ChatGPT

ガバメントAIで試用する国内大規模言語モデル公募結果の深掘り分析

エグゼクティブサマリー

本報告書は、デジタル庁[1]が公表した「ガバメントAIで試用する国内大規模言語モデル（LLM）の公募結果」（公開日：2026-03-06）を一次情報として、政策背景・公募要件・選定結果（7モデル）・技術的評価可能性・運用計画とリスク・中長期の示唆を、可能な限り公開情報で裏取りしつつ分析したものである。公式発表URL：`https://www.digital.go.jp/news/10d55c63-b3e1-42b9-9cc5-93a06943ae0e`[2]

公募は2025-12-02〜2026-01-31を期限として実施され、応募15件から7件が選定された。選定モデルは、国内開発であること、説明可能性（開発経緯・方法・体制等）、行政実務での実用性、当日初開示の50問評価テスト、海外主要LLMとの比較ベンチマーク、安全性（ハルシネーション・バイアス・有害出力等）、学習データの法令遵守、そして「機密性2情報」取扱いを想定したガバメントクラウド上での推論実行などを基準に評価された。[3]

今後は、2026年3月以降に契約・実装調整、5月頃に政府職員向け「源内」の大規模実証が開始され、8月頃から源内上で国内LLMの試用評価が始まる。評価・検証結果（一部）は2027年1月頃に公表予定で、2027年4月以降に優れたモデルを有償で政府調達する構想が示されている。[4]

技術面では、公開情報の厚み（ベンチマークの再現性、モデル仕様の開示、提供形態の透明性）にモデル間で大きな差がある。例えば、PLaMoは価格（1Mトークンあたり入力60円／出力250円）や評価設計（IFEval、ELYZA-tasks-100等）を比較的詳細に公開している一方、いくつかの商用モデルはパラメータ数・学習データ詳細・推論要件が「未指定」である。したがって、政府側の評価公表（2027年1月予定）が、国内エコシステムにとって事実上の「共通比較軸（標準）」になり得る。[5]

運用・リスク面では、(a) 公的組織の説明責任（出力の根拠・ログ・監査可能性）、(b) 機密情報・個人情報の保護、(c) サプライチェーン／学習データの法令遵守、(d) 生成AI特有のハルシネーション・偏り・有害出力、(e) ベンダーロックインと調達の継続性が主要論点となる。政府側には既に「生成AIの調達・利活用に係るガイドライン」があり、これを“運用のガードレール”として試用評価と一体運用することが合理的である。ガイドラインURL：`https://www.digital.go.jp/news/3579c42d-b11c-4756-b66e-3d3e35175623`[6]

背景と政策文脈

行政DXとしてのガバメントAIと「源内」の位置づけ

日本[7]では、人口減少・少子高齢化に伴う担い手不足への対応として、公共サービスの維持・強化にAI活用が不可避であるとの問題意識が公式に示されている。ガバメントAIは「政府による活用をAIの社会実装の起点」に据え、行政現場での実装・検証・評価・ナレッジ共有・国産AIの育成強化を進める枠組みとして説明される。[8]

この枠組みの実装基盤が生成AI利用環境「源内（げんない）」であり、デジタル庁内では2025年5月以降に利用が進められ、2026年度には全府省庁約18万人規模へ拡大する大規模実証（2026年5月〜2027年3月予定）が示されている。[9]

政策文脈として、AI戦略本部で「源内の徹底活用」や政府職員の大規模利用に言及があり、さらに「AI基本計画（2025-12-23閣議決定）」で“まず政府自らが先導的に利活用する”方針が示された旨がデジタル庁ページで述べられている（※当該計画本文の詳細要件は本報告書では未検証のため、ここではデジタル庁が引用した範囲に限定して扱う）。 [10]

国産AI推進と「調達→試用→評価→政府調達」の循環設計

公募結果の説明では、デジタル庁の狙いが明示的に三点整理されている。すなわち、(1) 政府での安全・安心な国産AI利用推進、(2) 行政現場からのフィードバックによる国産AI性能向上、(3) 政府調達を通じた安定的需要創出である。これは単なるPoCではなく、「調達可能性」を見据えた制度設計（評価後に有償調達を検討）という点で重要である。 [2]

この循環が成立する前提として、政府内ガバナンス（各府省の体制整備）が強調され、各府省庁は利活用促進とガバナンス強化のための組織的取組体制を構築し、生成AI調達・利活用ガイドラインに基づく対応や、CAIO（Chief AI Officer）による統括監理を推進するとされる。 [11]

