2026 年 3 月 3 日　全 9 頁

人工知能基本計画が目指す信頼できる AI

「世界で最も AI を開発・活用しやすい国」は実現できるのか

経済調査部　主任研究員　田邉 美穂

[要約]

■ 2025 年 12 月に「人工知能基本計画（以下、基本計画）」が閣議決定され、「信頼できる AI」を中心概念に据えつつ、「世界で最も AI を開発・活用しやすい国」を目指す方針が強く打ち出された。「信頼できる AI」を OECD の AI 原則から整理すると、人間中心の価値観・公平性や頑健性・安全性を重要視していることがわかる。米国や EU の取り組みを見ると、日本とは重要視している項目やアプローチが異なる。日本が引き続き国際協調の枠組みづくりを主導していくためには、評価基準の整合など課題も多い。

■ 基本計画では「信頼できる AI」の実現に向けた日本の強みとして「質の高いデータ」をあげ、医療・研究・産業分野を例示している。しかし、整備や利活用の進展度合いには分野差が見られる。医療・研究分野は、公共性が高く、政府や研究コミュニティが基盤整備を主導しやすいため、データや利用環境が比較的整いやすい。他方、産業分野では、分野・業界を超えた連携の技術的整理が進みつつある一方で、企業間連携や中小企業を中心とする DX の遅れから、現時点で「質の高いデータ」が産業全体に一定程度存在するとまでは言い難い。

■ また基本計画では、AI 関連の開発や投資の出遅れについて問題意識を示している。汎用基盤モデルの国際的上位層は米国や中国が主導し、日本の存在感は現時点で限定的である。さらに、計算資源（GPU クラスター）の国際的偏在や、民間投資規模の差も踏まえると、日本は競争力ある AI モデルを継続的に開発・運用するための前提条件が十分に整っていない可能性が示唆される。

■ 「信頼できる AI」という趣旨に照らし合わせると、日本の文化や習慣に適合した AI 基盤モデルの開発と一定の自律性確保、ならびに信頼性が求められる領域で質の高いデータを用いて社会実装を積み上げるという基本計画の方向性は望ましい。一方で、基本計画が示すデータ整備などの課題解決には一定の時間を要し、実現に向けた具体的な施策や実行力、スピード感については不安が残る。政府は 2026 年春をめどに官民投資ロードマップを取りまとめる方針を示しており、引き続き動向を注視していく必要があるだろう。

株式会社大和総研 丸の内オフィス　〒100-6756 東京都千代田区丸の内一丁目 9 番 1 号 グラントウキョウ ノースタワー

1. はじめに

2025 年 12 月に「人工知能基本計画（以下、基本計画)」が閣議決定された。日本の AI に関する国家戦略として「信頼できる AI」を中心概念に据えつつ、「世界で最も AI を開発・活用しやすい国」を目指す方針を強く打ち出している。その一方で、国内では AI 開発や投資の遅れ、社会実装の不足が重なり、諸外国と比較すると AI 分野で後れをとっている状況にある。基本計画においても、こうした現状を踏まえた反転攻勢の色合いが強く出ているものの、その実現に向けては多くの課題が残されている。

本レポートでは、基本計画の方向性を踏まえながら、日本が今後どのような AI 戦略をとるべきかについて考察する。

2. 人工知能基本計画で目指す「信頼できる AI」

日本政府が目指す「信頼できる AI」とは何か

基本計画では、「信頼できる AI」を目指し、AI 利活用の加速的推進、AI 開発力の戦略的強化、AI ガバナンスの主導、AI 社会に向けた継続的な変革という 4 つの基本方針を掲げている [1]。ここで鍵となる「信頼できる AI」については、基本計画から、価値・倫理・安全性の要件を満たし、適正性が確保された状態の AI を指す概念と読みとれるものの、厳密な定義が明示されているわけではない。ただ、「信頼できる AI」自体は国際的にも広く用いられている概念であり、重視する要素や表現は国や機関により差異があるのが現状だ。

