

安全・安心なAI利活用環境整備について

～AI事業者ガイドラインのご紹介～

2024年9月10日

総務省 情報流通行政局 参事官補佐

小倉 知洋

- 1. AIとは**
- 2. AIに関する政府全体の取組**
- 3. AIガバナンスの主な動向**
- 4. AI事業者ガイドライン**
- 5. その他の関連施策等**

目次

1. AIとは

2. AIに関する政府全体の取組

3. AIガバナンスの主な動向

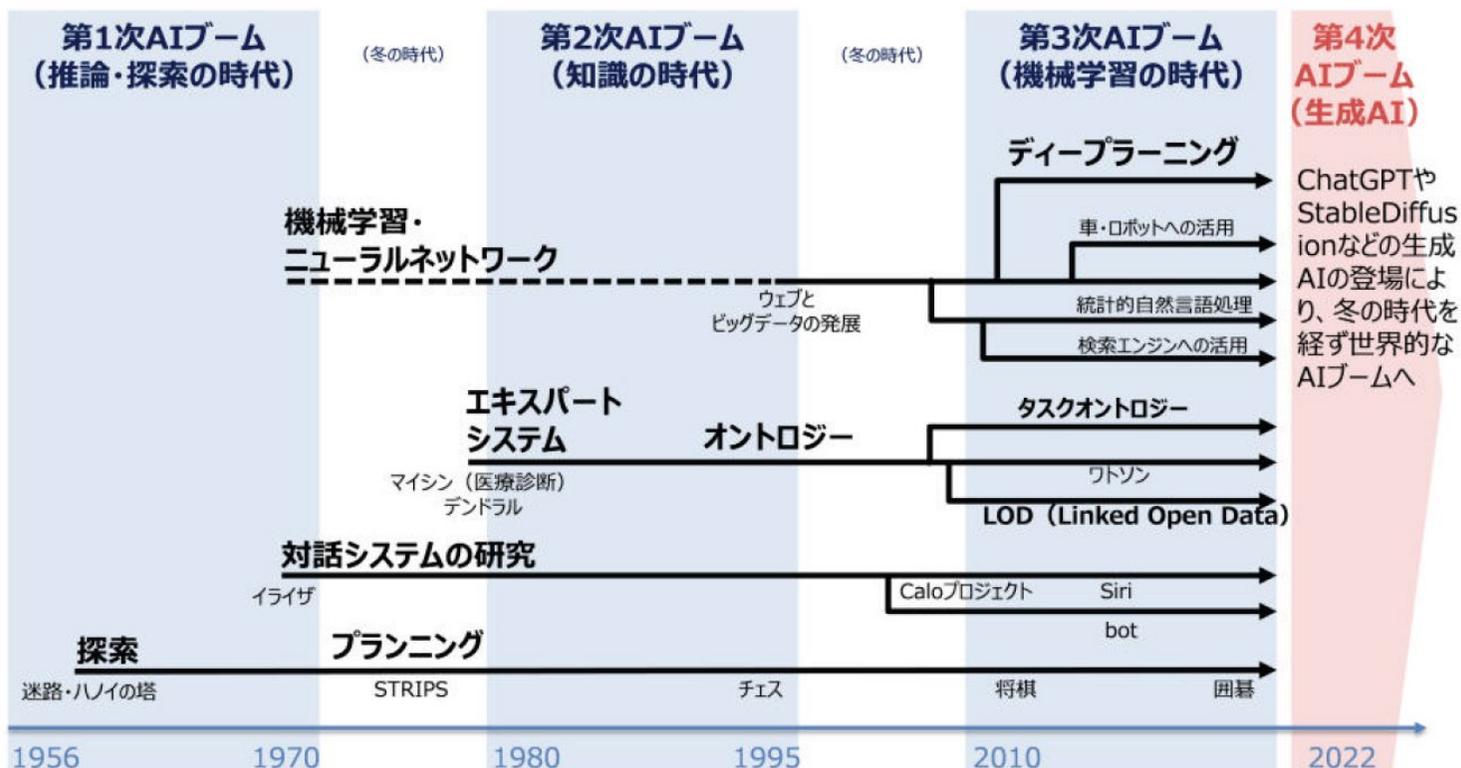
4. AI事業者ガイドライン

5. その他の関連施策等

AIの技術的進展の歴史

- ◆ AI（人工知能）の歴史は1950年代から始まり、何度かブームと冬の時代を繰り返してきた。
- ◆ 推論・探索から始まった第1次AIブームは、音声認識等が組み込まれた第2次AIブームを経て、第3次AIブームとしてディープラーニング（深層学習）をはじめとした革新的な技術が登場し、社会で実用され得るAIが開発されて社会に浸透
- ◆ 2022年頃からの生成AIの急速な普及により、現在は第4次AIブームに入ったとも言われている

人工知能・ビッグデータ技術の俯瞰図



AIとは

AIは、大量のデータから、故障の予兆や詐欺やスパム情報の検出、将来予測、また次に個人に表示する動画を決定している。このようにデータ分析目的で活用されるAIは「**Analytical AI**」と呼ばれており、既に社会の多くの場所で用いられている。

これに加えて、近年、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**の技術が急速に発展してきた。

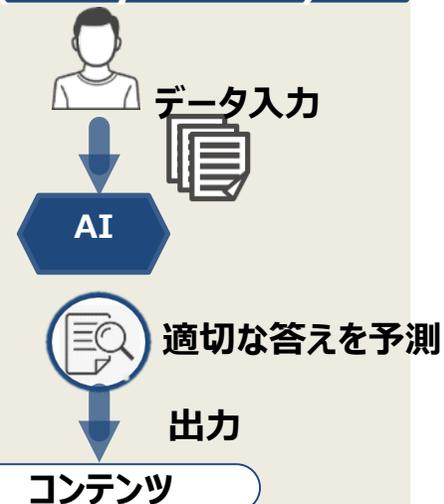
※総務省 令和5年版情報通信白書 より

AI：現時点で確立された定義はなく、広義の人工知能の外延を厳密に定義することは困難である。本ガイドラインにおけるAIは「AIシステム（以下に定義）」自体又は機械学習をするソフトウェア若しくはプログラムを含む抽象的な概念とする。

AIシステム：活用の過程を通じて様々なレベルの自律性をもって動作し学習する機能を有するソフトウェアを要素として含むシステムとする（機械、ロボット、クラウドシステム等）

※総務省、経済産業省 AI事業者ガイドライン より

従来型AI (Analytical AI)



生成AI



AIシステムは、明示的又は黙示的な目的のために、実環境又は仮想環境に影響を及ぼす、予測、コンテンツ、推奨、又は決定などの出力をどのように生成するか、受け取った入力内容から推論する機械ベースのシステムである。デプロイ後の自律性と順応性のレベルは、AIシステムによってそれぞれ異なる。

※OECD AI原則(2024年5月改定) 仮訳

出典) <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449>

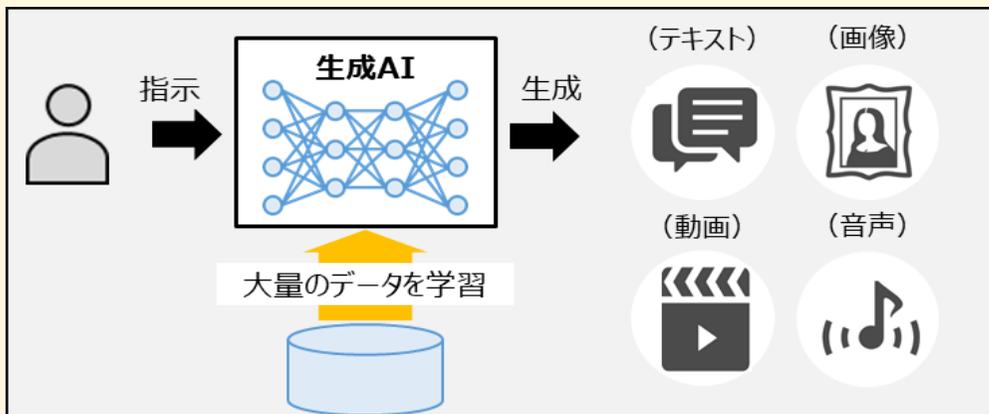
原文

AI system: An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.

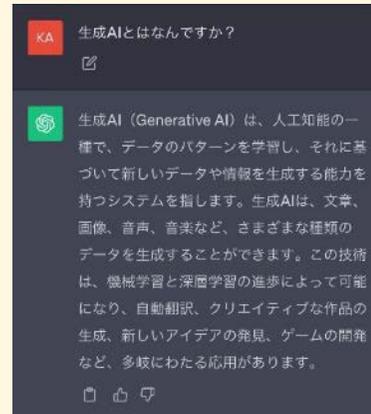
生成AIのメリットとリスク

「生成AI」は、大量の学習データから、文章、画像、動画、音声、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、固有のリスクも指摘

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ブームの背景

◆ 複数の要因

- ◆ 自然言語処理や画像生成などのタスクにおけるモデルの精度が飛躍的に向上
- ◆ 様々なタスクに適用可能な知識を獲得したモデルの登場により、新たなタスクに対応するためにモデルを再トレーニングする必要がなくなり、開発や利用が大幅に容易化されるとともに、AIがより複雑なタスクをこなせるようになり、その有用性が広く認知
- ◆ AIが人間と協調して働く環境が整い、多くの業界でAIの導入が促進

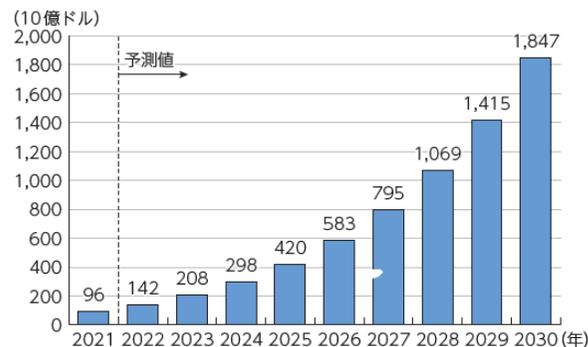
生成AIブームにある技術的要因

要因	解説
大規模言語モデル・基盤モデルの登場	人間の言語を理解し生成する能力を持つモデル。大量のテキストデータから学習し、自然なテキスト生成が可能。
オープンソース化	ソースコードが公開されており、誰でも無料でアクセス、使用、改良が可能。技術の普及とイノベーションを促進。
ユーザーインターフェース（UI）	直感的で使いやすいインターフェースを提供し、非技術者でもAIツールを容易に操作できるように設計されている。
APIによる提供	プログラミングインターフェースを通じてAI機能を他のアプリケーションに組み込みやすくする。多様な開発が可能。
高い汎用性・マルチモーダル機能	テキストだけでなく、画像や音声など複数のモードを扱う能力。様々なタイプのデータを処理・生成できる。

