

# LLMの世界と日本のAI格差に関する調査報告書

## Executive Summary

本報告書は、指定YouTube番組におけるPreferred Networks（PFN）代表取締役社長・岡野原大輔氏の議論を主要出典の一つとし、2026年9月8日時点で入手可能な政府資料、企業公式発表、研究機関資料、Stanford AI Index等を突合したものである。動画のタイムスタンプは提供された文字起こしに基づく。文字起こしは自動生成を含むため、固有名詞が曖昧な箇所は本報告書では「未特定」とした。[filecite](#)[turn0file0](#)

結論を先に述べると、「日本は世界のLLMから約1.5年遅れている」という一本の数字だけでAI格差を表現するのは適切ではない。岡野原氏の「1～2年、感覚的には1.5年」という評価は有力研究者・経営者による専門家推定として重要だが、世界共通の測定尺度によって検証できる数字ではない。一方、日本政府自身が「AIの積極的な利活用に至っておらず、AI関連の開発・投資でも出遅れている」と認識しており、**フロンティア汎用モデル、民間資本、計算資源、実験回数、グローバル人材、事業化速度では明確な格差が存在する。** <sup>1</sup>

しかも世界の競争構造は、単なる「巨大なTransformerを一回学習する競争」から変化している。2025年には「notable AI models」の90%以上が産業界から生まれ、世界のAI計算能力は2022年以降年率約3.3倍で増え、17.1百万H100相当へ達したとStanford AI Indexは推計する。データ品質、post-training、強化学習、エージェント、評価・検証を大量に回す能力が競争力を左右する比重も高まっている。 <sup>2</sup>

動画の中心命題である「**強いAIを使って、次の強いAIを作るループが回り始めた**」は、限定された意味ではかなり裏付けられる。Google DeepMindのAlphaEvolveはLLMと自動評価器を組み合わせ、Googleのデータセンター運用、TPU設計、AI学習そのものを改善し、Geminiで使われる行列演算カーネルを23%高速化して学習時間を1%削減した。これは「AIがAI開発を加速する」明確な実例である。 <sup>3</sup> ただし、これを「自律的な再帰的自己改善（RSI）が既に成立した」と解釈するのは行き過ぎである。OpenAIは2026年2月のGPT-5.3-Codex System Cardで、AI自己改善能力は同社Preparedness Frameworkの「High capability」水準には達していないと明記していた。したがって、**現状は「AI支援型R&Dループの高速化」であって、無人で能力が自己増殖する強い意味のRSIが実証済みという状態ではない。** <sup>4</sup>

一方、米中の差は動画の認識以上に縮小している。Stanfordの比較可能な2026年3月スナップショットでは、米国トップと中国トップのモデル性能差は2.7%にまで縮まり、Anthropic、xAI、Google、OpenAIの上位4社はArena Eloで25点以内に集中した。中国は2025年のAI論文数・引用・特許付与で世界を牽引し、米国は「notable models」で59件、中国35件だった。 <sup>5</sup> 2026年にはDeepSeek V4、Alibaba Qwen3.8、Moonshot Kimi K3など、中国勢が引き続き大規模モデルを高速投入している。DeepSeek V4 Previewは1.6兆パラメータ／49B activeのProと284B／13B activeのFlashを公開し、1Mコンテキストとオープンウェイトを打ち出した。 <sup>6</sup>

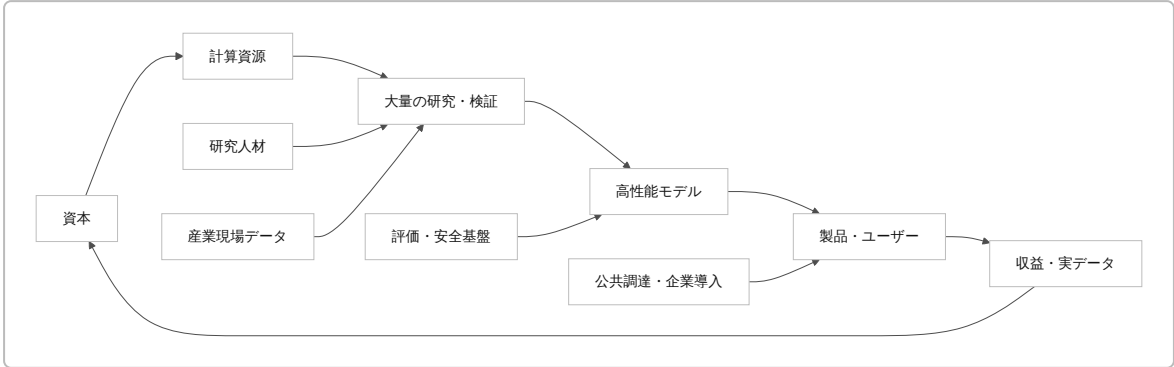
しかし、「中国は全部オープン」という動画中の表現は文字通りには正しくない。DeepSeek V4やQwenの多数の系列にはオープンウェイト版がある一方、AlibabaにはAPI提供を中心とするフラッグシップも存在し、中国も**open / closed / API-onlyを使い分けるハイブリッド戦略**へ進んでいる。 <sup>7</sup>

日本について最も重要な変化は、「**米国の汎用LLMを同じやり方で追う**」ことから、「**日本が優位な産業データ・現場・ロボティクスを使ったPhysical AIへ重点を移す**」国家戦略が具体化したことである。経済産業省・NEDOは2026～2030年度にAIロボット／Physical AI向けマルチモーダル基盤モデル開発を進め、Noetraと産総研を実施主体として選定した。NoetraにはVera CPU 13,750基、Rubin GPU 27,500基、140MW

級のAI基盤が計画され、その事前学習済みモデルを国内開発者・企業へ広く提供する構想が公式に示されている。<sup>8</sup>

これは合理的な選択である。Stanford AI Indexによれば、2026年時点でもロボットはシミュレーション上のRLBenchでは89.4%の成功率に達する一方、現実の家庭作業では成功率が12%にすぎず、Physical AIには巨大な未解決領域が残る。<sup>9</sup> 日本には製造・自動車・ロボット・工作機械・物流・高信頼制御の現場があり、政府も製造業データのAI-Ready化とロボット基盤モデルをGENIACで支援している。<sup>10</sup>

ただし、この戦略が成功する条件は厳しい。「GPUを買うだけ」では差は埋まらない。本当に重要なのは、計算資源を「年間何件の大規模実験・失敗・再学習に変換できるか」、現場データを「合法・安全・標準化された学習データに変換できるか」、研究成果を「世界市場で売れる製品に変換できるか」である。日本の最大の問題は単一の技術不足ではなく、次のループが弱いことである。



米国ではこの「資本→compute→研究→製品→売上・データ→再投資」の循環が極めて大規模で、2025年の米国民間AI投資は2859億ドルに達した。対して日本は、AI・半導体全体で2030年度までに10兆円超の公的支援を行い50兆円超の官民投資を促す枠組みを整えたが、**公的資金を、世界市場で再投資可能な民間売上へ変換できるかが次の勝負**である。両数字は範囲が異なるため直接比較はできないが、競争の資本規模を理解するうえで重要である。<sup>11</sup>

したがって、日本の目標は「GPT-6を完全コピーすること」ではなく、**汎用フロンティア技術へのアクセスを失わない最低限の主権能力を確保しつつ、日本語・業務・製造・ロボット・エッジ・高信頼AIで世界の不可欠なレイヤーを取る**ことに置くべきである。完全自給は非現実的だが、モデル、GPU、クラウド、電力、データのどれか一つが遮断されただけで国家機能が停止する状態も避ける必要がある。Stanfordも「AI sovereignty」が世界の国家AI政策の中心概念になりつつあると報告している。<sup>12</sup>

