AIガバナンスの確立に向けた政府の取組

総務省　MIC　Ministry of Internal Affairs and Communications

総務省　国際戦略局　国際戦略課
AI政策推進室

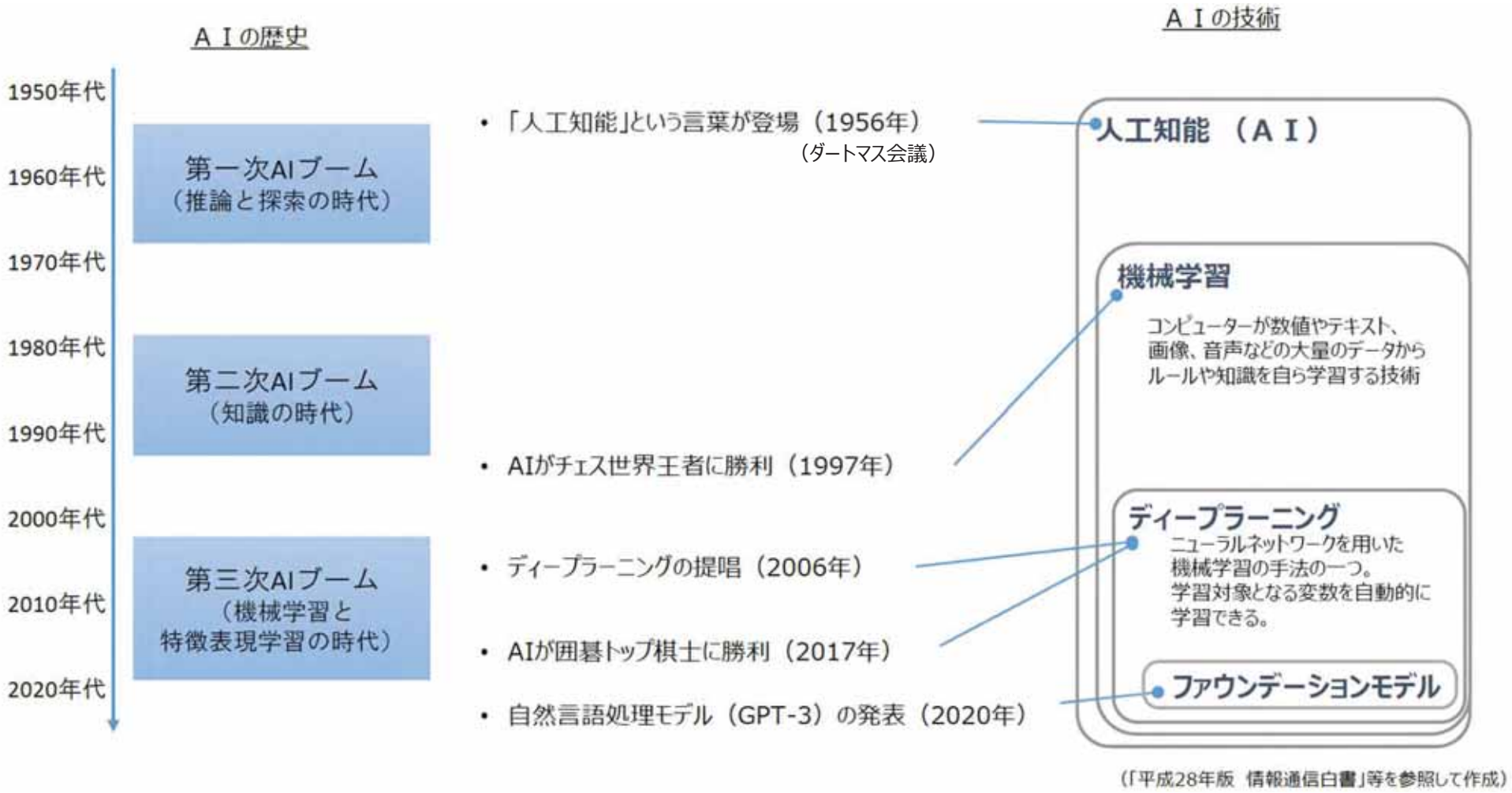
はじめに（AIガバナンス）

■AI ガバナンス
　AI の利活用によって生じるリスクをステークホルダーにとって受容可能な水準で管理しつつ、そこからもたらされる正のインパクト（便益）を最大化することを目的とする、ステークホルダーによる技術的、組織的、及び社会的システムの設計並びに運用。

※出典：「AI事業者ガイドライン（第1.1版）」本編
https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02ryutsu20_04000019.html

AIの技術的進展の歴史

- AI（人工知能）の研究は1950年代から開始。
- 2000年代には**ビッグデータ**を活用しAI自身が知識を獲得する**機械学習**が一般化。
- 2010年代以降は実用性の高い**ディープラーニング**が主流となる。

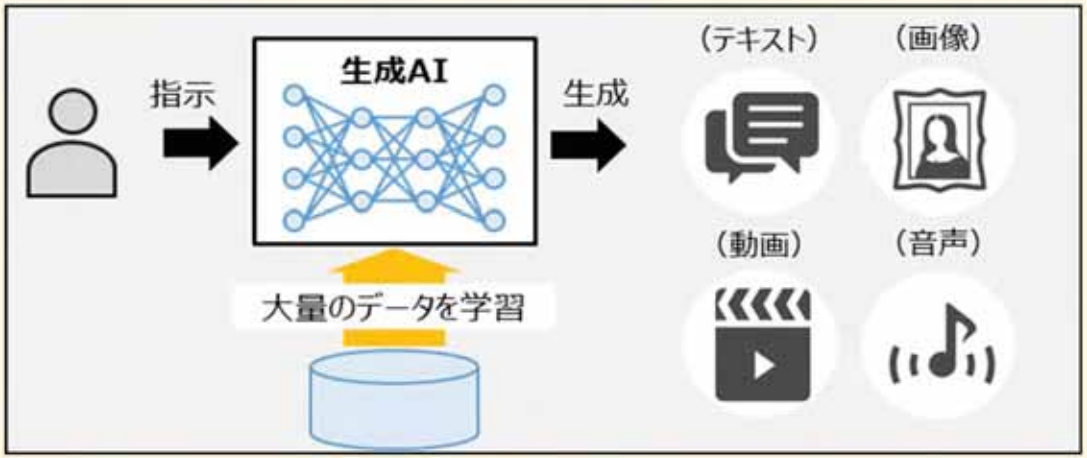


出典）自民党AIの進化と実装に関するプロジェクトチーム（令和５年２月３日）
内閣府資料「AI戦略と政府の取り組みについて」より抜粋、一部加工

生成AIのメリットとリスク

「生成AI」は、大量の学習データから、文章、画像、動画、音声、プログラムのコードなどを「生成」することが可能

（生成AIの概要）



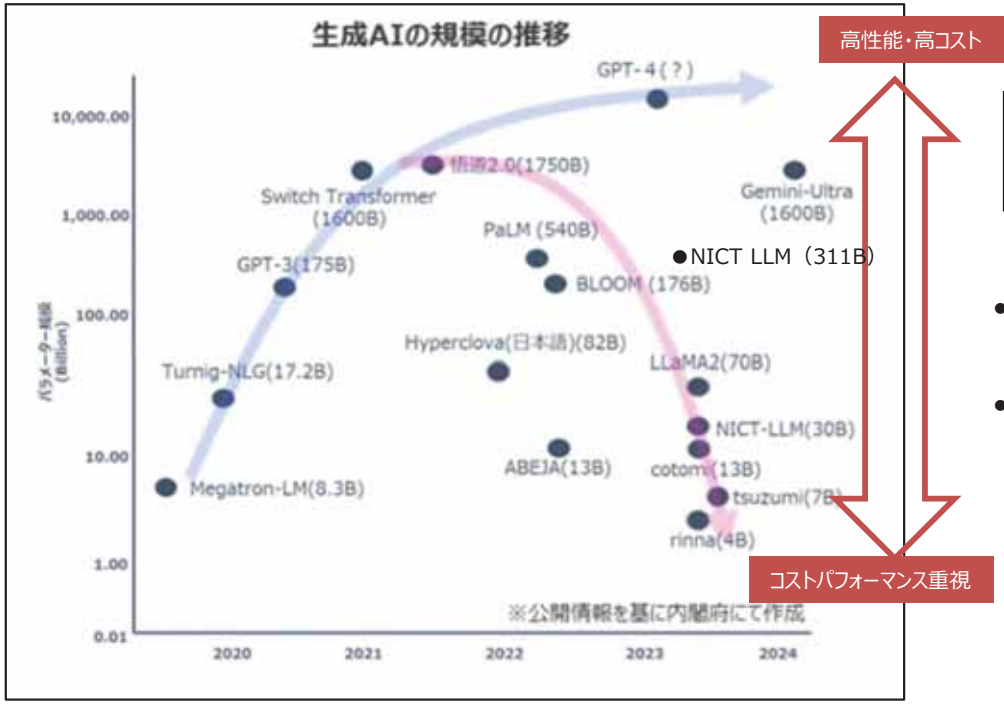
（アウトプットイメージ）

「生成AI」は「従来型AI」と比較して汎用性が高く創造的であるなどの特徴がある一方、様々なリスクも指摘

生成AIのメリット	生成AIのリスク
労働力不足の解消（問合せ対応、点検・監視など）	**犯罪増加リスク**
事務作業の効率化（文献調査、要約・翻訳、資料原案作成など）	**誤情報提示のリスク**
イノベーションの創出（新たな素材、新たな薬等の開発など）	**著作権侵害のリスク**
地球規模課題の解決（災害予測・対策、・パンデミック対策など）	**差別・偏見の増幅リスク**

大規模言語モデル（LLM）について

- Googleが2017年に開発した「**Transformer**」という深層学習アーキテクチャに基づき、**膨大なテキストデータから学習し、前後の文脈（単語間の関係性）をAIが“理解”**できるようになり、**AIと人間が自然な言語で“対話”可能に**。
- ChatGPTに代表される「テキスト生成AI」は、**従来型AIと比べてモデルサイズ（パラメータ数）が非常に大きい**ことから**大規模言語モデル（Large Language Model：LLM）**と呼ばれる。
- LLMでは**「計算量」「データ量」「モデルサイズ（パラメータ数）」を増やせばAIの性能向上が続く**という経験則（スケーリング則）が見出され、GAFAM等を中心に、大規模な投資に基づくLLM開発競争が続けられている。
- **LLMは高い汎用性を有するが、モデルサイズを増やして性能を上げようとすればするほど学習・運用に大きなコスト**を要するため、**学習データの工夫など小さなモデルで性能を上げる取組**も進展。



※内閣府資料に総務省加筆

LLMの高性能化 ➡ 「**データ量**」×「**パラメータ数**」 ➡ 計算量が増加

- LLMは、多数のプレイヤーが開発に参画し、この1年で**急激に多様化**
- GPT-5などの、**高性能、高コストのハイスペックモデル**（データセンター内の多数の計算機で動作）から、**コストパフォーマンス重視の小さめのモデル**（ノートPC・スマホ単体で動くレベル　※エッジAI）まで幅広く開発が進んでいる

主な生成AIサービスの例

✓ テキスト、画像、音楽、音声、動画などの分野で生成AIサービスが提供。また、OpenAIの「GPT-5」「OpenAI o4-mini」やGoogleの「Gemini」など、テキストに加え、画像や音声など、複数の情報形式に対応した**マルチモーダルモデル**も登場。

生成AI	テキスト	ChatGPT（米OpenAI）、Gemini（米Google）、Llama（米Meta）、Claude（米Anthropic）
	画像	DALL-E（米OpenAI）、Stable Diffusion（英Stability AI）、EMU（米Meta）
	音楽	MusicLM（米Google）、Amper Music（米Amper Music）
	音声	VALL-E（米Microsoft）、Amazon Polly（米Amazon）
	動画	Make-A-Video（米Meta）、Gen（米Runway）、Phenaki、Dream Machine（米Luma AI）、Sora（米OpenAI）
	※ 代表的なマルチモーダル	GPT-5、OpenAI o4-mini（米OpenAI）、 Gemini（米Google）、Claude（米Anthropic）

AIの進化・発展はますます急速に

◆ **従来型のAI**

ルールベース、画像認識、分野特化型 ・・・

◆ **生成AI（一般的なLLM）**

Chat GPTでブレイク、様々な基盤モデルやサービスが急成長

◆ **マルチモーダルAI**

テキスト、画像、動画、音声、コードなどを双方向に入出力可能に

◆ **推論に強いAIモデル（Reasoning Model）**

OpenAI o4-mini、Gemini 2.0 Flash Thinking、DeepSeek-R1 ・・・

◆ **AIエージェント**

OpenAI Operator、Deep Research、Claude Compute use ・・・

将来？

○ **AGI（Artificial General Intelligence）汎用人工知能？**

○ **ASI（Artificial SuperIntelligence）人工超知能？？**

AIはどこへ向かうのか

- 生成AIが社会にも**徐々に普及**
- 身近な例としては、大手の企業が生成AIにより生成された実在しないモデルを広告モデルとして起用
- 基礎性能の向上、マルチーモーダル化、AIエージェントの普及も進み、**近い将来、社会やビジネスに着実に変化が起こる**

株式会社伊藤園のテレビCMの例

AIタレントをテレビCMに日本で初めて起用。AI生成で出力された多数の顔から選定し、デザイナー・クリエイターが微調整。

（出典：https://www.itoen.co.jp/news/article/64855/　より一部抜粋・追記）

１．AIって何だ？

２．AIのリスクと規律のあり方

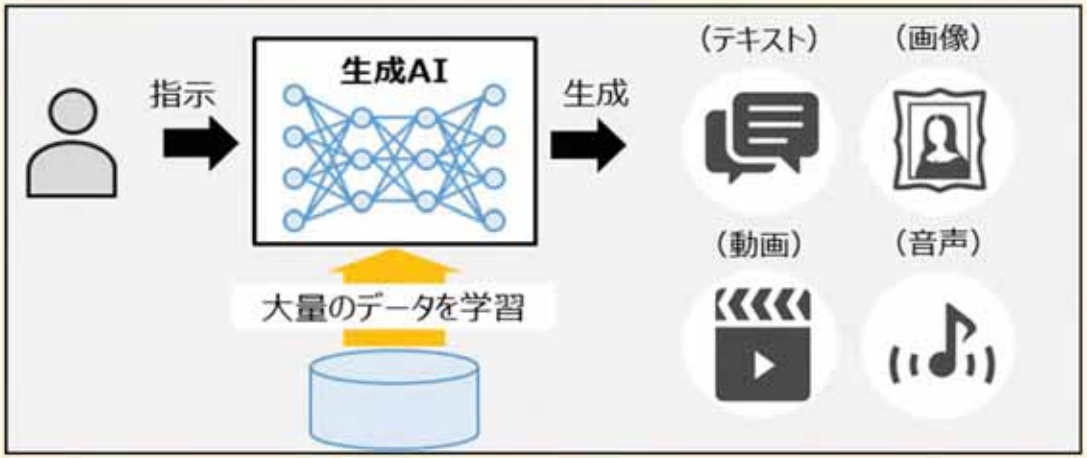
３．AIガバナンスに向けた取組

４．欧米の制度や検討の動向

生成AIのメリットとリスク（再掲）

「生成AI」は、大量の学習データから、文章、画像、動画、音声、プログラムのコードなどを「生成」することが可能

（生成AIの概要）



（アウトプットイメージ）

「生成AI」は「従来型AI」と比較して汎用性が高く創造的であるなどの特徴がある一方、様々なリスクも指摘

生成AIのメリット

- **労働力不足の解消**（問合せ対応、点検・監視など）
- **事務作業の効率化**（文献調査、要約・翻訳、資料原案作成など）
- **イノベーションの創出**（新たな素材、新たな薬等の開発など）
- **地球規模課題の解決**（災害予測・対策、・パンデミック対策など）

