

## 小特集「AI原則から実践へ：国際的な活動紹介」にあたって

江間 有沙

(東京大学)

### 1. 「原則から実践へ」

本特集のタイトルにある「原則から実践へ (From Principle to Practice)」とは、2019年3月に米国電気電子学会 (IEEE) の「倫理的に調和したデザイン第1版 (Ethically Aligned Design, 1st edition)」1章の小題でもある。現在、AIの倫理などに関する原則やガイドラインが数多く発行されているが、具体的にどのように取り組めばよいのか、どのような組織的な取り組みが行われているのか、そもそも誰が主要なステークホルダなのかは実は外からは見えにくい。

国際的でマルチステークホルダの議論が重要とされつつも、その多様なステークホルダが実際は固定化されている、との批判もある。筆者自身いくつかの国際的なマルチステークホルダの会議に参加しているが、毎回ほぼ同じ人達の集まりである場合もある。また、国際的な議論の場に参加している日本人はごく少数でもある。

そこで本小特集では主に国際的にAIの倫理やガバナンスのコミュニティやネットワーク形成を主宰・実働している方々に、彼女らの活動を紹介してもらうと同時に、現在の実践への取り組みとその課題、今後の展望として日本や日本の研究者コミュニティへ期待することを記述してもらった。

### 2. 本小特集の構成と著者・組織紹介

各論文の論点は別稿でコメントリとして紹介したので、「特集にあたって」では、簡単にではあるが本小特集の執筆者達と、彼女達の活動の背景を紹介したい。

第一稿の執筆者であるガル氏の現在の所属はケンブリッジ大学であるが、彼女は同時にさまざまな国際・学際ネットワークのハブとなっている。まず彼女は「倫理的に調和したデザイン」を公開したIEEEの実行委員会委員を務めている。またIEEEはこれに関連するいくつかの規格を議論しているが、そのうちP7009自律的・反自律的なシステムのフェールセーフ設計標準の副議長も務めている。また、北京AIアカデミーのAI4SDGs協力ネットワーク<sup>\*1</sup>の発起人の一人でもある。さらには2020年12月末まで国連の事務次官・特別顧問室の技術アドバイザとしてAIガバナンスに関する議論のとりま

とめも行ってた。2017年度人工知能学会全国大会「公開討論：人工知能学会倫理委員会」でもパネリストとしてご参加いただいている<sup>\*2</sup>。

二稿目の著者であるバク氏はPartnership on AI (PAI)で現在、Methods for Inclusionプロジェクトをリードしている。PAIはGoogle社やAmazon社、Facebook社といった巨大IT企業が初期設立メンバであったが、現在の参加メンバの過半数は非営利団体であり、公式Webサイトは13の国から100以上のパートナーが参加していると報告している。日本からもソニー、東京大学次世代知能科学研究センター、ソフトバンクなどが加盟している。PAIはAIにおける公平性の問題や雇用への影響、安全性といった個別な論点に関する調査・研究を行っている。このうち、バク氏がリードするMethods for Inclusionは最近始まったプロジェクトであり、積極的に調査や活動を行っている。

三稿目の著者であるランキスト氏が所属するThe Future Societyも近年その存在感を増している非営利団体である。2014年にハーバード大学ケネディスクール発祥のシンクアンドドゥタンクとして設立され、近年では国際機関や各国政府とともにAI政策の研究や教育プロジェクト、アドバイザリサービスなどを展開している。最近では、フランスとカナダのイニシアティブで開始された国際的なAIに関する議論をするネットワークであるGPAI (Global Partnership on AI) の「責任あるAI」分科会でレビューや分析のアドバイザもしている。

四稿目の著者であるシャーダンド氏はLighthouse3というAI利活用におけるコンサルティング企業のCEOを務める傍ら、Women in AI Ethics™というネットワークを立ち上げた。彼女は、AI倫理やガバナンス領域で活躍している女性の支援を思い立ち、2018年にたった一人で100人の活躍している女性を選んで公開した。それが反響を呼び、現在はAI倫理やガバナンス領域で活動をする女性やマイノリティの人達のイベント支援のほか、コロナ禍で支援を必要としていた学生や研究者への支援プログラムを打ち出すなど、さまざまな企画を展開している。

