



2026年における日本の国産LLM採用状況：包括的調査報告

エグゼクティブサマリー

2026年、日本の国産大規模言語モデル（LLM）は「実験的プロジェクト」から「戦略的投資対象」へと転換し、金融・医療・自治体・製造業などミッションクリティカルな領域での本格導入が進んでいる。官民合わせて総額3兆円規模の投資が始動し、1兆パラメータ級の世界最先端クラスモデル開発を国家目標に掲げるなど、かつてないスケールでの取り組みが展開されている。市場は海外製LLMとの明確な「使い分け」フェーズに移行しており、機密性・安全性が求められる領域では国産LLM、創造的・汎用的タスクでは海外製モデルというハイブリッド戦略が定着しつつある。[1][2][3][4]

政府の取り組み

ガバメントAI「源内」と国産LLM公募

デジタル庁は2025年12月、政府職員向け生成AI基盤「源内（げんない）」に組み込む国産LLMの公募を開始した。ChatGPTやGeminiなど海外モデルへの依存を減らし、日本語・行政文書・公的業務に特化した国産AIの育成と実装を本格化させる狙いがある。[5][6]

公募スケジュールは以下の通りである。

- **2025年12月～2026年1月30日**：公募期間[6]
- **2026年2～3月**：国産LLM選定[7][5]
- **2026年5月頃**：他府省庁への源内展開（政府職員10万人超が利用可能に）[7]
- **2026年夏頃**：国産LLMの試験導入開始[8][5]
- **2027年度以降**：本格提供・ライセンス契約[9]

さらに、PFN開発の「PLaMo翻訳」がガバメントAIに先行導入され、翻訳業務での国産AI活用が既に始まっている。[10]

官民3兆円プロジェクト

2025年12月、ソフトバンクを中心とする日本企業十数社が出資する新会社設立構想が明らかになった。経済産業省は2026年度から5年間で約1兆円を支援し、2026年度予算案には3,000億円超が計上されている。ソフトバンクは6年間でデータセンターに2兆円を投じ、北海道苫小牧市・大阪府堺市の施設を2026年度までに稼働させる計画である。[4][11][12]

この新会社は約100人規模の技術者（ソフトバンク+PFNのAI技術者中心）が所属し、まず「1兆パラメータ」規模の基盤モデル開発を目指す。開発モデルは日本企業に開放され、最終的にはロボットに搭載できる「フィジカルAI」の開発につなげる方針である。[11][4]

AI基本計画とGENIAC

2025年12月、日本政府は初のAI基本計画を採択し、技術革新とリスク管理のバランスを取りながら「信頼できるAI」の創出を掲げた。経済産業省・NEDOが推進する「GENIAC」プロジェクトは第3期を迎え、楽天、PFN、Stockmark、Turingなど多様なプレイヤーが採択され、汎用モデルだけでなく創薬・自動運転・製造といった特定領域の基盤モデル開発が支援されている。[13][1]

主要国産LLMの現状

主要モデル比較

開発元	モデル名	パラメータ数	主な特徴	商用化状況
NTT	tsuzumi 2	30B (7B/0.6Bも)	軽量・省電力・1GPU動作	商用展開済、引き合い2,000件[14]
NEC	cotomi	130億	高速推論 (GPT-4oの2.2倍)、MCP対応	商用展開済、cotomi Act提供[15][1]
富士通	Takane	130億	Cohere共同開発、高い日本語能力	中央省庁実証成功、2026年度サービス提供目標[16][17]
ソフトバンク	Sarashina 2	MoE 8×70B	大規模MoE、通信業界特化	社内利用開始、商用化準備中[18][1]
楽天	Rakuten AI 3.0	約7,000億 (MoE)	国内最大規模、アクティブ40B	2026年春オープンウェイト公開予定[19][20]

ELYZA (KDDI)	ELYZA-LL M-Diffusion	1,300億	Llama系ベース日本語特化	WAKONX基盤で商用展開中 [21][22]
PFN	PLaMo 2.0	310億	フルスクラッチ国産、NICT連携	限定公開、2026春後継モデル予定[23]

