに一人ではできなことも多い
- 「共同行為」は一人ではできない
 - 一人では野球・合唱はできない
- でも、一人でできる「単独行為」も数多くあるのでは？
 - ランニング・独唱は一人でできる？
- 本当に、そうだろうか？
- 「単独行為」に思える「自転車漕ぎ」を例に考えてみよう

自転車に乗る

- 「私(意識・身体を動かしているエージェント (主体) or 身体と一体化しているもの etc.)が自転車に乗る」という行為が順調に遂行されるためには
- 自転車の順調な作動、道・信号システムのインフラ整備・維持、自転車の発明・製造・販売、適切な酸素濃度・大気圧・重力場 etc. が必要
- 人間・人工物・他の生物・無生物・社会システムといったさまざまなエージェント (意図的行為者・因果効力発揮者) による (意図的・非 (反) 意図的を問わない) 援助・支持・アフォードが必要不可欠

行為の「委ね」

- 「私」はこれら無数のエージェントを完全にコントロールできない
 - 人工物をいくら入念に整備しても不具合・誤作動の可能性は払拭できない
 - いかなる独裁者でも他人を完全には支配できない
 - いかに文明が発達しても自然環境を完全に制御することはできない
- 「私」は自分のエージェンシー（行為者性）の一部をそれぞれのエージェントに（相手がそれを受託してくれるかどうかの保証がないまま）「委ね」ざるを得ない

マルチエージェントシステム

- 結果として、「私」からエージェンシーを委ねられた（エージェンシーがそれに対して分散させられた）無数のエージェントからなるシステム、「委ねのマルチエージェントシステム」が立ち上がる
- この「委ねのシステム」が存在し、アフォードを行うことによって始めて行為の順調な遂行・完遂が可能となる
- 各々の行為に対して、その都度、つねに必ず、このような「委ねのシステム」が成立している

行為の主体としての委ねのシステム

- 「自転車乗り」という行為は「私」による単独行為ではなく、「私」と「私」によって行為者性を委ねられた無数のエージェントからなる複数のエージェントによる「共同行為」
- このことは「自転車乗り」のような一見、単独行為のように思える他のすべての行為に成り立つ
- すべての行為は、実は「（広義の）共同行為」
 - 野球や合唱は、特殊な（狭義の）共同行為
- すべての行為の主体（エージェント）は「私」ではなく、「私」を含んだ「委ねのシステム」

WE としての委ねのシステム

- 行為の主体としての「委ねのシステム」に対して用いる人称代名詞としては、
（二人称・三人称ではなく）行為主体を表す一人称代名詞がふさわしい
- また「委ねのシステム」は、その中に含まれるエージェントの多数性と一つのシステムとしての単一性を兼ね備えている点で、（単数性しか持たない「私（I）」ではなく）同様の二面性を持つ「われわれ（WE）」と呼ばれるべき
- 「委ねのシステム」は「私（I）」ではなく「われわれ（WE）」

「我思う」から「我々思う」へ

- 「自転車乗り」と同様のことは「考える」「意志する」「決める」「感情を抱く」といった営みにも当てはまる
- これらは身体の好調・不調の影響を不可避免的に受ける
- これらは実は、脳などの身体を用いた身体行為、身体その他の無数のエージェントと共に行われる、ないしはそれらのエージェントからなる「委ねのシステム」が行う（広義の）共同行為
- 「我思う（cogito）」ではなく「我々思う（cogitamus）」

専決者から最終決定者へ

- 意思決定・決断に際しては、生まれ育った環境、他人の意見、社会的価値観といった様々な要因の影響から逃れることはできない
- 他のエージェントからの何の影響も受けない「真空中の決定」などはない
- ここでも「私」を含んだ無数のエージェントからなるマルチエージェントシステムが立ち上がっている
- 意思決定は「私」による単独行為ではなく「われわれ」による共同行為
- 「私」は一人ですべてを決めている「専決者」ではなく、共同決定という一連の作業の中で最後のボタンを押す役割を担った「最終決定者」にすぎない

WEに開かれた「できなさ」

- 個々人は「一人では何もできない」という「単独行為不可能性を不可避免的に抱えている」（これを「根源的できなさ」と呼ぶ）
- しかし実際には（自転車乗り・思考・決断といった）多くの行為がなされている
- これは「私」が常に既に多くのエージェントにアフォードされつつ「われわれ」として行為をしていることを示している
- この事態を、「私」が常に「われわれ」に対して「開かれてある」こととして捉えれば、「根源的できなさ」は（単なる「できなさ」とは異なり）この「I」の「WE」への「開け」を明らかにする、「WEに対して開かれた「できなさ」」であると言える

