

NTTの大規模言語モデル「tsuzumi 2」調査レポート

1. 技術性能（日本語処理能力・推論精度・長文対応力・1GPU動作・省電力性・モデルサイズ）

図1: GPT-5クラスの超大型モデルとtsuzumi 2のベンチマーク比較（MT-bench, 日本語）。赤線がtsuzumi 2、青線がGPT-5であり、多くの評価項目でtsuzumi 2がGPT-5に近いスコアを示している^①。これは**tsuzumi 2が超大型モデルに迫るトップクラスの性能を持つこと**を示唆している。特に日本語での知識問答、内容解析、指示遵守（フォーマット順守など）、安全性といったビジネス上重要な基本性能で、**数倍大きいフラッグシップモデルに匹敵する水準**を達成しており、サイズ当たりの**コストパフォーマンスに優れる**^②。

モデルサイズと効率: tsuzumi 2のモデル規模は約**300億パラメータ**と公表されており、初代tsuzumi（約70億パラメータ）から大幅な増強が行われた^③。それにもかかわらず、推論時の要求メモリは抑えられていて**1枚のGPU（40GB以下メモリ搭載）で動作可能な省リソース設計**を維持している^④。最新の大型モデルが複数GPUや数百GBメモリを必要とする中、tsuzumi 2は**1GPU環境で動作し**、電力消費を抑えつつ高性能を発揮することが特徴だ^⑤。この軽量性により、**従来モデル(tsuzumi 1)の強みだった低コスト運用性**も継承されている^{⑥ ⑦}。NTTの発表によれば、tsuzumi 2は同規模クラス他モデル中で**世界トップレベルの日本語性能**を記録し^⑧、日本語文脈の理解力・生成力で**GPT-4相当の大型モデルにも匹敵**する結果を示したとされる^⑨。実際、日本語のベンチマーク評価でtsuzumi 2は同サイズ帯モデルを上回り、数倍規模の海外モデル（GPT-4クラス）と同等以上の日本語理解力を確認しています^⑩。

推論精度と長文対応: tsuzumi 2は高度な推論タスクや専門知識が要求される質問にも対応できるよう性能強化されています。社内評価では、例えばNTTの財務システムに関する問い合わせ対応業務への適用テストで、**他社の先進モデルと同等以上の性能**を示したと報告されています^⑪。特に金融・医療・公共といった専門知識領域では知識量を大幅に増強したため、複雑な質問への回答精度が向上しています。**長文入力への対応力**についても改善が図られました。東京通信大学での検証では、「複雑な文脈理解や長文ドキュメント処理が安定し、複雑なタスクの複合工程でも実用水準に達している」と評価されています^⑫。これは初代モデルに比べ長大な文章や文書の要約・分析処理において**安定性・正確性が向上**していることを意味します。実演デモでは、企業が公開した長文のニュースリリース文章に対し、その改善点を高速にリストアップするタスクが紹介されました。OpenAI提供の20B規模モデルでは指示に反し表形式で回答してしまった一方、tsuzumi 2は**指示通り箇条書きのリスト形式で素早く出力し**、原文の要点もきちんと維持できることをアピールしています^⑬。このように**指示遵守や長文からの要点抽出**といった実用面でも優れた性能を発揮しています。

省電力性と環境負荷: 大規模な生成AI活用時に問題となる消費電力について、tsuzumi 2はモデル自体の軽量化により対処しています。NTTはChatGPTなど海外LLMの普及に伴う**電力消費増大**を課題視しており^⑭、tsuzumi 2の開発背景には**計算資源効率の向上**がありました。1GPUで動作する設計は無駄な電力を消費する巨大モデルとは一線を画し、利用あたりの電力コスト低減に寄与します。NTT発表資料のGPUコスト比較によれば、tsuzumi 2を動作させるサーバの初期投資額は、桁違いに巨大なモデル(GPT-4クラスの数百億〜1兆パラメータ規模)を動かす場合に比べ極めて小さく、**およそ1/10以下のサーバコストで済む**とのこと^⑮。このように**モデルサイズの割に低リソースで運用**できるtsuzumi 2は、環境負荷の低減と運用コスト削減の両面で優れた特徴を持っています。

2. 特徴（純国産モデル、セキュリティ、オンプレミス対応、プライベートクラウド対応など）

純国産・フルスクラッチ開発: tsuzumi 2最大の特徴の一つは「純国産」である点です。NTTがゼロから独自開発したモデルであり、他社の公開モデルを流用していません¹⁶。そのため、学習データの選定から権利処理・品質管理・バイアス対策までNTTが自社でコントロールしており、開発プロセスの透明性とモデルへの信頼性を高めています¹⁶。昨今は生成AIの学習データ無断利用に関する訴訟も起きており、モデル開発段階の著作権や倫理面の問題がクローズアップされています。NTTはそうしたリスクを回避し、利用企業側も安心して使えるモデルを目指して、あえてフルスクラッチでtsuzumiシリーズを構築しています¹⁶。初代tsuzumi同様、tsuzumi 2も日本の国内法規に準拠して開発されたモデルであり、日本企業のデータガバナンス要件にも適合しやすいとされています¹⁶。

