# AI ホワイトペーパー 2024

## ステージⅡにおける新戦略

## ― 世界一 AI フレンドリーな国へ ―

2024 年 4 月 11 日

自由民主党デジタル社会推進本部

AI の進化と実装に関するプロジェクトチーム

# 目次

# 第１章　「ステージⅡ」に臨む日本

## ○ 自民党ホワイトペーパー（2023 年）

2022 年晩秋、ChatGPT が話題になり始めた当初から、自由民主党「AI の進化と実装に関するプロジェクトチーム（以下、**自民党 AI プロジェクトチーム**）」は生成 AI の圧倒的なインパクトを認識し、2023 年 4 月、**ホワイトペーパー**をいち早く国内外に発信した。これが全ての始まり、いわば**日本の AI のビッグバン**である。

このホワイトペーパーでは、新たな戦略や司令塔の必要性、AI 開発基盤の強化、法規制の検討など、**大胆で先見性に富む施策を包括的に提言**した。

ホワイトペーパー発表から一年経過したが、これまでに、国内外から多くの問合せを受けるとともに、他国で類似の戦略も発表されている。また、ホワイトペーパーの**提言のほとんどは実現**している。例えば、政府は、AI 戦略会議・AI 戦略チームの設置、研究開発力強化に向けた計算資源の確保、学習用データの整備、モデル開発力の強化等に直ちに着手した。そして、岸田首相は広島 AI プロセスの立ち上げを提唱し、国際指針・国際行動規範をとりまとめた。また、政府は AI 事業者ガイドラインの策定やその履行確保の在り方の検討等に取り組んでいる。

こうした動きは、**ホワイトペーパーの先見性の高さと、政府の反応の早さの双方**を示すものであり、ともに特筆すべきものである。

加えて、**昨年 12 月に自民党 AI プロジェクトチームは緊急提言**を発出した。これに沿って、政府は **AI セーフティ・インスティテュートを設立**した。英米に続き、世界で 3 番目の設置、アジアでは最初の設置であり、高く評価したい。

## ○ 現在（2024 年４月）の景色

この一年の世界と日本の動きは、官民問わず、他の分野で例を見ない速さだった。その結果、現在の景色は一年前とは全く異なり、**一年前には誰も予測できなかった景色**となっている。

米国ビッグテック等による**高性能・大規模の汎用基盤モデルの進化・社会実装**が進む中、日本を含む各国は、**小規模・高性能モデルや複数モデルの組合せなど、多様な戦い方**で活

路を見出そうとしている。スタートアップ等による新たなモデルの開発、大学・研究機関での先進的な研究により、日本は世界と競い合っている。

**オープンソース**の AI が登場し、セキュリティ面での不安の声もあるが、多様な主体が開発に参加できる等の利点もあり、**誰でも AI を開発できる**状況になっている。

生成 AI が扱うデータは、テキストに加え、画像、動画、音声、音楽、プログラムなど、**マルチモーダル化**が進んでいる。創薬、材料開発等の研究に AI を用いる「**AI for Science**」、ロボットへの AI 搭載など様々な分野に AI を導入する「**AI for ALL**」の広がりの兆しも見られる。

AI の利活用は、中央官庁、地方自治体の行政サービス、金融、教育機関等での顧客サービス、事務効率化等においても進みつつある。**自由民主党は、生成 AI をいち早く自ら試行的に開発し、実証的に利用**している。

一方で、生成 AI の性能向上と利用拡大に伴い、サイバー攻撃や詐欺の巧妙化、偽情報・誤情報の流布、ハルシネーション、著作権・知的財産権の侵害、個人情報漏洩等の**多様なリスクに対する懸念が広がっている**。安全保障や偏見・差別等のリスクを巡る議論も活発化した。今年は世界的な「選挙イヤー」であり、**AI を用いた選挙妨害**も世界的に懸念されている。

リスクへの対応に関しては、**人権尊重や差別・偏見の排除等を軸に包括的な法規制に進む EU**、AI 開発大手による自主的なコミットメントに加えて**安全保障の観点を中心に既存の法令を用いて対応する米国**など、様々なアプローチが見られる。

## ○ ステージⅡの戦略　― 世界一 AI フレンドリーな国へ ―

現在の光景と一年前の光景が全く異なるのと同様に、一年後の光景は現在とは全く異なるものになるであろう。我々は今、**誰も予測できない「ステージⅡ」の入口**にいる。AI の開発・利用に関わる者が増え、ステージⅡにおいては、技術、サービス、利用形態、規制等のあらゆる面での世界の動きがこれまで以上にダイナミックに進む可能性がある。

変化が速く、先が見えない中で、我々は国民の安心・安全を守りながら産業競争力を強化するとともに、より良い世界を築いていくための貢献もしていかなければならない。その戦略を練るためには**多様なステークホルダーとの対話**が重要であり、我々はこの一年間、OpenAI 社のサム・アルトマン CEO、NVIDIA 社のジェンスン・ファン CEO、モントリオール大学のヨシュア・ベンジオ教授をはじめ、80 人を超える様々な分野の有識者と率直な意見交換を重ねてきた。（開催実績の詳細は、別紙 1 参照）

ステージⅡにおいて日本は「**世界一 AI フレンドリーな国**」を目指すべきである。世界各地で政治的・経済的な思惑から AI 開発・規制に関して様々な駆け引きが見られるが、**日本**

**は世界で最も AI に理解があり、AI を実装しやすい国を目指す**。そして、**国民のリスクは最小化しつつ**、**利益を最大化**する。また、日本は、**広島 AI プロセスの実績をベースに、国際的にさらに大きなリーダーシップ**を発揮すべきである。

世界一 AI フレンドリーな国を実現するためには、**競争力強化と安全性確保の両面からの新戦略**が必要である。**競争力強化に関しては、AI を研究開発する側と AI を利活用する側の双方のイノベーション**を同時に創出していく必要がある。**安全性確保は、競争力強化と相互補完関係にあり、一体的に進める**ことが大事である。

AI はデータを学習して進化するため、データ戦略が重要である。ビジネスや社会課題解決に**有益なデータが**、個人情報保護、セキュリティ、知的財産権等に関する**信頼を確保しながら、グローバルに自由に流通することと相まって、AI は健全な発展**を遂げる。

また、AI は様々な分野で利用されるため、例えばスタートアップ、半導体、ロボットなど、**様々な分野の戦略・政策との連携、施策の有効活用**も重要である。

今後も AI に関する政策は広がり続けると見られ、政府の**司令塔機能の強化**が必要である。

これらを踏まえ、以下を提言する。

- 「**世界一 AI フレンドリーな国**」、すなわち、世界で最も AI に理解があり、AI の研究開発・実装がしやすい国を、官民をあげて実現すること。
- 官民をあげて AI による**国民のリスクを最小化しつつ、利益を最大化**すること。
- **相互補完関係にある競争力強化と安全性確保**を官民をあげて**一体的に推進**すること。
- 日本は、**広島 AI プロセスの実績をベースに**、安心・安全で信頼性のある AI に関する**国際的なルールメーキングを引き続き主導**すること。
- また、**アジア諸国やグローバルサウスとの協調関係を強化**するとともに、AI に関する日米共同研究など**国際共同研究や利活用促進についても、世界の中で強いリーダーシップを発揮**すること。
- 政府は、AI 政策の司令塔である「AI 戦略チーム」、「AI 国際戦略推進チーム」**を支える事務局の体制を強化**すること。

## ○ AI を活用した日本の競争力強化のための戦略（第２章）

**日本の競争力の強化に向けては、高度な人材やインフラ**（計算基盤、通信基盤等）を基に、**AI 研究開発力の強化と AI の利活用促進を一体的に**、かつ、**グローバルな視点で進めていく**エコシステムが重要である。それを構築・推進する戦略を第 2 章で記述する。

AI 戦略のプレイヤーは官民の多岐にわたる。

**政府は、**多様な**企業が AI に関する競争力を高め、それぞれの勝ち筋を見出していけるよう、環境整備、人材育成、国際協調**を進める役割を担っている。

技術も環境も激しい変化が続くため、**各プレイヤーは**様々な可能性に即応できる**複数のシナリオを描いておき、臨機応変に対応**していくことも時に必要である。**環境変化に屈するのではなく、環境変化を味方にする**柔軟性が重要である。