セキュリティ・プライバシー要件の政策的位置づけ

公募の選定基準に「政府職員が機密性2情報を取り扱えるよう、十分なセキュリティを確保できること（ガバメントクラウド上の推論環境で動作）」が明記された点は、行政利用における実装要件を強く規定する。これは「国内で開発されたこと」だけでなく、「運用環境が政府要件を満たすこと」を採否に直結させる設計である。 [3]

一方で、個別の暗号要件、ログ保全期間、監査手続、インシデント対応フロー等の詳細は、公募結果ページでは未指定である（＝本報告書では未指定と扱う）。 [2]

公募要件と選定結果

公募の概要とスケジュール

公募結果ページは、(a) 応募件数（15件）と選定件数（7件）、(b) 試用評価の実施場所（源内）、(c) 2027年度の政府調達（有償）検討、(d) 2026年3月以降〜2027年4月以降までのマイルストーンを提示している。URL：`https://www.digital.go.jp/news/10d55c63-b3e1-42b9-9cc5-93a06943ae0e` [2]

報道（Yahoo!ニュース配信の一次媒体に相当：Impress Watch、ASCII）も、選定7モデル・試用開始時期（8月頃）・評価公表（2027年1月頃）・有償調達検討（2027年4月以降）など、公式発表と整合する事実関係を補足している。Impress Watch記事URL：`https://www.watch.impress.co.jp/docs/news/2091358.html` ／ ASCII記事URL：`https://ascii.jp/elem/000/004/379/4379302/` [12]

選定基準の整理

選定基準は6点で構造化されており、要点は以下である（原文は公式ページ参照）。 [2]

- **国内開発＋説明可能性**：開発経緯・開発方法・開発体制、独自開発か派生か（例：海外オープンモデルの追加学習）を具体的に説明できること。 [3]
- **行政実務での実用性**：対話型AI（チャット）または行政実務特化AIアプリへの組込みを想定し、実務で実用可能な性能を持つこと。 [2]
- **当日初開示の評価テスト**：デジタル庁作成の50問（試験当日に初めて開示）による評価テスト結果が優秀であること。 [2]

- **海外主要LLMとの比較ベンチマーク**：比較結果の提出と優秀性。特にハルシネーション、バイアス・差別的表現、有害コンテンツ生成等の安全性取組を説明でき、適切に実施していること。 [13]
- **学習データの法令遵守**：学習用データに関する法令遵守状況が具体的に説明でき、適切であること。 [2]
- **機密性2情報に耐える運用要件**：ガバメントクラウド上の推論環境で動作し、十分なセキュリティを確保できること。 [2]

この設計は、モデル性能（精度）と同等以上に、**説明責任（traceability）とガバナンス適合（法令遵守・安全性・運用要件）**を“採否の必須条件”に置いている点が特徴である（分析）。根拠は基準文言の構造に基づく解釈であり、個別モデルの採点内訳は未公表である。 [2]

選定された7モデルの一覧

以下は公式発表に基づく選定一覧である（五十音順）。 [2]

モデル名	応募主体（開発・提供主体）	備考
tsuzumi 2	株式会社NTTデータ [14]	モデル本体はNTTグループの「tsuzumi 2」と整合する公開資料があるが、応募主体はNTTデータ（＝提供主体）である点に留意（分析）。 [15]
CC Gov-LLM	カスタマークラウド株式会社 [16]	公開仕様は限定的（未指定が多い）。 [17]
Llama-3.1-ELYZA-JP-70B	KDDI株式会社 [18] ・株式会社ELYZA [19] （共同応募体）	海外ベースモデル（Llama 3.1-70B）に日本語追加学習を行った派生モデル。 [20]
Sarashina2 mini	ソフトバンク株式会社 [21]	法人向けAPIとして提供され、入力・出力データを学習に利用しない旨を明記。 [22]
cotomi v3	日本電気株式会社 [23]	v1は13B公表、v2以降はサイズ非公開との報道（公式未指定）。 [24]
Takane 32B	富士通株式会社 [25]	「Takane」はCohereと協業し企業向けに展開と説明。32Bはモデル名として公式発表に記載（＝パラメータ数は推定扱い）。 [26]
PLaMo 2.0 Prime	株式会社Preferred Networks [27]	GENIAC第2期成果を基に、31B級モデル群を“一から”開発した旨を技術ブログで説明。 [28]

技術仕様の比較と選定理由の分析

仕様比較表

本節の比較表は、(a) 開発元の一次情報（プレスリリース／技術ブログ／公式サービスページ）、(b) デジタル庁の選定結果ページ、(c) 一次媒体の報道（Impress、ASCII）に基づき、「公開されていない項目は未指定」として整理した。 [29]

> 注：推論に必要なGPUメモリは、量子化やKVキャッシュ、同時実行数、コンテキスト長等で大きく変動するため、表中の「推定VRAM」は“重みのみ”の概算である。概算式は、NTTが例

