

人工知能に関する問題発掘型対話基盤と 新たな価値観の創出

研究代表者 江間 有沙 | 東京大学教養学部附属教養教育高度化機構特任講師

Acceptable Intelligence with Responsibility (AIR)は人文・社会科学、情報学とその両者を有機的に結び付ける科学技術社会論を専門とする異分野協同研究グループである。分野を超えて人工知能の社会的影響を議論し、政府による干渉や産業による利益誘導に左右されない、異分野間の対話・交流を促すための媒体や基盤を構築する。また、人工知能の目指すべき共通アジェンダや社会の未来ビジョンを設計し、技術開発・実装時の新設計基準や規範・倫理・制度に関する価値観を提案する。



参加研究者リスト

氏名	所属・役職
江間 有沙	東京大学教養学部附属教養教育高度化機構特任講師
秋谷 直矩	山口大学国際総合科学部助教
市瀬 龍太郎	国立情報学研究所情報学プリンシプル研究系准教授
大家 慎也	神戸大学大学院人文学研究科博士課程
大澤 博隆	筑波大学システム情報系助教
大谷 卓史	吉備国際大学 アニメーション文化学部准教授
神崎 宣次	滋賀大学教育学部准教授
久保 明教	一橋大学大学院社会学研究科専任講師
久木田 水生	名古屋大学大学院情報科学研究科准教授
駒谷 和範	大阪大学産業科学研究所教授
西條 玲奈	京都学園大学経営経済学部非常勤講師
田中 幹人	早稲田大学政治経済学術院准教授
服部 宏充	立命館大学情報理工学部准教授
本田 康二郎	金沢医科大学一般教育機構講師
宮野 公樹	京都大学学際融合教育研究推進センター准教授
八代 嘉美	京都大学iPS細胞研究所特定准教授
山下 倫央	産業技術総合研究所 サービス工学研究センター主任研究員
吉澤 剛	大阪大学大学院医学系研究科准教授

研究目的と方法

現在、人工知能に関する話題は多く、人工知能の社会的影響についての異分野間対話の必要性が認識されている。情報学系の研究者には、未来への影響を見据えた技術設計への期待が増化しており、人文・社会科学系研究者には、科学者・技術者とともに問題を設定、定義していくことが求められている。そしてそのような場を可能にするような対話の場の形成も求められている。

本研究プロジェクトの意義は、まさに今、現場で技術を作り出している若手・中堅の人工知能研究者が人文・社会科学の研究者と対話することによって、技術の設計・思想段階から有機的に結び付き、新たな概念や技術へのアプローチ法を模索することにある。このような観点から、本研究プロジェクトは、1) 人文・社会科学系研究者による倫理的・法的・社会的問題(Ethical, Legal and Social Implications: ELSI) 調査グループ、2) 人工知能研究者によるAI社会応用調査グループ、3) 科学技術社会論や科学コミュニケーションを専門とする対話基盤設計グループを設け、人工知能の社会的影響を議論し、(1) 政府による干渉や産業による利益誘導に左右されない、異分野間の対話・交流を促すための媒体や研究拠点をボトムアップで構築し、(2) 対話を通して、人工知能の目指すべき共通アジェンダや社会の未来ビジョンを設計し、技術開発・実装時の新設計基準や規範・倫理・制度に関する価値観を提案することを目的とする。

2015年度実績報告

2015年度は各研究分野の問題意識や全体像を共有し、それぞれの価値観のすり合わせや整理を行うために、(1) Acceptable Intelligence with Responsibility (AIR)が提案するコンセプト・ペーパーの構想、(2) 人工知能研究者や人文・社会科学系研究者などの人工知能に対する態度や考えを示せる尺度の作成、そして(3) 過去の文理融合プロジェクトに係わった先生方へのインタビュー調査の実施を開始することで合意にいたった(AIR Newsletter Vol.2 (1)参照: <http://sig-air.org>)。 (1)のコンセプト・ペーパーについては最終年度に向けて議論を継続しており、2015年度は主に(2)尺度の作成を具体化するためのWEBアンケート調査と、(3)インタビュー調査を行った。

(1)コンセプト・ペーパーの構想

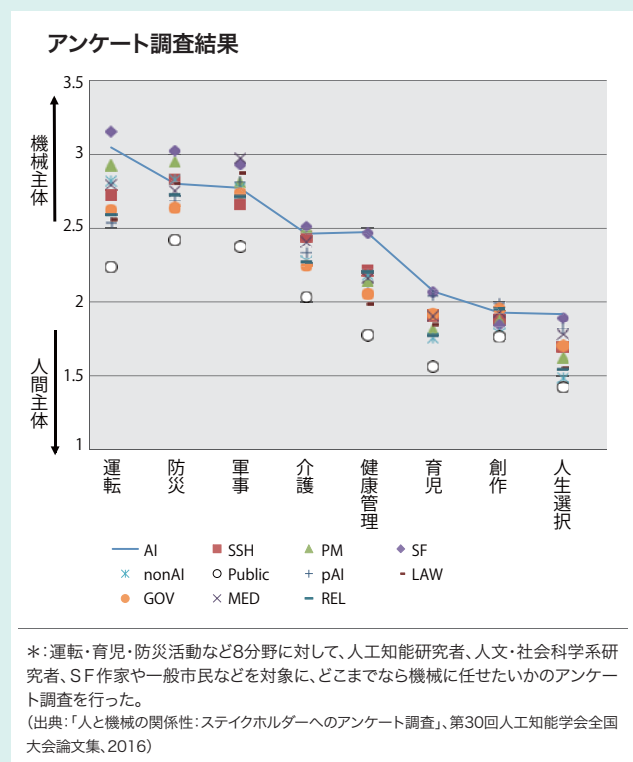
本プロジェクトメンバーは7月の国立情報学研究所での合宿において、新たな判断や選択肢を生み出す手助けをする発想支援システムの実装とその倫理的課題を検討する基盤となる「コンセプト・ペーパー」の議論を行った。2016年2月に行われた合宿では、コンセプト・ペーパーを執筆するために、システムのプロトタイプを情報学系研究者と人文・社会科学系研究者の協同で製