NTT 「tsuzumi 2」

2025年10月にリリースされたtsuzumi 2は、30Bパラメータを搭載しつつ1枚のGPUで動作可能な軽量設計を維持している。NTT島田社長によれば、2025年4月～12月のAIビジネス受注額はグループ全体で約1,478億円に達し、tsuzumi 2だけでも約2,000件の引き合いがある。特に自治体・金融・医療からの問い合わせが多い。[24][14][25]

tsuzumiの戦略的優位性は、NTTのIOWN構想との連動にある。将来的には光ネットワークで結ばれた多数のtsuzumiが連携し、巨大モデルに匹敵する「AIコンステレーション」への進化が計画されている。[1]

NEC 「cotomi」

cotomiはGPT-4oと比較して2.2倍以上の推論速度を実現し、128kトークンのロングコンテキストに対応した。2026年1月からは「cotomi Act」の提供を開始し、AIが業務ノウハウを自動抽出・組織資産化するエージェント技術への展開を進めている。NECは生成AI関連事業で500億円の売上目標を掲げ、ハードウェアからコンサルティングまでの一気通貫提供で差別化を図る。[1]

富士通 「Takane」

2025年中に中央省庁と協働で実施したパブリックコメント業務の実証実験では、約12万文字の意見データに対し、賛否分類・要約を10分程度で完了させた。法令案と各意見の整合チェックでは、全体の8割超の意見について正しい参照を回答できることが確認された。富士通は、この成果をもとに政策立案・法律制定プロセス向け生成AIサービスを2026年度中に提供開始する方針である。[26][16][27][17]

楽天 「Rakuten AI 3.0」

2025年12月に発表されたRakuten AI 3.0は、約7,000億パラメータ（MoEアーキテクチャ）を持ち、アクティブパラメータを約400億に抑えることで推論効率と知識容量を両立している。日本語MT-Benchで8.88というスコアを記録し、GPT-4o（8.67）を上回ったとされる。2026年春をめどにオープンウェイトモデルとして公開される予定で、日本のAI研究者やエンジニアが自由に活用できるようになる。[19][20][28][1]

ソフトバンク 「Sarashina」 と通信業界特化

ソフトバンクは2025年10月、国産LLM「Sarashina」を活用した通信業界向け生成AI基盤モデル「Large Telecom Model」を発表し、社内利用を開始した。データの学習から運用まで国内で完結する安全な生成AI基盤を実現しており、経済安全保障の観点から注目されている。[18]

PFN・さくらインターネット・NICT連携

2025年9月、PFN・さくらインターネット・NICTの3者が完全国産LLM開発で基本合意し、2026年春を目標に「PLaMo 2.0」後継モデルを共同開発している。NICTが保有する700億ページ超の日本語Webページデータは他に類を見ない資産であり、国産LLMの学習において決定的な優位性をもたらすとされる。さくらインターネットの「生成AIプラットフォーム」上でモデルが提供され、クラウドからアプリケーションまで完全に国内で完結する体制が構築される。[23]

採用が加速する分野

行政・公共

ガバメントAI「源内」を通じた国産LLM導入が最も注目される分野である。2026年夏には試験導入が始まり、2027年度の本格展開を目指す。富士通Takaneの中央省庁パブリックコメント実証は、行政AI活用の具体的成果として高く評価されている。[16][5][9]

金融・医療

NTT tsuzumi 2への引き合い2,000件のうち、金融・医療は特に問い合わせが多い分野である。金融機関の顧客データや医療カルテといった機密性の高い情報を国内で処理できる点が、国産LLM選択の最大の理由となっている。NTTのtsuzumiはファイナンシャルプランニング技能試験2級レベルの専門知識追加学習においても、少ないデータ量で合格基準に到達できることが実証されている。[14][1]

製造業

NECのcotomiは設計図面・部品表の構造理解や熟練技術者のノウハウ継承に特化したモデルを展開している。Stockmarkは製造業特化の1,000億パラメータ級LLMを開発し、パナソニックなど大手製造業での採用実績を持つ。[1]

通信

ソフトバンクのLarge Telecom Modelは通信品質の予測精度90%以上を達成し、通信インフラのAI活用で先行している。[18][1]

技術的アーキテクチャの分岐

2026年の国産LLM開発は、大きく二つの技術思想に分かれている。[1]

