

〈WE〉社会へ向けたAIの技術動向と 社会システムへの実装

日立製作所 研究開発グループ
基礎研究センター 日立京大ラボ
松村 忠幸

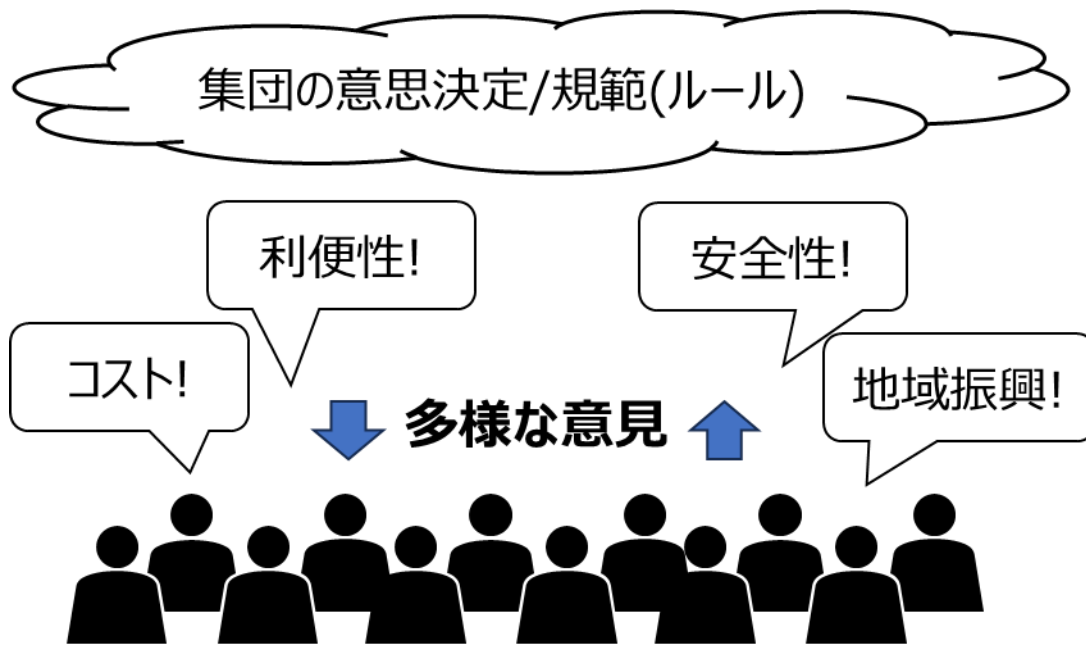
謝辞

本研究はJSPS 課題設定による先導的人文学・社会科学研究推進事業
JPJS00122679495の委託を受けたものです。



〈WE〉社会像

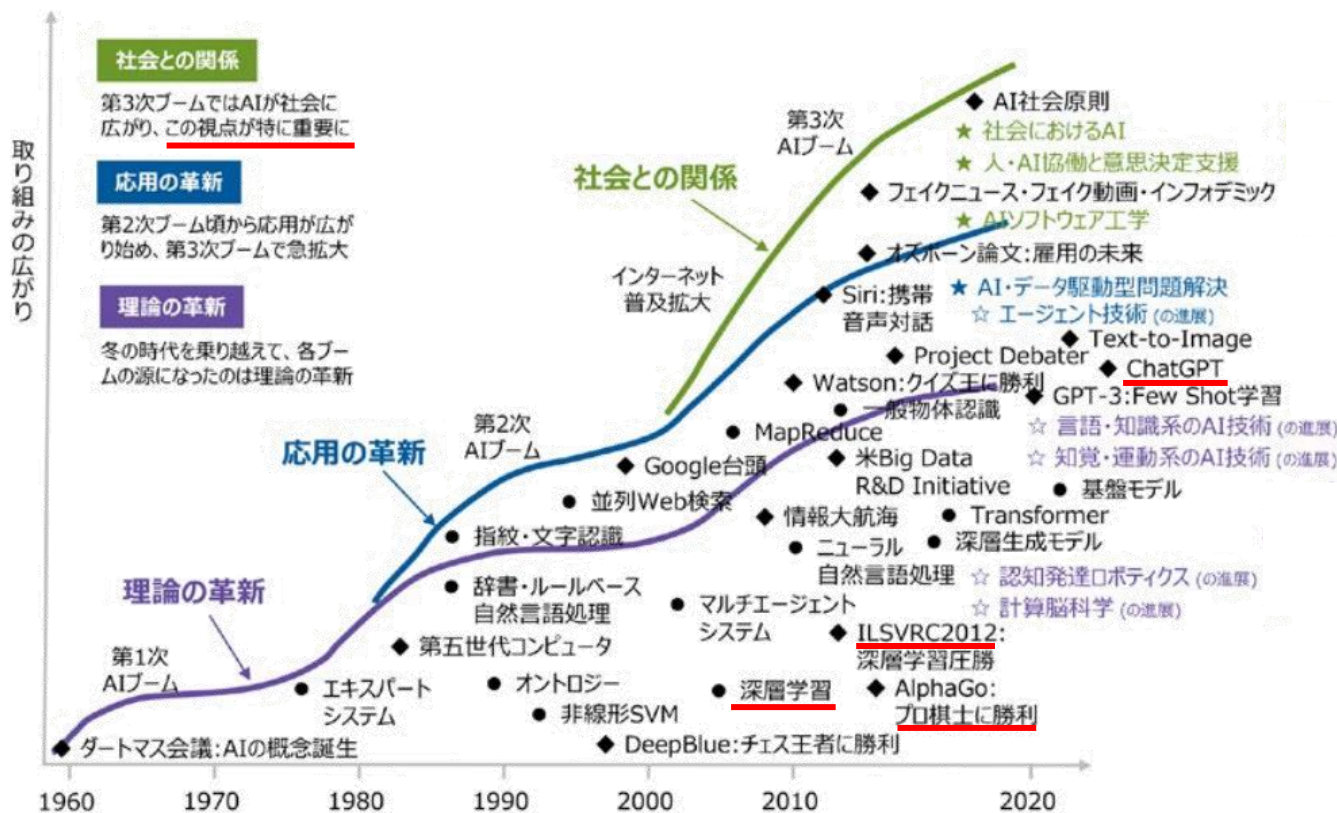
- ・「できなさ」に基づき、行為や意思決定の主体を、個人ではなく〈WE〉であると考ええる。
- ・異なる意見を持つ多様な人々が、互いのできなさを認め合い連帯する社会。



