

## AIプラットフォームの実装

## AIソリューションの複雑さを克服して価値創出を加速



AIと生成AIが持つ変革の力を活用するには、迅速な対応が必要です。多くの組織は、旧式のITインフラストラクチャ、非効率な運用モデル、スキルギャップの拡大に苦慮しています。さらにAIワークロードの需要の高まりに伴い、AIの実装に伴う状況はますます複雑化しています。テクノロジーを実装および統合するにあたってスピード、コスト効率、拡張性について妥協したくない場合は、技術エキスパートの経験と知識が不可欠です。

ユースケースとビジネスオブジェクトに基づいてカスタマイズされ、セキュリティと信頼性に優れたソリューションを実現するには、信頼できるサービスプロバイダーが必要です。このような課題を速やかに克服できる組織は、AIの可能性を最大限に引き出して、測定可能な価値と長期的な成功を促進することができます。

92%

AIプラットフォームにアクセスできる従業員が今後1年間で増加すると予測している組織の割合<sup>1</sup>

50%

ビジネス要件を満たすAIソリューションの調査、選択、導入、管理の能力に自信がないと回答している組織の割合<sup>2</sup>

84%

AIまたは生成AIソリューションの実装のためにサービスプロバイダーへのアウトソーシング/提携を希望している組織の割合<sup>3</sup>



Dell ServicesはAIソリューションの設計、実装、運用に関する課題に対処してパフォーマンス、拡張性、効率性を高めます。エキスパートがお客様のニーズを評価して最適なソリューションをお届けします。

- ✓ 運用の合理化
- ✓ 導入の向上
- ✓ ROIの最大化
- ✓ 柔軟に拡張可能

# 実装の複雑さ

# ROIへの影響

最適なAIソリューションが明確でない		目標に沿ったビジネス成果が得られない
実装に必要な技術的スキルのギャップ		価値創出までに時間がかかる
既存のワークフローへの統合		運用の非効率性
拡張して継続的に改善する能力		導入の遅れと競争力の後退
投資に関する不確実性		成長の機会を逃す

コスト効果に優れたオンプレミスソリューションのタイムトゥバリューを短縮

- ✓ ビジネスニーズに合わせてカスタマイズ
- ✓ 最適化されたパフォーマンス
- ✓ スケーラビリティを最大化
- ✓ 導入と活用の向上
- ✓ オーケストレーションと自動化による運用の合理化
- ✓ 統合データからの強化された出力
- ✓ オープンエコシステムの専門知識から得られる成果の最大化

Dell Servicesのエキスパートが、実装の複雑さを克服し、卓越したAIソリューションを導入できるようお客様をサポートします。インフラストラクチャのセットアップとプラットフォームの選択、統合と最適化から、拡張の余地のある運用効率まで対応します。

Focus Areas		
Use Cases		
Application Infrastructure	Training Frameworks	Inference
	Foundational Models / Model Training	
	MLOps / LLMOps	
Software Infrastructure	Workload Manager	
	Driver / Compiler / Toolchain	
	Operating Systems	
	Deployment Tools	
Hardware Infrastructure	Infrastructure from device to data center Dell compute, storage and networking	
		Partner GPUs

86%

ターンキーのフルスタックソリューションとサービスによりタイムトゥバリューが短縮される割合<sup>4</sup>



## 設計 & 計画

- AI使用パターン、GPU/ネットワーク最適化戦略、ベンダー選択のためのAIスタックコンポーネント、パフォーマンスに関する考慮事項などを特定する
- データの準備状況とコンプライアンスのニーズを確認する
- AI機能に対するデータセンターの準備状況を包括的に評価する
- ソリューションが目標に沿っており、実稼働と複数のユースケースに向けて簡単に拡張可能であることを確認する
- 具体的な統合ニーズを定義するワークショップを実施する



## インフラストラクチャの導入

- 工場でのラック構築と統合
- ラック間接続とGPUラック/スイッチ間のケーブル接続
- 電源管理と冷却システムの接続（水冷または空冷）
- クラスタ構築とフルラックテストおよび検証
- ネットワークファブリックの構成とテスト
- クラスタ全体の検収テスト（HPL、NCCLなど）
- 導入されたハードウェアに関する知識の引き継ぎ



## プラットフォームの実装

- 選択したソフトウェアスタックの実装と構成
- 推論サーバーの実装とデモ（事前トレーニング済みLLMを使用）
  - モデルトレーニング（ファインチューニング）ワークロードの例を作成し、その環境での運用を実証
  - 目指す成果に基づいて追加のツール/構成を提供
  - 機能を実証するためのテストスクリプト
- AIプラットフォーム導入計画



## 導入と拡張

- ストレージ、可観測性、ログ記録の各プラットフォームと統合し、最適なGPU割り当てを可能にするモニタリングによりシステムパフォーマンスを向上
- 本番運用の向上をサポートするために運用手順とワークフローを作成
- コンテナ化されたツール、モデルチューニング、トラブルシューティングによる高度な開発
- AIワークロードの規模拡張に必要な専門知識でスキルレベルを向上
- 専任の常駐エキスパートのサポートで、使用率を向上
- フルマネージドソリューションで運用を合理化

Day 0

Day 1

Day 2以降

どこから始めたらよいかわかりますか？ すぐに開始できるように [Accelerator Workshop](#) とユースケースの試験導入をご用意しています。詳しくは、Dellの担当者に今すぐお問い合わせください。

1. ESG [There's no time to waste: Confidently moving ahead with Enterprise AI] (Mike Leone, 2024年11月)

2. ESG [The AI Maturity Playbook: Driving Business Impact and Innovation] (Mike Leone, Dell・テクノロジーからの委託。実際の結果は異なる場合があります)

3. Dell Servicesの委託を受けてForrester Consultingが実施した調査 (2024年9月)

4. 推定値はDellの分析 (2024年5月) に基づきます。汎用LLM用の2ノードKubernetes®クラスタのセットアップに自動化スクリプトを使用した場合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。

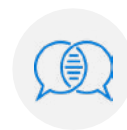
セットアップ時間には、基本インストールのみが含まれます。実際のセットアップ時間は、ソリューションの構成に応じて異なります



Dell [AIサービスの詳細](#)



デル・テクノロジーズのエキスパートに[問い合わせる](#)



#DellTechnologiesで会話に参加する