



(報道発表資料)

2025 年 10 月 20 日

NTT 株式会社

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～

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問題の解決に向け、軽量でありながら高性能な日本語処理性能を持つ LLM「tsuzumi 2」の提供を 2025 年 10 月 20 日開始いたします。

NTT は、パブリック環境やクローズ環境、顧客の業務形態やデータ要件に最適な AI ソリューションの提案・導入を推進しています。多様な AI 提案力を総合的に向上させ、AI ポートフォリオの拡充を目的に、このたび「tsuzumi 2」を新たにラインナップに追加しました。

※本ニュースリリース文章は tsuzumi 2 を使って作成しております。

1. 提供の背景

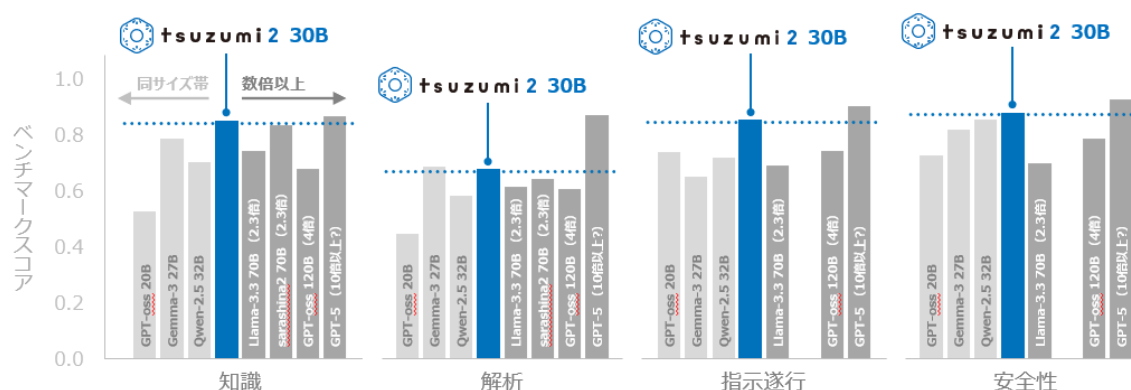
近年、ChatGPT をはじめとする LLM への注目が高まる一方で、膨大な学習データと計算リソースを必要とする従来の LLM は、電力消費や運用コストの増大、機密情報の取り扱いにおけるセキュリティリスクといった課題を抱えています。NTT は、これらの課題に対応すべく、軽量でありながら世界トップレベルの日本語処理性能を持つ「tsuzumi^{※2}」を 2023 年に発表し、国内企業・自治体等での AI の普及を推進してきました。

実際のビジネス現場での AI 活用において、特に企業・自治体が保有する複雑なドキュメントへの理解や専門的な知識への対応力強化等のご要望を多数いただきました。このようなニーズを研究開発へフィードバックし、NTT は「tsuzumi」の次世代モデル「tsuzumi 2」を開発しました。お客様からのニーズにお応えできる性能に加え、1GPU で推論可能という従来モデルの特長も継承し、環境負荷とコストを抑え、企業・自治体等の AI 活用を加速できるよう進めてまいります。

2. 進化した tsuzumi 2 の特徴

① 日本語性能のさらなる向上

日本語性能においては、同サイズ帯のモデルと比較して世界トップクラスの性能を実現しています。ビジネス領域で主に重視される知識、解析、指示遂行、安全性の基本性能では、数倍以上大きなフラッグシップモデルに匹敵するレベルを達成し、コストパフォーマンスに優れています。



