

戦略的創造研究推進事業
(社会技術研究開発)
令和元年度研究開発実施報告書

「人と情報のエコシステム」

研究開発領域

「人と情報テクノロジーの共生のための

人工知能の哲学2.0の構築」

鈴木 貴之

(東京大学大学院総合文化研究科、准教授)

目次

1. 研究開発プロジェクト名	2
2. 研究開発実施の具体的内容	2
2-1. 研究開発目標	2
2-2. 実施内容・結果	2
2-3. 会議等の活動	5
3. 研究開発成果の活用・展開に向けた状況	6
4. 研究開発実施体制	6
5. 研究開発実施者	7
6. 研究開発成果の発表・発信状況、アウトリーチ活動など	8
6-1. シンポジウム等	8
6-2. 社会に向けた情報発信状況、アウトリーチ活動など	8
6-3. 論文発表	8
6-4. 口頭発表（国際学会発表及び主要な国内学会発表）	9
6-5. 新聞／TV報道・投稿、受賞等	10
6-6. 知財出願	10

1. 研究開発プロジェクト名

人と情報テクノロジーの共生のための人工知能の哲学2.0の構築

2. 研究開発実施の具体的内容

2-1. 研究開発目標

- ①人工知能の可能性と限界を考察するうえで鍵となる徳を中心とした諸概念の関係を明らかにし、哲学研究者以外にもその成果が容易に理解できるように、コンセプト・マップなどを作成する。
- ②人工知能の社会実装可能性を考えるための手がかりとなる概念枠組を構築する。具体的には、人間と人工知能のとりうる関係を類型化し、各類型の実現可能性や長所・短所などを明らかにしたチャートなどを作成する。
- ③人工知能の可能性と限界を検討するための新たな理論的枠組（人工知能の哲学2.0）を構築する。その内容を、人工知能研究者や人工知能の社会実装に携わる人々もアクセスできる教科書や概説書などの形で公刊する。
- ④情報テクノロジー研究開発者へのインタビューや研究会の開催などを通じて、情報テクノロジーの研究開発において、哲学をはじめとする人文諸科学に（倫理的問題の検討以外に）どのような貢献の可能性があるかを明らかにする。
- ⑤学会ワークショップやシンポジウムの開催などを通じて、情報テクノロジーの研究開発者と人文科学研究者との交流を促進する。さらに、一般向けのトークイベントの開催や新聞・雑誌における記事の執筆などを通じて、研究者や技術者と一般の人々との問題関心の共有可能性を探る。

2-2. 実施内容・結果

(1) スケジュール

実施項目	2018年度	2019年度	2020年度	2021年度
1. 人工知能研究の歴史と現状のレビュー	→	→	→	
2. 人工知能の哲学の体系的再検討	→	→		
3. 徳概念を手がかりとした各論点の関係の分析		→	→	
4. 人工知能の可能性と限界の解明			→	
5. 成果物の作成				→

6. 徳に関する哲学的議論のレビュー	→	→		
7. 経験的知見をふまえた徳理解のアップデート		→		
8. 人工知能による徳の実現可能性			→	
9. 成果物の作成				→
10. 拡張された心に関する先行研究のレビュー	→	→		
11. 情報テクノロジーによる人間知性の拡張可能性の検討		→		
12. 人間知性と人工知能の協働可能性の検討			→	
13. 成果物の作成				→

(2) 各実施内容

今年度の到達点①

(目標) 人工知能研究の歴史と現状に関する整理と分析

実施項目1: 人工知能研究の歴史と現状のレビュー (全体計画書の実施項目1)

実施内容: ①東京大学の大学院生の協力を得て、人工知能の哲学に関する日本語文献のリストを作成した。②東京大学で深層学習に関する研究会を開催した。③2019年8月に、北海学園大学で人工知能研究者を講演者に招いた研究会を開催した。④2019年8月に、東京大学の堀浩一教授へのインタビューを実施した。