アプローチ	代表企業	特徴	適した用途
規模追求型 (MoE)	ソフトバンク、楽天	巨大パラメータ数（数千億～7,000億）で汎用能力を追求、推論時はアクティブパラメータを限定	クラウド大規模サービス、AGI的振る舞い
効率特化型 (Adapter/Edge)	NTT、NEC、ELYZA	軽量ベースモデル（7B～30B）＋タスク特化アダプタ（LoRA等）	オンプレミス、エッジ、閉域環境、低コスト運用

特に効率特化型は、高価なGPUを大量調達できない中小企業や自治体にとって現実的な選択肢となっており、日本の計算資源制約から生まれた「軽量化技術」はグローバルでも通用する独自の競争力になりつつある。[1]

国産LLMの強みと課題

強み

- **日本語処理精度**：敬語・謙譲語・業界専門用語の適切な使い分けにおいて海外モデルに優位[3][1]
- **データ主権・安全性**：機密データを国内で完結して処理可能、カントリーリスクの回避[1]
- **軽量モデルの優秀さ**：省エネ・オンプレミス向きで、エッジAI普及にも適合[3]
- **業務特化の最適化**：製造・建設・医療・法務など「日本型業務」への対応力[3]
- **コスト効率**：円安環境下で海外APIコスト負担が増す中、国内完結型の運用が経済的に有利[1]

課題

- **汎用性能**：総合力ではGPT-5系、Claude 4系、Gemini 3系に及ばない[3]
- **1兆パラメータ級モデルの不在**：世界最先端の「怪物モデル」はまだ存在しない[3]
- **GPU調達の制約**：NVIDIA製GPUの争奪戦が続き、円安も相まって調達コストが上昇[1]

- **高品質日本語データの確保**：国産LLM開発において学習用データの質と量が喫緊の課題[29]
 - **人材不足**：AI研究者・エンジニアの育成が追いついていない[3]
-

今後の展望

2026年のマイルストーン

2026年は、国産LLMが実証実験フェーズを完全に脱却し、本格的な社会実装に移行する「AI社会実装元年」と位置づけられている。主な注目イベントは以下の通りである。[1]

- **2026年春**：楽天AI 3.0のオープンウェイト公開、PFN新モデルリリース、ソフトバンク主導新会社設立[23][4][19]
- **2026年夏**：ガバメントAI「源内」への国産LLM試験導入[5]
- **2026年度中**：富士通Takane行政向けサービス提供開始、NEC cotomi Actによるエージェント実装本格化[17][1]
- **2026年度以降**：NTT tsuzumi 2の電力業務特化LLM実用化[30]

中長期的見通し

「チャットボット」型のAI利用から、AIが自律的にタスクを実行する「エージェント」への進化が2026年の決定的なトレンドとなる。NECのcotomi ActやソフトバンクのAI-RAN構想に見られるように、AIが計画立案・システム操作・外部連携を自律的に行う世界観が現実化しつつある。[1]

市場全体としては、国産LLMと海外LLMの「使い分け」がさらに進み、基幹的・機密的領域は国産AI、創造的・実験的領域は海外AIという棲み分けが定着していく見込みである。3兆円プロジェクトの成果が2027年後半に初期成果として登場し、2030年には日本が「AI主権国」としての地位を確立できるかどうか問われている。[3][1]

References

1. [\[PDF\] 2026年日本国産 AIモデル \(LLM\) の到達点と 未来展望](#) - 2026年1月現在、日本の人工知能 (AI) 産業、とりわけ大規模言語モデル (LLM) の開発競争は、かつてないほどの活況と構造的な転換期を迎えている。2023年から2024年 ...
2. [国産LLM開発競争の現状レポート | パラメータ数・学習 ... - AI](#) - 日本の大規模言語モデル開発競争を徹底分析。ELYZA、Preferred Networks、rinna、サイバーエー

ジェントなど主要プレイヤーのパラメータ数、学習データ量、投資額を比較。国産LLMの技...

