

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	南山大学外国語学部教授
久木田 水生	名古屋大学大学院情報科学研究科准教授
久保 明教	一橋大学大学院社会学研究科准教授
駒谷 和範	大阪大学産業科学研究科教授
西條 玲奈	京都大学園地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
吉澤 剛	大阪大学大学院医学系研究科准教授
吉添 衛	立命館大学情報理工学部学部長



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。

2016年度実績報告

2016年は2015年からの継続調査として(1) オーラルヒストリー調査を行ってきたほか、(2) 多様な価値観への気づきを与える対話型システムの構築という具体的な技術設計についての議論を行ってきた。また(3) 研究ネットワークを広げるため、学会でのオーガナイズドセッションの企画やアイデアブレインストーミングワークショップ等を開催した。

(1) 温故知新プロジェクト

オーラルヒストリー調査として、6月29日に東京大学駒場キャンパスで中島秀之氏(東京大学情報理工学系研究科・特任教授: 人工知能)にインタビューを行い、主に1980年代AIUEOの活動など第二次人工知能ブームについてお伺いした。また、9月23日に京都大学吉田キャンパスにて水谷雅彦氏(京都大学大学院文学研究科・教授: 倫理学)にインタビューを行い、主に1990年代に行われたFINEプロジェクトについてお伺いした。本調査は「AIR温故知新プロジェクト: オーラルヒストリーシリーズ」として書き起こしデータをウェブページで公開する予定であり、中島氏のデータは2016年2月に公開された。

(2) 対話システムの構築

AIRは異分野間の対話・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てきた。また対話を通して、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して掲げてきた。しかし、AIRの研究グループだけではなく、多くの人たちが対話に加わるようなシステムを構築する必要がある。特に価値が多様化してタコツボ化している現在、異なる価値を持つ人やコミュニティが会おうと、想定外の対立や炎上が起きる。そのため、研究者が技術の社会実装を行う前に、多様に変化する価値に気づき予防的に安心して議論や試行錯誤ができる場が必要であるという問題意識のもと、研究者が研究開発の過程からフィードバックを得ることで、異なる価値観に気づく手助けをするシステムのプロトタイプを製作することを目的と掲げ、科学技術振興機構社会技術研究開発センター(JST-RISTEX)「人と情報のエコシステム」研究開発領域へ応募し、採択された。

(3) ネットワークの拡大

AIRの活動の基本は人である。そこで今年も北海道大学で開催された科学技術社会論学会にて「人工知能が浸透する社会を考える」オーガナイズドセッションを開催した。詳細はAIR Newsletter3(1)を参照されたい。また、人工知能と社会の間で生じる様々な問題に対し、学生や脚本家等も巻き込んだアイデアブレインストーミングワークショップも2回開催した。



オーガナイズドセッションの様子

(4) 2016年度 高等研究会実施報告

- 1) 2016年5月12日 グループミーティング@国立情報学研究所
・対話システムの構想や設計と、今年度スケジュールについて打ち合わせ
- 2) 2016年6月23日 グループミーティング@京都大学
・対話システムの構想と、中島秀之氏未来大学学長へのインタビュー打ち合わせ
- 3) 2016年6月29日 グループミーティング@東京大学駒場キャンパス
・オーラルヒストリー調査: 中島秀之氏へのインタビュー
- 4) 2016年7月22日 グループミーティング@東京工業大学キャンパス・イノベーションセンター
・対話システムの構想や設計と、今年度スケジュールについて打ち合わせ
- 5) 2016年8月23日 グループミーティング@京都大学
・対話システムの構想や設計と、英論文について打ち合わせ
- 6) 2016年9月23日 グループミーティング@京都大学
・オーラルヒストリー調査: 水谷雅彦氏へのインタビュー
- 7) 2016年10月19日 グループミーティング@立命館大学東京オフィス
・システムの実装とフィールド調査、学会発表の打ち合わせ
- 8) 2016年11月30日 グループミーティング@名古屋大学
・システムの実装とフィールド調査の打ち合わせ
- 9) 2017年1月9日 グループミーティング@会議室のルビコン
・アイデアブレインストーミングワークショップについての打ち合わせ
- 10) 2017年1月14日 グループミーティング@政策研究大学院大学
・アイデアブレインストーミングワークショップ
- 11) 2017年2月2日 グループミーティング@京都大学
・アンケート調査の進め方と内容と今後の企画についての打ち合わせ

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「情報学と社会の異分野協働方法論の構築: 1980年代AIブームを事例として」(平成28年度)によって実施されてきた。また、2016年11月より、平成28年度科学技術振興機構社会技術研究開発センター(JST-RISTEX)「人と情報のエコシステム」研究開発領域の「多様な価値への気づきを支えるシステムとその研究体制の構築」に採択された。

今後の計画・期待される効果

2017年度は、これまでの議論を受け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の検証と、それを具体的に技術実装のプロトタイプまで落とし込む予定である。そのためにはシステム開発だけではなく、運営・設計論の評価システムの考察を行うことも重要となる。例えばAIRは2016年春に2本の論文を投稿したが、両方とも著者数が多すぎると雑誌側より再考を求められた。しかし本報告は「異分野の研究者が知見を持ち寄り」た成果である。「ギフト・オーサーシップ」が「縦のつながり」を利用したシステムであるのに対し、協同研究、特に異分野間のそれは「横のつながり」を意識的に作り上げようとする試みであり、共著論文は異分野の弱いつながりを公的につなげる橋の役割を果たす。このような主張を受け入れていただき、「情報管理」は共著者全員を、「情報処理」は「AIR」を著者としていただいた。現在、異分野協同研究が推奨されているが、異分野協同研究は参加する研究者だけでなく、所属機関や同僚の理解、研究助成や評価システムの支援があって可能となることを改めて実感する機会となった。AIRの活動をきっかけとして異分野研究グループの活動がさらに活性化されることを期待したい。