



広島AIプロセスとAIガバナンスの国際的動向

2025年3月

総務省 情報通信国際戦略特別交渉官

OECD デジタル政策委員会 議長

G7 広島AIプロセス・フレンズグループ 議長

飯田 陽一



AIに関する議論の国内外における大きな流れ

2016

2017

2018

2019

2020

G7/G20

**G7香川・高松
情報通信大臣会合**

AIの国際的議論
の必要性を提起

G7 情報通信・産業
大臣会合
(イタリア)

基本理念の共有
「人間中心のAI」

G7 シャルルボワ
サミット (カナダ)

AIに関する
共通ビジョン

G7 デジタル大臣
会合 (フランス)

G7 科学技術大臣
会合 (アメリカ)

GPAI

(2020年6月～)

G20大阪サミット

G20茨城つくば
貿易・デジタル経済
大臣会合
(2019年6月)

G20AI原則の合意

ONE AI

(2020年2月～)

OECD/デジタル経済政策委員会 (CDEP)

AIGO (AI専門家グループ)
(2018年5月～)

AIに関する原則の検討・策定



OECD AI原則

日本からのインプット

総務省

内閣府

AI開発ガイドライン案 (2017年7月)

人間中心のAI社会原則 (2019年3月)

AI利活用原則案 (2018年8月)

その他の国際機関

- UNESCO (2018年～)
- 欧州評議会 (2019年～)

等

AI開発原則の策定に関する提案

① 透明性の原則

A I 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

A I ネットワークシステムが利用者を支援するとともに、利用者を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A I 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

A I 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

A I 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

A I 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A I 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A I 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。



AIの開発・実装が進む中、**予測可能で、安定的かつ柔軟な開発・利用環境**が求められている。そのため、「**人間中心**」の考え方を基本にAIに関わる全ての人に適用される**実用的な指針**が必要である。

原則	説明
包摂的な成長、持続可能な開発及び幸福	AIに関わる全ての人、人間にとって有益な成果を追求するために、AIの責任ある管理・運用に積極的に取り組むべき。
人間中心の価値及び公平性	AI開発・運用者は、法の支配、人権及び民主主義的価値観を尊重すべき。その文脈に適合したメカニズムと予防措置を備えるべき。
透明性及び説明可能性	AI開発・運用者は、AIシステムへの一般的な理解やAIの影響を受ける人の理解を促進するため、意味のある情報を提供すべき。
頑健性、セキュリティ及び安全性	AI開発・運用者は、AIシステムの入力データ、処理過程及び決定に関し、検証可能なものとすべき。また、体系的なリスク管理を行うべき。
アカウントビリティ	AI開発・運用者は、AIシステムの適切な作動や上記の原則を尊重していることについて、アカウントビリティを果たすべき。



推奨される政府の取組

研究開発への投資

AIエコシステムの発展

イノベーションと競争を促すための政策環境

人材育成労働市場変革への備え

国際協力

OECD AI原則の改定について

- 2016年G7香川・高松情報通信大臣会合における、AIの研究開発等に関する国際的なガイドラインの必要性に係る我が国の提案を契機としてOECDは調査・分析に着手。策定過程において、総務省のAI開発ガイドラインやAI利活用原則案など国内の検討成果をインプットし、2019年5月、AIに関する最初の政府間スタンダードとなるOECD AI原則を公表。
- 同原則は、履行状況を採択後5年以内に報告することとされており、2023年中頃より、OECDは改定に向けた作業に着手。改定案には、**広島AIプロセスの成果を踏まえ、生成AIによる偽・誤情報への対処に関する追記等**が含まれ、我が国が議長を務めた2024年5月のOECD閣僚理事会において採択・公表。