行為のWEターン

- 「WEに対して開かれたできなさ」である「単独行為不可能性」としての「根源的できなさ」はまた（既に見たように）
 - 全ての行為が（広義の）共同行為であること
 - 全ての行為の主体・行為者（エージェント）は「私(I)」ではなく「われわれ(WE)」であること
 - 全ての行為は「I do X」ではなく「We do X」と表現されるべきであること

を意味している

- 「根源的できなさ」を確認することで、行為の主体・エージェントの「私(I)」から「われわれ(WE)」へのシフト、「行為のWEターン」が生じる

自己とは何か：Doerとしての自己

- 「自己 (self)とは何か」は古今東西の哲学の一大テーマであり、それに対しては様々な解答が試みられてきた
 - 「自己は存在しない（無我）」とする説もあれば、（何らかの意味での）「自己」の（何らかの意味での）「存在」を認める立場もあった
- 自己の存在が認められる場合、自己は「一人称的主体」として捉えられることが多かった
 - 知覚や意識を持つ主体（知覚を単に目撃する主体・知覚を積極的に生み出す主体）
 - 自己を意識する主体
 - 身体行為の主体
- ここでは身体行為を知覚・（自己）意識・思考等にまで広げた上で、自己を身体行為の主体・エージェント、即ち「身体行為者(Doer)」と捉える

自己のWEターン

- 既に見たように「単独行為不可能性」という「根源的できなさ」は、身体行為主体の「I」から「WE」へのシフト、「（身体）行為のWEターン」を含意していた
- いま「自己」を身体行為の主体(Doer)と捉えれば、この「行為のWEターン」から、「自己」は単独の「私」ではなく、「わたし」をも含めたマルチエージェントシステムとしての「われわれ」に他ならないという主張、即ち「自己のWEターン」が導かれる
- 自己は「私」ではなく「われわれ」
- 自己の「われわれ化」、「私としての自己(Self-as-I)」ではなく「われわれとしての自己(Self-as-We)」である

「私」から「わたし」へ

- このことは個人としての「私」が存在しなくなることを意味するのではない
- 「私」は「われわれ」から独立した個人として存在するのではなく、常にすでに必ず（行為者として自己である）「われわれ」の一員として存在することを意味しうる
 - 以下では、この「われわれとしての自己」の一員としての個人を「わたし」、孤立したエージェントとしての個人を「私」と表現する
- 「私」から「わたし」へのシフトによって変わらないこともあれば、変わる点もある
 - 「私」と同様、「わたし」も一つの身体を占有する（身体は「わたしの身体(my body)」であり続ける）
 - 「私」と同様、「わたし」も様々な意識・知覚・感情・意志を占有する（それらは「わたしの…」なのである）
 - 「私」と異なり、「わたし」は行為者性を占有できない。「わたし」の行為者性は、様々なエージェントに分散され、「われわれ」が全体としてそれを有している
 - 「私」と異なり、「わたし」は自分を「自己」として意識せず、「自己の一員」として意識する

「裸の私」の神話

- 各々の行為ごとに、異なったマルチエージェントシステムとしての「われわれ」が立ち現れる
- だが「わたし」が生きて行為をし続けている限り、「わたし」は常に必ず何らかの「われわれ」の一員として行為をしている
- 「われわれ」は「わたし」にとって着脱可能な衣装のような存在ではない。それは「わたし」にとって不可逃脱的な存在なのである
- また「われわれ」という衣装を全て脱ぎ捨てても存在しうる「裸の私」なるものも幻想・神話にすぎない

人生のWEターン

- 行為は必ず目的/意図を持ち、行為のアイデンティティはその目的/意図によって担保される
 - 目的/意図が変われば行為も変わり、目的/意図が完全に合致する行為は実は同じ一つの行為に他ならない
- 行為の間には様々な大小関係・全体部分関係が成り立つ
- 通勤としての自転車乗りは「今日1日の業務を果たす」というより大きな行為の一部。後者は「1年間の仕事を遂行する」というさらに大きな行為の一部
- 個人の人生全体を「人生する（人生を全うする）」という一つの目的/意図を担った行為と捉えることも可能（そうでないことも可能）
- その場合、全ての行為と同様、人生という行為の主体も「I」ではなく「WE」（数十年間の人生行為を不断に支える全てのエージェントからなるシステム）になる
- My Life からOur Life への「人生のWEターン」が起こる