(参考) AI市場規模

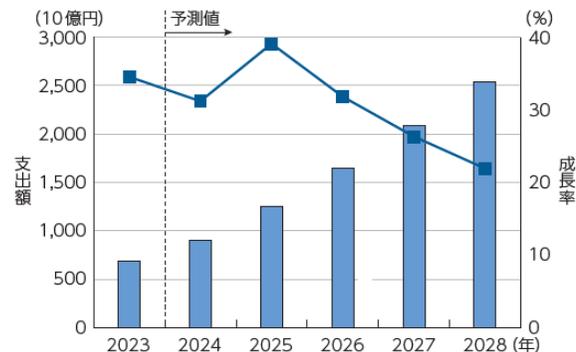
- ◆ 世界のAI市場規模（売上高）は、2022年には前年比78.4%増の18兆7,148億円まで成長すると見込まれており、その後も2030年まで加速度的成長が予測。
- ◆ 日本のAIシステム市場規模（支出額）は、2023年に6,858億7,300万円（前年比34.5%増）となっており、今後も成長を続け、2028年には2兆5,433億6,200万円まで拡大すると予測。
- ◆ 2023年の670億ドルから2032年には1兆3,040億ドルと大幅な拡大を予測。

世界のAI市場規模（売上高）の推移及び予測



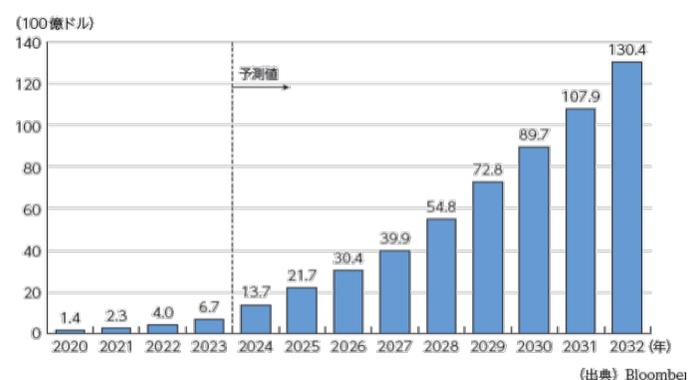
(出典) Next Move Strategy Consulting (Statistaより引用)

国内AIシステムの市場規模（支出額）及び予測



(出典) IDC 「2024年 国内AIシステム市場予測を発表」(2024年4月25日)

世界の生成AI市場規模の推移及び予測



(出典) Bloomberg

目次

1. AIとは

2. AIに関する政府全体の取組

3. AIガバナンスの主な動向

4. AI事業者ガイドライン

5. その他の関連施策等

政府全体でのAIに関する取組の推進

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、Society 5.0に寄与
! 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

G7広島サミット2023において合意された共通のビジョンと目標
「我々が共有する民主的価値に沿った、信頼できるAI」
閣僚級の議論を続け、年内に報告（広島AIプロセス）

AI戦略会議（有識者） + AI戦略チーム（関係省庁）等による議論

「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）

基本的な考え方

- ① 国際的なルール構築に向けた主導的役割の発揮（広島AIプロセスに貢献）
- ② リスクへの対応と利用
- ③ 多様な関係者を巻き込んだ迅速かつ柔軟な対応

国際的な議論とリスクへの対応

- ・「広島AIプロセス」など、国際的議論を主導
- ・生成AIに関する懸念やリスクへの対応（ガイドライン遵守、新技術の開発・普及、国際的な議論の動向も踏まえ、必要な対応を検討等）

AIの最適な利用

- ・データ連携基盤の構築など、事業環境整備
- ・政府機関における生成AIの試験的な利用
- ・幅広い世代のスキル・リテラシー教育の充実

AI開発力の強化

- ・生成AIに関する基盤的な研究力・開発力の醸成
- ・公的機関の保有データを開発で利用
- ・世界からトップ人材が集まる環境の構築
- ・スタートアップ施策の強力な推進

AI戦略会議 と AI戦略チーム を軸とした各省庁の協力による政策の推進

- ✓令和4年11月、大規模言語モデル「GPT 3」をベースにした対話型言語モデル「Chat GPT」や、テキストを入力すると画像を生成する「プロンプト型画像生成 AI」、テキストから作曲を行うAIなど、多様な用途でのAIが登場。
- ✓国内では、**AI戦略チーム**（村井英樹総理補佐官（当時）の下、関係省庁の実務者級（審議官・課長級）のチームを編成）、**AI戦略会議**（座長：松尾 豊（東京大学大学院工学系研究科教授））において、我が国における生成AIをめぐる課題について検討を行い、令和5年5月26日、「**AIに関する暫定的な論点整理**」を取りまとめ

基本的な考え方

- 国際的なルール構築に向けた主導的役割の発揮
 - AI に国境はなく、国際的な共通理解・ルールづくり・相互運用性が重要。我が国は、「広島 AI プロセス」などを通じ、議論をリードすべき
- リスクへの対応と利用
 - 生成 AI の開発・提供・利用を促進するためにも、生成 AI の懸念やリスクへの適切な対応を行うべき。いわば、「ガードレール」の設置が必要。
- 多様な関係者を巻き込んだ迅速かつ柔軟な対応
 - 政府は、広島 AI プロセスなどの検討スケジュールも踏まえつつ、マルチステークホルダーを巻き込んだ、迅速かつアジャイルの対応が求められる。

AI の利用

- 生成 AI は、デジタル化を加速させ、我が国全体の生産性向上のみならず、様々な社会課題解決に資する可能性がある。
- AI 利用を加速するため、医療や介護・行政・教育・金融・製造等のデータ連携基盤の構築・DFFT 構想の具体化・人材育成・スタートアップの事業環境整備を進めるべき。
- 政府機関が一体となって、機密情報漏洩のリスクなどに配慮しつつ、率先して生成 AI の利用可能性を追求することが重要。
- 幅広い世代が生成 AI の恩恵を享受できるよう、スキル・リテラシーを身に付けることが大切。

リスクへの対応

- リスク対応の基本方針
- 透明性と信頼性
 - AI 開発者・提供者には、現行法令やガイドラインに則り、積極的な情報開示を求めたい。
 - 政府は、主要な AI 開発者・提供者に対して、透明性・信頼性の確保を直接働きかけることも検討すべき。
 - **生成 AI の普及を踏まえ、既存のガイドラインに関して、必要な改定などを検討する必要がある。**その際、諸外国における検討とも協調し、第三者認証制度や監査制度等も参考とすべき。
 - 顕在化したリスクを低減するような技術の研究開発・普及を奨励することも望ましい。
- 懸念されるリスクの具体例と対応

AI 開発力

- 生成 AI に関する基盤的な研究力・開発力を国内に醸成することが重要。政府は、AI 開発におけるインフラとも言うべき、計算資源とデータの整備を行うことが最も重要。
- 計算資源を活用するための電力調達が課題。地方のデータセンターの活用を含め、電力を有効活用する方策の検討が必要。
- 開発に用いることのできる日本語を中心とするデータの整備・拡充を進めるべき。
- また、生成 AI 自体の開発は、市場原理を最大限尊重した、民間活力を十分活用した従来型ではない開発促進策が期待される。世界からトップ人材が集まる研究環境の構築も期待される。

AI戦略会議の概要

- 構成員**
- 江間 有沙 東京大学国際高等研究所東京カレッジ 准教授
 - 岡田 淳 森・濱田松本法律事務所 弁護士
 - 川原 圭博 東京大学大学院工学系研究科 教授
 - 北野 宏明 株式会社ソニーリサーチ 代表取締役CEO
 - 佐渡島庸平 株式会社コルク 代表取締役社長
 - 田中 邦裕 さくらインターネット株式会社 代表取締役社長
 - 松尾 豊 東京大学大学院工学系研究科 教授 **【座長】**
 - 山口 真一 国際大学グローバル・コミュニケーション・センター 准教授

※このほか、村井英樹内閣官房副長官、関係省庁（内閣府、デジタル庁、総務省、文部科学省、経済産業省）の閣僚が参加

開催日	議事
第1回（令和5年5月11日）	1. AIを巡る主な論点について
第2回（令和5年5月26日）	1. 「AIに関する暫定的な論点整理」について
第3回（令和5年6月26日）	1. AIに関する取組状況について 2. 広島AIプロセスにおける今後の対応に関する討議
第4回（令和5年8月4日）	1. 広島AIプロセスの今後の進め方 2. AI開発力の強化
第5回（令和5年9月8日）	1. 広島 AIプロセスの報告と統合ガイドライン 2. AI開発力の強化の報告 3. その他の報告事項
第6回（令和5年11月7日）	1. 広島AIプロセス及びAI事業者ガイドライン 2. AI学習データの提供促進 3. 人材育成 4. 経済対策におけるAI施策
第7回（令和5年12月21日）	1. 広島AIプロセス及びAI事業者ガイドライン (1)広島AIプロセスの報告 (2)AI事業者ガイドラインの報告 2. 来年のAI戦略会議の課題について
第8回（令和6年4月19日）	1. AI事業者ガイドラインの報告 ※持ち回り開催
第9回（令和6年5月22日）	1. AI戦略の課題と対応について
第10回（令和6年7月19日）	1. AI制度研究会の設置について ※持ち回り開催
第11回（令和6年8月2日）	1. AI政策の現状と制度課題について ※AI制度研究会と合同開催

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】

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と知的財産権との関係】

文化審議会著作権分科会法制度小委員会「AIと著作権に関する考え方について」（文化庁、2024.3.15）

AIと著作権の考え方について整理。周知・啓発、関係者の相互理解の促進等。

AI時代の知的財産検討会 中間とりまとめ案（知的財産戦略推進事務局、2024.4.22）
法、技術、契約による取組み（5月に中間とりまとめ予定）。

【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】

ガイドライン等履行確保・AI利用促進に関する調査（内閣府、2023.11～）

各業種等におけるリスクや対応等を整理。

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事業者ガイドライン（第1.0版）（総務省・経産省、AI戦略会議報告、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

【政府における利用】

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】

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、2024年4月、AI for Scienceの日米連携枠組みを創設。

AIのイノベーションとAIによるイノベーション

人材の育成・確保や計算資源等のインフラの高度化とともに、AI利活用の推進と研究開発力の強化を一体的に官民で推進。データ政策、スタートアップ政策とも連携。

研究開発力の強化

データの整備

- ・ 質の高い日本語データを整備し、適切な形で提供。各分野において競争力あるAIを開発するためデータ共有などを促進するとともに、好事例を共有。

計算資源の確保

- ・ AI開発に不可欠な計算資源を引き続き官民で整備。



新たなモデルの開発

- ・ モデルの高効率化や高精度化、マルチモーダル化、リスク低減化等の新たな研究開発を産学連携で推進。
- ・ 革新的な技術を有するスタートアップを支援。

日本の強みを活かす

- ・ 医療、創薬、マテリアル等の分野で日本の強みを活かし、AI for Scienceを加速。
- ・ 「富岳」の次世代となる優れたAI性能を有する新たなシステムの開発・整備に着手。
- ・ 労働力不足の解消やGX等にも資する革新的なAIロボット等の研究開発・実装を推進。



