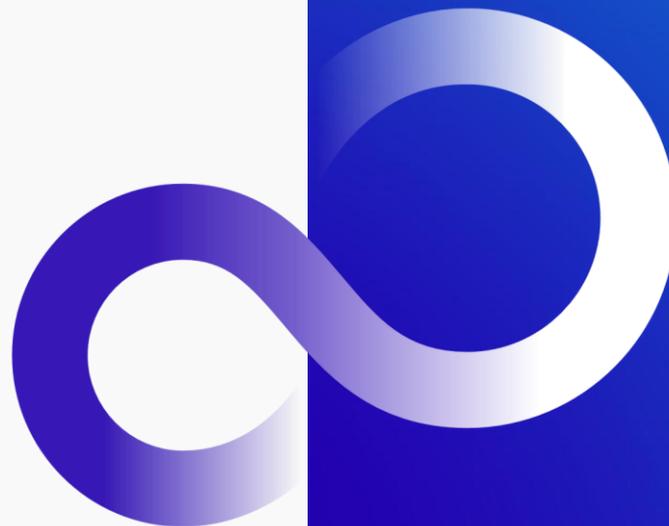


Fujitsu クラウドサービス Generative AI Platform ご紹介資料

2025年10月

富士通株式会社



1. サービス概要
2. サービス提供機能
3. トライアル利用について

Appendix. Takaneのご紹介

Appendix. RAGとファインチューニングの位置づけ

1. サービス概要

生成AIは、**文章・画像・音声・動画など多様なコンテンツを自動生成でき、創造性を拡張**する技術です
ビジネスでは効率化やパーソナライズの向上、研究開発ではシミュレーションや自動設計など
多岐に渡る活用が**試行・実践**されています
今後は、**専門知識との統合**や**人間との協調**によって、さらなる**価値創出**が期待される領域です

生成AI・深層学習・機械学習の違い

生成AI

膨大なデータを分析し、テキストや画像など、新しいデータやコンテンツを自ら生成
例：市場調査、プログラム作成、画像出力

深層学習(DL)

多層のニューラルネットワークを用いて特徴抽出や認知
人間の脳活動を模倣
例：画像認識、音声認識、言語翻訳

機械学習(ML)

データからパターンを学習し、明示的なプログラムなしにタスクを改善
例：需要予測、スパムフィルター、購買レコメンデーション

人工知能

人間の知的活動を模倣する技術の総称
AGI (Artificial General Intelligence)
ASI (Artificial Superintelligence)