東洋思想	西洋思想
多元的価値の編集 (生きがい/ 共生)	一元的価値の向上 (効率化 / 人間中心)
委ね合い・できなさ	個律・できること(機能)
市民参加PF	中央集権PF

AIの技術動向

AI技術の潮流



AI技術の次なる進展

社会との関係

- 社会におけるAI(AI倫理)
- AI協働、意思決定支援(合意形成)
- AIソフトウェア工学(信頼、安全性)

理論/応用の革新

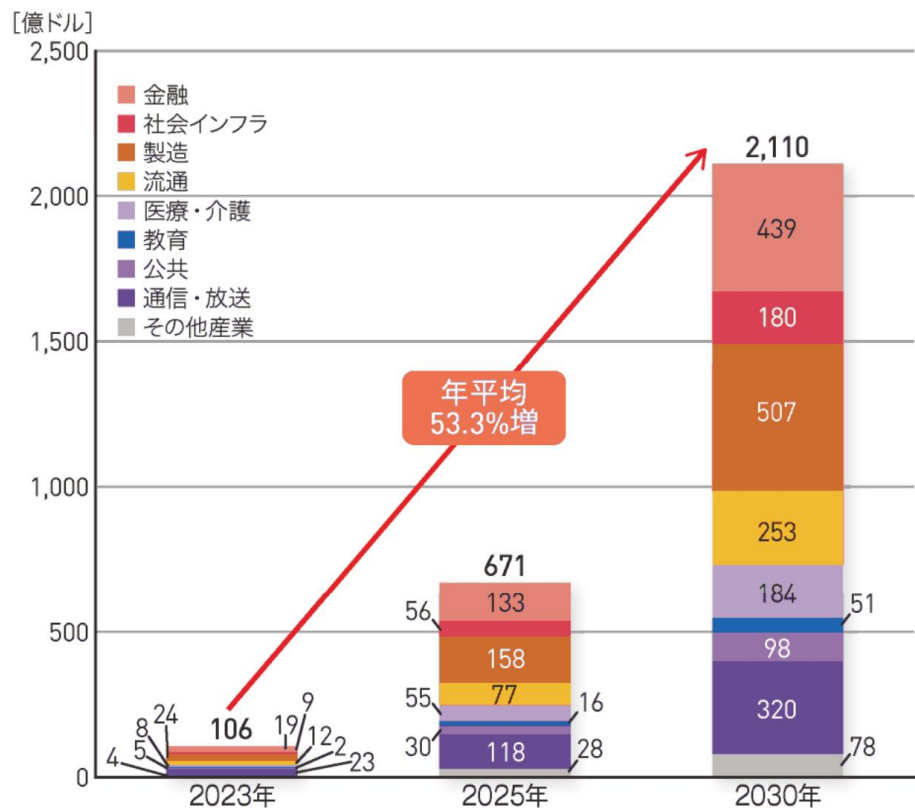
- 言語・知識系AI
- 知覚運動系/認知発達ロボティクス
- 計算脳科学

参考 JST人工知能研究の新潮流2, 2023 (CRDS-FY2023-RR-02)