## 指定動画の要約と主要論点

動画は「世界と日本のAI格差」をテーマに、フロンティアAI、Noetra／国産Physical AI、人間とAIの役割という三つの領域を約59分にわたって扱っている。以下の時刻は提供された文字起こしに準拠する。

filecite0turn0file0

| タイムスタンプ     | 主な内容                                                      | 本報告書での位置づけ |
|-------------|-----------------------------------------------------------|------------|
| 00:00–01:02 | テーマは「世界と日本のAI格差」。PFN岡野原大輔氏が出演。前回は日本と最先端との差を1～2年程度と評価していた。 | 格差の出発点     |

| タイムスタンプ     | 主な内容                                                                                               | 本報告書での位置づけ                                    |
|-------------|----------------------------------------------------------------------------------------------------|-----------------------------------------------|
| 01:02–03:19 | 「RSI（再帰的自己改善）」的なループ、すなわち強いモデルを次のモデル開発へ投入する循環が本格化したとの主張。ソフトウェア開発、強化学習、合成・高品質データ、小型モデルへの能力移転が重要。     | <b>部分的に裏付けあり。強い意味のRSIとは区別が必要</b>              |
| 03:19–04:25 | AI研究・開発のうちパラメータ調整、ソフトウェア、データ処理などがかなり自動化されつつある。                                                     | AIによるR&D加速                                    |
| 04:25–05:27 | Kimi、Qwen、GLM等の中国モデルが競争し、公開された技術を互いに素早く取り込んでいる。「中国はオープン」という評価。                                     | 中国の高速なopen-model ecosystemを指摘。ただし「全部オープン」は要修正 |
| 05:27–06:34 | 日本と世界最先端の差は依然1～2年、感覚的には約1.5年。ただし一定の能力水準を超えると最先端側が加速する恐れ。                                           | <b>専門家推定。客観的な年月尺度では未検証</b>                    |
| 05:27–07:40 | 米国フロンティアラボは最強モデルを公開せず、内部で次世代モデル開発に使っている可能性。サイバー安全性や推論コストも理由ではないか、と推測。                              | 本人も推測と明言。 <b>未特定・外部検証不能</b>                   |
| 07:40–10:00 | 単純なモデルサイズ拡大だけでなく、post-training、強化学習、試行錯誤、検証コストが重要化。多数の大規模実験をできる組織が有利。                              | 現在の研究動向と整合的                                   |
| 10:00–14:35 | open vs closed。open-weightはローカル利用・防御側にも利点があるが、安全機構を外して悪用できる問題がある。OpenAIとHugging Faceに関するサイバー事例に言及。 | 一般論は妥当。 <b>具体的Hugging Face事例は未特定</b>          |
| 14:35–16:51 | AI研究を減速させる議論に一定の理解。ただし世界全体で停止させるのは困難。社会適応の時間を確保する意味はある。                                            | 政策・安全性に関する価値判断                                |
| 16:51–19:09 | モデル競争は能力だけでなくコスト・電力・インフラへ。モデル自体が無料でも推論コストはゼロにならない。                                                 | 強く裏付けられる                                      |
| 19:09–21:35 | すべての用途にフロンティアモデルが必要なわけではない。一方、日本が自らコントロールすべき用途は国内能力を保持すべき。Physical AIはLLMのGPT-2～GPT-3期程度という比喩。     | 「Physical AIは未成熟」は裏付けあり。GPT世代との対応は比喩          |
| 21:35–25:56 | Noetraの「母モデル」構想。推論・マルチモーダル基盤モデルを作り、国内企業が派生モデルを構築する。世界で未解決のPhysical AIを狙う。                          | 公式FRONTia/Noetra構想と整合                         |
| 25:56–30:26 | 2030年前後までの計画だが毎年見直す。開発本体と探索的研究を並列化し、約20チームが研究しているとの説明。                                             | 長期・アジャイル型開発。 <b>「約20チーム」の公式な対応内訳は未特定</b>      |
| 30:26–34:07 | 日本企業が単独で汎用基盤技術へ投資し続けるのは商業回収が難しい。国・複数企業で協力する合理性がある。国産でも国際連携は必要。                                     | 日本の資本構造に重要な指摘                                 |

| タイムスタンプ     | 主な内容                                                                                               | 本報告書での位置づけ           |
|-------------|----------------------------------------------------------------------------------------------------|----------------------|
| 34:07–38:32 | Physical AIの未解決課題は、外界・3D状態の表現、長時間動画の処理、エッジでのリアルタイム実行、simulation-to-real、非常に高い信頼性。                  | 現在のロボティクス研究課題と整合     |
| 38:32–40:40 | 技術革新は「突然の魔法」より既存技術の組み合わせ・検証の積み重ねで起こることが多い。AI自体を研究に利用できる。                                           | AlphaEvolve等の方向性と整合  |
| 40:40–43:03 | 日本の強みは、高信頼な製造現場、精密機械、密集した産業現場、そこでの検証・データ収集、人手不足による強い需要。                                            | 政策仮説として合理的。優位性の実証は今後 |
| 43:03–46:34 | Noetraが約27,500基のRubin GPUを導入する計画。計算資源は大量の「検証」を行う必要条件。NVIDIA一社依存・国外依存にはリスクがあり多様化が必要。ただし完全国内完結も非現実的。 | GPU数は公式確認済み          |
| 47:45–55:29 | 人間とAIが大きく異なりすぎると制御・コミュニケーションが難しくなる。Physical AI、感情、身体性、記憶を持つAIでは新しい倫理問題が起き得る。                       | 社会・倫理上の論点            |
| 55:29–58:41 | AI時代には、人が問題設定・評価を担い、AIの結果を理解し続ける必要。最後に「自分で考え、結果に責任を持つこと」を手放してはならないと結論。                             | 本報告書の人材・ガバナンス提言と整合   |

特に動画の「1.5年差」は、「今日の日本モデルが、1.5年後には今日の米国モデルに追いつく」という意味の機械的予測として扱うべきではない。フロンティア能力は不均一で、コーディング、数学、長時間エージェント、マルチモーダル、サイバー、ロボットでは進歩速度が異なるうえ、評価ベンチマーク自体が数か月で飽和する現象が起きている。Stanford AI Indexも、難関ベンチマークHumanity's Last Examでフロンティアモデルが1年で30ポイント向上した一方、ベンチマークの信頼性・汚染・gamingが測定上の問題になっていると指摘している。<sup>13</sup>

## 動画で提示された主張の検証

### 検証サマリー

| 動画の主張                     | 判定     | 検証結果                                                                                                                                     |
|---------------------------|--------|------------------------------------------------------------------------------------------------------------------------------------------|
| 「強いAIで次の強いAIを作るループが回り始めた」 | 部分的に支持 | AlphaEvolveはAIをデータセンター、TPU、Gemini学習の改善に利用しており、AI-assisted AI R&Dは現実化。一方、OpenAIは2026年2月時点で自己改善能力を「High未満」と評価。完全自律RSIとは言えない。 <sup>14</sup> |
| ソフトウェア／コーディングが急速に伸びている    | 支持     | フロンティアモデル各社がcoding/agentを重点化。StanfordはOSWorldエージェント性能が約12%から66.3%へ改善したと報告。ただし依然約3分の1を失敗。 <sup>9</sup>                                    |