作り、そのプロセスそのものをコンセプト・ペーパーとして論文化していくこととし、実現に向けた工程表の作成を行った。

(2) アンケート調査

WEBアンケート調査は、国立研究開発法人科学技術振興機構社会技術研究開発センター（JST-RISTEX）と協同して行うことで、2015年10月に合意が得られた。具体的には、AIRメンバーがアンケート調査項目を作成し、情報学系研究者と人文・社会学系研究者など多様なステークホルダーを対象として調査を行った。ほぼ同様の調査項目をJST-RISTEXが一般の人々を対象として調査を行い、研究者間、また研究者と一般の間での人工知能に対する態度や考え方についてのベースライン調査を行った。本調査は、2015年12月と2016年2月に行った。

本調査の結果の一部は2016年2月17日の「JST-RISTEX第13回社会技術フォーラム～新領域に関する社会との対話～人と情報のエコシステム：情報技術が浸透する超スマート社会の倫理や制度を考える」にて公表されたほか、2016年度人口知能学会全国大会で発表し、「情報管理」2016年8月号に掲載される。



(3) インタビュー調査

2015年9月18日に大学評価・学位授与機構教授の土屋俊氏、また2016年3月1日に東京大学教授の堀浩一氏にインタビューを行い、1980年代に人工知能・認知科学などの異分野研究がいかに行われたかのオーラルヒストリーと、現在の人工知能の異分野間協同研究についての示唆を得た。本調査は来年度も継続して行っていく。

また、現在人工知能やロボットなどが導入されている現場へのフィールド調査として、2016年2月にハウステンボス「変なホテル」を訪問し、総支配人にインタビューを行った。近未来の社会の働き方や価値観の一考となる内容であり、本インタビュー報告は「情報処理」2016年11月号に掲載される。

(4) 2015年度 高等研研究会実施報告

- 1) 2015年8月21日：全体会合に向けた打ち合わせ
- 2) 2015年9月3日：2015年度の活動計画の設計／

アンケート調査とインタビュー調査に関する打ち合わせ

- 3) 2015年9月18日：土屋俊氏インタビュー
- 4) 2015年10月26日：アンケート調査項目最終確認／科学技術社会論学会公開ワークショップの最終打ち合わせ
- 5) 2016年1月6日：第2回アンケート調査設計の最終確認／2月研究会での議題の検討
- 6) 2016年2月8日-9日：長崎県佐世保市のハウステンボス「変なホテル」視察／アンケート調査の結果確認／3月1日インタビュー調査に関する打ち合わせ／コンセプト・ペーパーと来年度の活動についての意見交換
- 7) 2016年2月25日：第2回インタビュー調査の最終確認／今後の方針についての検討
- 8) 2016年3月1日：堀浩一氏インタビュー
- 9) 2016年3月16日：JST-RISTEXとアンケート調査結果報告会でディスカッション

なお、本年度の研究は、国際高等研究所のプロジェクト以外に、科学研究費補助金(挑戦的萌芽研究)「人工知能の規範・倫理・制度に関する対話基盤と価値観の創出」(平成27年度-28年度)、国立情報学研究所・公募型共同研究「情報と社会の系譜学」(平成27年度)によって実施されてきた。国際高等研究所に採択されたのは7月以降であるため、それ以前の活動についてはAIR Newsletter vol.2 (1)を参照されたい。



グループミーティングの様子

今後の計画・期待される効果

人工知能やロボットに対し、「責任ある研究とイノベーション (Responsible Research and Innovation: RRI)」を考えるためには、一般市民も含め多様な利害関係者を巻き込んだ研究体制を作っていくことが各国で課題となっている。そのためには、まだ見ぬ技術に対して問題設定の段階から異分野・ステークホルダー間での視点を把握、共有することが重要である。本プロジェクトでは異分野間での議論とネットワーク形成によって、現在課題となっていることを整理し、新たな学問的テーマが生まれること、また多様な利害関係者を結びつけることによって、人工知能と社会について考えるような土壌を形成することを目的としている。

2016年度は、人工知能の目指すべき共通アジェンダや社会の未来ビジョンの検証と、それを具体的に技術実装の設計標準や、規範・倫理・制度の価値観創出へとつなげていくために必要なことの議論を行う。人工知能を社会実装していくときの社会的規範・倫理・制度設計標準も組み込んだプロトコルの提案などを想定し、具体的なシステムの開発や倫理ガイドラインなどを模索していく予定である。そのためには、システム開発だけではなく、運営・設計論についての評価尺度の考察を行うことも重要となるだろう。