

人工知能と社会

人工知能 (Artificial Intelligence: AI) は、多くの恩恵をもたらす一方で、グローバルな協調と包摂的なガバナンスをもって対処すべき重大な懸念も惹起しながら、我々の世界を作り変えようとしている。2023年10月、G7首脳は、高度な AI システムを開発、利用する公的機関及び民間組織向けの「広島 AI プロセス国際指針」及びそれに付随する「行動規範」を発表した。我々は、この共同声明において、AI がもたらす技術的・社会的懸念に関し、政策立案者への提言リストを提示することにより、G7サイエンスアカデミーの立場を表明する。

1. 強力な AI システムは、悪用されると社会に危険を及ぼす可能性があるため、サイバー攻撃や物理的な攻撃から適切に保護されなければならない。機密性の高い AI アプリケーションの場合、リスクのレベルに応じたデータ保護の枠組み（例えば、欧州連合 (EU) の一般データ保護規則）やプライバシーを強化する技術の活用が、個人データを保護するためには不可欠である。AI システムとやりとりする際に、個人データがどのように、またどの程度の期間、利用・再利用・保存されるのかを把握できるよう、利用者はデータ保護のための明確なガイダンスを必要としている。
2. AI システムの性能が更に向上し、人間による制御により大きなリスクを与えるようになるにつれて、システムの導入を行う前に、安全性についての公的保証が求められるようになるだろう。要件の検証、システムの妥当性確認及びテスト、並びに導入後の監視に関する厳格な基準に従うことは必須であり、隠れたバイアスのチェックについても独立機関が行うことが理想的である。AI アプリケーションの性能、限界及び適切なドメインについて利用者が理解できるよう、文書化の必要性を義務づける法整備を推進すべきである。AI モデルの構築に関する基礎となるデータのトレーサビリティは極めて重要である。AI は医療、科学、法律、金融等の分野において重大な意思決定に影響を与えるため、透明性と説明責任が何よりも重要である。AI システムが開発・運用される際には、説明責任のための明確な道筋が整備されるべきである。説明可能性、アルゴリズムに対する是正措置、是正のための手段等の機能が備われば、AI の決定に異議を唱える個人の能力を強化できる。
3. AI システムの信頼性は極めて重要である。人間が AI システムとのやりとりに一層時間を費やすようになり、かつ我々の情報エコシステムにおいて AI が生成するコンテンツの割合が大きくなると、民主主義・社会・人間の理解に対するリスクが極めて重要になる。各国政府は、AI が生成するコンテンツを対象にした、法的拘束力のある基準を推進しなければならない。正確性と信憑性を促進するための制度的な仕組みがますます重要になる。さらに、AI が生成するコンテンツをいかに識別し、それらといかに関わり合うべきかを理解するためのリテラシーを、人々が生涯にわたって高めることが肝要となろう。技術開発者とオンライン・メディア・プラットフォームは、社会技術評価（例えば、レッドチーム演習）にドメインの専門家を参加させることはもちろん、デジタル・コンテンツの出所を示すシステムの有用性も検討すべきである。