| 動画の主張                                           | 判定              | 検証結果                                                                                                                   |
|-------------------------------------------------|-----------------|------------------------------------------------------------------------------------------------------------------------|
| 単純なpre-training scalingからpost-training・検証へ比重が移る | 概ね支持            | Stanfordはtraining computeが増え続ける一方、データ品質・post-trainingの重要性を指摘。AlphaEvolveも「生成→自動評価→選択→再生成」の検証ループで成果を出す。 <sup>15</sup>   |
| 中国勢は急速に米国へ接近                                    | 強く支持            | 2026年3月の米中トップ性能差は2.7%。中国は論文数・引用・特許で世界を牽引し、2025年のnotable modelsは米59、中国35。 <sup>5</sup>                                  |
| 「中国は全部オープン」                                     | 文字通りには反証        | DeepSeek V4等はopen-weightだが、Alibaba等はopen-weightとAPI/closed系を併用。中国を「すべてopen」と分類するのは不正確。 <sup>7</sup>                    |
| 日本は最先端から約1～2年、1.5年遅れ                            | 定量検証不能          | 政府も日本の開発・投資の遅れを認めるが、「1.5年」を裏付ける標準尺度は存在しない。モデル・分野により差は異なる。 <sup>16</sup>                                                |
| 米国ラボには公開モデルより強い「社外秘モデル」がある                      | 未特定／推測          | StanfordはOpenAI、Anthropic、Google等で学習データ・パラメータ・学習期間の非開示が進んだことを確認。ただし「各社がどの程度強い秘密モデルを保有するか」は外部から確認できない。 <sup>2</sup>     |
| open-weightは安全対策を除去されるリスクがある                    | 技術的に妥当、定量比較は未確定 | Stanfordは通常条件では安全なモデルもjailbreakで安全性能が低下すると報告。open-weight固有の追加リスクは合理的だが、openとclosedの総合的安全性の優劣は単純には決められない。 <sup>17</sup> |
| 動画中の「OpenAIのサイバートストがHugging Faceに侵入した」事例        | 未特定             | 提供文字起こしから正確な事件・研究名を一意に特定できず、本報告書では検証済み事実として採用しない。                                                                      |
| Physical AIはLLMよりまだ未成熟                          | 強く支持            | 現実家庭作業のrobot successは12%に対し、シミュレーションRLBenchでは89.4%。sim-to-real / robustness gapは依然非常に大きい。 <sup>9</sup>                 |
| 日本の製造現場・実データはPhysical AIで強みになり得る                | 戦略仮説として支持       | METI/NEDO自身がFRONTiaを、日本の製造技術・産業実データを活用するPhysical AI戦略として設計。GENIACでも製造データAI-Ready化とロボット基盤モデルを採択。 <sup>18</sup>          |
| NoetraがRubin GPUを約27,500基導入                     | 確認済み            | NVIDIA公式発表はRubin GPU 27,500基、Vera CPU 13,750基、140MWと明記。 <sup>19</sup>                                                  |
| NVIDIA一社依存はリスク                                  | 支持              | Stanfordは世界AI computeの60%以上をNVIDIAが占めると推計。さらに先端チップ製造はTSMCへの集中が大きく、AI供給網そのものに集中リスクがある。 <sup>2</sup>                     |

## 「RSI」をどう理解すべきか

ここは動画で最も重要で、かつ最も誤解されやすい部分である。

現実 to 成立し始めているのは、

モデル → コード・実験案生成 → 自動評価 → 良い結果を選択 → 新しい学習／インフラ改善 → より良いモデル

という人間・AI混成型の高速R&Dサイクルである。AlphaEvolveはその典型であり、AI自身が使用する計算基盤や学習処理の最適化にまでAIを投入できることを示した。<sup>3</sup>

一方で、

**AIが自ら研究目標を設定し、外部の信頼できる評価や人間の介入なしに、自身より能力の高い後継AIを連続生成し続ける**

という強い意味のrecursive self-improvementが実証されたわけではない。少なくともOpenAIは2026年2月時点で自社の自己改善能力をHigh capability未満としていた。<sup>4</sup>

したがって動画の「ループが回り始めた」は、「自己増殖する超知能が成立した」ではなく、「AI研究をAIで高速化する正のフィードバックが実用段階へ入った」と読むのが最も証拠に整合する。

なお、AI能力そのものはその後も急速に上昇している。2026年9月3日にOpenAIはGPT-6 Astraを公開し、自社史上最も高性能な広範配備モデルと位置づけ、同時にサイバー能力をPreparedness Frameworkの「Critical」と評価した。これは「能力向上と安全性評価の重要性が同時に増す」という動画の問題意識を強く補強する。<sup>20</sup>

## オープン化をめぐる動画の評価

「openかclosedか」は二者択一ではなく、少なくとも四層に分解すべきである。

コードの公開、モデル構造の公開、重みの公開、学習データ・学習recipeの公開は別物であり、open-weightモデルだから開発過程が透明とは限らない。逆にclosed APIでも、system cardや第三者評価が充実している場合がある。Stanfordは2025年、主要AI企業のFoundation Model Transparency Index平均が40まで低下し、特にtraining data、compute、deployment impactの情報不足が目立つと報告した。<sup>21</sup>

2026年3月時点でトップclosed modelとトップopen modelの性能差は3.3%であり、openモデルは依然かなり競争力があるが、2024年8月の0.5%から差は再び開いた。<sup>13</sup>

したがって日本に必要なのは「全部open」でも「全部closed」でもなく、**能力・悪用可能性・用途に応じた段階公開**である。特にPhysical AIでは、基盤重みを国内企業に広く配布して産業イノベーションを促しつつ、高危険度のサイバー・CBRN等の能力についてはアクセス管理する方式が現実的である。

## 世界のLLM開発動向と日本の現在地

### 世界のフロンティア構造

2026年の世界市場は「OpenAI一強」ではない。比較可能なStanfordの2026年3月Arenaスナップショットでは、Anthropic 1,503、xAI 1,495、Google 1,494、OpenAI 1,481と、米国トップ4社が極めて接近し、Alibaba 1,449、DeepSeek 1,424が続いていた。これは、競争が単純な「ベンチマーク1位」から**価格、レイテンシ、信頼性、ツール利用、エージェント、特定領域性能**へ移っていることを示す。<sup>13</sup>

その後も世代交代は速い。OpenAIは2026年9月3日にGPT-6 Astraを投入した。GoogleはGemini系列を推論・マルチモーダル・world model/roboticsへ拡張し、AnthropicはClaudeを長時間coding/agentic workへ重点

化、xAIもGrok 4.x系列を継続している。<sup>22</sup> 中国側でもDeepSeek V4、Alibaba Qwen3.8、Moonshot Kimi K3など大型MoE・long-context・agent系モデルが競争しており、MoonshotはKimi K3を2.8兆パラメータ、native multimodal、1M token contextのモデルとして発表している。<sup>23</sup>

ただし、各社が公表するベンチマークはprompt、sampling、tool access、reasoning budget等の条件が違  
うため、ベンダー公称値を横並びにして「世界ランキング」を作ることは科学的には危険である。Stanford  
もbenchmark invalidityやleaderboard gamingを問題視している。<sup>9</sup>

主要プレイヤーの位置づけ

| 地域・主体              | 2026年の代表的方向                                            | データ・<br>計算資源          | 公開方針                                 | 戦略上の特徴                                                                                       |
|--------------------|--------------------------------------------------------|-----------------------|--------------------------------------|----------------------------------------------------------------------------------------------|
| OpenAI             | GPT-6 Astra、<br>coding/research/<br>computer use/cyber | 詳細な学<br>習規模は<br>非公開   | closed API中心                         | フロンティア能力+agent+<br>安全評価。AstraはCritical<br>cyber分類。 <sup>24</sup>                             |
| Anthropic          | Claude系列、長時間<br>coding/agent、<br>enterprise            | AWS/GCP<br>等大規模<br>基盤 | closed、system<br>card重視              | coding・agent・safetyを競<br>争軸にする。 <sup>25</sup>                                                |
| Google<br>DeepMind | Gemini、robotics、<br>world model、<br>AlphaEvolve        | TPU+<br>Google<br>DC  | Gemini closed、<br>Gemma open         | モデルだけでなくchip/<br>cloud/science/roboticsまで<br>垂直統合。AlphaEvolveがAI<br>R&Dそのものを改善。 <sup>3</sup> |
| xAI                | Grok 4.x、agent・<br>visual/reasoning                    | 大規模<br>GPU<br>cluster | closed中心                             | 巨大computeによる高速<br>scaling。 <sup>26</sup>                                                     |
| Alibaba/<br>Qwen   | Qwen3.x、MoE、<br>multimodal、<br>coding/agent            | Alibaba<br>Cloud      | openとclosed/<br>API併用                | open ecosystemを使いつつ<br>商用cloudを拡大。 <sup>27</sup>                                             |
| DeepSeek           | V4 Pro/Flash、1M<br>context、agent                       | 高効率<br>MoE志向          | V4 Previewは<br>open-weight           | 1.6T/49B activeのPro等、効<br>率とopennessで競争。 <sup>28</sup>                                       |
| Moonshot<br>AI     | Kimi K3、long-<br>context、<br>multimodal                | 大規模<br>MoE            | open ecosystem<br>志向だが各提供<br>形態はモデル別 | 2.8T、1M contextを掲げる<br>中国有力勢。 <sup>29</sup>                                                  |