生成AIをエンタープライズ向けに提供するには FUJITSU

Outputの 信頼性



- 誤情報生成リスク
 - ハルシネーション
- 生成内容の一貫性制御
 - 企業ブランド/ポリシーに準拠した回答

セキュリティ データプライバシー



- 機密情報の漏洩リスク
- データ管理と法規制
 - GDPR、CCPA等

投資対効果[ROI] の難解性



- 膨大な導入/運用コスト
- 具体的なユースケース選定
- スケーラビリティと
インフラ管理の煩雑性

AIリテラシーと 実践スキル

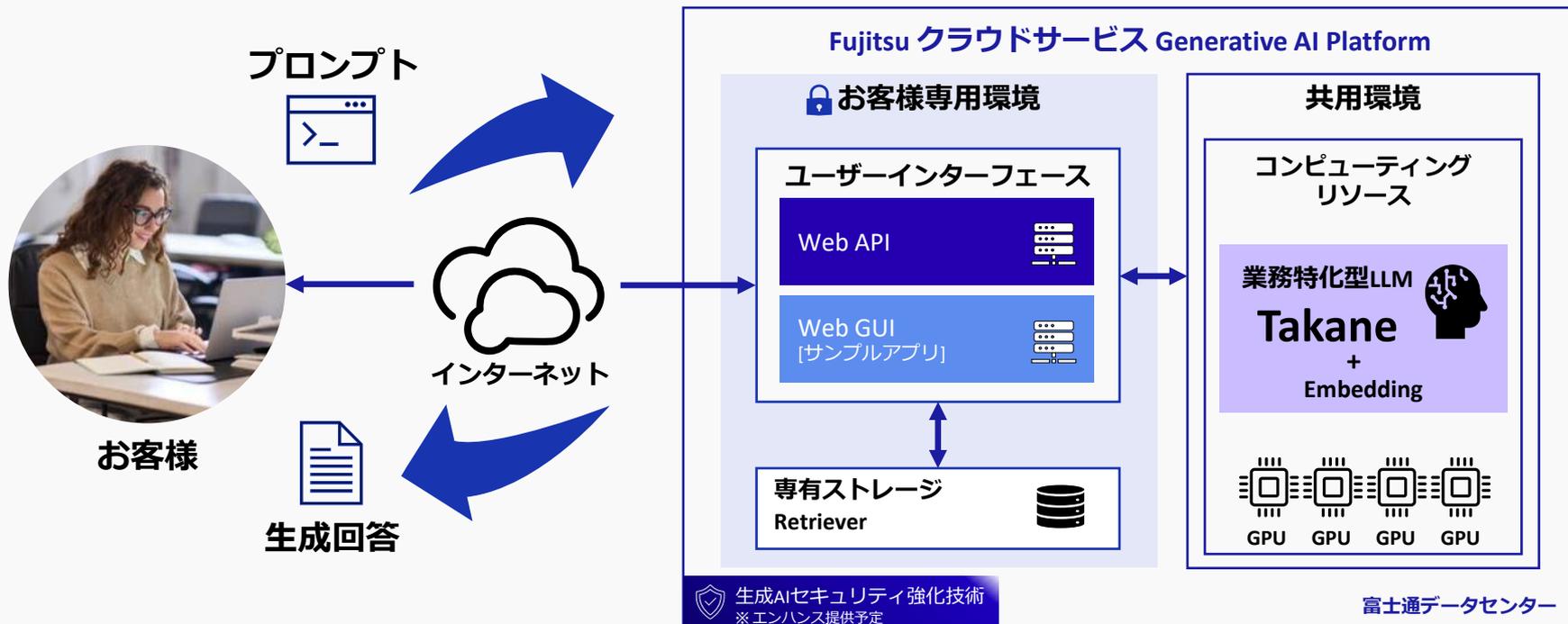


- AIエンジニアの確保
- 利用部門への適用/教育
- 自社開発が必要

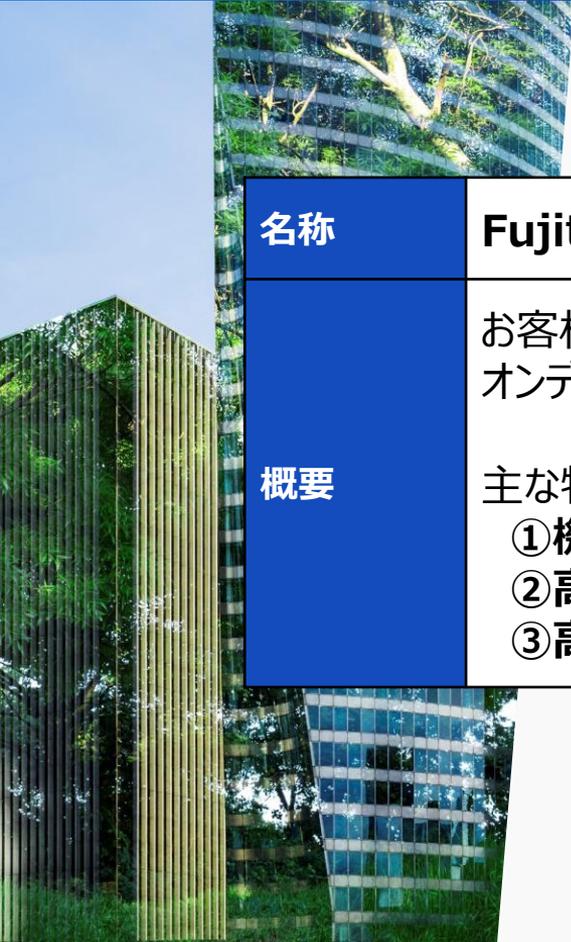
信頼性、セキュリティ、データプライバシー、投資対効果、AIリテラシー、誤情報リスクなど