AI原則(改定版)の項目	主な内容
1.1 包摂的な成長、持続可能な開発及び幸福	AIに関わる全ての人々、創造性の向上、包摂性の促進、格差の改善、持続可能な開発といった人々と地球への有益な結果を追求するために、AIの責任ある管理・運用に積極的に取り組むべき。
1.2 法の支配、人権並びに公平性及びプライバシーを含む民主主義的価値の尊重	AI展開・運用者は、法の支配、人権及び民主主義的・人間中心的価値観を尊重すべき。これにはプライバシーとデータの保護や AIにより増幅される偽・誤情報への対処 も含まれる。さらに、目的外や不適切な利用に起因するリスクへの対処のためのメカニズムと予防措置を備えるべき。
1.3 透明性及び説明可能性	AI展開・運用者は、AIシステム自体やAIから受ける影響、AIが予測や推奨、決定 やコンテンツ を出力する仕組みや入力データへの理解を促進し、異議申立を可能とする情報を提供すべき。
1.4 頑健性、セキュリティ及び安全性	AIシステムは、通常時のみならず悪条件下における利用においても安全やセキュリティ上のリスクをもたらさず正常に機能すべき。不当な危害や望ましくない挙動を安全に制御し、表現の自由を尊重しつつも 情報インテグリティを強化する仕組みが具備されるべき 。
1.5 説明責任	AI展開・運用者は、AIシステムの適正な作動や上記原則の尊重について説明責任を果たすため、AIシステムの出力や応答を分析可能とする追跡可能性を確保すべき。また系統化されたリスク管理手法を適用し、リスクに対応する責任ある企業行動を採用すべき。

推奨される
政府の取組

2.1 AIの
研究開発
への投資

2.2 包摂的な
AIを推進する
エコシステムの整備

2.3 **相互運用可能な
ガバナンス**及び
政策環境の形成

2.4 人材育成及び
労働市場の変化
への備え

2.5 信頼できる
AIのための
国際協力

GPAI (Global Partnership on AI) 1.0

- 設立趣旨 : 人間中心の考え方に立ち、**OECD AI原則に基づき「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**するために設立された、政府・国際機関・産業界・有識者等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**。
- 設立経緯 : 2019年G7ビアリッツサミット (仏) においてGPAIの立ち上げが提唱され、2020年5月のG7 科学技術大臣会合 (米議長) において立ち上げに関するG7の協力に合意。同年6月15日に設立。
- 構成 : 4つのWGを設置。①責任あるAI、②データガバナンス、③仕事の未来、④イノベーションと商業化
- 議長国 : カナダ(2020~21), フランス(21~22), 日本(22~23), インド(23~24)

GPAI 2.0

- **統合パートナーシップ** : 2024年GPAIは**OECDのAIコミュニティ (AIGO)と統合**され、新たにGPAI2.0のコンセプトの下で、OECDのプロジェクトとして機能強化が図られた。
これによりOECD事務局が正式にGPAIの事務局を務めることとなり、また財政基盤が強化された。
- 構成国 : OECDのコミュニティとの統合により、GPAIの**参加国は44 + EU**に増加した。(2025年3月時点)
(日本、アルゼンチン、オーストラリア、オーストリア、ベルギー、ブラジル、カナダ、チリ、コロンビア、コスタリカ、チェコ、デンマーク、エストニア、フィンランド、フランス、ドイツ、ギリシャ、ハンガリー、アイスランド、インド、アイルランド、イスラエル、イタリア、韓国、ラトビア、リトアニア、ルクセンブルク、メキシコ、オランダ、ニュージーランド、ノルウェー、ポーランド、ポルトガル、セネガル、セルビア、シンガポール、スロバキア、スロヴェニア、スペイン、スウェーデン、スイス、トルコ、英国、米国、EU) (下線は当初メンバー国)
- 議長国 : 2024年末にインドから**セルビア**に引き継がれた

名称 G7群馬高崎デジタル・技術大臣会合

開催日程、場所 2023年4月29～30日
群馬県高崎市

G7デジタル・技術大臣会合 主な成果

以下、6つのテーマについて議論が行われ、成果として、

「G7デジタル・技術閣僚宣言」を採択。

- (1) 越境データ流通と信頼性のある自由なデータ流通(DFFT)の推進
- (2) 安全で強靱性のあるデジタルインフラ
- (3) 自由でオープンなインターネットの維持・推進
- (4) 経済社会のイノベーションと新興技術の推進
- (5) 責任あるAIとAIガバナンスの推進**
- (6) デジタル競争