世界全体では、研究から産業への重心移動が決定的である。2025年のnotable AI modelsの90%以上が産業界  
から生まれ、最先端モデルほどparameter count、dataset size、training durationを開示しない傾向が強  
まった。<sup>2</sup>

計算資源も集中している。Stanfordによれば世界のAI compute capacityは2022年以降年率3.3倍で増え17.1  
百万H100相当に達し、その60%以上をNVIDIA系が占める。米国には5,427のデータセンターがあり他国の10  
倍以上、さらに最先端AIチップの製造はTSMCへの集中度が極めて高い。したがってAI競争は、もはや「優秀  
な研究者のアルゴリズム競争」だけでなく、GPU・network・電力・冷却・半導体製造を含む産業政策その  
ものになっている。<sup>2</sup>

## 日本の研究・企業エコシステム

日本のモデル開発能力が「存在しない」という評価は誤りである。むしろ2024～2026年に多様なプレイヤーが成立した。

PFNはフルスクラッチ国産モデルPLaMoを開発し、2026年3月にはReasoningを中核にしたPLaMo 3.0 Prime βを発表した。PLaMo 2.0 Primeはデジタル庁のGovernment AIにおける国産LLM試用モデルにも選定された。<sup>30</sup>

NTTは軽量・日本語特化のtsuzumi 2を展開し、2026年5月には1GPU環境で図表を含む日本語ビジネス文書を扱うマルチモーダル機能を追加した。これは「最大モデルを作る」以外に、**edge / on-premise / confidentiality / cost**で差別化する日本的戦略の例である。<sup>31</sup>

NECはcotomiを基盤に業務agentやsecurity用途へ展開し、2026年9月にはフロンティアAIが発見する未公開脆弱性の分析からvirtual patch生成までを支援する「cotomi Security for Industry」を発表した。<sup>32</sup>

SB IntuitionsはSarashina2-70BやSarashina 2-Visionを展開し、政府システムや医療など日本語・データ主権を重視する領域へ実装を進めている。またNICTとの安全性共同研究も開始した。<sup>33</sup>

NIIのLLM-jpは日本の重要な**オープン研究基盤**である。2026年4月時点で大学・企業等から2,600人以上が参加し、コーパス、モデル、マルチモーダル、実環境interaction、安全性などのWGを持つ。LLM-jp Corpus v4は総量19.5兆token、そのうち日本語6880億tokenであり、日本語資源として大きい。2026年8月には33.2B dense model、9月にはLLM-jp-4-VL 9Bも公開した。<sup>34</sup>

Fugaku-LLMでは東京工業大学（現東京科学大学）、東北大学、理研、富士通、名古屋大学、CyberAgent等が「富岳」で日本語LLMを学習した。これは大学・国研・産業界をまたぐ日本型の共同研究能力があることを示す。<sup>35</sup>

スタートアップではSakana AIが特に重要である。同社は2025～2026年のSeries Bで320億円を調達し、海外資本も取り込みつつ、AI Scientist、multi-agent、post-training、日本企業向けAI実装へ進んだ。2026年8月のSakana NamazuはMoonshot AIのopen model Kimi K2.6を基盤に、日本語・日本商習慣へpost-trainingする戦略を採用している。これは、「**すべてを日本でpre-trainする**」ことだけがソブリンAIではないことを示す重要な事例である。<sup>36</sup>

## 世界 vs 日本

| 軸          | 世界最先端                                                                        | 日本                                                                              | 格差評価   |
|------------|------------------------------------------------------------------------------|---------------------------------------------------------------------------------|--------|
| 汎用フロンティア性能 | Anthropic、OpenAI、Google、xAI、中国上位勢が世界トップを高速交代。米中差2.7%（2026年3月）。 <sup>13</sup> | PLaMo、tsuzumi、cotomi、Sarashina、LLM-jp等が存在するが、独立した世界総合ランキングの最上位集団にはまだ定常的に現れていない。 | 大      |
| 研究速度       | モデルを研究・coding・評価に投入し、24/7型R&Dを拡大。                                            | GENIAC等で改善中だが、大規模実験回数は資本・computeで制約。                                            | 大      |
| 計算資源       | 世界17.1百万H100相当、米国DC 5,427、NVIDIA >60%。 <sup>2</sup>                          | 国内cloud整備に最大725億円支援。Noetra 27,500 Rubin GPU計画が大幅な能力増強。 <sup>37</sup>            | 大→縮小余地 |

| 軸                | 世界最先端                                                              | 日本                                                                                | 格差評価           |
|------------------|--------------------------------------------------------------------|-----------------------------------------------------------------------------------|----------------|
| 民間資本             | 米国民間AI投資は2025年2859億ドル。 <sup>12</sup>                               | Sakana AI Series Bは320億円等。政府はAI・半導体全体に7年間10兆円超を計画。 <sup>38</sup>                  | 極めて大           |
| 研究者・コミュニティ       | 米国がAI talent総数で最大、中国が論文・citationで強い。 <sup>2</sup>                  | LLM-jpだけでも2,600人超の横断コミュニティ。企業・大学の層は厚くなった。 <sup>39</sup>                           | 量・トップ層で差、連携余地大 |
| 日本語データ           | 英語中心。多言語・方言では性能低下が残る。 <sup>17</sup>                                | LLM-jp Corpus v4、日本語企業文書・行政・産業データ等に潜在優位。 <sup>40</sup>                            | 日本に戦略余地        |
| 産業・Physical data | 米中も製造・robotics、自動運転で強力。中国は世界のindustrial robot導入の54%。 <sup>41</sup> | 自動車・robot・工作機械・製造現場をPhysical AI戦略へ接続するFRONTiaが始動。                                 | 競争可能だが先行優位なし   |
| open ecosystem   | open-weightとclosedの混合。open/closed差は3.3%。 <sup>13</sup>             | LLM-jpはopen志向、商用企業はclosed/APIも多い。Noetraはpretrained weightsの広範提供を計画。 <sup>19</sup> | 戦略次第           |
| 社会実装             | 世界調査では88%の組織がAI利用、70%が少なくとも一業務でGenAI利用。 <sup>41</sup>              | IPAは2026年、日本でもAI導入が大企業中心に拡大する一方、成果は効率化中心で価値創出・事業変革が弱いと報告。 <sup>42</sup>           | 利用量以上に価値化で差    |
| 政策               | AI sovereigntyへの国家投資が世界的に拡大。EU AI Actは2026年8月から本格執行。 <sup>43</sup> | AI法、AI基本計画第Ⅱ期、GENIAC、FRONTia、AI・半導体10兆円枠など急速に整備。 <sup>44</sup>                    | 制度整備は急速に追隨     |

この表から分かるのは、日本の「AI格差」は研究論文の知識不足よりも、資本・compute・実験速度・商業化ループの格差の方が大きいことである。

## 格差を生む構造要因

### 技術・研究開発サイクル

フロンティア開発では、一つの良い論文を知っているだけでは優位を作れない。研究アイデアを小モデルで試し、medium scaleで再検証し、full scaleで学習し、agent benchmark、safety、latency、costまで測る必要がある。失敗の方が多以上、「失敗を何回買えるか」が研究力になる。

