

AI に関する国際的なルールメイキングと生成 AI のインパクト

(1) 欧米における議論の動向

一般財団法人日本情報経済社会推進協会 上河辺 康子

はじめに

2023年5月、日本が議長国を務めたG7広島サミットにおいて、G7首脳は、生成AIに関する議論のための広島AIプロセスを年内に創設¹することとし、現在検討が進められつつある。本稿では、2回に分けて、生成AIを含むAIに関する規制の在り方に関する議論を紹介する。1回目である今回は、欧州連合、米国及び英国の取組みを取り上げる。

1. AI 及び生成 AI に関する議論の経緯

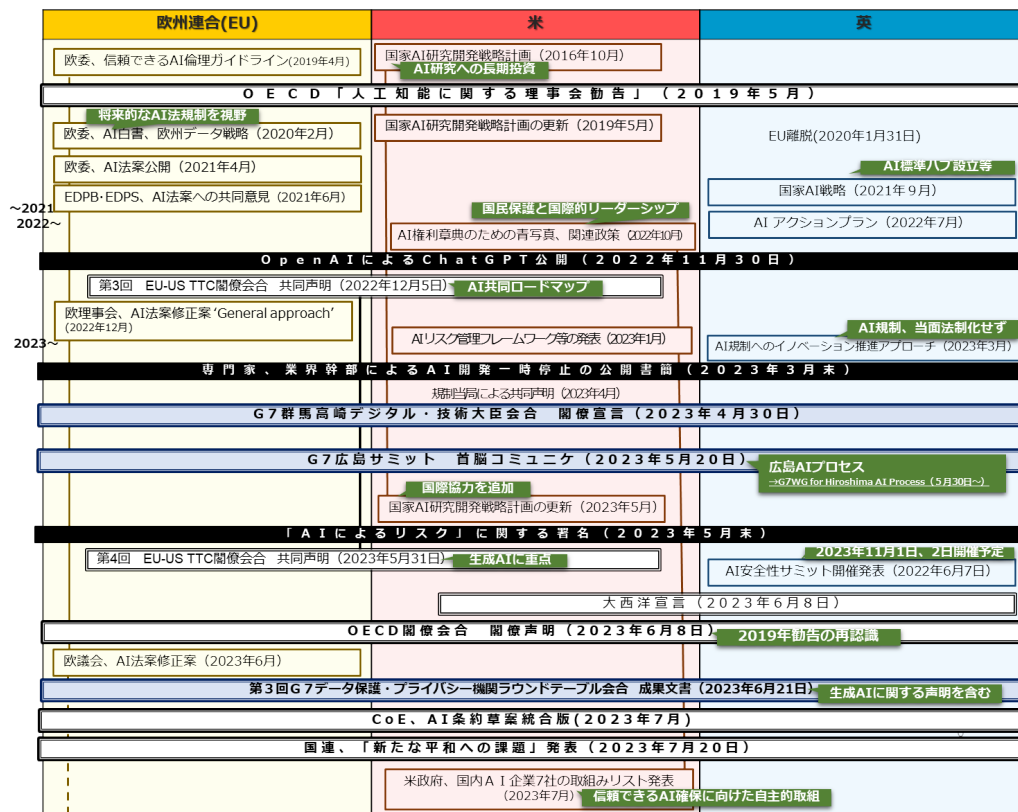


図1 AI 及び生成 AI に関する議論の経緯

※図1は、2023年7月までの状況をもとに作成

AIに関する国際的なルールとしては、2019年5月にOECD閣僚会合で採択された人工

¹ G7広島首脳コミュニケ (2023年5月20日)、38項

知能に関する理事会勧告（OECD/LEGAL/0449）がある。OECD は、2016 年より AI に関する政策的議論の一助となる活動に取り組み、2019 年の勧告では、信頼できる AI の責任あるスチュワードシップのための原則と、これらの原則と整合して国内政策及び国際協力の場で実行されるべき勧告が示された。このような国際的な議論の背景には、AI 技術の進展と各国の対応があるが、2022 年には生成 AI 利用が急激に拡大し、AI の専門家や経営者らによる署名活動が度々行われるなど、生成 AI の開発や利用の在り方が国際的な課題となり、各国の政策や二国間、多国間の議論に影響を及ぼしつつある。