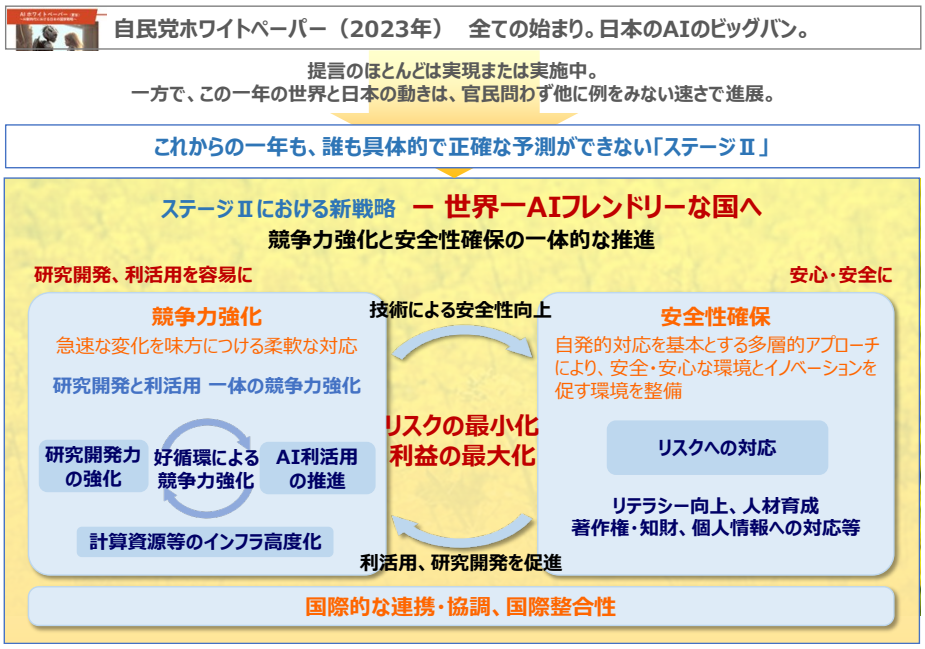
## ○ 安全性確保のための戦略（第３章）

リスクを抑えるためには、**AI が適切に開発・提供され、適切に使われる必要がある。**技術の変化の速さ、複雑さ、多様性等を踏まえると、まずは**AI に関わる全ての者が、リテラシーを高め、自発的に一定の規律（ソフトロー）を守る**必要がある。

加えて、国民の安心・安全、安全保障等の観点からは、**極めて大きなリスクを伴う AI** に関しては、**法規制（ハードロー）による必要最小限の対応も検討していく必要がある。**当 PT においても WG 有志が法的枠組みの試案を報告した。今後は、この試案や諸外国の動向等も踏まえた対応が必要である。

**規律・規制はイノベーションと対立するものではなく、安心・安全な環境の構築によって利活用や研究開発を促進する**ものでもある。第３章では、このような AI の安全性確保に関する戦略を記述する。

**AIホワイトペーパー2024 の骨格**




（AI ホワイトペーパー2024 の骨格と主な提言については、別紙２参照）

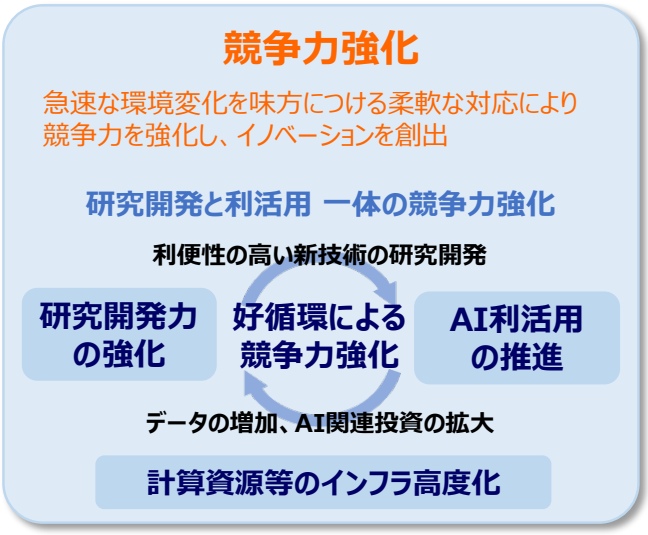
# 第２章　AI を活用した日本の競争力強化のための戦略：急速な環境変化を味方につける柔軟な対応

AI 戦略のプレイヤーは官民の多岐にわたる。

**企業が AI に関する競争力を高め、それぞれの勝ち筋を見出していけるよう、政府は環境整備、人材育成、国際協調**の役割を担う。

**AI を利活用する者と開発・提供する者の双方の競争力強化**に向けて、高い性能の AI を獲得するには、**計算基盤、データ、モデルの連携**が重要となる。（計算資源、データ、モデルの各全体像は、別紙３参照）そのため、AI の**利活用促進、研究開発力強化、インフラ（計算基盤・通信基盤等）の高度化**に、官民連携して取り組み、エコシステムを構築していくことが重要である。

特に、自動車・ロボット・材料開発など**日本の強みを活かせる分野**や、医療・金融・農業など**安全保障上も重要な分野**において重点的に取り組み、**国内外の社会課題の解決等に貢献**していくことが重要である。




その際に必要なことは、刻々と変化する**環境を鋭敏に感知**し、可能な限り正確に将来を**予測し、対応**していくことである。それでもなお、AI を巡る環境はとてつもない速さで大きく変わっていく。**時には抜本的に自らの戦い方を見直す柔軟性**も必要である。**環境変化に屈するのではなく、環境変化を味方につける**のである。

このような認識に基づき、以下、利活用の促進、研究開発力の強化、インフラの高度化について提言する。

## ○利活用の促進

日本では、文章の要約やアイデア出し等を中心に、組織・個人による生成 AI の利活用は拡大している。しかしながら、組織における利活用は、「AI で何ができるか分からない」「AI を導入する体制が整っていない」「リスクや規制への抵触が心配」「AI サービスがまだ不十分」等の課題を抱えている。このままでは、**多くの組織が、DX やクラウド導入の時と同じように出遅れる懸念**がある。

AI の潜在力を引き出し、人手不足等に対応し、AI の利活用を通じて幅広い産業の競争力を強化していくには、上記のような課題への対応が必要である。その際、**行政における率先的な利活用の推進も、先行事例の一つとして有効**である。

各分野における**利活用を通じたデータの獲得やニーズの把握**は、後述する研究開発力強化にとっても必要である。

そこで、以下を提言する。

- **官民で AI の先進的な活用事例や定量・定性的な導入効果を共有**すること。
- 特に、**中央官庁や地方自治体**は、画期的なアイデアを集めるためハッカソンやアイデアソンを**率先して**実施すること。また、行政における更なる利活用を進めるため、機密情報の扱い等を踏まえた**ガイドラインを新たに策定**すること。さらに、活用事例を蓄積し、中央官庁や地方公共団体等への展開を支援するチームを強化すること。
- DX 銘柄への格付けや三位一体の労働市場改革等を通じ、広範な組織において AI 導入を主導する **CAIO（Chief AI Officer）の設置**、人事制度の適正化・早期明確化、レガシーシステム刷新等を進めること。
- 社内外で AI 利活用に向けたスキルの習得に励む人材を支援するため、**習得すべきスキル指針の周知・運用**、**教育コンテンツの充実・共有**、スキルの可視化を様々な業種において進めること。
- AI リスクに対して、リスクベースで事業者が環境変化に迅速に対応できるよう、**「AI 事業者ガイドライン」を速やかに策定・公表**するとともに、これを**広く周知徹底**して、各組織による **AI の適切な利活用を促進**すること。また、各組織におけるリスク対応の事例を、随時、蓄積・共有すること。
- 個人情報保護法、著作権法との関係、リーガル、医療領域をはじめ各種業法で事業者のサービス提供が制限されている分野など AI に関して留意すべき規制の運用について、**事業者が萎縮せずにチャレンジできるよう**、政府は、**明確化やその周知**を図ること。また、事業者からの要望や問合せにスピーディに対応する仕組みを構築すること。必要に応じて、**規制に関わる当事者間の対話**を促進することにより、規制の更な

る明確化を図り、事業者に不必要なリスクをとらせることなくイノベーションを推進すること。

・ デジタル原則に基づくアナログ規制の見直しが行われている中、**規制の AI 対応**を進めるため、必要に応じてサンドボックスのような試行環境等を官民で活用すること。