今年度の到達点②

(目標) 先行研究の分析のアップデートと深化

実施項目2-1: 徳概念を手がかりとした各論点の関係の分析 (全体計画書の実施項目3)

実施内容: ①2019年11月に開催された日本科学哲学会第52回大会において、ワークショップ「機械学習・深層学習の哲学的意義」(オーガナイザー: 鈴木貴之)を開催した。②2019年11月にスペイン・グラナダ大学で開催された国際ワークショップ Japanese-European Meeting on Artificial Intelligence and Moral Enhancementにおいて、鈴木貴之が発表を行った。

実施項目2-2: 経験的知見をふまえた徳理解のアップデート (全体計画書の実施項

目7)

実施内容：①昨年度に引き続き、徳に関する哲学的議論のレビュー（実施項目6）を実施した。②グラナダ大学を中心とした研究グループとの上記国際ワークショップなどにおいて、立花幸司が経験的知見をふまえた徳理解のアップデートの可能性に関する研究発表を行った。③日本科学哲学会における上記ワークショップにおいて、植原亮が、人工知能が広義の徳の一種である創造性に対してどのような可能性を持つかを検討する発表を行った。

実施項目2-3：情報テクノロジーによる人間知性の拡張可能性の検討（全体計画書の実施項目11）

実施内容：①技術哲学および科学技術人類学における関連文献の調査と分析を実施した。②その成果をふまえ、グラナダ大学における上記国際ワークショップにおいて、染谷昌義と上杉繁が人工知能による知性の拡張性に関する発表を行った。③応用哲学会やいくつかの国際会議において、上杉繁がロボット技術と人間の関係についての発表を行った。

（3）成果

今年度の到達点①

（目標）人工知能研究の歴史と現状に関する整理と分析

実施項目1：人工知能研究の歴史と現状のレビュー（全体計画書の実施項目1）

成果：人工知能の限界に関する従来の哲学的な議論は、中国語の部屋の議論に由来する意味理解の問題と、フレーム問題の2つを中心的な論点とすること、深層学習を初めとする近年の人工知能研究の進展によって、従来の議論の見直しが必要となることが明らかとなった。

今年度の到達点②

（目標）先行研究の分析のアップデートと深化

実施項目2-1：徳概念を手がかりとした各論点の関係の分析（全体計画書の実施項目3）

成果：人工知能の限界に関する従来の哲学的批判のうち、意味理解の問題に関しては、深層学習を初めとする近年の人工知能技術が解決の手がかりをもたらしているが、フレーム問題は解決がより困難であること、深層ニューラルネットはいくつかの点で記号計算的なアプローチによる補完を必要とするかもしれないことが明らかになった。

実施項目2-2：経験的知見をふまえた徳理解のアップデート（全体計画書の実施項目7）

成果：経験的知見をふまえた徳理解のアップデートの具体例として、ブレイン・マシン・インターフェイス技術による徳の増強の可能性とその倫理的問題について検討

した。また、人工知能を用いた徳の実現可能性を検討しているグラナダ大学（スペイン）を中心とした研究グループと国際ワークショップを実施し、共同研究に向けた調整を行った。さらに、ブダペスト工科経済大学（ハンガリー）の研究者とも研究交流の準備を開始した。

実施項目2-3：情報テクノロジーによる人間知性の拡張可能性の検討（全体計画書の実施項目11）

成果：人間を人工知能など外部のエージェントとのハイブリッドとして考えることが可能であること、それによって人間の能力はさまざまに変容しうること、そのようなハイブリッド化の技術を開発する際には、設計原理の検討が不可欠であることを明らかにした。また、創造性を例として考えたときには、創造性のある種の側面は機械学習や深層学習による分業が可能であり、人工知能をある種の認識論的なツールとして用いることができることを明らかにした。