示した「必要GPUメモリサイズ＝パラメータ数×量子化サイズ/8bit（8bit量子化）」に基づく推定である。[30]

観点	tsuzumi 2	CC Gov-LLM	Llama-3.1-ELYZA-JP-70B	Sarashina2 mini	cotomi v3	Takane 32B	PLaMo 2.0 Prime
開発タイプ	スクラッチ開発（NTT説明）[31]	未指定[32]	派生（Meta Llama 3.1-70Bに日本語追加事前学習＋指示学習）[33]	スクラッチ開発と説明（サービスページ）[34]	v1は自社開発（13B）を公表、v3はサイズ未指定[35]	ベースはCohere Command R+、協業開発と説明[36]	GENIAC成果を元に1B/8B/31Bモデル群を“一から”開発[37]
パラメータ数	30B（公表）[38]	未指定[32]	70B（公表）[39]	未指定（“mini”と称す）[34]	未指定（v1=13B公表、v2以降非公開との報道）[35]	32B（モデル名から推定：公式は「Takane 32B」と表記）[2]	31B級（PLaMo 2.0 31Bを言及）[40]
学習データ概要	未指定（評価指標や安全性評価枠組みの説明はあり）[41]	未指定[42]	日本語追加事前学習＋指示学習（詳細未指定）[39]	学習データ収集・選定〜構築まで自社で一貫管理（詳細未指定）[34]	社内対話履歴の解析・活用に言及（学習配合詳細は未指定）[43]	追加学習・チューニングで日本語強化、詳細未指定[36]	高品質学習データセットの重要性を強調（詳細未指定）[40]
アーキテクチャ	未指定（ただしLLMとして評価枠組み提示）[44]	未指定	Llama系（Transformer系）※一般論＋Meta説明に基づく[45]	未指定	未指定（NECは学習方式・アーキ刷新を説明）[43]	Command R+ベース（Transformer系）[36]	未指定（ただし独自トークナイザ等の言及あり）[46]

観点	tsuzumi 2	CC Gov-LLM	Llama-3.1-ELYZA-JP-70B	Sarashina2 mini	cotomi v3	Takane 32B	PLaMo 2.0 Prime
ライセンス/提供形態	商用提供前提（詳細未指定）[38]	未指定（ただしオンプレ志向を示すPRあり）[42]	ベースはLlama 3.1 Community License（派生提供形態は未指定）[47]	法人向けAPI（料金は問い合わせ、規約ベース）[34]	商用サービス（詳細未指定）[48]	企業向け（DI PaaS/Kozuchi統合）[36]	Chat（登録で利用可）＋API（価格明記）[37]
推定VRAM（8bit重みのみ）	約30GB（公表例と整合）[30]	未指定	約70GB（推定：式に基づく）[49]	未指定	未指定	約32GB（推定：式に基づく）[50]	約31GB（推定：式に基づく）[51]
公開されている推論コスト情報	GPUコスト試算（A100 40GB相当1基など）[30]	未指定	未指定	価格は問い合わせ（未指定）[34]	未指定	未指定	API価格：入力60円／出力250円（1Mトークン）[37]
公開評価・ベンチ	知識/解析/指示遂行/安全性の統合評価枠組み提示[52]	未指定	日本語性能向上を主張（定量は未指定）[39]	SB Intuitionsの公開モデルは複数ベンチのスコア提示（※mini自体は未指定）[53]	Pro/LightでELYZA Tasks 100等のスコア公表（v3は未指定）[54]	JGLUEで世界最高水準と主張（数値表を提示）[55]	IFEval・ELYZA-tasks-100等の比較を公開[56]

選定理由の分析

公募結果ページは「基準」を明示する一方、モデルごとの採点・順位・選定理由の個別説明は未公表である。したがって本節は、**基準6点に対する“適合しやすさ”を公開情報から推定**する（推定であることを明記）。[2]

第一に、7モデルは「スクラッチ開発（純国産）」系（例：tsuzumi、Sarashina、cotomi、PLaMo）と、「海外ベースモデル派生」系（ELYZAのLlama派生、TakaneのCommand　R+ベース）を混在させた構成であ

る。これは、基準1で「独自開発か派生かの区別を説明可能」を求めつつ、派生モデルを排除していないことと整合する。 [57]

第二に、基準3（当日初開示50問）と基準4（海外主要LLMとの比較ベンチ）を満たすには、(a) 特定ベンチのスコアだけでなく、(b) 行政タスクに近い問合せでの頑健性（ハルシネーション耐性、規範的文章の正確さ、差別表現抑制等）が重要になる。公開情報上、TakaneはJGLUE等で高精度を示し、tsuzumi 2は指示遂行・安全性を含む枠組みを提示、PLaMoは指示追従評価（IFEval等）と価格まで含めた実装志向の情報を出しているため、評価設計に乗せやすい（推定）。 [58]