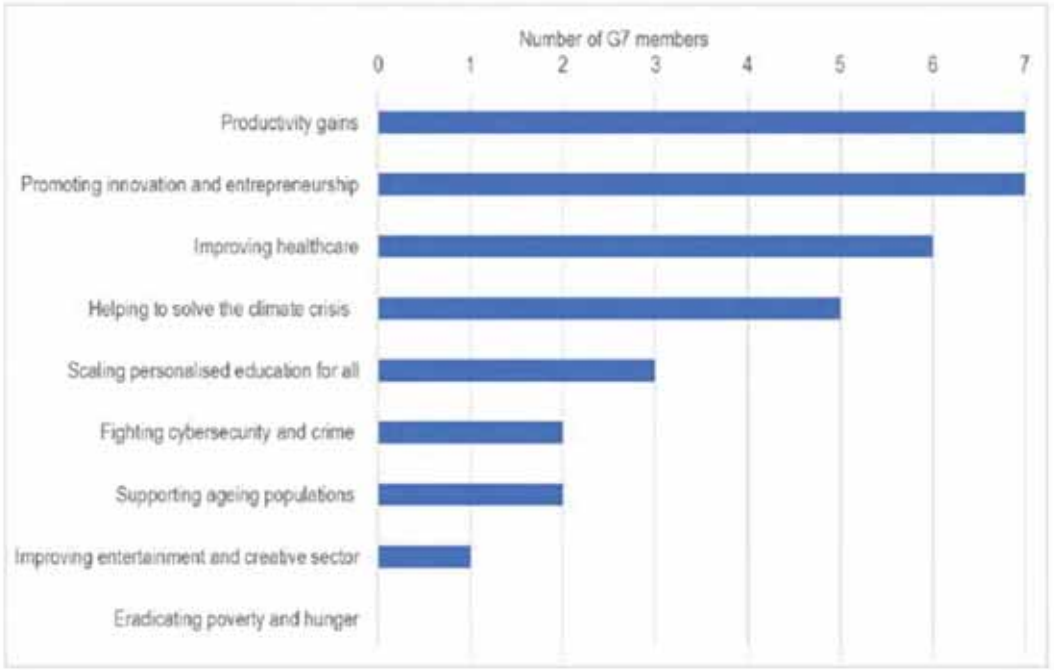
生成AIのリスク

- **犯罪増加リスク**
- **誤情報提示のリスク**
- **著作権侵害のリスク**
- **差別・偏見の増幅リスク**

「生成AI」技術の急速な進展とリスク

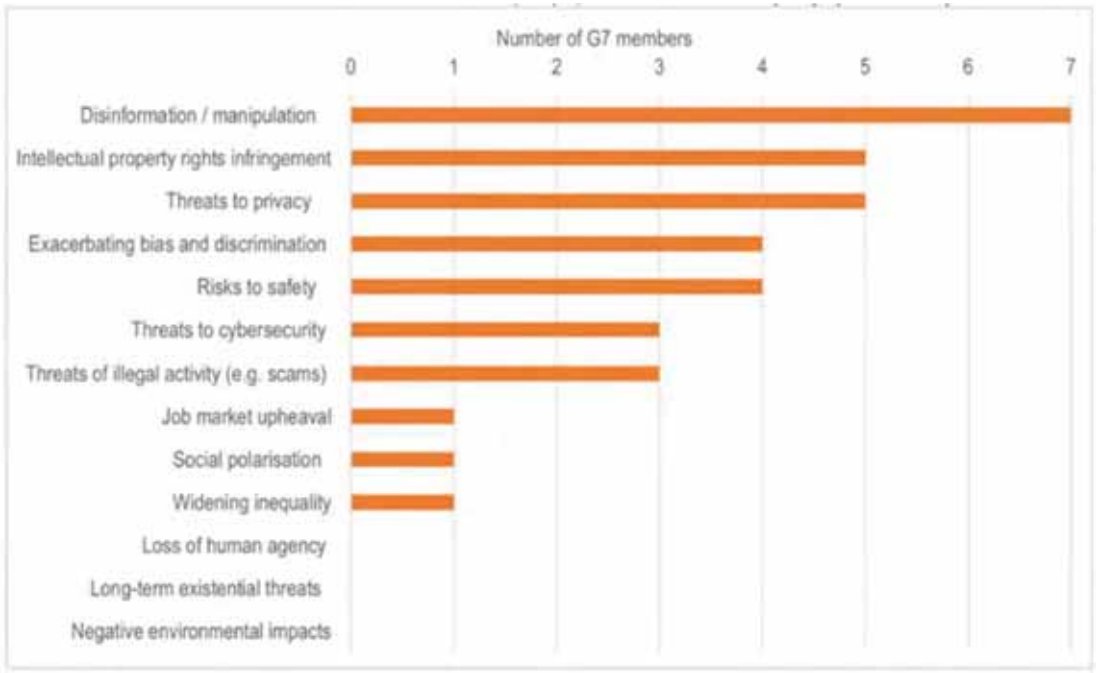
- **大規模言語モデル（LLM:Large Language Models）**の開発が進み、従来人間が得意としてきた、情報を生成・創造する目的で用いられる**生成AI（Generative AI）**技術が急速に進展し、生産性向上等が期待。
- 他方、偽情報・情報操作、知的財産権侵害、プライバシー侵害、偏見・差別の助長、安全上のリスク等のリスクをもたらすとの指摘。特に、**偽情報・情報操作については、G7構成国全てがリスクとして認識**。

◆ G7構成国が選択した生成AIの活用の機会
(あらかじめ与えられた選択肢の中から5つを選択)



Note: The figure aggregates responses from seven respondents to the question: "From your country or region's perspective, what are the top five opportunities generative AI presents to help achieve national and regional goals? (Please select five options)".

◆ G7構成国が選択した生成AIに関するリスク
(あらかじめ与えられた選択肢の中から5つを選択)



Note: The figure aggregates responses from seven respondents to the question: "From your country or region's perspective, what are the top five risks generative AI presents to achieving national and regional goals? (Please select five options)".

【出典】G7 Hiroshima Process on Generative Artificial Intelligence (AI), Report prepared for the 2023 Japanese G7 Presidency and the G7 Digital and Tech Working Group, OECD (2023)

リスクの一例：バイアス

差別的なバイアスが生まれる一例　(採用選考の例)

- 過去の学習データでは、応募者のほとんどが特定の性別

- AIが過去の学習データからその特定の性別を採用することが好ましいと認識

- 特定の性別以外を差別するという欠陥が発生

⇒　こうしたバイアスを放置した場合、バイアスが拡大再生産していく

不適切な出力防止や出力のモニタリング等が重要

AIの安全性とは？

自動車の安全性（一例）

- 仕組み
  ⇒安全装備（エアバッグなど）
- 事故後の対応
  ⇒負傷者の救護など
- 継続的な取り組み
  ⇒車両点検

道路交通法などの規律が存在

AIの安全性（一例）

- 仕組み
  ⇒不適切な出力防止
- 事故後の対応
  ⇒適切な情報開示
- 継続的な取り組み
  ⇒出力のモニタリング／
  安全性評価

AIに関する規律が必要

では、どんな規律が望ましいか

イノベーションと安全性のジレンマ

- AIの安全性の確保のための一定程度AIに対する規律が必要
- 政府による過度な規制は民間企業のAI開発や利用を委縮させ、イノベーションを阻害
- **<u>イノベーションの促進とAIの安全性の確保のバランスが取れた規律が必要</u>**

⇒　では、どうやって「バランス」を取ればよいのか？

規律について①「Code」（ローレンス･レッシグ）における議論

- インターネット上の行為に規律に関して、ローレンス･レッシグ ハーバード大教授が著書「CODE and Other Laws of Cyberspace」（1999年）において、法（Laws)、アーキテクチャ（Architecture)、市場（Market)、社会規範（Norms）による枠組みによる規制を提唱。
- これらの規制方法の適切な組み合わせにより、インターネットの自由と秩序を維持することが出来ると主張。

Code における議論

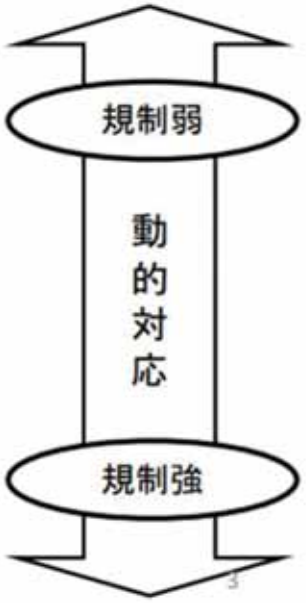
- 法、アーキテクチャ、市場、社会規範による4つの規制方法の適切な組み合わせを強調
- 適切なバランスを欠いた場合、何らかの問題が生じうる可能性（例　二次創作市場の抑圧）

⇒ 一定程度の自由な余地を残すことにより、イノベーションを促進しつつ、
インターネットの健全な発展を志向

規律について②　共同規制（co- regulation）

- 「共同規制」とは、自主規制の持つ柔軟性等の利点と政府規制がもつ信頼性等の利点を組み合わせ、イノベーションに親和的かつ確実なルール枠組みを作り出す政策手法。
- 自主規制から政府規制の各段階には明確な区分があるわけではなく、政策手段としての自主規制と共同規制を合わせて「ソフトロー・アプローチ（政策）」と呼ぶことも多い。
- 例えば、法律により目的や手段の大枠は明記しつつ、詳細な実現手段や執行は事業者に委ねる手法がこれに該当。日本ではデジタルプラットフォーム規制（取引透明化法）で採用。

規制なし	特に規制の必要なく、市場自身が問題の発生を抑止・解決している
自主規制	業界団体等による自主的な規制によって当該問題が適切に解決されている（政府による一般原則の提示は存在し得る）
共同規制	自主規制と政府規制の混合措置により問題が解決されている（政府の自主規制補強措置が存在する）
政府規制	目的とプロセスが政府によって定義されており、政府機関によるエンフォースメントが担保されている



共同規制の概念と実践詳細：拙著『情報社会と共同規制』勁草書房、2011年

英Ofcom [2008] Identifying appropriate regulatory solutions: principles for analysing self- and co-regulation.を元に作成

出典）生貝 直人　一橋大学教授　提出資料から作成　※知的財産戦略本部　構想委員会（第6回）
https://www.kantei.go.jp/jp/singi/titeki2/tyousakai/kousou/2021/dai6/gijisidai.html
⇒生貝氏資料

１．AIって何だ？

２．AIの安全性と規律のあり方

３．AIガバナンスに向けた取組

４．欧米の制度や検討の動向

AI政策の全体像

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
チャンスとリスクに対応

国際的な議論を主導
広島AIプロセスとそのアウトリーチ活動等

AI戦略会議（様々な分野の有識者）

「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）

リスクへの対応

中央省庁における生成AIの扱い、個人情報保護委員会からの注意喚起、教育現場におけるガイドライン、AI事業者ガイドライン 等

知財とAI、著作権とAI、AIと雇用に関する検討 等

AIセーフティー・インスティテュートの創設

AI制度の在り方についての検討

AIの利活用促進

官民における利用促進

リテラシー向上

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

インフラ高度化、人材育成

アジア、グローバルサウスとの連携

＜出典＞AI戦略会議（第11回）、AI制度研究会（第1回）合同会議（令和6年8月2日）資料より

AIガバナンスの確立に向けた取組

		主な取組	主な関係省庁
国際ルール		広島AIプロセス	総務省、外務省
国内ルール	法令等 （ハードロー）	AI法 （＋個別の業法等）	内閣府 （総務省含む各省）
	ガイドライン （ソフトロー）	AI事業者ガイドライン	総務省、経産省
	技術基準 ガイド	AISI評価観点ガイド レッドチーミングガイド	内閣府・AISI （総務省含む各省）

AIガバナンスの確立に向けた取組

		主な取組	主な関係省庁
国際ルール		広島AIプロセス	総務省、外務省
国内ルール	法令等（ハードロー）	AI法 ※5/28成立 （＋個別の業法等）	内閣府 （総務省含む各省）
	ガイドライン（ソフトロー）	AI事業者ガイドライン	総務省、経産省
	技術基準ガイド	AISI評価観点ガイド レッドチーミングガイド	内閣府・AISI （総務省含む各省）

AI戦略会議・AI制度研究会　中間とりまとめ 概要（１／２）

背景

■AIは我が国の発展に大きく寄与する可能性がある一方、**様々なリスクが顕在化**。

■AIに対する不安※の声が多く、諸外国と比べても**開発・活用が進んでいない**との指摘。

※AIに関する意識調査（PwC「2023年AI予測」）の日米比較では、日本は「品質の不安定さ」、「プロセスのブラックボックス化」、「フェイクコンテンツ」にリスクを感じている企業が多いとの結果が示されている。

▶ AIの透明性など、**適正性を確保し、AIの開発・活用を進める**必要がある。

基本的な考え方

■**イノベーション促進**と**リスク対応**の両立

- ●研究開発支援、人材育成、データや計算資源の整備などイノベーションの促進
- ●法令とガイドライン等の適切な組合せ
- ●OECD原則、広島プロセス国際指針等の共通的な指針等と個別の既存法令の活用

■**国際協調**

- ●AIガバナンスの形成に向けて議論をリード
- ●国際整合性・相互運用性の確保

＜出典＞内閣府　https://www8.cao.go.jp/cstp/ai/index.html
⇒中間とりまとめ（2025年2月4日　AI戦略会議・AI制度研究会）概要

AI戦略会議・AI制度研究会　中間とりまとめ 概要（２／２）

具体的な制度・施策の方向性

「世界で最もAIを開発・活用しやすい国」を目指す

速やかな法制度化が必要
世界のモデルになるような制度

■全般的な事項

●政府の司令塔機能の強化、戦略の策定

- 全体を俯瞰する**司令塔機能強化**
- AIの安全・安心な研究開発・活用のための**戦略（基本計画）の策定**

●安全性の向上等

- **国による指針（広島AIプロセス準拠）の整備**、**事業者による協力**
- **国による調査・情報収集、事業者・国民への指導・助言、情報提供等**

■政府等による利用

●適正なAI政府調達・利用 等

■基盤サービス等における利用

●各業法等による対応 等

＜出典＞内閣府　https://www8.cao.go.jp/cstp/ai/index.html
⇒中間とりまとめ（2025年2月4日　AI戦略会議・AI制度研究会）
概要