参加国等

- 我が国から河野デジタル大臣、松本総務大臣、西村経済産業大臣が共同議長として参加。
- G7各国（仏、米、英、独、伊、加）、EUに加え、招待国（印、インドネシア、ウクライナ）、国際機関（OECD、ITU、世銀、国連、ERIA）が参加。



責任あるAI とAI ガバナンスの推進

- G7は、民主主義の価値に基づく、信頼できるAIという共通ビジョンを推進するため、国や地域により異なる**AIガバナンスの相互運用性を促進**することの重要性を認識。
- 「AIガバナンスのグローバルな相互運用性を促進等するためのアクションプラン」を承認
- 生成AI技術による機会と課題を早急に把握し、技術の発展に際して、**安全性と信頼性を促進する必要性を認識**。
- 生成AIの急速な伸張の中で、OECDやGPAIなどの国際機関等も活用した**AIガバナンス、知的財産権保護、透明性確保、偽情報への対策**とともに、**責任ある形での生成AIを活用する可能性についてのG7における議論の場を設置**。

【AIガバナンスのグローバルな相互運用性を促進等するためのアクションプラン】

- ◆ 信頼出来るAIのための相互運用可能なツールの役割を支持し、責任あるAIイノベーションのためのオープンで実現可能な環境を構築する役割を認識
- ◆ 国際標準化機関による国際標準開発への多様なステークホルダーの参加を支持
- ◆ 信頼できるAI のためのツールに関するG7 ワークショップやラウンドテーブルを通じた対話の促進
- ◆ OECD、GPAI、UNESCO等を含む国際機関やイニシアティブとの協力を歓迎。
G7やOECD等における生成AIを含むAIに関する今後の機会や課題に関する将来の政策や社会のニーズについて、関連する全てのステークホルダーとの協力を強化
- ◆ 発展途上国や新興国との関与を強化し、世界的にAIイノベーションを実現可能とする環境を支援するとともに、世界中のAIガバナンスの枠組み間の相互運用性を促進するための努力を歓迎

1 広島AIプロセス

- ✓ 2023年の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A Iを実現するためのルール作りを日本が主導。
- ✓ AIの開発・利用について守るべき原則や具体的な行動例を定めた**「国際指針」及び「国際行動規範」**を取りまとめた。



G7広島サミット

国際指針の骨子

【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

2 アウトリーチ活動

- ✓ G7を越えて**開発途上国を含む多くの国との連携強化を図るため**、2024年5月のOECD閣僚理事会の際、49の国・地域の参加を得て、広島AIプロセスの精神に賛同する国々の自発的な枠組である**広島AIプロセス・フレンズグループ**を立ち上げ。現在は、55の国・地域が参加。



OECD閣僚理事会

- 安全、安心、信頼できるAIの実現に向けて、高度なAIシステムの設計、開発、導入、提供及び利用に亘るライフサイクル全体の**全ての関係者**に向けた指針。全ての関係者がそれぞれの立場で責任を有するが、12番目の項目は偽情報の拡散等の生成AIに特徴的なリスクに関するリテラシーの向上や脆弱性検知への協力等、利用者にフォーカス。

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. 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、学术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。

- 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。**
- 以降、G7で議論を進め、企業が履行状況を公開するための質問票の試行版を作成。参加企業からの意見も踏まえ、**「報告枠組み」の基本的な運用方法及び質問票の最終版に各国が合意**(12月23日にイタリア議長国が公表)。
- **2025年2月7日、正式運用開始。**(AIアクションサミットのサイドイベントとして、運用開始イベントをOECD本部で実施。同イベントでは、日米の13社※の企業が本枠組みへの参加の意思を表明している旨が共有された。)

〔※日(7社)：NTT、NEC、KDDI、ソフトバンク、楽天グループ、富士通、プリファードネットワークス
米(6社)：アマゾン、グーグル、マイクロソフト、オープンAI、アンソロピック、セールスフォース〕

「報告枠組み」の概要

(ロゴデザイン)



※イタリアのオリーブの木と日本の桜をイメージしたデザイン

- 「国際行動規範」に沿って作成されたAI開発企業への質問票をOECDのウェブサイト上(<https://transparency.oecd.ai/>)で公開し、回答を依頼。
- **質問票に回答した企業は、OECDウェブサイト上で一覧化(リスト化)されるとともに、回答内容も全て公開。**