そこで、日本政府が志向する「信頼できる AI」の特徴をみるため、OECD の AI 原則（包摂的成長・持続可能性／人間中心の価値観・公平性／透明性・説明可能性／頑健性・安全性／アカウンタビリティ）[2]に当てはめて整理すると、**図表 1** のようになる。この枠組みに照らすことで、日本の AI 戦略が「信頼できる AI」の実現に向けて、どの要素を重視し、どのような方向性をとっているのかが明確になる。

具体的には、**人間中心の価値観・公平性**の実現にあたり、日本の文化・習慣・現場への適合を重視し、国内の制度・現場で“使える信頼”を確保していこうとする姿勢がうかがえる。また、**頑健性・安全性**の面では、AI セーフティ・インスティテュート（AISI）[3]の機能強化を通じて、評価・基準・監査等の PDCA を具体的に回すことで、“信頼性を担保する仕組み”を構築していこうとする姿勢が、日本の「信頼できる AI」の特徴として示唆される。なお、OECD は政策立案者向けの提言も併記しており、基本計画における施策群はそれらとも整合的な内容が多く、国際協調を意識した設計である点も付言しておきたい。

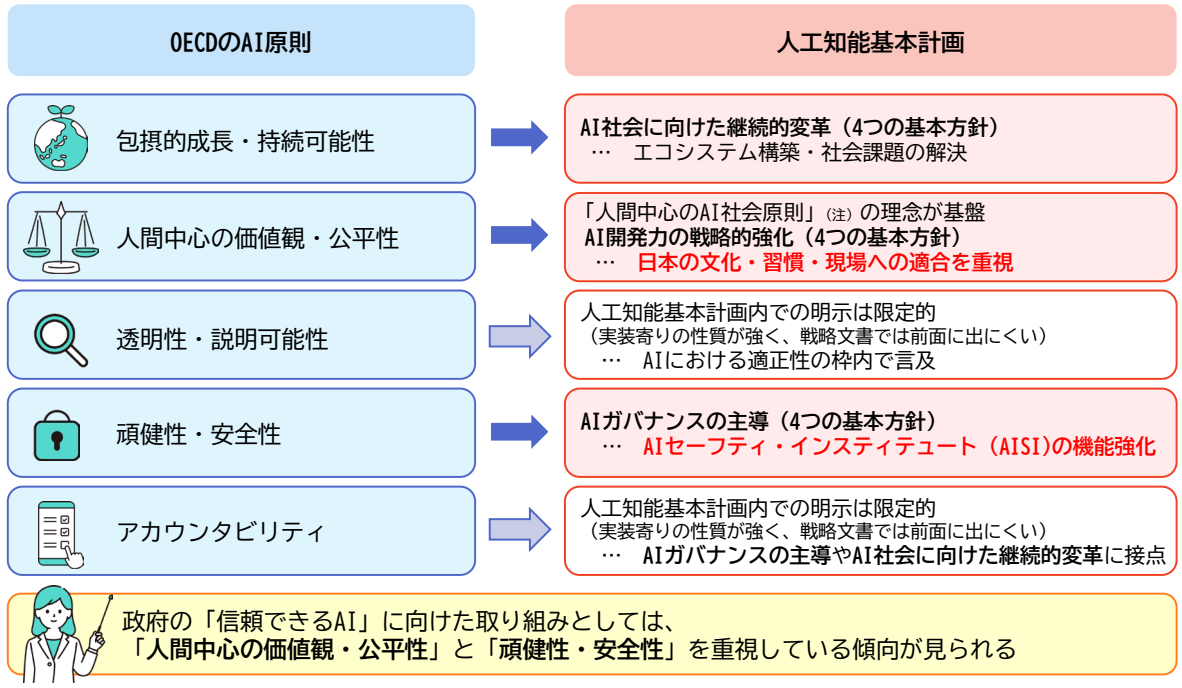
---

[1]基本計画の 4 つの基本方針については、田邉美穂「世界が再注目、実現に動き出すソブリン AI：日本の AI 基本計画が描くデータ主権と技術的自立への道筋」大和総研レポート（2025 年 11 月 19 日）を参考にされたい。