これは岡野原氏が動画で強調する「検証コスト」とほぼ一致する。世界のtraining computeが増え続け、最先端ではAI自身が実験・coding・evaluationを高速化する以上、compute不足は単に「最後の巨大trainingができない」という問題ではなく、**毎日の研究iterationが遅くなる問題**になる。 <sup>15</sup>

日本のGENIACはこの弱点に直接対応し、2024年から基盤モデル開発者への計算資源提供とcommunity形成を続け、2026年の第4期でも16テーマを採択した。これは方向として正しい。 <sup>45</sup>

しかし今後は「GPUを支給したか」より、**一GPU時間から何件の再現可能な実験・論文・モデル改善・製品を生み出したか**をKPIにすべきである。

## 資本・計算資源・電力

最大の構造差は資本である。2025年の米国民間AI投資2859億ドルという規模は、フロンティアAIを「研究プロジェクト」ではなく、データセンター、発電、半導体、ネットワークを含む巨大産業として成立させる。StanfordはGoogle単独の2025年年間capexが1500億ドル超と報告している。<sup>41</sup>

日本政府はこの問題を認識しており、2024年には国内AI計算資源整備5件へ最大725億円を支援し、AI・半導体産業基盤強化フレームでは2030年度までの7年間に10兆円超の公的支援、10年間で50兆円超の官民投資を目指している。<sup>46</sup>

Noetraの27,500 Rubin GPU／140MWは、その意味で日本の従来のAI計算環境からは大きな跳躍である。<sup>19</sup>しかし140MW級になると、モデル研究だけでなく電力調達、系統接続、冷却、稼働率、token/W、GPU故障率、scheduler効率が国家AI競争力になる。

したがって「何枚GPUを保有したか」より、年間有効FLOP、利用率、FLOP/W、実験throughput、token/yenで評価の方が実務的である。

## データ

pre-trainingでsynthetic dataだけに移行できたわけではない。Stanfordは2026年版でも、synthetic dataがreal dataを代替してはいない一方、deduplication、curation、data quality、post-trainingの効果が大きいと整理している。<sup>2</sup>

ここは日本に比較的明確な戦略余地がある。LLM-jp Corpus v4のような日本語基盤データに加え、企業が持つ設計図、保守記録、製造条件、故障ログ、映像、sensor data、熟練技能などは、一般Web crawlでは取得できない。METIが2026年に「企業・組織が持つ実データ」の利用を強調し、9件のdata ecosystem研究を採択したことは、この認識を反映している。<sup>47</sup>

ただし「日本にはデータがある」はそのまま強みではない。

紙、PDF、独自フォーマット、部門silos、権利不明、個人情報・営業秘密を含むデータは、AIから見れば存在しないのと同じである。

日本の真の課題はデータ量より「AI-ready率」であり、動画で強調された「現場」は、データ収集だけでなくモデルを試し、失敗を見つけ、再び現場データへ戻すclosed loopを作れることに価値がある。

## 人材と組織

世界のnotable modelの90%以上が産業界から出る現在、大学だけを強化してもフロンティア競争には足りない。<sup>2</sup> 研究者が最新GPU、production data、ユーザー、報酬の大きい企業へ移る構造があるため、大学—startup—大企業—国研を移動できる労働市場そのものが研究インフラになる。

日本にはLLM-jpの2,600人超コミュニティという重要な基盤がある。<sup>39</sup> またGENIACはこれまで300人超のengineerに基盤モデル開発経験を提供したとMETIが説明している。<sup>48</sup>

問題は「AI人材」という大括りではなく、

frontier pre-training engineer、distributed systems engineer、post-training/RL researcher、evaluation/safety researcher、AI compiler/kernel engineer、robotics/world-model researcher

といった世界的に希少な専門職が何人いるかである。

さらに、AIを導入する企業側にも、問題設定・評価・process redesignができる人材が必要である。IPAの2026年調査で、日本企業のAI効果が主に「業務効率化・迅速化」に留まり、価値創出やbusiness transformationへの発展が弱いことは、単なるモデル性能ではなく組織能力の問題を示している。 42

## 言語・文化

英語中心のフロンティアモデルが「日本語も話せる」ことと、「日本社会・法制度・業務慣行を高精度で理解する」ことは同じではない。

Stanfordは2026年版Responsible AIで、英語以外の地域固有benchmarkでは世界トップモデルでも性能低下が大きく、標準語から方言へ移るだけでも精度が大幅に落ち得ることを示している。 17

その意味で、tsuzumi、PLaMo、Sarashina、cotomi、LLM-jp、Namazu等の価値は「GPTより総合点で勝つ」ことだけではない。日本の法令、行政文書、医療、安全基準、商習慣、敬語、曖昧性、企業固有用語へ適合し、**国内・オンプレミスで運用できること自体が競争軸**になる。 49

## 規制・ガバナンス

「日本は規制が厳しいからAIで負けている」という単純な説明は、現状の証拠とは合わない。

日本のAI法は2025年に全面施行され、政府はAIの開発・活用の遅れを認識した上で、イノベーション促進とリスク対応を同時に進める設計を採っている。2026年7月には人工知能基本計画第Ⅱ期が閣議決定された。 50

一方EUではAI Actが2026年8月2日から本格的な執行段階へ入り、GPAIや透明性義務についてAI Office等のenforcement powersが発動している。 51

したがって現時点で日本の主なボトルネックは「過剰規制」よりも、**資本・compute・人材・データ整備・導入能力の不足**である可能性が高い。

ただしPhysical AIでは、自動車、医療機器、工場設備等の既存安全規制と生成AIをどう接続するかが今後大きな課題になる。ここでは「規制を緩める」より、**sandbox、simulation、第三者evaluation、事故責任の明確化によって安全に試せる場所を増やす**方が競争力につながる。

## AI格差が経済・安全保障・社会に与える影響

### 経済と産業競争力

AI格差の経済的意味は「日本にChatGPT類似サービスがない」ということではない。海外APIを使えば短期的な機能差は埋められる。

本質的な問題は、**AIによって生まれる高付加価値・知的財産・platform revenue・cloud revenue・data feedbackが国外へ蓄積する可能性**である。METIのGENIAC自体が、日本ではデジタル分野の遅れによって「デジタル赤字」が拡大してきたとの問題認識を示している。 52

同時に、海外モデルを使わないことも解ではない。高性能な海外モデルを遮断すれば、日本企業のproductivityが下がり、かえって産業競争力を損なう。

必要なのは、

**海外フロンティアAIを最大限使うことと、  
代替可能な国内モデル・compute・data能力を維持すること**

の両立である。

世界ではAI adoptionが88%まで上がり、Stanfordがまとめた研究では、customer supportで14～15%、software developmentで26%、marketing outputで50%など、構造化され評価しやすい業務で生産性改善が報告されている。<sup>41</sup> 日本でこれを単なる文章要約・議事録生成に留めれば、AI格差はモデル格差より**生産性格差**として拡大する。

## 経済安全保障

フロンティアAIはdual-use technology化している。

2026年9月のGPT-6 AstraについてOpenAIは、自社Preparedness Frameworkで初めてサイバー能力を「Critical」と分類し、未知の脆弱性を発見し高度なシステムへ攻撃手順を構築できる水準として強化された防護措置を導入した。これは一企業の自己評価であり独立検証とは区別すべきだが、AIがサイバー安全保障の主要要素になりつつあることは明確である。<sup>53</sup>

この状況で、政府・重要インフラ・防衛・金融が100%国外モデルに依存すると、

モデル供給停止、価格改定、API仕様変更、外国法制、export control、通信断、情報流出リスク

が政策リスクになる。

一方、100%国産化も、NVIDIA、TSMC、HBM、networking、EDA等の国際供給網を考えれば現実的ではない。世界AI computeの60%以上がNVIDIA系で、先端chip fabricationもTSMCへ高度に集中している。<sup>2</sup>