数多くの課題が存在

オンデマンドで利用できるセキュアな クラウド型生成AIサービス



※日本国内での実績を踏まえ、展開国に則したモデルを検討の上グローバル展開予定



名称	Fujitsu クラウドサービス Generative AI Platform
概要	<p>お客様データをプライベート領域で管理し、生成AIで機密情報を扱えるオンデマンド利用可能なクラウド型生成AIサービス</p> <p>主な特長は以下3点</p> <ul style="list-style-type: none">①機密性の高い情報扱う業務での生成AI活用が可能②高価なGPUリソースを共有することでコスト抑制③高度な生成AIセキュリティ強化技術(※)等により安全性が高い

※ エンハンス提供予定

お客様の課題

- 社内規定などにより、パブリッククラウド上の生成AIでは機密性の高い情報は活用できない
- 自社用に独立した生成AIシステムをオンプレで構築・維持運用するにはコストがかかる

オンプレ個別構築
生成AIシステム

独立性

Fujitsu クラウドサービス
Generative AI Platform

パブリック
生成AIサービス

利用しやすさ

Fujitsu クラウドサービス Generative AI Platform はオンプレとパブリックの利点を融合しデータの独立性と利用しやすさのバランスが取れたクラウド型生成AIサービスを提供

サービスの特長（詳細）

Fujitsu クラウドサービス Generative AI Platform



専有ノード Web API

Web GUI サンプルアプリ



専有領域
ストレージ



RAG

共有領域

コンピューティング
リソース

Takane



LLM
Embedding



生成AIセキュリティ強化技術(※)

富士通DC

1 お客様データはプライベート領域で管理 機密性の高い情報扱う業務での活用が可能

RAG用データをお客様ごとのプライベート領域に保管可能なため、機密情報を扱う業務においても生成AIに自社データを活用可能。

2 高価なGPUリソースを共有することで コストを抑制

推論用のGPUサーバは共有領域で提供。
個別構築と比較し、初期導入コストを抑制。

3 高度な生成AIセキュリティ強化技術(※)により 高い安全性を実現

例えば利用者がLLM経由で意図せずに参照権限の無い情報を引き出してしまうことを防ぐことができます。
脆弱性対応や第三者監査を通して安全性を担保したJDCCティア4相当の富士通DCにより、安全・安心かつオンデマンドでLLMを利用できるクラウドサービスです。