高セキュリティ・データ保護: tsuzumi 2は機密データを安全に扱える点も大きな特徴です。上述の純国産ゆえのデータ管理だけでなく、提供形態もオンプレミスやプライベートクラウドでの運用に対応しています¹⁵。つまり、利用企業や自治体は自社のサーバ内や閉域環境でtsuzumi 2を動かし、データを外部に出さずAI活用することが可能です¹⁰。クラウド上の外国製LLMでは社外へのデータ送信による情報漏洩リスクや、利用規約上データを学習に使われる懸念があります。それに対しtsuzumi 2はデータ流出リスクを抑え、国内拠点内で処理を完結できるため、金融機関や自治体など高いセキュリティ基準を持つユーザにも採用しやすい設計です¹⁰。実際、東京通信大学がtsuzumi 2を採用した理由も「学生・教職員データを学内に留める」という要件を満たすクラウド依存のない国産LLMだったからであり¹⁷、このように国内企業・機関のセキュリティニーズに応える特徴が評価されています。

オンプレ/プライベートクラウド対応と運用効率: tsuzumi 2は1GPUで動作可能な軽量モデルであるため、企業の自社サーバや私有クラウド上での手軽なデプロイを可能にしています¹⁸¹⁵。NTTによれば、超大型モデル利用時に課題となるAPI利用料の高騰やクラウド課金の負担に対し、ある程度利用頻度が高い顧客であれば「自前でモデルサーバを運用した方が安上がり」なケースが増えるだろうと指摘しています¹⁹。tsuzumi 2はそうした需要に応えるために開発された低コスト・高効率運用モデルであり、自社運用によってランニングコストを抑制できる点を強みとしています¹⁹⁷。また、オンプレ対応はクラウド接続が難しい環境（例えば閉域ネットワーク内システムやネット非接続環境）でもAIを活用できることを意味し、産業インフラ系や地方自治体などクラウド利用に制約のある領域への導入も視野に入ります²⁰²¹。NTTはtsuzumi 2を「高性能・高セキュア・低コストな純国産LLM」と位置付けており、この三拍子が日本の企業DX推進に資するものとアピールしています²⁰。

業務ニーズに合わせた最適化: tsuzumi 2は企業での業務利用を念頭に、実ビジネスで頻出するユースケースに特化した強化が行われています。特にNTTの調査では企業利用の約8割を占めるという「ドキュメントに対するQA（検索＋要約）」「長文からの情報抽出・要約」といったタスクに対し、集中的に性能向上が図られました²²。そのため、ユーザ企業は自社ドキュメントをtsuzumi 2に読み込ませてQ&Aや要約させる際、高精度な回答や要約結果を得やすくなっています。また、出力形式の柔軟さも特徴です。例えば定型フォーマットの報告書に沿った要約文を生成したり、文書から特定情報を抽出してJSON形式で出力するといった細かな指示にも従えるよう、モデルの指示遂行能力を強化しています²³。これは企業利用で「フォーマットを守った出力」や「構造化データ化」が求められるケースが多いことを踏まえた設計で、tsuzumi 2はユーザの指示通りに出力様式を整える能力が向上しています²³²⁴。実例として、あるニュースリリース文の要約タスクでtsuzumi 2は指定通り箇条書きリストで改善点を出力し、別の契約書チェックタスクでは入力したチェックリスト項目に照らして契約書草案の不足条項や問題点を表形式で抽出するなど、実務で有用な応用例が報告されています¹³²⁵。

3. 他のLLMとの比較評価（性能、コスト、安全性、RAG・Fine-tuning性能など）

tsuzumi 2の位置付けを理解するため、主要な他のLLM（GPT-4、Google Gemini、Anthropic Claude、Meta社Llama2など）と**性能・運用面の比較**を以下にまとめます。

モデル名	開発企業・提供時期	パラメータ規模	日本語対応・性能	長文入力対応	導入形態・環境	コスト構造	安全性・ガバナンス	カスタマイズ性 (RAG/FT)
tsuzumi 2	NTT（日本）2025年発表	約 300億 ³	日本語特化 ：同規模モデル中トップクラス ⁸ 。主要指標で数倍大きいモデルに匹敵 ¹⁰	数千トークン級？長文処理の安定性向上 ¹²	オンプレ/プライベートクラウド対応 （自社サーバで運用可） ¹⁵	初期導入＋自前GPU 。大量利用時はクラウドより低コスト ¹⁹	国内開発 でデータ管理徹底。初期学習で安全性考慮（不適切発言抑制） ²	高度なチューニング性 ：RAGで外部知識補強、FTで業界特化容易 ^{26 27} （少量データで高精度） ²⁸
GPT-4	OpenAI（米）2023年発売	非公開 （推定数千億～1兆以上）	多言語対応 ：英語で最高水準。日本語も高精度（GPT-4はWASEDAベンチで東大合格レベルとの報告もあり）。ただし日本特有文脈や制度理解は限定的。	8k/32kトークン （標準版8k、拡張コネクスト版32k）	クラウドAPIのみ提供 （オンプレ不可）	従量課金制 （1000トークンあたり数円～十数円。高精度だが利用量が多いと高コスト）	厳格な安全策 ：OpenAIによるRLHF調整済みで有害発言や個人情報出力を抑制 ² 。ただし制御不能時の責任分界が課題。	ユーザーによる微調整不可 （Fine-tuning一般開放されず※一部GPT-3.5のみ）。RAG用途ではプラグインやベクタDBを併用。