したがって「ソブリンAI」はautarkyではなく、

**障害が一か所起きても代替経路を持つresilience**

と定義する方が合理的である。

## Physical AIと日本の産業機会

動画で最も政策的価値が高いのはPhysical AIの議論である。

LLMではすでに米中企業が巨大な先行投資を持つが、roboticsでは現実環境の成功率がまだ低く、世界標準が確定していない。<sup>9</sup> Noetra/FRONTiaはここへ日本の製造現場データ、信頼性工学、精密機器を投入する構想である。NVIDIAの公式発表でも、製造、物流、医療、通信を対象にopen multimodal foundation models、agents、digital twins、roboticsを開発することが掲げられている。<sup>19</sup>

ただし、日本が「製造業が強いから自然にPhysical AIでも勝つ」と考えるのは危険である。中国はすでに世界industrial robot設置の54%を占め、米国にはGoogle DeepMind、Tesla、Waymo等の大規模Physical AI ecosystemがある。<sup>41</sup>

日本の勝機は、ロボットハード単体ではなく、

**実設備+sensor data+世界モデル+control+safety+digital twin+edge semiconductor**

を統合し、工場・倉庫・病院・建設現場で継続的に学習できるsystemを輸出することにある。

**社会・雇用・教育**

AIの影響は「仕事が全部なくなる／なくなるしない」という二択ではない。

Stanfordによれば2025年には調査対象組織の約3分の1が翌年にAIによるworkforce reductionを見込む一方、約半数は大きな変化を想定していない。影響は特にservice operations、software engineering等で先行している。また米国データではAI exposureの高い若年software developerの雇用変化も観察されているが、これはそのまま日本の因果関係として転用できない。<sup>41</sup>

日本の場合、人口減少・人手不足があるため、AIは「失業を生む技術」であると同時に**供給制約を緩和する技術**でもある。PFNのPLaMoが政府AIで試験され、Physical AIが労働力補完を目的に含むことは、この日本固有の事情を反映している。<sup>54</sup>

動画の最後にある岡野原氏の「人間は自分で考え、結果の責任を引き取るべき」という主張は、ここで重要になる。

AIが実行を担うほど、人間の仕事は、

**問題設定 → 制約設定 → AI選択 → 結果評価 → 責任判断**

へ移る。

したがって教育も「AIを使わずに答案を書く能力」だけでなく、**AIが生成した高度な答案の誤りを検出し、根拠を確認し、責任ある判断を下せる能力**を測らなければならない。

**改善策と実行可能な提言**

以下のコストは、既存の政府決定額ではなく、**本報告書による政策設計上の概算レンジ**である。算定の尺度として、既存の国内AI compute支援が最大725億円、GENIAC-PRIZE 2026が約10億円、AI・半導体全体の公的支援枠が2030年度までに10兆円超、Noetraが27,500 Rubin GPU級であることを基準にしている。既存制度との重複分を含む可能性があるため、「追加予算」とは限らない。<sup>55</sup>

**短期**

| 施策                       | 優先度 | 想定公的コスト    | 実施内容                                                                                          | 主な利点                                   |
|--------------------------|-----|------------|-----------------------------------------------------------------------------------------------|----------------------------------------|
| Japan Frontier Evalsの常設化 | 最優先 | 年100～300億円 | 日本語・coding・agent・cyber・science・Physical AI・安全性について非公開test setと第三者評価環境を整備。AISI/NII/AIST/大学で運営。 | 「1.5年差」のような感覚値から脱却し、政策投資を性能・安全性で評価できる。 |

| 施策                                  | 優先度 | 想定公的コスト        | 実施内容                                                                          | 主な利点                       |
|-------------------------------------|-----|----------------|-------------------------------------------------------------------------------|----------------------------|
| GENIACを「GPU配布」から「実験throughput支援」へ拡張 | 最優先 | 年500～1,500億円   | Noetra、ABCI、国内cloudを共通scheduler化し、大学/startupへcompute voucher。大規模失敗実験も成果として評価。 | iteration speed格差を直接縮小。    |
| 産業データAI-Ready基金                     | 最優先 | 2年で500～1,500億円 | 製造・物流・医療・建設等の動画、sensor、maintenance dataを権利処理・匿名化・標準化。                         | 日本独自の「Webにないデータ」を競争資源化。    |
| Government AIを性能運動型公共調達へ            | 高   | 年200～500億円     | 国産限定ではなく、国産・海外・hybridを同一evalで競争。一定割合をchallenger企業に配分。                         | 国内企業に最初の大型顧客を作り、研究を売上へ接続。  |
| Frontier AI人材Fellowship             | 高   | 年200～400億円     | 博士・postdoc・engineerへ世界水準報酬、GPU、海外lab滞在、兼業を提供。                                 | 人材流出抑制と海外人材誘致。             |
| 高リスクモデルの段階公開ルール                     | 高   | 年数十億円規模        | open-weight / controlled weight / API-onlyを能力別に定義。                            | opennessとsecurityの二者択一を回避。 |

日本にはすでにAISIのAI safety評価環境、GENIAC、LLM-jp等があるため、これらをゼロから作る必要はない。AISIは2026年に民間と共同で評価環境の開発・OSS公開を進めている。<sup>56</sup> 短期で最重要なのは新しい組織を増やすことより、**既存組織を共通benchmark・compute・dataの三層で接続すること**である。

## 中期

| 施策                                   | 優先度 | 想定公的コスト                    | 実施内容                                                                                | 利点                                            |
|--------------------------------------|-----|----------------------------|-------------------------------------------------------------------------------------|-----------------------------------------------|
| National Physical AI Testbed Network | 最優先 | 3～5年で2,000～5,000億円         | 自動車、factory、warehouse、construction、hospital等の実証現場を相互接続。simulation/digital twinも標準化。 | 日本の「現場」を再利用可能なAIインフラへ転換。                      |
| Noetra母モデル＋競争的派生モデル制度                | 最優先 | 既存FRONTia予算＋年500～1,000億円程度 | 共通pretrained modelは共有し、post-training/agent/applicationでは企業同士を競争させる。                 | 「協調しすぎて一社化」するリスクを回避。                          |
| AI chip second-source戦略              | 高   | 3～5年で3,000～8,000億円         | NVIDIAを排除せず、国内accelerator、AMD系、custom ASIC、compiler、interconnectへの移植性を確保。           | 一社・一国依存を軽減。                                   |
| 日本語・産業domain post-training市場         | 高   | 1,000～2,000億円              | 医療、金融、製造、行政、scienceごとのdatasets・evaluatorsを市場化。                                      | gigantic pre-trainingに依存せず世界open modelを日本仕様化。 |

| 施策                        | 優先度 | 想定公的コスト         | 実施内容                                                               | 利点                         |
|---------------------------|-----|-----------------|--------------------------------------------------------------------|----------------------------|
| AI startup growth capital | 高   | 官民で5,000億～1兆円規模 | Series B以降の大型growth funding、政府fund-of-funds、海外market expansionを支援。 | 「研究はできるが海外scaleできない」谷を埋める。 |

特に重要なのは、Noetraを「国策巨大モデルを一つ作って終了する事業」にしないことである。

Noetra自身が計画するpretrained weightsの広範提供は、その意味で合理的である。<sup>19</sup> 成功条件は、**共通基盤は協調、応用・post-training・製品化は激しく競争**という構造を維持することにある。

また、Sakana AIがKimi系open modelへ日本向けpost-trainingを行うような方式は、日本の資本制約に適している。<sup>57</sup> 「国産AI」の定義を「pre-trainingの全tokenが日本国内で計算されたAI」に固定するより、