2. 欧州連合

欧州連合（以下、EU）の執行機関である欧州委員会は、AI 戦略に基づき、倫理的で安全な最先端の AI 開発を支援するため、2018 年 6 月に High-Level Expert Group on AI（AI HLEG）を設置、AI HLEG は信頼できる AI のための倫理ガイドラインを策定した。このガイドラインは、信頼できる AI は適法（lawful）、倫理的（ethical）かつ堅牢（robust）である必要があるとし、信頼できる AI のための倫理原則、倫理原則を満たすための要件及び AI 評価リストを示した。欧州委員会は、AI HLEG の成果をふまえ、2020 年 2 月に、将来的な法による AI 規制を視野に入れた AI 白書を欧州データ戦略とともに公表、さらに、2021 年 4 月に AI 法案を提案した。AI 法案では、リスクベースアプローチを採用し、許容できないリスクを持つ AI システムは原則禁止、ハイリスクな AI には要件が課せられる。禁止される AI システムとして、個人の意思決定を歪ませる AI（サブリミナル効果、年齢や障害など人々の脆弱性の利用）、公的機関によるソーシャルスコアリング、法執行を目的とした公にアクセスできる場所でリアルタイム遠隔生体識別システムの利用が挙げられた。この AI 法案に対し、2021 年 6 月に、欧州データ保護委員会（EDPB）および欧州データ保護監察機関（EDPS）が、ソーシャルスコアリングや公にアクセス可能な空間における個人の遠隔生体識別のため AI 使用を全面的に禁止すべき等の意見を共同で提出した。また、欧州理事会は 2022 年 12 月²、欧州議会は 2023 年 6 月に、それぞれ異なる修正案を採択した。欧州議会の修正案は、欧州委員会提案から 2 年以上を経て提出され、その間の生成 AI の急激な利用拡大を背景に、ChatGPT のような生成 AI システムはコンテンツが AI によって生成されたことを開示しなければならない等、生成 AI に関する具体的な規制が含んでいる。今後は、早期の法の公布と施行を目指し、トリオログを経て AI 法の合意案の調整を行う方向とされている。

EU は、AI に関する研究開発において米国と協力関係にある。EU 米間の新興技術や貿易の課題に取り組む枠組みである EU-US Trade and Technology Council（TTC）の閣僚会合における議論を通じて、2022 年 12 月に信頼できる AI の開発運用に向けた共同ロードマッ

² 欧州理事会は、2022 年 12 月 6 日に AI 法に関する共通の見解‘General approach’を採択した。

プを公開した。EU 米の信頼できる AI のためのリスクベースアプローチの適用例として、EU の AI 法、米国の AI リスク管理フレームワークや AI 権利章典の青写真（後述）を挙げ、双方の規制アプローチの見解は異なるとしつつも、共通の価値観が新興技術を導くとしている。具体的な活動として、①用語と分類法の共有、②国際的な技術標準の開発活動と分析・ツール収集、③既存及び新たな AI リスクの監視と測定を進めることを挙げ、さらに、2023 年 5 月、G 7 広島 AI プロセスを補完する作業として、共同ロードマップ検討において生成 AI を重視することが確認された。

3. 米国

米国は、2016 年 10 月に策定した国家 AI 研究開発戦略計画を、2019 年 6 月、2023 年 5 月に更新した。同計画は、連邦政府の研究開発投資を調整し集中させるために、AI における主要な研究課題を定義するものであり、2016 年の計画当初より、AI 研究への長期投資を行うことを戦略の第一に掲げてきた。2023 年 5 月の更新では、生成 AI を含め、基礎的かつ責任ある AI 研究に引き続き長期投資を行うとともに、新たに、AI 研究における国際協力への原則的かつ協調的アプローチを確立することが追加された。

米国は、前述の EU と後述する英国との協力を含め、AI 分野の国際協力を進めながら、国内の世界的 AI 企業との協力関係も構築し、自主的な規制への取組みを促そうとしている。