五稿目の著者であるアフアナシェワ氏はGoodAIの

\*1 <http://www.ai-for-sdgs.academy/ai4sdgs-cooperation-network>

\*2 <http://ai-elsi.org/archives/583>

\*3 <https://www.partnershiponai.org/partners/>

COOである。GoodAIは、2014年にマレク・ローザ氏が1,000万ドルを投資して汎用人工知能(AGI)の開発を目的として設立した。まだ見ぬ汎用人工知能は、実現すれば多大な利益をもたらすだろう一方、だからこそ開発や運用、利用方法に関しては慎重に議論していかなければならない。このような目的意識から、GoodAIは「汎用AIチャレンジ」を開催し、AI競争が起こらないようにするための知恵や知識を募っている。アフアナシュワ氏とローザ氏は2017年秋に東京で開催された「AIと社会シンポジウム」\*4にも登壇しており、当時は構想段階であった「汎用AIチャレンジ」の結果を今回、うかがうことができた。

AIの諸課題に対応するためにはマルチステークホルダーでの議論が不可欠になる。それを表すかのように、彼女達の所属はアカデミア、国際機関、非営利団体や企業など多様であり、かつ一つに縛られない。組織的な流動性や活動範囲の広さが本小特集で扱っているような問題を扱っていくため重要であることを示唆している。

一方で時間の関係上、執筆者が筆者の知り合いに今回は限られてしまい、欧米に集中してしまったことは特記しておくべきであろう。AIに関する議論は中東、東欧、アフリカ、東アジア、東南アジア、オセアニアでも素晴らしい人達がリーダーシップをもって議論を行っており、議論主体の多様性も増してきている。

なお多様性という観点では、気付いた方もおられるかもしれないが、本小特集の執筆者は全員女性である。本小特集の構想段階で候補をあげていくうち、大半が女性であったこともあり、筆者の独断と偏見ですべて女性の執筆者で構成することに決めた。シャーダンド氏の100人リストをはじめ、世界各国でAI倫理やガバナンスの領域で活躍している女性が多い。「多様性」という観点からは逆に偏った特集の構成になってしまったが、将来的にはこの偏りをあえてつくらずとも本学会誌の全体としての執筆陣の多様性が確保されるような状況になってほしいと願っている。

### 3. おわりに

最後に、本小特集をまとめるのにご尽力いただいた方々に感謝申し上げたい。第一に、執筆者の方々には、年末年始を挟む1か月程度という短い期間で執筆をご快諾いただき感謝申し上げます。快く引き受けてくれた背景としては日本に対する期待があるからだろう。AIがもたらすさまざまな課題のうち、多様性や包摂性に関する議論は国際的にも深刻である。「多様性や包摂性を重要視する会議」の参加者の多様性がないとする批判は、日本の「人間中心のAI社会原則」をはじめ、さまざまな組織に対して行われている。特に、国際的な議論においては、欧米諸国の参加者が多い中、アジアの島国で世界第3位の経済大国かつ技術大国でもある日本が参加することへの期待は大きい。

また時間の都合上、1週間(のうち全日は使えないので、毎日夜の2~3時間と土・日曜日)で全翻訳をしなければならなかった筆者の助けとなったのは、複数の機械翻訳サービスと検索エンジンであった。原文を複数の機械翻訳サービスにかけ、専門用語に関しては筆者の知識・経験を総動員するほか論文や記事の検索をしたり、時に音声認識ソフトも駆使したりしながら翻訳原稿をつくり上げていった。ある意味「人とAIの協同作業」であったといえる。また、清田陽司編集委員長には、原稿の整形のみならず表現の確認訂正、引用文献の確認など週末の夜中までお時間をいただいご対応いただき、大変お世話になった。

本特集を通じて、より多くのAI関係者が国際的な議論へと参加していただく足掛かりを提供できれば幸いである。なお、本小特集の原文(英語)は東京大学のWebサイト(<https://ifi.u-tokyo.ac.jp/en/projects/future-ai/>)から閲覧可能となる予定である。

\*4 <http://www.aiandsociety.org/>、なお本シンポジウムの報告書は、本誌 Vol. 33, No. 2, pp. 189-222 (2018) にて特集として掲載されている。