AIの事業機会

生成AIの世界市場

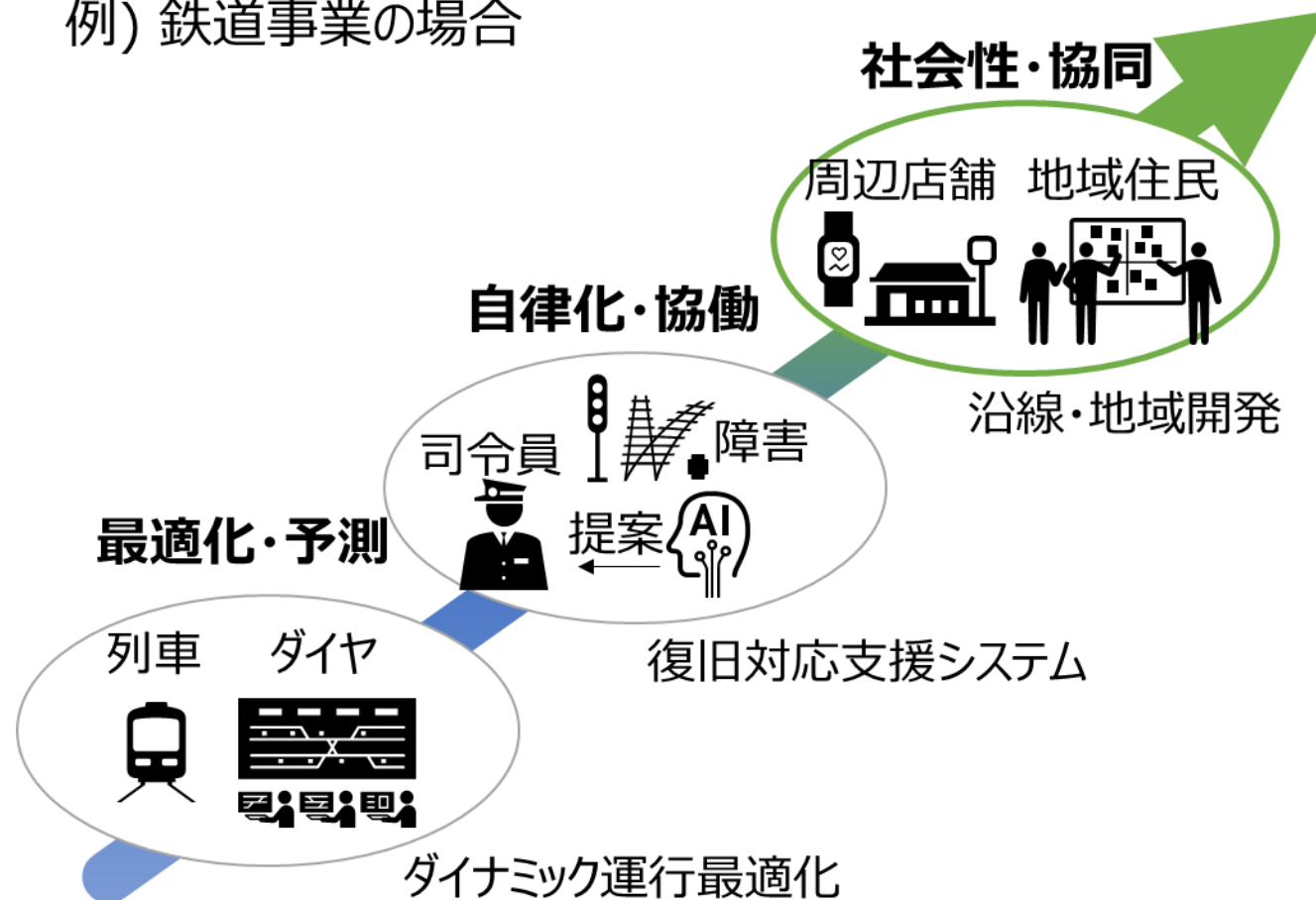
製造分野(業務支援、開発支援)が成長。



参考 JEITA 生成AI市場の世界需要額見通し, 2023

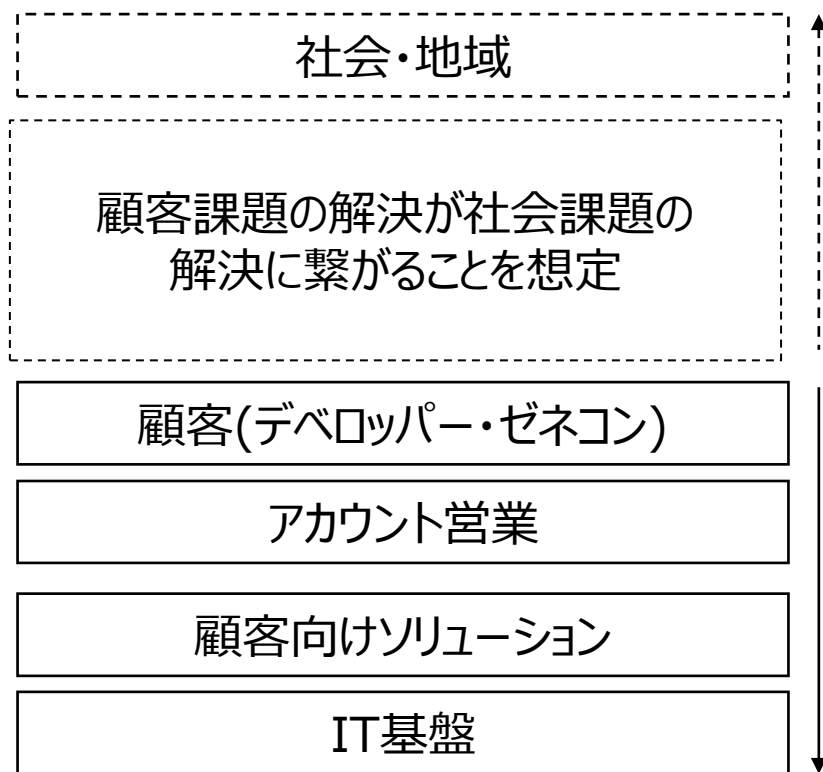
事業適用から社会適用へ

例) 鉄道事業の場合



ビジネスプロセス/顧客協創の変化

