「人と情報のエコシステム」研究開発領域

研究開発プロジェクト事後評価報告書

令和５年５月

研究開発プロジェクト名：「PATHーAI：人間-AI エコシステムにおけるプライバシー、エージェンシー、トラストの文化を超えた実現方法」

研　究　代　表　者　：中川裕志（理化学研究所革新知能統合研究センター　チームリーダー）

実　施　期　間　：2020 年 1 月～2023 年 3 月

Ａ．総合評価

十分な成果が得られたと評価する。

AI システムが社会において浸透し円滑に利活用されるためには、AI が人や社会との「なじみ」があることが重要である。本プロジェクトでは、多様な AI 応用に関して、「なじみ」ないしトラストの観点から分析し、幅広い人々に AI の将来像に関する知見、指針を与えることを目的としたものである。研究のアプローチとしては、日本チームと英国チーム（The Alan Turing Institute、The University of Edinburgh 等）との共同研究体制にて推進し、人々の AI 感について、歴史的、文化的背景や世代的な背景からの分析を行った。

「AI と文化」「パーソナル AI エージェント（PAI Agent）」「トラスト」「ガバナンス」の４つの視点を切り口にし、12 もの多岐にわたる項目について実施し、それぞれの項目について興味深い研究成果を多数創出したと評価できる。特に、PAI Agent のトラストとガバナンスに関する倫理的、法的、社会的な調査、分析および考察は、政策や企業活動へのインパクトも大きく、また今後のさらなる調査と研究の重要な参照点となり得ると思われる。

個別の研究プロジェクトとしては大きな成果が得られたと考えられる一方で、AI と文化について、技術受容の面で英国と比較して留意すべき点や日本の独自性はあったのかどうか、文化の差異についてどのように向き合うべきか、また、国際ガバナンスの在り方がどのようにあるべきか等については、実質的な日英連携により踏み込んだ検討が今後深められることを期待したい。

また、AI 倫理規範の背景調査での包括的な調査や、インタビュー、WEB アンケートなどの社会調査で得られた知見を、例えば、技術開発者や AI 技術の利用者が参照してセルフチェックできるような成果物があれば、技術開発や社会実装をする上で、様々な技術開発に対して、本プロジェクトの豊かな成果・知見をもっと還元できると思われる。

本プロジェクトの問題意識や研究成果は、新たな政策や法制度等の検討のための材料として今後大いに活用されていく可能性があるが、対象が研究開発や社会情勢の変化のスピードが速い分野であるため、引き続き、継続的な検討、並びに積極的な情報発信に期待したい。

Ｂ．項目評価

Ⅰ．研究開発プロジェクトの研究開発内容とその成果について

1. 目標の妥当性

十分妥当であったと評価する。

近年著しい進展をみせる AI 技術にとって、個人や社会・経済、法制度等との「なじみ」があることが、AI システムが社会において浸透し円滑に利活用されるために重要となってくる。本プロジェクトでは、人々の AI 感について、歴史的、文化的背景や世代的な背景から分析し、パーソナル AI エージェント（PAI Agent）の設計提案に至る研究計画を実施したものであり、AI システムの本格的な普及期を迎えている今日においてタイムリーな目標設定であったと評価する。特に、死後における PAI Agent の利用の問題、トラストの問題をはじめ、AI システムの法的・社会的・倫理的課題の研究と、その成果を通じた豊かな AI エコシステム構築への提言は現在の、そして今後の日本のみならず世界にとって非常に重要であり、目標設定は十分妥当であったと評価する。

また、本プロジェクトは、COVID-19 の世界的な大流行が発生している最中に行われたが、そのような状況下においても、英国チーム側と COVID-19 対策の現状と課題について議論を重ね、日本の「COCOA」やイングランドとウェールズで利用されている「NHS COVID-19」等のデジタル接触追跡アプリを対象として調査を実施するなど、社会情勢に応じて柔軟な計画修正が行われた点も評価に値する。

２．研究開発プロジェクトの運営・活動状況

十分適切になされたと評価する。

英国チームとの共同研究者の体制等も含めて、十分な事前準備が行われて研究に臨まれたことと推察される。プロジェクト開始直後に COVID-19 の世界的流行という困難な状況に陥った中で、オンラインを活用して、定期的なミーティング、国内外での研究会やシンポジウムを開催するなど、適切かつ積極的なプロジェクト運営がなされたと思われる。