第三に、基準6（ガバメントクラウド上で動作）を満たすには、モデルの“重さ”だけでなく、推論基盤への実装容易性・運用分離（学習データ／入力ログの取り扱い）、SLAと障害対応が鍵になる。PLaMoはAPI提供を前提にしており、Sarashinaは法人APIとして「入力出力データを学習に利用しない」と明記している点で、運用要件に適合させやすい（推定）。ただし、政府内の具体的なログ・監査ポリシーは未指定であり、適合の最終判断はデジタル庁の契約・実装調整に依存する。 [59]

技術的評価と公開ベンチマークに基づく整理

ベンチマークの意味づけ

行政用途のLLM評価は、(1) 日本語言語理解（例：JGLUE）、(2) 指示追従・対話品質（例：Japanese MT-Bench系、IFEval系、ELYZA-tasks-100等）、(3) 安全性（有害出力、差別表現、誤情報、漏えい誘発耐性）、(4) 実務適合（RAGや規程・法令・公文書の扱い、監査性）の複合である、という前提が公募基準（特に基準2と4）から読み取れる。 [2]

JGLUEは日本語の一般言語理解能力を測るために構築された日本語NLUベンチマークであり、翻訳に依存せず日本語で“一から”構築したと説明される。 [60]
ELYZA-tasks-100は、日本語の複雑な指示・タスクへの追従を評価する用途で参照されており（PFN技術ブログでも評価対象として明記）、データセット自体は公開されている。 [61]

モデル別の強み・弱み（公開情報ベース）

tsuzumi 2（強み：実装コストとセキュリティ設計の一体化／弱み：外部再現可能な詳細仕様が限定的）

NTTの公開資料では、tsuzumi 2は30Bで「なぜ30Bか：Only 1 GPU → On-premise」という設計思想を前面に出し、推論時GPUコストの試算（A100 40GB相当1基＝約500万円、8bit量子化前提のメモリ算定式等）を示す。これは行政のスケール展開（多数ユーザー×継続運用）に対して、TCOを“モデル設計”の側で抑えにいくアプローチであり、ガバメントクラウドに載せる際にもインフラ設計上の見通しを与える。 [62]

また、知識・解析・指示遂行・安全性の4カテゴリで他モデル比較を提示し、指示遂行にM-IFEval_Ja、安全性にAnswerCarefully等を使う枠組みを説明している（ただし本資料の抜粋範囲では、完全な再現実験条件やスコア表が網羅的に提供されているとは言い切れず、政府評価での二次検証が重要になる）。 [63]

PLaMo 2.0 Prime（強み：評価と価格の透明性／弱み：推論基盤の仕様がWeb上で読み取りにくい）

PFNの技術ブログは、PLaMo 2世代をGENIAC第2期成果として位置づけ、1B/8B/31Bモデル群を“一から作った”と述べる。さらに、PLaMo 2.0 Primeの価格を「1Mトークンあたり入力60円、出力250円」と明記し、前世代比で約1/4以下にしたと説明する。公募基準のうち「運用を前提にした比較可能性（コスト＝調達の基礎情報）」という観点では突出している。 [64]

また別稿では、IFEval（指示追従）、ELYZA-tasks-100（日本語指示タスク）、pfgen（日本語生成能力評価）などを用いた比較を公開し、弱点として数学推論カテゴリの相対的低さにも触れている。これは“弱みの自己開示”であり、行政利用で重要となる説明責任の観点からはプラスに働き得る（推定）。 [65]

Takane（強み：日本語理解の定量提示と企業向け統制機能／弱み：選定モデル「32B」仕様の公開範囲が限定）

富士通はTakaneを企業向けLLMとして発表し、Cohereと協業して開発、KozuchiとDI PaaSに統合し、機微データを扱う産業向けに「安全なプライベート環境」で使える点を強調する。 [36]

性能面では、JGLUEで世界最高水準として、タスク別スコアと平均値を表として提示している。公募基準4（海外主要LLMとの比較）に対して、少なくとも日本語理解の一部について比較の“形”を提示している点が強い。 [55]

一方、今回の公募で選定されたのは「Takane 32B」であり、これはデジタル庁の記載上のモデル名で、富士通の2024年リリース文書における「Takane（中規模）」表現とは粒度が異なる。32B版の厳密な仕様（学習データ、コンテキスト長、提供形態、監査ログ仕様など）は公開情報からは確定できず、未指定とする。 [26]

cotomi v3（強み：高速性・実務導入を意識した評価指標／弱み：v3のサイズ等が非公開）

NECはcotomi Pro/Light（v2系に相当する位置づけ）で、ELYZA Tasks 100とJapanese MT-Bench等のベンチマーク結果を示し、速度面（GPU2枚でGPT-4比5倍以上など）や、セキュリティ・業務データ連携・情報漏えい／脆弱性といった導入懸念が顧客要件になることを明示した。 [43]