（4）当該年度の成果の総括・次年度に向けた課題

今年度は、新型コロナウイルス問題の影響で、2020年2月以降に計画していたいくつかの活動（グラナダ大学の研究チームとの東京大学での国際ワークショップ、東洋大学におけるシンポジウムの発表、一橋大学の久保明教氏を講演者とした研究会など）を実施することができなかった。また、令和2年度に開催を計画していたトークイベントなどの交渉を進めることができなかった。対面での研究活動が難しい状況がしばらく続くと考えられるため、次年度は研究会やインタビューにZoomなどのビデオ会議システムを積極的に活用したい。また、これまでの研究活動の成果を論文化できていないため、今年度は論文としての公刊を積極的に進めたい。

2-3. 会議等の活動

年月日	名称	場所	概要
2019年5月10日	2019年度第1回全体研究会	東京大学駒場キャンパス	立花幸司が研究発表を行った。
2019年8月29日、30日	2019年度第2回全体研究会	北海学園大学	北海道大学の飯塚博幸氏、小野哲雄氏による講演と、植原亮および鈴木貴之の研究発表を行った。

3. 研究開発成果の活用・展開に向けた状況

プロジェクトの研究活動の記録および一般向けの資料を下記プロジェクトウェブサイト
公開している。ウェブサイトでは、人工知能をめぐる哲学的な議論の成果を、人工知能研
究者や一般市民など、哲学研究者以外に向けて発信することを目指している。

プロジェクトウェブサイトURL：<http://updatingphilosophyofai.net/>

4. 研究開発実施体制

(1) 人工知能の哲学2.0の構築グループ

- ①鈴木貴之（東京大学大学院総合文化研究科、准教授）
- ②実施項目：人工知能研究の歴史と現状のレビュー、徳概念を手がかりとした各論点の
関係の分析

(2) 徳と人工知能グループ

- ①立花幸司（熊本大学大学院人文科学研究部、准教授）
- ②実施項目：経験的知見をふまえた徳理解のアップデート

(3) 拡張された心と人工知能グループ

- ①染谷昌義（高千穂大学人間科学部、教授）
- ②実施項目：情報テクノロジーによる人間知性の拡張可能性の検討

5. 研究開発実施者

人工知能の哲学2.0の構築グループ

氏名	フリガナ	所属機関	所属部署	役職 (身分)
鈴木貴之	スズキタカユキ	東京大学	大学院総合文化研究科	准教授

徳と人工知能グループ

氏名	フリガナ	所属機関	所属部署	役職 (身分)
立花幸司	タチバナコウジ	熊本大学	大学院人文社会科学部	准教授
植原亮	ウエハラリョウ	関西大学	総合情報学部	教授

拡張された心と人工知能グループ

氏名	フリガナ	所属機関	所属部署	役職 (身分)
染谷昌義	ソメヤマサヨシ	高千穂大学	人間科学部	教授
柴田崇	シバタタカシ	北海学園大学	人文学部	教授
上杉繁	ウエスギシゲル	早稲田大学	創造理工学部	教授
中澤栄輔	ナカザワエイスケ	東京大学	大学院医学研究科	講師

6. 研究開発成果の発表・発信状況、アウトリーチ活動など

6-1. シンポジウム等

年月日	名称	場所	参加人数	概要
2019年 11月10 日	「機械学習・深層学習の 哲学的意義」	慶応技術大 学三田キャンパス	約80名	日本科学哲学会第52回大会におけるワークショップ。人工知能の哲学、統計学の哲学、知識の哲学の観点から機械学習・深層学習の哲学的意義を検討した。
2019年 11月26 日、27 日	Japanese-European Meeting on Artificial Intelligence and Moral Enhancement	グラナダ大 学（スペイン）	約15名	グラナダ大学を中心とした研究グループと、ヨーロッパおよび日本における人工知能関連の哲学研究の現状を確認した。