[2] OECD[2024]“Recommendation of the Council on Artificial Intelligence”（OECD/LEGAL/0449）

[3] AI セーフティ・インスティテュート（AISI）とは、信頼できる AI の実現に向けて、AI の安全性に関する評価手法や基準の検討・推進を行うための機関。

図表 1 日本が目指す「信頼できる AI」の特徴



（注）統合イノベーション戦略推進会議決定「人間中心の AI 社会原則」（平成 31 年 3 月 29 日）

（出所）内閣府「人工知能基本計画～『信頼できる AI』による『日本再起』～」（令和 7 年 12 月 23 日閣議決定）および脚注 2 より大和総研作成（イラストはソコスト（https://soco-st.com/））

諸外国における「信頼できる AI」の位置づけと取り組みの方向性

先述の通り、「信頼できる AI」という概念自体は国際的に広く用いられているものの、その具体的な定義や重視する点、政策手段は国・地域によって異なる。そこで、この分野において先行している米国と EU の動向から日本の AI 戦略を相対化する。

米国では、イノベーションの促進と実装の加速を重視する姿勢をとっており[4]、「信頼できる AI」についても、一律の包括的規制よりガイドラインや枠組み（ガイダンス）を通じたリスク管理を軸としつつ、企業や機関の自主的な取り組みを促すアプローチが中心となっている[5]。OECD の AI 原則に当てはめると、**透明性・説明可能性**や**アカウンタビリティ**といった要素を重視しつつ、**包摂的成長・持続可能性**の観点では、民間主導のイノベーションを通じた経済成長への寄与が強く意識されていると考えられる。全体として、AI の安全性確保は重要視されているものの、規制による事前統制よりも、事後的な是正や責任の明確化を通じて対応していく方向性がうかがえる。

これに対し EU では、AI を社会に実装する上でのリスクをより強く意識し、法制度を通じた信頼性の確保を重視する傾向が見られる。EU AI 法（AI Act）に象徴されるように、用途やリスクの程度に応じたルールを明確に定め、事前に一定の要件を満たすことを求める枠組みが整

[4] THE WHITE HOUSE “REMOVING BARRIERS TO AMERICAN LEADERSHIP IN ARTIFICIAL INTELLIGENCE”, January 23, 2025.

[5] ガイドラインの例としては、National Institute of Standards and Technology (NIST) “AI Risk Management Framework”, January 2023. 等を参照されたい。

備されつつある[6]。OECD の AI 原則との対応関係で見ると、特に**人間中心の価値観・公平性**や**頑健性・安全性**を重視しており、基本的人権の保護や社会的影響への配慮を制度的に担保しようとする姿勢がうかがえる。透明性や説明可能性についても法的義務として位置づけられる場面が多く、信頼性を遵守すべき要件として捉えている。

このように、どちらも OECD の AI 原則に掲げられた要素は広くカバーしているものの、米国はイノベーション重視・ガイダンス中心、EU はリスク管理重視・法規制中心と、アプローチには違いが見られる。日本のアプローチは、ガイドラインやフレームワークを通じた実装の後押しという点では米国に近く、人間中心の価値観・公平性の強調という点では EU に通じる側面を持つが、いずれかに単純に重なるわけではない。

また、日本は 2023 年の G7 広島サミット以降、「広島 AI プロセス」などを通じ、国際協調の枠組みづくりを主導してきた。この国際協調の観点でみると、概ね OECD の AI 原則が共通言語として機能し得るため、相互運用のための共通土台は一定程度整っているといえる。しかしながら、アプローチが異なる中で相互運用性を確保するためには、評価基準の整合や制度・枠組みの互換性の整理をはじめ、課題は多い。