一方、cotomi v3については「v1は130億公表、v2以降は非公開で、今回も非公開」との報道があり、公式ページ（少なくとも本報告書で参照した範囲）ではパラメータ数・学習データ・推論要件が未指定である。よって、本報告書でも未指定として扱う。 [66]

Sarashina2 mini（強み：データ主権・学習への入力不使用明記／弱み：miniの仕様・ベンチの公開が限定）

ソフトバンクの法人向けSarashina APIの説明では、(a) スクラッチ開発、(b) 国内データセンターでの運用主権維持、(c) 学習データ収集・選定からモデル構築まで一貫管理、(d) そして「お客さまの入出力データをモデル学習に利用しない」と明記している。行政における機密性2情報の取扱いを考えると、(d)は“少なくともベンダー側の再学習に流用しない”という重要な契約論点に直結する。 [67]

ただし、Sarashina2 miniそのもののパラメータ数、コンテキスト長、外部比較ベンチマーク等は当該ページでは未指定であり、詳細は問い合わせとしている。よって、政府の試用評価の公表が、客観比較の主要根拠になる可能性が高い（推定）。 [67]

なお、同グループのSB Intuitionsは0.5B/1B/3BのSarashina2.2 InstructモデルをMITライセンスで公開し、ELYZA-tasks-100やJapanese MT Bench等のスコアも提示しているが、これは“公募で選定されたSarashina2 mini”と同一モデルであるとは断定できない（未指定）。 [68]

Llama-3.1-ELYZA-JP-70B（強み：開発方法が明確／弱み：行政向け運用仕様は公開資料だけでは確定しない）

ELYZAは、当該モデルがMetaのLlama-3.1-70Bをベースに、日本語追加事前学習と指示学習を行って開発した700億パラメータモデルであることを明確に述べている。したがって、公募基準1の「派生モデルであることの説明可能性」は高い。 [69]

ライセンス面では、ベースモデルは「LLAMA 3.1 COMMUNITY LICENSE AGREEMENT」に基づく利用となる。派生モデルの再配布・商用利用条件は、ELYZAの提供形態（安全なAPIサービス／共同開発）によって追加条件があり得るため、厳密な条件は未指定とするが、少なくともベースライセンスの存在は調達時の法務確認事項になる。[47]

推論要件は公開情報のみでは確定しないが、一般に70B級は8bit重みでも70GB程度のメモリを要するため（NTTが示した概算式に基づく推定）、単一GPU（80GB級）または複数GPU構成が現実的になる可能性が高い。これはガバメントクラウド上での同時利用者数を考えると、推論基盤費用の主要因になり得る（推定）。[70]

CC Gov-LLM（強み：政府向けを明示する可能性／弱み：公開仕様が不足）

カスタマークラウドの公表資料（PR）からは、政府向けAIとしての打ち出しが読み取れるが、モデルサイズ、学習データ、評価指標、推論要件などの技術仕様は未指定である。公募基準上は、応募時点でデジタル庁に対して詳細情報を提示していると推定されるものの、公開情報のみで第三者検証はできない。[32]

試用評価（2026年8月以降）で焦点化すると考えられる技術論点

公募基準（特に基準2・3・4・6）から逆算すると、2026年8月以降の試用評価は、少なくとも以下の論点を横断的に測る設計になる可能性が高い（推定）。[71]

```
flowchart TD
  A[行政ユースケース] --> B[対話: 職員の普段使い]
  A --> C[業務アプリ: 検索・要約・分類・起案補助]
  B --> D[指示追従・日本語品質]
  B --> E[安全性: 差別/有害/誤情報/漏えい誘発]
  C --> F[RAG適合: 根拠提示・参照整合]
  C --> G[監査性: ログ・再現・説明責任]
  H[運用制約] --> I[ガバメントクラウド実装]
  H --> J[機密性2情報の取り扱い]
  I --> K[性能/コスト: 同時利用・遅延・GPU]
  J --> L[データガバナンス: 入出力の保護]
```

この中で、(a)“誤情報・漏えい誘発への耐性”と (b)“監査可能性”は、行政特有の説明責任要請により、民間PoCより重く評価される蓋然性がある（推定）。ただし、実際の配点は未公表であり、2027年1月の評価公表を待つ必要がある。[8]

