

「人工知能と未来」プロジェクトから見る現在の課題 Artificial Intelligence and Future Society: Projects and Prospect

江間 有沙*¹

Arisa Ema

*¹ 東京大学教養学部附属教養教育高度化機構
The University of Tokyo

Various projects and institutions that mission is to draw a big-picture of humanity's future have founded in England and in the USA recently. Artificial Intelligence has been considered as one of the big key research topics that may have great potential of benefits as well as pitfalls. This paper introduces five projects and institutions in England and the USA and then suggests that Japan also have to construct appropriate discussions about social values of AI and responsibility of researchers including both AI and Social Science and Humanities.

1. はじめに

近年、欧米を中心として「人工知能が仕事を奪う」、「シンギュラリティは近い」などの問題意識から、その社会的影響について技術の初期段階から多様な分野の専門家とともに考えていくことを目的とするプロジェクトが設立されている。個別のプロジェクトを見ていくと、哲学的な概念の構築に力を入れているもの、実装研究を行っているもの、主に研究者ネットワークの形成に力を入れているものなど、それぞれに特色があることがわかる。しかしプロジェクトの目的や人的ネットワークを詳細に見ていくと、実は中心となっている人的ネットワークはそれほど複雑ではなく、互いにゆるい関係性を保持しながら、プロジェクトごとに目的のすみわけが形成されている様子が伺える。本稿では、英米の主要なプロジェクトを紹介したうえで、今後日本が考えていくべき点について提案を行う。

2. 「人工知能と未来」関係プロジェクト

2.1 イギリス: Future of Humanity Institute (FHI)

イギリスのオックスフォード大学では哲学科の下部機構として、Future of Humanity Institute (FHI)が2005年に設立された。数学、哲学、科学を用いて技術変化によってもたらされる倫理的ジレンマやリスクなどの検討を行い、長期的な未来のためには人類がどのような選択をしていく必要があるのかを明らかにしていくことをセンターの目的として掲げている。そのため、人工知能だけではなく気候変動や経済破綻などのリスクも扱っている。

所長には、哲学者でありトランスヒューマニスト協会の設立者の一人でもある Nick Bostrom が任命され、彼の著作“Superintelligence: Paths, Dangers, Strategies” (2014)では、人工知能が人類にもたらすかもしれない脅威に言及し、危機を避けるためには研究者たちが取り組んでいかねばならないと警鐘を鳴らしている。その他の研究メンバーには、Stuart Armstrong や Eric Drexler なども名を連ねている。また、2013年に話題となった“The Future of Employment: How Susceptible are Jobs to Computerisation?” (あと10年で「消える職業」「なくなる仕事」702業種を調査した報告書)の著者の一人である Carl Benedikt Frey は Oxford Martin School 所属であり、謝辞には、Stuart Armstrong, Nick Bostromら FHIメンバーの名前がある。

2.2 イギリス: Cambridge Center for Existential Risk (CSER)

イギリスのケンブリッジ大学では、Cambridge Center for Existential Risk (CSER)が、Jaan Tallinn (Skype CEO)らの寄付を受けて人文社会学研究所の下に設立された。オックスフォードの FHI と同様、人文・社会科学部局の下に設立されているのが特徴的である。設立の構想段階では、人間に脅威をもたらす技術について研究をする「ターミネーターセンター」などと言われていたが、実際には人工知能、バイオテクノロジー、ナノテクノロジーなどの最先端技術や科学を概念化し、そのリスクについての哲学的、倫理的な研究が行われている。ワークショップや会議を産業や政策決定者も巻き込みながら開催していくことによって、人工知能の安全なプロトコルや適切な規制の在り方についての議論を推進していくことを目的として掲げている。

研究メンバーには、オックスフォード大学 FHI の Bostrom も名を連ねている。CSERで掲げられている Existential Risk という概念は、Bostrom の2002年の論文によると、“threats that could cause our extinction or destroy the potential of Earth-originating intelligent life (人類を絶滅させる脅威、もしくは地球起源の知的生命体の生じる可能性をなくす脅威)”と定義されており、問題設定の段階でも両機関が相互に連携している様子が伺える。また、ほかにも Stuart Russell, David Chalmers, Dana Scot などアメリカの研究者も名を連ねている。