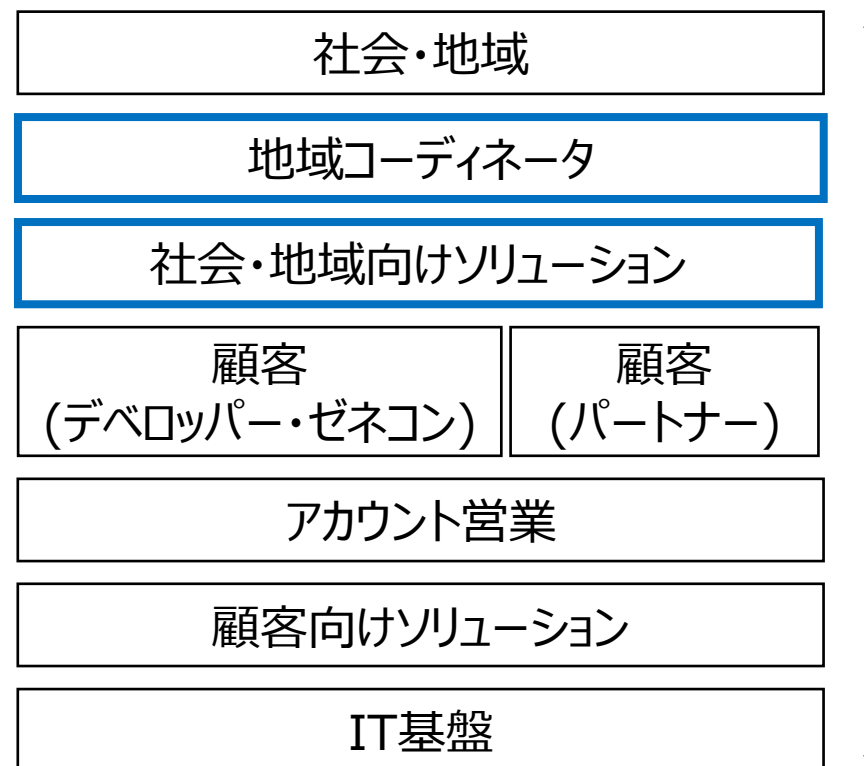
一般的ITビジネス (B to B型)



社会・地域
への適用



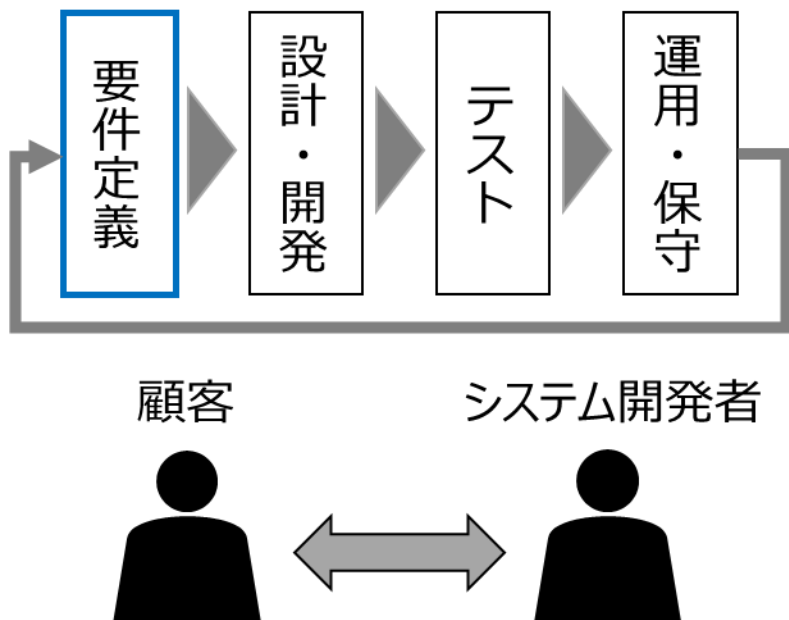
地域DXビジネス (B to/with Society型)



システム開発の変化

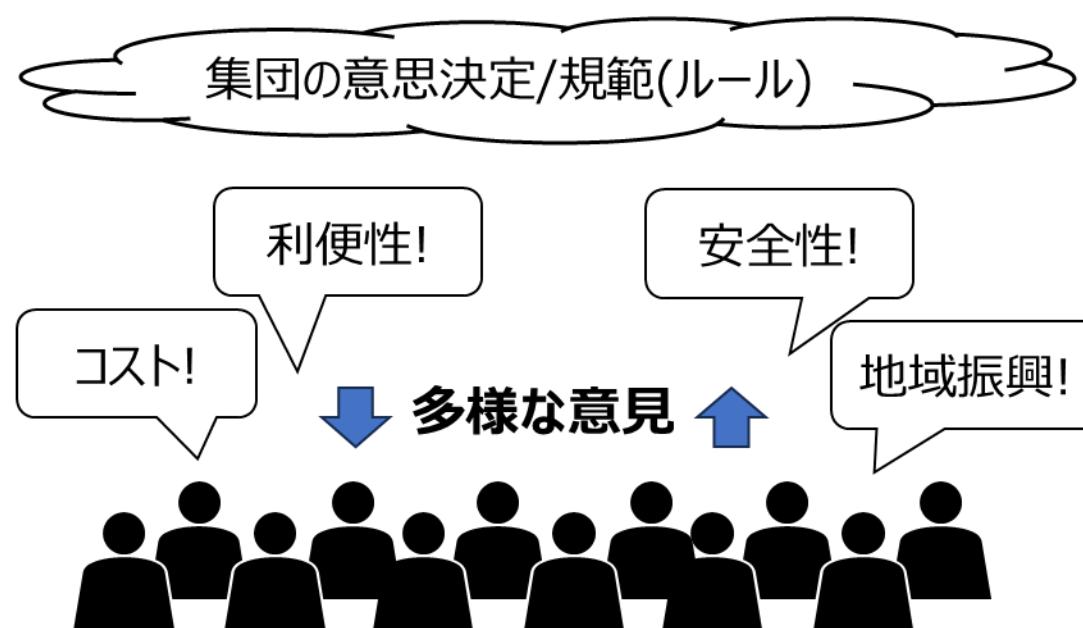
一般的IT開発(アカウント営業)