モデル名	開発企業・提供時期	パラメータ規模	日本語対応・性能	長文入力対応	導入形態・環境	コスト構造	安全性・ガバナンス	カスタマイズ性 (RAG/FT)
Gemini (推定)	Google (米) ※2024年公開予測	非公開 (超大規模モデルと推定)	多言語・マルチモーダル対応 予定：詳細未公開だが画像も処理可能な次世代モデルと報道あり。 日本語対応も強化と予想。	不明（おそらく長大コンテキスト対応）	クラウド提供（GCP等）予定。オンプレ提供の情報なし。	未公開 (Google Cloud サービス内組み込みか、API課金か)	GoogleのAI原則に準拠。検証済みの安全策導入と予想。	未公開（企業向けにカスタム可能API用意の可能性）。
Claude 2	Anthropic (米) 2023年発表	非公開 (推定70億×複数頭=総計数千億?)	英語特化：会話や創作で高性能。 日本語は非最適化で精度中程度との指摘も。	最大100kトークン (超長文チャットに対応)	クラウドAPI提供 (オンプレ不可)	従量課金制（入力1Mトークンあたり\$11など比較的高額）	憲法AIによる独自安全策：人権尊重や有害回避の原則を組込。安全だが過度応答拒否の指摘も。	Fine-tuning非公開 (Anthropic側でのみ実施可能)。 RAG用途には長文プロンプトで対応可能（大容量コンテキスト活用）。

モデル名	開発企業・提供時期	パラメータ規模	日本語対応・性能	長文入力対応	導入形態・環境	コスト構造	安全性・ガバナンス	カスタマイズ性 (RAG/FT)
Llama 2	Meta (米) 2023年公開	7B/ 13B/ 70Bなど	多言語対応 (日本語可) : Chatモデルは日本語も一定可読。ただし生成品質は英語ほど高くなく、和文は要微調整。	4kトークン前後 (長文版やcommunity版で拡張可能)	オープンソース (モデル重み公開。オンプレ自由)	無料 (モデル自体無償利用可。計算リソースは自前負担)	基本的な安全調整済み : 有害出力を抑えるチューニング済。ただオープンモデルのため再調整や変更も可能。	非常にカスタマイズ容易 : 社内データでのFine-tuningやLoRA適用が可能。用途に応じ変更・再配布も許諾範囲で可能。

※表中の**RAG** (Retrieval-Augmented Generation) は外部データ検索併用による生成、**FT** (Fine-Tuning) は追加学習調整を指します。

上表より、tsuzumi 2は特に**日本語能力とオンプレ運用性で優位性**が見られます。他方、**多言語対応やマルチモーダル対応**は現時点で未実装であり²⁹、グローバル展開するには今後の課題と言えます³⁰。GPT-4やClaudeなど海外LLMは英語圏では卓越した性能を持ちますが、**利用コストやデータ主権**の面で企業に悩みを与えています¹⁴。tsuzumi 2はそれらに対し**日本語特化の高性能とオンプレミス運用による低コスト・高セキュリティ**を武器に差別化を図っています¹⁰¹⁸。また**RAGやFine-tuningによる業務特化の容易さ**も特徴で、NTTの比較では金融資格試験対策に必要な追加学習データ量が、Google系27Bモデルでは1900問必要だったのに対しtsuzumi 2は**10分の1の200問で済んだ**とされています²⁷。このように**少ないデータで高精度化**できる点は企業が自社知識を組み込む上で大きなメリットです。

4. 評価・評判 (導入事例・コメント、SNS・メディアの反応、専門家評価)

企業や団体からの評価: tsuzumi 2は発表時点で具体的な導入事例が報告されており、初期評価は上々です。特に教育分野では、**東京通信大学**が国内初の導入事例として挙げられます。同大学は「クラウドに頼らず学内でLLMを運用したい」という要件から国産LLMを検討しており、複数モデルの事前実証を経てtsuzumi 2の採用を決定しました¹⁷。評価ポイントは「**複雑な文脈理解や長文ドキュメント処理が安定し**、複雑なタスクの複合工程でも実用水準に達した」ことであり、授業Q&A高度化、教材・試験作成支援、履修・進路相談のパーソナライズといった用途で実践投入される予定です¹⁷³¹。これは教育現場から見てもtsuzumi 2の**日本語理解力・安定性**が十分実用に耐えるとの評価と言えます。またNTTによれば、他にも金融機関や自治体などから**事前検証の問い合わせが多数寄せられており**、既に2件の具体的事例が生まれているとされています³²。金融・自治体・医療分野での専門知識強化を評価し、自社システムへの組み込み検討を進める企業も多いようです。NTTグループ社内でも、tsuzumi 2を用いた社内問合わせ対応への試験導入で**先行モデルより大幅な精度向上 (回答精度4倍向上)**を確認したとされ³、社内関係者から高い評価を得ています。