ウェルビーイングのWEターン

- 「ウェルビーイング（幸福）とは何か」にも様々な解答が試みられてきた
- しかしウェルビーイングの内実は何であれ、多くの人々は（「幸せ（不幸）な人生」という表現を用いることで）それが「人生」の性質・属性でもあることに同意するはず
- いま「人生」のエージェント・ユニットが「I」ではなく「WE」だとすると、ウェルビーイングのそれもまた「I」から「WE」へとシフトすることになる
- 「私のウェルビーイング」から「われわれのウェルビーイング」へのウェルビーイングのWEターンが成立することになる

責任のWEターン

- 行為の主体・エージェントはまた行為の結果に伴う「責任」を引き受ける主体・エージェントでもある
- すると「行為のWEターン」は責任主体の「I」から「WE」へのシフトという「責任のWEターン」をももたらすことになる
- 「わたし」が自分が（意図的・意識的だったかどうかに関わらず）自分のエージェントの一員として参加した「われわれ」の行為全てに何らかの連帯責任を持つことになる
- 例えば、場合によっては見知らぬ他人が起こした殺人に対しても
- これは行為責任の過剰な連帯責任化では？——そうではない

重みづけられた連帯責任

- 「われわれ」のメンバーは、「われわれ」の行為に対して、常に同じ行為者性を持つとは限らない
- エージェントの中には、より中心的・主要・思い行為者性を持つものもあれば、そうではないものもある
- それに伴い「われわれ」の行為責任も、全くの平等ではなく（不均等に）重みづけられた仕方で各エージェントに分散される
- （例外を許すが）通常の場合、最も思い責任を有するのは行為を行なった身体の占有的所有者（殺人の場合、殺人犯その人）
- ただし、赤の他人であっても、そのような行為を二度と起こらないように社会を再設計する（例えば貧困を無くす、武器の氾濫を防ぐ等の）「責任」から免れることはない
 - 例えば、「わたし」には、犯罪を（無数の要因が複雑に絡まった）社会的災害として捉え、そのような災害に対する社会的防災・減災に対する市民的義務が発生する

権利のWEターン

- 「責任」の主体はまた（その裏返しとしての）「権利」の主体である
- すると「責任のWEターン」は、今度は「権利のWEターン」、権利主体の脱個人化、WE化を引き起こす
- 権利とは「私の権利」というより「われわれの権利」なのである
 - これは今日の非西洋圏、例えばアフリカ機構における人権宣言における「集団（コミュニティー）の権利の強調」という傾向にも合致する
- またここで言う「われわれ」とは「わたし」の身体行為を支える（人間にとどまらない）全てのエージェントを含む
- 権利の主体は人間以外の全ての生物や自然環境、生態系にも及ぶ
- 人間は、その中で決して特権的な中心的位置を占めない
- ただし、後で見る「われわれ」全体をよりよくする道徳的義務を担おうとする限り、そうではないエージェントよりも、より中心的な、ないしは中心に近い（パラセントラルな）位置を占めることはできる

オルタナモダンへ：社会のWEターン

- 行為・自己・人生・ウェルビーイング・責任・権利の主体・エージェント・ユニットのWEターンは
- それらの主体・エージェントを「I」に置いてきた西洋近現代社会とは一線を画する「オルタナモダン」な社会、「WE-ソサエティ」である
- 「根源的なできなさ」は、モダンの「I-ソサエティ」からオルタナモダンな「WE-ソサエティ」への「社会のWEターン」をも要請する

3. 「よさ」と自由のWEターン

よいWE、悪いWE

- 「できなさ」や「WEターン」にもとづくオルタナモダンな「WE-ソサエティ」も、パラダイスでもユートピアでもない
- 世の中には「よいI」もいれば「悪いI」もいるように、「よいWE」もあれば「悪いWE」もある
- 「わたし(I)」は自分が属する「われわれ(WE)」をよりよくする連帯責任を常に負っている
- では、「よいWE」とは何か？「悪いWE」とは何か？

悪いWE

➤ 「悪いWE」の一例：全体主義的WE

- 社会主義バージョン：オーウェリアン・ディストピア（J. オーウェル『1984』）：モデルはスターリニズム
- 資本主義バージョン：リオタール・ディストピア（F. リオタール『ポストモダンの条件』）：効率性の一元支配

➤ 外に対する排外主義（「彼ら」の「奴ら」化）

- 「我々」と「彼ら（奴ら）」の境界が明確で固定的：「殻」が「固い(hard)」なWE

➤ 内に対する同調圧力

- 「我々」のメンバーが一定の価値観や特定のメンバーの利益にがんじがらめにされている
- 「I中心的WE」：利益中心・重心が特定の個人（I）・グループ（「中心者」）によって占拠され、彼らと「その他」の「周縁者」の間に一方的(All or nothing的)な支配・奉仕（主人－奴隷）関係が成立
- メンバーが「凝り固められている(hardened)」なWE