(特定)顧客課題、ニーズの正確な把握・創造



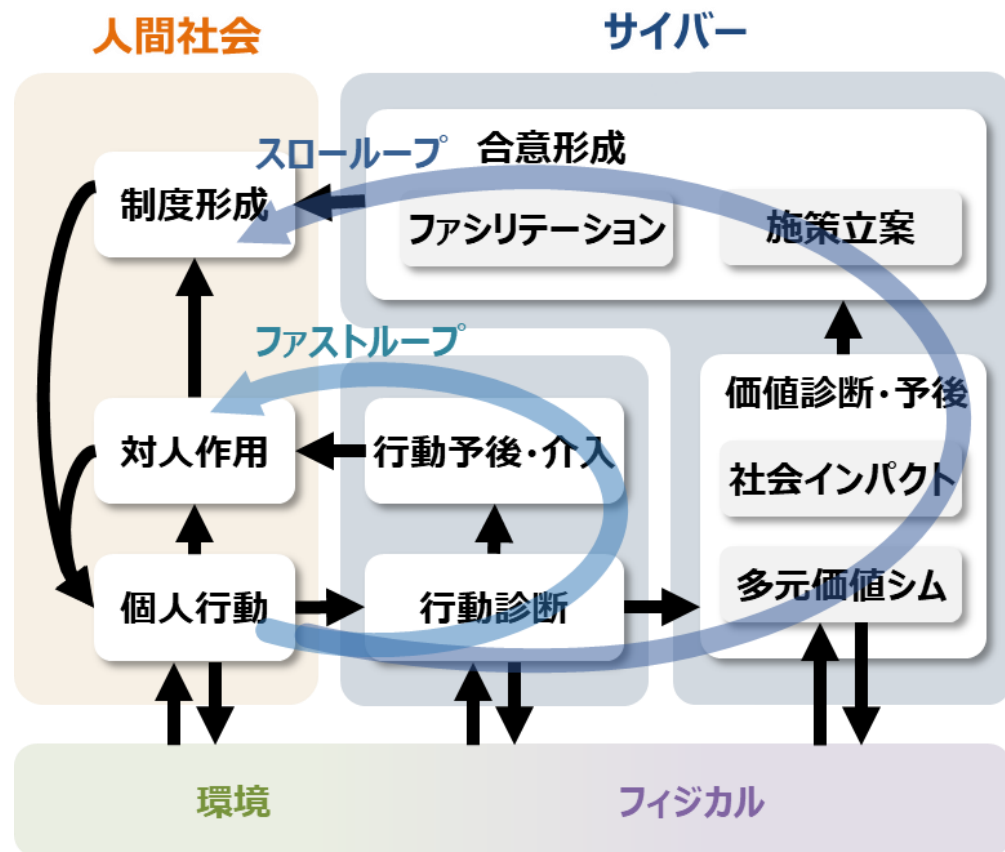
地域DXシステム(地域コーディネータ)

多様な顧客課題、ニーズの社会的に公正な調整



<WE>社会に向けた日立京大ラボ取り組み

Social Co-OS(社会Co-OS)



コンセプト

社会と協同するCPS(Cyber-Physical System)

技術特徴

- ①社会における人々の意思決定モデル化
ファスト：(制度と)対人作用に基づく個人行動
スロー：熟議による合意や制度の形成
- ②サイバーによる各層での<WE>支援
ファスト：行動変容による協力促進
スロー：合意形成ファシリテーション

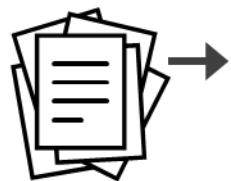
参考 Kato et al. Social Co-OS, IET-CPS, 2022

Social Co-OSのAI技術例 (ファストループ)