※ エンハンス提供予定

© 2025 Fujitsu Limited

生成AI関連
ソリューションサービス

生成AIをエンタープライズするためのSI/上流サービス

e.g. アセスメント/コンサル、役務サービス等

生成AI関連
アプリケーション

生成AI基盤を活用したアプリケーション層

e.g. 開発用プラットフォームを含む：Langchain/Dify等

生成AI基盤モデル

基盤モデル自体

e.g. Takane/Cohere等

インフラ基盤

基盤モデルの稼働を支えるインフラ

e.g. 高度な演算が可能なGPU、安全なデータ保管ストレージ等

提供範囲

生成AI活用に必要な4要素の内「インフラ基盤、生成AI基盤モデル」をサービス提供

生成AIセキュリティ強化 LLM脆弱性スキャナー&ガードレール

※エンハンス提供予定

LLMの脆弱性を網羅的な評価で判定し、最適な防御技術を自動的に適用します

LLM脆弱性スキャナー



LLMガードレール



業界トップクラスの脆弱性対応

- 富士通独自の手法を含む**7,700種超の脆弱性**に対応
- 従来検出困難だった**対話型の攻撃**も高精度に分析

専門知識不要

- 説明機能により、**非専門家でも**影響を容易に理解可能
- 学習コスト**を抑えたセキュリティ対策強化が可能

防御ルールを自動生成

- レポートに基づき最適な防御ルールを**自動生成**
- LLM本体に手を加えず、**外付けでの防御**を実現

サンプルとして提供する「Web GUI」

プログラミングの知識が無くても、
直感的なUIを使って独自のチャットアプリケーションを作成することが可能です

チャット画面



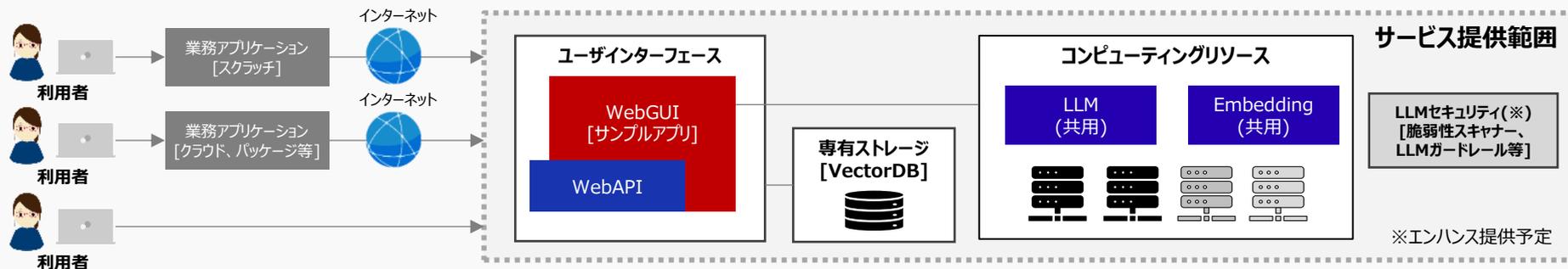
RAG登録



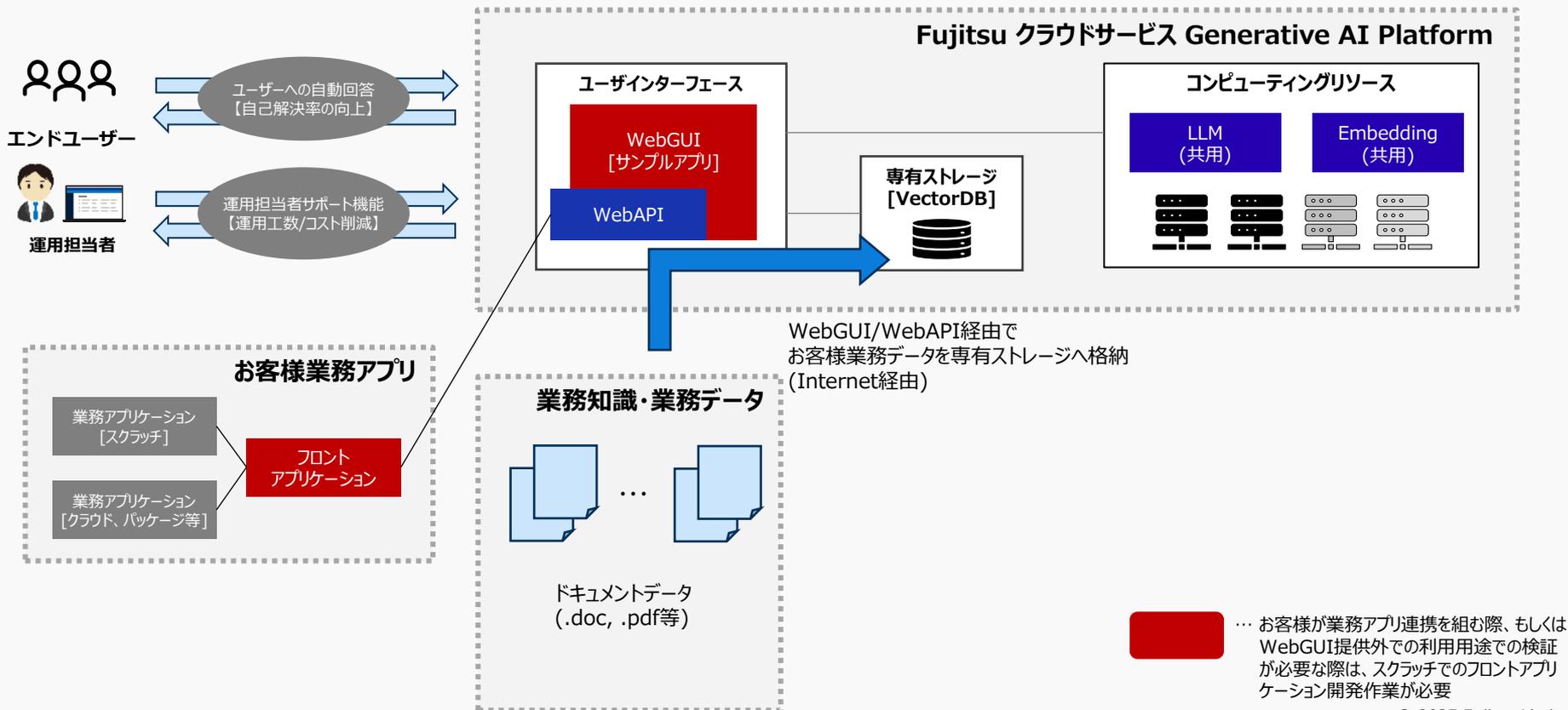
2. サービス提供機能

● 提供機能概要

- 生成AI基盤モデル(Takane等)をオンデマンドかつセキュアに利用できる環境を提供
- 提供物
 - 生成AI基盤モデル(LLM)提供：共用LLM
