

要 旨

1 作成の背景

急速に進展する生成 AI には、あらゆる学術分野、産業分野、そして社会全体に大きな影響を持つという包括性、将来的には人間と共存する知的レベルとなり得る革新性、さらに、それが加速度的に進展するという加速性などの特徴がある。それゆえに、脅威や課題が存在するとともに、社会への大きな波及効果があり、人類社会の重要課題に対して解決策を提供する可能性がある。

このような状況のもとで、生成 AI の現状と動向、脅威と課題、活用による波及効果について、学術の立場から深く洞察し、生成 AI を受容・活用する社会の実現に向けてどのような施策をとるべきかについて提言をまとめる。

2 現状および問題点

生成 AI は、2020 年代以降急速にその技術が発展・普及し、特に大規模言語モデル (Large Language Model : LLM) を基盤とする ChatGPT は、2022 年 11 月に公開後 2 ヶ月でアクティブユーザー数が 1 億人に達した。LLM は Transformer と呼ばれるニューラルネットワークモデルが基盤となっており、モデルサイズ (パラメータ数) と学習データ量に対して、対数スケールで性能が向上することが知られている。言語だけでなく、画像・映像、音声・音楽などと統合するマルチモーダル処理、ロボティクスへの応用も進んでいる。

生成 AI には様々なリスクが存在する。生成 AI の活用においては、ハルシネーション (事実と異なる情報を出力すること)、高品質な生成メディアによる詐欺や世論操作、非社会的な回答の生成、機密情報漏洩などの懸念がある。また、著作権侵害、名誉毀損、芸術的活動への脅威、社会の価値観や文化への影響などの懸念もある。これらの問題に対処するために、生成 AI モデルには正確性、指示追従性、頑健性、透明性、説明可能性などが求められる。

一方で、生成 AI には人類社会に資するプラスの面が大いに期待されている。科学技術においては、生成 AI の活用により仮説生成、仮説検証、論文による知識流通など科学技術の各ステップが大きく進展しようとしている。知識の細分化・専門化による学問領域の分断が指摘される中、生成 AI の活用による分野横断的な新たな知の創造も期待される。産業界においても、生成 AI の活用による業務効率化・業務量削減が、労働力不足や長時間労働などの社会課題を解決する強力な一手となり得る。さらには、教育の効率化や高度化、文章の推敲・要約・翻訳、挨拶文やメールの作成支援、旅行計画の提案、投資のアドバイス、音楽・アート・デザインの生成など、我々の日常的な活動への波及も始まっている。

3 提言の内容

生成 AI の世界的進展が留まる気配のない中で、我が国は、リスク対策についても十分に工夫をしながら、生成 AI の研究開発や社会での活用を積極的に進め、人類と AI の共存社会のデザインで世界をリードすべきである。

(1) 生成 AI 研究開発の望ましい体制

- ① 日本の技術競争力を強化するため、国家戦略として生成 AI の研究開発を推進すべきである。特に、オープンな研究開発の取り組みへの支援を重視・強化することが必要である。
- ② 日本国内の生成 AI 研究者コミュニティの強化と国際的研究連携の推進が必要である。プライバシーやセキュリティに配慮したデータインフラの構築を支援するとともに、公共データの開放や産業界とのデータ共有プロジェクトを奨励すべきである。
- ③ 生成 AI による判断や行動が、人間の価値観や倫理観に合うことが極めて重要であり、学習データや学習手法を含む開発プロセスの透明性を確保すること、AI の設計・開発・評価においてガイドラインを作成してリスクを最小化すること、AI ガバナンスの国際的ルールメイキングに日本の考え方を反映させる体制を構築することが必要である。

(2) 生成 AI モデルの適切な運用

- ① 生成 AI モデルがサイバー攻撃や物理的攻撃から適切に保護される必要があり、これらの攻撃を検知・回避する頑健なシステムが構築されるべきである。
- ② AI 技術に起因する問題が発生した場合に、迅速かつ適切に対処できる体制を整えることが必要である。また、国際的な協力を通じて、AI 技術の標準化やベストプラクティスを共有し、グローバルな視点での AI の発展と運用を推進することが重要である。
- ③ 人間中心の原則に基づく持続可能な社会の実現に向けて、市場原理や競争原理に任せるのではなく、地球規模の課題や社会・経済にとって最重要な問題に対して AI の利活用・運用を優先すべきである。

(3) 責任ある生成 AI 実装に向けた制度設計

- ① 従来型の規制モデルでは、複雑で変化の速い AI がもたらす様々なリスクに適切に対処することができない。アジャイル（迅速かつ反復的）でマルチステークホルダー型のガバナンスを志向すべきである。
- ② 政府は、オープンな場でのルール形成の促進、事故調査への関係者の積極的な協力を促す制度設計、事故被害者に対する迅速な救済制度の設計などを行う必要がある。

- ③ 民間主体は、政府を含むステークホルダーに十分な質と量の情報開示を行うとともに、ステークホルダーからのフィードバックを得て、常にガバナンスのあり方を改善する必要がある。

(4) 生成 AI モデル以降の教育とリテラシー

- ① 社会全体で生成 AI の教育やリスキリングに取り組んでいくことが必要であり、それを推進するためのリテラシーを持つ人材の養成と教育プログラムの推進、リスキリング支援があまねく必要である。この際、地域格差に配慮し、むしろ地域格差を解消することを目指すべきである。
- ② 慎重な議論を行った上で、AI の活用を前提として AI との共存を目指した新たな教育への転換を図るべきである。従来の知識の伝達に偏重するのではなく、AI を批判的に利用し、課題を解決し、創造する能力を高める教育・カリキュラムが必要である。また、新たな教育について情報共有・議論する場を国として支援することも重要である。
- ③ AI の活用は、学術を学際的に深め、複合的な社会課題の解決につながるが、そのためには、科学者が高い AI リテラシーを身につけることが必要であり、学術分野間および産学間の対話・連携の促進が必要である。