・ **新たな AI アプリケーションの開発・利活用を推進するため、政府は、「中小企業向け IT 導入補助金」の支援対象サービスへの AI の追加**、「**イノベーションボックス税制**」（AI 関連ソフトウェアの提供等に応じた税制優遇）、スタートアップ等の創意工夫を引き出す全国的な**ハッカソンやアイデアソン**、ニーズが認められる分野別データの整備・更新等に取り組むこと。

## ○ 研究開発力の強化

**AI はいまだ黎明期**である。その潜在力やリスクを踏まえると、**研究開発力の強化によって**、**日本での利活用**（日本の言語や文化・商慣行等）**に合った AI** の開発、**日本が強い分野を通じて世界展開できる AI** の開発、**高度な研究人材の涵養**、**リスク管理**等に取り組むことが重要である。また、研究開発力を持てる日本としては、今後 AI の利活用が進む**グローバルサウス**、特にアジア圏のデジタル基盤の構築に対し、**積極的に貢献**すべきである。

**逆に、研究開発力を持たなければ**、AI 利活用の遅れから、**幅広い産業における競争力の低下**、日本が強い分野の競争力の喪失、大幅な成長が見込まれる AI 提供ビジネスへの参入機会の逸失、**学術界の地盤沈下**、**リスクのブラックボックス化**を招きかねない。

ここ一年で、日本の複数の大企業やスタートアップ、研究機関が独自の LLM を開発・発表したことは明るい兆しである。単にモデルを大規模化して性能を向上させるだけでなく、**小型のモデルの組合せ**や**マルチモーダル化**など、様々な方向性が追求されている。

**米国が先行している大規模汎用モデルのキャッチアップを図りつつ**、民間の創意工夫に基づく**様々な AI モデル・システムの開発**が重要である。また、未成熟だが波及効果が大きく重要な技術は、政府主導での開発も必要である。その際には、予見可能性を持った安定的な研究開発とともに、技術動向の急激な進展に応じた臨機応変な研究開発も可能とするよう、制度上の隘路を改善することも重要である。ただし、競争は激化し、残された時間は少ない。

日本では急速に少子高齢化が進み、労働力不足対応のための**革新的な AI ロボット**等の活用が期待される。実用化に向けて、様々な課題を克服する必要がある。

**AI モデルの開発には、大規模・高性能な計算基盤と、大量かつ多様で良質なデータ**が重要となる。このうち、大規模・高性能な計算基盤には、様々なタイプが存在するが、AI 向

けには、CPU のみならず、GPU 等の AI の計算に適した半導体等を具備していることが重要となる。

また、データについては、AI モデルの開発への有用性を踏まえつつ、WEB 等を通じた公表データ、公的機関の保有データ、民間の保有データといった特性の違いに応じ、利活用を進めていくことが重要である。「データドリブン社会」の到来が指摘され、今般の生成 AI のインパクトにより、それがさらに認識される中、データの保有者には、そのデータを活用していくことの重要性を改めて訴えたい。

そこで、以下を提言する。

- AI 開発に不可欠だが世界的に需給が逼迫している**計算基盤**を、幅広い開発者が利用可能となるよう、 AI への適性を踏まえつつ、**官民で整備・拡充**すること。
- 様々な公的機関が保有するデータを AI 開発に活用できるよう、ニーズを踏まえつつ持続的に**政府等が保有するデータを提供・管理するスキーム**を新たに構築すること。また、WEB から収集した**質の高い日本語データの整備・拡充**を進め、適切な開発者に提供すること。
- **民間データを AI 開発に活用**し、特定分野で競争力を持つ AI を開発できるよう、AI 開発者とデータ保有者の連携に当たってのボトルネック解消に向けた対価還元等の事例の共有や、AI 開発に有用かつ基盤的なデータの新たな整備などを、官民をあげて実施すること。
- 特に、自動車・ロボット・材料開発など**日本の強みを活かせる分野**や、医療・金融・農業など**安全保障上も重要な分野**において、産業の競争力や自律性の確保に向け、幅広い主体が参画して AI の開発・利活用をしっかり進められるよう、**データの収集・整備・更新と AI の開発・利活用の相乗的な取組**を、官民で実施すること。
- **グローバルサウスへの貢献**の観点から、各国の固有の言語・文化に関するデータを日本に集め、学習した AI について、様々なアプリケーションを支えるデジタル基盤として、それぞれの国に還元していくことも官民で検討すること。
- AI 開発者同士でのノウハウ共有や、**グローバルテック企業との交流**、本格的な利活用を志す組織と開発者のマッチングなど、コミュニティ活動を官民で進めること。
- 基盤モデルに関し、国内の産学連携をベースとする開発を政府が支援することにより、その開発力の底上げを図ること。また、モデルの高効率化や高精度化、マルチモーダル化、リスクの低減化等に取り組むスタートアップを政府は重点的に支援すること。このため、自由民主党におけるスタートアップ関係の議論※・提言を踏まえつつ、政府は、**AI スタートアップを対象とした支援プログラムを取りまとめる**こと。

※ 大学発 AI スタートアップの支援、AI と応用分野双方に精通した人材育成、ふるさと納税を活用した AI 研究の拡充など。

- 労働力不足等の社会課題を解決するため、変化する環境に柔軟に対応するなど、現在の AI では実現できない**革新的な AI を搭載したロボット等の研究開発**を、官民で抜本的に強化すること。
- 医療・創薬、材料開発等の分野での日本の強みを活かした **AI for Science** を官民で加速すること。**米国等の有志国と強固に連携した開発体制**や、大学・国立研究機関等のポテンシャルも活用した**産学連携体制**を構築すること。
- AI for Science を含む最先端の AI に係る競争力を飛躍的に強化させるため、政府は、**国立研究機関等におけるデータ基盤の整備**とともに、安定的かつ臨機応変な研究開発を可能とするため、技術進展に応じ柔軟な計算基盤を利用できるよう、クラウド利用料を複数年度に渡って切れ目なく支弁できるような措置を講ずること。
- AI をはじめとする ICT 分野の研究を大学等で行う**若手研究者やポストドクター等について、研究開発等に専念できるよう経済的サポート**を手厚くすることや、独創的なアイデア等を持ち、ビジネスや社会課題の解決をリードしようとする若手人材の活動を支援することなどを通じ、政府は、国際競争力のある人材の育成を強化すること。また、**世界中から高度な研究人材を引き寄せる**ことのできる魅力あるトップレベル研究者の育成・支援を強化すること。

## ○ インフラの高度化

AI を支えるインフラのうち、サーバー及びストレージは、世界需要が 2030 年まで年平均 12.9%増加するのに対し、日本は年平均 15.8%増加し、2030 年単年で約 1 兆円に達する見通しがある。爆発的に増加するデータ処理、求められる処理時間、電力制約等を踏まえ、**国内有数の中心的なデータセンター ～ 利活用地点に近い分散拠点 ～ 無数の端末**といった形で、**計算基盤は通信基盤とともに大規模化・分散化**する見込みである。

その際、消費電力の増加に対応したインフラ全体の省電力化をはじめ、情報処理スピードの高速化や、科学研究等の分野における AI とシミュレーションを高度に組み合わせた利活用等への対応、分散化するインフラ機能や AI 間の連携を支える基盤として期待される超大容量・高信頼・低遅延な情報通信ネットワークの実現等、計算基盤や通信基盤の高度化が求められる見込みである。

こうした**インフラ整備とその利活用・高度化に向けた研究開発**が、世界的な戦略物資となっている**半導体も含め、今後の強靭なデジタル産業基盤の構築**につながる。

データの処理や保存を行う計算基盤は、重要なデータの安全管理、処理時間などサービス品質の向上（利用者の近くでの処理）、緊急時における安定供給等の必要性から、**国内での整備が必要**である。また、計算基盤の運用には高度なノウハウが求められ、社会インフラ化の進展と将来の発展性も踏まえると、**海外に過度に依存せず、専門家人材を確保し、自律性を確保**することも必要である。

そこで、以下を提言する。

- 世界一 AI フレンドリーな国となるためのデータセンター等のインフラを確保できるよう、政府は、**金融面等での政策的支援**を行い、**必要な民間投資を促す**こと。これにより、AI の利活用及び開発・提供双方の競争力強化に向けた礎を築くこと。
- 強靭なデジタル産業基盤の構築に向け、上記インフラ整備と合わせて、政府は、ニーズをしっかり踏まえつつ省電力化・高度化を目指す**新たなコンピュータシステムやネットワークシステム、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、人材育成を支援すること。
- 政府は、「AI 橋渡しクラウド（ABCI）」の拡充・高度化を図るとともに、CPU によるシミュレーション性能だけでなく AI 性能も備えた形で「富岳」の次世代の整備に着手し、AI とシミュレーションを高度に組み合わせた計算ニーズに応える**世界最高水準の AI 利用環境**を実現すること。
- 今後のインフラ整備に必要な電力（特に脱炭素電力）について、政府は、日本の様々な制約の中で**迅速かつ安価な量的確保に向けた最大限の環境整備**を検討すること。