 - データ領域：追加学習(RAG構成)に活用可能なデータ格納領域を専有環境としてご提供
 - ユーザインターフェース：専有環境として下記をご提供
 - WebGUI[Chat機能を有するサンプルアプリケーション] ※サンプルアプリのため保証対象外
 - WebAPI[エンドポイント/APIキーをお客様毎に発行]
 - アクセス制限/認証：
 - IPアドレス制限：ご指定のGlobal IPアドレスからのみアクセス可に制限



活用イメージ：サポート業務への適用



3. トライアル利用について

下記表のとおり、トライアル環境をご用意しております
ご希望のお客様は弊社担当Salesへご相談ください

	内容
利用料	無償 ※基本構成のみご利用いただけます（ストレージ拡張オプションは提供対象外）
利用可能期間	2週間
利用可能時間帯	メンテナンス時間、土日・祝日および当社休業日を除く平日9:00から17:00まで
メンテナンス	計画メンテナンス（事前通知あり）／緊急メンテナンス（事前通知なし）
SLA	トライアルのため保証なし
サポート	<ul style="list-style-type: none">・ E-mailで問い合わせ窓口へ質問可能（ベストエフォートの応答）・ 対応時間は、土日・祝日および当社休業日を除く平日9:00から17:00まで・ 受付可能な質問は、本サービスの仕様・利用方法、本サービスが正常動作しない場合の回避措置
責任分界点	<ul style="list-style-type: none">・ 当社責任範囲：本サービスを構成するためのデータセンター設備、ハードウェア、仮想化基盤、エントリーポイント、APIおよびサービス運用の管理・ 契約者責任範囲：本サービス提供環境内に格納したデータ、当該データに対する以下の項目についての作業<ul style="list-style-type: none">・ 格納したデータの新規登録、変更、削除などのデータに対する操作
トライアルデータ削除	<ul style="list-style-type: none">・ トライアル終了日までにお客様にて削除・ 終了日を超えての保管は不可であり、富士通裁量で削除可能とする