3. 「世界で最も AI を開発・活用しやすい国」は実現できるのか

ここからは、基本計画における日本の AI 戦略が実効性を持つかを検討するため、まずは基本計画があげる日本の強み・弱みをもとに、日本の現在地を把握する。

日本の強みである「質の高いデータ」とは

基本計画では、日本の強みとして「質の高いデータ」をあげている。医療・研究および産業のデータが具体例としてあげられている。これらについてみると、データ整備が進む分野がある一方、まだ課題がうかがえる分野も存在することがわかる。

医療分野では、厚生労働省の NDB（National Database）が、保険診療のレセプトと特定健診データを全国規模で継続収集し、毎年「NDB オープンデータ」として公表しているなど、データを提供する環境は既にいくつか存在する[7]。「質の高いデータ[8]」について、この NDB の例で考えてみると、出所が明確である点で信憑性が高いといえる。また、制度改正等に伴い項目の追加や変更は避けられないものの、同一観点で継続的にデータが収集されていることは、長期的な観点での正確性や一貫性、最新性といったデータ品質の基本的な要件を下支えする要素となる。さらに、データ利用の観点においても、オープンデータとして公開されており、アクセシビリティを高めている。このことから、医療分野においては、日本には質の高いデータが存

[6] EUR-Lex “Artificial Intelligence Act” （2024 年 7 月 12 日）
[7] 厚生労働省ウェブサイト「【NDB】NDB オープンデータ」（2026 年 2 月 18 日最終閲覧）
[8] 本稿における「質の高いデータ」は、データ品質モデルの国際規格 ISO/IEC 25012 に基づく整理であり、基本計画における用法と必ずしも一致しない可能性がある点に留意されたい。

在し、その活用基盤も一定程度整っているといえるだろう。

研究分野においても、国立情報学研究所が 2021 年から NII 研究データ基盤（NII Research Data Cloud、NII RDC）の本格運用を開始している[9]。オープンサイエンスと研究公正を支え、データ駆動型研究を推進する情報基盤として、研究データのライフサイクルに沿って管理基盤・公開基盤・検索基盤の３つの基盤を提供している。「質の高いデータ」の観点でみると、研究データ自体は研究者や研究プロジェクトの裁量で管理されるものの、データガバナンス機能として、データの品質を機械的に検証する仕組みなどが整備されている。また、データ利用の観点においても、検索基盤によるアクセシビリティの確保はもちろん、公開範囲や公開時期の設定、秘匿解析機能など機密性にも配慮がされている。このことから、研究分野においても、研究データを質の高い形で管理・公開・利活用していくための基盤は一定程度整いつつあるといえるだろう。

産業分野については、分野や業界を越えたデータ利用と個別の業界や企業の現場でのデータ利用の２つのレイヤーが想定される。まず、分野や業界を越えたデータ利用については、DATA-EX やウラノス・エコシステム[10]など、分野・業界を越えたデータ連携に向けた取り組みが進められてきた。2025 年 10 月には、これらの国内の主要なデータスペースの技術的取り組みを取りまとめる共通仕様として「Open Data Spaces」を共同で推進することが合意され、技術面での整理が進みつつあるといえる[11]。一方で、現時点では、プロジェクト制度や実証事業を通じて社会実装を進める段階にとどまっており、医療・研究分野と比べると、質の高いデータを利活用していく基盤が整っているとは言い難い。

次に、個別の業界や企業の現場でのデータ利用についてみると、企業間連携に限ってみても、日本は「他社とのデータ連携やデータ提供は行っていない」とする割合が高く、米国やドイツとの間に差が見られる（**図表 2**）。