SNS・メディアの反応: 2025年10月20日のtsuzumi 2発表は、日本のAI業界において大きな話題となりました。IT系メディア各社は「世界トップクラスの日本語性能を実現」「国産AI開発競争の新章」などを見出しを付けて報道しています³³³⁴。発表会に出席したNTT島田社長が「**“国産AI開発競争”に負けられない**」と述べたことも報じられ³⁴、国内AI開発への強い意気込みが伝わりました。この発言は各種ニュースサイトやSNSでも引用され、NTTがチャットGPTブームに対抗し**日本発のLLMで存在感を示す決意**を感じるとの反響がありました。専門家や研究者のSNSでも「日本語に特化した独自モデルの登場は重要」「海外モデル依存から脱却する第一歩」など肯定的な意見が見られ、生成AIの国産化に期待を寄せる声が多く上がりました。一方で「性能アピール通りなら凄いが、汎用性や英語対応は今後の課題」「実際に使ってみないと安全性や細かな品質は判断できない」といった慎重な意見も一部にあり、今後の検証結果を注視する姿勢も見受けられます。

専門家の評価・分析: AI専門メディアや有識者もtsuzumi 2の意義についてコメントしています。Business Insider Japanは「日本語性能や専門知識を強化しつつ軽量のAIモデル」とその**バランスの良さ**を評価し、特に**金融・医療・公共分野の知識強化**によるファインチューニング効率向上（追加学習データ削減）の点を注目しています²⁷。また、発表会デモで見せた**指示遵守能力の高さや出力速度の速さ**にも触れ、実ビジネスでの使い勝手向上を実感できる内容だったと伝えています¹³。一方、AI領域のブロガーによる分析では、tsuzumi 2の登場は「**国内モデル化**」の**象徴的一步**と評価しつつも、「モデル競争・活用状況・収益化までの道のりは簡単ではない」と指摘されています³⁵。具体的には、モデル性能そのもの以上に**運用体制やエコシステム構築、ビジネスモデル設計**が成否を分ける可能性があるとし、NTTが掲げる大きな市場目標（後述の5000億円規模）を達成するには競争環境や技術進展など多くの課題をクリアする必要があると論じられています³⁶³⁵。総じて、専門家筋からは「**日本語に強いLLM**」としてtsuzumi 2への期待は高く、同時に「今後、海外大型モデルとどこまで渡り合えるか」「日本市場での普及をどう実現するか」に注目が集まっています³⁷³⁵。

5. 用途・導入分野（金融・医療・行政・教育・製造などでの導入状況・ユースケース）

tsuzumi 2は幅広い業界での活用が見込まれており、既にいくつかの分野で導入・実証が進んでいます。以下、主要分野ごとのユースケースと導入状況を整理します。

- ・**金融分野:** 金融業では社内問い合わせ対応や文書分析、自動報告書作成などでLLM活用が注目されています。NTTは社内の財務システム問い合わせ対応にtsuzumi 2を試行し、**回答精度が従来比4倍向上**するなど有望な結果を得ました³。このことから、銀行・保険会社における**社内ヘルプデスク**や、証券会社での**目論見書・決算書の要約**、フィンテック企業での**顧客向けQ&Aチャットボット**などにtsuzumi 2が応用できる可能性があります。実際、NTTの生成AIソリューションは既にドイツの大手自動車メーカーBMWの社内活用や国内外の銀行で採用例があると報じられており³⁸、こうした導入先では今後tsuzumiシリーズへの置き換え・アップグレードも期待されます。金融分野は規制上クラウド利用が制約されるケースもありますが、tsuzumi 2はオンプレ運用可能なため**顧客データを外部に出せない銀行等でも導入検討が進みやすい**でしょう。
- ・**医療分野:** 医療・ヘルスケア領域でも、大量の医学文献や電子カルテ情報を処理できるLLMへの期待が高まっています。tsuzumi 2は**医療分野の知識を特に強化**して学習されており³⁹、臨床現場や医薬品開発での活用に向けた素地が整えられています。例えば製薬企業が持つ試験データの要約、新薬の承認申請資料の自動要約、医師の問い合わせ対応（症例ベースのQ&A）などに応用できるでしょう。現在具体的な医療機関での導入事例は公表されていないものの、NTTは医療分野をtsuzumi 2の重点領域と位置付けており、専門医学データを学習済みである点から**医療専門問答や診療ガイドラインの要約**といった高度なタスクでも高精度が期待されます³⁹。医療情報は機密性が高く海外クラウドに載せにくいですが、tsuzumi 2であれば病院内サーバで完結運用できるため、今後病院や研究機関での採用も十分考えられます。医療AIベンチャーとの協業や、NTT東日本が展開する地域医療クラウドでの組み込みなど、**医療データ×国産LLM**のユースケース創出に注目が集まっています。