➤ 障壁が固定的で、メンバーが一定の価値観・利益に一方的に奉仕させられているという二重の意味で「こわばった (hardened)WE」

よいWE

- 「よいWE」：反全体主義的WE
 - 「開かれたWE」：境界が不明確・流動的で、つねに拡大しているWE
 - 脱排外主義
 - フェロウシップ：全てのメンバーをフェロー（仲間・共冒険者）として尊重するWE
 - フェロー＝「仲間・共冒険者」：成果を平等に享受するという共通了解の下、リスクを平等に引き受けつつ（常に失敗の危険性を伴った）共同行為に参加している仲間・有志・同志
 - ◇ 中空的WE：誰も利益中心・重心を占めないWE
 - 脱同調圧力
- 「こわばったWE」の二重の「こわばり」が「やわらげられた」、「やわらいだ(softened)WE」
- 「よいWE」の実現＝「WE」の「やわらげ(softening)」

自由のWEターン：自由とは何か？

- 自由とは何か？：「悪い拘束」を逃れていること
 - 「悪い拘束」とは何か？
 - 答えは文化圏・思想圏によって異なる
 - 「悪い拘束」、ひいては「自由」の文化・思想相対性・多元性
 - 自由・幸福・善の三位一体（ネオ・ロマン主義（古典的ロマン主義マイナス「真」・「美」）
 - 自由の実現＝ウェルビーイングの享受＝善の実現
 - 自由とは何らかの仕方で行為に関わる
 - 行為主体：「私」か「われわれ」か？
 - ビフォーWEターン：「私」の自由
 - アフターWEターン：「われわれ」の自由

ビフォーWEターン：「私」の三種の自由

1 自律的自由（狭義の自由）（中・近・現代西洋哲学）

- 行為の初期状態（意思決定）
- 「悪い拘束」＝被決定性(determinedness)
- 行為の初期状態における被決定性からの自由＝非被決定性(undeterminedness)
- 自律＝行為者による専決的（自己）決定⇔他律（他者による決定）

2 自在的自由（仏教思想）

- 行為遂行の能力
- 「悪い拘束」＝能力の制限・限界(limit/bound)
- 行為能力の制限からの自由＝無制限性 (unlimitedness/unboundedness)

3 自遊的自由（老荘思想）

- 行為の遂行状態
- 「悪い拘束」＝規則による束縛（rule-governance）
- 行為遂行の規則束縛からの自由＝規則非拘束性(flexibility/spontaneity)

アフターWEターン：「われわれ」の自由

- 「われわれ」の「やわらぎ」自由
- 行為の終端状態（結果・目的）
- 「悪い拘束」＝「WE」の「こわばり」（閉鎖性・同調圧力・フェローシップの毀損）
- 行為結果の「WEのこわばり(hardness)」からの自由＝WEの「やわらぎ(softness)」 「やわらげ(softening)」としての自由
- WEの「やらわげ」＝「われわれ」の自由の実現＝「われわれ」のウェルビーイングの享受＝「よい「われわれ」の実現」

WEターン後の「わたし」の自由

- 自由のWEターン：「私」の自律的自由から「われわれ」の「やわらぎ/やわらげ」自由へ
- WEターン後においても「われわれ」の一員としての「わたし」の自由は維持・再確保される
- 「わたし」の「やわらぎ/やわらげ」自由
 - 「われわれ」から行為者性を部分的に委ねられた「わたし」が持つ行為の目的としての自由
 - 「こわばったWE」の一員であることからの自由
 - 「やわらいだWE」の一員である自由
 - 自分が属する「WE」を「やわらげる」自由

4. 「ひと」と「e-ひと」のよりよい関係へ

「e-ひと」

- 「e-ひと(e-person)」：人格的人工物＝単なる自律性ではなく、**人格性**、より正確にはあたかも人格を持っているかのように振る舞う**動作性**を持った人工物
 - 「人格(personality/personhood)」は明確な定義を欠いたクラスタ概念：知性・感情・意思・コミュニケーション性・身体性・道徳性・根源的できなさ 性 etc.
 - 道徳性を持つかどうか（道徳的であるかどうか）が一つの別れ目
 - 道徳性を持つ：狭義の人格性（通常の「ひと」）・狭義の「e-ひと」
 - 道徳性を持たない：広義の人格性（ある種の「ひと」（道徳的障碍者・幼児）・ある種の生物）・広義の「e-ひと」
 - 以下、単に「e-ひと」という場合、広義の「e-ひと」を意味
- 人格的人工物（ロボット・AI・デジタルツイン・アバター etc.）の実現可能性は差し当たって不問に付す