# 第３章　安全性確保のための戦略

生成 AI の普及に伴い、生成 AI の出力の不正確性のリスク（バイアス、ハルシネーション等）、開発者の設定を超えて不適切な出力が引き出されるリスク（プロンプト・インジェクション、ジェイル・ブレイク等）、悪意ある攻撃のリスク（データ・ポイズニング、サイバー攻撃等）、偽画像・偽動画・偽音声が詐欺等に悪用されるリスクも顕在化している。

さらに、著作権をはじめとする知的財産の侵害に関する懸念、金融・医療・交通など AI の誤作動がクリティカルな影響を及ぼす分野でのシステミックリスク、AI の導入に伴う失業者の増加、AI 依存症の増加に対する懸念等も顕在化しつつある。

将来的には AGI（汎用 AI）が実現し、人間が AI を制御できなくなり、破壊的な事態が生じる可能性についても、世界の専門家がすでに警鐘を鳴らしている。

このような中、リスクへの懸念が AI の利活用・開発を躊躇させる要因となってはならない。**リスクを最小化し、安心・安全な環境の下で、過度に委縮することなくイノベーション創出に取り組み、利益を最大化**することが重要である。

リスクへの対応は、**適時に抜かりなく最善の手を打ち、アジャイルに**行っていく必要がある。それによって、世界で最も合理的な AI 規律とイノベーション創出のバランスを確保したガバナンスを日本において実現できる。

AI 技術の急激な変化や多様性に対して、**硬直的な制度による対応では後手**に回るおそれがある。このため、日本においては、**ガイドライン（ソフトロー）に基づく事業者等の自発的な対応**を基本とする。一方、リスクの大きさや諸外国の動向も踏まえ、**必要に応じて最小限の法規制（ハードロー）も適用する柔軟で多層的なアプローチ**によって、安心・安全な環境を整備し、イノベーションを促す必要がある。

生成 AI の登場によって世界的に懸念が拡大している**偽・誤情報対策**に関しては、**対策技術の高度化・実用化**が期待される。また、**AI 開発者や提供者・利用者だけでなく、オンラインプラットフォーマー等も含めて対策**を進める必要がある。

**著作権等の知的財産と AI の関係**についても、**法的な考え方の整理やその周知、関係者間の意思疎通、関連する技術の開発**等が必要である。

そして、日本が**世界で最も合理的な規律とイノベーション促進のバランスがとれた AI フレンドリーな国**となり、**世界から優れた人材や投資**を呼び込むことを目指す。

## ○ 適切なガバナンス

AI ガバナンス（AI 開発者が開発する AI の安全性・信頼性等を保つガバナンスと、AI 提供者・利用者等が AI を適切に提供・利用するためのガバナンスがある）に関する対応が世界各国・地域で急速に進んでいる。**EU では**本年 3 月、AI に関する**包括的な法案**（AI Act）が欧州議会本会議において採択され、EU 理事会で承認された後、成立予定である。**米国では**昨年 7 月及び 9 月に、**開発大手による自主的なコミットメント**が発表され、10 月には**大統領令**が出され、関係省庁による対応が進められている。**中国では**、昨年 8 月に「**生成 AI サービス管理暫定弁法**」が制定された。

昨年 11 月には、**英国で「AI セーフティサミット」**が開催され、それに合わせ、米国及び英国が **AI セーフティ・インスティテュート（AISI）**の設置を発表するなど、**AI の安全性を巡る主導権争いの様相**も垣間見える。

そのような中、**日本は広島 AI プロセスを立ち上げて世界のルール作りを主導**し、昨年 12 月には**世界で初めてとなる包括的な政策枠組み**が合意された。このような日本主導の国際的な取組は高く評価できる。

広島 AI プロセスの国際指針・国際行動規範を、法制度によって遵守するか、自主規制によって実践するか等の制度の選択は、各国に委ねられている。日本は、**技術の変化やリスクの多様性等に対して、リスクベースで迅速・柔軟に対応可能なガイドラインに基づく対応**を選択した。**AI の開発・提供・利用に関する AI 事業者ガイドライン**は、パブリックコメントの結果を踏まえ、まもなく第 1.0 版が策定・公表される予定である。今後、様々な分野への AI 事業者ガイドラインの浸透を図り、それをベースに各業界での対応が進むことが重要である。

**AI 事業者ガイドライン等に基づき事業者等が自発的・継続的にリスクを評価し、低減を図ることが日本の戦略の基本**である。一方、ガイドライン等のリスク低減策の社会実装を進めるため、政府と事業者等との間で密に意思疎通を図り、共働して実効的な取組を行う必要がある。AI が社会において重要な役割を担いつつある中、**AI のブラックボックス化は避けなければならない**場面がある。**国民の安心・安全を揺るがすリスクがある AI** の開発に関しては、欧米の法制度も参考にしつつ、**安全性や透明性に関する必要最小限の法的義務（ハードロー）も必要**である。今年 2 月、**当 PT の WG 有志は「責任ある AI 推進基本法（仮称）」を試案として提案**した。（「責任ある AI 推進基本法（仮称）」の骨子は、別紙４参照）

そこで、以下を提言する。

- **AI 事業者ガイドライン等に基づき事業者等が自発的・継続的にリスクを評価し、低減を図ることを日本の AI ガバナンスの基本**とすること。幅広い業種において、その**周知・浸透、各分野に応じた具体的な実装・実行等の取組を推進**すること。また、技

術やビジネスの変化に応じて、**不断の更新**を行うこと。

- **当 PT の WG 有志による「責任ある AI 推進基本法（仮称）」**の考え方等を踏まえ、政府は、**極めて大きなリスクがある AI モデルに対し、必要最小限の法的枠組みを整備**すること。
- 政府は、法的枠組みの対象や法的義務の内容について、海外の動向や広島 AI プロセスの議論も踏まえ、**十分に検討・検証し、明確化**を図ること。
- 医療・金融・自動運転など、AI の誤動作やシステミックリスクによる影響が大きいと考えられる分野に関しては、政府は、**既存の業法の見直しの必要性等を十分に検討**すること。

## ○ 生成 AI を利用した偽・誤情報対策

生成 AI の高性能化により、あたかも現実であるかのような街並や風景、有名人に酷似した画像・動画等を容易に作成できるようになった。そのようなコンテンツが意図的か否かに関わらずネット上で広く拡散される状況も顕在化している。

AI モデル側での対応も進み、例えば、銃、爆発物、生物化学兵器等の製造方法やテロへの使用といった違法な出力が出ないようなセーフガード措置も進んではいるが、プロンプト・インジェクション等の**悪意ある利用とのいたちごっこ**になっている。

また、生成 AI を利用して画像・動画等を生成・流通する際に、**AI 生成物であることを明示する取組**が進みつつある。コンテンツの**出典・作成者を示す技術の開発・実証**も進められている。偽情報は、一度流通し、広まってしまうと、収拾が難しい面があるが、**ネットに流通している情報の真偽を判定する技術・仕組み**も見られる。災害や事故のような場面での偽情報は、命にかかわる危険性もある。偽情報対策は喫緊の課題である。

このような課題に対応するためには、関係事業者だけでなく、情報の受信者、送信者ともに、幅広い世代・立場の利用者一人一人が、生成 AI の仕組みや生成 AI を利用した偽・誤情報の扱いに対する正確な知識を持ち責任ある行動をとることができるよう、リテラシー向上の取組も一層重要になっている。

AI 生成物に限ったことではないが、**偽・誤情報が民主主義の根幹である選挙に及ぼす負の影響**についても深刻に受け止める必要がある。今年は世界的「選挙イヤー」であり、**AI を用いた選挙妨害**も世界的に懸念されている。