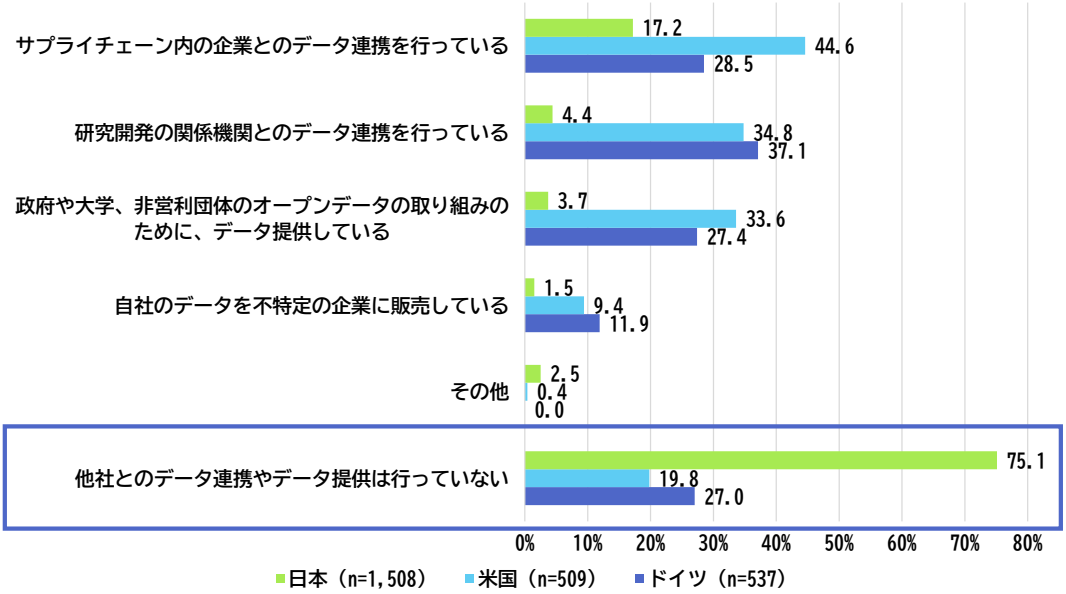
---

[9] 国立情報学研究所オープンサイエンス基盤研究センター「NII 研究データ基盤（NII Research Data Cloud：NII RDC）」（2026 年 2 月 18 日最終閲覧）

[10]データ社会推進協議会「『DATA-EX』とは」（2026 年 2 月 18 日最終閲覧）、経済産業省「Ouranos Ecosystem（ウラノス・エコシステム）」（2026 年 2 月 26 日更新）

[11] 情報処理推進機構、データ社会推進協議会、ロボット革命・産業 IoT イニシアティブ協議会、東京大学大学院情報学環「プレス発表 データスペースの技術コンセプト『Open Data Spaces』の共同推進を合意」（2025 年 10 月 15 日）