同ウェブサイト上には2023年イタリア議長国下で作成のロゴも併せて掲示。

- 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企業等
- 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度の頻度で更新**すること、③**全ての質問に回答**すること等が求められる。

- 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？

- 2024年、OECD閣僚理事会の場で、岸田総理（当時）が立ち上げを宣言。広島AIプロセスの精神に賛同する49の国・地域が参加。
- 2025年3月現在、**55か国・地域**が広島AIプロセス・フレンズグループに参加。



Argentina



Australia



Austria



Belgium



Brunei



Bulgaria



Cambodia



Canada



Chile



Colombia



Costa Rica



Croatia



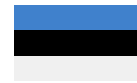
Cyprus



Czech



Denmark



Estonia



Finland



France



Germany



Greece



Hungary



Iceland



India



Ireland



Israel



Italy



Japan



Kenya



South Korea



Laos



Latvia



Lithuania



Luxembourg



Malta



Mexico



Netherlands



New Zealand



Nigeria



Norway



Poland



Portugal



Romania



Serbia



Singapore



Slovakia



Slovenia



Spain



Sweden



Thailand



Türkiye



UAE



United Kingdom



United States



Viet Nam



European Union

2025年2月27日～28日、東京で初の対面会合となる**広島AIプロセス・フレンズグループ会合**を開催。G7を含むフレンズグループ参加国約40か国代表や国際機関のほか日米のAI企業等が参加。

概要

- G7を含む約40の国・地域の高級実務者（局長級）や国際機関等が参加。2024年5月のフレンズグループ立ち上げ後の**初の対面会合**。
- 開会式では、**石破総理大臣からビデオメッセージにより挨拶**。2日間の会合で、総務省の阿達副大臣、今川総審に加え、城内内閣府大臣、平デジタル大臣※、生稲外務大臣政務官他が参加。※ ビデオメッセージ
- 松尾東京大学教授や原山GPAI東京センター長が講演を行った他、OECD、世界銀行及びJICAがプレゼンを実施。参加者間で、AIの機会とリスク、AIガバナンスの在り方、国際連携等について議論。

会合の意義・狙い

- **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。
- 開発途上国に対しては、**人材育成（キャパシティビルディング）の機会**として、AI政策や広島AIプロセス等に対する理解を促進。
- **「報告枠組み」への各国AI開発者の参加を促進**し、広島AIプロセスの**「国際行動規範」等の実効性を確保**する。

会合結果

- 参加各国で**フレンズグループを通じた、国際連携の強化の重要性が再確認**された。フレンズグループの活動を含む国際的なAIガバナンスにおける**日本の貢献・リーダーシップに参加各国より強い支持・賛意が示された**。
- 民間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である**「広島AIプロセス・フレンズグループ パートナースコミュニティ」を立ち上げ**。マルチステークホルダーによる**協力体制を強化し、広島AIプロセスを一層促進**。



開会式の様子

- 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. 参加組織

➤ 2月28日の立ち上げ時においては、以下の16組織が参加を表明(今後も拡大予定)。

**Amazon、富士通、Google、KDDI、Microsoft、NEC、NTT、OpenAI、Preferred Networks、SaferAI、Salesforce、ソフトバンク、楽天グループ
経済協力開発機構(OECD)、国際協力機構(JICA)、世界銀行**

- 2月10日～11日、パリにてAIアクション・サミットが開催。2023年11月の英主催AI安全性サミット、2024年5月の韓・英共催AIソウル・サミットに続く、3回目のAIサミット。次回はインドが開催を表明。
- マクロン仏大統領及びモディ印首相の共同議長の下、**G7、中、伯等の約90の各国政府代表、国連・EU等の国際機関代表、主要企業の代表等**が参加。日本からは松本外務政務官、今川総審等が参加。
- 11日の首脳セッションは各国首脳等が登壇し、**ヴァンス米副大統領**の演説が注目を集めた。サミットの焦点は、従前のAIの「safety」から、「opportunity」(イノベーション、研究開発、投資、公益等)へシフトした印象。

