個人的見解を含みます

AI政策の動向

2024.12.9

渡邊昇治

目次

AI政策 全体像

内閣府作成資料から編纂。

これまでの基本戦略・理念 「AI戦略2022」「人間中心のAI社会原則」

生成AIなどの技術の変化

自然な対話が可能、精巧な画像生成が容易など

！大きな便益・イノベーション

！一方で、AIに関するリスクはより切迫したものに

国際的な議論

広島AIプロセス国際指針等に2023年末にG7合意

アウトリーチ（現在54ヶ国が賛同）、モニタリング等の議論

国連、欧州評議会（条約）、EU法、米大統領令等の動き

AI戦略会議、AI制度研究会（有識者）＋ AI戦略チーム（関係省庁）

「AIに関する暫定的な論点整理」（2023年5月26日　AI戦略会議とりまとめ）

「統合イノベーション戦略2024」（2023年6月4日 閣議決定）

国際的な議論とリスクへの対応

- 中央省庁における扱い、個人情報保護委員会からの注意喚起、教育現場のガイドライン、AI事業者ガイドライン
- 知財・著作権とAIに関する検討、デジタル空間の健全性確保の在り方に関する検討、AI制度の在り方に関する検討
- AIセーフティインスティテュートの創設
- 偽情報対策の技術開発など

AIの最適な利用

- 官民における利用の促進
- 幅広い世代のスキル・リテラシー教育

AI開発力の強化

- 計算資源の確保、データ整備、モデル開発、基礎研究
- トップ人材が集まる環境整備
- スタートアップ施策の推進

国際的な政策動向

広島AIプロセス

G7広島サミット　広島AIプロセスを提唱（2023.5）

閣僚級会議を経て、包括的政策枠組みを承認（2023.12.6）

米国

大手AI開発者によるボランタリーコミットメント（2023.7）

大統領令（2023.10.30）

大規模汎用のAIについては国防生産法による報告義務も。

日米

日米首脳会談、共同声明（2024.4.10）

AIセーフティに関する協力推進、政府発表文書の改ざん防止等。

EU

AI法が成立（2024.5.21）

英国

AI安全性サミットを英国で開催（2023.11.2）

そのフォローアップ会議としてAIソウルサミットを韓国で開催（英国・韓国共催）（2024.5.21）

IGF（Internet Governance Forum）

IGF京都2023においてAI特別セッション（2023.10.9）

GPAI（Global Partnership on AI）

2022.11～1年間日本が議長国

GPAIサミットをインドで開催し、閣僚宣言（2023.12）

GPAI東京センターを立ち上げ

OECD

OECD閣僚理事会 生成AIに関するサイドイベント（2024.5.2）

広島AIプロセスのアウトリーチ。広島AIプロセス・フレンズ・グループ創設。

欧州評議会

AI条約案の交渉が妥結（2024.3.14）

国連

AIに関する国連総会決議（米国提案）
国連総会で採択（2024.3.21）

AIに関する高級諮問機関 報告
「人類のためのAI統治」（2024.9.19）

広島AIプロセス

G7広島サミット（2023年5月）で提唱され、閣僚級の議論を開始。
2023年12月にG7首脳声明。

(1) 広島AIプロセス包括的政策枠組み

- 生成AIに関する G7の共通認識のためのOECDレポート
- 広島AIプロセス **全てのAI関係者向けの国際的な行動指針**
- 広島AIプロセス **高度なAIシステム開発者向けの国際的な行動規範**
- 偽情報対策技術などに関するプロジェクトベースの取組み

(2) 広島AIプロセスを前進させる行動計画

- パートナー国へのアウトリーチ
- モニタリングツールとメカニズムの導入提案
- 広島AIプロセスのためのウェブサイト
- マルチステークホルダーとの対話
- 議長国イタリアの下での協働継続

「全てのAI関係者向けの広島プロセス国際指針」の概要

「AI開発者向けの広島プロセス国際指針」の11の項目が、高度なAIシステムの設計、開発、導入、提供及び利用に関わる全ての関係者に適宜適用し得るものとして整理した上で、利用者に関わる内容が12番目の項目として追加。