6-2. 社会に向けた情報発信状況、アウトリーチ活動など

(1) 書籍・冊子等出版物、DVD等

なし

(2) ウェブメディアの開設・運営

- ・プロジェクトウェブサイト（URL：<http://updatingphilosophyofai.net/>、2019年2月開設）

(3) 学会（6-4.参照）以外のシンポジウム等への招聘講演実施等

鈴木貴之「人工知能と自然知能—代用品？上位互換？それとも新種？」第28回自然科学研究機構シンポジウム「SF／未来／科学技術」2019年8月24日、国際交流会議場
柴田崇「サイボーグ論の転回：『拡張』思想の射程の考察から」早稲田大学創造理工学研究科総合機械工学科専攻主催講演会、2019年10月3日、早稲田大学

6-3. 論文発表

(1) 査読付き（0件）

- 国内誌（0件）

・

- 国際誌（0件）

(2) 査読なし（1件）

・ 染谷昌義「アフォーダンスからの希望」『臨床心理学』第20巻第2号、金剛出版、136-141頁、2020年3月

6-4. 口頭発表（国際学会発表及び主要な国内学会発表）

(1) 招待講演（国内会議0件、国際会議0件）

・

(2) 口頭発表（国内会議3件、国際会議9件）

- ・ 上杉繁（早稲田大学）「ロボット技術と人間との関係におけるジレンマへのアプローチ」応用哲学会第11回年次研究大会、京都大学吉田キャンパス、2019年4月21日
- ・ 鈴木貴之（東京大学）「深層学習の哲学的意義：認知科学の哲学と人工知能の哲学の場合」日本科学哲学会第52回大会ワークショップ「機械学習・深層学習の哲学的意義」慶應義塾大学三田キャンパス、2019年11月10日
- ・ 植原亮（関西大学）「機械学習・深層学習と知的創造性」日本科学哲学会第52回大会ワークショップ「機械学習・深層学習の哲学的意義」慶應義塾大学三田キャンパス、2019年11月10日
- ・ Shigeru Wesugi. Waseda University. Designing approaches addressing dilemma in relating to robots. The 21st Conference of the Society for Philosophy and Technology. Texas A&M University in College Station. 2019/5/20-22.
- ・ Koji Tachibana. Kumamoto University. AI-based moral enhancement and the future of human virtue. Third International Workshop On Ethics And Human Enhancement. University of Granada, Spain. 2019/6/3.
- ・ Shigeru Wesugi. Waseda University. Analysing and Solving the Reduced-ability and Excessive-use Dilemmas in Technology Use. The 22nd International Conference on Engineering Design. Delft University of Technology, The Netherlands. 2019/8/5-8.
- ・ Koji Tachibana. Kumamoto University. Artificial Intelligence and Epistemic Virtues. Virtue, Media, and Democracy. University of Genova, Italy. 2019/9/26.
- ・ Koji Tachibana. An extended reply to Lara and Deckers (2019). Japanese-European Meeting on Artificial Intelligence and Moral Enhancement. University of Granada, Spain. 2019/11/26.
- ・ Takayuki Suzuki. The University of Tokyo. Toward an Update of Philosophy of Artificial Intelligence. Japanese-European Meeting on Artificial Intelligence and Moral Enhancement. University of Granada, Spain. 2019/11/27.
- ・ Ryo Uehara. Kansai University. Could AI be a creative machine? Japanese-European Meeting on Artificial Intelligence and Moral Enhancement. University of Granada, Spain. 2019/11/27.
- ・ Someya Masayoshi. Takachiho University. What the 21st century's philosophy of AI should consider: Human-machine hybrid nature. Japanese-European Meeting on Artificial Intelligence and Moral Enhancement. University of Granada, Spain. 2019/11/27.
- ・ Shigeru Wesugi. Waseda University. Design tool for analyzing human-AI technology relation. Japanese-European Meeting on Artificial Intelligence and

Moral Enhancement. University of Granada, Spain. 2019/11/27.

(3) ポスター発表 (国内会議0件、国際会議0件)

.

6-5. 新聞／TV報道・投稿、受賞等

(1) 新聞報道・投稿 (0件)

.

(2) 受賞 (0件)

.

(3) その他 (0件)

.

6-6. 知財出願

(1) 国内出願 (0件)

.

(2) 海外出願 (0件)

.