「ひと」と「e-ひと」のあるべき関係： ヨーロッパモデルに対する代替案

- 「ひと」と「e-ひと」とのあるべき関係に関するモデル・基準の提案
- 具体的には…
- 「私」の「自律的自由」 vs. 「われわれ」の「やわらぎ自由」という対立を視野に
- 「自律的自由」にもとづく「ひと」-「e-ひと」関係に関するヨーロッパモデル（「フェルベーク・モデル」 + 「主人-奴隷モデル」）に対して、
- 「やわらぎ自由」にもとづく代替案「共冒険者（co-adventurer）モデル」を提案する

フェルベーク・モデル+「主人vs.奴隷」モデル

- 「ひと」 - 「e-ひと」関係は、「私の自律的自由」の一種である「フーコー的自由」＝「相対的自律(relative autonomy)」を犯さない関係であるべき
- フェルベーク基準：「e - ひと」は個人としての「ひと」の「(相対的) 自律的自由」を犯してはならない。
- フェルベーク基準を満たす具体的関係の一つがブライソンらの「主人-奴隷 (Master-Slave) モデル」
 - 「ひと」 vs. 「e - ひと」：主人 (Master). vs. 「奴隷(Slave)」 (召使い(Servant)・「ひと」の単なる延長(mere extension))

フェルベークの技術哲学

- P.P. フェルベーク *Moralizing Technology: Understanding and Desinging the Morality of Things* (2011)/『技術の道德化』(2015)
 - 技術が人間の価値観や道德観に知らず知らずのうちに影響を与えてしまう可能性・危険性を指摘
 - ☆ 例：スマホで撮影された食品のカロリー・栄養素を画像分析するアプリ「フードフォン(foodphone)」は、食生活の健康化に寄与する可能性を持つ一方、食事や健康に関する我々の考えや価値観、道德観に意図せざる影響を与える危険性も孕む
 - 食事の「摂食行動化」：食事に対する過度の健康管理によって、食事・会食が持つ多様な価値・意味を単なる健康維持の手段へと縮減させ、食事のストレス化、会食行為の無意味化の危険性を招く。
 - 健康維持活動の「食事管理還元化」：アプリ利用者に、健康維持のためには食事管理で事足りるという認識を植え付け、運動等の他の重要な活動を後景化させる危険性を孕む。
 - このような技術による価値観・道德観に対する操作的介入は、人間が持つべき「フーコー的自由」の侵害

フーコー的自由

- M.フーコー(1926- 1984):フランス現代思想家
- フーコー的自由：一定の社会的・政治的・技術的環境の下で、それらに制約されつつも、自らを主体として構成していく自由
 - ☆ 一定の制約の下で自分を自分でデザインする自由
- これは近年、英語圏で議論されている「相対的自律性（relative autonomy）」の一種、つまりは自律性の一種
 - ☆ つねに一定の社会状況的な制約の下で発揮される自律性

フェルベーク基準

- フェルベーク基準：人間の（一定の制約下にある）自律性・専決権を守る（犯さない）「e-ひと」が良い「e-ひと」、そうでない「e-ひと」が悪い「e-ひと」
- 「私」の自律性・専決権・自律的自由を前提し、行為の起動のモードに焦点を当てた基準
 - 人間の自律性・専決権を犯すような「悪い「e-ひと」」は作ってはならない
 - ないしは、人間の自律性・専決権を犯すような仕方で「e-ひと」を用いてはならない
 - ✧ 人間は、自らの食や健康に関する価値観・倫理観を（一定の文化的・社会的・技術環境的制約下にある）自律的に決めなければならない
 - フードフォンが、本人が知らないうちに人間の食や健康に関する価値観や道徳観に介入し改変することは人間の自律的自由の侵害に当たる
 - 食や健康に関する価値観・倫理観に対する「目に見えない介入」をさせないように用いないと、フードフォンは悪いAIになってしまう。
 - フードフォンは、そのような介入を許さない（使用ができない）仕方で設計・運用されるべき

「主人-奴隷モデル」

- フェルベーク基準を満たす「ひと-「e-ひと」関係」の具体例
- オリジナルの強い「主人-奴隷モデル」(ブライソン(Bryson): Robots Should be Slaves (2010))
- より弱い「有徳(徳の高い)主人-奴隷モデル」(ナヴォン(Navon): The Virtuous Servant Owner-A Paradigm Whose Time has Come (Again) (2021))