首脳セッションの登壇者

- 【冒頭スピーチ】 マクロン・仏大統領(閉会スピーチを含む)、モディ・印首相
 - 【基調スピーチ】 ヴァンス・米副大統領、フォン・デア・ライエン・EU委員長、グテーレス・国連事務総長
 - 【パネリスト】 ファキ・アフリカ連合委員長、ピチャイGoogle社CEO、ブヴロー・仏大統領特使、オルポ・フィンランド首相、ビロル・IEA事務局長、スヨン・NAVER社CEO、ボグダン・マーティン・ITU事務総局長
- ※ この他、トルドー・加首相、張・中国副首相など、計31カ国の首脳(正副大統領・首相)が出席。また、12の国際機関の長、企業トップ(アルトマン・OpenAI CEO、伊藤・サカナAI CEO、オードリー・タン氏等)が出席。



成果文書の概要

- 【サミット声明】
 - 「人類と地球のための包摂的で持続可能なAIに関する声明」を、日、仏、独、伊、加、印、中、EU等、62の国と地域の支持により採択。
- 【その他】
 - 高度なAIの安全性に関する国際学術報告※を公表。
 ※英主催AI安全性サミットで策定が決定されたもので、事務局は英AISi。内容の検討には、G7、中印、アフリカ諸国、国連等の専門家が参加。今後も、同様の年次報告書を策定予定。
 - 仏主導で、公益、雇用、持続可能性等の観点からのAIに関する文書等を公表。

サミット声明概要

- 人権・倫理に適い、**安全、安心で信頼できるAI**の確保
- **広島AIプロセス**や国連機関等の既存の多国間の取組を踏まえる
- SDGs等への貢献による公益の追求
- AIに関するイノベーションの促進
- 多様な利害関係者が関与する包摂的アプローチ
- デジタル格差の縮小のため、途上国支援の必要性を強調

J・D・ヴァンス米副大統領

- AIの安全性ではなく、機会を議論したい。
- 米の半導体を用いた最強のAIシステムが米で構築されるようにする。
- AI分野の規制緩和を進める。28年までのAI投資推定7000億ドルの半分超が米で投資される。AIの未来は、安全性の規制ではなく、安定的な発電所や半導体製造施設の構築により実現される。
- 米で開発されたAIが言論の自由を制限しないことを保証。一部の権威主義体制によるAIの盗用、監視能力の強化、外国データの取得、他国の国家安全保障を脅かすプロパガンダの作成等の取組を阻止する。
- AI政策で米の労働者を常に中心に据える。

フォン・デア・ライエン欧州委員会委員長

- 欧州が米や中国に遅れを取っていると耳にするが、AI競争は始まったばかり。
- 欧州には独自のアプローチが必要。異なる国・分野から才能を結集し、オープンソースの力を活用。さらなるイノベーション加速に取り組む。
- AI法の目的はEU全体で一つの安全規則を提供すること。安全はビジネスの利益にもつながる。
- 計算能力には大規模な資本が必要。総額2000億ユーロの欧州におけるAI投資動員を目指す。
- 誰もが恩恵を受ける善のためのAI実現が欧州の道。



1 AI法における汎用AI規定

◆ 2024年8月1日に発効したEU・AI法は、汎用AIモデル※の提供者に対し、技術文書・学習データ等に係る情報提供等の義務を規定。汎用AIがシステミックリスク※※をもたらす可能性が認められる場合にはリスク管理と重大事故の監視等のより厳しい義務を規定している。こうした汎用AI規定については、2025年8月から適用予定。

※ 生成AI、大規模なデータを学習したAI等、汎用性の高いAIモデル
※※ 公衆衛生、安全性、治安、基本権や社会全体へ大規模な悪影響を及ぼし得るリスクを指す

2 汎用AIに関する行動規範

◆ 欧州委員会下のAIオフィスは、2025年5月までにAI法の下での汎用AI提供者の義務履行の指針となる行動規範を策定予定。これを踏まえ、欧州委員会は2025年8月の汎用AI規定適用開始までに実施規則を準備予定。
◆ 本規範には、著作権、安全性とセキュリティ、透明性に関する各種措置等が記載される予定。