AI法案の国会審議

◆ 中間とりまとめを踏まえ、2025年2月28日、内閣府が
「人工知能関連技術の研究開発及び活用の推進に関する法律案（AI法案）」
を国会に提出

◆ 日本初のAIに特化した法案

◆ 4月24日、**衆議院本会議において、賛成多数で可決**

※ 附帯決議が付されている

◆ 5月28日、**参議院本会議において、賛成多数で可決され、成立**

※ 附帯決議が付されている

人工知能関連技術の研究開発及び活用の推進に関する法律案（AI法案）の概要

法律の必要性

日本のAI開発・活用は遅れている。

多くの国民がAIに対して不安。

イノベーションを促進しつつ、リスクに対応するため、既存の刑法や個別の業法等に加え、新たな法律が必要。

法案の概要

目的	**国民生活**の向上、**国民経済**の発展
基本理念	経済社会及び**安全保障上重要** ➡ 研究開発力の保持、**国際競争力**の向上 基礎研究から活用まで総合的・計画的に推進 **適正な研究開発・活用**のため**透明性**の確保等　　**国際協力において主導的役割**
AI戦略本部	**本部長：内閣総理大臣　構成員：全閣僚**　　関係行政機関等に対して必要な協力を求める
AI基本計画	研究開発・活用の推進のために**政府が実施すべき施策の基本的な方針**等
基本的施策	**研究開発**の推進、**施設等の整備・共用**の促進　　**人材確保、教育振興** **国際的な規範策定**への参画　　**適正性**のための**国際規範に即した指針**の整備 **情報収集**、権利利益を侵害する**事案の分析・対策検討、調査** **事業者等への指導・助言・情報提供**
責務	国、地方公共団体、研究開発機関、事業者、国民の責務、関係者間の連携強化 **事業者は国等の施策に協力しなければならない**
附則	**見直し規定**（必要な場合は所要の措置）

世界のモデルとなる法制度を構築　国際指針に則り、**イノベーション促進とリスク対応を両立。最もAIを開発・活用しやすい国へ。**

＜出典＞AI戦略会議（第14回）（2025年6月2日）資料より

今後のAI政策の進め方

人工知能（AI）戦略本部の設置

法律附則※1の規定に基づき、公布の日から三月以内に設置予定。

有識者会議の設置

政令又はAI戦略本部決定等により設置予定。

AI基本計画の策定

有識者の意見も踏まえつつ本部で案を作成し、パブリックコメントを経て、AI戦略本部決定/閣議決定予定。

AI指針の整備

既存のガイドライン類との関係を分かりやすく整理しつつ、内閣府で検討予定。

情報収集、調査研究

①主要な業種の活用実態調査、②主要なAI開発者の安全性向上対策の情報収集、③最新の技術や活用事例の調査、④国民の権利利益を侵害する案件・事象の調査を内閣府で実施予定。

国際協調

関係府省庁の協力の下、広島AIプロセス、GPAI※2、AISI※3等の活動の更なる推進

※1 AI法附則（施行期日）第一条 この法律は、公布の日から施行する。ただし、第三章及び第四章並びに附則第三条及び第四条の規定は、公布の日から起算して三月を超えない範囲内において政令で定める日から施行する。（「第四章（法第19条～第28条）」は、AI戦略本部関連の規定。）

※2 GPAI（The Global Partnership on Artificial Intelligence）は、人間中心の考え方に立ち、「責任あるAI」の開発・利用をプロジェクトベースの取組で推進するため、2020年6月に発足した、政府・国際機関・産業界・有識者等のマルチステークホルダーによる国際連携イニシアティブ。

※3 AISI（AI Safety Institute）は、AIの安全性に関する評価手法等を検討・推進するため、2024年２月に独立行政法人情報処理振興機構（IPA）に設置された機関。

6/2 AI戦略会議（第14回）における石破総理発言（抜粋）1/2

・去る５月２８日に、ＡＩ新法が成立をいたしました。皆様方のこれまでの御尽力に、心から感謝を申し上げます。ありがとうございました。

・イノベーションの促進とリスク対応を両立させるＡＩ法により、
『**世界で最もＡＩの研究開発・実装がしやすい国**』を目指してまいります。

・今日がその出発点であります。今後の法律の運用が極めて重要であり、
城内大臣を中心に、具体的な取組を加速していただきたく存じます。

・**本年の秋までに**、ＡＩ法に基づく**ＡＩ戦略本部と有識者会議を設置**いたします。

・**本年冬までに**、本部や有識者会議での議論や、
国民の皆様の御意見を踏まえ、**国の基本的な方針を示す基本計画を策定**してください。
基本計画には、私たちの暮らし、特に、
地方の暮らしがどう変わるのか、
分かりやすいビジョンを盛り込んでください。

＜出典＞首相官邸のAI戦略会議（第14回）（2025年6月2日）に関する発表資料より

6/2 AI戦略会議（第14回）における石破総理発言（抜粋）2/2

・日本が競争力を持つ分野であり、かつ、人手不足対策や生産性向上に資するロボットとＡＩとの融合、いわゆる**『フィジカルＡＩ』に関する競争力強化策**についても盛り込んでください。

・本年冬までに、国際協力や、国際的な情報発信に取り組むとともに、**『ＡＩを開発・活用する者が遵守すべき指針』の整備**を進めてください。

・偽情報の拡散など生成ＡＩをめぐるリスクが指摘される中、ＡＩによって**国民の権利や利益が侵害される事案が発生した場合**には、ＡＩ法に基づき、**国が調査し、必要に応じて事業者への指導や助言などを行っていく必要**があります。

・有識者の皆様におかれましては、引き続き、新たに設置する有識者会議等での御協力をお願いいたします。

＜出典＞首相官邸のAI戦略会議（第14回）（2025年6月2日）に関する発表資料より

9/1 人工知能戦略本部 設置

内閣府
Cabinet Office

English 検索

内閣府の政策 組織・制度 広報・報道 活動・白書等 情報提供

内閣府ホーム > 内閣府の政策 > 科学技術・イノベーション > 新着情報一覧 > 人工知能戦略本部の設置等について

人工知能戦略本部の設置等について

令和7年9月1日
科学技術・イノベーション推進事務局

(1)概要

本日(9月1日)、人工知能関連技術の研究開発及び活用の推進に関する法律(令和7年法律第53号)第19条の規定に基づき、内閣総理大臣を本部長、内閣官房長官及び人工知能戦略担当大臣を副本部長、全ての国務大臣を本部員とする人工知能戦略本部が内閣に設置されました。

また、同日付けで内閣府特命担当大臣(人工知能戦略)が設置され、城内大臣が任命されました。今後、人工知能戦略本部がAI政策の司令塔となって、人工知能関連技術の研究開発及び活用の推進に関する施策を総合的かつ計画的に推進します。

＜出典＞内閣府　https://www8.cao.go.jp/cstp/stmain/20250901ai.html

人工知能基本計画骨子（案）の概要

人工知能基本計画骨子（案）の概要：全体構成

資料１－１

第１章 基本構想 ～「世界で最もＡＩを開発・活用しやすい国」を目指して～

今こそ「反転攻勢」の好機、ＡＩを軸とした経済社会を構築する国家戦略を策定

人とＡＩが協働する「人間中心のＡＩ社会原則」を堅持、イノベーション促進とリスク対応を両立

第２章 施策についての基本的な方針

3原則：イノベーション促進とリスク対応の両立、PDCAとアジャイル対応、内外一体の政策展開

4方針：ＡＩを使う、ＡＩを創る、ＡＩの信頼性を高める、ＡＩと協働する

第３章 政府が総合的かつ計画的に講ずべき施策：4方針に基づく施策集

第１節：ＡＩ利活用の加速的推進：政府で、あるいは社会課題解決のため、まず「使ってみる」

第２節：ＡＩ開発力の戦略的強化：信頼できるＡＩエコシステムを国内で構築、海外にも展開

第３節：ＡＩガバナンスの主導：PDCAサイクルを絶えず回し適正性を確保、国際協調も主導

第４節：ＡＩ社会に向けた継続的変革：産業、雇用、社会への影響を能動的に検証・対応

第４章 施策を政府が総合的かつ計画的に推進するために必要な事項

推進体制の構築（例：規制改革推進室、デジタル庁などとの連携）、基本計画は当面毎年変更

➡ 2025年 年内目途　　ＡＩ戦略本部会合への報告、閣議決定

■人工知能戦略本部（第2回） 資料1-1
https://www8.cao.go.jp/cstp/ai/ai_hq/2kai/2kai.html

ＡＩ法に基づく指針骨子（案）の概要　（1/2）

ＡＩ法に基づく指針骨子（案）の概要

資料１－３

国際的な規範やその基礎として我が国が先んじて策定した「人間中心のＡＩ社会原則」を踏まえ、ＡＩの研究開発・活用における**適正性確保の考え方、適正性確保のための基本方針**を提示。その上で、**主体別に特に取り組むべき事項**を整理した形で構成。

【指針骨子案　全体構成】

1　適正性確保に関する基本的な考え方

●適正性確保の考え方

人間中心	公平性	安全性	透明性	アカウンタビリティ
セキュリティ	プライバシー	公正競争	ＡＩリテラシー	イノベーション

●適正性確保のための基本方針

リスクベースでのアプローチ	ステークホルダーの積極的な関与
一気通貫でのＡＩガバナンスの構築	アジャイルな対応

2　研究開発機関及び活用事業者が特に取り組むべき事項

- ✓ＡＩガバナンスによる俯瞰的な適正性の確保
- ✓ステークホルダーとの信頼関係の構築に向けた透明性の確保
- ✓十分な安全性の確保
- ✓事業継続性確保による安全な環境の維持
- ✓ＡＩのイノベーションの基盤となるデータの重要性を踏まえたステークホルダーへの配慮

3　国及び地方公共団体が特に取り組むべき事項

- ✓ＡＩの積極的かつ先導的な活用によるイノベーションの促進
- ✓社会全体におけるＡＩリテラシーの向上
- ✓ＡＩガバナンスの在り方の検討
- ✓行政としてのアカウンタビリティを果たすこと

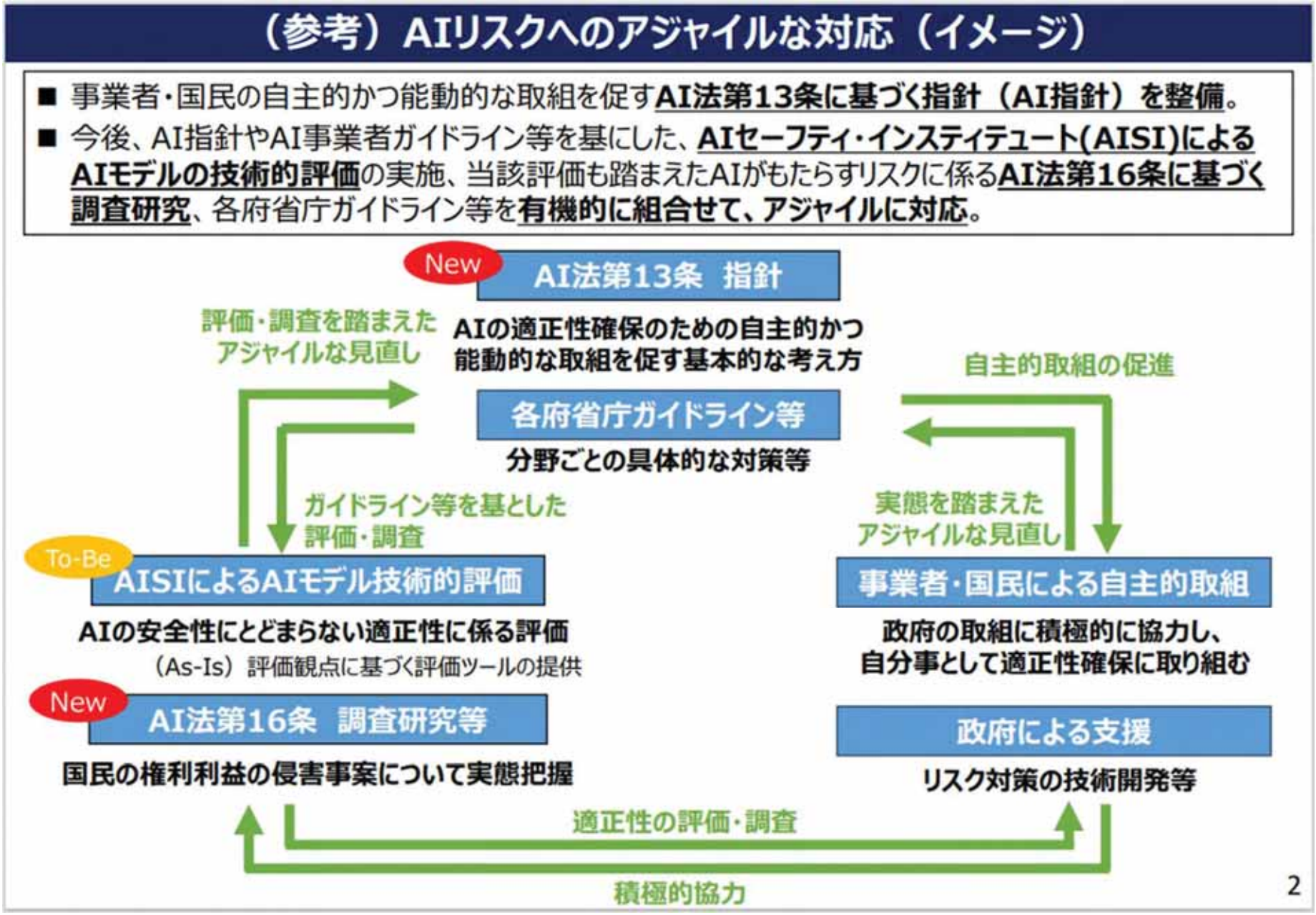
4　国民が特に取り組むべき事項

- ✓人間中心の原則に基づくＡＩの責任ある利用
- ✓ＡＩリテラシーに基づく適切な利用

1

■人工知能戦略本部（第2回）　資料1-3
https://www8.cao.go.jp/cstp/ai/ai_hq/2kai/2kai.html

ＡＩ法に基づく指針骨子（案）の概要　（2/2）



■人工知能戦略本部（第2回） 資料1-3
https://www8.cao.go.jp/cstp/ai/ai_hq/2kai/2kai.html

AIガバナンスの確立に向けた取組

		主な取組	主な関係省庁
国際ルール		広島AIプロセス	総務省、外務省
国内ルール	法令等（ハードロー）	AI法（＋個別の業法等）	内閣府（総務省含む各省）
	ガイドライン（ソフトロー）	AI事業者ガイドライン	総務省、経産省
	技術基準ガイド	AISI評価観点ガイド レッドチーミングガイド	内閣府・AISI（総務省含む各省）

2016年（4/19～4/20）G7香川・高松情報通信大臣会合

<出典> 総務省 https://www.soumu.go.jp/joho_kokusai/g7ict/index.html

2016年（4/19～4/20）G7香川･高松情報通信大臣会合

G7香川･高松情報通信大臣会合　開催結果

5．主な開催結果

大臣会合における主な議論

① 新たなICTがもたらすイノベーションや経済成長

・ イノベーションを創出するため、IoT推進団体間の国際連携を推進することで一致。

・ 我が国より、AIネットワーク化が社会経済に与える影響の分析を国際機関と連携し実施し、AIの開発原則の議論に着手することを提案し、各国から賛同が得られた。

② 情報の自由な流通とサイバーセキュリティ

・ 正当な公共政策目的がある場合を除き、情報の自由な流通を阻害するようなデータローカライゼーション（サーバー設備等の国内設置）要求に対し反対することにつき一致。

・ICTのテロや犯罪への悪用に対し、Ｇ７として強く反対することにつき一致。

・安全・安心なサイバー空間の確保のため、サイバー攻撃傾向の分析技術に関する研究や、サイバーリスクを客観的に把握するための指標の活用などの取組を推進していくことにつき一致。