全てのAI関係者向けの広島プロセス国際指針の12項目

1．高度なAIシステムの市場投入前及び、高度なAIシステムの開発を通じて、AIライフサイクルにわたる**リスクを特定、評価、低減するための適切な対策**を実施する。

2．市場投入後に**脆弱性、インシデント、悪用パターンを特定し、低減**する。

3．十分な透明性の確保や説明責任の向上のため、高度なAIシステムの**能力、限界、適切・不適切な利用領域を公表**する。

4．産業界、政府、市民社会、学術界を含む関係組織間で、**責任ある情報共有とインシデント報告**に努める。

5．リスクベースのアプローチに基づいた**AIのガバナンスとリスク管理ポリシーを開発、実践、開示**する。特に高度AIシステムの開発者向けの、プライバシーポリシーやリスクの低減手法を含む。

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及び内部脅威対策を含む強固なセキュリティ管理措置に投資し、実施**する。

7．AIが生成したコンテンツを利用者が識別できるように、**電子透かしやその他の技術等、信頼性の高いコンテンツ認証および証明メカニズムを開発**する。またその**導入が奨励**される。

8．社会、安全、セキュリティ上の**リスクの低減のための研究を優先し、効果的な低減手法に優先的に投資**する。

9．**気候危機、健康・教育などの、世界最大の課題**に対処するため、**高度なAIシステムの開発を優先**する。

10．**国際的な技術標準の開発と採用を推進**する

11．適切な**データ入力措置と個人情報及び知的財産の保護**を実施する。

12．**偽情報の拡散等のAI固有リスクに関するデジタルリテラシーの向上**や**脆弱性の検知への協力と情報共有**等、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。

米国

- バイデン-ハリス政権は、AIがもたらす大きな可能性を捉え、リスクを管理し、米国人の権利と安全を守るためのコミットメントの一環として、アマゾン、グーグル、Meta、マイクロソフト、OpenAIなど大手AI開発者からボランタリー・コミットメントを確保（2023年7月～）。
- 米国は大統領令を発出（2023年10月）。既存の法令・予算を活用し、イノベーションを促進し、リスクに対応することを各省庁に対して指示。

ボランタリー・コミットメントのポイント

安全性	AI製品のリリース前にセキュリティ・テスト 業界、政府、社会、学界との情報共有
セキュリティ	モデルウエイトの保護 脆弱性対策
信頼性	AI生成物であることをユーザーに伝える技術 AIシステムの能力、限界、使用領域等の公表 AIシステムの社会的リスクに関する研究 社会課題に対処する高度なAIシステムの開発

大統領令のポイント

- AI技術の安全性とセキュリティの確保（ガイドライン・標準・ベストプラクティス、安全で信頼できるAI（デュアルユース基盤モデルの開発企業に報告義務）、重要インフラ・サイバーセキュリティ、化学・生物・放射性物質・核兵器脅威の軽減、合成コンテンツのリスクの軽減、デュアルユース基盤モデルに関する意見募集、AI学習のための連邦データの安全な提供等）
- イノベーションと競争の促進（AI人材誘致、イノベーション・競争促進）
- 労働者の支援
- 公平性と公民権の推進（刑事司法制度におけるAI利用と公民権の保護等）
- 消費者、患者、乗客、学生の保護
- プライバシーの保護
- 連邦政府によるAI利用促進（AI管理指針、政府におけるAI人材確保）
- 海外における米国のリーダーシップの強化
- 実施（ホワイトハウスAI評議会の設置等）

EU

EUのAI法は、広くAI全体をカバー。人権、差別・偏見問題への対応を重視。

禁止されるAI

- センシティブな特性（政治的、宗教的、哲学的信条、性的指向、人種など）を使用するバイオメトリクス分類システム
- 顔認識データベースを作成するために、インターネット等の映像から顔画像をかき集める
- 職場や教育機関における感情認識
- 社会的行動や個人的特徴に基づく社会的スコアリング
- 人間の行動を操作して自由意志を奪うAIシステム
- 人々の脆弱性を悪用するために使用されるAI