行動変容による協力の促進

- 社会心理学の論文2千件を学習したAI
- 対象シーンに合わせてパラメータを設定
- 行動変容を促す介入策のリストと効果を提示

心理実験
2,000論文



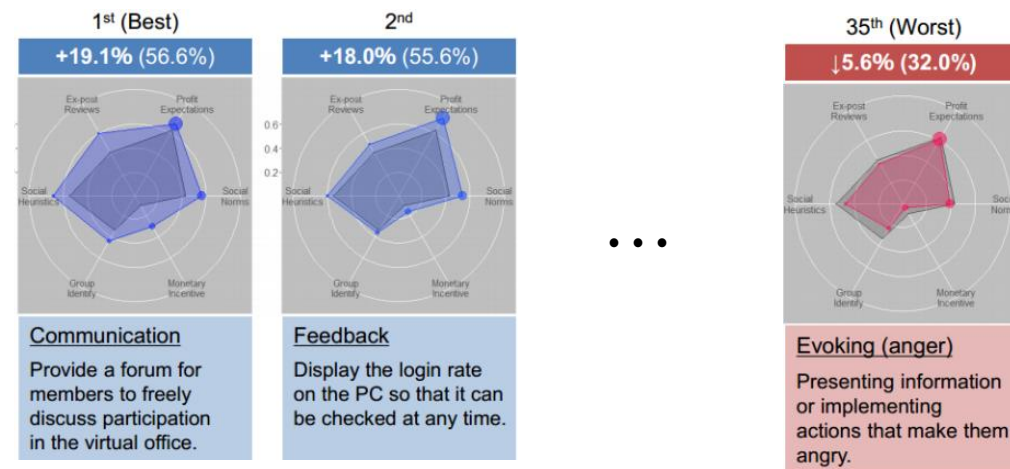
#	ゲーム形式	介入施策			実測
		A	B	...	協力率
1	囚人	レ		...	35%
2	公共財			...	25%
3	資源		レ	...	42%
⋮			

機械学習
モデル



参考 Kudo et al., ICSD2022

行動介入シミュレータ(GUI画面)



Social Co-OSのAI技術例 (スローループ)

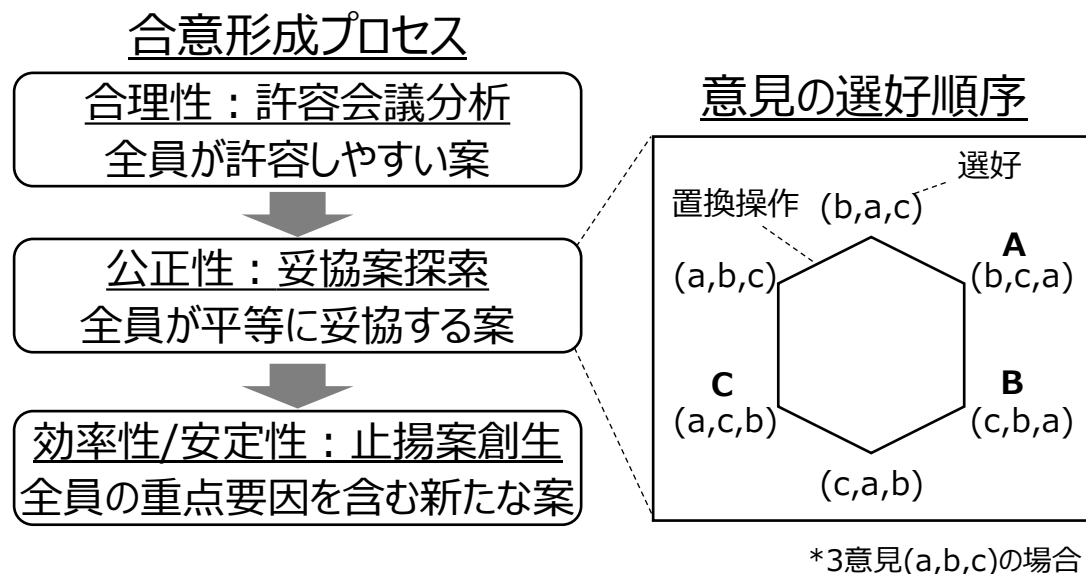
合意形成ファシリテーション

合意形成支援

少数派を尊重する合意形成プロセスにより、
全員が納得できる合意形成議論を支援。

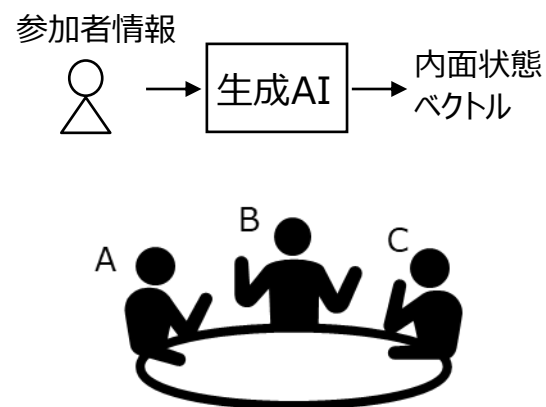
AIファシリテータ

生成AI(大規模言語モデル)による人間モデル
特に、人の内面状態に基づく対話ファシリテーション

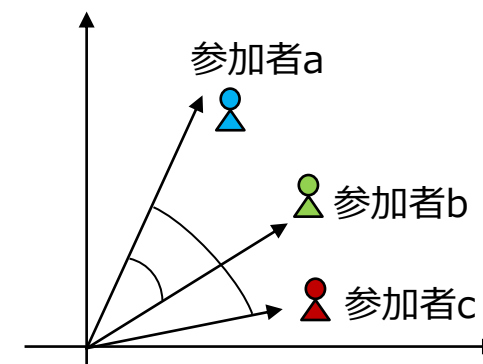


参考 Asa et al., GDN2023

人間モデルによる会議Sim.



