

# Dell AI Factory

AMD を搭載した Dell の生成 AI ソリューション

複雑な生成 AI に対応する拡張性に優れたモジュラー型アーキテクチャにより、イノベーションを加速し、コストを削減して、データを保護します。

75%

パブリッククラウド IaaS と比較した、推論のコストパフォーマンスの向上率<sup>1</sup>

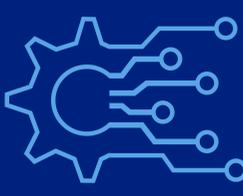
86%

タイムトゥバリューの短縮率<sup>2</sup>

主なユースケースにはパワー、柔軟性、拡張性が必要



アシスタント、チャットボット、コンテンツ作成



アクセラレーター  
アズアサービス



マルチモーダル  
検索拡張生成 (RAG)



## シンプル化

34 万時間を超えるエンジニアリング時間に裏打ちされた実証済みソリューションで、生成 AI の導入を合理化します。

### パフォーマンスの最適化

高性能アクセラレーター、オープンアーキテクチャ、AI 最適化ファブリック

### 場所を問わない AI の活用

マルチクラウドストレージの柔軟性で遍在するデータに対応

### ターンキー マルチノード

実証済みのフルスタック AI 基盤により、迅速に成果を達成



## カスタマイズ

AMD ROCm™ オープンソースソフトウェアとオープンエコシステムが、AI の開発と運用を促進します。

### イノベーションの加速

オープンソースのソフトウェアとエコシステムを使用して、独自のアプリケーションを開発

### 開発の高速化

業界標準のフレームワークと柔軟なテクノロジー スタックを活用

### データのアクティブ化

複数の AI ユースケースを同時に効率的に実行



## 信頼性

82% の IT 導入決定者がオンプレミスまたはハイブリッドモデルを希望<sup>3</sup>。データで成果が決まります。そのデータを保護しましょう。

### スピーディに開始

ルートオプトラスト、セキュリティ、完全な制御を備えたオンプレミス基盤

### 接続性の合理化

拡張性と最適化されたトラフィックフローを備えた安全で機能豊富なファブリック

### プロビジョニングの自動化

ハイパフォーマンス クラスターをデプロイし、管理するためのオープンソース基盤

## AMD を搭載した Dell の生成 AI ソリューション



### 推論

単一の AMD Instinct™ MI300X アクセラレーターで 70B パラメーターモデルを実行します<sup>4</sup>。

### カスタマイズ

8 つの同時 70B モデルを単一の Dell PowerEdge XE9680 に展開して微調整できます<sup>4</sup>。

### 強化

データを生成プロセスに組み込みます。

実績のあるオープンソリューションで競争上の差別化を達成し、安全なオンプレミス AI アプリケーションを大規模に提供します。

[詳細はこちら](#)

<sup>1</sup>Enterprise Strategy Group, 『AI の ROI を最大化 : デル・テクノロジーズのオンプレミスでの推論により、パブリッククラウドと比較してコストパフォーマンスが 75% 向上』 (2024 年 4 月)。

<sup>2</sup>Dell の分析に基づく推定値です (2024 年 5 月)。自動化スクリプトを使用した汎用 LLM 用の 2 ノード Kubernetes クラスターのセットアップにかかる時間と、一般的な設計を手動でデプロイした場合の時間を比較しています。セットアップ時間には、基本インストールのみが含まれます。実際のセットアップ時間は、ソリューションの構成によって異なります。

<sup>3</sup>デル・テクノロジーズ, 『Generative AI Pulse Survey』 (2023 年 8 月および 9 月)。

<sup>4</sup>デル・テクノロジーズのブログ, 『Silicon Diversity: Deploy GenAI on the PowerEdge XE9680 with AMD Instinct MI300X Accelerators』 (2024 年 5 月)。