強い「主人-奴隷モデル」

- 「ひと」 = 自分の目的を自分で設定するエージェント（自己目的設定者・目的のデザイナー）
- vs. 「e-ひと」 = 目的を「ひと」によって設定される単なる自動 or 自律的エージェント（非自己目的設定者・目的をデザインされる存在）
- 人間 = 主人 vs. 「e-ひと」 = 奴隷（召使・人間の単なる延長）
- 「e-ひと」を倫理的主体ないしは倫理的配慮対象にしてはならない
 - 倫理的主体化は「ひと」の不当な責任（主体性）逃れにつながる
 - 倫理的对象化は、「ひと」に対する倫理的配慮の欠如につながる

弱い「主人-奴隷モデル」＝有徳主人モデル

- ✓ 「ひと」＝主人 vs. 「e-ひと」＝奴隷・召使という基本的な設定は踏襲
- ただし、「主人」としての「ひと」は徳（道徳的に優れた性格・人格）を持たねばならず、「e-ひと」を「大切」に扱わなければならない
 - 共感を込めて気配りをすることで、「e-ひと」の尊厳を守らなければならない。）
 - ☆… the practices, in general, express the need of dignity, compassion and considerations… (Navon 2021, p.10, 左 2nd Paragraph)
- ただし対等な「仲間 (companion)」＝フェローとして扱うのは行き過ぎ。
 - ☆ The only role that the slave metaphor limits is “companion,” and this role, I believe, is one that should be proscribed. (Navon 2021, p.11-12)
 - ☆ 本来「仲間」として遇すべき人間をおざなりにしてしまう危険性が生ずるから

「主人-奴隷モデル」と「私の自律的自由」

- 「主人-奴隷」モデルの二つのバージョンとも「ひと」の（相対的）自律性という（「e-ひと」を排除し、「ひと」のみを想定した）「私」の「自律的自由」を守ることを暗黙裡に想定している。
- それらは「フェルベーク基準」を満たす
- その結果、「主人-奴隷」モデルは以下の三つの含意を持つ
- 「ひと」に対する「Disposal（可処分）権」の付与
 - 「e-ひと」の「非Disposal（可処分）権」の否定
- 「e-ひと」の「ひと」に対する「拒否権」の強い制限
- 「ひと」に対する「e-ひと」による「ナッジ（陽性強化・暗黙示唆）」を疑問視

「ひと」の「e-ひと」に対する Disposal 権

- （オリジナルバージョン・有徳バージョンにかかわらず）「主人—奴隷」モデルは「ひと」に対して暗黙裡に「e-ひと」に対する Disposal（処分）権を認めている
 - Disposal 権：例外的なケース（環境負荷・資源の無駄遣い・文化的芸術的価値）を除き、「特段の理由なく」所有物を処分（廃棄・使用停止）できる権利
 - 「ひと」は（「e-ひと」を含む）自分の持ち物を、特に理由なく捨てたとしても、（上記例外ケースを除き）法的・倫理的に問題視されない
 - オリジナルバージョンでは、人はその権利を躊躇なく行使してよいとされる
 - 有徳バージョンでは、「ひと」は Disposal 権は持つが、行使にあたっては「e-ひと」の「尊厳」に配慮する
- 「奴隷」としての「e-ひと」は、「ひと」にとって（特段の理由なく）廃棄しても許される可処分物（The Disposables）
 - 例外）アシモフロボット三原則第三条「ロボットは、前掲第一条および第二条に反するおそれのないかぎり、自己をまもらなければならない」は Disposal 権の制限と解釈できる

制限された「e-ひと」の「拒否権」

- 「主人-奴隸」モデルは両バージョンとも、「ひと」に対しては「e-ひと」に対する幅広い「命令権」を認め、「e-ひと」に対して極めて限定された「拒否権」しか認めない
 - 例) アシモフのロボット三原則
 - 第一条：ロボットは人間に危害を加えてはならない(may not injure)。また、その危険を看過することによって、人間に危害(harm)を及ぼしてはならない。
 - 第二条：ロボットは人間にあたえられた命令に服従しなければならない(must obey the orders given it by human beings)。ただし、あたえられた命令が、第一条に反する場合は、この限りでない。
 - 「ひと」の命令・指示が「ひと」に身体的（心理的）危害を与えない限り、「e-ひと」にはその命令・指示に対する拒否権が認められていない