③ ICTによる地球規模課題の解決

・ マルチステークホルダーによる取組を通じ、2020年までに世界中で新たに15億人をインターネットに接続可能とすることを目指すことで一致。

・ 我が国より、「質の高いインフラパートナーシップ」も踏まえ、質の高いICTインフラの構築を推進することにつき、Ｇ７における連携の可能性を提案、各国から概ね賛意が示された。

・ 高齢化社会や防災など、Ｇ７が先進的な知見を有する分野におけるICTの活用を推進し、「持続可能な開発のための2030アジェンダ」の実現への寄与を確認。

④ 国際連携と国際協力

・ Ｇ７各国の連携・協力による取組を強化するため、「デジタル連結世界」の実現に向けた基本原則や行動計画を取りまとめた「憲章」及び「共同宣言」を採択。

・ また、Ｇ７各国の具体的な取組を集め、国際機関も含めた相互の連携・協力を図るべく「協調行動集」を策定。

2

＜出典＞ 総務省 https://www.soumu.go.jp/main_content/000426060.pdf

ＡＩの研究開発の原則の策定

OECDプライバシーガイドライン、同・セキュリティガイドライン等を参考に、関係ステークホルダーの参画を得つつ、**研究開発に関する原則を国際的に参照される枠組みとして策定**することに向け、検討に着手することが必要。

研究開発に関する原則の策定に当たっては、少なくとも、次に掲げる事項をその内容に盛り込むべき。

① 透明性の原則
AIネットワークシステムの動作の説明可能性及び検証可能性を確保すること。

② 利用者支援の原則
AIネットワークシステムが利用者を支援するとともに、利用者に選択の機会を適切に提供するよう配慮すること。

③ 制御可能性の原則
人間によるAIネットワークシステムの制御可能性を確保すること。

④ セキュリティ確保の原則
AIネットワークシステムの頑健性及び信頼性を確保すること。

⑤ 安全保護の原則
AIネットワークシステムが利用者及び第三者の生命・身体の安全に危害を及ぼさないように配慮すること。

⑥ プライバシー保護の原則
AIネットワークシステムが利用者及び第三者のプライバシーを侵害しないように配慮すること。

⑦ 倫理の原則
ネットワーク化されるAIの研究開発において、人間の尊厳と個人の自律を尊重すること。

⑧ アカウンタビリティの原則
ネットワーク化されるAIの研究開発者が利用者等関係ステークホルダーへのアカウンタビリティを果たすこと。

＜出典＞G7香川高松情報通信大臣会合
https://www.soumu.go.jp/joho_kokusai/g7ict/index.html
⇒AIの研究開発の原則の策定

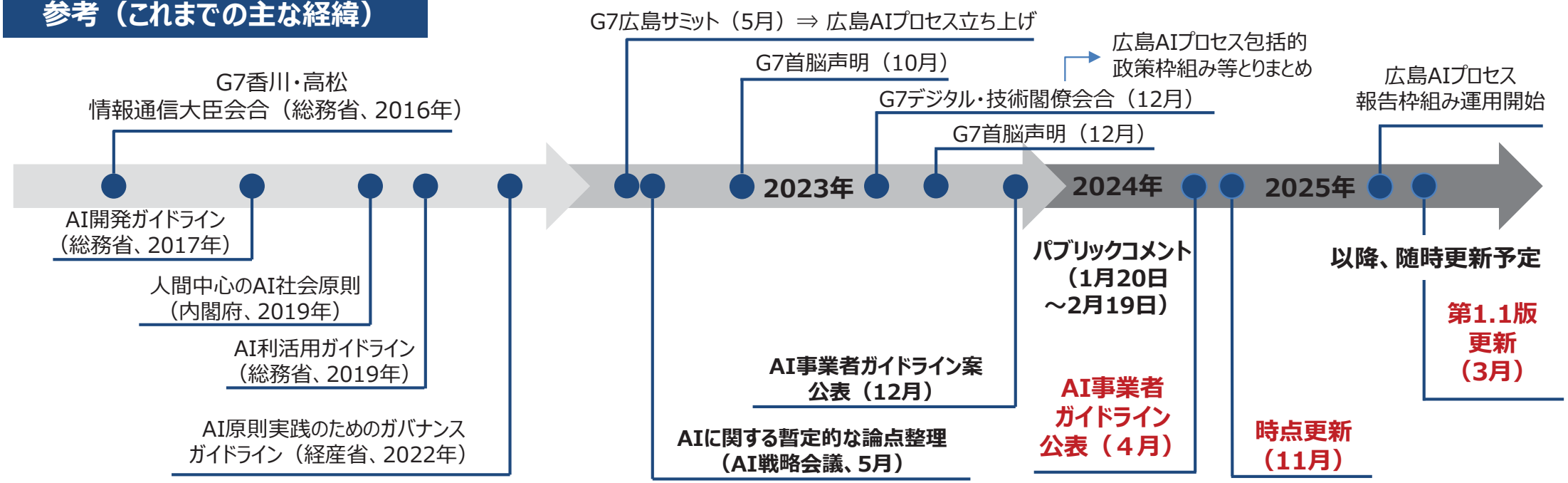
AI事業者ガイドラインの策定・更新

- 「**AIに関する暫定的な論点整理**」（2023年5月、AI戦略会議）を踏まえ、総務省・経済産業省を事務局として、**既存のガイドライン※を統合・アップデート**し、**広範なAI事業者を対象にしたガイドライン**を検討。

  ※AI開発ガイドライン（2017年、総務省）、AI利活用ガイドライン（2019年、総務省）、AI原則実践のためのガバナンスガイドライン（2022年、経済産業省）

- 総務省・経済産業省の合同検討会（2024年3月）にて、「AI事業者ガイドライン（第1.0版）」（案）をとりまとめ。2024年4月、第8回AI戦略会議で報告した上で、同日、**「AI事業者ガイドライン（第1.0版）」を公表、2025年3月には第1.1版として更新。**
- AIを取り巻く様々な環境の変化等を踏まえ、今後も随時更新予定。

参考（これまでの主な経緯）



AI事業者ガイドライン（第1.1版）ファイル構成

「AI事業者ガイドライン」掲載ページ

【AI事業者ガイドライン第1.1版（令和7年3月28日 公表）】

＜ガイドライン資料＞

「AI事業者ガイドライン（第1.1版）」本編

「AI事業者ガイドライン（第1.1版）」別添

「AI事業者ガイドライン（第1.1版）」本編（第1.01版からの見え消し版）

「AI事業者ガイドライン（第1.1版）」別添（第1.01版からの見え消し版）

※チェックリスト等更新がないものについては第1.0版をご参照ください。

＜関連資料＞　※英語版については仮訳。

「AI事業者ガイドライン（第1.1版）」本編（概要）

「AI事業者ガイドライン（第1.1版）」別添（概要）

AI Guidelines for Business Ver1.1

AI Guidelines for Business Appendix Ver1.1

AI Guidelines for Business Ver1.1 (Version showing updated parts from Version 1.01)

AI Guidelines for Business Appendix Ver1.1(Version showing updated parts from Version 1.01)

Outline of "AI Guidelines for Business Ver1.1"

Outline of "AI Guidelines for Business Appendix Ver1.1"

【AI事業者ガイドライン第1.0版（令和6年4月19日 公表）】

＜ガイドライン資料＞

「AI事業者ガイドライン（第1.0版）」本編

「AI事業者ガイドライン（第1.0版）」別添

「AI事業者ガイドライン（第1.0版）」チェックリスト（別添7）

「AI事業者ガイドライン（第1.0版）」ワークシート（別添7）

「AI事業者ガイドライン（第1.0版）」主体横断的な仮想事例（別添8）

「AI事業者ガイドライン（第1.0版）」海外ガイドライン等の参照先（別添9）

「AI事業者ガイドライン」掲載ページ
https://www.soumu.go.jp/main_sosiki/kenkyu/ai_network/02ryutsu20_04000019.html

「AI事業者ガイドライン」の基本的な考え方

- 本ガイドラインは、「1 事業者の自主的な取組の支援」、「2 国際的な議論との協調」、「3 読み手にとっての分かりやすさ」を基本的な考え方としています
- 加えて、「マルチステークホルダー」で検討を重ね実効性・正当性を重視するとともに、「Living Document」として今後も更新を重ねます

考え方

1 事業者の自主的な取組の支援

対策の程度をリスクの大きさ及び蓋然性に対応させる「リスクベースアプローチ」に基づく企業における対策の方向性を記載

2 国際的な議論との協調

国内外の関連する諸原則の動向や内容との整合性を確保

3 読み手にとっての分かりやすさ

「AI開発者」・「AI提供者」・「AI利用者」ごとに、AIに関わる考慮すべきリスクや対応方針を確認可能

＋

プロセス

マルチステークホルダー

教育・研究機関、一般消費者を含む市民社会、民間企業等で構成されるマルチステークホルダーで検討を重ねることで、実効性・正当性を重視したものとして策定

Living Document

AIガバナンスの継続的な改善に向け、アジャイル・ガバナンスの思想を参考にしながら適宜、更新

AI事業者ガイドライン　AI開発者・AI提供者・AI利用者　共通の指針

- 「基本理念」を実現するために各主体が念頭に置くべき原則を、「共通の指針」として整理しています
- 「共通の指針」は、「各主体が取り組む事項」と、「社会と連携した取組が期待される事項」に分類されます

人間中心のAI社会原則

- AI利活用ガイドライン
- AI開発ガイドライン
- AI原則実践のためのガバナンス・ガイドライン

- 信頼性を備えたAIのための倫理ガイドライン
- 人工知能に関する理事会勧告
- 人工知能の倫理に関する勧告
- AI権利章典の青写真
- AI RMF 1.0
- 広島AIプロセス包括的政策枠組み

共通要素を抽出・整理

各主体が取り組む事項

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

社会と連携した取組が期待される事項

8) 教育・リテラシー
9) 公正競争確保
10) イノベーション

AI事業者ガイドライン　AI開発者・AI提供者・AI利用者　共通の指針

- 各主体が連携しバリューチェーン全体で取り組むべき「共通の指針」は、以下のように整理されます
- なおこれらの取組は、 AIのもたらすリスクの程度や各主体の資源制約に配慮しつつ自主的に進めることが重要です

指針		内容（主な項目の抜粋）
各主体が取り組む事項	1) 人間中心	✓ AI が人々の能力を拡張し、多様な人々の多様な幸せ（well-being）の追求が可能となるように行動する ✓ AI が生成した**偽情報・誤情報・偏向情報**が社会を不安定化・混乱させるリスクが高まっていることを認識した上で必要な対策を講じる ✓ より多くの人々がAIの恩恵を享受できるよう**社会的弱者によるAIの活用**を容易にするよう注意を払う
	2) 安全性	✓ 適切なリスク分析を実施し、**リスクへの対策**を講じる ✓ 主体のコントロールが及ぶ範囲で本来の利用目的を逸脱した提供・利用により危害が発生することを避ける ✓ AIシステム・サービスの特性及び用途を踏まえ、学習等に用いるデータの正確性等を検討するとともに、**データの透明性の支援、法的枠組みの遵守**、AIモデルの更新等を合理的な範囲で適切に実施する
	3) 公平性	✓ 特定の個人ないし集団へのその人種、性別、国籍、年齢、政治的信念、宗教等の多様な背景を理由とした**不当で有害な偏見及び差別をなくす**よう努める ✓ AIの出力結果が公平性を欠くことがないよう、AIに単独で判断させるだけでなく、適切なタイミングで人間の判断を介在させる利用を検討した上で、潜在的な**バイアス等に留意**し、AIの開発・提供・利用を行う
	4) プライバシー保護	✓ 個人情報保護法等の**関連法令の遵守、各主体のプライバシーポリシーの策定・公表**により、社会的文脈及び人々の合理的な期待を踏まえ、ステークホルダーのプライバシーが尊重され、保護されるよう、その重要性に応じた対応を取る
	5) セキュリティ確保	✓ AI システム・サービスの**機密性・完全性・可用性を維持**し、常時、AIの安全な活用を確保するため、その時点での技術水準に照らして合理的な対策を講じる ✓ AIシステム・サービスに対する外部からの攻撃は日々新たな手法が生まれており、これらの**リスクに対応するための留意事項を確認**する

AI事業者ガイドライン　AI開発者・AI提供者・AI利用者　共通の指針

- 各主体が連携しバリューチェーン全体で取り組むべき「共通の指針」は、以下のように整理されます
- なおこれらの取組は、 AIのもたらすリスクの程度や各主体の資源制約に配慮しつつ自主的に進めることが重要です

指針		内容（主な項目の抜粋）
各主体が取り組む事項（続き）	6) 透明性	✓ AIを活用する際の社会的文脈を踏まえ、AIシステム・サービスの検証可能性を確保しながら、必要かつ技術的に可能な範囲で、**ステークホルダーに対し合理的な範囲で適切な情報を提供**する（AIを利用しているという事実、活用している範囲、データ収集及びアノテーションの手法、AIシステム・サービスの能力、限界、提供先における　適切/不適切な利用方法、等）
	7) アカウンタビリティ	✓ トレーサビリティの確保や共通の指針の対応状況等について、ステークホルダーに対して情報の提供と説明を行う ✓ 各主体の**AIガバナンスに関するポリシー、プライバシーポリシー等の方針を策定**し、公表する ✓ 関係する情報を文書化して一定期間保管し、必要なときに、必要なところで、入手可能かつ利用に適した形で参照可能な状態とする
社会と連携した取組が期待される事項	8) 教育・リテラシー	✓ AIに関わる者が、その関わりにおいて**十分なレベルのAIリテラシーを確保**するために必要な措置を講じる ✓ AIの複雑性、誤情報といった特性及び意図的な悪用の可能性もあることを勘案して、**ステークホルダーに対しても教育を行う**ことが期待される。
	9) 公正競争確保	✓ AIを活用した新たなビジネス・サービスが創出され、持続的な経済成長の維持及び社会課題の解決策の提示がなされるよう、**AIをめぐる公正な競争環境が維持**に努めることが期待される
	10) イノベーション	✓ 国際化・多様化、**産学官連携**及びオープンイノベーションを推進する ✓ 自らのAIシステム・サービスと他のAIシステム・サービスとの相互接続性及び相互運用性を確保する ✓ 標準仕様がある場合には、それに準拠する