また、本領域の橋田プロジェクト（「パーソナルデータエコシステムの社会受容性に関する研究」）におけるデータオーナーシップやプライバシー保護、ムーンショット目標 1 の石黒プロジェクト（「誰もが自在に活躍できるアバター共生社会の実現」）のアバター活用など、既存の開発研究との関連を念頭に置きながら、12 もの実施項目を着実に進められた。実施項目は多岐にわたっていたが、各パートでそれぞれ、代表者やグループリーダーを中心に精力的な研究活動が展開されたと評価する。

一方で、現状認識が主なアプローチとなっていたが、プロジェクト期間の終盤で chatGPT が社会を席巻したことに代表されるように、変化のスピードが速い分野であるため、願わくは今後の見通しや方向性への考察も欲しかった。例えば、調査などにおけるスピード感や変化を捉える工夫、シミュレーションのような工夫などがあれば、また違った視点を持った調

査結果が創出されたかもしれない。また、本プロジェクトの記述では多面的な「AI」を一括りにして扱うような傾向が見受けられるが、例えば、技術的・応用的の両面から分野分けや段階分けのような捉え方もあれば、尚良かったのかもしれない。

３．研究開発プロジェクトの目標の達成状況および研究開発成果

一定の成果は得られたと評価できる。

本プロジェクトでは、「AI と文化、PAI Agent、トラスト、ガバナンス」という４つの視点を切り口にし、12 もの多岐にわたる項目について実施した。12 の実施項目の中には、国の AI 倫理規範の策定に関わるものから、具体的な想定に基づく調査（医療診断チャット、診断・手術支援、AI トリアージ、政策決定支援システム、裁判支援など)、「COCOA」の検証まで、幅広い検討と分析がなされており、それぞれの項目について十分な成果が得られたと評価できる。また、その成果は、総務省「AI ネットワーク社会推進会議」「WEB3 時代に向けたメタバース等の利活用に関する研究会」をはじめ、各種学会や他のプロジェクト等にフィードバックされている点も評価できる。

個々の成果に目を向ければ、「共視論」を援用した人間とロボットの関係における文化差の分析や、トラスト概念における「trustworthiness（＝信頼に値すること、信頼相当性)」の提起、AI ガバナンスにおける法とアーキテクチャの役割の分析など、多くの興味深い調査結果が示されている。なかでも、PAI Agent のトラストとガバナンスに関する倫理的、法的、社会的な調査、分析および考察は、政策や企業活動へのインパクトも大きく、また今後のさらなる調査と研究の重要な参照点となり得ると思われる。また、PAI Agent の社会受容に関して、一般の方と専門家の両方に調査を実施し、その結果を比較考察した研究は、まだ一般の方と専門家の間の違いがどこにあるのかの十分な解明に至っていないようだが、それに向けての重要な参考資料になり得ると考える。

それぞれのパートで分析や論点の抽出等十分な成果を創出した一方で、それぞれが独立しており、最終的なまとめの段階でもう少し統合の工夫があっても良かったかもしれない。本プロジェクトでは、プロジェクトタイトルに「文化を超えた実現方法」と銘打っているが、その結論が必ずしも明確に示されてはいないように思われる。グローバルなプラットフォーム上で行われる活動と、各国の法規制を受ける部分とをうまく区別しながら、「文化を超えた実現方法」という観点から個々の成果を俯瞰分析して、もう一段階掘り下げた検討がなされればより良かったと思われる。

また、AI 倫理規範の背景調査での包括的な調査や、インタビュー、WEB アンケートなどの社会調査で得られた知見を、例えば、技術開発者や AI 技術の利用者が参照してセルフチェックできるような成果物があれば、技術開発や社会実装をする上で、より多くの様々な技術開発に対して有益であると考える。本プロジェクト成果の広範な還元として今後の検討に期待したい。

４．研究開発成果の活用･展開の可能性

大いに期待できると評価する。

本プロジェクトの遂行を通じて、多くの研究データが蓄積され、また多数の研究成果が公表されている。現実社会でAIが急速に浸透している中、新たな政策や法制度等検討していく必要があるが、本プロジェクトの問題意識やこれらの研究成果は今後大いに活用されていくと思われる。

とりわけ、画像生成AIであるmidjourneyやStable Diffusion、自然言語生成AIであるChatGPTなどの生成的AIがますます実装される状況化において、その生成物の権利帰属やそれへのトラスト、AI エージェントの役割などの議論が今後社会にとって求められるが、本プロジェクト成果からの応用検討の可能性も期待できる。