法執行機関の免除

公共空間におけるバイオメトリクス識別システムは、事前の司法認可を条件とし、重大な犯罪を犯して有罪判決を受けた、あるいはその疑いのある人物を対象とした捜査にのみ使用。

高リスクなAIシステム

- 高リスク（健康、安全、基本的権利、環境、民主主義、法の支配、有権者の行動に重大な害を及ぼす可能性がある）と分類されたAIシステムについては、市場投入前に影響評価、適合性評価（リスクマネージメントシステム、データガバナンス、技術文書、記録保持、透明性と利用者への情報提供、サイバーセキュリティ等）を義務化。
- 市民は苦情を申し立てる権利、説明を受ける権利を持つ。

限定的なリスクのAIシステム

透明性義務。例えば、コンテンツがAIによって生成されたものであることの開示。

汎用AIシステムの扱い

汎用AI（GPAI、General Purpose AI）システムとそのベースとなるGPAIモデルは、透明性要件を遵守（技術文書の作成、EU著作権法の遵守、学習に使用したコンテンツに関する開示など）。

システミック・リスクを伴う影響力の大きいGPAIモデルについては、より厳しい義務（モデル評価の実施、システミックリスクの評価と軽減、敵対的テストの実施、重大インシデントに関する欧州委員会への報告、サイバーセキュリティの確保、エネルギー効率の報告など）。

イノベーションと中小企業支援策

革新的なAIを開発・訓練するため、市場投入前にいわゆる規制のサンドボックスと実証試験を促進。

罰則

違反企業の前会計年度における全世界の年間売上高に対する割合、または、あらかじめ決められた金額のいずれか高い方。
禁止AIの違反に対しては3500万ユーロ（7%）、不正確な情報の提供に対しては750万ユーロ（1.5%）など。

発効

AI法は発効から2年後に適用されることになっているが、例外の条項もある。

国内の政策動向

AIに関する暫定的な論点整理（AI戦略会議、2023.5.26）

【個人情報保護】 リスクへの対応

OpenAIに対する注意喚起（個人情報保護委員会、2023.6.1）
生成AIサービスの利用に関する注意喚起等（個人情報保護委員会、2023.6.2）

【AIと知財・著作権との関係】

文化審議会著作権分科会「AIと著作権に関する考え方について」（文化庁、2024.3.19）
著作権侵害になり得る場合について明確化、周知・啓発。関係者の相互理解の促進。

AI時代の知的財産検討会 中間とりまとめ（知的財産戦略推進事務局、2024.5）
法、技術、契約による取組み。

【偽・誤情報等】

デジタル空間における情報流通の健全性確保の在り方に関する検討会（総務省、2023.11～）

【雇用への影響】

雇用政策研究会中間整理「新たなテクノロジーが雇用に与える影響について」（厚労省、2023.12.21）
生産性向上の一方で仕事内容は変化。労使コミュニケーション深化、キャリア形成支援などが重要。

【ガイドライン・履行確保等】

AI事業者ガイドライン ver1.0（総務省・経産省、AI戦略会議了承、2024.4.19）
（参考）不正競争防止法「秘密情報の保護ハンドブック」（2024.2）生成AIに関して記載

AIセーフティ・インスティテュート創設（内閣府等、2024.2.14）
海外の同種の機関とも連携し、安全性確保に向けた調査研究、基準作成等に取り組む。

AI制度研究会（内閣府、2024.8～）
法制度の要否も含め、制度の在り方を検討。

【政府における利用】 利用促進

ChatGPT等の生成AIの業務利用に関する申合せ（デジタル社会推進会議幹事会）
第一版（2023.5.8）で機密性1情報から試験的に利用。第二版（2023.9.15）で機密性2情報についても試験的に利用。