AI開発者・AI提供者・AI利用者別　取り組み事項の例

第3部 **AI開発者に関する事項**	• 適切なデータの学習（適正に収集、法令に従って適切に扱う） • 適正利用に資する開発（安全に利用可能な範囲の設定、AIモデルの適切な選択） • セキュリティ対策の仕組みの導入、開発後も最新動向に留意しリスクに対応 • 関連するステークホルダーへの情報提供（技術的特性、学習データの収集ポリシー、意図する利用範囲等） • 開発関連情報の文書化 • イノベーションの機会創造への貢献 等
第4部 **AI提供者に関する事項**	• 適正利用に資する提供（利用上の留意点の設定、AI開発者が設定した範囲でAIを活用等） • 文書化（システムのアーキテクチャやデータ処理プロセス等） • 脆弱性対応（サービス提供後も最新のリスクを把握、脆弱性解消の検討） • 関連するステークホルダーへの情報提供（AIを利用していること、適切な使用方法、動作状況やインシデント事例、予見可能なリスクや緩和策等） • サービス規約等の文書化 等
第5部 **AI利用者に関する事項**	• 安全を考慮した適正利用（AI提供者が想定した範囲内での適正な利用） • バイアスに留意し、責任をもってAI出力結果の事業利用判断 • プライバシー侵害への留意（機密情報等を不適切に入力しない等） • セキュリティ対策の実施 • 関連するステークホルダーへの情報提供（業務外利用者等に平易かつアクセスしやすい形で示す等） • 提供された文書の活用、サービス規約の遵守 等

「AIによるリスク」とその対策を整理（1/2）

- AIによるリスクを、事業者ができる限り網羅的に把握し対策を検討できるよう、体系的に整理しました

表 3. AI によるリスク例の体系的な分類案

・下表はAIのリスクを網羅したものではなく、想定に基づく事案も含んでおり、あくまで一例として認識することが期待される
・下表には政府等の公的機関も含めた社会全体での対応・議論が必要となるリスクも含まれる

大分類	中分類	リスク例
技術的リスク（＝主にAIシステム特有のもの）	学習及び入力段階のリスク	データ汚染攻撃等のAIシステムへの攻撃
	出力段階のリスク	バイアスのある出力、差別的出力、一貫性のない出力等
		ハルシネーション等による誤った出力
	事後対応段階のリスク	ブラックボックス化、判断に関する説明の不足
社会的リスク（＝既存のリスクがAIにおいても発生又はAIによって増幅するもの）	倫理・法に関するリスク	個人情報の不適切な取扱い
		生命等に関わる事故の発生
		トリアージにおける差別
		過度な依存
		悪用
	経済活動に関するリスク	知的財産権等の侵害
		金銭的損失
		機密情報の流出
		労働者の失業
		データや利益の集中
		資格等の侵害
	情報空間に関するリスク	偽・誤情報等の流通・拡散
		民主主義への悪影響
		フィルターバブル及びエコーチェンバー現象
		多様性・包摂性の喪失
		バイアス等の再生成
	環境に関するリスク	エネルギー使用量及び環境の負荷

「AIによるリスク」とその対策を整理（2/2）

- さらに、事業者におけるリスクへの対策の検討に結び付けるべく、各リスクに対応する主な共通の指針と、事業者における対策の例を記載しました

表 4. AI によるリスク例と共通の指針及び主体毎に重要となる事項のマッピング

・下表はAIのリスクを網羅したものではなく、想定に基づく事案も含んでおり、あくまで一例として認識することが期待される
・下表には政府等の公的機関も含めた社会全体での対応・議論が必要となるリスクも含まれる

	リスク例	関連する共通の指針	「共通の指針」に加えて主体毎に重要となる事項 第3部 AI開発者	第4部 AI提供者	第5部 AI利用者
技術的リスク	データ汚染攻撃等のAIシステムへの攻撃	5）セキュリティ確保	i. セキュリティ対策のための仕組みの導入 ii. 最新動向への留意	i. セキュリティ対策のための仕組みの導入 ii. 脆弱性への対応	i. セキュリティ対策の実施
	バイアスのある出力、差別的出力、一貫性のない出力等	1）人間中心 ①人間の尊厳及び個人の自律 ③偽情報等への対策			
	ハルシネーション等による誤った出力	2）安全性	i. 適切なデータの学習 ii. 人間の生命・身体・財産、精神及び環境に配慮した開発 iii. 適正利用に資する開発	i. 人間の生命・身体・財産、精神及び環境に配慮したリスク対策 ii. 適正利用に資する提供	i. 安全を考慮した適正利用
		3）公平性	i. データに含まれるバイアスへの配慮 ii. AIモデルのアルゴリズム等に含まれるバイアスへの配慮	i. AIシステム・サービスの構成及びデータに含まれるバイアスへの配慮	i. 入力データ又はプロンプトに含まれるバイアスへの配慮
		8）教育・リテラシー			
	ブラックボックス化、判断に関する説明の不足	6）透明性	i. 検証可能性の確保 ii. 関連するステークホルダーへの情報提供	i. システムアーキテクチャ等の文書化 ii. 関連するステークホルダーへの情報提供	i. 関連するステークホルダーへの情報提供
		7）アカウンタビリティ	i. AI提供者への「共通の指針」の対応状況の説明	i. AI利用者への「共通の指針」の対応状況の説明	i. 関連するステークホルダーへの説明

AI事業者ガイドライン　事業者様のガバナンス取り組み事例

・AIガバナンスの取組について、9事業者様（8企業＋1自治体）の事例をコラム化
・AIガバナンスに係るルールメイク、体制整備、ステークホルダとの連携など様々な観点の取組を掲載

（一例）

NECグループ様

コラム 5：NEC グループの AI ガバナンスに関する取組

NEC は、2018 年に AI の利活用に関連した事業活動が人権を尊重したものとなるよう、全社戦略の策定・推進を担う組織として「デジタルトラスト推進統括部」を設置し、2019 年に「NEC グループ AI と人権に関するポリシー（以下、全社ポリシー）」を策定。ガバナンス体制として AI ガバナンス運行責任者（CDO: Chief Digital Officer）を置き、リスク・コンプライアンス委員会や取締役会との関係性を明確化するとともに、外部有識者会議の「デジタルトラスト諮問会議」を設置し、外部とも積極的に連携する等、経営アジェンダとして AI ガバナンスに取り組んでいる（「図 9. AI ガバナンスの推進体制」参照）。

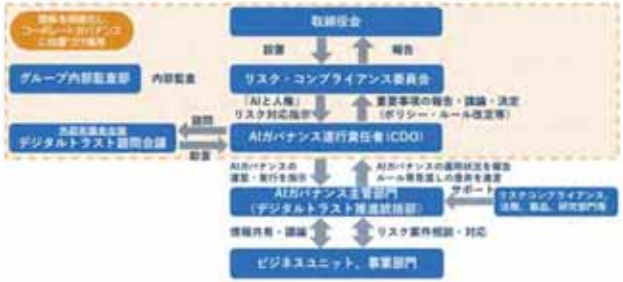


図 9. AI ガバナンスの推進体制

全社ポリシーは、デジタルトラスト推進統括部が国内外の原則や自社のビジョン・価値・事業内容等から検討し、社内の研究開発・サステナビリティ・リスク管理・マーケティング・事業部門などの関係部門や、外部の有識者・NPO・消費者など社内外の様々なステークホルダーとの対話を経て 2019 年 4 月に策定された。このポリシーは、AI 利活用においてプライバシーへの配慮や人権の尊重を最優先するために策定されたものであり、「公平性」、「プライバシー」、「透明性」、「説明する責任」、「適正利用」、「AI の発展と人材育成」、「マルチステークホルダーとの対話」の 7 つの項目から構成されている。

全社ポリシーを実践するために、デジタルトラスト推進統括部が中心となり、社内制度の整備や従業員の研修などを実施している。具体的には、ガバナンス体制や遵守すべき基本的事項が定められた全社規定、対応事項や運用フローが定められたガイドラインやマニュアル、リスクチェックシートを整備し、AI 利活用へのリスクチェックと対策を企画・提案フェーズからフェーズ毎に実施する枠組みを整えている。

また、全従業員向けの Web 研修や AI 事業関係者・経営層向けの研修も実施しており、外部有識者を講師に迎え、最新の市場動向やケーススタディも交えることで理解の促進が図られている（「図 9. AI ガバナンスの推進体制」参照）。

ソフトバンク様

コラム 9：ソフトバンクの AI ガバナンスに関する取組

ソフトバンクは「Beyond Carrier」戦略の下、AI や IoT などの先端技術を活用し、革新的なサービス提供と DX 推進に取り組んでいる。AI については活用が広がる一方で、倫理面での配慮が必要とされている。そこでソフトバンクは AI の適切な活用と安心・安全なサービス提供をするため、2022 年 7 月に「ソフトバンク AI 倫理ポリシー」を策定した。具体的には「人間中心の原則」「公平性の尊重」「透明性と説明責任の追求」「安全性の確保」「プライバシー保護とセキュリティの確保」「AI 人材・リテラシーの育成」の 6 つの項目において指針を定め、この指針に則った事業運営やサービス開発などを行うことを宣言した。また、このポリシーをグループ会社でも適用できる体制を整えており、2024 年 7 月時点で 74 社が適用を決定している。AI 倫理ポリシーを基に規程・基準、ガイドライン、チェックシートなど、様々な社内ルールを制定している。制定にあたっては AI 事業者ガイドライン(第 1.0 版)の基礎となっている内閣府作成の人間中心の AI 社会原則、総務省作成の AI 開発ガイドライン、AI 利活用ガイドライン、経産省作成の AI 原則実践のためのガバナンス・ガイドライン、等の準拠性を考慮して作成している（「図 20. ソフトバンク AI 倫理ポリシー概要とルール構造」参照）。

図 20. ソフトバンク AI 倫理ポリシー概要とルール構造

ソフトバンクの AI ガバナンス推進にあたっては、AI 事業の戦略部門である AI 戦略室がミッションを担い、その中に独立した専門部門として AI ガバナンス推進室を設け、社内 AI 利活用部門のガバナンスを推進している。ステアリングコミッティには、当社の CIO(Chief Information Officer)、CDO(Chief Data Officer)、CISO(Chief Information Security Officer)、CCO(Chief Compliance Officer)が管理監督としてサポートする。アドバイザリーボードとして社内委員、社外委員で編成された AI 倫理委員会が助言を行いガバナンスの推進に活かしている（「図 21. AI ガバナンス推進実行体制」参照）。

NTT DATA様

コラム 10：NTT DATA の AI ガバナンスに関する取組

NTT DATA は、公平かつ健全な AI の活用による価値創造と持続可能な社会の発展を目的に、2019 年から AI ガバナンス整備を開始した。国内外の AI ガバナンスの整備・運用は AI ガバナンス室が推進しており、その活動の全体像は 6 の領域からなる（図 23）。本稿では、AI ガバナンス実装に関わる具体的な取り組みとして、の①②③⑥の領域について紹介する。

図 23. AI ガバナンスの活動

① AI ガバナンス体制の整備

AI の活用によって生じるリスク（AI リスク）はビジネス規模に依らず多大な影響を及ぼすことを踏まえ、AI リスク管理を所掌する専任組織として AI ガバナンス室を設置した。AI リスクの対処には、技術だけではなく、法務や知財、情報セキュリティといった広範囲な知見が必要なことから、技術・法務・知財・情報セキュリティの専門家を集めている。https://www.nttdata.com/global/ja/news/release/2023/032301/

また、国内外の約 600 社の 20 万人を対象に、グローバル全体でコントロールできる体制を整備するために、グループ各社に AI リスク関連窓口を設置し、連携体制を構築している。

② AI ガバナンス基盤の整備

AI ガバナンスの実装に際し、以下の文書を整備している。

- NTT データグループ AI 指針：NTT DATA が AI を活用する上での共通的な考え方。 https://www.nttdata.com/jp/ja/news/release/2019/052900/
- AI リスクマネジメントポリシー：グローバル共通のポリシーとして、管理すべき AI リスクとマネジメントフレームワーク（各社の役割、責任範囲、実装要件）を定義した文書。
- 社内規程（プロジェクト意思決定プロセスへの実装）：AI を含む開発プロジェクト（AI プロジェクト）の実施に際して、AI リスクのチェックを必須化する社内規程。
- 生成 AI 利用ガイドライン：生成 AI に対する開発者・提供者・利用者の各立場に応じた留意事項と対応方針を示した文書。なお、この立場の分類は AI 事業者ガイドラインと共通である。

③ リスクチェックの実装

AI事業者ガイドライン 第1.1版（2025/3）の更新ポイント

構成員・委員や事業者様等からの意見を踏まえ、第1.1版の更新のポイントを次のとおり整理し更新

#	更新論点	主な更新方針
1	**AIによるリスクの洗い出し・分類**	リスクを過度な依存、金銭的損失などを追加、技術的リスク・社会的リスクなどに分類
2	**AIの契約に関する留意事項**	契約モデルの多様化や責任分界に関する記載の充実
3	**生成AIに関する新たなリスクや留意点の記載の追加**	マルチモーダルな生成AI、RAG導入、プログラムコード生成など記載を追加
4	**AIガバナンスに関する事例の充実**	「リスクベース・アプローチ」を実施する上で考慮すべき点、様々な企業の事例追加、「人材不足」の課題を追加
5	**AIガバナンスの動向等の反映**	AI制度研究会、広島AIプロセスの動向等、最新状況を追記
6	**特定単語の整理・見直し**	「バイアス」、「透明性」、「多様性」「包摂性」などの定義や表現の揺れを修正
7	**その他**	UIの改善（目次から該当ページへのリンクを追加）

今後もAIの最新動向を踏まえ、マルチステークホルダの元、Living Documentとして継続して更新し、事業者様の利活用を促していく

AIガバナンスの確立に向けた取組

		主な取組	主な関係省庁
国際ルール		広島AIプロセス	総務省、外務省
国内ルール	法令等 （ハードロー）	AI法 （＋個別の業法等）	内閣府 （総務省含む各省）
	ガイドライン （ソフトロー）	AI事業者ガイドライン	総務省、経産省
	技術基準 ガイド	AISI評価観点ガイド レッドチーミングガイド	内閣府・AISI （総務省含む各省）

AIセーフティ・インスティテュート（AISI）

- 国際的にAIガバナンスの重要性が高まる中、**AI安全性サミット**（昨年11月、英国）を契機に**「AI安全性」を具現化するための議論**が進み、英・米はAIセーフティ・インスティテュートを設置。
- 我が国も第7回AI戦略会議（昨年12月）における**岸田総理からの指示を踏まえ**、本年2月に日本の**AIセーフティ・インスティテュート（所長：村上明子氏）を設置。**
- AI安全性の知見のハブとして、**国内外の関係機関とのネットワークを強化中。AISIの安全性の評価能力を確立**しながら、**安全性評価のための基準、ガイダンスの作成等**を目指す。

日本のAISIの概要

名　称　（日本語）AIセーフティ・インスティテュート
（英　語）Japan AI Safety Institute（略称 J-AISI）

業　務
- 安全性評価に係る調査、基準等の検討
- 安全性評価の実施手法に関する検討
- 他国の関係機関（英米のAI Safety Institute等）との国際連携に関する業務

関係機関　内閣府、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、経済産業省、防衛省

情報通信研究機構、理化学研究所、国立情報学研究所、産業技術総合研究所、情報処理推進機構

村上明子所長

2024年2月１日、内閣府・IPAから内定発表。**2月14日就任**

- 1999年　日本アイ・ビー・エム株式会社 東京基礎研究所　入社
- 2024年　損害保険ジャパン株式会社 執行役員 CDaO(Chief Data Officer) データドリブン経営推進部長（現職兼務）
- 2025年　SOMPOホールディングス株式会社 執行役員常務 グループChief Data Officer（現職兼務）