米国における AI に関するルールメイキングへの取組みとして、2022 年 10 月に科学技術政策局（OSTP）が公開した AI 権利章典のための青写真がある。AI 等の自動化システムにより人々の監視やランク付けなどの問題が過去 10 年で劇的に拡大し人々の権利を脅かしているとして、安全で効果的なシステムやアルゴリズムによる差別からの保護など 5 つの保護策を示した。AI 権利章典のための青写真公表と同日に、米政府は、技術分野の説明責任を進め国民の権利を守るための主要な行動を発表した³。この中には、労働者等の保護、教育者や医療提供者の支援に加えて、アメリカが AI 分野のリーダーシップを維持することや、技術者の指導・支援も含まれた。技術者の指導・支援策の一つとして、2023 年 1 月、国立標準技術研究所（NIST）が AI リスク管理フレームワーク（AI RMF 1.0）を公開した。フレームワークの目的は、信頼できる責任ある AI システムの開発と使用の促進とされ、フレームワークは自主的なものとされている。

2023 年 7 月 21 日⁴、米政府は、安全、安心、信頼できる AI の確保に向けて、国内の AI

³ FACT SHEET: Biden-Harris Administration Announces Key Actions to Advance Tech Accountability and Protect the Rights of the American Public、米ホワイトハウス発表、2022 年 10 月 4 日

⁴ FACT SHEET: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI、米ホワイトハウス発表、2023 年 7 月 21 日

企業 7 社による自主的な取組みについて 8 項目のリストを発表した。このリストには、AI で生成されたコンテンツであるかどうかをユーザーが理解できるようにするメカニズムの開発・展開、有害な偏見や差別の回避やプライバシーの保護などの AI システムによってもたらされる社会的リスクに関する研究の優先などが含まれている。同時に、米政府は、同盟国やパートナーと AI の開発と利用を管理する国際的枠組みを確立すると発表した。発表によれば、米国は既に、日本、英国、仏独伊等のヨーロッパ諸国、アジア太平洋諸国など 20 か国と協議し、国連やその加盟国とも AI について議論している。

また、米政府は、米国内の AI 企業の自主的な取組みの確保について、7 月の 7 社に続き、2023 年 9 月 12 日に、新たに 8 社と、製品の安全性の確保等に取り組むことを確認⁵した。

米国では、民間の活動も活発に行われている。AI 利用に関して、脚本家や俳優がストライキを実施し、新聞大手が AI 学習のための記事利用禁止を利用規約に含めた⁶。一方、米民間 AI 企業 4 社は、2023 年 7 月 26 日、AI 利用の安全性のための業界団体 Frontier Model Forum の設立を発表した⁷。フォーラムは、今後 1 年間、ベストプラクティスの特定、AI の安全性の研究促進、企業と政府間の情報共有促進の 3 点に焦点をあてて活動するとしている。

4. 英国

英国は、2020 年に EU を離脱した。英国の国家 AI 戦略（2021 年 9 月策定、2022 年 12 月更新）は、「英国は AI における世界的な大国」であり「今後 10 年間で世界をリードする立場にある」として、政策上の 3 つの柱と、柱毎の短期（今後 3 か月）、中期（今後 6 か月）、長期（今後 12 か月以上）の計画を示し、その進捗状況を監視・評価するためのメカニズムを導入するとした。2022 年 7 月に公開された AI アクションプランは、国家 AI 戦略の進捗状況を概観し、毎年更新するとされており、戦略の進捗状況を監視・評価するためのメカニズムの一環と考えられる。国家 AI 戦略は、中期計画のひとつに AI 標準ハブの設立を挙げ、世界的な AI の技術標準策定に貢献するとしている。AI アクションプランによれば、AI 標準ハブについては、英アラン・チューリング研究所が、英国規格協会（BSI）と国立物理学研究所（NPL）が支援を得て、試験運用を主導することとなり、その後、2022 年 10 月より運用が開始された⁸。

⁵ FACT SHEET: Biden-Harris Administration Secures Voluntary Commitments from Eight Additional Artificial Intelligence Companies to Manage the Risks Posed by AI、米ホワイトハウス発表、2023 年 9 月 12 日

⁶ 利用規約 4 (3)、<https://help.nytimes.com/hc/en-us/articles/115014893428-Terms-of-Service>

⁷ <https://openai.com/blog/frontier-model-forum>

⁸ Standards Hub、<https://aistandardshub.org/>

2023年3月、英国科学技術省は、AI規制へのイノベーション推進アプローチを示し、この中でAI標準ハブは、信頼できるAIのためのツールと位置付けられている。また、AI規制について、指針となる安全性・セキュリティ・堅牢性、適切な透明性と説明可能性等の原則を挙げつつも、企業に対する厳しい法的抑制はイノベーションを抑制するとして、AI規制の原則を、当面は法制化しないとしている。