【教育分野での利用】

初等中等教育段階における生成AIの利用に関する暫定的なガイドライン ver.1.0（文科省、2023.7.4）

【人材育成】

デジタルスキル標準の普及、ITSS（ITスキル標準）レベル３相当の教育訓練を認定制度の対象へ

【公的機関によるデータの整備】 開発力の強化

NICTによる日本語データの整備

政府データのAI学習への提供アクションプラン ver.1.0（内閣府、2023.11.7）
政府データに関するニーズ調査、データ形式変換方法等を検討。

【計算資源の整備】

AI用計算資源について、産総研ABCIの拡充や民間における整備を支援。

【モデル開発支援】

スタートアップ等によるAIモデルの開発を促進させるべく支援を行っているところ。

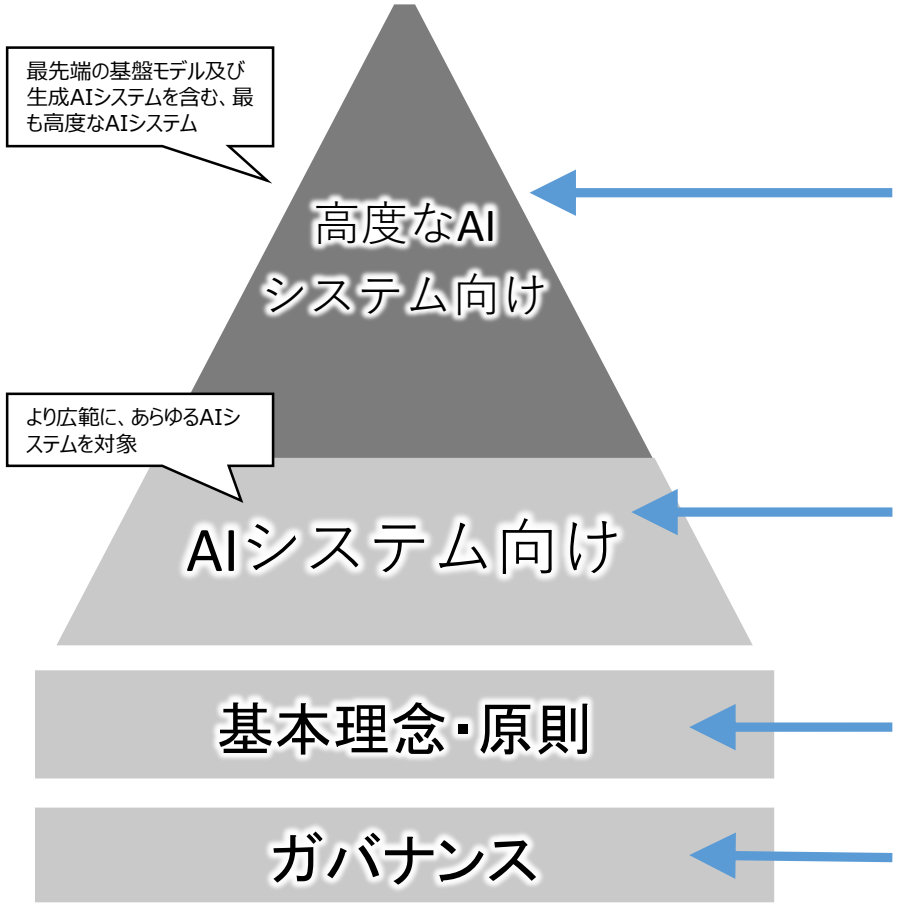
【基礎研究】

NIIにおいて、生成AIモデルの透明性・信頼性の確保等の研究開発を開始。
理研において、科学研究向けAI基盤モデルの開発に着手するとともに、 AI for Scienceの日米連携枠組みを創設（2024.4）。

日本のAI事業者ガイドライン（総務省・経産省）

広島AIプロセスの成果（高度なAIシステムに関する国際指針、国際行動規範）を反映しつつ、一般的なAIを含む全てのAIシステム・サービスを広範に対象。

AI開発・提供・利用においては、本ガイドラインを参照し、各事業者が適切なAIガバナンスを構築するなど、具体的な取組みを自主的に推進することが重要。



広島AIプロセスの成果を反映
- すべてのAI関係者向け及び高度なAIシステムを開発する組織向けの広島プロセス国際指針
- 高度なAIシステムを開発する組織向けの広島プロセス国際行動規範

[第2部　D.高度なAIシステムに関する事業者に共通の指針、第3部 AI開発者に関する事項]