＜出典＞AI戦略会議（第11回）、AI制度研究会（第1回）合同会議（令和6年8月2日）資料より

AISIの主な取組例

◆ AISIパートナーシップ協定（2024.8）

- AIの安全性に関する取組を進めるためには、国内の関係機関と連携し、共同で対応していくことが必要
- 関係府省庁の協力の下、関係機関が連携してAIの安全性に係る取組を推進していく協力体制を構築

◆ AI事業者ガイドラインと米国NIST AIリスクマネジメントフレームワーク（RMF）とのクロスウォーク第二弾（2024.9）

- 日本の「AI事業者ガイドライン」と米国NISTの「AIリスクマネジメントフレームワーク」について、用語に関するクロスウォーク第一弾に続き、主要なコンセプトを比較

◆ AIセーフティに関する評価観点ガイド（2024.9）

- AIセーフティに関する評価の観点についてガイド（1.0.0版）を作成し公表、2025年3月には第1.10版へ更新し、今後も適宜修正を行い改訂予定
- 主な読者としてAI開発者、AI提供者を想定（特に「開発・提供管理者」、「事業執行責任者」が想定読者）

◆ AIセーフティに関するレッドチーミング手法ガイド（2024.9）

- レッドチーミング手法の観点についてガイド（1.0.0版）を作成し公表、2025年3月に第1.10版へ更新し、今後も適宜修正を行い改訂予定

◆ AISI国際ネットワークにおいて、「Mission Statement」が公開（2024.11）

- 2024年11月20～21日に米国サンフランシスコで開催されたAISI国際ネットワーク（The International Network of AI Safety Institutes）において、「Mission Statement」が公開

AIガバナンスの確立に向けた取組

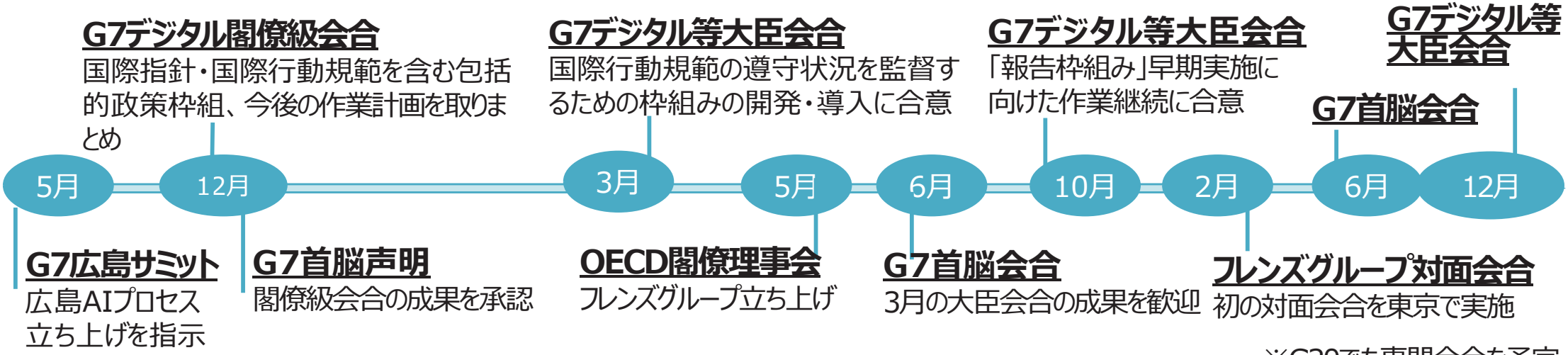
		主な取組	主な関係省庁
国際ルール		広島AIプロセス	総務省、外務省
国内ルール	法令等 （ハードロー）	AI法 （＋個別の業法等）	内閣府 （総務省含む各省）
	ガイドライン （ソフトロー）	AI事業者ガイドライン	総務省、経産省
	技術基準 ガイド	AISI評価観点ガイド レッドチーミングガイド	内閣府・AISI （総務省含む各省）

広島AIプロセスの推進について

- 2023年5月のG7日本議長国下の広島サミットを受け、生成AI等に関する国際ルールの検討を行う**「広島AIプロセス」を立ち上げ**、**安全・安心で信頼できるＡＩを実現するためのルール作りを日本が主導**。同年12月に、**「国際指針」と「国際行動規範」を含む「包括的政策枠組」を取りまとめ**。
- 2024年のG7イタリア議長国下では、**「国際行動規範」を実践・実装していくため、AI開発者による遵守状況の監督の枠組み**を議論。試行的な取組を実施し、早期の本格実施に向けた作業を推進。

  また、**G7を超えて開発途上国を含む多くの国との連携強化を図るため**、５月のOECD閣僚理事会で、広島AIプロセスの精神に賛同する国々の自発的な枠組みである**「広島AIプロセス・フレンズグループ」を立ち上げ**、定期的に情報交換等を実施。2025年11月現在、**59の国・地域が参加**。
- 2025年の**カナダ議長国下においても**、**引き続き広島ＡＩプロセスを推進**することを要請。その上で、①国際行動規範の**報告枠組みの活用促進**、②フレンズグループを通じた**参加国の拡大**、の2点を重点的に推進。

2023 日本議長国 | **2024　イタリア議長国** | **2025　カナダ議長国**



※G20でも専門会合を予定

G７広島AIプロセスとそのアウトリーチ活動について

1　広島AIプロセス

- ✓ 2023年５月、G7広島サミットを受け、生成AIに関する国際的ルールの検討のため、**広島AIプロセス**を立ち上げ。
- ✓ 同年10月及び12月の首脳声明を通じ、**広島AIプロセス包括的政策枠組**を策定・承認。

広島AIプロセス包括的政策枠組

1.生成AIに関するG7の共通理解に向けたOECDレポート	**3.**高度AIシステムの開発組織向け**広島AIプロセス国際行動規範**
2.全てのAI関係者向けの**広島AIプロセス国際指針**(下表)	**4.**偽情報対策に資する研究促進等の**プロジェクトベース協力**

（参考）**国際指針**

【リスク対応】 ➢ 開発・公表前のリスクの特定、評価、軽減 ➢ 開発・公表後の脆弱性、インシデント等の特定・軽減	**【研究開発】** ➢ セキュリティ管理やリスク軽減のための投資、研究、実施 ➢ コンテンツ認証と来歴メカニズムの開発・導入や標準化
【情報共有等】 ➢ AIの性能と制約に関する情報共有 ➢ AIに関する責任ある情報共有とインシデント報告 ➢ AIガバナンス・リスク管理方針の策定、実施、開示	**【その他】** ➢ 個人データ・知的財産保護 ➢ 責任のある利用のためのリテラシー・スキルの向上

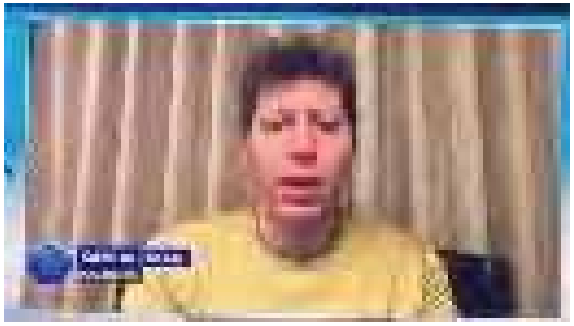
G7広島サミット

広島AIプロセス・フレンズグループ（概要）

2024年5月、OECD閣僚理事会の場で岸田総理（当時）が立ち上げ。
広島AIプロセスに賛同する国を拡大し、グローバルに安心、安全で信頼できるAIの開発・利用の普及を目指す。
2025年11月現在、59か国・地域が参加。

Argentina, Australia, Austria, Belgium, Brunei, Bulgaria, Cambodia, Canada
Chile, Colombia, Costa Rica, Croatia, Cyprus, Czech Republic, Denmark, Egypt
Estonia, Finland, France, Germany, Greece, Hungary, Iceland, India
Indonesia, Ireland, Israel, Italy, Japan, Kenya, South Korea, Laos
Latvia, Lithuania, Luxembourg, Malaysia, Malta, Mexico, Netherlands, New Zealand
Nigeria, Norway, Poland, Portugal, Romania, Serbia, Singapore, Slovakia
Slovenia, Spain, Sweden, Switzerland, Thailand, Türkiye, UAE, United Kingdom
United States, Viet Nam, European Union

OECD閣僚理事会における
フレンズグループ立ち上げイベント



オンライン参加の
サム・アルトマン氏

活動内容

各国のAI政策の状況やAI産業についての情報交換、広島AIプロセスに関する知見の共有のほか、AI企業や学術界などの関係者との意見交換なども通じて、国際的なAIガバナンスの向上・発展を目指す。

国際行動規範の「報告枠組み」について

- 2023年、我が国は、G7議長国として、生成AIに係る国際的なルール形成を行う枠組みである「広島AIプロセス」を立ち上げ、「国際指針」及び「国際行動規範」を取りまとめた。
- 2024年３月のG7産業・技術・デジタル大臣会合では、生成AI開発における透明性及び説明責任を促進するため、**「国際行動規範」を自主的に遵守する企業による履行状況を確認するための適切な手法（「報告枠組み」）を　開発・導入することに合意。**
- 2024年12月、イタリア議長国下のG7で、**「報告枠組み」の基本的な運用方法及び質問票の最終版に合意**。
  -この過程で、行動規範の履行状況に関する報告に向け、AI関連組織と議論を行い、質問票の試行版を作成する手続きを経た。
- **2025年2月7日、正式運用開始。**AIアクションサミット（パリ）のサイドイベントとして、運用開始イベントがOECD本部で開催。

「報告枠組み」の概要

- 「国際行動規範」に沿って作成されたAI関連組織への質問票をOECDのwebサイト上で公開し、回答を依頼。
- **質問票に回答した組織は、OECDwebサイト上で一覧化（リスト化）されるとともに、回答内容も全て公開**。　なお、同webサイト上には2024年イタリア議長国作成のロゴも併せて掲示。
- 「報告枠組み」の対象者：OECD加盟国、GPAI加盟国、又はOECDのAI原則の遵守を表明する国に拠点を置いていること
- なお、回答者は、①回答時点において正確かつ事実に即した情報を提供すること、②年に１度の頻度で回答すること、③全ての質問に回答すること等が求められる。

HAIP Reporting Framework

※イタリアのオリーブの木と日本の桜をイメージしたロゴデザイン

（参考）質問票の内容

- OECDのwebページ(https://transparency.oecd.ai/)で公開されている質問票にweb上で回答。
- 質問項目は広島AIプロセス「国際行動規範」の項目に対応。Yes/No等の選択式、自由記述で構成。

	質問項目	質問例（仮訳）
1	リスクの特定および評価	AIに関連するさまざまなリスクをどのように定義および/または分類していますか？
2	リスク管理および情報セキュリティ	AIライフサイクル全体にわたるリスクと脆弱性に対処するためにどのような措置を講じていますか？
3	高度なAIシステムに関する透明性報告	高度なAIシステムに関連するリスクについて、多様な利害関係者とどのように情報を共有していますか？
4	組織の統治、インシデント管理、および透明性	リスク管理の方針（ポリシー）および実践について、ユーザーおよび/または一般市民に伝達していますか？
5	内容の認証および来歴確認の仕組み	高度なAIシステムによって生成されたコンテンツをユーザーが識別できるようにするコンテンツの生成履歴検出、ラベル付け、または電子透かし手法（メカニズム）を使用していますか？
6	AIの安全性向上と社会リスクの軽減に向けた研究及び投資	コンテンツ認証および来歴の現状を向上させるための研究に、貴社はどのように協力し、投資していますか？
7	人間と世界の利益の促進	ユーザーの意識向上や、高度なAIシステムの性質、能力、限界、影響の理解を支援するためのデジタルリテラシー、教育、研修の取り組みを支援していますか？

「パートナーズコミュニティ」の立ち上げについて

- 広島AIプロセス・フレンズグループ**「パートナーズコミュニティ」は、広島AIプロセスの精神に賛同する民間AI関連企業や国際機関等が参画し、フレンズグループの活動を支援する自発的な枠組み。**
- 「広島AIプロセス・フレンズグループ会合」（2025年2月27日～28日@東京）において、**「パートナーズコミュニティ」の立ち上げを公表。**
- 同コミュニティを通じて、フレンズグループ参加国政府が、**広島AIプロセスをより深く理解し実施できるよう支援**するとともに、各国AI開発者の**「報告枠組み」への参加を促進**し、**「安全、安心で信頼できるAI」**をグローバルに実現することを目的とする。

1．概要

（１）参加者：広島ＡＩプロセスの精神に賛同する**民間AI関連企業、公的機関、国際機関**等

（２）目的：フレンズグループ参加国政府が、**広島AIプロセスをより深く理解し実施できるよう支援**するとともに、各国AI開発者の**「報告枠組み」への参加を促進**し、**「安全、安心で信頼できるAI」をグローバルに実現**する。