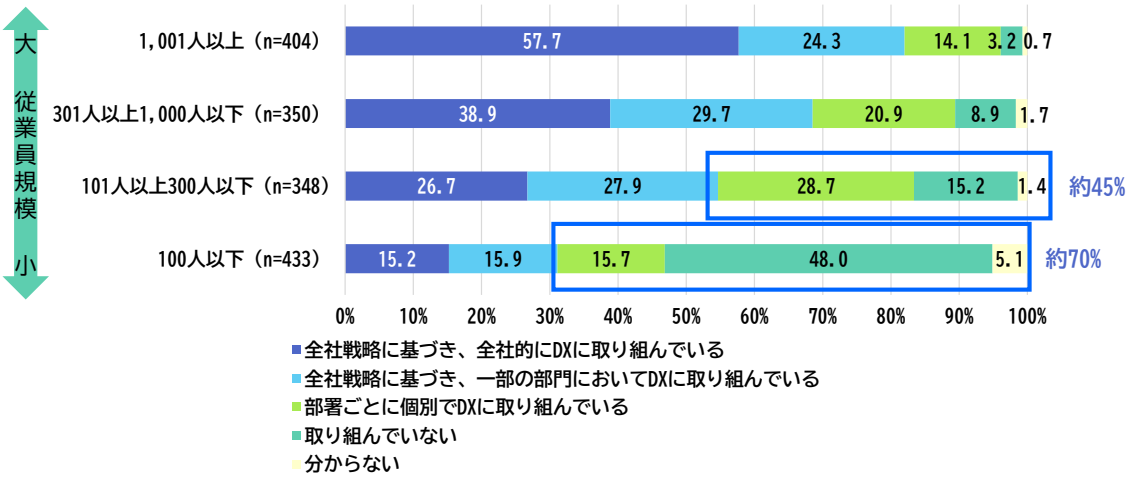
図表 2　データの企業間連携の状況（2024 年度）



（出所）情報処理推進機構「DX 動向 2025：日米独比較で探る成果創出の方向性『内向き・部分最適』から『外向き・全体最適』へ」（2025 年 7 月 9 日更新）P. 34 より、大和総研作成

さらに、各企業でのデータ利用は、アナログ情報のデジタル化や社内データの統合といった取り組み（デジタイゼーション）を起点に段階的に進むため、企業の DX 取り組み状況を手がかりに現在地を確認する。**図表 3** に示す従業員規模別の DX 取り組み状況を見ると、従業員規模が小さい企業ほど DX への取り組みが限定的である傾向がうかがえる。日本では企業数ベースで中小企業が大部分を占めることを踏まえると、産業分野全体としては、利活用を前提とした質の高いデータの水準まで整備が十分に進められている企業は必ずしも多くない可能性がある。また、こうした企業規模による取り組み状況の差は、AI 利活用に要する人材・投資余力・運用ノウハウの面で成熟度のばらつきが影響していると考えられる。

図表 3 DX への取り組み状況（従業員規模別・2024 年度）



（出所）情報処理推進機構「DX 動向 2025：日米独比較で探る成果創出の方向性『内向き・部分最適』から『外向き・全体最適』へ」（2025 年 7 月 9 日更新）P. 3 より、大和総研作成

このように、基本計画で「質の高いデータ」としてあげられた分野を見ると、実際のデータや利活用環境の整備度合いには差が見られる。医療・研究分野のように公共性が高く、政府や研究コミュニティが基盤整備を主導しやすい領域では、信頼できるAIに向けたデータや利用環境が比較的整いやすい。一方、産業分野では対象範囲が広いことに加え、企業が保有するデータ自体がAIの登場により競争力の源泉として重要性を増していることもあり、業界横断/業界別にデータを共有・連携できる環境の整備は容易ではないと考えられる。さらに、企業間連携の実態や企業規模が小さいほどDXへの取り組みが限定的である状況を踏まえると、現時点で「質の高いデータ」が、産業全体に一定程度存在するとまでは言い難い。したがって、産業分野については、基本計画が示す強みが十分に成立しているかは不確実であり、「質の高いデータ」が乏しい可能性も排除できないことを踏まえて慎重に捉える必要がある。

AI関連の開発・投資の出遅れに課題

また基本計画では、弱みとまでは明示していないものの、AI関連の開発や投資が出遅れているとの問題意識を示している。この「AI関連の開発・投資の出遅れ」を、①AI基盤モデル開発、②計算資源・インフラ、③投資規模の3点に分解して整理する。

1つ目のAI基盤モデル開発について、例えば、近年のAI競争を象徴する大規模な汎用AIモデルの動向をみると、性能競争における国際的な上位層は米国や中国のプレイヤーが主導している状況にある。また、直近では中国の台頭によりオープン型AIモデル（技術情報等が公開されているモデル）にも勢いが出てきている。それにより、フランスや韓国など米中以外の国・地域で、汎用AIモデルの分野において国際的に存在感を示すモデルが登場するなど、諸外国においても国産モデルの競争力強化がうかがえる。こうした状況の中、現時点で日本は性能指標や論文など国際的に広く参照される場での存在感が限定的であることは否めず、国際的に見て主導的な位置にあるとは言い難い。また、フィジカルAIに利用されるロボット基盤モデルについても、米中をはじめ海外においては実証実験等の取り組みが先行しており、この分野においても現時点で日本は存在感を示せてはいない。