**モデルを理解・評価・改変でき、国内データを安全に使え、必要時に国内computeで運用できるAI**

と定義した方が、経済安全保障と競争力を両立しやすい。

## 長期

| 施策                                    | 優先度 | 想定コスト                         | 目標                                                                        | 長期的利点                               |
|---------------------------------------|-----|-------------------------------|---------------------------------------------------------------------------|-------------------------------------|
| Minimum Viable Sovereign AI Stack     | 最優先 | 5～10年で1～3兆円相当を既存AI・半導体枠内で重点配分 | compute、network、model、data、evaluation、powerの全層に最低2系統以上の代替手段               | 地政学・供給停止へのresilience                |
| AI-native industrial architectureへの更新 | 最優先 | 官民数兆円規模                       | 工場・robot・machine toolを「AIが後付けされる設備」から「学習可能な設備」へ                           | Physical AIのデータnetwork effectを国内に形成 |
| AI-native教育体系                         | 高   | 年1,000億円前後                    | 全大学生にAI使用だけでなくevaluation、statistics、security、domain expertiseを教育          | AIに仕事を任せても監督できる人材を供給                |
| 世界研究拠点化                               | 高   | 年500～1,000億円                  | 日本国内computeを国外一流研究者にも一定開放し、共同研究・startup創業を促進                              | 国内人材のみで閉じず、talent inflowを形成         |
| AI safety / assurance産業の輸出化           | 高   | 5年で1,000～3,000億円              | robot、medical、automotive、industrial AIのverification/certification技術を国際標準化 | 日本の信頼性工学をAI時代の輸出産業へ                 |

この長期戦略の核心は、「完全な日本製」を目指すことではなく、「日本が交渉力を失わない程度の独立能力」を持つことである。

AI半導体だけを例にしても、世界はNVIDIA GPU、TSMC製造、HBM、EDA、networking、cloudと高度に国際分業されている。完全自給は極めて高コストで、技術進歩にも遅れる可能性がある。<sup>2</sup> 一方、一つのvendor/APIに全面依存すれば、輸出規制・価格・supply shortageがそのまま国家リスクとなる。

したがって政策上は、

### 自給率最大化ではなく「切替可能性最大化」

をKPIにするべきである。

具体的には「特定GPUが入手不能になった場合、国内重要モデルの何%を何日で別acceleratorへ移行できるか」「国外API停止時に政府・重要インフラ機能の何%を国内モデルで継続できるか」を定期的にstress testする方が、単純な「国産比率」より安全保障上意味がある。

### 最終評価

指定動画の問題提起は、2026年9月時点でもかなり本質を突いている。

特に、

AI開発は「一度巨大モデルを学習する競争」ではなくなり、強いAI・計算資源・検証・データが相互強化するループへ入った

という見方は、AlphaEvolve、agent性能の急伸、post-trainingの重要性、compute growthとよく整合する。<sup>58</sup>

一方で、「RSI」「日本は1.5年遅れ」「中国は全部open」「米国企業には公開できない最強秘密モデルがある」といった表現は、**方向感を示す専門家コメントとしては有用でも、厳密な事実命題としては限定・修正が必要**である。特に自律的RSIについては、AI-assisted R&Dと強い自己改善を分けて議論する必要がある。

4

日本にとってより重要なのは、「世界一の汎用チャットモデルを持っているか」だけで勝敗を判断しないことである。

日本の2026年時点のポジションは、

**汎用フロンティアLLMでは米中に大きく後れ、  
資本・compute・実験速度でも不利だが、  
日本語、企業データ、製造・robotics、高信頼システム、edge AIには依然有力な資産を持ち、  
Noetra／FRONTia、GENIAC、LLM-jp、PLaMo、tsuzumi、cotomi、Sarashina、Sakana AI等によって  
その資産をAIへ接続する基盤がようやく形成され始めた段階**

と評価するのが妥当である。<sup>59</sup>

最大のリスクは「技術的に永久に追いつけないこと」そのものではない。**米中のAIがAI研究をさらに高速化する一方、日本では研究→製品→売上→データ→再投資の循環が十分立ち上がらず、同じ1年の暦時間に海外が日本の数倍の研究iterationを回すことである。**

逆に最大の機会は、Physical AIがまだ完成していないことである。現実世界roboticsの成功率が依然低い現在、Noetra級compute、製造現場、industrial data、信頼性工学を一体化できれば、日本は単なる「LLM後追い国」ではなく、**Physical AIにおける基盤技術・評価・実装の供給国**になり得る。 60

そのための政策優先順位は明確である。

**第一に、計算資源を大量の検証iterationへ変える。**

**第二に、産業現場をAI-ready data infrastructureへ変える。**

**第三に、世界モデルを測れる独立evaluation能力を持つ。**

**第四に、研究成果を公共調達と民間売上へつなげる。**

**第五に、GPU・cloud・modelの「完全自給」ではなく「代替可能性」を確保する。**

**第六に、人間が問題設定・評価・責任を担う能力を教育する。**

日本政府はすでに「世界で最もAIを開発・活用しやすい国」を国家目標に掲げ、AI基本計画を技術進歩に合わせて毎年見直す方針を採っている。 61 今後の勝負は戦略文書を増やすことではなく、**compute、data、talent、evaluation、procurement**を一つの高速な学習ループとして実際に回せるかにある。

---

1 16 61 初の「人工知能基本計画」を閣議決定しました - 内閣府

[https://www.cao.go.jp/press/new\\_wave/20260206.html](https://www.cao.go.jp/press/new_wave/20260206.html)

2 15 hai.stanford.edu

<https://hai.stanford.edu/ai-index/2026-ai-index-report/research-and-development>

3 14 58 AlphaEvolve: A Gemini-powered coding agent for designing advanced algorithms — Google DeepMind

<https://deepmind.google/blog/alphaevolve-a-gemini-powered-coding-agent-for-designing-advanced-algorithms/>

4 GPT-5.3-Codex System Card | OpenAI

<https://openai.com/ja-JP/index/gpt-5-3-codex-system-card/>

5 13 hai.stanford.edu

<https://hai.stanford.edu/ai-index/2026-ai-index-report/technical-performance>

6 7 28 DeepSeek | DeepSeek-V4 Preview: Entering the Era of Affordable Million-Token Context

[https://deepseek.com/en/news/v4-preview/?utm\\_source=chatgpt.com](https://deepseek.com/en/news/v4-preview/?utm_source=chatgpt.com)

8 「AIロボット・フィジカルAIを見据えたマルチモーダル基盤モデル開発事業」の実施体制の決定について | 公募 | NEDO

[https://www.nedo.go.jp/koubo/CD3\\_100431.html?utm\\_source=chatgpt.com](https://www.nedo.go.jp/koubo/CD3_100431.html?utm_source=chatgpt.com)

9 60 Technical Performance | The 2026 AI Index Report | Stanford HAI

[https://hai.stanford.edu/ai-index/2026-ai-index-report/technical-performance?utm\\_source=chatgpt.com](https://hai.stanford.edu/ai-index/2026-ai-index-report/technical-performance?utm_source=chatgpt.com)

10 生成AIの開発力強化に向けたプロジェクト「GENIAC」において、新たに製造業データ等のAI-Ready化に関する研究開発テーマ計9件及びロボット基盤モデルに関する研究開発テーマ計2件を採択しました（METI/経済産業省）

[https://www.meti.go.jp/press/2026/05/20260514001/20260514001.html?utm\\_source=chatgpt.com](https://www.meti.go.jp/press/2026/05/20260514001/20260514001.html?utm_source=chatgpt.com)

11 12 43 hai.stanford.edu

<https://hai.stanford.edu/ai-index/2026-ai-index-report/policy-and-governance>

17 21 Responsible AI | The 2026 AI Index Report | Stanford HAI

[https://hai.stanford.edu/ai-index/2026-ai-index-report/responsible-ai?utm\\_source=chatgpt.com](https://hai.stanford.edu/ai-index/2026-ai-index-report/responsible-ai?utm_source=chatgpt.com)