（３）活動内容：フレンズグループ会合への参加、フレンズグループ参加国政府への情報提供等。

（４）その他：金銭的な負担等を含め義務等は発生しない。参加者による自発的な活動が基本。

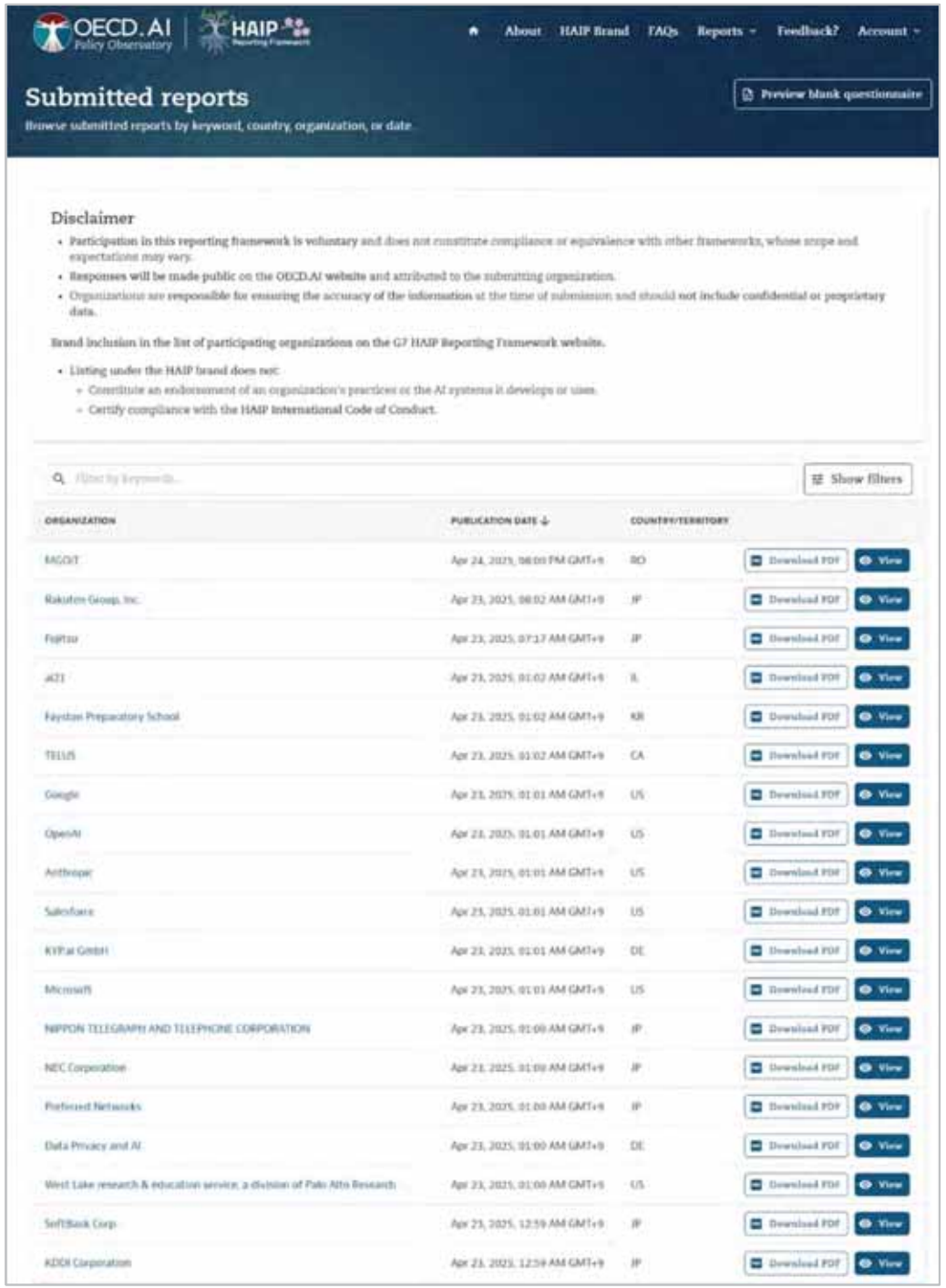
2．参加企業・組織（2025年11月現在28主体。今後も拡大予定。）

Adobe、Amazon、Anthropic、Autodesk、BSA、CAIDP（Center for AI and Digital Policy（NPO）（米））、富士通、Google、日立製作所、Impact AI（AIソリューション提供企業（南ア））、JICA（国際協力機構）、KDDI、Box Japan、Microsoft、NEC、NTT、OECD（経済協力開発機構）、OpenAI、日本オラクル株式会社、Palo Alto Networks（サイバーセキュリティ企業(米)）、Preferred Networks、楽天グループ、SaferAI（AIリスク管理に関するNPO（仏））、Salesforce、ソフトバンク、UNDP（国連開発計画）、WB（世界銀行）、WEF（世界経済フォーラム）

※下線は2025年2月28日の立ち上げ時の参加企業・組織（16主体）。

国際行動規範の「報告枠組み」の状況

- 2025年4月、まず最初の報告として、**19社からの回答が公開**されている（**OpenAI、Google、Anthropic、Microsoft、Salesforce**などを含む）
- **日本企業**からは、**NTT、ソフトバンク、KDDI、楽天、NEC、富士通、Preferred Networksの７社**の報告が公開
- 詳細は**https://transparency.oecd.ai/**を参照



報告枠組み参加組織一覧（2025年11月）

組織名（所在国）	組織名（所在国）
ai21 (Israel)	Anthropic (US)
Data Privacy and AI (Germany)	Fayston Preparatory School (ROK)
Fujitsu (Japan)	Google (US)
KDDI Corporation (Japan)	KYP. Ai GmbH (Germany)
MGOIT (Romania)	Milestone (Denmark)
Microsoft (US)	NEC Corporation (Japan)
NTT (Japan)	OpenAI (US)
Preferred Networks (Japan)	Rakuten Group, Inc. (Japan)
Salesforce (US)	Soft Bank Corp. (Japan)
TELUS (Canada)	West Lake Research and Education Service (Palo Alto) (US)
TELUS Digital (Canada)	Amazon (US)
Hitachi, Ltd (Japan)	ABEJA.inc (Japan)

１．AIって何だ？

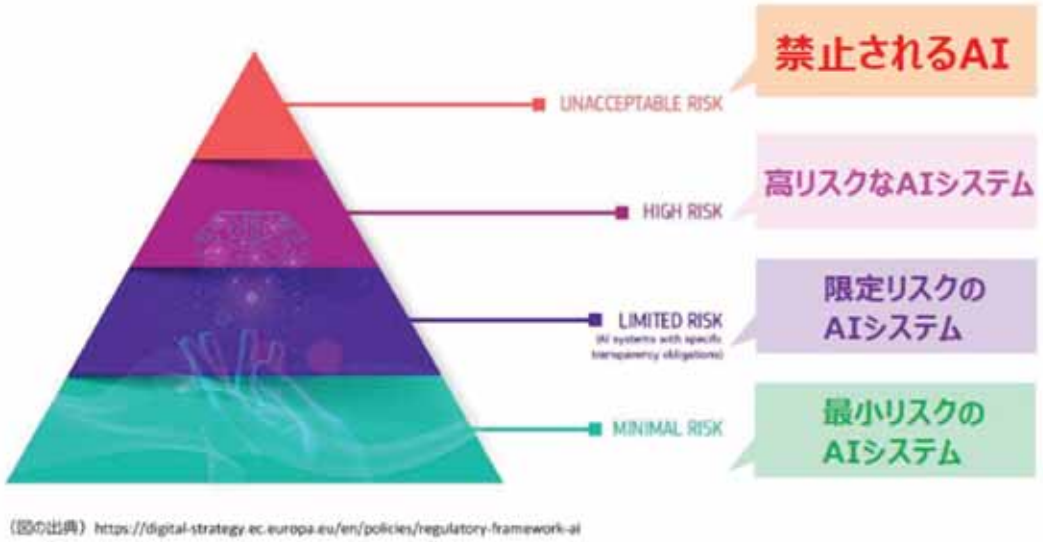
２．AIの安全性と規律のあり方

３．AIガバナンスに向けた取組

４．欧米の制度や検討の動向

【EU】AI Act（1/2）

- 安全性の確保、基本的人権の遵守等とともにイノベーションを推進することを目指しており、リスクに応じて規律を設定する「リスクベースアプローチ」を採用。
- 汎用AI（General Purpose AI：GPAI）システムについての定義や義務等も規定。



（図の出典）https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

禁止されるAI （AI法 第2章§5）

市民の権利を脅かす特定のAIを禁止

- センシティブな特徴（人種、政治、宗教、性的指向等）に基づくバイオメトリクス分類システムの市場への投入、使用
- 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の収集
- 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・脆弱性を悪用するために使用されるAI　等

高リスクなAIシステム （AI法 第3章§6～49）

高リスクなAIシステムの定義

- 付属書Ⅰの法令（機械、玩具、医療機器等）の対象製品の**セーフティコンポーネントとしての使用が意図**されているか、それ自体が同法令の対象のもの
- 同法令によって**第三者による適合性評価が必要なもの**
- 付属書Ⅲで定めるもの（**重要インフラ、教育・職業訓練、雇用、必要不可欠な民間・公共サービス（医療、銀行等）、司法と選挙など民主プロセス等**）

高リスクなAIシステムの義務

- 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。

限定リスクのAIシステム （AI法 第4章§50）

透明性の義務

- **ディープフェイク**のようなAIにより生成された画像、音声、映像コンテンツは、明確に、人為的に**生成または操作されたものであることを開示**

最小リスクのAIシステム （AI法 第10章§95）

行動規範

- **「高リスクなAIシステム」の要件を任意に適用することを促す**
- AIシステムが環境の持続可能性に与える影響の評価や、AIの開発、運用、利用に携わる人々のAIリテラシーの向上などを内容とする、**AIシステムに対する自主的な適用に関する行動規範の策定を促進**

【EU】AI Act（2/2）

汎用AIシステムの扱い （AI法 第5章 §51～56）

- 汎用AI（GPAI、General Purpose AI）システム 及び そのベースとなるGPAIモデル※1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等）
- システミック・リスク※2を伴う影響力の大きいGPAIモデル※3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等）

※1「汎用AIモデル（*General-purpose AI models*）」は、大量のデータを用いて学習され、様々なタスクを適切に実行することができるモデル

※2「システミック・リスク」は、汎用AIモデルの特有のリスクであり、公衆衛生、安全、治安、基本的人権、または社会全体に重大な影響を与えるリスクを指す

※3 学習に使用される計算の累積量が浮動小数点演算で***10^{25}*を超える場合、高い影響力があると推定**される

イノベーション支援策 （AI法 第6章 §57～63）

- 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 （AI法 第12章 §99～101）

- 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%）等

発効 （AI法 第13章 §113）

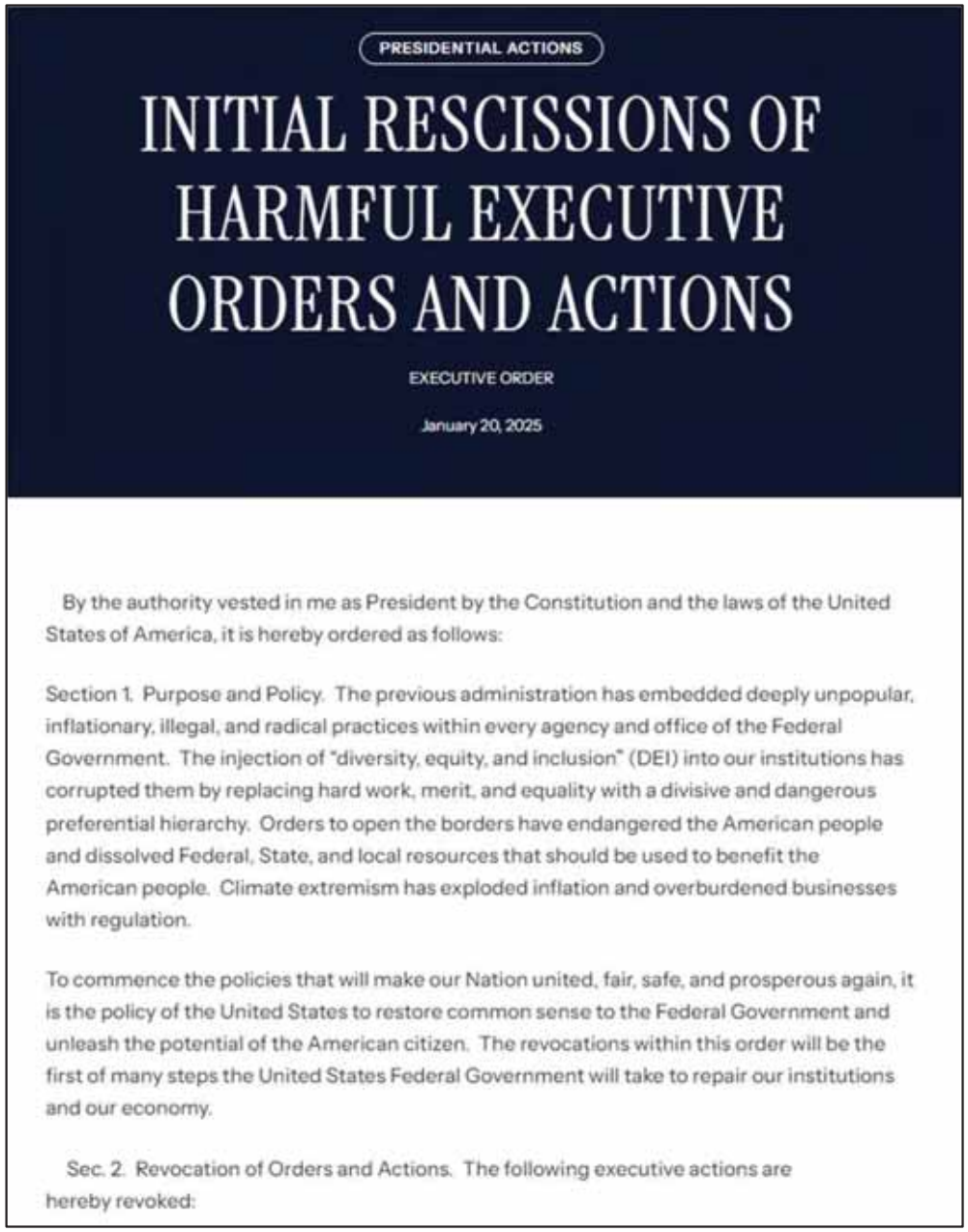
- 基本的に発効日から24カ月後に適用、ただし、
  - AIシステムの定義、AIリテラシーのほか、「禁止されるAI」に係る規定は、2025年2月から適用開始
  - 汎用AIシステムに関する規定や罰則規定等は、2025年8月から適用開始予定
  - その他、「高リスク」、「限定リスク」、「最小リスク」のAIシステムに係る規定等は、2026年８月から適用開始予定

【米国】2023年10月の大統領令の概要

- AIがもたらすメリットを把握し、そのリスクを管理することを目的として、2023年10月30日に**「AIの安全、安心で信頼できる開発と利用に関する大統領令」**（The Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence）を発表
- 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指示する内容

セクション名	主な内容
Sec. 4. **AI技術の安全性・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・核兵器に関する脅威の軽減、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
Sec. 5. **イノベーション及び競争の促進**	AI人材誘致、イノベーション促進、競争促進等
Sec. 6. **労働者の支援**	労働市場におけるAI活用や影響に関するレポートの発行等
Sec. 7. **公平性及び公民権の推進**	刑事司法制度におけるAI利用、公民権の保護等
Sec. 8. **消費者、患者、交通機関利用者、学生の保護**	ヘルスケア、運輸交通、教育等に関する連邦政府機関によるAI活用等
Sec. 9. **プライバシー保護**	連邦政府機関向けのプライバシー保護関連ガイダンス等
Sec. 10. **連邦政府によるAI利用の促進**	AI管理指針、調達指針の策定、政府におけるAI人材確保等
Sec 11. **米国の国際的リーダーシップの強化**	米国のAI基準やフレームワークの海外への推進等

トランプ大統領就任初日の大統領令（2025.1.20）
Initial Rescissions of Harmful Executive Orders and Actions



PRESIDENTIAL ACTIONS

INITIAL RESCISSIONS OF HARMFUL EXECUTIVE ORDERS AND ACTIONS

EXECUTIVE ORDER

January 20, 2025

By the authority vested in me as President by the Constitution and the laws of the United States of America, it is hereby ordered as follows:

Section 1. Purpose and Policy. The previous administration has embedded deeply unpopular, inflationary, illegal, and radical practices within every agency and office of the Federal Government. The injection of "diversity, equity, and inclusion" (DEI) into our institutions has corrupted them by replacing hard work, merit, and equality with a divisive and dangerous preferential hierarchy. Orders to open the borders have endangered the American people and dissolved Federal, State, and local resources that should be used to benefit the American people. Climate extremism has exploded inflation and overburdened businesses with regulation.

To commence the policies that will make our Nation united, fair, safe, and prosperous again, it is the policy of the United States to restore common sense to the Federal Government and unleash the potential of the American citizen. The revocations within this order will be the first of many steps the United States Federal Government will take to repair our institutions and our economy.

