

戦略的創造研究推進事業（社会技術研究開発）

2021(令和3)年度

研究開発実施報告書

科学技術の倫理的・法制度的・社会的課題（ELSI）への
包括的実践研究開発プログラム

研究開発プロジェクト

「人工知能の開発・利用をめぐる自律性および関係性の
理論分析と社会実装」

Autonomy and Affinity in AI Development and Use

2022(令和4)年5月31日

研究開発期間

(2021(令和3)年10月1日～2025(令和7)年3月31日)

研究代表者／Principal Investigator

宇佐美 誠

京都大学・大学院地球環境学堂・教授

USAMI Makoto

Professor, Graduate School of Global Environmental Studies,

Kyoto University

人工知能の開発・利用をめぐる自律性および関係性の理論分析と社会実装

■概要： 人工知能（AI）の加速度的な発展・普及に伴って、個人の自律性への新たな脅威、人間と機械との関係の変容、自律型機械から生じるリスクなど、多様な倫理的課題が発生し増大しつつある。本プロジェクトでは、法哲学・社会哲学・倫理学等の研究者と工学研究者との緊密な協働の下、人工知能やそれをういたビッグデータ分析の技術開発・利活用において生起する倫理的課題を発見・予見した上で、「人間・人工知能の自律性」と「人間－人工知能の関係性」という2つの観点から、理論分析を実施する。また、ステークホルダーとの連携・実践を通じて、技術者の行動指針を作成する。人工知能倫理学の先端的研究成果を活用しつつ、日本の特徴も考慮した精緻な分析を通じて、科学技術が人間・社会と調和する人間関係や社会のあり方を構想するとともに、日本の文脈に即した責任ある研究・イノベーションを根付かせる基礎となる行動指針の提案を目指す。

■参画・協力機関： 京都大学、九州大学、共立女子大学、慶應義塾大学、東京大学、名古屋大学、南山大学、日本大学、武蔵野大学、明治学院大学

■キーワード： AI、ビッグデータ、ロボット、倫理、行動指針

Autonomy and Affinity in AI Development and Use

■ Summary: As artificial intelligence (AI) has been rapidly developed, humanity has encountered an increasing number of philosophical and moral challenges including new forms of threat against individual autonomy, the transformation of relationships between humans and smart machines, and risks from autonomous machines. This research project, carried out by a multidisciplinary team consisting of legal, moral, and political philosophers as well as computer scientists and engineers, seeks to uncover or foresee ethical challenges raised in the contexts of AI development and use. The team then analyzes these challenges from the perspectives of individual and AI autonomy and of human-machine relationship, utilizing recent findings made by overseas theorists in AI ethics. Finally, it proposes the code of conduct for AI developers to actualize the principle of responsible research and development in the context of Japanese society.

■ Joint R&D Organizations: Kyoto University, Keio University, Kyoritsu Women's University, Kyushu University, Meiji Gakuin University, Musashino University, Nagoya University, Nanzan University, Nihon University, University of Tokyo

■ Key words: artificial intelligence, big data, robot, ethics, code of conduct

1. プロジェクトの達成目標

本プロジェクトは、人工知能（AI）と人・社会の調和という終極理念の下、次の2点を具体的達成目標とする。

①AIの倫理的課題の理論分析： AI・ビッグデータ分析の技術開発や企業の開発者との共同探索的会合を通じて、開発・利用上の倫理的課題を把握し、人間・AIの自律性および人間-AIの関係性という観点から構造化する。次に、AI倫理学の国際学界での先端的研究成果を活用するとともに、日本の状況の特徴に留意しつつ、新たな規範理論を構築し、理論分析を実施する。これらの作業を通じて、AIの倫理的課題について精緻な規範的分析を行うことが、本プロジェクトの理論的目標である。

②技術開発者の行動指針の策定： 倫理的課題の理論分析の成果を基盤に、ステークホルダーとの連携の下、AIの開発・利用の現実やそれを取り巻く日本社会の特徴を踏まえつつ、開発者がとるべき行動や、開発・利用を行う企業が設けるべき制度に関して、行動指針（CoC）を策定し公表する。これは、責任ある研究・イノベーション（RRI）の制度化に向けた本プロジェクトの実践的目標となる。