本ページ内容は規約等の抜粋情報であり、お申し込みに際してはトライアル契約関連書類（申込書・利用規約・サービス仕様書）を必ずご確認ください。
なお、本ページ内容と契約関連書類の間で不一致がある場合、トライアル契約関連書類を有効とします。

【事前準備】 以下についてご確認ください：

- トライアル目的が本サービス提供環境に合致しているか
 - 次頁「先行トライアル申込前確認シート」を活用ください
 - 確認の結果によりトライアル利用のご希望に沿えない場合がございます
- トライアル利用の開始日と終了日の確認
 - 提供可能な環境数に限りがあるため、提供状況によっては利用開始日の調整をお願いする場合がございます

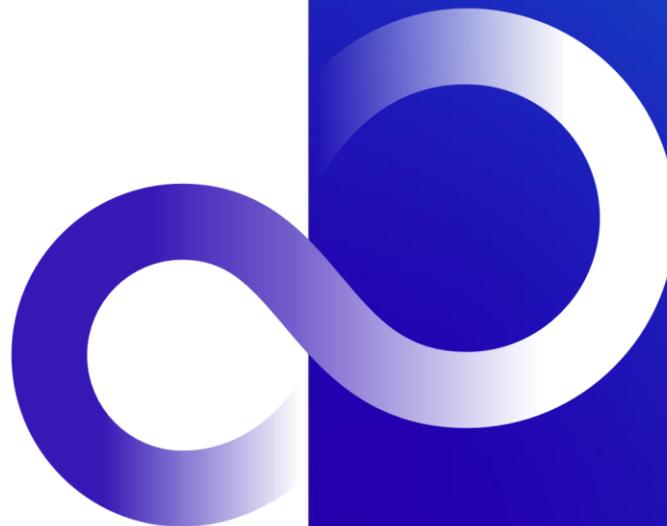
【お客様手続き】

- ① 弊社担当Salesより**トライアル申込書および環境設定ヒアリングシート**を受領
- ② 2点に必要な情報を記入、トライアル申込書については押印の上、**トライアル申込書および環境設定ヒアリングシートの記入済みExcelとトライアル申込書の押印済みPDF**の3点を弊社担当Salesへ提出
- ③ 弊社担当Salesより**請書**を受領
- ④ トライアルサービス提供開始1営業日前を目途に本サービスへのアクセス情報および利用者マニュアル、お問合せフォーマット（メール）を弊社担当Salesより受領
- ⑤ トライアルサービス提供開始日より**利用開始**

トライアル利用 申込前確認シート

- トライアルのお申し込みにあたり、お客様のご期待するところと提供物が合致するか以下の観点で事前確認をお願いいたします

分類	確認観点	結果
トライアルサービス仕様	本サービスのデータ格納用プライベート領域は、論理的に区切られたお客様用領域です。筐体ごと物理専有へのご要望には応えられません。	OK / NG
	本サービスへの接続方法はインターネット経由となります。	OK / NG
	本サービスのインターフェイスはWEB APIとなります。APIを呼び出すアプリなどはお客様にてご用意いただく必要があります。（簡単なサンプルGUIはご用意しています）	OK / NG
	Takaneは特に日本語処理に適したLLMです。評価ケースに適しているかご確認ください。	OK / NG
トライアルご利用条件	本サービスの利用料は無償とします。導入コンサルやチューニングなど役務サービスはスコープ外となり含まれておりません。	OK / NG
	トライアル開始時期については調整のうえ決定させていただきます。	OK / NG
	トライアル目的のみご利用可能です。商用利用など他の利用目的は不可となります。	OK / NG
	トライアル終了後、結果のフィードバック等を当社の求めに応じて提供いたします。	OK / NG
	セキュリティ担保のため、トライアル環境へのログイン認証にMicrosoft Entra ID（旧Azure AD）を使用します。お客様の所属される企業・団体等のアカウント管理がMicrosoft Entra IDであること・使用可否をご確認ください。※ご利用に際し、新規に取得いただいたMicrosoft Entra IDでもご利用可能です。準備手順のご案内が可能ですので必要に応じて弊社担当Salesへご相談ください。	OK / NG