原則を元に、各主体が取り組むべき指針や事項を整理
（AI開発ガイドライン、AI利活用ガイドラインも取り込み）
- 各主体が連携して、バリューチェン全体で取り組む事項：1)人間中心、2)安全性、3)公平性、4)プライバシー保護、5)セキュリティ確保、6)透明性、7)アカウンタビリティ
- 各主体が社会と連携して取り組むことが期待される事項：8)教育・リテラシー、9)公正競争確保、10)イノベーション

[第2部　C.共通の指針、第3-5部　AI開発者、AI提供者、AI利用者に関する事項]

「人間中心のAI社会原則」の基本理念を土台とし、OECDのAI原則等を踏まえ、原則を構成

[第2部　A.基本理念、B.原則]

AI原則実践のためのガバナンス・ガイドラインを元に整理

[第2部　E.ガバナンスの構築]

政府資料を編纂。

AIセーフティ・インスティテュート

AI開発者等による安全性確保をサポートするため、英国、米国、日本、シンガポール、韓国はAI Safety Institute（EUはAI Office）を創設。

日本は、IPA（独立行政法人 情報処理推進機構）にAIセーフティ・インスティテュートを設置し、関係省庁・関係機関と協力し、諸外国の機関ともネットワークを構築。

オーストラリア、フランス、カナダにおいても同様の機関を創設予定。

日本のAISIの業務内容

- 安全性評価に係る調査、基準等の作成
- 安全性評価の実施手法に関する検討
- 国際連携に関する業務

AI Safety Institute

村上明子所長

損害保険ジャパン株式会社
執行役員CDO・DX推進部長

関係省庁
内閣府（科学技術・イノベーション推進事務局）、国家安全保障局、内閣サイバーセキュリティセンター、警察庁、デジタル庁、総務省、外務省、文科省、経産省、防衛省

関係機関
情報通信研究機構、理化学研究所、産業技術総合研究所

※関係省庁、関係機関は現時点での予定。

国民の感覚

個人メモ

多くの国民がAIは便利だと思っているが、懸念もあると思っている。

	Global	Japan
AIには利点がある	85%	75%
AIを信用したい	39%	23%
AIには規制が必要	71%	77%

Survey of 17,000 people in 17 countries

Global Survey on Changing Public Perceptions of AI 2023
KPMG, University of Queensland

内閣府資料等を基に作成。

AI制度の在り方について

AI戦略会議の下に設置されたAI制度研究会において、法制度の要否も含め検討中。

基本的な考え方

- **リスクへの対応とイノベーション促進の両立。**
- **急速な技術・ビジネスの変化と多様性に対応。**
- **国際相互運用性の確保。広島AIプロセス国際指針等を遵守。**

刑法、個人情報保護法、著作権法等の法令、ガイドライン、技術の活用。リスクに応じた対応。

機密情報を扱う政府

適正な調達と利用。

重要なインフラ・製品安全等

経済安全保障法、各業法、製品安全に関する法令等による対応。

上記以外の領域

- 利用者のリテラシー向上。開発者・提供者による安全性・透明性確保。
- 実態把握、情報収集。
- 国際標準、第三者認証制度の検討・活用。

ISO/IEC、AISI等の活動

利活用促進・開発力強化

官民における利活用の促進

- 利用者のリテラシー向上、人材育成

計算資源の確保

- 官民のデータセンター拡充（学習用／推論用）
- 半導体、通信インフラの高度化

学習用データの整備

- 日本語データの整備
- 日本が競争力を有する分野のデータ整備

新たなモデル開発、基礎的な研究

- 新たなモデル（複数モデルの組合せ、RAG等）
- AI for X（AI導入の遅れによって得意分野の競争力を失わないように）
- AI for Science
- 高度AI人材の確保

AI関連予算の推移

- 令和７年度概算要求におけるAI関連予算の合計額は、約**1,780億円**※
- 令和６年度当初予算におけるAI関連予算の合計額は、約**1,176億円**※

※　内閣府SIP/BRIDGE、国立研究開発法人の運営費交付金等、AI関連予算額を抽出困難な施策分は含まず