そこで、以下を提言する。

- 生成 AI を利用して生成されるなりすまし動画への対応も含め、**違法（権利侵害）情報の削除の迅速化**や、**プラットフォーマーによる運用状況等の透明化**等のため、政府は、必要な制度整備に取り組むこと。
- **生成 AI を利用した偽・誤情報等のネット上に流通する偽・誤情報について**、政府は、慎重な配慮を払いつつ、制度面も含む**総合的な対策を検討**し、今夏を目処にとりまとめ、必要な措置を推進すること。
- **技術には技術で対処**する観点から、ネット上に流通する AI 生成コンテンツ（画像や動画等）を判別する**技術の開発・実証**等に官民で積極的に取り組むこと。
- 官民連携の取組により、子供から高齢者まで**幅広い層のリテラシー向上**や**ファクトチェック等を推進**すること。
- **AI による選挙への負の影響に適切に対処**するため、関係事業者は、**グローバル企業 20 社**による「2024 年選挙における AI の欺瞞的使用に対抗するための技術協定」（2024 年 2 月、**ミュンヘン・アコード**）**と同様の取組を日本国内でも実施**すること。

## ○ AI の安全性確保に向けた更なる取組

AI の安全性を確保するためには、国際整合性のある評価手法等の検討が重要であり、英米に次いで設置された **AI セーフティ・インスティテュート（AISI）への期待は高い。**

日本の AISI が迅速に立ち上がったことは大いに評価するが、関係機関が協力し、**専門人材の確保など必要な体制整備**を進めることで、**AISI に生命を吹き込み、本格的に国際連携・業務を加速**させる必要がある。

また、安全性確保とイノベーション促進の両立に向け、**AI そのものの安全性を高める基礎的・基盤的な研究開発**も重要である。

さらに、**個人情報やプライバシーの保護に配慮しつつ、データの利活用を適正に促進**することで、**多様な AI サービスを生み出す土壌**をより豊かにしていくことも重要である。その際、AI の開発と応用のためにも、プライバシーやセキュリティ等に関する信頼をしっかりと確保しながら、AI に関連する多様なデータが国際的に自由に行き来する、「ステージⅡ」に相応しい形での信頼性あるデータ流通（DFFT）の具体化について推進することが重要である。

そこで、以下を提言する。

- AI の安全性確保に向けた国際的な協調を図るため、**日本の AISI と諸外国の AISI 等のハイレベルのネットワーク**を構築すること。
- AISI は、 AI の安全性評価に関する**我が国の結節点としての役割**を担うべく、以下の取組を行うこと。
  - **チェックツール**や**レッドチーミングテストの実施手法**等を含め、**AI の安全性評価に必要な調査、基準等の検討**
  - 関係機関の協力のもとで、安全保障、サイバーセキュリティ、AI 技術など**幅広い専門人材の確保や育成、** AI の安全性に関する国内外の研究や国際標準化等の最新動向を含む**先進的な技術的知見の集約・提供**等
  - **官民の様々な組織による AI 利活用等を適切に進められる人材育成（CAIO 等）に資するよう、ドキュメント・教材等の制作**
  - 国際的な整合性に十分留意しつつ、 **オーディット（第三者認証）の在り方についての必要な検討**
- 政府は、AISI による上記の取組を支えるため、**必要な予算・人員を確保**すること。
- AI の脆弱性の悪用や、AI に脆弱性を埋め込む攻撃に対する研究開発、RAG（ Retrieval-Augmented Generation：検索拡張生成）のような外部知識を利用してハルシネーションを防止する技術等の最先端の研究開発を官民で進めること。
- プライバシー保護を強化した新しい AI モデルの開発技術（例えば PETs：Privacy Enhancing Technologies の活用等）を通じて、安全かつ性能の高い AI 分野の発展を官民で推進すること。
- 全ての国民が安心・安全に、多様で魅力的な AI サービスを享受できる環境を整備するため、政府は、個人情報を適切に保護しつつ、過度な委縮が起こらぬよう、規制内容の明確化等を今後も進めること。

## ○ 著作権などの知的財産との関係

生成 AI と著作権の関係については、**文化審議会著作権分科会**において、AI 開発・学習段階、生成・利用段階、さらに AI 生成物の著作物性についての懸念に対する考え方を整理し、**「AI と著作権に関する考え方について」**がとりまとめられた。

また、**AI 時代の知的財産権検討会**では、著作権法以外も含む知財（意匠、商標、不正競争防止法）との関係や、**技術による対応、契約による対価還元の在り方等**について「中間

とりまとめ骨子案」を示し、まもなくとりまとめられる予定である。

これらの検討が短期間で精力的かつオープンに進められ、**法令に基づく解釈を整理・明確化し、現行の法制度を維持**しつつ、今後も**マルチステークホルダーとの意見交換**等を続けていく旨の方向性が示されたことを高く評価したい。また、**法・技術・契約の各手段を組み合わせ**ながら、**生成 AI 技術の進歩と知的財産権の適切な保護が両立するエコシステム**の実現に向けて、AI 事業者等の各主体がアジャイルに取り組む方向性について議論されていることも高く評価したい。今後も、技術の進展等を踏まえ、AI の適切な利用環境を醸成していくため、**引き続き機敏な対応が必要**である。

そこで、以下を提言する。

- **著作権などの知的財産権**については、政府はこれらの権利を尊重しつつ、**AI 時代に即した対応を行い、適切な AI 利活用を促進**すること。
- 特に、「AI と著作権に関する考え方について」や AI 時代の知的財産権検討会の「中間とりまとめ」の周知や啓発を進めるとともに、関係者間のコミュニケーションを通じた相互理解を促進すること。

別紙１

# AI の進化と実装に関するプロジェクトチーム　開催実績（2023 年４月以降）

| No | 日時 | 議題 | 発表者 |
|---|---|---|---|
| | 2023 年 | | |
| 1 | ４月 10 日 | ChatGPT などの AI 利活用について | ・OpenAI, Inc. CEO　Sam Altman |
| 2 | ５月 11 日 | 言語翻訳 AI の進化について | ・DeepL GmbH |
| | | G7 デジタル・技術大臣会合の閣僚宣言における AI 関連のポイント | ・総務省 |
| | | AI 戦略会議について | ・内閣府 |
| 3 | ５月 25 日 | AI 新時代の規制のあり方について | ・BSA（ザ・ソフトウェア・アライアンス） |
| | | G7 広島サミットの報告について（AI 関係） | ・総務省<br>・外務省<br>・デジタル庁<br>・経済産業省 |
| 4 | ６月１日 | 民間における AI の利活用について | ・パナソニック ホールディングス株式会社<br>・Bain & Company, Inc. |
| 5 | ６月７日 | AI 技術とディスインフォメーション対策について | ・国際大学グローバルコミュニケーションセンター 山口真一准教授<br>・xID 株式会社<br>・アドビ株式会社 |
| 6 | ６月９日 | AI のデータ利活用：個人情報保護とデータバイアス | ・個人情報保護委員会<br>・国立国会図書館<br>・独立行政法人国立公文書館 |
| 7 | ６月 22 日 | 自治体業務における生成 AI 等の活用について | ・横須賀市<br>・株式会社 THE GUILD |
| 8 | ７月 18 日 | 広島 AI プロセスの概要と今後の進め方 | ・総務省 |
| | | AI リスクと第三者認証の最新動向について | ・Robust Intelligence, Inc |
| 9 | ７月 27 日 | AI の発展を支えるための計算資源確保の課題について | ・経済産業省<br>・アマゾンウェブサービスジャパン合同会社（Amazon Web Services Japan G.K.）<br>・日本マイクロソフト株式会社（Microsoft Japan Co., Ltd） |
| 10 | ９月７日 | 民間企業における LLM の取組について | ・Meta Platforms, Inc.<br>・日本電気株式会社<br>・株式会社 ABEJA |