3 行動規範策定プロセス

◆ AIオフィスは、汎用AI関連企業、関連業界・市民団体、研究者等の利害関係者の参加を得た行動規範会議(全体会合及び作業部会)を複数回開催予定。2024年9月にキックオフ会議が開催。同11月に第一草案、同12月に第二草案を公表。2025年3月、行動規範会議及びAIアクションサミット等での議論を反映した第三草案が公表。
◆ **本会議参加者は、AIオフィスから事前共有される草案に意見提出可能。**最終案は2025年5月公開予定。

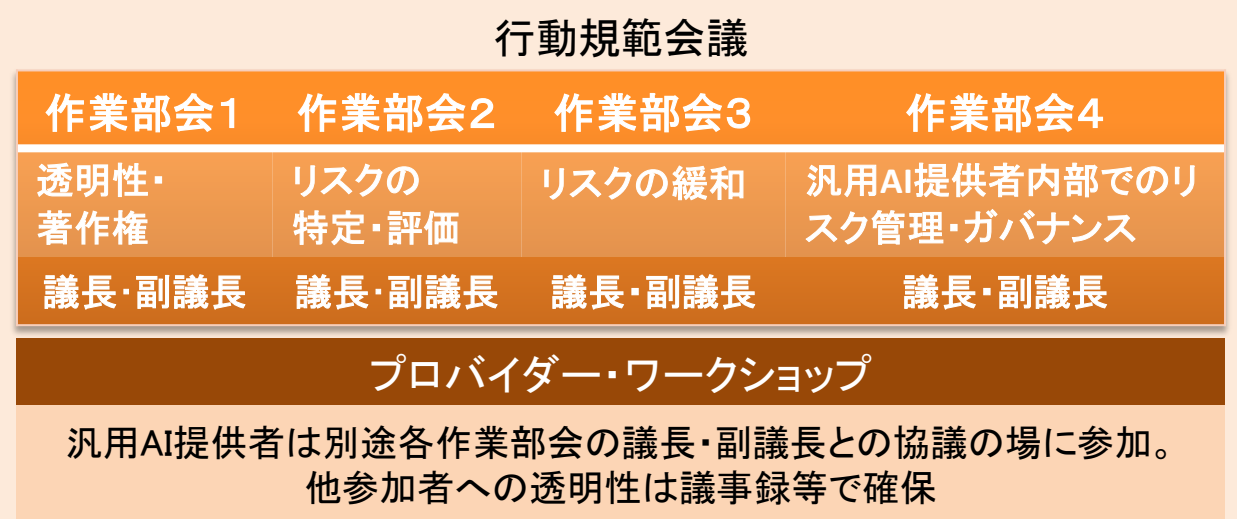
4 行動規範策定会議の構成

会議参加者公募
2024年8月に実施済。約1000団体・個人の参加表明あり。

各作業部会の構成員となる
※日本からも有識者・企業が参加

パブコメ公募
2024年7月-9月に実施済。約430の意見あり。

ドラフトに反映される



「人工知能（AI）、人権、民主主義、法の支配に関する枠組条約」について

欧州評議会（Council of Europe（CoE））於 仏ストラスブール（1949年～）

- 加盟国：46カ国（EU27カ国及びEU外の欧州諸国） + オブザーバー5カ国（日本、バチカン、米国、カナダ、メキシコ）
- 目的：**欧州人権条約等に基づく人権、民主主義、法の支配の保護・推進**
- 200超の多数国間条約（欧州域外国も締結可能）を作成。
 - 個人データ保護条約（Convention 108：1981）：欧州、中南米等、55か国締結
 - サイバー犯罪条約（Budapest Convention：2001）：G7等、76か国締結（日本：2012）



「人工知能（AI）、人権、民主主義、法の支配に関する枠組条約」

- 交渉参加国・地域：57か国 + EU（欧州評議会加盟46カ国 + オブザーバー5カ国、オーストラリア、イスラエル、中南米4か国、EU）
- 基本的内容
 - **AIシステムのライフサイクルにおける活動が人権、民主主義、法の支配に合致することを確保**
（例）
 - ・ AIシステムのライフサイクルにおける活動に関する原則：人間の尊厳、透明性、説明責任、平等・無差別、プライバシー・データ保護、信頼性、安全性を確保したイノベーション推進等
 - ・ リスク・影響評価：人権、民主主義、法の支配に関するリスク・影響の特定、評価、予防、緩和のための措置をリスクベースアプローチに基づき採用/維持。