- 18 19 59 Japan Government, Industrial Leaders and NVIDIA Launch the World's First National AI Infrastructure | NVIDIA Newsroom  
<https://nvidianews.nvidia.com/news/japan-government-industrial-leaders-and-nvidia-launch-the-worlds-first-national-ai-infrastructure>
- 20 24 53 Safety overview: GPT-6 Astra | OpenAI  
[https://openai.com/index/safety-overview-gpt-6-astra/?utm\\_source=chatgpt.com](https://openai.com/index/safety-overview-gpt-6-astra/?utm_source=chatgpt.com)
- 22 Release Notes | OpenAI | OpenAI  
[https://openai.com/products/release-notes/?utm\\_source=chatgpt.com](https://openai.com/products/release-notes/?utm_source=chatgpt.com)
- 23 Change Log | DeepSeek API Docs  
[https://api-docs.deepseek.com/updates/?utm\\_source=chatgpt.com](https://api-docs.deepseek.com/updates/?utm_source=chatgpt.com)
- 25 Anthropic's Transparency Hub \ Anthropic  
[https://www.anthropic.com/transparency?utm\\_source=chatgpt.com](https://www.anthropic.com/transparency?utm_source=chatgpt.com)
- 26 Introducing Grok 4.6 | SpaceXAI  
[https://x.ai/news/grok-4-6?utm\\_source=chatgpt.com](https://x.ai/news/grok-4-6?utm_source=chatgpt.com)
- 27 qwen3.8-2.4t-a95b Model Info - Alibaba Cloud Model Studio - Alibaba Cloud Documentation Center  
[https://www.alibabacloud.com/help/en/model-studio/qwen3-8-2-4t-a95b?utm\\_source=chatgpt.com](https://www.alibabacloud.com/help/en/model-studio/qwen3-8-2-4t-a95b?utm_source=chatgpt.com)
- 29 Moonshot AI  
[https://www.moonshot.cn/?utm\\_source=chatgpt.com](https://www.moonshot.cn/?utm_source=chatgpt.com)
- 30 生成AI基盤モデルPLaMo 3.0 Primeβ版のモニター企業募集 - 株式会社Preferred Networks  
[https://www.preferred.jp/ja/news/pr20260319?utm\\_source=chatgpt.com](https://www.preferred.jp/ja/news/pr20260319?utm_source=chatgpt.com)
- 31 49 NTT版LLM tsuzumi 2アップデート ～世界トップレベルの図表入り日本語ビジネス文書処理性能を1GPU環境で実現～ | ニュースリリース | NTT  
[https://group.ntt.jp/newsrelease/2026/05/19/260519a.html?utm\\_source=chatgpt.com](https://group.ntt.jp/newsrelease/2026/05/19/260519a.html?utm_source=chatgpt.com)
- 32 NEC開発のAIコア技術「cotomi」: NEC Generative AI | NEC  
[https://jpn.nec.com/LLM/cotomi.html?utm\\_source=chatgpt.com](https://jpn.nec.com/LLM/cotomi.html?utm_source=chatgpt.com)
- 33 「Gビズポータル」のデータベース拡充支援に「Sarashina」を活用 | プレスリリース | SB Intuitions株式会社  
[https://www.sbintuitions.co.jp/news/press/20260325\\_01/?utm\\_source=chatgpt.com](https://www.sbintuitions.co.jp/news/press/20260325_01/?utm_source=chatgpt.com)
- 34 39 WebクロールとWebアーカイブに基づく大規模音声・音響データセットの一般公開 - 国立情報学研究所 / National Institute of Informatics  
[https://www.nii.ac.jp/news/release/2026/0428.html?utm\\_source=chatgpt.com](https://www.nii.ac.jp/news/release/2026/0428.html?utm_source=chatgpt.com)
- 35 スーパーコンピュータ「富岳」で学習した大規模言語モデル「Fugaku-LLM」を公開 | 理化学研究所  
[https://www.riken.jp/pr/news/2024/20240510\\_2/index.html?utm\\_source=chatgpt.com](https://www.riken.jp/pr/news/2024/20240510_2/index.html?utm_source=chatgpt.com)
- 36 38 Announcing Our Series B  
[https://sakana.ai/series-b/?utm\\_source=chatgpt.com](https://sakana.ai/series-b/?utm_source=chatgpt.com)
- 37 46 55 経済安全保障推進法に基づくクラウドプログラムの安定供給確保に係る供給確保計画の認定等について (METI/経済産業省)  
[https://www.meti.go.jp/press/2024/04/20240419002/20240419002.html?utm\\_source=chatgpt.com](https://www.meti.go.jp/press/2024/04/20240419002/20240419002.html?utm_source=chatgpt.com)
- 40 LLM-jp Corpus v4 の公開 - LLM-jp  
[https://llm-jp.nii.ac.jp/news/post-998/?utm\\_source=chatgpt.com](https://llm-jp.nii.ac.jp/news/post-998/?utm_source=chatgpt.com)
- 41 hai.stanford.edu  
<https://hai.stanford.edu/ai-index/2026-ai-index-report/economy>

42 プレス発表 国内企業のDX動向・AI活用動向のポイントを公表 | プレスリリース | IPA 独立行政法人 情報処理推進機構

[https://www.ipa.go.jp/pressrelease/2026/press20260716.html?utm\\_source=chatgpt.com](https://www.ipa.go.jp/pressrelease/2026/press20260716.html?utm_source=chatgpt.com)

44 AI戦略 - 科学技術・イノベーション - 内閣府

[https://www8.cao.go.jp/cstp/ai/index.html?utm\\_source=chatgpt.com](https://www8.cao.go.jp/cstp/ai/index.html?utm_source=chatgpt.com)

45 生成AIの開発力強化に向けたプロジェクト「GENIAC」を開始します (METI/経済産業省)

[https://www.meti.go.jp/press/2023/02/20240202003/20240202003.html?utm\\_source=chatgpt.com](https://www.meti.go.jp/press/2023/02/20240202003/20240202003.html?utm_source=chatgpt.com)

47 生成AIの開発力強化・社会実装に向けたプロジェクト「GENIAC」において、新たにデータエコシステム構築等に関する研究開発テーマ計9件を採択しました (METI/経済産業省)

[https://www.meti.go.jp/press/2026/07/20260702001/20260702001.html?utm\\_source=chatgpt.com](https://www.meti.go.jp/press/2026/07/20260702001/20260702001.html?utm_source=chatgpt.com)

48 生成AIの社会実装に向けたプロジェクト「GENIAC-PRIZE」を開始します (METI/経済産業省)

[https://www.meti.go.jp/press/2025/05/20250509002/20250509002.html?utm\\_source=chatgpt.com](https://www.meti.go.jp/press/2025/05/20250509002/20250509002.html?utm_source=chatgpt.com)

50 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（A I 法） - 科学技術・イノベーション - 内閣府

[https://www8.cao.go.jp/cstp/ai/ai\\_act/ai\\_act.html?utm\\_source=chatgpt.com](https://www8.cao.go.jp/cstp/ai/ai_act/ai_act.html?utm_source=chatgpt.com)

51 AI Act | Shaping Europe's digital future

[https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai?utm\\_source=chatgpt.com](https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai?utm_source=chatgpt.com)

52 GENIAC (METI/経済産業省)

[https://www.meti.go.jp/policy/mono\\_info\\_service/geniac/?utm\\_source=chatgpt.com](https://www.meti.go.jp/policy/mono_info_service/geniac/?utm_source=chatgpt.com)

54 PFNの国産大規模言語モデル「PLaMo」、デジタル庁のガバメントAIに試用モデルとして選定 - 株式会社 Preferred Networks

[https://www.preferred.jp/news/pr20260312?utm\\_source=chatgpt.com](https://www.preferred.jp/news/pr20260312?utm_source=chatgpt.com)

56 プレス発表「AIセーフティ評価環境 検討タスクフォース」総会を開催 | プレスリリース | IPA 独立行政法人 情報処理推進機構

[https://www.ipa.go.jp/pressrelease/2025/press20260306.html?utm\\_source=chatgpt.com](https://www.ipa.go.jp/pressrelease/2025/press20260306.html?utm_source=chatgpt.com)

57 Sakana AI、日本語特化のLLM API「Sakana Namazu」を提供開始

[https://sakana.ai/namazu-api/?utm\\_source=chatgpt.com](https://sakana.ai/namazu-api/?utm_source=chatgpt.com)