| | | | |
|---|---|---|---|
| 11 | 9 月 27 日 | 各国の AI 規制について | ・株式会社国際社会経済研究所<br>・株式会社野村総合研究所 |
| 12 | 10 月 10 日 | Google の責任ある AI の取組について | ・グーグル合同会社（Google Japan G.K.） |
| 13 | 11 月 8 日 | 広島 AI プロセスの報告・今後の予定 | ・総務省 |
| | | アメリカにおける大統領令に関する報告 | ・内閣府 |
| 14 | 11 月 17 日 | AI モデル開発支援について | ・経済産業省 |
| | | AI モデル開発について | ・ソフトバンク株式会社<br>・日本電信電話株式会社（NTT） |
| 15 | 11 月 22 日 | AI 関係経済対策（補正予算）について及び AI 学習データの提供促進に向けたアクションプラン | ・内閣府 |
| | | AI 関係省庁内での生成 AI 活用事例について | ・デジタル庁<br>・経済産業省<br>・総務省<br>・農林水産省 |
| 16 | 11 月 28 日 | 世界の中での日本の勝ち筋について | ・千葉工業大学 伊藤穣一学長<br>・東京大学大学院 工学系研究科 松尾豊教授 |
| 17 | 11 月 29 日 | 顧客企業向け各種 AI サービスの取組について | ・Salesforce, Inc. |
| | | 英国「セキュア AI システム開発ガイドライン」について | ・内閣府<br>・内閣サイバーセキュリティセンター |
| 18 | 12 月 4 日 | NVIDIA の AI 戦略について | ・NVIDIA Corporation CEO ジェンスン・ファン |
| 19 | 12 月 7 日 | 自治体における生成 AI 利活用の取組について | ・大阪市 |
| | | 生成 AI 時代の社会変革 | ・株式会社グラファー |
| | | 自民党提言申入れ　広島 AI プロセス<br>G7 デジタル・技術大臣会合の結果概要 | ・総務省 |
| 2024 年 | | | |
| 20 | 1 月 26 日 | AI 事業者ガイドラインについて | ・総務省<br>・経済産業省 |
| | | NICT が保有する AI 学習用言語データの提供について | ・総務省 |
| | | AI 事業者ガイドラインについてセーフティ・インスティテュート | ・内閣府 |
| 21 | 1 月 30 日 | 最新の製品・サービスの概要、活用事例と今後の開発の方向性について | ・グーグル・クラウド・ジャパン合同会社（Google Cloud Japan）<br>・日本アイ・ビー・エム株式会社<br>・株式会社 Preferred Elements |

| | | | |
|---|---|---|---|
| 22 | 1 月 31 日 | AI に関する欧米諸国の情報法について | ・一橋大学大学院法学研究科<br>生貝直人教授 |
| | | AI に関する諸外国の動向と規制全体像について | ・内閣府 |
| 23 | 2 月 8 日 | 最新の製品、サービスの概要、活用事例と今後の開発の方向性について② | ・アマゾンウェブサービスジャパン合同会社（Amazon Web Services Japan G.K.）<br>・OpenAI, Inc. |
| 24 | 2 月 16 日 | 「責任ある AI 推進基本法（仮）」案について | ・AIPT 有志 WG<br>（殿村桂司弁護士、岡田淳弁護士、<br>生貝直人一橋大学教授、<br>丸田颯人弁護士、<br>小谷野雅晴弁護士） |
| 25 | 2 月 21 日 | AI ガバナンスについて | ・AI ガバナンス協会<br>（大柴行人氏、羽深宏樹氏、<br>生田目雅史氏） |
| 26 | 2 月 29 日 | AI に関する最新の研究開発動向と AI 技術の今後の動向を踏まえた日本の戦略等について | ・東京工業大学 情報理工学院情報工学系 岡崎直観教授<br>・国立情報学研究所 黒橋禎夫所長<br>・理化学研究所 AIP センター<br>杉山将センター長 |
| 27 | 3 月 1 日 | AI に関する計算資源等の取組について | ・産業技術総合研究所<br>辻井 潤一フェロー<br>・さくらインターネット株式会社<br>・理化学研究所　計算科学研究センター（R-CCS）松岡聡センター長 |
| 28 | 3 月 5 日 | 金融、保険業界における AI に関する民間の利用促進について | ・株式会社みずほフィナンシャルグループ<br>・株式会社三菱ＵＦＪ銀行<br>・明治安田生命保険相互会社<br>・東京海上日動火災保険株式会社 |
| 29 | 3 月 7 日 | Government Policy for Advanced AI Systems（先進的 AI システムに対する政府の政策について） | ・東京大学大学院 工学系研究科<br>松尾豊教授<br>・モントリオール大学<br>ヨシュア・ベンジオ教授 |
| 30 | 3 月 8 日 | 民間企業における AI の利活用について | ・株式会社ベネッセコーポレーション、一般社団法人 Generative AI Japan<br>・Degas 株式会社 |
| 31 | 3 月 13 日 | 偽情報対策について | ・富士通株式会社 富士通研究所<br>・国立情報学研究所 情報社会相関研究系 越前功 研究主幹・教授 |

| | | | |
|---|---|---|---|
| | | AI データにおける個人情報保護及び権利関係について | ・個人情報保護委員会<br>・文化庁 |
| | | AI 事業者ガイドラインについて | ・総務省<br>・経済産業省 |
| 32 | 3 月 14 日 | 地方自治体における AI 導入について | ・西川町<br>・AI ガバナンス自治体コンソーシアム |
| 33 | 3 月 21 日 | 欧州における AI 関係検討状況について（報告） | ・内閣府<br>・外務省 |
| | | 民間企業における AI の利活用について | ・デロイト トーマツ コンサルティング 合同会社<br>・弁護士ドットコム株式会社<br>・株式会社 CoeFont |