・**行政（自治体）分野**： 地方自治体や行政機関でも、住民対応や事務効率化のため生成AI導入の動きがあります。ただし行政文書や個人情報扱う関係上、データをクラウド外に出せない制約があり、オンプレ実行できる国産LLMであるtsuzumi 2は**自治体ニーズと親和性が高い**と評価されています²⁰。NTTによると、初代tsuzumiは既に国内の自治体で試験導入され、AI活用のすそ野を広げてきました⁴⁰。tsuzumi 2ではさらに公共分野の知識を強化したことで、法律文や条例、申請書類など**行政特有の文書理解**にも対応力が増しています¹⁰。想定ユースケースとしては、**市民からの問い合わせ対応チャットボット**（例：「〇〇手続きには何が必要？」等に即答）、**議事録や報告書の自動要約**、職員向け**法令の要点解説**、さらには**自治体内部システムと連携した申請内容チェック**などが考えられます。実例として名は挙がっていませんが、NTTはすでに複数自治体とAI活用の共同実証を進めており、今後tsuzumi 2の本格導入も加速するとみられます。総務省や内閣府も生成AI活用のガイドライン策定や実証実験を推進しているため、**自治体×国産LLM**のモデルケースとしてtsuzumi 2が採用されることは、他の自治体にも波及効果をもたらすでしょう。

・**教育分野**： 教育現場での生成AI活用は、教材作成や個別学習支援など多岐にわたります。前述の**東京通信大学**のケースでは、学内限定LLM基盤の中核としてtsuzumi 2を導入し、**授業Q&Aの高度化（学生からの質問にAIが即回答）**、**教材・試験問題作成の支援**、**履修・キャリア相談のパーソナライズ**といったユースケースに活用する計画です¹²。同大学では既に試行運用でtsuzumi 2の有用性を確認しており、2025年度中にも本格展開する予定です¹⁷。他の教育機関でも、例えば中高生向けの**対話型学習アシスタント**や、大学における**研究文献の要約・翻訳支援**、教育委員会での**校務文書の作成補助**などにLLMが応用可能です。教育データは慎重な取り扱いが必要ですが、tsuzumi 2なら校内サーバで運用できプライバシー面の不安も低減できます。NTTは今後、教育分野へのAI導入支援にも注力するとしており、国産モデルの強みを生かして**学校現場で安心して使えるAIツール**を提供していく構想です⁴¹。現場教員からも「生成AIを使いたい情報漏洩や外部サービス利用規約が心配」という声があり、そうしたニーズに応えるソリューションとして期待されています。

・**製造・建設分野**： 製造業や建設業では、社内に蓄積された**技術文書・マニュアル**の活用や、**設計図面・CADデータの説明生成**など、専門知識と長文処理が必要なシーンが多く存在します。tsuzumi 2はこれまでに金融・医療・公共を重点学習しましたが、基本となる日本語理解・知識推論能力が高いため、製造業のドメイン知識を追加学習させることで優秀な**エンジニアリング支援AI**になり得ます。例えば、工場の故障対応マニュアルを読み込ませておけば、現場作業員の質問（「〇〇エラーの対処方法は？」等）にAIが即座に手順を答える**対話型マニュアル**が実現できます。また、設計部門では過去の設計レビュー資料や技術レポートを要約・検索することで、**ベテラン技術者の知見を新人に継承**したり、新製品の仕様策定時に類似事例をAIが教えてくれるといった使い方も考えられます。建設業でも膨大な施工基準書・図面からAIが適切な部分を抜粋して示すことで、**現場の調査・検討作業を効率化**できるでしょう。製造・建設分野は安全管理上データを社外に出せないケースも多いですが、tsuzumi 2のオンプレ運用なら**工場や現場オフィスのサーバ内で閉じたAIアシスタント**を構築できます。現時点で具体的企業名は出ていないものの、NTTグループは富士フイルムビジネスソリューション社との連携を発表しており、同社が持つ紙文書・図面のデジタル構造化技術（AIブランド「REiLL」）とtsuzumi 2の推論能力を組み合わせ、企業内の文書・画像データ分析を高度化するソリューション提供を目指すとしています⁴²⁴³。この取り組みは**製造業における設計図・仕様書のAI解析**にも通じるもので、今後、製造現場での国産LLM活用事例が創出される可能性があります。

以上のように、tsuzumi 2は**金融・公共・医療**で先行して知識チューニングされていますが、その他の**産業領域（製造・建設・流通など）**でも用途開拓が期待されています³⁰。NTT自身「多様な業界・分野での活用を通じ、さらなる導入実績を創出したい」と述べており⁴⁴、今後ユーザー企業やパートナーとの協業によりユースケースが拡大していく展望です。