2. 研究開発実施内容・成果の概要

当初の計画通り、実施項目のうち「AI・ビッグデータ分析の倫理的論点の把握」および「理論分析の枠組みの構築」について研究開発を実施し、以下の成果を得た。

項目1：AI・ビッグデータ分析の倫理的論点の把握

個人の自律性に影響を及ぼす行動誘導型情報提供、自律型AIの一典型である致命的自律型兵器システム、人間-AIの関係性が問われる介護ロボットなどについて、開発現場や利用の場面で生じる倫理的課題を探索した。

開発グループでは、各実施者による技術開発や、企業の技術開発者・システム開発者らとのネットワークに基づくトランス・ディシプリナリーな共同探索を通じて、倫理的課題を把握する作業を開始した。実施者による技術開発および課題把握としては、次の3種の研究開発作業を実施した。第1に、個人の自律性や人間とAIの関係性を築く上で重要となりうる要素として、説明責任への研究関心が近年高まっており、特に説明可能性を担保するAI（XAI）が注目されている。こうした技術動向を調査するとともに、説明可能な人工知能の実装手法として、Interpretable Smart Computingを提案し、その実現によって自律性や人間とAIの関係性が変化するかを検討した。第2に、研究従事者へのヒアリングを通じて倫理的課題を把握する作業を、IT/AI分野、特に「機械学習自体に関する研究」および「AIを主要な要素として内包するシステム（“AIシステム”）に関連する研究」を遂行中の大学院生博士課程の研究従事者を対象に行い、AIの学習自体、AI（モデル）に対する攻撃、AIを使った人間行動をめぐる諸問題を導出した。第3に、交通分野における行動誘導型情報提供としては、公共交通オープンデータの活用事例として行動変容をめざしたアプリ開発の事例などを調査するとともに、データ活用による交通計画について研究を進めた。また、AI信号システムを開発し、その機能や性能とともに利用者の受容性をフィールド実験により検討した。

上記の研究開発の諸作業と並行して、開発グループを中心に、企業の開発者との共同探索会合に向けた準備会を2回実施し、招聘対象者の人選および討論テーマについて審議した。それに基づき、共同探索会合を2回開催した（講師：松田雄馬氏（株式会社オンギガンツ代表取締役）、佐々木由樹氏（株式会社リンクアンドコミュニケーション））。その結果、行動誘導型情報提供システムをめぐる倫理的課題の一部を同定できた他、人間の生物的情報処理システムに対するAIの含意について検討を深化させることができた。

自律性グループ・関係性グループにおける主な調査的活動としては、次の2点を挙げられ

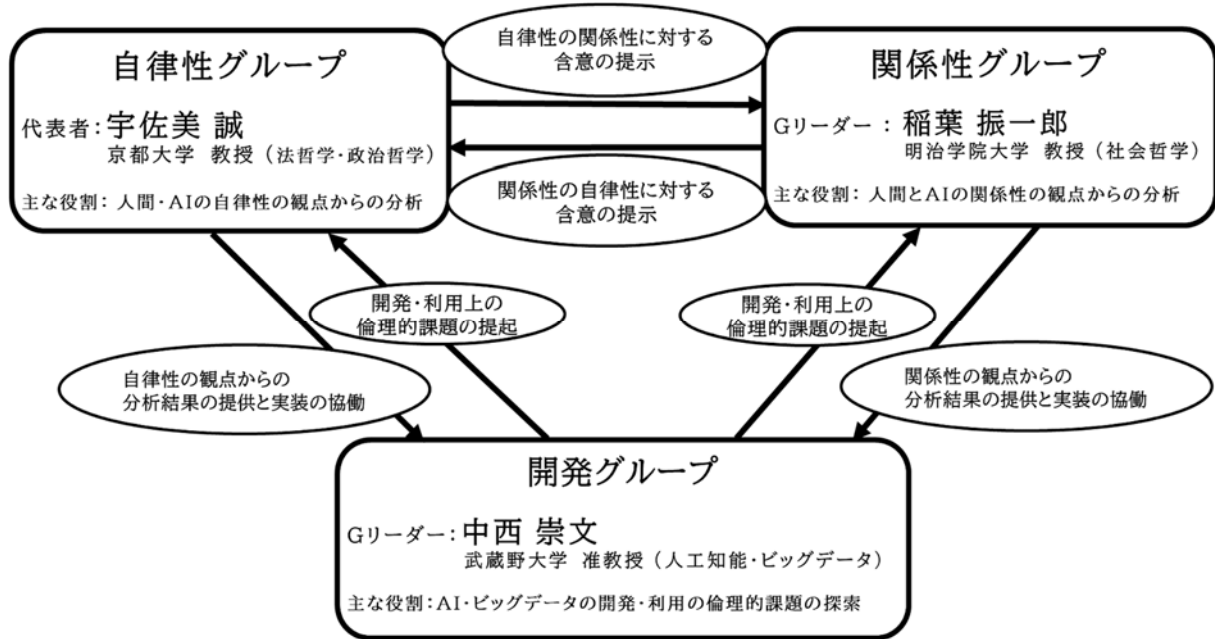
る。第1に、自律的AIの実装の一場面である致命的自律型兵器システム(LAWS)に関して、LAWSをめぐる市民意識を探查するため、調査票作成、標本抽出、調査の実施時の課題把握などに必要な情報を得るための予備的調査を、Web調査で実施した。先行研究を踏まえたシナリオに基づく選択肢選択型調査を行い、今後の本調査に向けての検討課題を明らかにするとともに、尊厳に関する市民意識や、今般のロシアのウクライナ軍事侵攻による安全保障意識の変化の可能といった発展的研究の端緒を見出すことができた。これは、AI・ビッグデータ分析の倫理的論点の把握を進めることにつながるものである。第2に、関係性をめぐる倫理的論点に関わる基礎的調査としては、見守り用途やコミュニケーション用途で使用可能な普及価格帯の機器を2種(Bocco Emo、およびEcho Show 10)を購入し、現段階で介護施設ではない自宅等での普及および活用が可能かを検討した。