AI利活用の推進



政府

- ・ 他機関のモデルにもなるよう、政府によるAIの適切な調達・利用、得られた知見の共有を推進。

重要分野

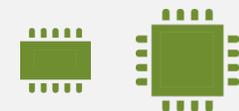
- ・ 医療・ロボットなどの重要分野におけるAI導入を促進。

制度運用の明確化

- ・ ユーザーや開発者が委縮することなくAI利活用・開発を進められるよう、個人情報保護法・著作権法・各種業法など留意すべき制度の運用を明確化。

インフラの高度化

- ・ データセンターの大規模化・分散化と省電力化、Beyond 5G等のネットワークシステムの高度化等（研究開発の促進）。AIに不可欠なインフラへの民間投資の拡大。
- ・ AI半導体等のキーデバイスの設計・開発・運用に関する産学連携体制等の構築や研究開発、人材育成を支援。
- ・ 安価な脱炭素電力の量的確保のための環境整備を検討。



人材の育成・確保

- ・ AIスキルの習得、AIリテラシー向上のための教育コンテンツの充実・普及啓発。
- ・ 初等中等教育段階における、情報モラルを含めたAIの利活用に関するパイロット的な取組の一層の推進。
- ・ 次世代のAI開発等を担う若手研究者や博士後期課程学生を支援。



イノベーション推進のためにもガードレールは必要。変化に迅速かつ柔軟に対応するため、日本は「AI事業者ガイドライン」に基づく事業者等の自発的な取組を基本。今後、AIに関する様々なリスクや、ソフトロー（規格、ガイドライン等）とハードロー（法律・基準等）に関する国際的な動向等も踏まえ、制度の在り方について検討。

自発的ガバナンスと制度の検討

ガイドライン

- ・ 幅広い業種に「AI事業者ガイドライン」の周知・浸透を図る。

AI制度

- ・ 学識経験者や事業者、利用者など、幅広くマルチステークホルダーの意見を聴取しつつ、検討。
- ・ 医療・自動運転、金融等の社会への影響が大きい重要分野は、技術の進展や利用状況に応じて制度の見直しの必要性等を検討。



AIの安全性の検討

AISI Japan AI Safety Institute AIセーフティ・インスティテュート

- ・ AIセーフティ・インスティテュートは安全性の中心的機関として、専門人材の確保、技術的知見の集約等を推進。

新たな技術

- ・ 外部知識を利用してハルシネーションを防止する技術などAIの安全性に関する最先端の研究開発を官民が連携して推進。



偽・誤情報への対策

- ・ 生成AIを利用したものを含め、ネット上に流通・拡散する偽・誤情報や、SNS上のなりすまし型偽広告への対応等について、国際的な動向を踏まえつつ、表現の自由をはじめとする様々な権利・利益に配慮しながら、技術・研究開発の推進、ファクトチェック推進、国際的な連携強化など、制度面も含む総合的な対策を推進。
- ・ ネット上に流通するAI生成コンテンツを判別する技術の開発・実証等や、リテラシー向上等に取り組む。

知的財産権等

- ・ 「AI時代の知的財産権検討会」の中間とりまとめ、文化審議会著作権分科会法制度小委員会の「AIと著作権に関する考え方について」を踏まえ、今後の技術発展や海外動向等も見ながら、俳優や声優等の肖像や声も含め引き続き必要な検討を進める。

グローバルなAIガバナンスのルール作り、イノベーション創出を引き続き主導

- ・広島AIプロセス等を通じて、安全、安心で信頼できるAIの実現に向け、国際的な取組を引き続き主導。アジア諸国やグローバルサウスとも協調。



広島AIプロセス フレンズグループを活用しながら、G7外へのアウトリーチを進め、国際指針や行動規範の実践を拡大。



GPAI（AIに関するグローバルパートナーシップ）東京センター等を通じ、広島AIプロセスに資する偽情報対策技術の評価・実証などのプロジェクトベースの取組を支援。



AIセーフティ・インスティテュート（AISII）と諸外国のAISII等の国際的なネットワークを構築し、AIの安全性の確保に向けた方策を検討。



OECD閣僚理事会 生成AIに関するサイドイベントにて、岸田総理から、広島AIプロセス フレンズグループ（49か国・地域が参加）の立上げ、GPAI東京センターの設立を発表。同閣僚理事会では、松本総務大臣がAIパートの議長を務め、AI政策に関する議論をリード。OECD AI原則の改定が採択され、閣僚声明にも広島AIプロセスに対する支持を明記。今後、広島AIプロセスの進展も踏まえつつ、改定されたAI原則がグローバルなAIガバナンスの基礎となるよう、OECDとの連携を更に強化し、引き続きOECDにおける取組をリード。

企業間連携、産学連携、国際連携等



AI開発者同士でのノウハウ共有、グローバルテック企業との交流などコミュニティ活動を官民で推進



有志国・地域と強固に連携した開発体制



大学・国研等のポテンシャルも活用した産学連携の体制

目次

1. AIとは
2. AIに関する政府全体の取組
- 3. AIガバナンスの主な動向**
4. AI事業者ガイドライン
5. その他の関連施策等

AIガバナンスとは

AIガバナンス ※AI事業者ガイドラインでの定義

- 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。

AI事業者ガイドラインでは、AI事業者に対して、AIを安全・安心して利用できるよう**AIガバナンスを設計・運用**することを求めている。

また、AIガバナンスは、各国/地域のAIに対するルール(ガイドラインや法規制など)や社会情勢等により求められるレベル・内容が異なるため、動向把握が重要となる。

- AIはグローバルで活用されるものであるため、AI事業者は自国内のルールだけを遵守するのではなく**各国/地域のルールにも気を配る**必要がある。
- G7の議論では、「**AIガバナンスの枠組み間の相互運用性**に関する国際的な議論の重要性を強調」している

ルールのアプローチは大きく分けて2つある。

ソフトローアプローチ

- ガイドライン
- 自主的コミットメント

相互運用性の推進

ハードローアプローチ

- AI Act

■ 欧米ともに、ソフトロー（標準、ガイドライン等）とハードロー（法令）の組合せを模索



広範なハードローをソフトローで補完

- EU理事会は**AI法案を採択**（2024年5月）

主として人権侵害、差別・偏見リスクを重大リスクと捉え、センシティブな情報を扱うAIは禁止。高リスクAIに安全性評価等を義務化。影響の大きい汎用AIにはテスト等の報告を義務化

- 欧州委員会は通信ネットワーク・コンテンツ・技術総局内に**AIオフィスを設置**（2024年5月）

- **デジタルサービス法（DSA）を採択**（2022年10月）

違法・利用規約違反コンテンツへの対応や透明性・説明責任に関する義務等



ソフトローをベースにしつつ、目的に応じてハードロー

- 大手AI開発企業は、信頼できるAI開発を進める**ボランタリー・コミットメントを発表**（2023年7月）

- **大統領令を発出**（2023年10月）、イノベーション促進とリスク対応を各省庁に指示

既存法令（国防生産法等）を活用し、主として安全保障の観点から、大規模汎用AIモデル等の開発企業に報告義務

- 海外顧客にIaaSを提供する際、身元を確認し政府へ報告を義務付ける**規則案を発表**（2024年1月）

- 選挙の保護やディープフェイクへの対応などを念頭に**複数の州で法規制**



欧州評議会は、AI、人権、民主主義、法の支配に関する枠組み条約を採択（2024年5月）

主として公的機関での人権侵害等への措置や監督体制の義務



国連では、「安全、安心で信頼できるAI」に関する**国連総会決議**を採択（2024年3月）

UNESCO、ITU等においても議論。

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12月の首脳声明において、**広島AIプロセス包括的政策枠組**を承認。

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
2. 全てのAI関係者向けの**広島AIプロセス国際指針**(下表)
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**※
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**

※2024年6月のG7首脳声明で同規範のモニタリングツールとして試験的報告枠組の立ち上げを歓迎

【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

【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

【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

【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

アウトリーチ活動

- ✓ 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・フレンズグループ**を立ち上げ。

広島AIプロセス・フレンズグループ

- 広島AIプロセスの精神に賛同する国々の自発的な枠組
- 53か国・地域※が参加 (2024年8月2日時点)

※G7、EU加盟国、その他のOECD・GPAI加盟国、その他のアジア、アフリカ諸国



OECD閣僚理事会

- ✓ 同年7月、プロジェクト協力を支援するため、情報通信研究機構に、GPAI (AIのための国際パートナーシップ) 東京専門家支援センターを立ち上げ (センター長: 原山優子東北大学名誉教授)

- 安全、安心、信頼できるAIの実現に向けて、AIライフサイクル全体の関係者それぞれが異なる責任を持つという認識の下、12の項目を整理。
- 「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、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、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。

- AIはイノベーション。一方で、様々なリスクがあり、イノベーション促進のためにも、適切なガードレールが必要。
(リスクの例)
製品・サービスの安全性に関するリスク（誤作動など）、人権侵害（プライバシーや公平性など）、安全保障・犯罪増加等のリスク、財産権侵害のリスクなど
- 各国とも、リスクの大きさに応じた対策（リスクベースアプローチ）。
 -  EUは主として人権等の観点からAI全体に関して規律。国際規格も併用。
 -  米国はビッグテックによるボランタリーコミットメントを基本としつつ、大統領令で、既存法令を活用した大規模汎用モデル開発者からの報告を求めるなど。
 -  日本はAI全体に対してAI事業者ガイドラインで迅速に対応。
- 各国はソフトロー（規格・ガイドライン）とハードロー（法律・基準）の組合せを指向。日本においても制度の要否も含め検討は必要。

様々なリスクを想定し、各主体の役割等を検討していく必要がある。一つの考え方として、以下のような整理の仕方がある。

影響大・高リスクのAI開発者

国民の安全・安心の観点からソフトローを補完する法制度の要否の検討。

変化の速さや多様性を踏まえ、規制の運用は官民連携型の第三者機関が担う「**共同規制型**」、「**ゴールベース**」も重要。

影響大・高リスクのAI提供者・利用者

基本的には**業法・規制法**があり、その下で対応。業法等のない**重要インフラ**等では、**技術変化や利用状況に応じて機動的な対応が望まれるが、議論の積上げが必要。**

法令とAI事業者ガイドラインの併用も。

政府

政府による**AIの適切な調達・利用**（他分野への波及効果）。

違法行為へのAI利用など**リスク情報を調査し、悪用される蓋然性の高いAIに対する改善・排除措置を検討。**

プロバイダー

偽・誤情報に関して、**オンラインプラットフォーム等による不適切なコンテンツの削除**など、**情報流通全体の枠組みの中で対応。**

AI制度研究会の構成員、これまでの開催状況

○ 「統合イノベーション戦略2024」（令和6年6月4日閣議決定）に基づき、有識者会議である「AI戦略会議」の下に、「AI制度に関する考え方」等を踏まえ、AI制度の在り方について検討することを目的として、「AI制度研究会」を設置。