6. 開発経緯や戦略的位置づけ（国産AIとしての意義、政府政策・経済安全保障との関係）

開発の背景と経緯: tsuzumi 2開発の発端には、ChatGPTに代表される海外製LLMが急速に普及する一方で、日本企業・自治体がそれらを利用する際に直面する課題がありました¹⁴。具体的には、巨大モデルゆえの**計算コスト・電力消費の増大**、クラウドサービス利用による**機密情報取り扱い上のリスク**、そして英語主体のモデルのため**日本語や国内制度・文化への最適化不足**といった点です⁴⁵。NTTはこれらを解決すべく、まず2023年に初代モデル「tsuzumi」を発表しました¹⁴。初代tsuzumiは「**軽量でありながら世界トップレベルの日本語性能を持つLLM**」と位置付けられ¹⁴、国内企業・自治体へのAI普及を促す役割を果たしてきました⁴⁰。その導入を通じて得られたユーザからのフィードバック（「**複雑なドキュメントを理解させたい**」「**専門知識対応力を上げてほしい**」等）を受け、NTTは次世代モデルtsuzumi 2の開発に着手⁴⁶。2024年3月には初代tsuzumiの商用提供を開始し、同年夏頃には決算説明会にてtsuzumi 2の構想（パラメータ増強や性能4倍向上など）を発表³、そして**2025年10月20日**に正式にtsuzumi 2のリリースに至りました³⁴。こうした経緯から、tsuzumi 2は**ユーザーニーズ主導で進化した実践志向のモデル**と言えます。NTT経営陣も「**お客様のニーズに応えられる性能に加え、1GPUで推論可能という従来モデルの特長も継承した**」とtsuzumi 2を総括しており¹⁸、企業DXを加速する現場目線のモデルとして位置づけています。

国内AI戦略における位置付け: tsuzumi 2は、日本発の基盤的AI技術として**戦略的な意義**を持ちます。NTTの島田社長は発表会で「**“国産AI開発競争”に負けれられない**」と発言し³⁴、海外ビッグテック主導の生成AI潮流に対抗しうる国産LLMを育てる決意を示しました。これは単なる企業の競争心に留まらず、日本全体の**経済安全保障と技術主権（AI主権）**の文脈でも重要視されています。生成AIの基盤を国外に依存し続ければ、将来的に利用制限やコスト高騰、データ主権侵害のリスクがあります。NTTは「国産技術による**AI主権確立を進める狙い**」がtsuzumiシリーズにあると明言しており¹⁰、自前のAI技術を持つことが国家レベルでも戦略的アセットになるとの考えです。日本政府も2023年以降、生成AI研究への支援策を打ち出し、NEDOや総務省を通じた大型言語モデル開発プロジェクトを推進しています。またデジタル庁も「日本における信頼性の高いAI基盤整備」を掲げ、官民連携での国産LLM活用実証を進めています。tsuzumi 2はそうした国策とも合致し、**国内AIインフラの一翼**を担う存在として注目されます⁴⁷⁴⁸。専門家の見方では、「tsuzumi 2リリースは日本のAIインフラ・ビジネスモデルにおける**“国内モデル化”の象徴的な一歩**」と評されており³⁵、この動きが政府の国産AI支援政策と連動すれば普及が国策レベルで加速する可能性も指摘されています²⁹。

NTTのAI事業戦略: NTTグループにとってtsuzumi 2は、自社のAI事業拡大戦略の中心的存在です。NTTはAIを含む新事業での収益拡大を中期経営目標に掲げており、島田社長は「**2027年度には生成AI関連で年間1000億円規模の売上を目指す**」と語っています⁴⁹。実際、NTTの生成AI関連受注額は2024年度約436億円から、2025年度は1500億円規模に達する見通しで、**2027年度には5000億円以上の受注を見込む急成長トレンド**にあります⁵⁰。この中でtsuzumiシリーズが占める割合はまだ小さいものの、NTTは**自社のAIポートフォリオ拡充**の一環としてtsuzumi 2をラインナップに追加し、国内外でのソリューション提案力強化を図るとしています⁵¹。NTT研究所が長年培った日本語処理技術を集大成したtsuzumi 2は、NTTグループ各社（NTTデータ、NTTコミュニケーションズ改めNTTドコモBusiness他）のサービスにも組み込まれ、**グループ横断でDXソリューション提供**に活かされる計画です⁵²。海外展開にも意欲的で、欧米やアジアのNTT顧客企業に対し、日本で磨いたLLMソリューションとして売り込む狙いもあります。実際NTTのAI関連受注では海外企業からの案件が国内を上回っており（2025年Q1で海外1300件、国内521件）⁵⁰、グローバル企業のニーズに応える中でtsuzumiの強み（データ秘匿性や企業内チューニング性）が評価されれば、海外で国産LLMが活躍する場面も出てくるでしょう。NTTにとってtsuzumi 2は、単なる技術開発成果にとどまらず**グループの成長戦略を担うプロダクト**であり、日本のAI産業競争力の象徴として戦略的に位置付けられています⁵³。