項目2：理論分析の枠組みの構築

自律性グループでは自律性およびその関連概念を主題として、また関係性グループでは人間-AIの関係性をめぐる諸論点について、それぞれ理論分析の枠組みを構築する作業を開始した。自律性に関しては、倫理学史上の代表的思想であるカントの道徳哲学とその現代的影響について包括的レビューを行うとともに、個人の日常的意思決定における自律性の構造について理論枠組みの基礎を検討した。個人の自律とAIの自律の関係について、個人の自律は、規範的な概念であり、それ自体に内在的な価値が認められるのに対し、AIの自律は、記述的・経験的な概念であり、手段的・道具的な価値を有するに留まるため、個人の自律を促進する手段としてAIの自律を理解することが適切であるとの見通しを示した。その上で、IoB(インターネット・オブ・ブレインズ)の発展などにより両者の関係が変容する可能性について考察した。また、自律型兵器については、兵器使用に関する道徳的是非の論点整理を行い、その判断能力や責任能力、閾値の低下、殺傷における人間の尊厳などを主要な論点として析出した。法システム上、自律性を促進するためのアプローチとしては、権利アプローチとガバナンス・アプローチの二つが考えられる。ガバナンス・アプローチの観点からAIシステムを開発する組織や主体の問題を考察した。自律性と関係性の双方に対して変容を迫るAI利用領域として、裁判官AIや法令調査用人工知能ROSSの導入の問題性を論じる文献を調査し、これを評価する規範理論を検討して権威と自律の両立可能性という問題設定を行った。

さらに、本プロジェクトに関連する諸論点についても、理論的検討を進めた。応用倫理学の観点からは、食農倫理学における家畜・農作物の位置づけとAI倫理学におけるロボット・AIシステムとのそれとの並行性・類似性について検討を行った。また、AI・ビッグデータ分析の開発・利用をめぐる論点において、データ流通の世界における現実の課題として、巨大なプラットフォーム事業者の市場における経済活動に対する規制・取組を整理し、特に、米国反トラスト法の訴訟実態を取り上げ、市場画定、独占力の推定など5つの観点から検討した。

3. 研究開発実施体制



〈実施体制図〉

4. 今年度までの活動実績

4-1. 研究論文

(1) 国際誌

- T. Nitta, S. Hagimoto, A. Yanase, R. Okada, V. Sornlertlamvanich, T. Nakanishi (2022). Realization for Finger Character Recognition Method by Similarity Measure of Finger Features. International Journal of Smart Computing and Artificial Intelligence, 6(1). <https://doi.org/10.52731/ijscai.v6.i1.684> (査読あり)
- A. Ikegami, R. Okada, T. Nakanishi (2022). The Discovery of Historical Transition in Aesthetic Notions Through Changes in Co-occurrence Words Mainly Used in Waka Poetry in Three Major Poetry Anthologies. In: Lee R. (eds), Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing. SNPD 2021. Studies in Computational Intelligence, vol 1012. Springer, Cham. https://doi.org/10.1007/978-3-030-92317-4_12 (査読あり)

