

2023.5.26

AI 戦略会議用資料

今後の議論の方向性案

背景：主要な国内外の AI ガバナンスの議論（下線はハードロー）

	国内	国際
2016		G7 香川・高松情報通信大臣会合：AI 研究開発原則ガイドライン案提案
2017	総務省：AI 開発ガイドライン	
2018		
2019	総務省：AI 利活用ガイドライン 内閣府：人間中心の AI 社会原則	OECD：AI に関する OECD 原則 G20 茨城つくば貿易・デジタル経済大臣会合： G20AI 原則の合意
2020		
2021	経産省：AI 原則実践のためのガバナンス・ ガイドライン	UNESCO：AI の倫理勧告 <u>欧州委員会：AI 規制案の発表</u>
2022		<u>欧州評議会：AI 条約のゼロドラフト公開</u>

- 2023 年現在、既存の AI ガバナンスの論点に、生成 AI や LLM がどのような影響をもたらすかが議論されている。既存の AI 原則で生成 AI のみに特有の議論はあるのか？
- 国際的な議論に日本が貢献するためには、既存の議論や整理を踏まえたうえで整理、提案をする必要がある

論点整理のためのいくつかの視点

1. 定義して概念を共有する

- AI、生成 AI、大規模言語モデルなどの技術的な関係性と可能性と課題の整理
- 生成 AI で生成できるコンテンツごとの可能性と課題の整理
- ユーザーインタフェースデザインの可能性と課題の整理

2. 生成 AI や LLM のみに特有な課題があるかを整理する

- 生成 AI や LLM に関する具体的な国内外の利用例を収集し共有する
- 生成 AI や LLM に関する各組織のガイドラインや取り組みを収集し共有する
- 既存の「AI 原則」と「原則を実践するための方法論」でカバーしきれない論点、あるいは強調される論点が生成 AI にあるのかを検討する
- 例) 認識 AI や予測 AI では公平性、プライバシー、人間による監督、透明性や説明可能性、アカウントビリティなどの原則が重視されたが、生成 AI や対話 AI では著作権や偽情報や誤情報などのほか、個人の尊厳、感情操作、機械への依存など人と機械の関係性（インターフェース）に関する議論に重点がある

3. 各関係者がそれぞれ取るべき対策を整理する

- 様々な角度からの関係者を整理
 - 役割（開発者・サービス提供者・利用者等）、領域や業界、職務等
- 各関係者が参照できる原則や対策の整備や公開
 - 特にリスク評価や影響評価に関して自己評価ができるようなガイドの提供
- 他の国や地域におけるガバナンス動向の調査と協調の道筋を検討
 - 事例を収集したり、対策を提供したりする組織として GPAI や OECD などとの連携強化、国際的な議論の場への貢献
- インシデントやクライシス時の対応や救済措置をガバナンスの仕組みに組み込む
- 責任ある AI 開発、サービス提供、利用者のためのリテラシー向上