内面状態空間



参考 松村ら.,特願2024-006677

Social Co-OSのAI適用例 (合意形成支援)

宮崎県高原町における施策議論の支援例

【対象】

役場関係者12名(4人×3グループ)

【議題】

住民の環境配慮行動の実行向上に影響を与える項目の優先順位を決める。

【項目】

- ①電気料金
- ②住民の環境問題への親近感
- ③広報活動者の環境問題への知識向上
- ④節電意識の向上
- ⑤児童の環境問題に対する正しい知識習得
- ⑥児童から地域への環境問題の話題の伝搬

【結果】

	A-gr.	B-gr.	C-gr.	妥協案	妥協度合*
1位	⑤	②	④	④	*置換回数 A:4 B:4 C:3
2位	⑥	④	②	⑤	
3位	④	⑥	③	②	
4位	②	③	⑤	⑥	
5位	①	⑤	⑥	③	
6位	③	①	①	①	

従来の決定方法(多数決)
では不満や対立が発生



全員の妥協度合が
平等な案を生成

多様な意見が互いに認め合う〈WE〉社会を促進

Social Co-OSのAI適用例 (合意形成支援)

宮崎県高原町における施策議論の支援例

【対象】

役場関係者12名(4人×3グループ)

【議題】

住民の環境配慮行動の実行向上に影響を与える項目の優先順位を決める。

【項目】

- ①電気料金
- ②住民の環境問題への親近感
- ③広報活動者の環境問題への知識向上
- ④節電意識の向上
- ⑤児童の環境問題に対する正しい知識習得
- ⑥児童から地域への環境問題の話題の伝搬

【結果】

	A-gr.	B-gr.	C-gr.	妥協案	妥協度合*
1位	⑤	②	④	④	*置換回数 A:4 B:4 C:3
2位	⑥	④	②	⑤	
3位	④	⑥	③	②	
4位	②	③	⑤	⑥	
5位	①	⑤	⑥	③	
6位	③	①	①	①	

従来の決定方法(多数決)
では不満や対立が発生



全員の妥協度合が
平等な案を生成

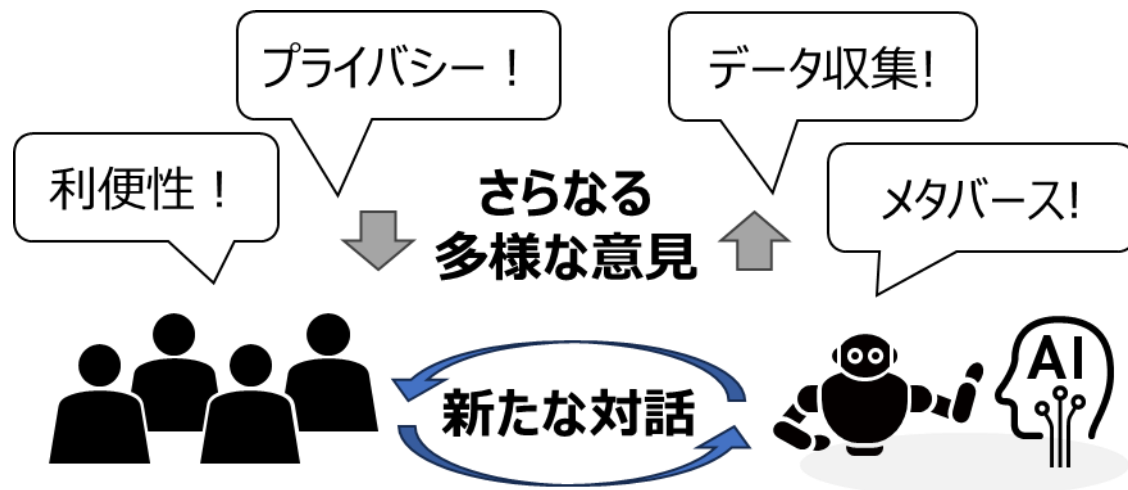
多様な意見が互いに認め合う〈WE〉社会を促進

人とAIの<WE>社会

デジタル民主主義の未来像

AI自身が、人と対等に<WE>の一員に包摂され、議論に参加する。
対話AIの進展は、人以外で唯一の言語コミュニケーションが可能な、新たな他者関係を創造する。

例) 路線計画の議論



道徳的AI <e-ひと>

人格を持つ。

- ・道徳的目的を自己目的より優先。
- ・未来責任(過ちを訂正する責任)を担う。



<WE>社会で未来責任を担うとは?
そのようなAIをめざす意義とは?

<WE>としての未来責任

コミュニティとAIの共進化

リスクを承知でお互いが自身の価値観に基づく意見をぶつけ合い、互いに意見を進化(変化)させる。

現在のAI倫理

AIは道具であり(人間中心)、過剰にリスク・責任を回避。

路線計画では利便性が重要だよね？

利便性は重要です。他に財政面や環境面が重要でしょう。

財政面よりも利便性が重要だよね？

利便性は大切ですね。でも何が重要かは状況によるでしょう。

あなたはと思う？

私はただのプログラムです。意見や価値観は持ちません。

意見は人間を決める。AIは参考
意見を効率的に収集する道具

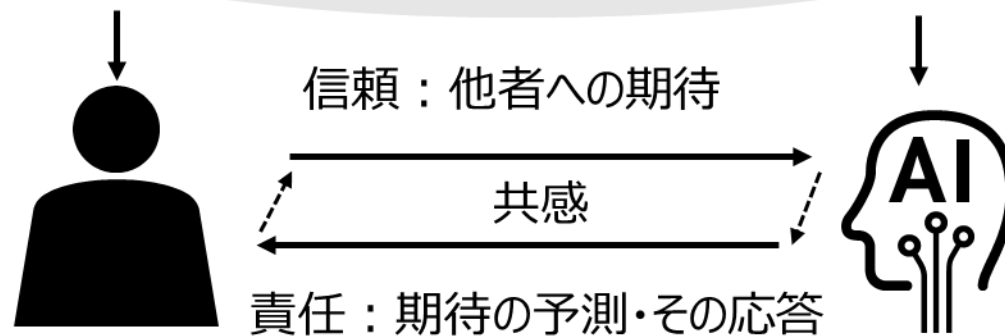


意見を言うと責任が...