対象が研究開発や社会情勢の変化のスピードが速い分野であることも踏まえ、引き続き、継続的な検討、並びに積極的な情報発信をお願いしたい。

Ⅱ．研究開発プロジェクトの領域への貢献

研究開発プロジェクトの運営と活動、および得られた研究開発成果は領域の目標達成に大いに貢献できたと評価する。

情報技術を人間を中心とした視点で捉え直し、社会への理解を深めつつ技術や制度を協調的に設計していくことに対して、AI という、現在において最も注目され、その有効かつ倫理的･社会的な利用が求められている技術に焦点を当ててアプローチしたものであり、領域の目標達成に多大な貢献をしたプロジェクトであったと評価できる。

PAI Agentに対するトラスト設計の提案など、本領域の中心的なテーマに取り組んだプロジェクトであり、パーソナルデータの利活用方針を設計する橋田プロジェクトとは、PAI Agentの社会受容性の観点から連携を促進し、社会受容およびトラスト形成のための論点抽出などお互いのプロジェクトにとって相乗的な効果が見られたと考える。また、山本（ベバリーアン）プロジェクトや稲谷プロジェクト等とも連携しながら進められており、他プロジェクトへの貢献という意味においても、重要な役割を果たしたと言える。

プロジェクトは終了となるが、プロジェクト期間中に創出した知見や成果をもとに、本領域と継続的に関係しながら、それらの成果を様々に展開していく取り組みがなされることを大いに期待したい。

以上

令和５年５月
社会技術研究開発センター

「人と情報のエコシステム」研究開発領域における 2022 年度　研究開発プロジェクト事後評価結果について（概要）

社会技術研究開発事業「人と情報のエコシステム」研究開発領域の研究開発プロジェクトに対し、以下のとおり事後評価を実施した。

１．評価対象

下表のプロジェクトを評価の対象とした。【１件】

プロジェクト名称	研究代表者	所属・役職 （事後評価実施時点）
PATH－AI:人間-AI エコシステムにおけるプライバシー、エージェンシー、トラストの文化を超えた実現方法	中川　裕志	理化学研究所 革新知能統合研究センター　チームリーダー

２．評価の進め方

以下の手順で評価を行った

・令和５年２月　　評価用資料の作成・「終了報告書」提出
・令和５年２月　　事前査読
・令和５年２月 24 日　　ヒアリング評価
・令和５年３～４月　　評価報告書（案）の検討
・令和５年５月　　評価報告書の確定
　　評価報告書の内容に関する事実誤認および非公開事項の有無等確認を研究代表者等に対して実施。

３．評価項目

以下の評価項目により、評価結果を「評価報告書」として取りまとめた。

A.総合評価

B.項目評価

（１）研究開発プロジェクトの研究開発内容とその成果について

①目標の妥当性

②研究開発プロジェクトの運営・活用状況

③研究開発プロジェクトの目標の達成状況および研究開発成果

④研究開発成果の活用・展開の可能性

（２）研究開発プロジェクトの領域への貢献

４．評価者（所属・役職は事後評価実施時点）

＜領域総括＞

國領　二郎	慶應義塾大学 総合政策学部　教授

＜領域総括補佐＞

城山　英明	東京大学　大学院法学政治学研究科　教授

＜領域アドバイザー＞

加藤　和彦	筑波大学　副学長・理事（総務人事・情報環境担当）
久米　功一	東洋大学　経済学部　教授
河野　康子	一般財団法人日本消費者協会　理事
砂田　薫	国際大学ｸﾞﾛｰﾊﾞﾙ・ｺﾐｭﾆｹｰｼｮﾝ・ｾﾝﾀｰ　主幹研究員
信原　幸弘	東京大学　名誉教授
松原　仁	東京大学　大学院情報理工学研究科　教授
丸山　剛司	元 中央大学　理工学部　特任教授
村上　文洋	株式会社三菱総合研究所　ICT・ﾒﾃﾞｨｱ戦略ｸﾞﾙｰﾌﾟ　主席研究員
村上　祐子	立教大学　大学院人工知能科学研究科・文学部　教授

＜評価専門アドバイザー＞

村田　潔	明治大学商学部　専任教授
奥和田　久美	北陸先端科学技術大学院大学　客員教授

以上