Appendix1. Takaneのご紹介

企業ニーズを満たす大規模言語モデル(LLM)の開発
およびサービス提供に向けた戦略的パートナーシップを締結



セキュリティとデータプライバシーに特化した エンタープライズAIのリーディングカンパニー[スタートアップ]



企業向けLLMの提供
にフォーカスし急成長（2019年創業）



企業内データを高精度に
利用可能とする**保有技術**



CEO **Aidan Gomez** 氏
Transformer開発者のひとり

企業価値

55億USD

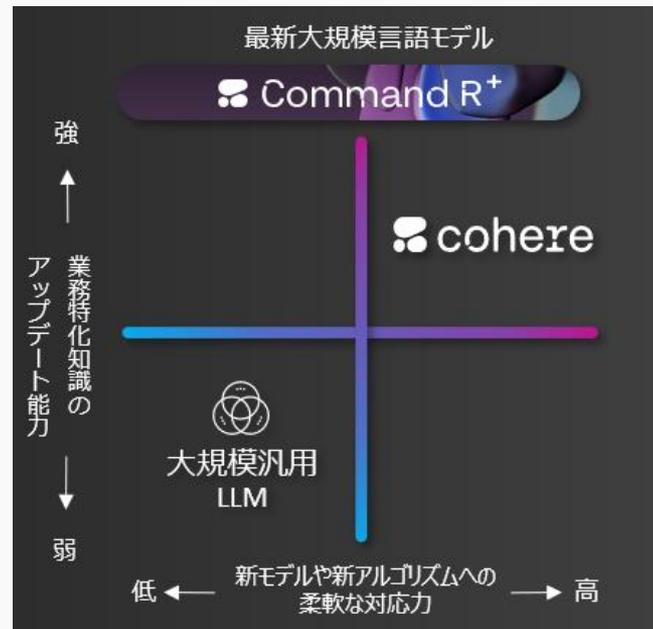
拠点



**世界トップクラスの
RAG技術**

エンジニア
111名

研究者
67名



エンタープライズ向け大規模言語モデル「Takane」 FUJITSU

2024年9月より富士通のAIサービス「Fujitsu Kozuchi」より先行スタート

cohere

共同開発
業務特化型LLM

Takane

高嶺

FUJITSU

- 企業の業務向けに最適化されたLLM※1
- 企業内データを高精度に利用可能とする世界トップクラスのRAG※2技術
- 業界をリードするEmbed※3、Rerank※4モデル
- 主要ビジネスタスクへの高カスタマイズ性

最新大規模言語モデル
「Command シリーズ※5」

- 日本語特化追加学習／ファインチューニング技術
- ナレッジグラフ拡張RAG
- 生成AI混合技術
- プライベート環境での提供
- 基幹業務システム構築で積み重ねてきたナレッジ

※1 Large Language Model：大規模言語モデル

※2 Retrieval-Augmented Generation：検索拡張生成

※3 テキストやその他のデータを数値ベクトルに変換する機械学習モデル

※4 検索システムの精度を向上させるためのモデル

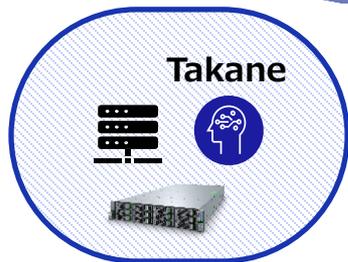
※5 Cohere Inc.が開発した大規模言語モデルシリーズ

✓ エンタープライズユースに特化したハイセキュリティなLLM

秘匿性の高いデータを LLMで扱いたい

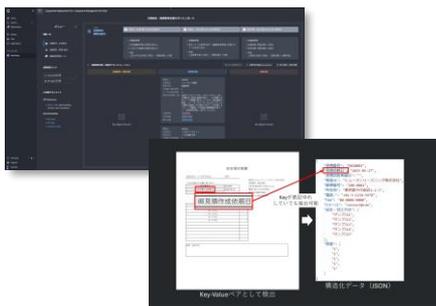
オンプレミス環境にTakaneをデプロイ可能であり、クラウド環境を利用できない秘匿性の高い業務領域のデータを安全にLLMで扱うことが可能