(2) 国内誌

- 吉添衛, 服部宏充, 江間有沙, 大澤博隆, 神崎宣次, 久木田水生, 小川祐樹 (2021) 「グループディスカッションにおける可視化情報提示に基づく気づき支援」『ヒューマンインタフェース学会論文誌』 23(4), pp. 501-512. https://doi.org/10.11184/his.23.4_501 (査読あり)
- 成原慧 (2021) 「媒介者責任の再検討: プロバイダ責任制限法改正および関連する取組の意義と課題」『法学セミナー』 66(12), pp. 45-51. (査読なし)
- 松尾陽 (2021) 「アーキテクチャ論から新型コロナ禍の対応を考える」『法哲学年報 2020』 pp. 80-92. (査読あり)

- 松尾陽 (2022)「規制手法の多様化と法哲学の課題」『法律時報』1174, pp. 52-57 (査読なし)
- 松尾陽 (2022)「コロナ禍における多様な規制手法とその制御のあり方:ポスト／ウィズ・コロナの法哲学」『学術の動向』2022年3月号, pp. 40-46. (査読あり)

4-2. 総説・書籍など

(1) 国際

- 該当なし

(2) 国内

- 宮下萌編著, 明戸隆浩, 石川優実, 川口泰司, 上瀧浩子, 曹慶鎬, 唐澤貴洋, 佐藤佳弘, 金尚均, 成原慧, 佐藤暁子 (2022)『テクノロジーと差別: ネットヘイトから「AI」による差別」まで』解放出版社.
- 松元雅和 (2022)「平和・安全保障」山岡龍一・大澤津編著『現実と向き合う政治理論』放送大学教育振興会, pp. 124-138.

4-3. 講演・発表

(1) 招待講演

- 伊藤昌毅 (2021)「公共交通オープンデータ第2幕:「静的データは出来た、次はリアルタイム」と決めつける前に考えること」福井県主催『令和3年度 新モビリティサービス検討会』福井県 (国内)
- 伊藤昌毅 (2022)「行動変容と交通インフラの動的制御によるスマートな都市交通基盤を目指して」『電子情報通信学会 第26回ヒューマンプローブ研究会 (HPB)』オンライン (国内)
- 神崎宣次 (2022)「倫理の観点からみた EdTech の ELSI 論点」『シンポジウム「学習データ利活用 EdTech (エドテック) の ELSI (倫理的・法的・社会的課題)」』オンライン (国内)
- Satoshi Narihara (2022) Does Trimming of a Photograph Using AI in Retweets Infringe Moral Rights of the Author of the Photograph?, Future science & technology and legal policy joint International Conference, ハイブリッド (国際)
- 成原慧 (2021)「デジタル社会におけるアルゴリズム／アーキテクチャと法」『GLOCOM 六本木会議オンライン#32』オンライン (国内)
- 成原慧 (2022)「AI・デジタル時代の プライバシー・個人情報保護」『第二東京弁護士会情報公開・個人情報保護委員会定例研修会』オンライン (国内)

(2) 口頭発表

- 寺園結基, 伊藤昌毅, 大口敬 (2021)「時刻表データを用いた公共交通網評価手法の提案」『第64回土木計画学研究発表会』オンライン (国内).
- 荒井弘毅 (2021)Data Distribution and Competition Law Issues. The 4th International Workshop on Cross-disciplinary Data Exchange and Collaboration (CDEC2021)オンライン (国際)<https://www.panda.sys.t.u-tokyo.ac.jp/CDEC/2021/index.htm> 採録資料 (doi: 10.1109/BigData52589.2021.9671329. 2021.) (査読付き)
- H. Nakata, T. Nakanishi (2021) Music Impression Extraction Method by Chord Impressions and

Its Application to Music Media Retrieval. Proceedings of 22nd IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD 2021-Fall), pp. 68-73. (国際)

▪

(3) ポスター発表

- 該当なし

4-4. WEB・プレス発表・メディア

(1) Web サイト

- プロジェクト Web サイト：
宇佐美研究室の共同研究
<https://www.envpolicy.ges.kyoto-u.ac.jp/current-projects>

(2) プレス発表

- 該当なし

(3) 報道・投稿

- 松尾陽, 朝日新聞 (2021 年 10 月 14 日) 「憲法季評：ワクチンの奨励政策 緩やかな規制、生かすには」

4-5. 会議・イベント

- 該当なし

4-6. 知的財産権

- 該当なし

4-7. 受賞

- 該当なし

4-8. その他

- 該当なし