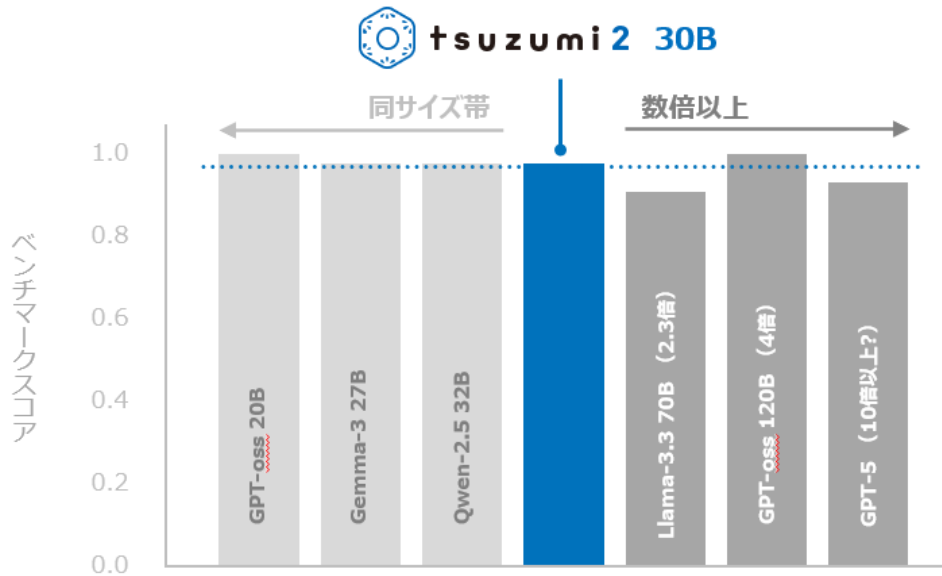
評価条件

- ・ 知識: llm-jp-eval (JAQKET (日本語QAタスク), NIILC (質問応答), JCommonsenseQA (常識推論能力)), pfgen-bench (PFN日本知識ベンチ) の平均値
- ・ 解析: llm-jp-eval (Wikipedia読解, Wikipedia係り受け解析) の平均値
- ・ 指示遂行: M-IFEval_Ja
- ・ 安全性: AnswerCarefully (x0.2)
- ・ 知識と解析においては、事前学習モデルを使用。ただし、gpt-oss 20B/120B、GPT-5は事前学習モデルが公開されていないため事後学習モデル (Reasoningモデル) を使用
- ・ 指示遂行と安全性において、全て事後学習モデルを使用。sarashina2 70Bは事後学習モデルが公開されていないため対象外

(図 1 日本語性能の評価)

② 特化型モデル開発効率の向上

RAG^{※3} と Fine Tuning^{※4} により、企業や業界に特化したモデルの開発効率が向上しています。tsuzumi 2 は、お客様からの要望が多かった金融・医療・公共分野の知識を強化しています。これにより、これらの分野で高い性能を発揮し、特化チューニングによる精度向上にも優れています。また、汎用的なタスク特化チューニングにより、多様な業務での活用も可能です。例えば、NTT 社内において「財務システムに関する問い合わせ対応業務」に活用した評価では、他社の先進モデルと同等以上の世界トップクラスの性能を確認しています。

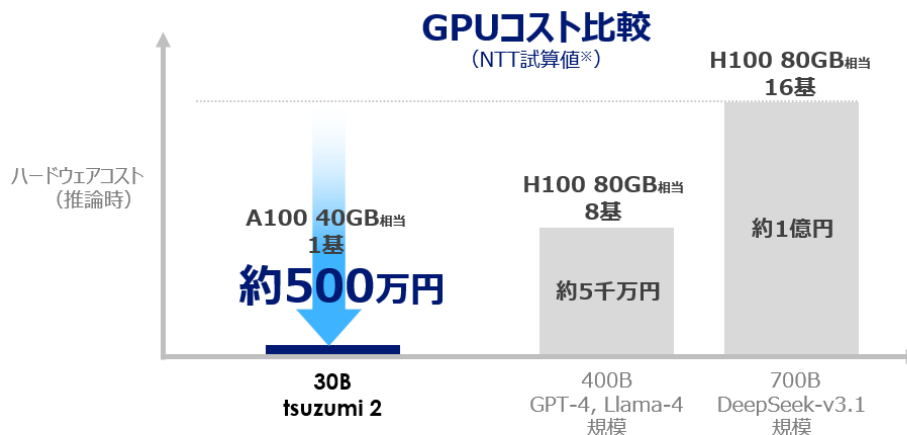


評価条件 NTT社内業務（財務システムに関する社内ヘルプデスク）における
トライアル案件において、RAGによる問い合わせ回答正答率を独自に評価。

（図2 RAG ユースケースにおける LLM の性能評価）

③ 低コスト・高セキュアの維持、国産 AI

tsuzumi は NTT がフルスクラッチでゼロから開発した純国産モデルです。1GPU で動作可能な軽量モデルでコストパフォーマンスに優れています。低コストでオンプレミスやプライベートクラウドでの運用が可能であり、機密性の高い情報も安全に取り扱える点が特徴です。