プライベート環境



特定の業務で的確に動作する LLMをオーダーメイドしたい

富士通研究所の長年にわたる研究ナレッジをもとに、各業種のお客様向けに様々なユースケースをリリース。専門知識や業種特有の挙動が要求される場合にLLMのファインチューニングが可能



導入からアフターサポートまで ワンストップで提供してほしい

富士通研究所をはじめとして、弊社内の様々なAI関連組織と有機的に連携。生成AI活用上流工程からお客様と伴走して導入をサポート

お客様



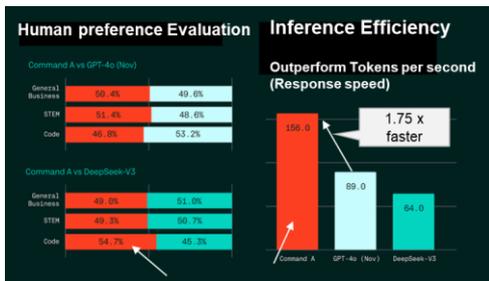
FUJITSU

なぜTakane? - 仕様/性能

✓ 少ないインフラリソースで高い日本語性能を発揮

日本語に強く、レスポンスの良いLLMを使いたい

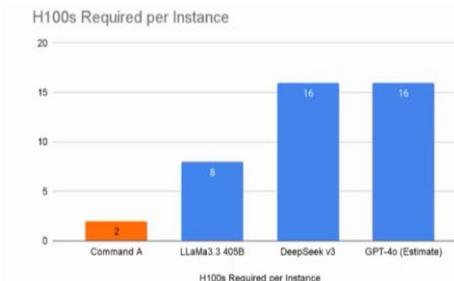
富士通研究所の知見を活かし、**日本語特化のファインチューニングを実施**。JGLUEにおいて**世界最高性能**を達成



Source: <https://cohere.com/blog/command-a>

インフラへの投資コストを最小限に抑えたい

NVIDIA H100使用時、GPT-4o比にて、たった**1/8のリソースで動作可能**。最小のインフラ投資で最大のパフォーマンスを発揮



Source: <https://cohere.com/blog/command-a>

必要なパラメータサイズに合った製品を選定したい

1~30B、30~100B、100B~とバラエティに富んだサービス群を用意。性能の過不足やファインチューニング時のコストを抑制

※2025.9 当社調べ

Parameter Size	Cohere/Takane	NTT tsuzumi	NEC cotomi	IBM granite	Ricoh	H2O
1-30B	Command R7B, Takane 2.0-7B	tsuzumi 8.6B, tsuzumi 7B	tsuzumi 13B	granite 3.3 2B v3, granite 3.3 8B v3		Danube10.5B, MessageGPT2.3.5B, Danube21.8B, Danube34B
30-100B	Command R, Takane 2.0-32B		cotomi Fast v2 #1, cotomi Pro v2 #2		70B Llama 3.3 70B	
100B+	Command R+, Takane 1.1				gpt-oss-120B	
	Command A, Command A MM					

Differentiation

※1 As of November 2024, it is being evaluated as the medium-sized version of Tsuzumi.
 ※2 The parameter size has not been disclosed.
 ※3 In addition to the language model, Granite also provides vision and code models.

Appendix2. RAGとファインチューニングの位置づけ

生成AIを活用するために、その目的や用途と、必要な技術を使い分けが必要になります