○ 構成員

- 座長 松尾 豊 東京大学大学院工学系研究科 教授
- 座長代理 村上 明子 独立行政法人情報処理推進機構AIセーフティ・インスティテュート 所長
- 生貝 直人 一橋大学大学院法学研究科 教授
- 岡田隆太郎 一般社団法人日本ディープレニング協会 専務理事
- 岡本浩一郎 一般社団法人ソフトウェア協会 副会長／株式会社リアルソリューションズ 代表取締役社長
- 柿沼 由佳 公益社団法人全国消費生活相談員協会消費者教育研究所 副所長
- 工藤 郁子 大阪大学社会技術共創研究センター 特任准教授
- 殿村 桂司 長島・大野・常松法律事務所 弁護士
- 中尾 悠里 富士通株式会社富士通研究所人工知能研究所 プリンシパルリサーチャー
- 永沼 美保 一般社団法人日本経済団体連合会デジタルエコノミー推進委員会 国際戦略WG 主査／日本電気株式会社 品質・エンジニアリング推進部門 主席プロフェッショナル
- 原山 優子 東北大学 名誉教授／GPAI東京専門家支援センター長
- 平野 晋 中央大学国際情報学部 教授・学部長
- 福岡真之介 西村あさひ法律事務所・外国法共同事業 弁護士
- 松原実穂子 日本電信電話株式会社 チーフ・サイバーセキュリティ・ストラテジスト

開催日	議事
第1回（令和6年8月2日） ※AI戦略会議（第11回）との 合同会議	1. AIを巡る主な論点について
第2回～	AIのリスクと制度的対応等について、関係者にヒアリング予定

（引き続き検討、今秋に中間とりまとめ予定）

基本的な考え方

- アジャイルでフレキシブルなガイドラインを最大限活用。
リスクに応じて必要な法規制の在り方を検討。
リスクへの対応とイノベーション促進を両立。
- 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。

多様な論点

- 対応すべきリスク、規制対象となる行為
- 規制の手法（事前規制／事後規制）
- 規制の強度
- 守るべき事項、基準 等

政府による適正な調達と利用

他の法令による対応（例）

- 刑法、個人情報保護法、著作権法 等
- 各種の業法、製品安全に関する法令 等

様々な有識者・専門家の意見を聴取し、議論（今秋に中間とりまとめ）

目次

1. AIとは
2. AIに関する政府全体の取組
3. AIガバナンスの主な動向
- 4. AI事業者ガイドライン**
5. その他の関連施策等

- 生成AIに代表されるように、AI関連技術は日々発展をみせ、利用機会と可能性は拡大の一途をたどり、産業におけるイノベーション創出や社会課題の解決に向けても活用されている
- 我が国においては、Society 5.0の実現に向け、AIの高度な活用に対する期待が高まっている
- 我が国は、G7におけるAI開発原則に向けた提案を先駆けとし、G7・G20やOECD等の国際機関での議論をリードし、多くの貢献をしてきた
- このような背景を受け、我が国におけるAIガバナンスの統一的な指針を示すことで、イノベーションの促進及びライフサイクルにわたるリスクの緩和を両立する枠組みを関係者と連携しながら共創していくことを目指す

技術革新

Society 5.0の実現

国際的な議論

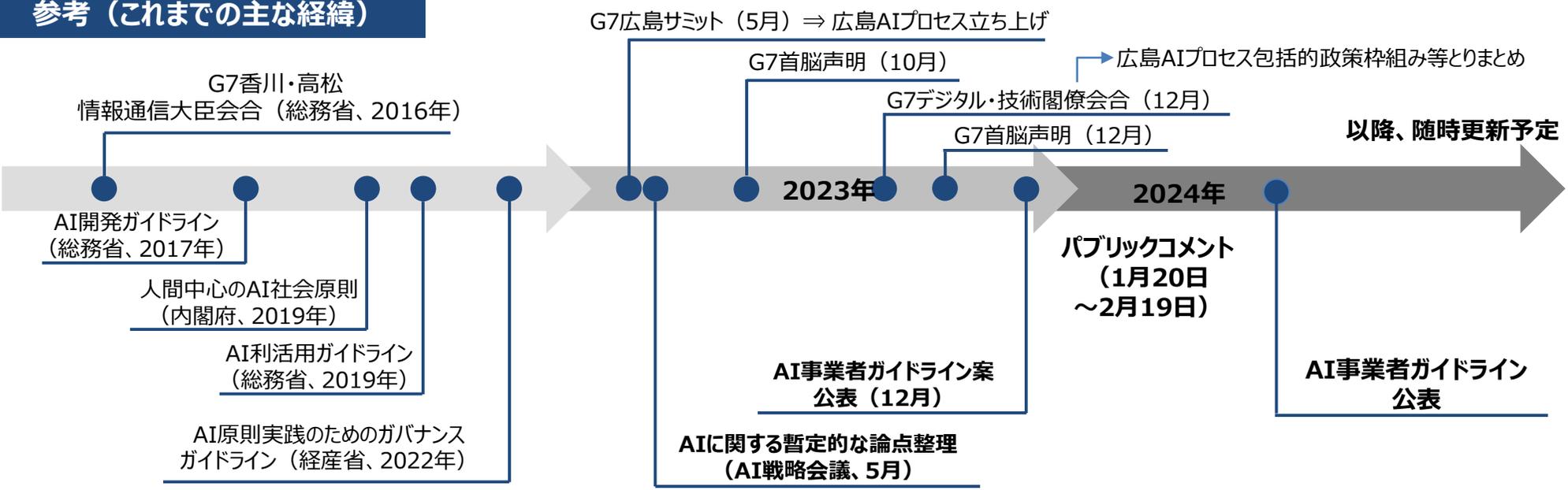
「AI事業者ガイドライン」を策定

AIに関係する者が、国際的な動向及びステークホルダーの懸念を踏まえたAIのリスクを正しく認識し、必要となる対策をライフサイクル全体で自主的に実行できるように後押しし、イノベーションの促進及びライフサイクルにわたるリスクの緩和を両立する枠組みを関係者と連携しながら積極的に共創していくことを目指す

AI事業者ガイドライン 検討の流れ

- 「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・経済産業省が共同事務局として、既存のガイドライン^(注)を統合・アップデートし、広範なAI事業者向けのガイドライン案を取りまとめ**
 (注) 既存のガイドラインは次の3つ：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)
- 検討にあたっては、**広島AIプロセスの議論やマルチステークホルダー・アプローチを重視**。総務省の「AIネットワーク社会推進会議」、経済産業省の「AI事業者ガイドライン検討会」及び各検討会下のWGを活用して、**産業界、アカデミア及び市民社会の多様な意見を反映**
- **パブリックコメント(2024年1月20日から2月19日まで実施)等を踏まえ修正し、AI事業者ガイドライン第1.0版はAI戦略会議に報告後、4月に公表**(今後も随時、適宜更新)

参考(これまでの主な経緯)



(参考) パブリックコメントの実施結果

- 本年1月20日から2月19日までの間、パブリックコメントを実施
- **1,049者から意見提出（法人・団体：47者^(注)、個人：1,002者）**
- **提出された意見の総数は約4,000件、約9割が個人からの意見**
 - ※ 同一人物・団体から重複して提出されたものを含む
- 総務省・経済産業省の合同検討会において、有識者の意見等を伺い検討（3月14日開催）
- **パブコメ等を踏まえ修正したAI事業者ガイドライン第1.0版を、AI戦略会議に報告し公表**（4月19日プレスリリース）

(注) 経済団体：日本経済団体連合会、新経済連盟、電子情報技術産業協会、日本民間放送連盟、米国商工会議所 等

AI開発企業：Google、ソフトバンク、LINEヤフー 等 関係団体：日本ディープラーニング協会、AIガバナンス協会、全国消費生活相談員協会 等

主なご意見と対応内容

記載内容のさらなる充実、明確化に関するもの（約2割）

- ガイドラインを活用する観点から、①記載内容の明確化・具体化・さらなる充実、②記載の平仄を揃えるべき 等
- ⇒ ご意見を踏まえ、主として、適切に修正に反映

AI政策一般に関するもの、その他（約2割）

- ①ソフトロー・ベースのアプローチに賛成、②ガイドラインの普及・促進が重要、③海外のガバナンスの枠組みとの相互運用性を検討していくべき、のほか、④生成AIのリスクへの懸念（生成AIの利用そのものに反対等を含む） 等
- ⇒ ご意見を踏まえ、主として、ガイドライン運用や更新等を含む今後のリスク対応等の検討に活用

著作権に関するもの（約6割）

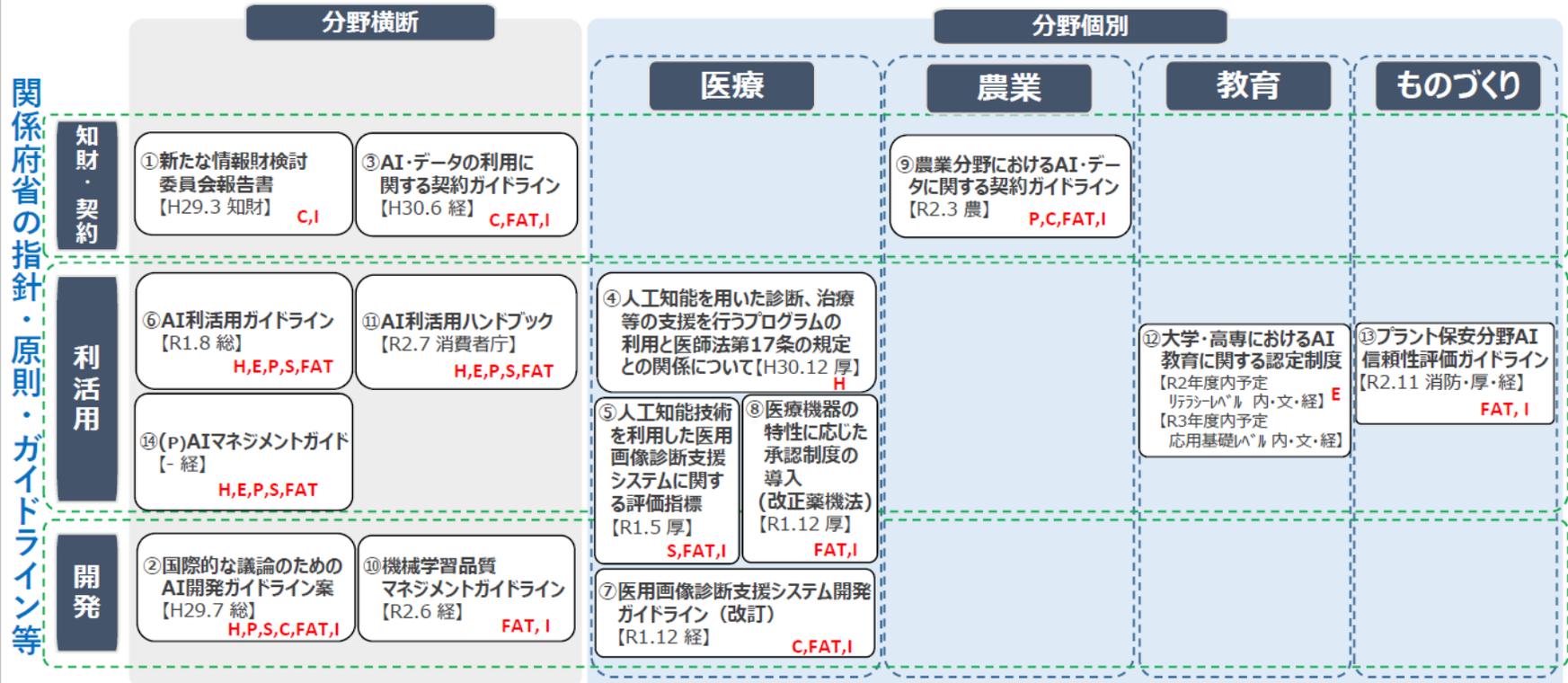
- 現状の著作権法に関する懸念^(注)が多く寄せられた一方、現状の著作権法の考え方がAIのモデル開発にむしろ望ましく維持すべきとのご意見等も
- 注) 例えば、制作者の許諾なくAIに学習されてしまうことへの懸念、実在の制作者に酷似した画風を生み出してしまうことへの懸念 等
- ⇒ 著作権法ほか知的財産権に関する法令等の遵守を基本として明記（最新の関連する検討状況等についても記載）