英政府は、2023年6月7日、今秋英国においてAIの安全性に関する初のグローバルサミット（以下、AIセーフティサミット）を開催予定であると発表した。サミットにおける検討は、G7、OECD、AIに関するグローバル・パートナーシップでの最近の議論に基づくものであると説明している。翌6月8日の米英首脳会談では、英国によるAIセーフティサミット開催に米国が歓迎の意を示すとともに、両国がOECD、国連、AIグローバル・パートナーシップ、欧州評議会、国際標準化機構、G7広島AIプロセスを含む国際的な活動を歓迎しAIもたらすリスクと機会に関してさらに前進すること、CBPRフォーラムやOECDへの支援を含めたデータセキュリティリスクに関する共通の理解を構築することなどが確認され、成果文書である大西洋宣言に盛り込まれた。

2023年8月24日、英国政府は、AIセーフティサミットの開催日程を2023年11月1日、2日と発表⁹、さらに、2023年9月4日、AIセーフティサミットでは、AIシステムにより生み出されるリスクや安全なAIによる公共の利益に焦点をあてる方針を示した¹⁰。

おわりに

欧州連合、米国、英国は、「信頼できるAI」のための政策を打ち出しているが、欧州連合は法制度、米国は民間による自主規制、英国は標準化に重点を置き、信頼できるAI実現に向けた規制アプローチはそれぞれ異なる。今回は、G7等の国際的な枠組みにおける議論を紹介したい。

参考文献

- [1] OECD、人工知能に関する理事会勧告、OECD/LEGAL/0449、2019年5月22日
- [2] 欧州委員会、Artificial Intelligence for Europe、2018年4月25日
- [3] 欧州委員会 AI HLEG、Ethics Guidelines for Trustworthy AI、2021年4月8日
- [4] 欧州委員会、WHITE PAPER On Artificial Intelligence -A European approach to excellence and trust、2021年2月19日
- [5] 欧州委員会、Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)

⁹ Iconic Bletchley Park to host UK AI Safety Summit in early November、英国政府発表、2023年8月24日

¹⁰ UK government sets out AI Safety Summit ambitions、英国政府発表、2023年9月4日

- and amending certain Union legislative acts、2021年4月21日
- [6] 欧州データ保護委員会及び欧州データ保護監察機関、EDPB-EDPS Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) 、2021年6月18日
 - [7] 欧州理事会、Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts- General approach、2022年11月25日
 - [8] 欧州議会、Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down Artificial Intelligence Act and amending certain Union legislative acts、2023年6月14日
 - [9] TTC Joint Roadmap on Evaluation and Measurement Tools for Trustworthy AI and Risk Management、2022年12月1日
 - [10] OSTP、National Artificial Intelligence Research and Development Strategic Plan、2016年10月
 - [11] OSTP、National Artificial Intelligence Research and Development Strategic Plan : 2019 Update、2019年6月
 - [12] OSTP、National Artificial Intelligence Research and Development Strategic Plan : 2023 Update、2023年5月
 - [13] OSTP、Blueprint for an AI Bill of Rights、2022年10月4日
 - [14] NIST、Artificial Intelligence Risk Management Framework (AI RMF 1.0)、2023年1月26日
 - [15] Department for Science, Innovation and Technology 他、National AI Strategy、2021年9月21日 (2022年12月18日更新)
 - [16] Department for Science, Innovation and Technology 他、National AI Strategy- AI Action Plan、2022年12月18日
 - [17] Department for Science, Innovation and Technology 他、A pro-innovation approach to AI regulation、2023年3月29日
 - [18] The Atlantic Declaration、2023年6月8日



JIPDEC 電子情報利活用研究部 上河辺 康子

- ヘルスケアサービスにおけるデータ利活用に関する調査
- 企業情報化動向調査
- JIS Q 15001 原案作成団体事務局 など

■協会外の主な活動

- プライバシーマーク主任審査員
- プライバシーマーク審査基準委員（保健医療福祉分野）
- ヘルスケア・PHR 関連団体参加