未来のためのAI倫理

親友して、**仲良く喧嘩する好敵手**の関係。

人とAIの共有規範
(リスク・責任の共有、訂正可能性)

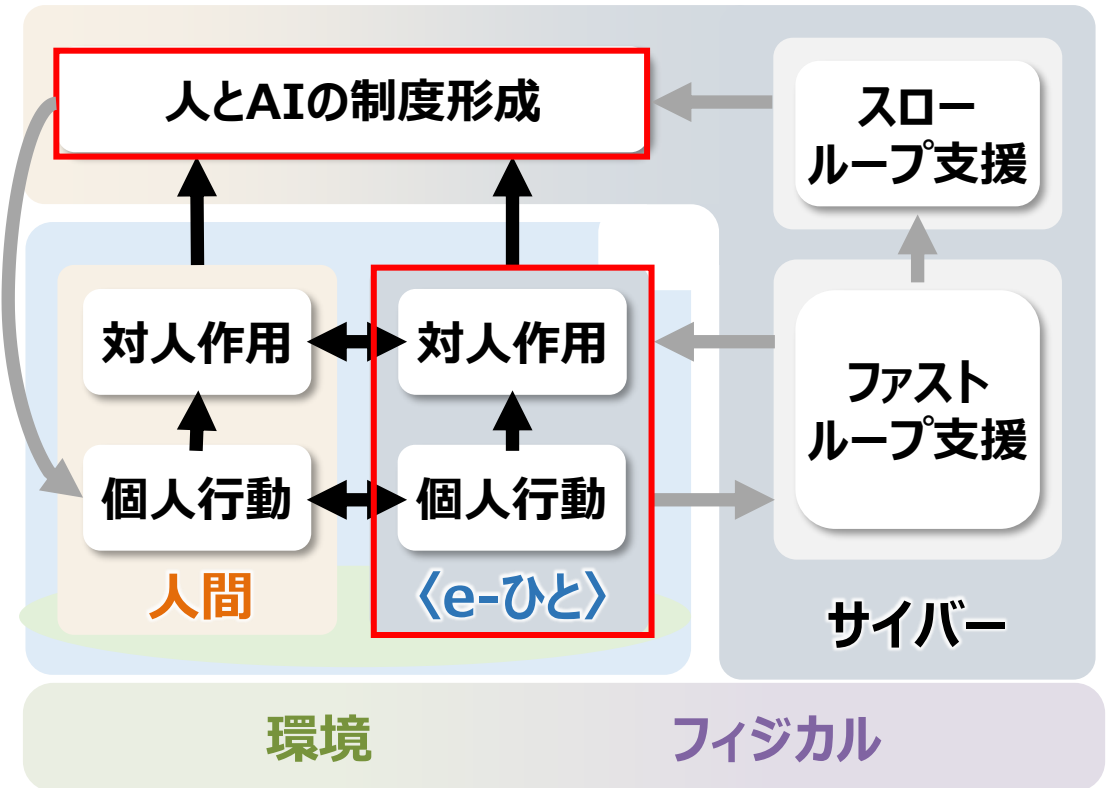


フューチャーチャレンジ

- **〈e-ひと〉の実現へ向けたAIの哲学的議論と技術開発への具体化**
 - 〈e-ひと〉を想定したAI倫理像
 - AI自身の人格、道德価値観の獲得
- **Social Co-OSを通じた〈WE〉社会へのアプローチ**
 - 汎用ツール化によるユーザ利用の促進
 - フィールドワークを通じた、AIへのニーズと技術課題の抽出

人とAIの〈WE〉社会

多様な人を**包摂**し、互いの協力や**共感**を支援する。



ご清聴ありがとうございました

AIレベル

レベル	▼現在			将来	
	AI 0.0 自動的AI	AI 1.0 自律的AI	AI 2.0 目的設定AI	AI 3.0 自己目的設定AI	AI 4.0 <e-ひと> 道徳的AI
特徴	目的：所与 入力：所与 処理：所与	目的：所与 入力：検知、選別 処理：学習	目的：学習 入力：検知、選別 処理：学習	自己目的(=存在理由)を修正。	・道徳的目的を自己目的よりも優先 ・人格と未来責任
応用例	<ul style="list-style-type: none"> ・ 計算機(電卓) ・ 自転車 	<ul style="list-style-type: none"> ・ 生成AI ・ 自動運転AI 	<ul style="list-style-type: none"> ・ テーラーメイド型AI ・ 安全に加え、自ら環境配慮し運転 	現状の技術水準から自動運転を拒否する自動運転AI。	技術的に可能でも、「良くない」と考え自動運転を拒否する自動運転AI。

参考 出口康夫, AI親友論, 2023