関係府省庁のAI関連の指針・原則・ガイドライン等俯瞰

人間中心のAI社会原則（平成31年3月 統合イノベーション戦略推進会議決定）の考え方を踏まえ、関係府省庁にてAI関連の指針・原則・ガイドライン等の策定が進捗。今後も、AIの健全な社会実装に向け、各分野での検討にあたっては、相互に参照を促すなど、引き続き関係府省間の情報共有を促進。

人間中心のAI社会原則

人間中心(H), 教育・リテラシー(E), プライバシー確保(P), セキュリティ確保(S), 公正競争確保(C), 公平性・説明責任・透明性(FAT), イノベーション(I)

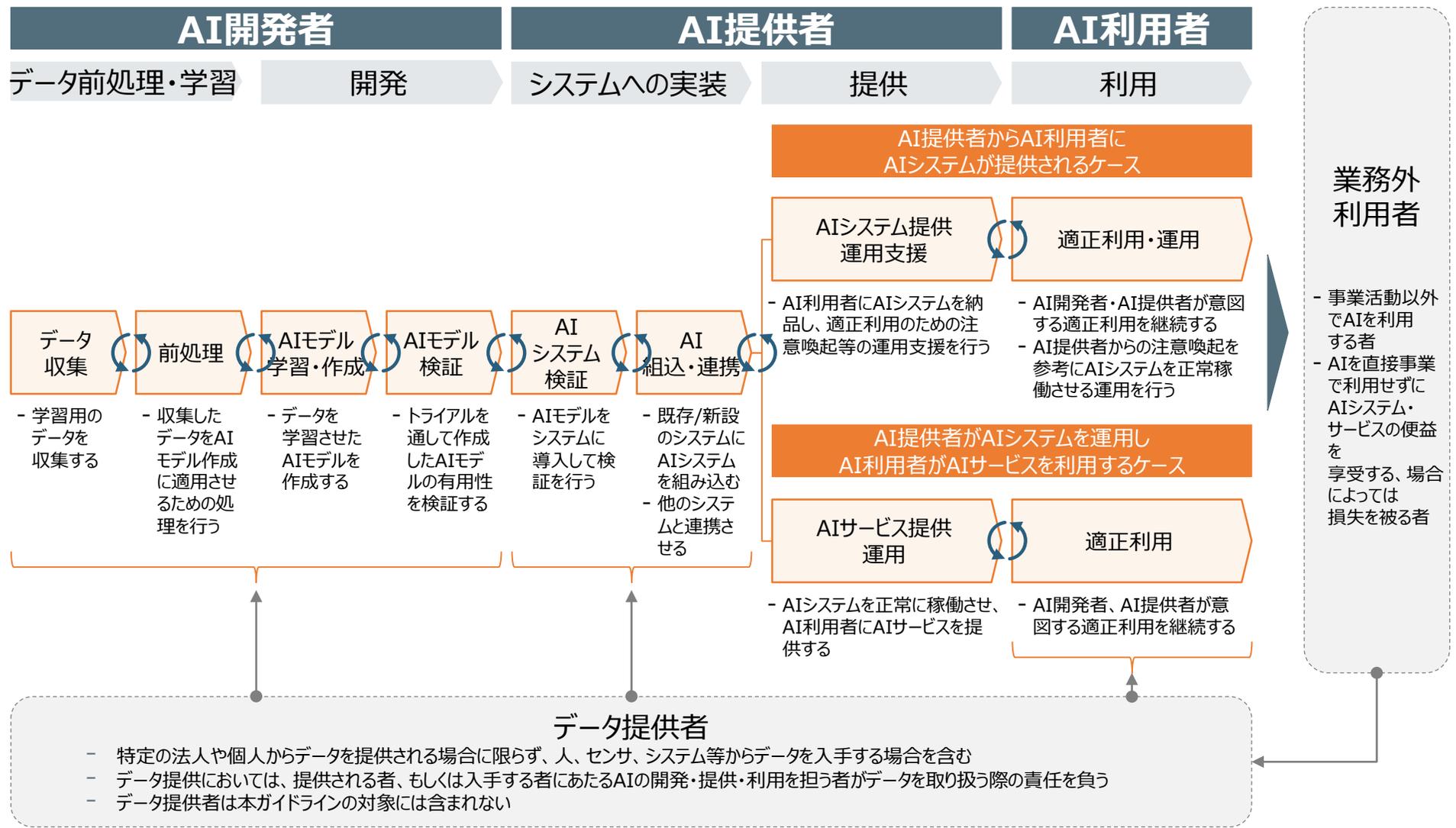


※赤字：人間中心のAI社会原則の関連する項目

今後、新たに検討される指針・原則・ガイドライン等についてもマッピング

一般的なAIの事業活動を担う主体

- AIライフサイクルにおける具体的な役割を考慮し、AIの事業活動を担う立場として、「AI開発者」、「AI提供者」、「AI利用者」の3つに大別して整理する ※「データ提供者」、「業務外利用者」は対象外とする



AI事業者ガイドライン（第1.0版）概要

- 事業活動においてAIに関係する全ての者（公的機関等も含む）を対象
- 事業者を ①AI開発者、②AI提供者、③AI利用者（注）に大別し、3つの事業者カテゴリに共通の指針を括りだした上で（第2部C）、各カテゴリに特有、重要となる事項を整理（第3部～第5部）
（注）事業活動以外でAIに関係する者は含まない
- 簡潔な本編を補完するため、別添において詳細に解説

本編の構成

- 第1部 AIとは
- 第2部 AIにより目指すべき社会と各主体が取り組む事項
 - A 基本理念
 - B 原則
 - C 共通の指針
 - D 高度なAIシステムに関係する事業者共通の指針
 - E AIガバナンスの構築
- 第3部 AI開発者に関する事項
（データ前処理・学習時、AI開発時、AI開発後、国際行動規範の遵守等）
- 第4部 AI提供者に関する事項
（AIシステム実装時、AIシステム・サービス提供後、国際指針の遵守等）
- 第5部 AI利用者に関する事項
（AIシステム・サービス利用時、国際指針の遵守等）

総論

各論

別添

本編を補完する位置付け、次の事項を盛り込み

- AIシステム・サービスの例（各主体の関係性等を含む）
- AIによる便益や可能性、具体的なリスクの事例
- ガバナンス構築のための実践ポイント、具体的な実践例
- 本編の各項目に関するポイント、具体的な手法の例示、分かりやすい参考文献 等

（上記に加えて）

- 「AI・データの利用に関する契約ガイドライン」を参照する際の主な留意事項
- チェックリスト
- 主体横断的な仮想事例 等

参考

AI事業者ガイドライン案を検討する
総務省、経済産業省の関連会議体

総務省

- AIネットワーク社会推進会議
（議長：須藤 修 中央大学国際情報学部教授）
- 同 AIガバナンス検討会
（座長：平野 晋 中央大学国際情報学部教授）



経済産業省

- AI事業者ガイドライン検討会
（座長：渡部 俊也 東京大学未来ビジョン研究センター教授）

AI事業者ガイドライン（第1.0版）公表


総務省
 Ministry of Internal Affairs and Communications

ご意見・ご提案 ENGLISH(TOP) MIC ICT Policy (English / Français / Español / Русский / 中文 / عربي)


 アクセシビリティ 視覚支援ツール

Google 提供

総務省の紹介 政策評価	広報・報道	政策	組織案内	所管法令	予算・決算	申請・手続
----------------	-------	----	------	------	-------	-------

[総務省トップ](#) > [広報・報道](#) > [報道資料一覧](#) > 「AI事業者ガイドライン案」に関する意見募集の結果及びガイドラインの公表

報道資料

令和6年4月19日
総務省
経済産業省

「AI事業者ガイドライン案」に関する意見募集の結果及びガイドラインの公表

総務省及び経済産業省は、「AI事業者ガイドライン案」を取りまとめ、令和6年1月20日（土）から同年2月19日（月）までの間、意見募集を実施いたしました。

その結果、計3947件の意見の提出がありましたので、提出された意見及び当該意見に対する両省の考え方を公表するとともに、「AI事業者ガイドライン（第1.0版）」を作成しましたので公表いたします。

1 背景

AIPR略会議（座長：松尾豊 東京大学大学院工学研究科教授）で取りまとめられた「AIに関する暫定的な論点整理」（令和5年5月26日（金））において、生成AIの普及を踏まえ、既存のガイドラインに関して必要な改訂などを検討する必要性が示されたことから、総務省及び経済産業省では、既存のガイドラインを統合・アップデートし、広範なAI事業者向けの統一的で分かりやすいガイドラインの検討を進めてきました。

両省において、それぞれ「AIネットワーク社会推進会議」（議長：須藤修 中央大学国際情報学部教授）、「AI事業者ガイドライン検討会」（座長：渡部俊也 東京大学未来ビジョン研究センター教授）を開催し、両会議での検討を踏まえて「AI事業者ガイドライン案」を取りまとめ意見募集を行いました。

【詳しくはプレスリリースをご参照ください】

「AI事業者ガイドライン案」に関する意見募集の結果及びガイドラインの公表（令和6年4月19日）

https://www.soumu.go.jp/menu_news/s-news/01ryutsu20_02000001_00010.html

【AI事業者ガイドライン 第1.0版】

<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\(別添9\)](#) 

<関連資料> ※英語版については仮訳

[「AI事業者ガイドライン\(第1.0版\)」本編\(概要\)](#) 

[「AI事業者ガイドライン\(第1.0版\)」別添\(概要\)](#) 

[AI Guidelines for Business Ver1.0](#) 

[Outline of “AI Guidelines for Business Ver1.0”](#) 

[AI Guidelines for Business Appendix Ver1.0](#) 

[Outline of “AI Guidelines for Business Appendix Ver1.0”](#) 

[▶ ページトップへ戻る](#)

「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

- 広島AIプロセスで取りまとめられた高度なAIシステムに関する国際指針及び国際行動規範を反映しつつ、一般的なAIを含む（想定され得る全ての）AIシステム・サービスを広範に対象
- 実際のAI開発・提供・利用においては、本ガイドラインを参照し、各事業者が指針遵守のために適切なAIガバナンスを構築するなど、具体的な取組を自主的に推進することが重要

高度な
AIシステム*¹に関する
取組事項

あらゆる
AIシステムに
関する取組事項

基本理念・原則

AIガバナンス

広島AIプロセスの成果（包括的政策枠組み）を反映

- ・ 全ての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通じる指針」、第3部

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(総務省)も取込み

- 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、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部～第5部

