

環境構築済みだから、導入後すぐに使える!

ローカルLLM / RAG構築サービス



ハードウェア

+ RAG・LLM環境

+ 運用サポート

研究開発向けに最適なHPC/ワークステーションを、
必要なシステムがインストール済みの状態でご用意できます!
さらに安心の運用サポートまで! 全て揃ったスタートアップセット!



インストール・設定ソリューションサービス 対象ツール

ハードウェアに、以下のような環境をインストールした状態でお渡しすることができます。

 <p>Difyは今すぐ実運用可能なAI Agentを簡単に構築できる、AIアプリケーション開発ツールです。</p>	 <p>Meta社が開発したオープンウェイトの大規模言語モデル (LLM) です。幅広い自然言語処理タスクに対応。</p>	 <p>ローカル環境で大規模言語モデル(LLM)を簡単に実行・管理できるAIモデル実行ツールです。</p>
 <p>Googleが開発した軽量で高性能なオープンな小規模言語モデル (SLM) ファミリーで研究・開発者向けに公開。</p>	 <p>様々なAIモデルの操作と統合を合理化するオープンソースプラットフォームです。</p>	 <p>OpenAIが公開した約200億パラメーターを持つオープンウェイトのLLMで、高性能と軽量を両立。</p>

- OS : Ubuntu 24.04 LTS / Windows11
- GPU : Nvidia RTX / GeForce
- LLM / RAG開発ツール : DifyAI

- AI実行ツール : LM Studio / Ollama / Xinference
- 複数AIモデル : ・ Llama / Gemma / gpt-oss-20B
- 各種API

インストール代行 ローカルLLM / RAG構築サービス

環境に応じてお見積りいたします

安心して始められる保守サービス・サポート

当社では、お客様が安心してLLM/RAGシステムを導入いただけるよう、充実したサポート体制を整えています。

■ 導入時のサポートもワンストップでご提供

システム環境構築 : ハードウェア・ソフトウェア専門エンジニアが予めセットアップして出荷
各ツール動作確認 : Dify・各種生成AIモデルの基本機能の動作確認実施済み



セキュアなローカル環境

個別情報や機密情報を含む組織内データの流出リスクがない安全なオンプレでの検索 / 生成AI環境

安心の導入サポート体制

導入後1カ月間の利用方法等のサポート
別途有償サポート契約を提供

導入後すぐにご利用可能

必要なツールのインストールおよび構築が完了した状態で納品いたします

有償サポート

講習会やエラー対応、バージョンアップなど

安心のローカルLLM / RAG 保守サービス

環境に応じてお見積りいたします

まずはお見積り致します! ご不明な点がございましたらお気軽にご相談下さい。

環境構築済みだから、導入後すぐに使える!

ローカルLLM / RAG構築済みHPC

用途に合わせて選べるハードウェアのご案内

NVIDIA DGX Spark / HPC Workstation / AI Server

お客様の用途に最適なお提案を致します。

DifyやOllamaを用いた小規模な検証環境から、LlamaやGemmaによるモデル評価、RAG構成を含む社内データ活用、さらにはXinferenceやGPTOSSを用いた本格的な研究開発まで、用途や運用規模によって求められるシステム要件は大きく異なります。特にGPU性能やVRAM容量、複数プロセスの並列実行可否は、応答速度や安定性に直結する重要な要素です。本チラシでは、想定される用途や使用規模に応じて、最適なGPU構成・拡張性を備えたPCをご紹介します。研究開発の目的に合わせた構成提案も可能ですので、お気軽にご相談ください。



ローカルのAI開発に革命を起こす小型AIスパコン。



- OS : NVIDIA DGX OS
- GPU : NVIDIA Blackwell Architecture
- CPU : Arm 20 core, 10 Cortex-X925 + 10 Cortex-A725
- メモリ : 128GB (統合システムメモリ)
- SSD : 1 or 4 TB NVME.M2
- 150mm x 150mm x 50.5mm (LxWxH)

RAG・LLM環境 セットアップ済

開発ツール : Dify
実行ツール : LMStudio/Ollama/Xinference
AIモデル : Llama/Gemma/gpt-oss-20B

MSI EdgeXpert

1TB 販売価格 648,000円(税別)
4TB 販売価格 753,000円(税別)

Xeon-WとRTX PRO6000 Blackwellを搭載したGPUワークステーション



- Ubuntu 24.04 LTS インストール代行
- インテル Xeon w3-2525 プロセッサ(8C/12T)
- 128GB (16GB x8) DDR5-5600 Registered ECC DIMM
- 2TB M.2 NVMe-SSD 高耐久仕様
- NVIDIA RTX PRO 6000 Blackwell Max-Q Workstation Edition 96GB | GDDR7
- 3年間センドバック方式ハードウェア保証
- 約 (W) 175 x (D) 500 x (H) 435 mm

RAG・LLM環境 セットアップ済

開発ツール : Dify
実行ツール : LMStudio/Ollama/Xinference
AIモデル : Llama/Gemma/gpt-oss-20B

CERVO Grasta TypeALIS25WC-PRO6000MQ

販売価格 4,078,000円(税別)

生成AIの常識を覆す、爆速推論。NVIDIA H200で次世代の計算速度を



- Ubuntu 24.04 LTS インストール代行
- インテル AMD EPYC 9365(36C/72T)
- 512GB (64GB x8) DDR5-5600 Registered ECC DIMM
- 3.84GB U.3 NVMe-SSD 高耐久仕様
- NVIDIA H200 NVL 141GB
- 3年間センドバック方式ハードウェア保証
- 約 (W) 438 x (D) 831 x (H) 176 mm

RAG・LLM環境 セットアップ済

開発ツール : Dify
実行ツール : LMStudio/Ollama/Xinference
AIモデル : Llama/Gemma/gpt-oss-20B

HPC AI Server Type-MSHP4UEP1S12B H200 NVL

販売価格 ※お問い合わせください

APPLIED アプライド株式会社 広域システム営業部

■関東営業所 東京都千代田区神田須田町1-2-7-8F ☎ TEL : 03-5280-9255
■東海営業部 名古屋市西区上名古屋三丁目25-28-5F ☎ TEL : 052-325-2782

■関西営業部 大阪市淀川区西中島2丁目14-6-5 ☎ TEL:06-6838-4123
■九州営業部 福岡市博多区上牟田1丁目6-23 ☎ TEL : 092-481-7812