# AIホワイトペーパー2024の骨格

自民党ホワイトペーパー（2023年）： 全ての始まり。日本のAIのビッグバン。

提言のほとんどは実現、または実施中。
一方で、この一年の世界と日本の動きは、官民問わず他に例をみない速さで進展。

**これからの一年も、誰も具体的で正確な予測ができない「ステージⅡ」**

## ステージⅡにおける新戦略 ー 世界一AIフレンドリーな国へ ー

**競争力強化と安全性確保の一体的な推進**

研究開発、利活用を容易に

### 競争力強化

急速な環境変化を味方につける柔軟な対応により競争力を強化し、イノベーションを創出

**研究開発と利活用 一体の競争力強化**

利便性の高い新技術の研究開発

研究開発力の強化 ／ 好循環による競争力強化 ／ AI利活用の推進

データの増加、AI関連投資の拡大

計算資源等のインフラ高度化

安全性を向上させる新技術

**リスクの最小化 利益の最大化**

安全性の確保がAI利活用、AI研究開発を促進

安心・安全に

### 安全性確保

自発的な対応を基本とする多層的なアプローチ（ソフトローと必要最小限のハードロー）により、安全・安心な利用環境とイノベーションを促す環境を整備

**リスクへの対応**

リテラシー向上、人材育成
著作権・知財、個人情報への対応等

**国際的な連携・協調、国際整合性**

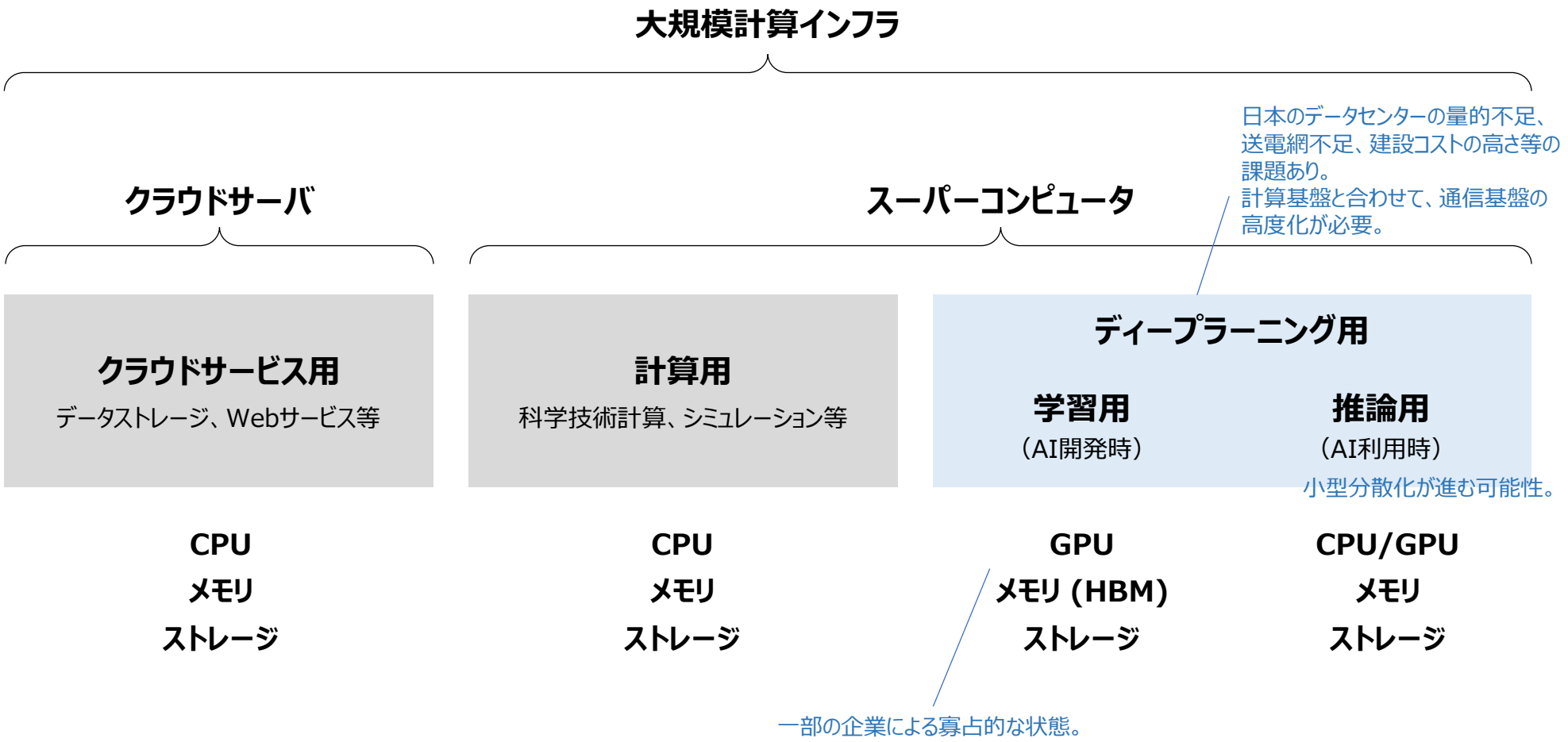


別紙２

# AIホワイトペーパー2024　主な提言

| | | |
|---|---|---|
| 第１章<br>「ステージⅡ」に臨む日本 | ステージⅡの戦略<br>-世界一AIフレンドリーな国へ- | ● **「世界一AIフレンドリーな国」**、すなわち、世界で最もAIに理解があり、AIの研究開発・実装がしやすい国を実現する<br>● AIによる**国民のリスクを最小化しつつ、利益を最大化**する<br>● **競争力強化と安全性確保を一体的に推進**する<br>● 日本は、**広島AIプロセスの実績をベース**に、安心・安全で信頼性のあるAIに関する**国際的なルールメーキングを引き続き主導**する<br>● **アジア諸国やグローバルサウスとの協調関係を強化**するとともに、**AIの国際共同研究や利活用促進**についても、**世界の中で強いリーダーシップを発揮**する |
| 第２章<br>AIを活用した日本の競争力強化のための戦略：<br>急速な環境変化を味方につける柔軟な対応 | 利活用の促進 | ● **行政における更なる利活用**を進めるため、機密情報の扱い等を踏まえた**ガイドラインを新たに策定**する<br>● AIリスクに対して、リスクベースで事業者が環境変化に迅速に対応できるよう、**「AI事業者ガイドライン」を速やかに策定・公表**するとともに、これを**広く周知徹底**して、各組織による**AIの適切な利活用を促進**する |
| | 研究開発力の強化 | ● **データのAI開発への活用**に向け、政府等保有データの提供スキームの構築や、民間データの活用事例の共有、開発に有用なデータの新たな整備等を実施する<br>● 自動車・ロボット・材料開発など**日本の強みを活かせる分野**や、医療・金融・農業など**安全保障上も重要な分野**において、AIの開発・利活用をしっかり進められるよう、**データの収集・整備・更新とAIの開発・利活用の相乗的な取組**を、官民で実施する<br>● 自由民主党における議論・提言を踏まえつつ、政府は、**AIスタートアップを対象とした支援プログラムを取りまとめる**<br>● AI for Scienceを含む最先端のAIに係る競争力の飛躍的な強化のため、政府は、**国立研究機関等のデータ基盤を整備**する |
| | インフラの高度化 | ● 世界一AIフレンドリーな国となるためのデータセンター等のインフラを確保できるよう、政府は、**金融面等での政策的支援**を行い、**必要な民間投資を促す**<br>● **「AI橋渡しクラウド（ABCI）」の拡充・高度化**や、AI性能も備えた**「富岳」の次世代**の整備に着手する |
| 第３章<br>安全性確保のための戦略 | 適切なガバナンス | ● **AI事業者ガイドライン等に基づき事業者等が自発的・継続的にリスクを評価し、低減を図ることを日本のAIガバナンスの基本**とする<br>● 当PTのWG有志による「責任あるAI推進基本法（仮称）」の考え方等を踏まえ、政府は、**極めて大きなリスクがあるAIモデルに対し、必要最小限の法的枠組みを整備**する |
| | 生成AIを利用した偽・誤情報対策 | ● **生成AIを利用した偽・誤情報等**について、制度面も含む**総合的な対策**を今夏を目途にとりまとめる<br>● **選挙への負の影響に適切に対処**するため、関係事業者は、**ミュンヘン・アコードと同様の取組を日本国内でも実施**する |
| | AIの安全性確保に向けた更なる取組 | ● AIの安全性確保に向けた国際的な協調を図るため、**日本のAISIと諸外国のAISI等のハイレベルのネットワーク**を構築する<br>● AISIは、AIの安全性評価に関する**我が国の結節点としての役割**を担う |
| | 著作権などの知的財産との関係 | ● **著作権などの知的財産権**については、政府はこれらの**権利を尊重**しつつ、AI時代に即した対応を行い、**適切なAI利活用を促進**する |

# 計算資源の全体像

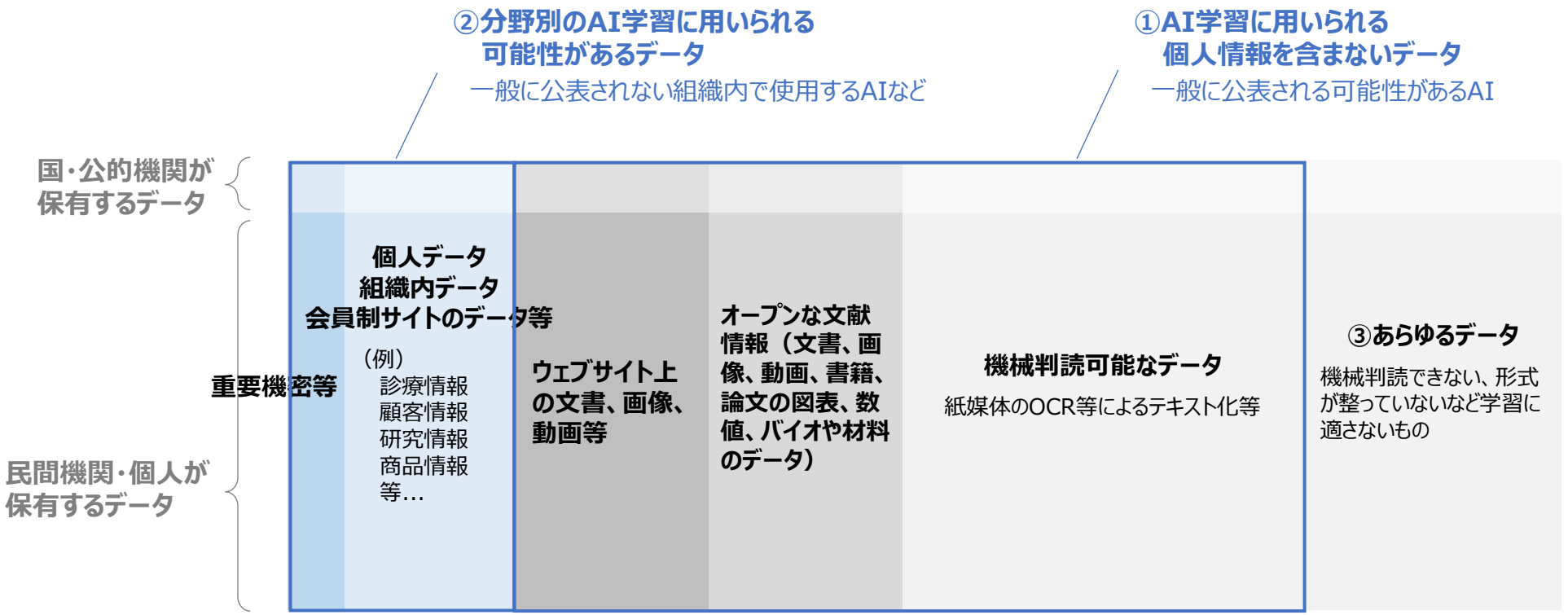





別紙３

# 学習用データの全体像

■ 世の中に存在するデータのうち、公表され、かつ個人情報を含まないデータが、一般的なAIの学習対象（①）。
■ 一方で、非公表データであっても、組織内のみで使用するAIの学習に用いられることがある（②）。
■ ①と②の境界は社会的なコンセンサスによって変化し得る。
■ 不定形の紙の資料など、機械判読できず、AIの学習に用いることができないデータも多く存在する（③）。
　➡ ③をAI学習可能な形に変換するための技術も期待される。
■ 日本語のデータは英語に比べて少ない。日本語の方がトレーニングが難しい。
　➡ 効率的な学習方法の開発が必要。画像、音声などでは言語のような壁が少ない。