3. [日本のAI：2026年1月時点のおさらいと未来予測 | MasaT. 武富 正人](#) - 現状のおさらい (2026年1月) 主要な国産LLMの状況 日本は現在、かなり活発にAI開発を進めています。代表的なプレイヤーは： 実用段階に入っているもの： ...
4. [官民で国産AI開発、26年度予算案に関連費用として3000億円 ...](#) - 官民による総額3兆円規模の国産AI (人工知能) 開発計画の全容が判明した。来春にもソフトバンクなどの日本企業十数社が出資して新会社を設立し、国内最大規模のAI基盤モデルの開発を目指す。AI開発は米中が性...
5. [ガバメントAIで試用する国内大規模言語モデル \(LLM\) の公募について](#) - 令和8年 (2026年) 2月から3月頃： 国内開発LLMの選定； 令和8年 (2026年) 5月頃： 源内の他府省庁展開； 令和8年 (2026年) 夏頃： 国内開発LLMの試験導入の開始 ...
6. [国産LLMを求めて動き出す—デジタル庁が「ガバメントAI」公募 ...](#) - ついに、日本の行政も自前のAIを求めて動き出しました。 デジタル庁は2025年12月2日、政府職員が使うAI環境「源内 (げんない)」に導入する国産大規模言語モデル (LLM) を公募すると発表。 Chat...
7. [【2026年最新】ガバメントAI「源内」とは？地方自治体が今から ...](#) - 源内本格運用開始. 2025年12月1日. 全国. 国産LLM公募開始 (~2026年1月30日). 2025年12月中. デジタル庁. PLaMo翻訳導入. 2026年1月以降. 各府省庁. 試験 ...
8. [そろそろしめきり！デジタル庁「ガバメントAI用国産LLMの公募](#) - そろそろしめきり！デジタル庁「ガバメントAI用国産LLMの公募」 こんにちはmakokonです。先月 (2025年12月2日) に発表された、デジタル庁の「ガバメントAIで試用する国内大規模原簿モデル (...
9. [デジタル庁、国内開発LLMの公募開始—ガバメントAI「源内」で26 ...](#) - デジタル庁は、ガバメントAI「源内」において試験導入する国内開発の言語モデル (LLMを中心に、小規模言語モデルを含む) の公募を開始した。2026年度中の試験導入・評価検証を経て、2027年度以降の本格...
10. [ガバメントAI「源内」での「PLaMo翻訳」利用開始について](#) - デジタル庁は、デジタル社会形成の司令塔として、未来志向のDX (デジタル・トランスフォーメーション) を大胆に推進し、デジタル時代の官民のインフラを一気呵成に作り上げることを目指します。
11. [【速報】ソフトバンク主導で国産AI新会社設立へ —経産省5 ...](#) - 2025年12月21日、日本のAI戦略が新たな局面を迎えました。ソフトバンクを中心とする日本企業十数社が来春にも新会社を設立し、ロボット向け国産AI (フィジカルAI) の開発に乗り出すことが明らかになり...
12. [官民で「国産AI開発」を本格化、5年で1兆円支援へ ソフトバンク ...](#) - 経済産業省が国産人工知能 (AI) の基盤モデル開発に向け、5年間で約1兆円規模の支援を計画していることが明らかになった。ソフトバンクやプリファードネットワークスなど、日本企業十数社が新会社設立を構想して...