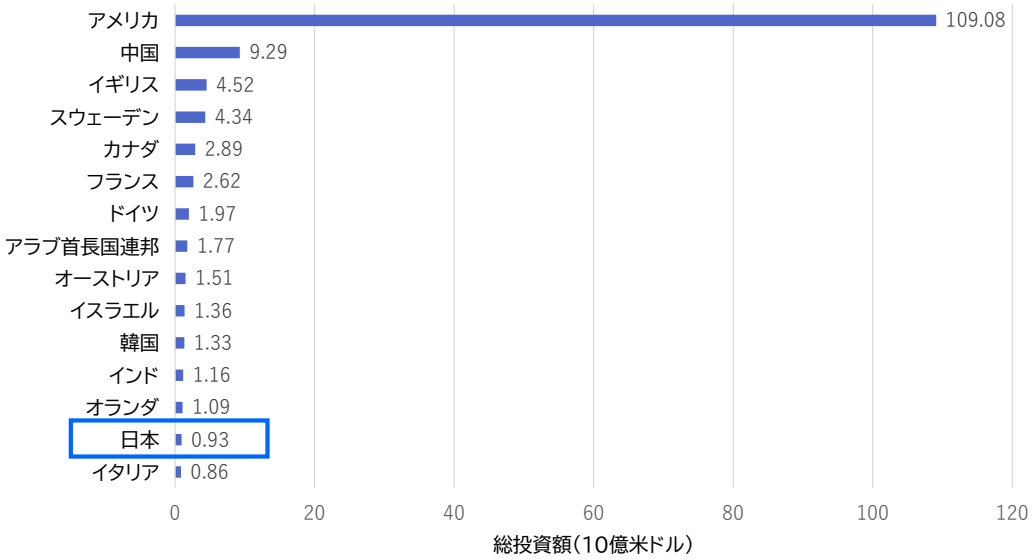
2つ目の計算資源・インフラについて、基本計画はAIインフラの構成要素として、計算資源／半導体／データセンター・クラウド／通信ネットワーク／安定的な電力供給をあげている。このうち、AIモデルの開発・運用において大前提となる計算資源に着目してみると、国際的にはGPUクラスター（AI計算資源）が米中に集中しているとの推計もある[12]。先述のAIモデルの開発状況を踏まえても、日本が競争力のあるAIモデルを継続的に開発・運用するうえで必要となる計算資源を十分に確保できていない可能性が示唆される。

3つ目の投資規模については、AIの民間投資額に着目すると、日本は諸外国と比べて投資規

[12] Konstantin F. Pilz, Robi Rahman, James Sanders, Luke Emberson, Lennart Heim “The US hosts the majority of GPU cluster performance, followed by China” in EPOCH AI, June 5, 2025.にて、国別のGPUクラスター性能シェアについて「2025年5月時点で米国が約4分の3、中国が約15%」（翻訳は大和総研による）とされている。ただし、同データセットは世界全体のGPUクラスター性能の推計10～20%程度をカバーするにとどまる旨が明記されており、推計値の解釈には留意が必要である。

模が小さい水準にとどまっている（**図表 4**）。こうした民間投資規模の差は、先述した AI 基盤モデルの開発や計算資源・インフラ整備の進展度合いとも無関係ではないと考えられる。このため、日本が競争力のある AI モデルの開発や利活用を進めていくうえでは、民間投資を補完および促進する観点から、国としての関与や投資が一定程度必要となる。実際に基本計画に関連して、政府として当面 1 兆円超を AI 関連施策の推進に投資する方針が示されている[13]。

