



2016年10月11日

人工知能の発展に伴う安全性と倫理責任問題

公益財団法人 国際通貨研究所
経済調査部 主任研究員 志波和幸

コンピューターの発達とともに研究が進められてきた人工知能（以下、「AI」）の進化は、ここ数年目を見張るものがある。1950年代以降のAI研究は何度かの興隆と冬の時代の浮き沈みを経つつ、近年の「ディープラーニング¹」の登場により急速に開発が進み、研究レベルから実用化へのステップに踏み出した。

実際、米国民間調査会社である Tractica 社は、ディープラーニングの登場により、金融機関の後方事務、画像解析、自動翻訳、自動運転等、AIの様々な応用が可能となったことから、多くの企業でAIの導入が進むと見ており、その世界市場は2015年の2億250万米ドルから2024年には111億米ドルまで拡大すると予想している²。

しかし、AIが人間と同じレベルに近づくにつれ、「人間がAIに職を奪われるのではないか」、「AIが人間に危害を加えるのではないか」、「ビッグデータ活用の際、個人情報（プライバシー）をきちんと守ることが出来るのか」という漠然とした不安が湧いてきているのも事実である。人間とAIとの関係を構築するにあたり、実際の生活やビジネスにおいて直面する社会倫理や責任問題について早急に指針を作る必要性が生じている。

こうしたなか、米国ではAIに対する上記懸念を払拭し、その普及を後押しするべく、2つのNPO（非営利団体）が設立されたことは注目されよう。

両NPOの出資母体（会社か個人か）の違いはあるが、いずれも

- (1) AIがいかに人類に貢献し、かつ安全であるかを社会全般に対し認知を向上させること。
- (2) 上記向上のため、AIが社会に及ぼす影響・懸念・課題について、様々な分野の専門家から助言を受けながら、業界の自主的な基準はルール作りに繋げること。

¹ 「ディープラーニング」の概略については、当研究所レポート「IIMAの目 2016年第44号(2016年6月27日付『Fintechは伝統的金融機関を破壊するのか？(ディープラーニングによる融資審査機能自動化の可能性と限界)』)」に詳述。

http://www.iima.or.jp/Docs/column/2016/0627_j.pdf

² <https://www.tractica.com/newsroom/press-releases/artificial-intelligence-for-enterprise-applications-to-reach-11-1-billion-in-market-value-by-2024/>

- (3) 一部の企業や政府機関が独占する強力な AI が悪用されないよう、参加企業の足並みを揃える (Partnership on AI) または将来的には AI に関する全ての特許を世界に開放する (Open AI) こと。

を目的としている。そして、この運動を通じ、世界各国の AI 政策に対する主導権・発言力を高めようとする狙いが透けて見える。

NPO名	Partnership on AI	Open AI
設立日	2016年9月28日	2015年12月11日
出資母体	Google Facebook Microsoft IBM Amazon DeepMind (Appleにも参加呼び掛け)	Elon Musk氏 (Tesla社CEO) Sam Altman氏 (Ycombinator社長) Peter Thiel氏 (Clarium Capital社長) Reid Hoffman氏 (LinkedIn共同設立者) 等
設立主旨	AIのアプリケーションが人と社会に有益であることを確認すること	人類全体に利益が最大限貢献される方向にAIを進化させること

一方、我が国に目を転じると、今年2月に総務省情報通信政策研究所で「AI ネットワーク化検討会議」が開催され、6月2日に公表された「AI ネットワーク化検討会議報告書 2016³⁾」で、AI ネットワーク化の進展が社会にもたらす影響を評価するための指標やリスク・シナリオ分析をするとともに、開発原則及びその指針(ガイドライン)の策定に向けた国内外の議論等の必要性を強調した。

また、4月にその中間報告書で提示した「当面の課題等」を踏まえ、同月29日・30日に「G7 香川・高松情報通信大臣会合」が開催された。そこで、日本の高市早苗総務大臣が、G7各国が中心となり、産学民官の関係ステークホルダーの参画を得て、AI ネットワーク化が社会・経済に与える影響や AI の開発原則の策定などについて国際的な議論を進めることを提案し、各国からの賛同を得た。

³⁾http://www.soumu.go.jp/menu_news/s-news/01iicp01_02000050.html

AIの研究開発に関する原則（開発原則）の策定

OECDプライバシーガイドライン、同・セキュリティガイドライン等を参考に、関係ステークホルダーの参画を得つつ、AIの研究開発に関する原則を国際的に参照される枠組みとして策定することに向け、検討に着手することが必要。

研究開発に関する原則の策定に当たっては、少なくとも、次に掲げる事項をその内容に盛り込むべき。

- ① 透明性の原則
AIネットワークシステムの動作の説明可能性及び検証可能性を確保すること。
- ② 利用者支援の原則
AIネットワークシステムが利用者を支援するとともに、利用者に選択の機会を適切に提供するように配慮すること。
- ③ 制御可能性の原則
人間によるAIネットワークシステムの制御可能性を確保すること。
- ④ セキュリティ確保の原則
AIネットワークシステムの頑健性及び信頼性を確保すること。
- ⑤ 安全保護の原則
AIネットワークシステムが利用者及び第三者の生命・身体の安全に危害を及ぼさないように配慮すること。
- ⑥ プライバシー保護の原則
AIネットワークシステムが利用者及び第三者のプライバシーを侵害しないように配慮すること。
- ⑦ 倫理の原則
ネットワーク化されるAIの研究開発において、人間の尊厳と個人の自律を尊重すること。
- ⑧ アカウンタビリティの原則
ネットワーク化されるAIの研究開発者が利用者当関係ステークホルダーへのアカウンタビリティを果たすこと。

このように、国内外においてAI及びAIネットワークシステムに関する技術開発や活用方法が日進月歩で進展するとともに、それらの高度化をめぐる社会的・経済的・倫理的課題に対する関心も急速に広がっている。

我が国としては、その技術開発を絶え間なく進めることは勿論のこと、海外の動向・情勢に十分注視し、世界共通のAIネットワーク開発・利用に関するガイドライン・ルール等の策定の際に一定のイニシアチブがとれるよう、産官学間の緊密なコミュニケーションを構築するとともに、人間とAIとの関わり方に関するビジョンを早急に明確にすることが必要となろう。

以上

当資料は情報提供のみを目的として作成されたものであり、何らかの行動を勧誘するものではありません。ご利用に関しては、すべてお客様御自身でご判断下さいますよう、宜しくお願い申し上げます。当資料は信頼できると思われる情報に基づいて作成されていますが、その正確性を保証するものではありません。内容は予告なしに変更することがありますので、予めご了承下さい。また、当資料は著作物であり、著作権法により保護されております。全文または一部を転載する場合は出所を明記してください。