2024年

- 3月11日～14日に行われた交渉会合にて**交渉妥結**
- 5月17日、閣僚委員会にて**採択** ※11月4日、モンテネグロの署名で10か国に拡大。
- 9月5日、欧州評議会司法大臣会合にて**署名開放**。同日、米英等9か国※とEU（27か国を代表）が署名

2025年

- 2月11日、**日本**とカナダがAIアクションサミット（@パリ）の際に**署名**。現在、**G7を含む13か国**※とEUが署名済。
※2月27日、リヒテンシュタインが署名。

※ CoE加盟3か国を含む5か国の批准により条約発効。その後、1年以内に締約国会議の規則を策定。

- **条約履行手段**：締約国は本条約の履行のため適切な立法上、行政上又はその他の措置を採用/維持。（1条2）
- **AIシステムの定義**：OECDによる改訂版AI原則における定義を採用。（2条）
AIシステム：明示的な又は黙示的な目的のために、受け取った入力から、物理的環境あるいは仮想環境に影響を及ぼす予測、推薦又は意思決定等の出力の方法を推論する機械ベースのシステム。各AIシステムは、導入後の自律及び適応の程度が異なる。
- **適用範囲**：公的部門への適用。民間部門のAIシステムのライフサイクルから生ずるリスクや影響に対処。民間への対処方法（条約の適用、または、それ以外の適切な措置）につき、署名/批准時に宣言。国防は対象外。（3条）
- **一般的義務**：人権保護、民主主義の健全性、法の支配の尊重に関する義務。（4条、5条）
- **AIシステムのライフサイクルにおける活動に関する原則**：人間の尊厳、透明性、説明責任、平等・無差別、プライバシー・データ保護、信頼性、安全性を確保したイノベーション推進等。（6条–13条）
- **救済**：人権侵害に対する実効的救済の確保。人権に悪影響を及ぼしうるAIシステムについての関連情報を文書化、当該情報へのアクセスを許可された団体に提供、適切・可能な場合、被害者に通知。権限ある当局への申立の確保（14条）
- **手続的セーフガード**：（救済に関する）効果的な手続の確保。適切な場合、AIシステムが対応していることを通知。（15条）
- **リスク・影響評価**：締約国は、AIシステム・ライフサイクルから生じる人権、民主主義、法の支配に関するリスク・影響の特定、評価、予防、緩和のための措置をリスクベースアプローチに基づき採用/維持。（16条）
- **一時的な停止又は禁止**：AIシステムの特定の使用が人権、民主主義、法の支配と合致しない場合、一時的な停止又は禁止等の措置の必要性を判断。（16条4）
- **報告**：公的部門における措置及び民間部門のAIシステムのライフサイクルから生ずるリスクや影響への対処について、締結から2年後、その後は定期的に報告。（24条）
- **監督枠組**：履行の監督のため、独立・公平で、適切な権限・リソースを備えた枠組みを設置/指定。（26条）

CoE46カ国

EU (加盟国を代表)

ドイツ	ポーランド	英国
ベルギー	ポルトガル	スイス
チェコ	スロバキア	ルウェー
デンマーク	スロベニア	アイスランド
エストニア	スペイン	トルコ
フィンランド	スウェーデン	
フランス	オーストリア	
リトアニア		
ルーマニア	ブルガリア	
マルタ	計27か国	
セルビア	ボスニア =	モナコ
北マケドニア	ヘルツェゴビナ	サンマリノ
ウクライナ	アルメニア	リヒテン
ジョージア	アンドラ	シュタイン
モンテネグロ		
アルバニア		
モルドバ		
アゼルバイジャン		

条約交渉参加57カ国

CoE オブザーバー

日本
米国
カナダ
メキシコ

バチカン

CoE・AI委 オブザーバー

イスラエル
コスタリカ
豪州

アルゼンチン
ペルー
ウルグアイ

カメルーン

OECD 38カ国

韓国
NZ
チリ
コロンビア

※カメルーンは交渉妥結後の2024年11月AI委に参加。
※赤字は署名済13か国 + EU
(25年3月7日現在)