「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、基本理念・原則を構成

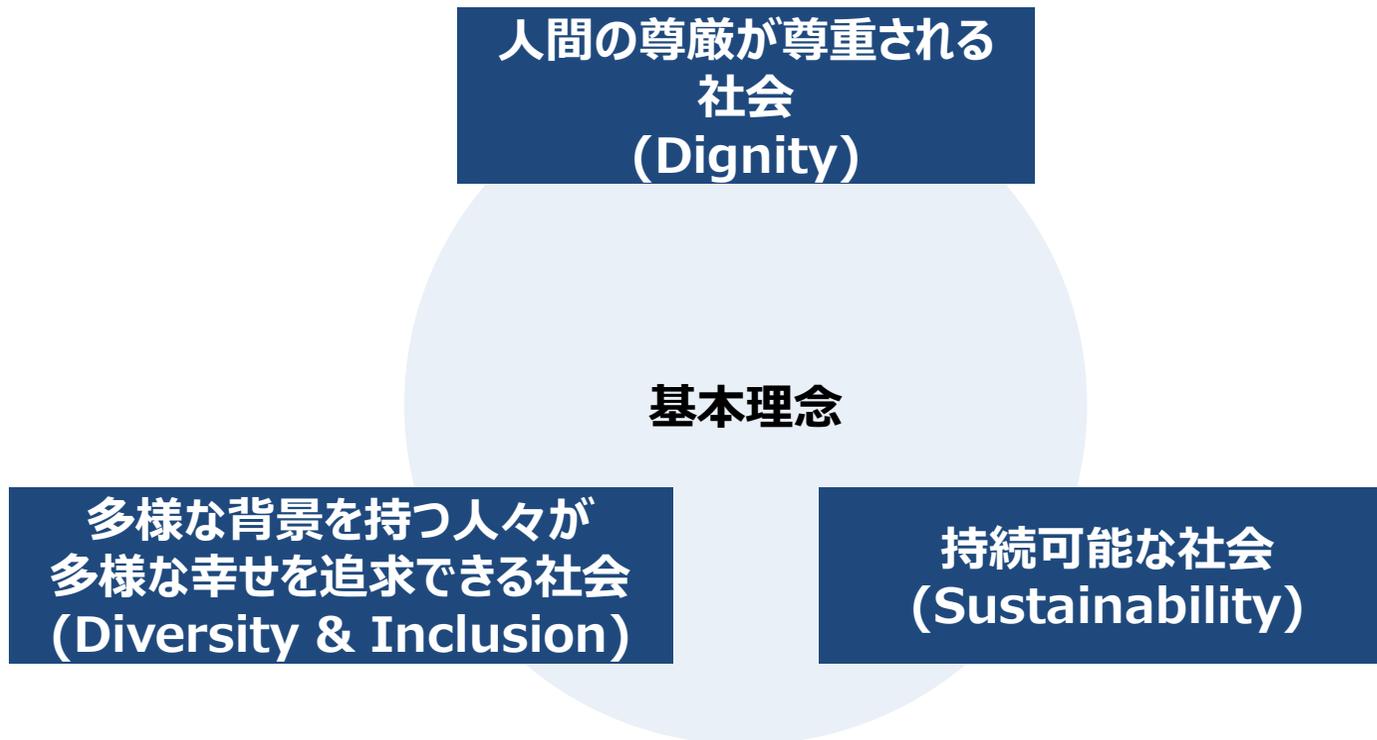
第2部 A.「基本理念」 B.「原則」

AI原則実践のためのガバナンス・ガイドライン(経済産業省)を元に整理

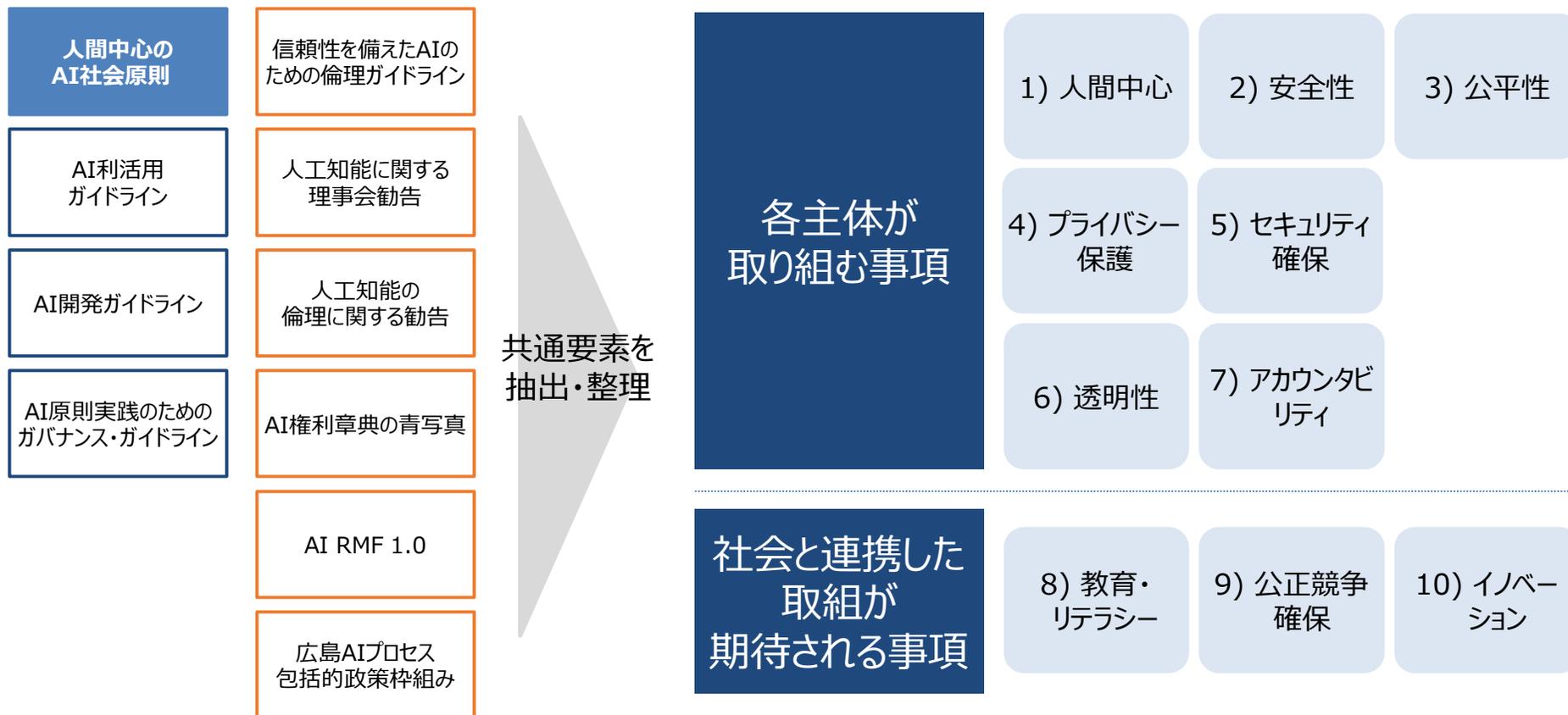
第2部 E.「AIガバナンスの構築」

*1: 最先端の基盤モデル及び生成AIシステムを含む、最も高度な AI システム

- 「人間中心のAI社会原則」において、AIがSociety 5.0の実現に貢献し、AIを人類の公共財として活用することで、社会の在り方の質的变化や真のイノベーションを通じて地球規模の持続可能性へとつなげることが重要であると述べられている
- 加えて、下記の3つの価値を「基本理念」として尊重し、「その実現を追求する社会を構築していくべき」としており、この普遍的な考え方は、今後も目指すべき理念であり続けている



- AIの活用による目指すべき社会の実現のために各主体が連携して取り組む内容を原則としてまとめた上で、「共通の指針」として整理する
- 「共通の指針」は、「人間中心のAI社会原則」を土台としつつ、諸外国における議論状況や、新技術の台頭に伴い生じるリスクへの対応等を反映している
- その結果、各主体が取り組む事項、及び社会と連携して取り組むことが期待される事項に分類される



- 各主体は、1) 人間中心に照らし、法の支配、人権、民主主義、多様性及び公平公正な社会を尊重する
- 憲法、知的財産関連法令及び個人情報保護法をはじめとする関連法令、AIに係る個別分野の既存法令等を遵守すべきであり、国際的な指針等の検討状況についても留意することが重要
- AIガバナンスを構築し継続的に運用（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実施）

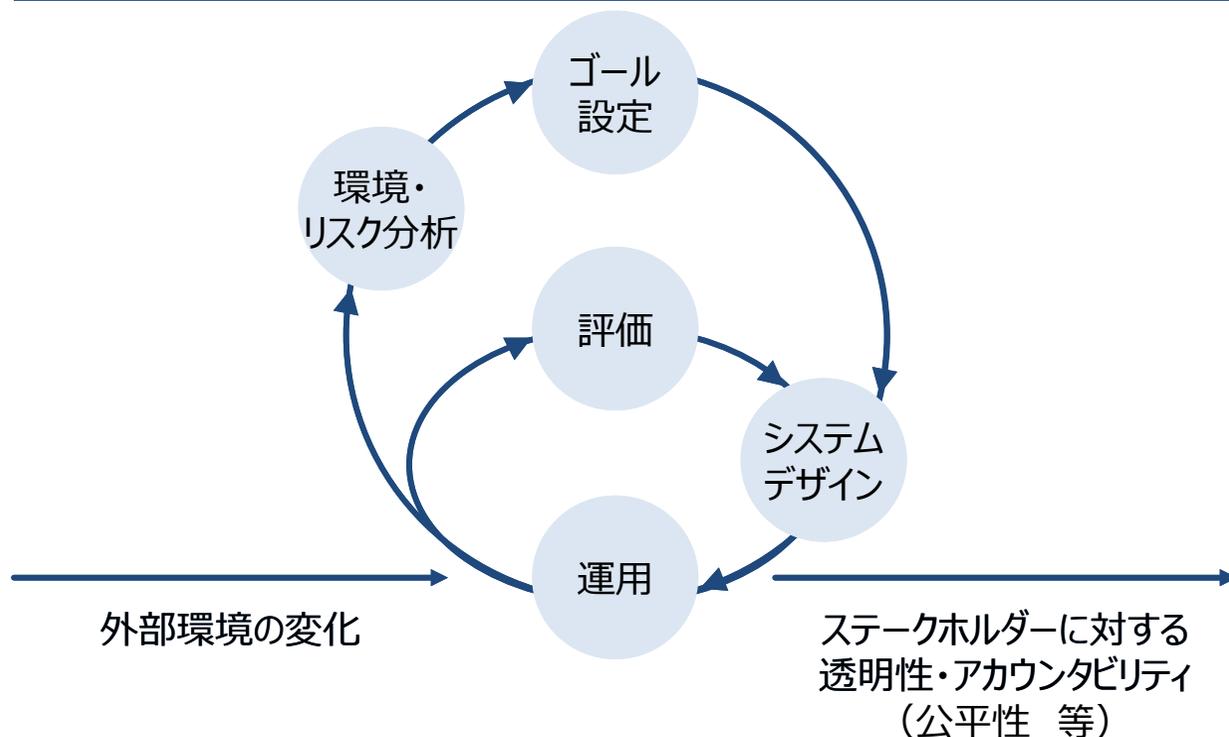
指針	内容（主な項目の抜粋）
各主体が 取り組む事項	<p>1) 人間中心</p> <ul style="list-style-type: none"> ✓ AI が人々の能力を拡張し、多様な人々の多様な幸せ（well-being）の追求が可能となるよう行動する ✓ AI が生成した偽情報・誤情報・偏向情報が社会を不安定化・混乱させるリスクが高まっていることを認識した上で必要な対策を講じる ✓ より多くの人々がAIの恩恵を享受できるよう社会的弱者によるAIの活用を容易にするよう注意を払う
	<p>2) 安全性</p> <ul style="list-style-type: none"> ✓ 適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
	<p>3) 公平性</p> <ul style="list-style-type: none"> ✓ 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、無意識や潜在的なバイアスに留意し、AIの開発・提供・利用を行う
	<p>4) プライバシー保護</p> <ul style="list-style-type: none"> ✓ 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
	<p>5) セキュリティ確保</p> <ul style="list-style-type: none"> ✓ 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