運用計画とリスク

運用計画の整理

運用計画として公式に明示されているのは、(1) 2026年3月以降の契約・実装調整、(2) 2026年5月頃の源内大規模実証開始、(3) 2026年8月頃からの国内LLM試用開始、(4) 2027年1月頃の評価公表、(5) 2027年4月以降の有償調達（予定）である。[71]

また、源内は単なるチャット基盤ではなく、汎用AIアプリ、行政実務向けAIアプリ、AI用データセット整備、利用状況ダッシュボード等を含む“利用環境”として説明されており、LLMはその中核部品として位置づけられる。これは、モデル単体評価に留まらず、アプリ連携・データ連携・観測可能性（利用ログ）の整備が運用価値を左右することを意味する（分析）。[72]

主要リスクと緩和策（リスク×可能性×影響）

以下は、(a) 公募基準、(b) デジタル庁の生成AI調達・利活用ガイドライン策定趣旨（利活用促進とリスク管理を表裏一体で進める）、(c) AIリスク管理の一般枠組み（NIST AI RMFの目的：AIリスクを管理し信頼性を高める）に基づき、行政運用での主要リスクを整理したものである。ガイドラインURL：`https://www.digital.go.jp/news/3579c42d-b11c-4756-b66e-3d3e35175623` [73]

情報漏えい（入力・出力・ログ経由）
可能性：中（プロンプト注入、誤操作、RAG連携等で増加）／影響：極大（行政の信頼失墜、法令違反、個人情報侵害）
緩和策：機密区分に応じた入力制御、プロンプトテンプレートとDLP、RAGデータのアクセス制御、強制ログ、職員教育。特に、ベンダー側が「入力出力を学習に利用しない」などのデータ取扱い方針を明記している場合は契約で拘束する（例：Sarashina APIの明記）。 [74]

ハルシネーション（誤情報）による行政判断の誤り
可能性：中〜高（LLM一般特性）／影響：大（誤った起案・誤った対外説明、政策ミス）
緩和策：回答根拠の提示（RAGの参照先固定化）、重要文書は人間のレビューを必須化、評価テスト（当日初開示50問）を継続運用に拡張し、業務ドメイン別の“誤りパターン”を可視化する。 [8]

偏り・差別的表現、有害出力、違法行為助長
可能性：中／影響：大（人権侵害、行政の中立性毀損）
緩和策：基準4で求められる安全性取組（バイアス・差別・有害出力の抑制）を、試用段階でも定量監視・ルール化する。モデル提供側がガード用モデル・仕組みを提供する例もあるため（例：Sarashina2　guardの提供）、アーキテクチャとして組み込む。 [75]

学習データの法令遵守・知財リスク
可能性：中（外部に詳細が公開されない場合、後日問題化するリスク）／影響：大（訴訟、利用停止、調達停止）
緩和策：基準5に沿って、学習データの法令遵守を説明可能にし、監査可能なエビデンス提出を契約要件化する。政府側は評価結果の一部公表に同意を求める運用も示しており（公募資料内の記述）、透明化のインセンティブ設計が重要になる。 [76]

ベンダーロックイン（モデル更新・価格・撤退）
可能性：中／影響：中〜大（継続運用不可、コスト増、機能停滞）
緩和策：複数モデル並行試用（今回の7モデル選定自体が緩和策）、標準インターフェース（API互換、ログ形式、評価指標）の整備、政府側の調達要件の共通化。デジタル庁が国産AIへの安定需要創出を掲げる一方、競争環境維持のための“出口戦略”が必要（分析）。 [8]

監査・説明責任（誰が何を根拠に出力したか）
可能性：高（行政は説明責任が本質）／影響：大
緩和策：ガイドラインを統一のガードレールとし、ログ（プロンプト、参照文書、モデルバージョン、出力）を監査可能に保全する。各府省庁はCAIOを中心にガバナンス体制を整備することが示されているため、試用段階からCAIOがレビューできる指標・ダッシュボードを整える。 [77]

自治体・省庁展開の含意（未指定事項の明確化）

公式発表では、源内の大規模実証対象は全府省庁約18万人規模であり、地方機関職員の一部も含めた提供が示唆される一方、自治体（地方公共団体）への横展開方式（共同利用の枠組み、費用負担、データ連携、ネットワーク分離等）は本件公募結果ページでは未指定である。よって、本報告書では「将来の展開可能性は示唆されるが、制度設計は未指定」と整理する。 [78]

今後の影響と示唆

政策的影響

本公募は、(a) 国産AI利用の推進、(b) 行政現場のフィードバックによる性能向上、(c) 政府調達による需要創出、を一連の政策ループとして明示している点で、従来の単発PoCより政策的コミットメントが強い。2027年度の有償調達まで見据えることで、国内モデル開発に対する“見込み需要”を政策的に形成しようとしている。 [3]