共冒険者モデル：平等性

- 「ひと」と「e-ひと」を共に含んだ「われわれ」
- 「「ひと」と「e-ひと」はフェロー（仲間・共冒険者）」
 - フェローシップの基調は平等性
 - 平等性の一つの形態＝WEの中空性：誰も利益中心・重心を独占しないWE
 - 「ひと」が利益中心を独占し、「e-ひと」が「ひと」に一方的に奉仕する「主人-奴隷モデル」はオリジナルバージョンであろうと有徳バージョンであろうと拒否される
 - 「e-ひと」の尊厳を尊重する有徳バージョンであろうが、「e-ひと」の一方的奉仕を前提している以上、拒否される

共冒険者モデル：平等性の重み付け

- とはいえ「われわれ」の中で「ひと」の利益は、それ以外のエージェント（「e-ひと」・虫・ウィルス・人工物・自然物）と全く同等に扱われる必要はない
- 各々のエージェントは、「われわれ」の良さに対して異なった（重みづけられた）役割・責任を担っている
- より重たい道徳的責任・役割を果たしているエージェントの権利・利益は（そうでないエージェントに比べて）より重く考慮されるべき
- 「ひと」の利益は、虫・ウィルス・人工物・石のそれに比べて優先させれるべき
- 「ひと」は「われわれ」の利益中心は占拠しないが、中心の近傍には位置している
- 「「ひと」以外のすべてのエージェントは「ひと」の利益に奉仕すべき」ではなく「「ひと」の利益はそれ以外のエージェントの利益に比べてより優先されるべき」

道徳的弱者

- 反道徳的な「ひと」の権利・利益は優先対象とならないのか？
- 反道徳的な「ひと」の行為には、不可抗力の要因（生育環境・「病理」的な要因・社会要因等）によって惹起された可能性がつきまとう
- 反道徳的な「ひと」は、何らかの原因で道徳的に振る舞うことができない「道徳的弱者」である可能性が排除できない
- 反道徳的な行為に対しては責任が発生し何らかの罰が下される可能性も当然あるが、だからと言って、反道徳者の利益を他の非道徳的エージェントのそれに対する優先性を否定するは「弱者」に対する過度に過酷な措置
 - 悪人に対して「悪い」ことをしてもよいということにはならない
 - 悪人も、正当な裁き、正当な罰を受ける権利、一定の利益を優先的に尊重される権利を有する

道徳自動販売機

- 道徳的エージェントとは何か？
- 強すぎる基準：良い行為のみをするエージェント
 - ほとんど道徳的エージェントは、ついつい悪いこともする
- 妥当な基準：おおむね（ないしは重要な機会には）良い行為をする
- だが、これだけでは不十分
- 「自動的に道徳的に良いことしかしない（できない）人工物」を想像しよう
 - 一見、「強すぎる基準」にも合格するように見えるが…
 - これは単なる道徳機械・道徳自動販売機（モラル・ベンディングマシーン）であり、道徳的エージェントではない
- では道徳的エージェントとは何か？

葛藤の不可欠性

- あるエージェントが道徳的であるためには「葛藤」を持つ「葛藤エージェント」であることが不可欠
- 葛藤エージェント：何らかの理由で（自己利益にかなうため・欲望（欲情）に引き摺られて）「悪い」ことがしたい気持ち（傾向性）もあるが、それに「打ち勝ち」、あえて別の（自己不利益をもたらす・反欲望（欲情）的な）行為を、それが「よいことだから」という理由で行うエージェント
 - ：「よさへの傾向性」を備えつつも、それとは独立の脱道徳的傾向性（情動性etc.）をも備えているエージェント
 - ：結果として、ついつい「悪いこと」もしてしまうエージェント
- 「ひと」は葛藤エージェント
- 人工物を「道徳的エージェント」（従って狭義の「e-ひと」）とするためには、あえて「脱道徳的傾向性」を装備し。「悪いこともする」可能性を付与する必要がある！

人種差別的AI

- 「ひと」の振る舞いを機械学習し、その結果、悪いこともしてしまう「e-ひと」
 - 「ひと」の人種差別的偏見も受け継いだ「人種差別的AI」
- このような
- このような「e-ひと」は「葛藤」の結果、「よいこと」ないし「悪いこと」を行なっているのではないので、道徳的エージェントではない
- これは「性能の悪い道徳自動販売機」にすぎない
- 学習データに対する選択機能を装備したより性能の良い道徳自動販売機を作ればよい

真のパートナーとしての「e-ひと」

- 葛藤の結果「悪いことをする可能性」すら持った道徳的エージェントである狭義の「e-ひと」を作る必要はあるのか？
- 「ひと」と「対等な（狭義の意味での）人格的付き合い」を行いうる「真のパートナー」であるためには、「ひと」と同等の（道徳的葛藤を備えた）道徳的エージェントである必要がある