7. 将来展望・強化予定（機能拡張、業界別チューニング、エッジ対応など）

機能拡張の方向性: tsuzumi 2の今後の発展として、いくつかの技術的強化が予想されます。まず指摘されるのは**マルチモーダル対応**です。NTTは以前より画像や音声も扱えるAIの研究を進めており、初代tsuzumiに関するデモでは人物写真を解析し感情を読み取るような試みも紹介されました⁵⁴。専門家も「次のステップとしてマルチモーダル対応強化が予想される」と分析しており⁵⁵、将来的にテキストだけでなく**画像や表・音声を入力して解析・生成**できるtsuzumiのバージョンが登場する可能性があります。例えば、企業内のPDF図表入り資料を入力し要約生成したり、監視カメラ映像の状況を文章化するなど、モーダル融合による新たな活用シーンが期待できます。

業界別チューニングとエコシステム: tsuzumi 2は既に金融・医療・公共の3分野に知識面でチューニングされていますが、今後は**他業界向けの特化モデル開発**や、ユーザー企業自身がカスタムモデルを作りやすいエコシステム整備が進むと考えられます。NTTは「多様な業界・分野での活用を通じ導入実績を創出する」としており⁴⁴、例えば製造業向けに技術文書コーパスで追加学習した版、法務分野向けに判例や六法を学習した版など、**用途特化型のtsuzumi派生モデル**を提供する可能性があります。またNTTはクラウドサービス提供企業ではないため、エコシステム面では**パートナー企業との協業**がカギとなります。既に富士フイルムBIとの協業を発表していますが⁴²、今後Sler各社や産業別リーディング企業と組んで、tsuzumiを組み込んだソリューションやチューニングサービスを展開するでしょう。NTTデータやNTTドコモBusinessが顧客企業向けに**tsuzumi 2のファインチューニングサービス**を提供したり、NTT研究所が**業界データセットを共有**してモデル精度向上を図るなど、ユーザーコミュニティと連携した精緻化も考えられます。専門家も「周辺サービス・運用体制の整備」が鍵と指摘しており³⁶、NTTがどれだけ迅速に企業ユーザーが使いやすい環境（モデル配布形態、管理ツール、問い合わせサポート等）を整えられるかが普及のポイントとなりそうです。

エッジ対応・小型モデル: tsuzumi 2は1GPU動作という意味で既に省リソースですが、さらに小型のモデル展開も将来の検討材料でしょう。現状30B規模はPCやモバイルデバイスでの実行は難しいものの、例えばIoT機器やモバイルアプリに組み込む**超小型版tsuzumi**（数億～数十億パラメータ）を用意できれば、エッジAIとしての活用範囲が一気に広がります。NTT研究所はこれまで日本語特化の小型言語モデルも研究しており、将来的にはtsuzumiシリーズのエッジ展開も視野に入れている可能性があります。もっとも、小型化による性能低下と本来の強みである高精度とのトレードオフがあるため、まずはデータセンタやオンプレサーバ上での利用シナリオを優先しつつ、必要に応じて**蒸留モデル**や**専門特化モデル**を開発する流れになるでしょう。

他AI技術との連携: NTTはtsuzumi 2を核として、他のAI技術との組み合わせにも取り組む計画です。公式発表では、**サイバーセキュリティ分野への応用**や、複数AIが自律的に連携・議論する「**AIコンステレーション**」（星座状にAIを配置した協調システム）の開発に言及しています⁵²。サイバーセキュリティ応用では、tsuzumi 2の自然言語処理力を用いて膨大なログや脅威レポートを分析し、攻撃兆候の検知や対策立案を支援するといったユースケースが考えられます。一方のAIコンステレーションとは、**小規模なAIエージェント同士がネットワーク越しに連携し、多様な視点で議論しながら複雑な課題を解決する**というNTT提唱の概念です⁵⁶。tsuzumi 2のような汎用LLMが司令塔となり、専門特化した他のAIエージェント（例えば画像認識AIや数理最適化AIなど）と対話・協働することで、一つの大きなタスクを柔軟かつ高精度にこなすことを目指しています⁵⁶。この構想は、単一の巨大LLMに何でもやらせるのではなく、**複数AIの分散知能**で問題解決に当たるアプローチであり、NTTらしいネットワーク×AIの戦略と言えます。今後NTTは研究開発フォーラム等でこのAIコンステレーションの成果を発表予定で、2025年11月開催のNTT R&Dフォーラムではtsuzumi 2を活用した最新ソリューションの展示・体験が予告されています⁵²。

政府政策・標準化への寄与: 将来的に、tsuzumi 2は日本政府のAI施策や国際標準化にも関与していく可能性があります。現在、日本では生成AIの社会実装に向けたルール整備（例えば著作権データの扱いやAI倫理ガイドライン）が進行中です。NTTは国内有数の技術企業としてこれら議論に参加しており、自社モデルtsuzumiを通じた**安全・信頼なAI活用モデル**を提示できる立場です。例えば、行政文書をAIが生成する際の品質検証や、日本語モデル評価基準策定への貢献など、tsuzumiの知見が政策立案に活かされる余地があります。ま