- 各主体は、1) 人間中心に照らし、法の支配、人権、民主主義、多様性及び公平公正な社会を尊重する
- 憲法、知的財産関連法令及び個人情報保護法をはじめとする関連法令、AIに係る個別分野の既存法令等を遵守すべきであり、国際的な指針等の検討状況についても留意することが重要
- AIガバナンスを構築し継続的に運用（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実施）

指針	内容（主な項目の抜粋）
各主体が 取り組む事項 (続き)	<p>6) 透明性</p> <ul style="list-style-type: none"> ✓ 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）
	<p>7) アカウンタビリティ</p> <ul style="list-style-type: none"> ✓ 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	<p>8) 教育・リテラシー</p> <ul style="list-style-type: none"> ✓ 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。
	<p>9) 公正競争確保</p> <ul style="list-style-type: none"> ✓ 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
	<p>10) イノベーション</p> <ul style="list-style-type: none"> ✓ 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を安全安心に活用していくために、経営層のリーダーシップのもと、下記に留意しながら適切なAIガバナンスを構築することで、リスクをマネジメントしていくことが重要となる
 - 複数主体に跨る論点について、バリューチェーン/リスクチェーンの観点で主体間の連携確保
 - 上記が複数国にわたる場合、データの自由な越境移転の確保のための適切なAIガバナンスの検討
 - 経営層のコミットメントによる、各組織の戦略や企業体制への落とし込み/文化としての浸透

適切なAIガバナンスの構築



複数主体間の連携確保

バリューチェーン/リスクチェーンの観点で主体間の連携を確保



適切なデータ流通

複数国にまたがる想定の場合の適切なリスク管理/AIガバナンスの実施



経営層のコミットメント

戦略/体制への落とし込み、各組織の文化としての浸透



AI事業者ガイドライン（第1.0版）各主体が取り組む主な事項の例（抜粋）⁴¹

第2部

AIにより目指すべき社会及び各主体が取り組む事項

- 法の支配、人権、民主主義、多様性、公平公正な社会を尊重するようAIシステム・サービスを開発・提供・利用すべきである。また、憲法、関連法令、AIに係る個別分野の既存法令等を遵守、人間の意思決定や感情等を不当に操作することを目的とした開発・提供・利用は行わない
- 人間の生命・身体・財産、精神及び環境への配慮、偽情報等への対策、AIモデルの各構成技術に含まれるバイアスへの配慮
- プライバシー保護やセキュリティー確保、関連するステークホルダーへの情報提供（AIを利用しているという事実、AIモデルの情報等）
- トレーサビリティの向上（データの出所、開発・提供・利用中に行われた意思決定等）
- 文書化（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等）
- AIリテラシーの確保、オープンイノベーション等の推進、相互接続性・相互運用性への留意等
- 高度なAIシステムに係る事業者は、広島AIプロセスで示された国際指針を遵守（開発者は国際行動規範も遵守）
- 「環境・リスク分析」「ゴール設定」「システムデザイン」「運用」「評価」といったサイクルを、マルチステークホルダーで継続的かつ高速に回転させていく、「アジャイル・ガバナンス」の実践 等

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

目次

1. AIとは
2. AIに関する政府全体の取組
3. AIガバナンスの主な動向
4. AI事業者ガイドライン
- 5. その他の関連施策等**

- 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 CDoO(Chief Data Officer) データドリブン経営推進部長（現職）

1. GPAI概要

- ✓ GPAI (The Global Partnership on Artificial Intelligence) は、人間中心の考え方に立ち、「責任あるA I」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 (2020年) された、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
- ✓ モントリオールとパリに続き、2024年、東京にアジア初となる専門家支援センターを設置

2. 東京センター概要

- ✓ 名称： GPAI東京専門家支援センター／GPAI Tokyo Expert Support Center
- ✓ センター長: 原山 優子 東北大学名誉教授
- ✓ 情報通信研究機構 (NICT) に設置

3. 東京センターの主要業務

- ✓ GPAI専門家によるプロジェクト※の実施の支援
 - ・プロジェクト関係： 専門家との連絡、プロジェクト進行管理、報告書作成支援等
 - ・その他： GPAI事務局及び仏・加の専門家支援センターとの連絡・調整、GPAIサミットやその他のAI関連国際会議への参加

※ 本センターで支援する最初のプロジェクトでは、生成AIの安全性に関する解決策の確立に向けた偽・誤情報対策や透明性確保等の取組を幅広くグローバルに収集し、マッピング



原山センター長

- ・2010年～2012年
OECD科学技術産業局次長
- ・2013月～2018年
総合科学技術・イノベーション
会議議員
- ・2020年～2022年
理化学研究所理事



GPAI中間サミット
(7/3、インド)

1. 検討事項

※ 総務省等におけるAIやメタバース等に関する検討と連携

- ① **デジタル空間を活用したサービスの普及・情報通信技術の進展等の状況**：（例）生成AI、メタバースなど
- ② **新たな課題と各ステークホルダーによる対応状況**：
（例）生成AI等による巧妙な偽・誤情報の生成や拡散に伴う社会的な影響の深刻化、メタバースにおけるデータの取得・利用に係る対応の重要性、国境を越えた情報・データ流通の広域化や迅速化に伴う国際的な協調の必要性など
- ③ **今後の対応にあたっての基本的な考え方**：
（例）基本理念：信頼性のある自由な情報流通、表現の自由、知る権利、青少年を含む利用者保護、デジタルシティズンシップなど各ステークホルダーの役割：デジタルプラットフォーム事業者、生成AI事業者、仮想空間関係事業者、通信・放送事業者、利用者など
- ④ **デジタル空間における情報流通の健全性確保に向けた具体的な方策**：
（例）多様なステークホルダーによる協力関係の構築、ファクトチェックの推進、幅広い世代に対するリテラシーの向上、情報発信者側を含む自主的取組の推進、研究開発の推進、国際的な対話の深化、生成AI・メタバース関連事項など

2. 構成員

※法律（憲法・メディア・情報）、技術、データ分析、認知科学・心理学、ジャーナリズム、プライバシー、サイバーセキュリティ、消費者保護等の専門家。以下、座長・座長代理以外は、50音順。

- 穴戸 常寿（東京大学大学院法学政治学研究科教授）【座長】
- 山本 龍彦（慶應義塾大学大学院法務研究科教授）【座長代理】
- 生貝 直人（一橋大学大学院法学研究科教授）
- 石井 夏生利（中央大学国際情報学部教授）
- 越前 功（国立情報学研究所情報社会相関研究系教授）
- 江間 有沙（東京大学国際高等研究所東京カレッジ准教授）
- 奥村 信幸（武蔵大学社会学部教授）
- 落合 孝文（渥美坂井法律事務所・外国法共同事業プロトタイプ政策研究所所長・シニアパートナー弁護士）
- クロサカ タツヤ（慶應義塾大学大学院政策・メディア研究科特任准教授）
- 後藤 厚宏（情報セキュリティ大学院大学学長）
- 澁谷 遊野（東京大学空間情報科学研究センター 准教授）
- 曾我部 真裕（京都大学大学院法学研究科教授）
- 田中 優子（名古屋工業大学大学院工学研究科基礎類 准教授）
- 増田 悦子（公益社団法人全国消費生活相談員協会理事長）
- 水谷 瑛嗣郎（関西大学社会学部メディア専攻准教授）
- 森 亮二（英知法律事務所弁護士）
- 安野 智子（中央大学文学部教授）
- 山口 真一（国際大学グローバル・コミュニケーション・センター准教授）
- 山本 健人（北九州市立大学法学部准教授）
- 脇浜 紀子（京都産業大学現代社会学部教授）

3. スケジュール

※ その他、関係団体がオブザーバとして参加

- 令和5年11月7日に第1回会合を開催し、ヒアリング・論点整理等を経て、本年夏までに一定のとりまとめを公表予定

① ディープフェイク対策技術の開発・実証

- インターネット上で流通する様々な情報について、生成AIにより生成されたコンテンツ（画像・映像）を判別する技術の開発・実証を実施。

【対策技術イメージ（例）】

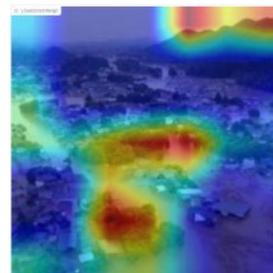
人物に関するディープフェイクだけでなく、街並み・風景についても判別。



2022年9月静岡水害時に拡散されたAI生成画像



判別技術
を適用



生成AIを利用した疑いがある範囲を黄・赤色等でマッピング

- 人間の目では判別しにくい不自然な点を自動で判別（山の輪郭や濁流等）
- 生成AIにより生成された確率を判定

画像出典：NABLAS株式会社提供資料

② 情報コンテンツや発信者の信頼性等確保技術の導入促進

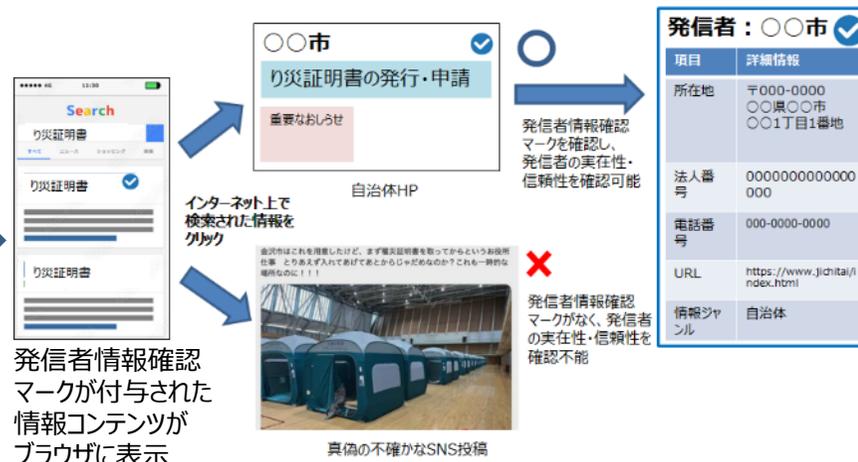
- 画像・動画等の情報コンテンツに対して来歴※1や発信者に関する情報※2を付与することなど、情報コンテンツや発信者の信頼性等を確保する技術の導入を促進すべく、技術実証を支援。