4. 生成 AI はコンテンツの作成やソフトウェア開発において大きな可能性をもたらすが、AI が既存の作品を基にして外見的には新しいコンテンツを創り出した場合、知的財産権の保護に関する諸々の問題が浮上する。生成 AI から大きな影響を受ける主な分野として、クリエイティブ産業や科学研究等が挙げられる。著作権、フェアユース、知的財産権保護に関する既存の法的枠組みは、AI 時代に備えて抜本的に見直す必要がある。このような諸問題には、明確な立法、産業界の自主規制、第三者による監視を組み合わせて対処しなければならない。
5. 大規模な AI システムは高価な計算資源を必要としており、それが参入障壁となっている。AI の公平な未来を築くためには、世界中から高性能コンピューティングにアクセスできる機会を広げるとともに、政府間の研究ハブを設立することが非常に重要である。AI の計算資源を拡充すると、エネルギー消費量が増大するため、環境に対する懸念が引き起こされる。AI と環境倫理を両立させるためには、エネルギー効率の高い新たな AI 技術に加え、持続可能なデータセンターの設計やエネルギー効率と冷却を考慮した戦略的立地が不可欠である。
6. 著作権問題の発生しない、適切に収集・選別・整理等が行われたデータセットのオープンソース化は、一般的には公益に資するものとして奨励されるべきであるが、大規模なオープンソース AI モデルは、(例えば、公共部門の研究者やそれを用いる小規模のアプリケーション開発者にとっての) 便益と人々が負わされることになる悪用から生じるリスクとの間に複雑なトレードオフ関係を生じさせる。強力な AI システムのオープンソースの普及の許可又は制限に関する決定は、民主的な監視下に置かれなければならない。また、プロプライエタリ・システム(独自の専用システム)に適用される安全規制は、オープンソースシステムの全ての事例にも適用されなければならない。
7. 自律型兵器システムの台頭は、急を要する倫理的かつ軍事的な懸念を生じさせている。その自律性は紛争の敷居を下げ、軍拡競争に拍車をかけ、ならず者政権又はテロリストによる悪用の危険性を招く恐れがある。このような兵器の禁止を主張する声もあれば、人間による監視を確実に行いつつ国際人道法に従って厳格に規制するよう求める声もある。社会による迅速な行動が、この重要な課題に取り組むための鍵となる。個人・地域・国家・国際レベルにおける物理的な安全保障と同様に、人間の価値観及び国際規範を守ることも目指しつつ、これらの課題への対処法を切り開くには、公的な透明性及び議論が極めて重要である。
8. AI システムが影響力を持つようになるにつれて、AI を人間の価値観や倫理規範に適合させる必要性が求められている。倫理的な AI の実装には、文化的多様性も考慮しながら、技術的側面と社会文化的側面を統合し、部門の垣根を越えて協力することが不可欠である。社会的弱者に悪影響を及ぼす可能性がある短期的リスクに対処することは、信頼を維持し、有益な技術の導入を促進するために非常に重要である。(AI を人間の価値観や倫理規範と) 適合させるための経済社会的コストは、透明化されるべきであり、AI システム利用のタイミング及び方法に関する意思決定過程に含めるべきである。

9. AI の恩恵を均等に分配するためには、効果的なステュワードシップ（管理運用の責任体制）が極めて重要である。ますます高性能になる AI システムは労働市場に多大な影響を与えることになるだろうが、その恩恵が公平に配分されるには、政府、産業界、教育者及び市民社会が共同で取り組むアプローチが不可欠である。政府は、既存の役割において単に人間に取って代わるシステムではなく、むしろ満たされていないニーズに応えるような社会的に価値のある AI システムへのインセンティブを検討すべきである。地球、社会、そして経済にとっての最重要課題に対処する AI システムの開発を優先する政策を策定すべきである。
10. 官民の協力が責任ある AI 開発を促進する。民間部門が迅速に技術革新を行う一方で、公的部門は規制及び倫理枠組を提供し、人類の利益のための独立した長期的研究を促進する。（官民の）人材交流によって、ベストプラクティスと目標を融合させることができる。この相乗効果によってイノベーションに説明責任が伴い、AI の社会的利益が最大限に生かされることになる。そして社会が、規制当局や民主的プロセスを通じて、何が許容できて何が許容できないかという、超えてはならない一線を引き、かつ開発者が自らのシステムがこの一線を越えていないことを証明する責任を負う場合に、AI は最適な役割を果たす。
11. AI の責任ある安全な開発と普及は、単に科学的、技術的な問題にとどまらない。AI が作動する複雑な社会技術システムを理解し構築し、適切なガバナンスの仕組みを考案するためには、人文科学や社会科学を含む多様な学術分野の分野横断的な協働アプローチが求められる。技術的な視点と非技術的な視点を融合することにより、包摂的で強固な、かつ透明性があり平等な AI の普及を確保しつつ、偏見や公平性といった AI の諸課題に対処することができる。
12. 教育は、AI の能力を明確にし、神話を否定し、AI の長所と短所を比較評価しなければならない。市民は技術的にリテラシーを高め、AI が示唆する社会的、倫理的な問題を認識すべきである。様々な利害関係者が促進する公開対話によって、個人が能動的に AI 政策を形成したり、AI を公共の利益に適合させたりすることが可能になる。教育の枠組みに強力な AI システムを組み込む際には、リスクと機会のバランスを慎重に取る必要がある。