

全ての AI 関係者向けの広島プロセス国際指針

1. 我々は、安全、安心、信頼できる AI を適切かつ関連性をもって推進する上での、全ての AI 関係者の責任を強調する。我々は、ライフサイクル全体にわたる関係者が、AI の安全性、安心、信頼性に関して、異なる責任と異なるニーズを持つことを認識する。我々は、全ての AI 関係者が、自らの能力とライフサイクルにおける役割を十分に考慮した上で、「高度な AI システムを開発する組織の向けの広島プロセス国際指針(2023 年 10 月 30 日)」¹を読み、理解することを奨励する。
2. 「高度な AI システムを開発する組織の向けの広島プロセス国際指針」の以下の 11 の原則は、高度な AI システムを開発する組織にのみ適用可能な要素もあることを認識しつつ、高度な AI システムの設計、開発、導入、提供及び利用をカバーするために、全ての AI 関係者に対し、適時適切に、適切な範囲で、適用されるべきである。
 - ① AI ライフサイクル全体にわたるリスクを特定、評価、軽減するために、高度な AI システムの開発全体を通じて、その導入前及び市場投入前も含め、適切な措置を講じる
 - ② 市場投入を含む導入後、脆弱性、及び必要に応じて悪用されたインシデントやパターンを特定し、緩和する
 - ③ 高度な AI システムの能力、限界、適切・不適切な使用領域を公表し、十分な透明性の確保を支援することで、アカウントビリティの向上に貢献する
 - ④ 産業界、政府、市民社会、学界を含む、高度な AI システムを開発する組織間での責任ある情報共有とインシデントの報告に向けて取り組む
 - ⑤ 特に高度な AI システム開発者に向けた、個人情報保護方針及び緩和策を含む、リスクベースのアプローチに基づく AI ガバナンス及びリスク管理方針を策定し、実施し、開示する
 - ⑥ AI のライフサイクル全体にわたり、物理的セキュリティ、サイバーセキュリティ、内部脅威に対する安全対策を含む、強固なセキュリティ管理に投資し、実施する
 - ⑦ 技術的に可能な場合は、電子透かしやその他の技術等、ユーザーが AI が生成したコンテンツを識別できるようにするための、信頼できるコンテンツ認証及び来歴のメカニズムを開発し、導入する
 - ⑧ 社会的、安全、セキュリティ上のリスクを軽減するための研究を優先し、効果的な軽減策への投資を優先する

¹ https://www.soumu.go.jp/main_content/000912746.pdf

- ⑨ 世界の最大の課題、特に気候危機、世界保健、教育等(ただしこれらに限定されない)に対処するため、高度な AI システムの開発を優先する
- ⑩ 国際的な技術規格の開発を推進し、適切な場合にはその採用を推進する
- ⑪ 適切なデータインプット対策を実施し、個人データ及び知的財産を保護する

3. また、AI 関係者は第 12 の指針に従うべきである

- ⑫ 高度な AI システムの信頼でき責任ある利用を促進し、貢献する。

AI 関係者は、高度な AI システムが特定のリスク(例:偽情報の拡散に関するもの)をどのように増大させるか及び／又は新たなリスクをどのように生み出すかといった課題を含め、自分自身そして必要に応じて他者のデジタル・リテラシー、訓練及び認識を向上させる機会を求めべきである。

全ての関連する AI 関係者は、高度な AI システムの新たなリスクや脆弱性を特定し、それに対処するために、必要に応じて、協力し情報を共有することが奨励される。