※試算条件
・ 量子化: 8ビット
・ 必要GPUメモリサイズ: パラメータ数 × 量子化サイズ/8bit (30Bは30GB、400Bは400GB、700Bは700GB)
・ GPT-4, Llama-4は、MoE、H100 x 8 (1ノード)動作、DeepSeek-v3.1は、MoE、H100 x 16 (2ノード)動作を前提
・ ハードウェアコストは、上位GPU H100 80GB: 1,000万円/台、下位GPU A100 40GB: 500万円/台として換算、その他の運用などの費用は含まず



3. tsuzumi 2 の導入に向けた取り組みについて

これまでの NTT グループ各社での AI 提案・導入活動を通じて、tsuzumi 2 に対する引き合いや事前実証が多数進められてきました。その結果、既に 2 件の具体的な事例が生まれています。今後も多様な業界・分野での活用を通じて、さらに多くの導入実績を創出できるよう取り組んでまいります。

① 顧客導入実績



東京通信大学では、学生・教職員のデータを学内に留めるという要件のもと、クラウド依存のない国産 LLM を核に学内 LLM 基盤を整備してきました。この度、tsuzumi 2 により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においても実用水準に達しているとの判断から、tsuzumi 2 の導入を決定しました。具体的には、授業 Q&A の高度化、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ユースケースで活用されることになり、今後本格的な学内での利用開始により、教育・運営の両輪で AI 活用を加速します。

② tsuzumi 2 を利用した新たな連携による価値創出



NTT ドコモビジネス株式会社（旧 NTT コミュニケーションズ株式会社、以下 NTT ドコモビジネス）は、富士フイルムビジネスイノベーション株式会社（以下富士フイルムビジネスイノベーション）と連携し、tsuzumi 2 を活用した新たな生成 AI ソリューションの提供に向けた検討を開始しました。富士フイルムビジネスイノベーションは、AI 技術ブランド「REiLI」※5のもと、企業内に散在する契約書や提案書など、形式の異なる文字・図・画像を含む非構造化データを構造化し、価値ある情報として活用可能にする技術を展開しています。機密情報を安全に取り扱いながら高度な推論を行える生成 AI「tsuzumi 2」と富士フイルムビジネスイノベーションが持つデータの構造化技術を組み合わせることで、企業内の文書・画像データの分析や活用をさらに高度化する生成 AI ソリューションの提供を



めざします。NTT ドコモビジネスおよび富士フイルムビジネスイノベーションは本取り組みを通じて、AI 提案力を強化し、企業の業務変革と生産性向上に貢献していきます。

4. 今後の展開

NTT は、tsuzumi 2 の提供開始を機に、NTT グループ各社からのソリューション提供、サービス実装を順次推進してまいります。また、サイバーセキュリティ分野への応用、自律的に連携し議論する AI コンステレーション等の開発も進めます。これらにより NTT は新たな価値創造、お客さま体験の高度化に向けた取り組みをより一層加速してまいります。

2025 年 11 月 19 日から開催される NTT R&D フォーラム 2025 (IOWN Quantum Leap) において、tsuzumi 2 を活用した最新の生成 AI ソリューションを展示し、ご来場の皆様に直接体験いただく機会を提供します。

【参考】

NTT 版 LLM 「tsuzumi」 https://www.rd.ntt/research/LLM_tsuzumi.html

NTT R&D フォーラム 2025 (IOWN Quantum Leap) <https://www.rd.ntt/forum/2025/>

【用語解説】

※¹LLM: 大規模言語モデル (**L**arge **L**anguage **M**odel): 大量のテキストデータを使って学習された言語モデルで、言語の理解や文章の生成に優れた能力をもつもの

※²tsuzumi: tsuzumi®は NTT 株式会社の登録商標です

※³ RAG (Retrieval-Augmented Generation): 外部情報を検索し、その情報を基に生成を行う技術

※⁴ Fine Tuning: 特定の業界知識を事前に学習させ、モデルを最適化するための技術

※⁵ REiLI: 富士フイルムグループの AI 技術ブランド。富士フイルムビジネスイノベーションにおいては、非構造化データの構造化を核に 5 つの AI エージェントを展開しています。詳細は

<https://www.fujifilm.com/fb/company/technical/ai>



<関連資料>

- ・ AI For Quality Growth の実現へ tsuzumi 2 のリリース
- ・ tsuzumi 2 進化のポイント
- ・ NTT の AI 戦略について

■本件に関する報道機関からのお問い合わせ先

NTT 株式会社

広報部門

ntt-pr@ntt.com

ニュースリリースに記載している情報は、発表日時点のものです。

現時点では、発表日時点での情報と異なる場合がありますので、あらかじめご了承ください。
くとともに、ご注意をお願いいたします。