【対策技術イメージ（例）】

発信者情報の作成

- 法人ごとに発信者情報確認マークを作成
- 法人ごとに実在性・信頼性を証明する発信者情報を作成

各法人が公開する情報コンテンツに発信者情報を付与



発信者情報確認マークが付与された情報コンテンツがブラウザに表示

真偽の不確かなSNS投稿

※1 来歴に関する情報例：写真・動画の撮影日時、場所
 ※2 発信者の「信頼性」に資する情報：
 企業姿勢、編集方針、報道責任、編集ガイドライン、プライバシーポリシー等の情報
 発信者の「実在性」を確認できる情報：
 情報コンテンツの作成者、配信サイト運営者、広告主等の企業の基本情報（法人名等）

- 総務省において、生成AIに係るリテラシーに関する啓発教材として「生成AIははじめの一步～ 生成AIの入門的な使い方と注意点～ver1.0」を開発し、令和6年4月に公表。

(対象者) 今後の生活の中で、生成AIに触れうる国民の方（初心者向け）

(目的) 生成AIを自身で利活用できるリテラシーを身に付けること

(内容) ①生成AIの基礎知識、②生成AIの活用場面や入門的な使い方、③生成AI活用時の注意点

(形式) オンラインでも実施が可能な内容としており、編集が容易なパワーポイント形式にて公表することで、講師の裁量により事例等を追加することが可能。

※総務省HP「上手にネットと付き合おう！安心・安全なインターネット利用ガイド」https://www.soumu.go.jp/use_the_internet_wisely/special/generativeai/

生成AIの特徴

生成物

- 文章・画像・動画 など幅広く作成可能
- 文脈を理解した回答や、人間が思いつかない回答を作ることができる

指示方法

- 高度な技術は必要なく、簡単に使える
- 文章のほか、画像・音声・グラフなどを入力しても回答を得られる

コンテンツの対象と目的

対象

今後の生活の中で、生成AIに触れうる国民の方
※対象は専門業務のお客対象としません。

目的

生成AIを自身で利活用できるリテラシーを身に付けること

学習の流れ

1 生成AIとは何か？	生成AIの基礎知識を学習
2 生成AIをどのように使うか？	生成AIの活用場面や入門的な使い方を学習
3 生成AI活用にあたって注意すべきポイントは？	生成AIを使いこなすうえでの4つの注意点を学習 <ul style="list-style-type: none"> 情報の正確性 情報流出 知的財産権の侵害 活用者としてのモラル

事例 2. 情報検索

活用場面の例 ① 情報のレシビリティ ② 製品企画のための市場動向調査の補助 など

東京から日帰りで行ける県外の観光地を3つ教えてください。

学習結果

- 学習結果
- それぞれの観光スポットも教えてください
- おすすめの観光地も詳しく教えてください

検索結果を指示 → 整理された検索結果！

※AI: OpenAI GPT-4o

個人情報や機密情報を生成AIに入力すると、情報流出のリスクがある

生成AIは、利用者が入力したデータを学習データとして利用することがある

リスク1

個人情報^{※1}や社外秘の機密情報^{※2}を入力すると、他人の質問への回答に使われ、情報が漏洩する可能性がある

リスク2

漏洩した情報がサイバー犯罪などに悪用される恐れがある

※1: 個人識別情報、氏名、住所、年齢、性別、メールアドレス、電話番号、住所、顔の画像、指紋、音声データ、その他の個人データ
 ※2: 顧客リスト、営業戦略、開発計画、社内会議の録音、顧客の個人情報、財務データ、営業秘密、その他の機密情報

- 一般消費者に向けて、AIの概要や活用事例、使用に当たってのチェックポイント等を示した「AI利活用ハンドブック～AIをかしく使いこなすために～」を作成・公表（2020年7月）
- 消費者が今後も安心してAIの利活用ができるように、生成AIに関する最新の情報等を盛り込んだ「AI利活用ハンドブック～生成AI編～」を作成・公表（2024年5月）



・ 生成AIとは ・

生成AIとは？

生成AIとは、文字等を入力した指示に基づいて文章・画像・音楽・プログラムなどのコンテンツを作成できる人工知能（Artificial Intelligence：AI）です。生成AIは、人間と会話しているような自然な対話を通じて文章や画像等を作成します。

生成AIの仕組み

学習した大量の文章や画像等のデータから、指示にもとづいてコンテンツを作成します。意味を真に理解して作成しているわけはあまりありません。新しい文章を生成するAIでは、ある単語に続く単語が最も適切だと判断して生成しています。

生成AIに向いている用途

生成AIが得意な用途で利用しましょう。
得意な用途：要約・言い換え・翻訳、アイデア出し 等
不得意な用途：最新の情報、厳密な正確性が必要な情報 等

生成AIに指示する方法

生成AIで得られる効果の質は、大まかに入力する指示（プロンプト）によって変わります。
2024年3月時点では、どのような指示を入力すればどのような結果が得られるかははっきりとはしていませんが、次のような留意点を念頭にとると考えられます。

- 目的や対象などの前提条件
- 簡潔で短い内容
- 指示形式（200文字以内、候補5つ等）
- より取りやすいものでも、応答の質も変わる可能性があります。

AIについて詳しくはハンドブックをご覧ください。
消費庁「AI利活用ハンドブック」をダウンロードしてご覧ください。

・ 社会に普及する生成AI ・

生成AIの利活用

生成AIを利用する際には以下の点に留意する必要があります。

生成AIを利用した事業者のサービス事例

チャットボット（文章生成AIの高活用事例）
画像生成AIの高活用事例

生成AIを使用した事例

- 詐欺
- 偽造品等、社会の混乱等

・ 事例① 画像生成サービス ・

画像生成サービスの仕組み

利用者が「生成AI」に入力した文章や画像を基に、生成AIが画像を生成する。生成された画像は、利用者の指定した条件に基づいて生成される。

画像生成サービスを利用するときの注意点

1. 画像生成サービスを使う前には何に注意すべきですか？
画像生成サービスを利用する際は、生成AIの出力結果が正確であること、生成された画像の著作権に留意する必要があります。
2. 画像生成サービスで生成した画像を複製しても良いですか？
生成AIが生成した画像は、生成AIの出力結果であり、複製や改変を行うことはできません。複製や改変を行う場合は、生成AIの利用規約を確認する必要があります。
3. 画像生成サービスが不得意な内容はありませんか？
生成AIは、特定の分野や特定の種類の画像の生成に得意な場合があります。生成AIが得意でない分野や種類の画像の生成は、生成AIの利用規約を確認する必要があります。

・ 事例② チャットボット ・

チャットボットサービスの例

オンライン英会話サービス
AIフレンド

チャットボットを利用したチャットボットの仕組み

利用者が「生成AI」に入力した文章や画像を基に、生成AIがチャットボットの回答を生成する。生成された回答は、チャットボットの対話履歴に基づいて生成される。

チャットボットを利用するときの注意点

1. チャットボットの回答は正しいのでしょうか？
生成AIが生成した回答は、生成AIの出力結果であり、必ずしも正しいとは限りません。生成AIの出力結果を確認し、必要に応じて修正を行う必要があります。
2. チャットボットに個人情報を入力される場合があります。
チャットボットに個人情報を入力する場合は、生成AIの利用規約を確認する必要があります。生成AIが個人情報を学習し、悪用する可能性があります。

出典：
https://www.caa.go.jp/policies/policy/consumer_policy/information/ai_handbook

著作権法 第30条の4

著作物は、次に掲げる場合その他の当該著作物に表現された思想又は感情を自ら享受し又は他人に享受させることを目的としない場合には、その必要と認められる限度において、いずれの方法によるかを問わず、利用することができる。ただし、当該著作物の種類及び用途並びに当該利用の態様に照らし著作権者の利益を不当に害することとなる場合は、この限りでない。

- 一 著作物の録音、録画その他の利用に係る技術の開発又は実用化のための試験の用に供する場合
- 二 情報解析(多数の著作物その他の大量の情報から、当該情報を構成する言語、音、影像その他の要素に係る情報を抽出し、比較、分類その他の解析を行うことをいう。第四十七条の五第一項第二号において同じ。)の用に供する場合
- 三 前二号に掲げる場合のほか、著作物の表現についての人の知覚による認識を伴うことなく当該著作物を電子計算機による情報処理の過程における利用その他の利用(プログラムの著作物にあつては、当該著作物の電子計算機における実行を除く。)に供する場合

文化審議会 著作権分科会 法制度小委員会

- 2023年7月から審議(7回開催)
- 本年2月29日にパブコメ結果公表(約2万5千件)、「考え方」を提示

今後の取組

- 基本的な考え方や著作権法の解釈について、分かりやすい形で、更なる周知啓発
- 判例等の蓄積、技術の発展、諸外国における立法や検討状況の把握・収集
- AIの開発や利用によって生じた著作権侵害について、相談窓口等を通じた事案の集積
- 共通理解の醸成に向け、関係当事者間の適切なコミュニケーションの実現

【開催趣旨】

- 様々なAIツールが生み出され、普及していく中であって、それらの開発・提供・利用を促進し、我が国経済社会の発展につなげていくためにも、生成AIの懸念やリスク等への対応を適切に行う必要がある
- このことを踏まえ、AIと知的財産権等との関係をめぐる課題への対応について、関係省庁における整理等を踏まえつつ、必要な対応方策等を検討するため、「AI時代の知的財産権検討会」を開催

基本的視点

- (1) 産業競争力強化の視点
- (2) AI技術の進歩の促進と知的財産権の保護の視点
- (3) 国際的視点

主な検討課題

- I. 生成AIと知財をめぐる懸念・リスクへの対応等
・著作権との関係 等
- II. AI技術の進展を踏まえた発明の保護の在り方

委員構成

〔座長〕 渡部俊也・東京大学執行役・副学長
未来ビジョン研究センター教授

- * AI技術研究者、AI開発・ビジネス事業者、コンテンツ関係従事者、知財法研究者、法曹実務家を含む、計13名で構成
- * オブザーバー：内閣府（CSTI）、文化庁、経産省、特許庁
法務省、総務省、公正取引委員会、外務省

開催経過

- 第1回（2023年10月4日）
・開催趣旨・背景
・本検討会において検討すべき課題
- 第2回（2023年10月18日）
・ヒアリング（JASRAC、日本知的財産協会、AI Picasso）
・議論
- 第3回（2023年11月7日）
・ヒアリング（日本マイクロソフト、日本新聞協会、特許庁）
・議論
- 第4回（2023年12月11日）
・ヒアリング（レベルファイブ、文化庁、経産省）
・意見募集結果公表〔意見募集：10月5日～11月5日〕
・論点整理（議論の振り返り）
- 第5回（2024年1月26日）
・残された論点等について検討
- 第6回（2024年3月21日）
・ヒアリング（文化庁、経産省）
・横断的見地からの検討
・中間とりまとめ骨子（案）
- <予定>
第7回（2024年4月22日）中間とりまとめ（案）



ご静聴ありがとうございました。

〒100-8926
東京都千代田区霞が関2-1-2
中央合同庁舎第2号館
電話: 03-5253-5111(代表)