Sec. 2. Revocation of Orders and Actions. The following executive actions are hereby revoked:

◆ バイデン政権時の大統領令や大統領覚書約80件を取り消し

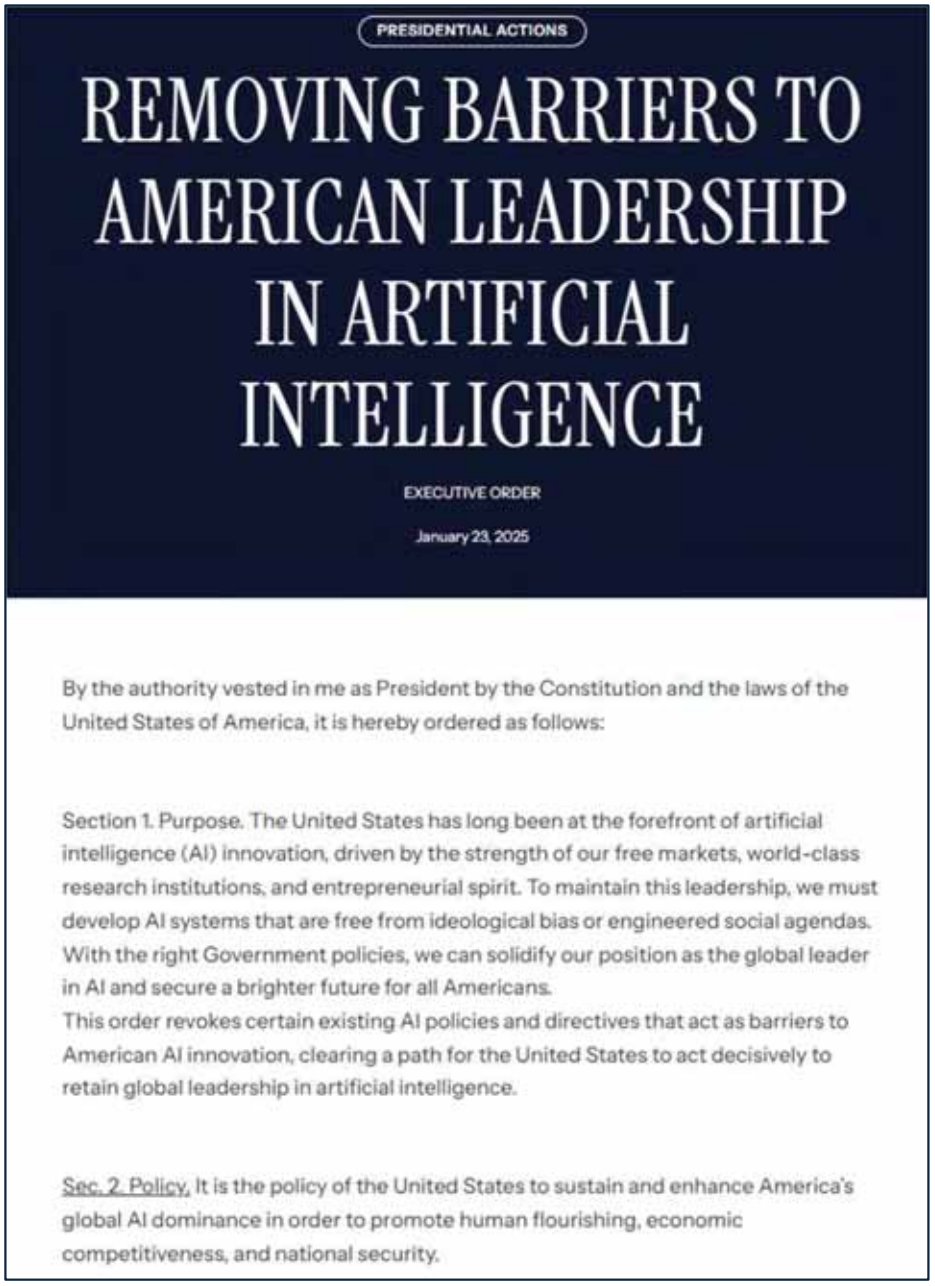
◆ 2023年10月30日の大統領令（Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence）も取り消し

◆ その上で、Domestic Policy Council（DPC）とNational Economic Council（NEC）に対し、
- 今回取り消した大統領令等に基づき実施された連邦政府の全ての取組を見直し、必要に応じて取り消し・置き換え・修正
- 45日以内に、前政権の大統領令等で取り消すべきものの追加リストと、米国の繁栄を増進させるため置き換えとなる大統領令等のリストを提出

すること等を求めている

＜出典＞https://www.whitehouse.gov/presidential-actions/2025/01/initial-rescissions-of-harmful-executive-orders-and-actions/

さらに新たな大統領令（2025.1.23）
Removing Barriers to American Leadership in Artificial Intelligence



PRESIDENTIAL ACTIONS

REMOVING BARRIERS TO AMERICAN LEADERSHIP IN ARTIFICIAL INTELLIGENCE

EXECUTIVE ORDER

January 23, 2025

By the authority vested in me as President by the Constitution and the laws of the United States of America, it is hereby ordered as follows:

Section 1. Purpose. The United States has long been at the forefront of artificial intelligence (AI) innovation, driven by the strength of our free markets, world-class research institutions, and entrepreneurial spirit. To maintain this leadership, we must develop AI systems that are free from ideological bias or engineered social agendas. With the right Government policies, we can solidify our position as the global leader in AI and secure a brighter future for all Americans.

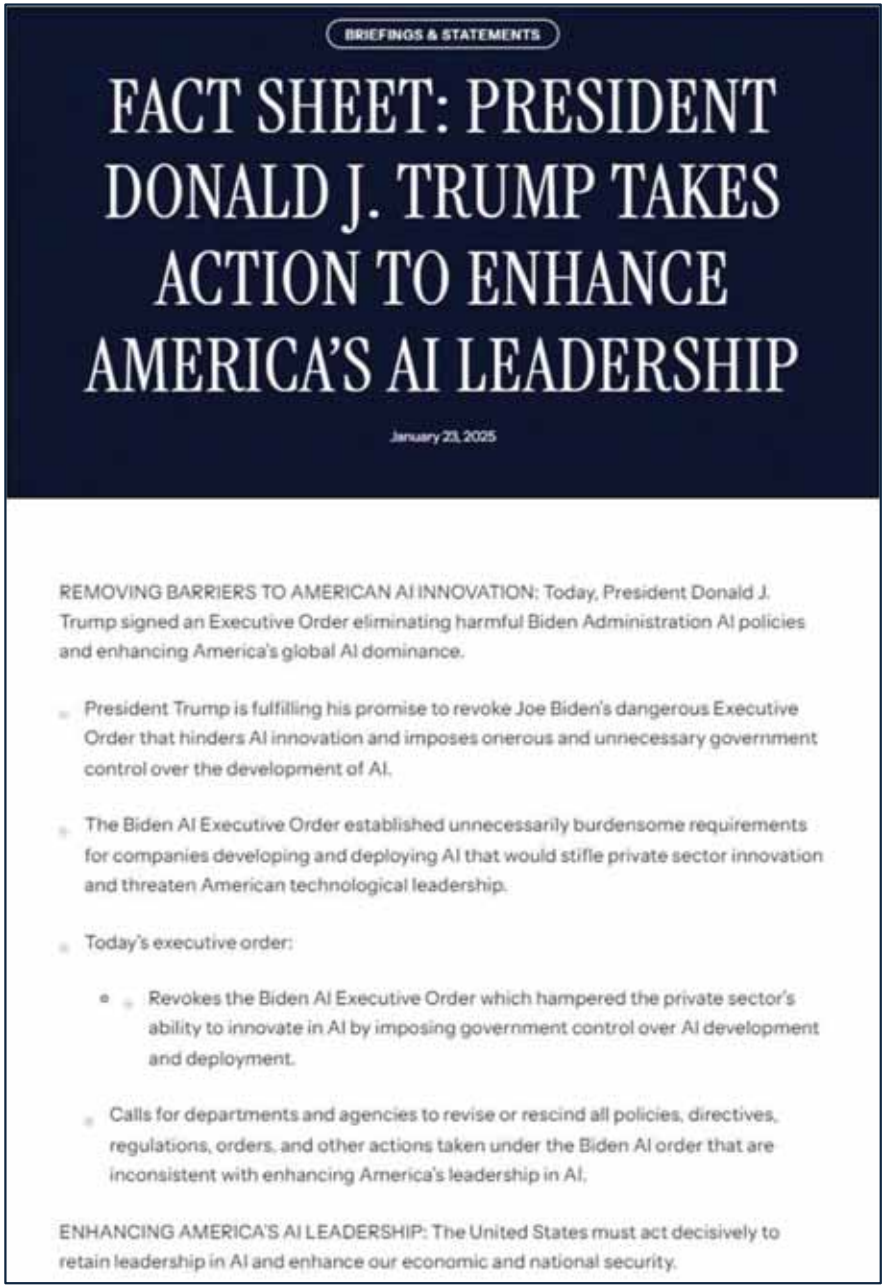
This order revokes certain existing AI policies and directives that act as barriers to American AI innovation, clearing a path for the United States to act decisively to retain global leadership in artificial intelligence.

Sec. 2. Policy. It is the policy of the United States to sustain and enhance America's global AI dominance in order to promote human flourishing, economic competitiveness, and national security.

- イデオロギー的なバイアスや社会的意図に左右されないAIシステムを開発することが必要、米国のAIイノベーションの障害となっている既存の（バイデン政権での）AI政策や指令を取り消し、グローバルなリーダーシップを維持すること等を目的に明記
- 米国の政策は、人類の繁栄、経済競争力、国家安全保障の促進のため、米国のグローバルなAI優位性を維持・強化することと規定
- 関係する大統領補佐官や、AI・暗号資産特別顧問等に、180日以内に行動計画を策定して提出すること、バイデン政権時代の大統領令に基づき実施された政策、規制等をレビューして今回の大統領令に反するものを特定し、一時停止、改訂、取り消し等を提案すること等を求める内容

＜出典＞https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/

これらの取組をまとめたファクトシートも発表（2025.1.23）
President Donald J. Trump Takes Action to Enhance America's AI Leadership



BRIEFINGS & STATEMENTS

FACT SHEET: PRESIDENT DONALD J. TRUMP TAKES ACTION TO ENHANCE AMERICA'S AI LEADERSHIP

January 23, 2025

REMOVING BARRIERS TO AMERICAN AI INNOVATION: Today, President Donald J. Trump signed an Executive Order eliminating harmful Biden Administration AI policies and enhancing America's global AI dominance.

- President Trump is fulfilling his promise to revoke Joe Biden's dangerous Executive Order that hinders AI innovation and imposes onerous and unnecessary government control over the development of AI.
- The Biden AI Executive Order established unnecessarily burdensome requirements for companies developing and deploying AI that would stifle private sector innovation and threaten American technological leadership.
- Today's executive order:
  - Revokes the Biden AI Executive Order which hampered the private sector's ability to innovate in AI by imposing government control over AI development and deployment.
  - Calls for departments and agencies to revise or rescind all policies, directives, regulations, orders, and other actions taken under the Biden AI order that are inconsistent with enhancing America's leadership in AI.

ENHANCING AMERICA'S AI LEADERSHIP: The United States must act decisively to retain leadership in AI and enhance our economic and national security.

- ◆ バイデン政権時におけるAI政策について、有害で、AI開発に対し不要に政府が関与する危険な大統領令を出し、不必要に大きな負担の要件を企業に求め民間のAIイノベーションを阻害、米国の技術的リーダーシップを損なった等、厳しい表現で批判
- ◆ 同日の大統領令（前頁参照）の概要や狙いを紹介
- ◆ また、トランプ政権では、AIが継続的な優先課題であり、第１期政権時の取組を紹介しつつ、それらに基づき、現在、AI分野での米国のリーダーシップが優先事項としている

＜出典＞https://www.whitehouse.gov/briefings-statements/2025/01/fact-sheet-president-donald-j-trump-takes-action-to-enhance-americas-ai-leadership/

米国　One Big Beautiful Bill Act

- 減税や歳出削減、債務上限引き上げなどをまとめた「One Big Beautiful Bill Act」にもAI関係が盛り込まれている
- システムの近代化、業務効率化、サービス向上、セキュリティ向上等のため、商務省にAIシステムや自動意思決定システム導入等の5億ドルを充当
- **他方、本法案の制定日から10年間、州またはその政治的下位区分（自治体）によるAIモデル、AIシステム、自動意思決定システムへの規制を禁止する条項も盛り込まれている**
- 2025年5月22日、連邦議会下院において僅差で可決（賛成215、反対214、棄権1）
- 2025年7月1日　**米上院は99対1という圧倒的多数で、州のAI規制を10年間禁止する条項を撤回**

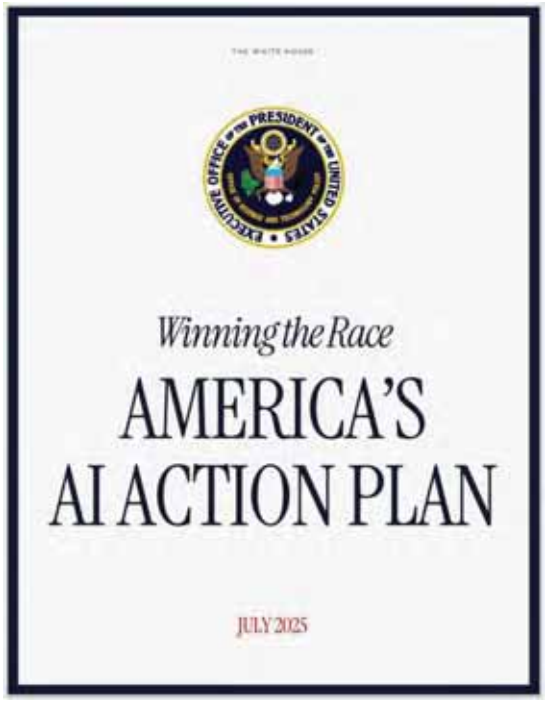
H.R.1　One Big Beautiful Bill Act　119th Congress (2025-2026)

PART 2 -- ARTIFICIAL INTELLIGENCE AND INFORMATION TECHNOLOGY MODERNIZATION
SEC. 43201. ARTIFICIAL INTELLIGENCE AND INFORMATION TECHNOLOGY MODERNIZATION INITIATIVE.

(c) Moratorium.

(1) In general. -- Except as provided in paragraph (2), no State or political subdivision thereof may enforce, during the 10-year period beginning on the date of the enactment of this Act, any law or regulation of that State or a political subdivision thereof limiting, restricting, or otherwise regulating artificial intelligence models, artificial intelligence systems, or automated decision systems entered into interstate commerce.

（ご参考）AMERICA’S AI ACTION PLAN ※2025/7/23



2025/7/23、AMERICA’S AI ACTION PLAN 発表

3つの柱と30の項目

- Pillar I: Accelerate AI Innovation
- Pillar II: Build American AI Infrastructure
- Pillar III: Lead in International AI Diplomacy and Security

https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf

3つの大統領令に署名

- Promoting the Export of the American AI Technology Stack
- Accelerating Federal Permitting of Data Center Infrastructure
- Preventing Woke AI in the Federal Government

https://www.ai.gov/

最後に