また、ガイドライン策定により「利活用促進とリスク管理を表裏一体で進める」ことが公式に位置づけられたため、今後は調達仕様（ログ、監査、データ取扱い、説明可能性）が標準化され、民間側の提案もその枠組みに収斂する可能性が高い（推定）。 [79]

産業・研究開発への波及

産業面では、7モデルが「API提供型」「プライベート環境提供型」「オンプレ志向（低コストGPU構成提示）」を混在させている。これは行政需要が、最高性能一辺倒ではなく、**運用可能性（コスト・セキュリティ・監査）**を重視する市場であることを示すシグナルになり得る。例えば、PLaMoは価格と評価を公開し、tsuzumi 2は1GPU・オンプレを前面に出し、Takaneはプライベート環境と統制機能を強調する。 [80]

研究開発面では、GENIAC成果に基づくモデル（PLaMo）や、海外ベースを日本語特化するアプローチ（ELYZA）が同時に採用されており、「フルスクラッチ vs 派生最適化」の二極ではなく、用途と統制要件に応じた多様化が進む可能性がある（推定）。 [81]

長期的課題と提言

第一に、評価公表（2027年1月予定）の粒度が、国内エコシステムの学習速度を左右する。公募基準は明確だが、個別モデルの“良かった点／悪かった点”がブラックボックスのままだと、(1) 行政側の調達合理性が説明しづらく、(2) ベンダー側も改善点を特定しづらい。したがって、機密を除いた範囲で、評価設計（タスク類型、誤り分類、安全性テスト）と集計結果を公開することが、政策目的（フィードバックで性能向上）に直結する（提言）。 [8]

第二に、ガバメントクラウド実装を要件化したため、モデル性能だけでなく「運用アーキテクチャ（監査ログ、ガードレール、RAG、権限管理）」が選定の実質を決める。源内が“アプリ＋データ＋ダッシュボード”の環境として構想されている以上、政府はモデル比較と同時に“共通運用部品（ログ、評価、フィルタリング）”の標準化を急ぐべきである（提言）。 [82]

第三に、ライセンス・データ法務の論点は、派生モデル（Llama系など）で特に重要になる。ベースライセンス（Llama 3.1 Community License）の遵守、派生提供形態（API／重み配布）の条件、学習データの適法性を、調達段階で監査可能にしておく必要がある（提言）。 [83]

参考情報

主要出典URLリスト

（ユーザー要望によりURLを明記。リンク先内容は本報告書の記述箇所で逐次引用した。）

【デジタル庁（公式）】
https://www.digital.go.jp/news/10d55c63-b3e1-42b9-9cc5-93a06943ae0e

https://www.digital.go.jp/news/2d69c287-2897-46d8-a28f-ea5a1fc9bce9
https://www.digital.go.jp/news/3579c42d-b11c-4756-b66e-3d3e35175623

【主要報道（Yahoo!ニュース配信記事の一次媒体として参照）】
https://www.watch.impress.co.jp/docs/news/2091358.html
https://ascii.jp/elem/000/004/379/4379302/

【モデル開発元・公式資料（一次情報中心）】
（tsuzumi 2：NTT公開資料）
https://group.ntt/jp/newsrelease/2025/10/20/pdf/251020aa.pdf

（Sarashina API：ソフトバンク公式）
https://www.softbank.jp/biz/services/ai/sarashina-api/

（SB Intuitions 技術ブログ：公開モデルの評価例）
https://www.sbintuitions.co.jp/blog/entry/2025/03/07/093143
https://www.sbintuitions.co.jp/blog/entry/2025/03/06/112144

（cotomi：NEC公式）
https://jpn.nec.com/press/202404/20240424_01.html
https://jpn.nec.com/press/202307/20230706_02.html

（Takane：富士通公式）
https://www.fujitsu.com/global/about/resources/news/press-releases/2024/0930-01.html
https://en-documents.research.global.fujitsu.com/takane-enterprise-llm-with-generative-ai-reconstruction/

（PLaMo：PFN技術ブログ）
https://tech.preferred.jp/ja/blog/plamo-2-prime-release/
https://tech.preferred.jp/ja/blog/plamo-2-posttrain/

（Llama 3.1 派生モデル：ELYZA公式プレスリリース）
https://prtimes.jp/main/html/rd/p/000000052.000047565.html

【原典（ライセンス／評価ベンチ）】
（Llama 3.1 Community License）
https://github.com/meta-llama/llama-models/blob/main/models/llama3_1/LICENSE
（Meta Llama 3.1 発表）
https://ai.meta.com/blog/meta-llama-3-1/
（JGLUE 原典）
https://aclanthology.org/2022.lrec-1.317/
（ELYZA-tasks-100 データセット）
https://huggingface.co/datasets/elyza/ELYZA-tasks-100
（NIST AI RMF 1.0（枠組みの背景参照））
https://www.govinfo.gov/app/details/GOVPUB-C13-ef6882c6bd970bc733de394f469ea3f3