2.3 アメリカ: The Future of Life Institute (FLI)

アメリカでは、2014年3月にボストン(MIT やハーバード)を中心にボランティアベースの The Future of Life Institute (FLI)が設立された。人工知能などの新技術を人類がうまくコントロールできるように研究を支援することを掲げる本団体は、シンポジウムや研究会を開催するほか、2015年1月に Elon Musk が寄付した1000万ドルを元手として、人工知能に関する研究プロジェクト/センター設立のファンドを設置したことで話題となった。本団体のアドバイザーボードには、FHI の Bostrom のほか、『機械との競争』(2013)や“The Second Machine Age” (2014)の著者の一人でもある Erik Brynjolfsson (MIT)、寄付者である Elon Musk (SpaceX, Tesla Motors)、CSERにも関わっている Stuart Russell (US Berkeley)、理論物理学者の Stephen Hawking (Cambridge University)など14名が名を連ねるほか、ボランティアも多い。また、設立者の一人である Jaan Tallinn は、CSERの設立にも助力するなど、他機関とも関係が深い。

連絡先: 江間有沙, 東京大学教養学部附属教養教育高度化機構, a.ema22@gmail.com

2.4 アメリカ: One Hundred Year Study on Artificial Intelligence (AI100)

One Hundred Year Study on Artificial Intelligence (AI100)はスタンフォード大学に設立され、人工知能の社会や経済に対する長期的(100年)な影響を学際的に考える材料を提供することを目標として掲げている。同センターの元となったのは2008年から2009年にかけてのAAAIの研究グループであり、人工知能については長期的な視点から考える必要があるとのことで、研究グループ期間中AAAI会長であったEric Horvitzと、スタンフォード大学のRuss Altmanが中心となり2014年12月に設立された。スタンフォード大学のマッカーシーが“Artificial Intelligence”という単語を作ったことなどスタンフォード関連の人工知能技術に言及しながら、スタンフォード大学は人工知能が生活に与える影響の研究に力を入れていることを特徴づけている。研究メンバーには法律家なども含まれ、想定しているトピックは、教育や健康など人工知能技術が社会にもたらす変革や法律制度や倫理、経済への影響、犯罪や兵器利用への懸念、人間と機械の共同の在り方や民主主義や自由に対する影響、一般の人々の人工知能への意識や報道の在り方のほか、哲学的な思索など多岐にわたる。

2.5 アメリカ: The Machine Intelligence Research Institute (MIRI)

2000年に設立されたカリフォルニア州パークレーに拠点を置くThe Machine Intelligence Research Institute (MIRI)は、NPO機関である。年度ごとに戦略プランを公開しており、2014年は人工知能がもたらす社会的影響について研究者や社会へ警鐘を鳴らす活動はFHI, CSER, FLIなどの他機関に任せ、MIRIとしてはより根本的な問題、すなわち「人工知能の技術的なアジェンダとは何なのか？」を考えるため、Friendly AI (FAD)の技術開発に注力を注ぐことを目的として掲げている。FAIは、MIRIの所長でもあるEliezer Yudkowskyが作り出した概念であり、人類に良い影響をもたらすSuperintelligenceをつくることを目的としている。この点からわかるようにMIRIは強いAIに関する研究に注目しており、“Intelligence Explosion”(I. GoodとVernor Vinge)やSingularityに関する言説に依拠して研究を行っている。MIRIはSingularity Universityなどと一緒にSingularity Summitなども開催をしている機関でもある。また、MIRIはオックスフォード大学FHIのStuart Armstrongの著書“Smarter Than Us: The Rise of Machine Intelligence”(2014)を出版しており、2015年のリサーチアドバイザーにはBostrom(オックスフォード大学:FHI), Max Tegmark(MIT: FLI)も名を連ねている。