た、経済産業省が進める**日本発のオープンAI基盤**構想において、将来的にtsuzumiを何らかの形でオープン提供する可能性もゼロではありません（現状tsuzumi 2はオープンソースではなくNTTのプロプライエタリですが、政府支援の下でコア技術共有が図られる可能性もあります）。いずれにせよ、「**日本語で高性能**」「**低運用コスト**」「**国内完結型**」というtsuzumi 2の特長は、日本社会に適したAIインフラの姿と言え、今後の産学官の動向次第で更なる強化・展開が見込まれます⁵⁷²⁹。

まとめ: NTTのtsuzumi 2は、技術性能・特徴において海外大型モデルに迫りつつ、日本ならではの要件（日本語・低コスト・セキュア運用）を満たす点で画期的な存在です。各方面での期待と関心は大きく、今後の導入事例の蓄積と機能拡張によって、国内AIエコシステムにおける地位を確立していくでしょう³⁵。もっとも、競合となる海外モデルや他の国内プレイヤーも存在する中、継続的な性能向上とユーザー視点の進化が求められます。NTT自身はtsuzumi 2を**企業DXを支える戦略製品**と位置づけ、巨額の市場目標を掲げていますが⁵⁸、その実現には技術力に加え**運用ノウハウやビジネスモデル**の確立も鍵となります³⁶。tsuzumi 2がこの先、日本の産業界に深く根付き、海外モデル依存からの脱却や新たな価値創造をもたらすことを期待しつつ、引き続きその進化を注視していく必要があるでしょう。

参考文献・情報源: 本調査レポートはNTT公式ニュースリリース¹⁴²¹⁵、NTT R&Dサイト記事⁵⁹¹⁶、ならびに各種専門メディアの記事⁸¹⁰²⁷を基に作成しました。各種性能比較値や評価コメントはこれら信頼性の高い情報源から引用・要約しています。

¹ ⁴ ¹⁶ ¹⁹ ²² ²³ ⁵⁹ NTT版大規模言語モデル「tsuzumi 2」 | NTT R&D Website

https://www.rd.ntt/research/LLM_tsuzumi.html

² ⁶ ⁷ ¹² ¹⁴ ¹⁵ ¹⁷ ¹⁸ ²⁶ ³⁹ ⁴¹ ⁴² ⁴³ ⁴⁴ ⁴⁶ ⁵¹ ⁵² 更なる進化を遂げたNTT版LLM tsuzumi 2の提供開始～日本の企業DXを支える高性能・高セキュア・低コストな純国産LLM～ | ニュースリリース | NTT

<https://group.ntt.jp/newsrelease/2025/10/20/251020a.html>

³ ⁹ ²⁹ ³⁰ NTT、純国産AI「tsuzumi 2」を10月公開 GPT-4o並の日本語力で業務支援強化へ | Plus Web3 media

https://plus-web3.com/media/latestnews_1000_4655/

⁵ ¹⁰ ⁴⁰ NTT、純国産LLM「tsuzumi2」提供開始、日本語性能を世界水準に | 財経新聞

<http://www.zaikei.co.jp/article/20251021/829740.html>

⁸ ¹¹ ³¹ ³² ³³ NTTが大規模言語モデル「tsuzumi 2」を提供開始 世界トップクラスの日本語性能を実現 - ケータイ Watch

<https://k-tai.watch.impress.co.jp/docs/news/2056478.html>

¹³ ²⁴ ²⁵ ²⁷ ²⁸ ³⁸ ⁵⁰ ⁵⁶ NTTが発表した国産LLM「tsuzumi 2」の実力。日本語性能や専門知識を強化しつつ軽量のAIモデル | Business Insider Japan

<https://www.businessinsider.jp/article/2510-ntt-tsuzumi-2-unveiled/>

²⁰ ²¹ ³⁵ ³⁶ ³⁷ ⁴⁵ ⁴⁷ ⁴⁸ ⁵³ ⁵⁵ ⁵⁷ ⁵⁸ 「NTTが切り拓く“国産LLM”時代——つづみ2で日本語AIインフラが変わる」 | 【楽楽好奇心note】気になること全部調査。雑談力×ビジネスのヒント集

<https://note.com/bostonmio/n/n4628f152497a>

³⁴ NTT、独自のAIモデル「tsuzumi 2」発表 “国産AI開発競争”に「負けられない」と島田社長 | Happyデジタル

<https://note.com/happydigital/n/nbcfe4c14262f>

⁴⁹ NTT版LLM「tsuzumi」に見る国産LLMの意義と戦略 - Impress Watch

<https://www.watch.impress.co.jp/docs/series/nishida/1544492.html>

⁵⁴ NTTの生成AI大規模言語モデル(LLM)「tsuzumi」驚きのデモ徹底 ...

<https://robotstart.info/2023/11/14/ntt-tsuzumi-demo.html>