13. [Japan Govt Adopts 1st Basic Plan on AI](#) - The basic plan stipulates that Japan will “create reliable AI” while balancing technological innovat...
14. [NTTの国産AI「tsuzumi 2」の滑り出しは？ 「国内で引き合い2000件 ... - 安全保障などの観点から国産AIモデルへの関心が高まる中、有力なモデルの一つとして注目を集めるNTTの「tsuzumi 2」。](#) 2025年10月、「（国産AI開発競争 ...
15. [NEC開発の生成AI「cotomi」](#) - 業務変革の実現のため、高速・高性能な生成AIの活用をお考えの方へ。NEC開発のAIコア技術「cotomi」の特長やメリットについてご紹介します。
16. [中央省庁がパブリックコメント業務を富士通のLLM「Takane」で ... - 2026年2月3日、富士通は、パブリックコメント業務に大規模言語モデル「Takane」を活用した実証実験を特定の中央省庁と協働で2025年中に実施し、有効性を確認したと発表 ...](#)
17. [パブリックコメント業務に大規模言語モデル「Takane」を活用し - パブリックコメント業務に大規模言語モデル「Takane」を活用し、中央省庁で業務効率化の実証実験を実施。政策立案や法律制定のプロセスに活用できる生成AIサービスの ...](#)
18. [通信業界向け生成AI基盤モデル「Large Telecom Model」が国産AI ... - ソフトバンクの公式ホームページです。企業・IRの「通信業界向け生成AI基盤モデル「Large Telecom Model」が国産AIモデルに発展、社内利用を開始～国産LLM「Sarashina」を活用...](#)
19. [楽天、7000億パラメータの日本語LLM「Rakuten AI 3.0」 - ... 2026年春を目途にオープンウェイトモデルとしての公開も予定している。Rakuten AI 3.0では、計算効率を高めるため、約7,000億個のパラメータのうち ...](#)
20. [新たなAIモデル「Rakuten AI 3.0」発表 約7000億のパラメータ持つ ... - 楽天グループは、経済産業省および新エネルギー・産業技術総合開発機構（NEDO）が推進する、日本の生成AIの開発力強化を目的としたプロジェクト「GENIAC（Generative AI Acceleration）」の一環として、約7,000億個のパラメータを持つ日本語LLM「Rakuten AI 3.0」を発表した。](#)
21. [KDDIが日本語最適化のELYZA-LLM-Diffusionを商用公開！ローカル ... - KDDI傘下ELYZAが日本語強化のLLM「ELYZA-LLM-Diffusion」を公開。商用利用可能でローカル実行に最適化された技術革新を解説。](#)
22. [NRIとKDDI、ELYZAが国産LLM基盤の提供で協業 - 野村総合研究所とKDDI、ELYZAは、生成AIの国産大規模言語モデル基盤の提供に向けて協業を開始すると発表した。](#)
23. [国産LLM「PLaMo 2.0」後継モデル開発へ：PFN・さくら ... - PFN・さくらインターネット・NICTが2026年春に向けて開発する国産LLM「PLaMo 2.0」後継モデルの詳細。700億ページの日本語データを活用した完全国産AIエコシステム構築の全貌を解説。](#)
24. [更なる進化を遂げたNTT版LLM tsuzumi 2の提供開始～日本 ... - NTT株式会社（本社：東京都千代田区、代表取締役社長：島田 明、以下「NTT」）は、大規模言語モデル（以下、LLM※1）の普及に伴う電力消費やコスト増加といった...](#)

25. [NTTが純国産LLMの新版「tsuzumi 2」提供開始、30Bパラメーター ...](#) - NTTは2025年10月20日、純国産のLLM「tsuzumi 2」の提供を開始した。2024年3月に商用提供を開始したtsuzumiに比べ、パラメーター数を70億（7B）から300 ...
26. [Fujitsu's Takane LLM successfully piloted in central ...](#) - Fujitsu's Takane LLM successfully piloted in central government agency to streamline public comment ...
27. [富士通と中央省庁がパブリックコメント処理を生成AIで自動化 ...](#) - 富士通は2026年2月3日、一部の中央省庁と共同で、国民から意見を募るパブリックコメント（意見公募手続）の業務に、生成AI/LLMを活用して、意見の分類や要約などの作業を自動化する実証実験を行ったと発...
28. [楽天、国内最大規模の高性能AIモデル「Rakuten AI 3.0」を開発 ...](#) - また2026年春頃を目安に、オープンウェイトモデルとしての公開も計画されています。本モデルは、Mixture of Experts (MoE) アーキテクチャを使用した ...
29. [2. AI「2026年問題」：日本の成長戦略への警鐘](#) - 日本のAI戦略においては、国産LLM(大規模言語モデル)の開発が加速する中で、特に高品質な日本語データの確保と生成が喫緊の課題となるでしょう。
30. [NTT版LLM\(大規模言語モデル\)「tsuzumi 2」を活用した電力業務特化 ...](#) - 本年1月から3月末にかけて次のプロセスで取り組み、構築したLLMの精度等を踏まえて2026年度以降の実用化を目指します。① データ収集・加工、中国電力が ...