3. 日本の今後の課題: 社会的影響と社会的責任

以上、概略を示した通り、英米の人工知能の未来についての研究プロジェクトやセンターは、依拠している概念や人的ネットワークを介して互いに連携しながら、来るべき未来社会の在り方について議論を進めている様子が伺える。特に、シンギュラリティやトランスヒューマンの議論が強いことも特徴的である。一方、日本においても、人工知能の社会的影響や、人工知能が浸透した未来について考える動きは研究者の間でも高まってきており、人工知能学会に倫理委員会が設立されたほか、人工知能の倫理規範や経済的影響について考える研究会がいくつか設立されている。しかし、学際性などの層の厚みやネットワーク、資金面、そして基盤とする概念の整備など、取り組むべき課題は多い。現在の問題に目を向けると同時に、既存の研究との縦と横のつながりを意識し、知の共有と継承を行っていく地盤を築き上げていくことが重要である。

また、異分野間での共同研究を行う時こそ、そもそも人工知能とは何なのかということも整理していく必要がある。1983年出版の『認知革命』においても「人工知能は、基本的な問題について同意が存在しないことで有名な分野である」と記されているが、それから10年、20年たつて現在においても、人工知能はまだ「体系化に至っていない」[松原 97]、「学会の対象物である「人工知能」がまだ見ぬものである」[松尾・山川 11]などと紹介をされている。このフロンティア精神あふれる考え方が、既存の枠に縛られない価値を生み出していると同時に、人工知能技術の社会的な影響を過小評価していないかを自問する時期に、現在差し掛かっていると言えるだろう。何か問題が起こるかもしれないということを研究者は知り得なくても、それについて知っているように注意すべきだった場合には、責任が問われる可能性もあるとの議論もある[フォージ 13]。研究者の社会的な責任についても、「技術の開発段階だから」という言い訳が通用しなくなってくることを理解し、事前に備えておく必要がある。

最先端の科学技術に対し、その社会的影響を技術の開発段階から考えていこうという考え方は、バイオテクノロジーやナノテクノロジーで2000年以降導入されている考え方であり、近年では、日本でもJST/CRDSの「知のコンピューティング」においてELSI(Ethical, Legal, Social Implications: 倫理的, 法的, 社会的影響)について人文・社会科学者とともに考える必要性などが提唱されている。1990年にヒゲノム計画を契機として米国で1990年に提唱されたELSIに変わり、欧州では2010年以降RRI(Research and Responsible Innovation: 責任ある研究・イノベーション)が学術論文や政策に現れるようになり、研究の初期段階での正と負の規範的課題に、人文・社会科学研究者だけではなく、政策決定者や一般市民も巻き込みながら議論していくことの重要性を掲げている[Owen 12]。情報技術の分野においても、人工知能やクラウドコンピューティング、量子コンピューティングやARなどの先端技術に対しては、技術予測や事前警戒原則を用いて、多様な人々を巻き込んだ議論を展開していくことが提案されている[Schomberg 11]。

人工知能という分野や言葉の持つ特殊性や汎用性を考慮に入れ、そのフロンティア精神に制御をかけず、しかし問題が起こりそうな事態への対応策を事前に考えていくことが、現在求められている。そこでは、未来の社会が「どうなるか」を漠然と考えるのではなく、研究者でありかつ一般市民でもある私たち自身が「どうしたいのか」を議論していくことが重要となってくる[堀 14]。

参考文献

- [フォージ 13] ジョン・フォージ: 科学者の責任—哲学的探究, 産業図書.
- [堀 14] 堀浩一: シンギュラリティへ向けてあなたと私はどうしたいか? 情報処理, 56(1), 41-43.
- [松尾・山川 11] 松尾豊, 山川宏: 人工知能学会 25周年特集「四半世紀を越えて」にあたって, 人工知能学会誌, 26(6), 553.
- [松原 97] 松原仁: 人工知能における「読んでおくべき本」, 人工知能学会誌, 12(1), 36-43.
- [Owen 12] Owen, RP, Macnaghten P and Stilgoe J. Responsible research and innovation; from science in society to science for society, with society, Science and Public Policy, 39, 751-760.
- [Schomberg 11] Schomberg VR. Towards responsible research and innovation in the information and communication technologies and security technologies fields, Publication Office of the European Union, Luxembourg.