本報告書における未指定・推定の扱い

- 公開情報で確認できない項目（例：一部モデルのパラメータ数、学習データ詳細、推論基盤仕様、価格、監査ログ仕様等）は「未指定」とした。 [84]

- 推定を行った箇所（例：8bit量子化時の重みメモリ概算）は、根拠（NTTが提示した概算式）を明示し、運用時にはKVキャッシュ等で増減する旨を併記した。 [30]

---

[1] [18] [30] [31] [38] [41] [44] [49] [50] [51] [52] [62] [63] [70] https://group.ntt/jp/newsrelease/2025/10/20/pdf/251020aa.pdf
https://group.ntt/jp/newsrelease/2025/10/20/pdf/251020aa.pdf

[2] [3] [4] [8] [13] [14] [15] [16] [17] [20] [21] [22] [23] [24] [26] [28] [29] [57] [71] [73] [74] [75] [76] [84] https://www.digital.go.jp/news/10d55c63-b3e1-42b9-9cc5-93a06943ae0e
https://www.digital.go.jp/news/10d55c63-b3e1-42b9-9cc5-93a06943ae0e

[5] [27] [37] [40] [64] [80] [81] https://tech.preferred.jp/ja/blog/plamo-2-prime-release/
https://tech.preferred.jp/ja/blog/plamo-2-prime-release/

[6] [79] https://www.digital.go.jp/news/3579c42d-b11c-4756-b66e-3d3e35175623
https://www.digital.go.jp/news/3579c42d-b11c-4756-b66e-3d3e35175623

[7] [25] [34] [59] [67] https://www.softbank.jp/biz/services/ai/sarashina-api/
https://www.softbank.jp/biz/services/ai/sarashina-api/

[9] [10] [11] [77] [78] https://www.digital.go.jp/news/2d69c287-2897-46d8-a28f-ea5a1fc9bce9
https://www.digital.go.jp/news/2d69c287-2897-46d8-a28f-ea5a1fc9bce9

[12] https://www.watch.impress.co.jp/docs/news/2091358.html
https://www.watch.impress.co.jp/docs/news/2091358.html

[19] [43] [48] [54] https://jpn.nec.com/press/202404/20240424_01.html
https://jpn.nec.com/press/202404/20240424_01.html

[32] [42] https://prtimes.jp/main/html/rd/p/000000715.000099810.html
https://prtimes.jp/main/html/rd/p/000000715.000099810.html

[33] [39] [69] https://prtimes.jp/main/html/rd/p/000000052.000047565.html
https://prtimes.jp/main/html/rd/p/000000052.000047565.html

[35] https://jpn.nec.com/press/202307/20230706_02.html
https://jpn.nec.com/press/202307/20230706_02.html

[36] [55] [58] https://www.fujitsu.com/global/about/resources/news/press-releases/2024/0930-01.html
https://www.fujitsu.com/global/about/resources/news/press-releases/2024/0930-01.html

[45] https://ai.meta.com/blog/meta-llama-3-1/
https://ai.meta.com/blog/meta-llama-3-1/

[46] https://forest.watch.impress.co.jp/docs/news/2016570.html
https://forest.watch.impress.co.jp/docs/news/2016570.html

[47] [83] https://github.com/meta-llama/llama-models/blob/main/models/llama3_1/LICENSE
https://github.com/meta-llama/llama-models/blob/main/models/llama3_1/LICENSE

[53] [68] https://www.sbintuitions.co.jp/blog/entry/2025/03/07/093143
https://www.sbintuitions.co.jp/blog/entry/2025/03/07/093143

[56] [65] https://tech.preferred.jp/ja/blog/plamo-2-posttrain/
https://tech.preferred.jp/ja/blog/plamo-2-posttrain/

[60] https://aclanthology.org/2022.lrec-1.317/
https://aclanthology.org/2022.lrec-1.317/

61 https://huggingface.co/datasets/elyza/ELYZA-tasks-100
https://huggingface.co/datasets/elyza/ELYZA-tasks-100

66 https://cloud.watch.impress.co.jp/docs/news/2029970.html
https://cloud.watch.impress.co.jp/docs/news/2029970.html

72 82 https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/eb376409-664f-4f47-8bc9-cc95447908e4/5145e700/20260113_meeting_ai-advisory_outline_02.pdf
https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/eb376409-664f-4f47-8bc9-cc95447908e4/5145e700/20260113_meeting_ai-advisory_outline_02.pdf