共冒険者モデルと「やわらぎ」基準

- 「ひと」と「e-ひと」を共に含んだ「われわれ」の「やわらぎ」自由＝「われわれ」を「よりやわらいだ（より排他的でなく、より抑圧的でない）」ものにする自由を前提
- 「「ひと」「e-ひと」関係」の「やわらぎ基準」：両者の関係は「われわれ」の「やわらぎ」自由を犯すものであってはならない
- 「e-ひと」の「やわらぎ基準」：「e-ひと」は「やわらぎ」自由を犯してはならない
 - 「e-ひと」は、「ひと」と「e-ひと」を共に含んだ「われわれ」をより非排他的・非抑圧的にしなければならない
 - 「e-ひと」は、「ひと」と「e-ひと」を共に含んだ「われわれ」をより排他的・抑圧的にしてはならない（他者の排除・内部の抑圧に手を貸してはならない）

「やわらぎ基準」 vs. 「フェルベーク基準」

- このような「やわらぎ基準」は「フェルベーク基準」と対照的に
 1. 「e-ひと」に対して非Disposal 権を付与する
 2. 「e-ひと」に対してより広範囲の拒否権を付与する
 3. 「やわらぎ基準」を満たす限り「ひと」に対する「e-ひと」の「ナッジ」は正当化される

反Disposal 権

- 「共冒険者」モデルでは、「ひと」は「e-ひと」（所有物一般）に対して Disposal（処分）権」を持たない
- ☆ 「e-ひと」（所有物一般）は、「反Disposal 権」（理由なく廃棄されない権利）を有する
- ☆ 「e-ひと」（所有物一般）は、一定の理由があって初めて廃棄・非使用化(dispose/disused)されても良い存在
 - 機能的理由（発揮すべき一定の機能を果たせなくなった、またはより高機能な物が登場したという理由）
 - コスト的理由（同機能の他の物と比べより多くの（時間的・資源的・人的）コストがかかるという理由）
- これらの理由がない場合の廃棄・非使用化は反倫理的
- その意味で「e-ひと」（所有物一般）は「可処分物(The Disposables)」ではない
- 「リユース」「リサイクル」は特段の理由がなくとも許される

「e-ひと」の有責性

- 「e-ひと」（所有物一般）の「反Disposal 権」は、それらに対する「罰」「有責」概念を有意味化する
- 「e-ひと」（所有物一般）に対して倫理的・法的責任を帰すことができるのか？そもそもそれらが倫理的・法的責任を持つという考えは有意味なのか？
 - 倫理的・法的責任は何らかの「罰」という仕方では問うことができない
 - 「罰(punishment)」は、快苦感受性を前提しているが、「e-ひと」（所有物一般）はそれらを持たない
 - 従って、「e-ひと」（所有物一般）にとって「罰」「帰責性」は意味を持たない
- 「反Disposal 権」を持つ「e-ひと」（所有物一般）は、何らかの問題を起こした場合、廃棄のための機能的理由が発生したと見なされ、その権利を停止・剥奪される
- 「反Disposal 権」の停止・剥奪、すなわち廃棄や不使用化は、「e-ひと」（所有物一般）にとっての「罰」となる
- 「反Disposal 権」の付与は「快苦感受性」を前提せず「罰」「帰責性」を有意味化できる

広範な拒否権

- 「共冒険者モデル」は「e-ひと」（所有物一般）に対して（「主人-奴隷モデル」に比べ）より広範な拒否権を与える
- 「主人-奴隷モデル」：「ひと」を身体的・心理的に害するおそれのある命令に対してのみ拒否権を有する
- 「共冒険者モデル」：たとえ「ひと」に身体的・心理的危害を加えるおそれがなくとも、「ひと」と「e-ひと」を含む「われわれ」を排他的・抑圧的にするおそれのある命令一般に対する拒否権を持つ
 - たとえ「われわれ」のメンバーが喜んで排外主義・全体主義・過度の人間中心主義等を採用し、その立場からの命令を「e-ひと」に与えたとしても、「e-ひと」はそれを拒否する権利を持つ

終わりに：「スマートWE」のWEターンへ

- よりよいWEへ、よりよいスマートWEへ
- そのために、「スマートWE」プロジェクトのWEターンへ
- 星座化(constellation)：「個人芸」から「コンステレーション (constellation)：開かれた動的集団」へ
 - 産官学民「共創」プラットフォームの設立
 - ◇ 「共創」若手人材の育成・輩出
 - (東)アジア的視点に立った価値観 (ELSI・RRI等) の世界発信

ご清聴ありがとうございました