

G 2 0 A I 原則

G 2 0 は以下のセクション 1 の信頼できる A I のための責任あるスチュワードシップに関する原則を支持し、セクション 2 の勧告について留意する。

1. 信頼できる A I のための責任あるスチュワードシップに関する原則

1. 1 包摂的な成長、持続可能な開発及び幸福

ステークホルダーは、人間の能力の増強や創造性の向上、少数派の包摂の促進、経済・社会・性別などにおける不平等の改善、自然環境の保護を通じ、包摂的な成長、持続可能な開発及び幸福の活性化のような人々と地球にとって有益な結果を追求することにより、信頼できる A I のための責任あるスチュワードシップに積極的に取り組むべきである。

1. 2 人間中心の価値観及び公平性

- a) A I のアクターは、A I システムのライフサイクルを通じ、法の支配、人権及び民主主義的な価値観を尊重すべきである。これらには、自由や尊厳、自主自律、プライバシーとデータの保護、無差別と平等、多様性、公平性、社会正義、国際的に認められた労働権が含まれる。
- b) このため、A I のアクターは、人間による最終的な意思決定の余地を残しておくことなど、状況に適した形で且つ技術の水準を踏まえたメカニズムとセーフガードを実装すべきである。

1. 3 透明性及び説明可能性

A I のアクターは A I システムに関する透明性と責任ある開示に取り組むべきである。このため、A I のアクターは下記の目的で、状況に適した形で且つ技術の水準を踏まえた有意義な情報提供を行うべきである：

- i. A I システムに関する一般的な理解を深めること。
- ii. 職場におけるものを含め、A I システムの関与をステークホルダーに認識してもらうこと。
- iii. A I システムに影響される者がそれから生じた結果を理解できるようにすること。
- iv. 要因に関する明快且つ分かりやすい情報、並びに予測、勧告あるいは判断の根拠となった論理に基づいて、A I システムから悪影響を受けた者がそれによって生じた結果に対して反論することができるようすること。

1. 4 頑健性、セキュリティ及び安全性

- a) A I システムは、通常の使用や予見可能な使用及び誤用あるいはその他の悪条件下においても正常に機能するとともに、不合理な安全リスクをもたらすことがないよう、そのライフサイクル全体にわたって頑健且つセキュリティが高く、安全なものであるべきである。
- b) このため、A I システムの出力に関する分析と問合せに対する対応を可能とするため、A I のアクターは、データセットやプロセス、A I システムがそのライフサイクルに行った決定に関することも含め、状況に適した形で且つ技術の水準を踏まえたトレーサビリティを確保すべきである。

c) A I のアクターは、その役割や状況、能力に基づき、系統化されたリスク管理のアプローチを A I システムのライフスタイルの各段階に適用することにより、プライバシーやセキュリティ、安全性、バイアスといった A I システムに関するリスクに絶え間なく対処していくべきである。

1. 5 アカウンタビリティ

A I のアクターは、その役割と状況に基づき、また、技術の水準を踏まえた形で、A I システムが適正に機能していることと上記の原則を順守していることについて、アカウンタビリティを果たせるようにすべきである。

2. 信頼できるAIのための国内政策と国際協力

2. 1 AIの研究開発への投資

- a) 信頼できるAIの実現に向けたイノベーションを促進するため、各国政府は、学際的な取組を含め、技術的に困難な課題やAIの社会的・法的・倫理的な影響と政策課題に焦点を当てた調査研究及び研究開発について、長期的な公共投資を検討し、また民間投資を奨励すべきである。
- b) また、各国政府は、不適切なバイアスがなく、相互運用性と技術標準の利用を増進するため、十分な代表性を有し、且つプライバシーとデータの保護を順守する開かれたデータセットについて、公共投資を検討し、また民間投資を奨励すべきである。

2. 2 AIのためのデジタル・エコシステムの整備

各国政府は、信頼できるAIのためのデジタル・エコシステムとそれへのアクセスを整備すべきである。このエコシステムには、デジタル・テクノロジーとデジタルインフラ、必要に応じてAI知識を共有するためのメカニズムが含まれる。これに関連し、各国政府は、データ・トラストのような、安全、公平、適法且つ倫理的にデータを共有するためのメカニズムに対する支援を検討すべきである。

2. 3 AIを推進するための政策環境の整備

- a) 各国政府は、信頼できるAIシステムが研究開発の段階から展開・稼働の段階への迅速な移行を支援するための政策環境を整備すべきである。このため、政府は必要に応じてAIシステムの実験と拡張のための制御された環境下での実証実験の活用を検討すべきである。
- b) 各国政府は、信頼できるAIの実現に向けたイノベーションと競争を奨励するため、必要に応じ、AIシステムに適用される政策及び規制の枠組みやその評価メカニズムの見直しと改正を行うべきである。

2. 4 人材育成及び労働市場の変化への準備

- a) 各国政府は、仕事の世界と社会全体の変化に備えるためにステークホルダーと緊密に協働すべきである。政府は人々に必要なスキルを習得させるなどして、人々が広い範囲で適用されるAIシステムを効果的に利用し、それとうまく関わることができるようすべきである。
- b) 各国政府は、就労期間を通じたトレーニング・プログラムや離職を余儀なくされた者への支援、労働市場における新たな機会へのアクセスなどを通じ、AIの普及がもたらす労働市場の変化が労働者にとって公平なものであるよう万全を期すため、社会的な対話などの措置を講じていくべきである。
- c) 各国政府はまた、職場におけるAIの責任ある利用の推進、労働者の安全及び仕事の質の向上、起業家精神と生産性の向上、AIの恩恵の幅広く且つ公平な共有が確保されるようにするため、ステークホルダーと密接に協働すべきである。

2. 5 国際協力

- a) 開発途上国を含め、各国政府は、ステークホルダーとともに、これらの原則を推進し、信頼できるAIのための責任あるスチュワードシップを促進するために積極的に協力すべきである。

- b) 各国政府は、必要に応じてO E C Dやその他の世界的及び地域的な国際場裏において、A I 知識の共有を推進するために協働すべきである。政府はA I に関する長期的な専門知識を蓄積するためには、国際的且つ分野横断的であり、マルチステークホルダーによる開放的な取組みを奨励すべきである。
- c) 各国政府は、相互運用性があり、且つ信頼できるA I の実現のため、マルチステークホルダーの合意に基づく世界的な技術基準の開発を推進すべきである。
- d) 各国政府はさらに、A I の研究開発や展開を測定すると共に、これらの原則の履行の状況を評価する際に根拠となる証拠を収集するため、国際的に比較可能な測定基準の開発とその利用を奨励すべきである。