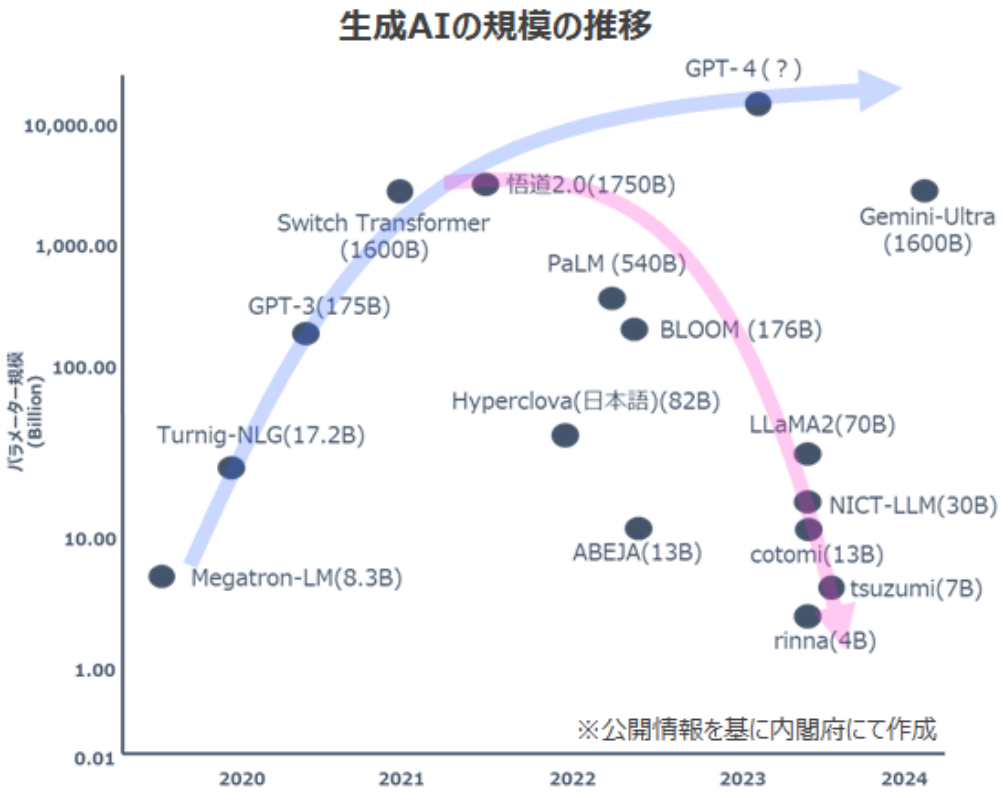



（出典）産業技術総合研究所の資料を基に作成。

# AI・モデルの全体像（2024年３月作成）

## 大規模汎用モデル＆追加学習 と複数モデルの組合せ＆学習効率化

大規模基盤モデルの構築方法はまだ確立されておらず研究段階。




## 分野別モデル、パーソナライズ化、エッジ化

各分野におけるAI導入が進展（〇〇分野×AI、AI for 〇〇）。
パーソナライズ化、各端末・機器への搭載が進む。

## マルチモーダル化

テキストだけでなく、画像、音声、プログラムなど多様なモデルが登場。

AI技術の適用先である製造や医療などに必要な画像、音声や音響などは、日本が質の良いデータを蓄えており、勝ち筋を見出しやすい。

## 外部知識の利用と学習の効率化

再学習することなく新たな知識を取り込み、それに基づく推論を可能にする技術（RAG※1など）や、複数モデルの統合により学習及び推論を効率化する技術（MoE※2など）などの構築を推進すべき。

※1 Retrieval Augmented Generation, ※2 Mixture of Experts

## オープン と クローズ

モデルの詳細を公開しないクローズドな開発手法に対して、オープンソースを指向する動きがある。

オープン化には、多くの人への機会提供、多様性、透明性などの利点があると言われる一方で、悪意ある者による活用、セキュリティ面の不安や知的財産が保護されないという声も。

# 責任あるAI推進基本法（仮）の骨子

自民党AIの進化と実装に関するPT
有志WG（2024年2月16日 資料から抜粋）

## 立法趣旨

**立法趣旨**：生成AIを含むAIの利活用により基本的人権をはじめとする国民の権利利益が侵害されるリスクを最小化しつつ、AIによるイノベーションを含むAIの健全な発展による利益を最大化するため、安全、安心で信頼できる責任あるAIの設計、開発及び導入並びに人間を中心としたAIの利用を可能とするような、開かれた環境の整備を促進する。

## ①責任あるAI利活用の促進

**国**：官民におけるAIの利活用を推進し、社会課題の解決を目指す
施策例：AIの技術革新を推進する官民パートナーシップの構築・強化

**国**：AI人材の育成・誘致と研究開発力の強化
施策例：AIの研究開発のための助成金・補助金等交付

**国**：先進的AIの安全性に関する研究機関の機能強化
施策例：今般創設されたAISIの機能強化

## ②特定AI基盤モデル開発者の指定

**国**：一定の規模・目的のAI基盤モデル開発者を「特定AI基盤モデル開発者」に指定する
論点
- ✓ 「基盤モデル」の「開発者」を規制の対象とする必要性・許容性の整理
- ✓ 「規模」「目的」を何を指標にして評価・区分するか（例：パラメータ数、汎用目的か否か）
- ✓ 指定は一方的に行うか、まず届出をさせるか。一方的に行う場合、指定のための調査権限を国に認めるか
- ✓ 届出すべきであるのに届出しない事業者に制裁するか
- ✓ 適用の地理的範囲（日本で提供されるサービスに「利用」されるモデルに限定するか。）

**民間**：届出義務を課す場合は、対象となる事業者は届出を行う

## ③特定AI基盤モデル開発者の体制整備義務

**国**：事業者に以下の項目を含む体制整備に関する義務を課す
- □ 特にリスクの高い領域におけるAIについては自社・外部による安全性検証（Red team test等）を行う
- □ リスク情報を企業・政府間で共有する
- □ 未公表の重み付けを守るサイバーセキュリティへの投資
- □ 第三者による脆弱性等の検出と報告
- □ 生成AIの利用を利用者に通知する仕組みの採用
- □ AIの能力、限界等の公表
- □ AIがもたらす社会的リスクの研究推進

## ③特定AI基盤モデル開発者の体制整備義務（続）

**民間**：各事業者又は業界団体が上記の義務内容を具体化する規格や行動規範を制定・公表する
論点
- ✓ EU AI Actの整合規格のように民間にAIの品質担保のための規格策定を委ねるか
- ✓ 利害関係者を含めた議論に基づく具体的な行動規範の制定の要否（例：EUデジタルサービス法では、欧州委員会が利害関係者を招請して行動規範を策定している）
- ✓ 民間機関による認証制度等を設けるべきか

## ④義務遵守状況の報告義務と監督

**国**：特定AI基盤モデル開発者に、定期的に③の義務の遵守状況を国または第三者機関（例：AISI）に報告する義務を課す
論点
- ✓ 国への報告にとどまらず対外的な開示まで求めるか

**国及び民間**：国は報告内容に基づき特定AI基盤モデル開発者のモニタリングレビューを行う。国は民間等の利害関係者の意見を聴取することができる

**国**：国は評価の結果を公表するとともに、一定の場合には是正を特定AI基盤モデル開発者に求める

**国**：特定AI基盤モデル開発者が義務を遵守していない場合やインシデントが発生した場合等に報告徴求や立入検査をできる

## ⑤罰則等

**国**：義務・命令違反に対して課徴金・刑罰を科す

**民間**：認証等の取消・一時停止等



別紙４