図表 4 AI に対する民間投資額の国際比較（地域別、2024 年）



（出所）Stanford University, Human-Centered Artificial Intelligence (HAI) “Artificial Intelligence Index Report 2025”, 2025, p. 252 より大和総研作成

日本政府が人工知能基本計画で示した方針は妥当なのか

ここまで日本の強みと弱みを整理した結果、日本には AI モデル開発において性能向上の重要な前提となり得る「質の高いデータ」が医療・研究分野で存在する一方、フィジカル AI を含め政府が AI 活用に期待を寄せる産業分野では、データが十分に利用できる形に整備されていない可能性があると言える。加えて、政府は大規模な投資方針を示しているものの、整備に時間を要する計算資源やインフラに遅れが見られるなど、強みを十分に活かすための準備が整っているとは言い難い。

これらを踏まえて、基本計画が掲げる「信頼できる AI」に向けた方向性が妥当なのか考察する。日本の文化や習慣に適合した AI 基盤モデルを開発し、一定の自律性を確保する方向性は、近年注目されるソブリン AI の潮流とも重なり、主権確保や安全保障の観点から一定の合理性がある。また、医療・研究分野や産業分野といった信頼性が求められる領域で、AI の性能向上が期待できる質の高いデータを用いて、当該分野に特化した AI 開発や利活用による社会実装を積み上げることは、基本計画が掲げる「信頼できる AI」の趣旨と整合的である。こうした領域は、例えばフィジカル AI も含めた産業分野での AI 利活用を通じて、労働力人口の減少による人手

[13] 首相官邸「人工知能戦略本部」（令和 7 年 12 月 19 日更新）

不足などの社会課題の解決に直結し得る分野でもある。そのため、信頼性を担保したAIの実装を進めることは、日本の社会課題の解決に資するのみならず、同種の課題を抱える国・地域への展開を通じて、国際的な課題解決にも貢献し得るという、基本計画の問題意識にも沿う。

他方で、実現に向けてはいくつか課題も見えてくる。基本計画はデータの集積・利活用・共有を促進し、データ連携基盤の構築を掲げているが、データを利活用可能な形に整備すること（品質管理・標準化・権利処理・メタデータ[14]整備等）を、どの主体がどのような仕組みで進めるかについては、現時点では具体化の余地が残る。先述の通り、産業分野では利活用を前提としたデータ管理の水準までの整備が不十分な企業も多いと考えられる。また、データ整備は企業の裁量が大きく、短期的には費用対効果が見えにくいため、投資が後回しになりやすい。したがって、連携基盤の整備だけでなく、ユースケース（活用事例）の創出、標準・品質の共通土台、ルール整備、投資インセンティブ等を組み合わせ、データ整備が進む環境を官民で設計することが重要となるだろう。

さらに、基本計画は国際協調の下でAIガバナンスを主導する方針を掲げているが、こうした対外的なリーダーシップを実効性ある形で確立するには、国内における「信頼できる AI」の具体的な実装実績と、それを支える評価体制の整備が前提となる。他方で、計算資源・インフラ整備、AISIによる評価体制の強化などはいずれも整備に一定の時間を要する。AIをめぐる競争環境の変化が速い中で、社会実装のタイミングが競争力に与える影響には留意が必要であり、フィジカルAI等の比較的新しい分野においても差が一段と拡大する可能性は残る。また、日本が早期に「信頼できる AI」の国際的プレゼンスを確立できなければ、国際協調やルール形成の場で主導的役割を果たすことが難しくなる。

このように、基本計画で目指す方向性は妥当であると思われる一方、実現に向けた具体的な施策や実行力、スピード感については不安が残る。政府は2026年春をめどに投資目標やデータ戦略等を含む官民投資ロードマップを取りまとめる方針を示している。今後もこれらの投資の効果が十分に顕在化しないリスクも含め、動向を継続的に注視していく必要があるだろう。

以上

[14] メタデータとは、データの特性や構造など、データを説明するためのデータを指す。