

AI-readyのデータプラットフォーム エンタープライズでの利用シーン



ここまでの流れ

- UDS事業におけるAIジャーニー
 - nVIDIAと取り組みと事例
- データ中心の世界を実現するために
 - 企業内の全てのデータを活用できる方法を提案
- PowerScaleのSuperPOD認定とDell Validated Designの紹介
 - サイジング

Contents

利用シーン

- オンプレミスでAIの利用や開発の基盤として
ストレージが利用されるシーンについて

1. お客様のタイプ別重点事項、業界別投資分野の分布
2. 企業・民間組織での潜在的なユースケース
3. ヘルスケア、大学、通信、製造での導入例
4. まとめ

お客様のタイプと導入に際しての重点事項

サービスプロバイダ

- **最高性能**のものを採用してインフラを整備（スペックの高さとそれに見合うネットワーク、ストレージ、スケールアウト性などを装備）し、ユーザの**要求のある都度柔軟に提供**できるサービスとして整える

公共・政府機関

- 政府機関は、**データ主権**に対する懸念から、AIに多額の予算を割り当てて、**国としての自前**の運営を目指す

民間企業

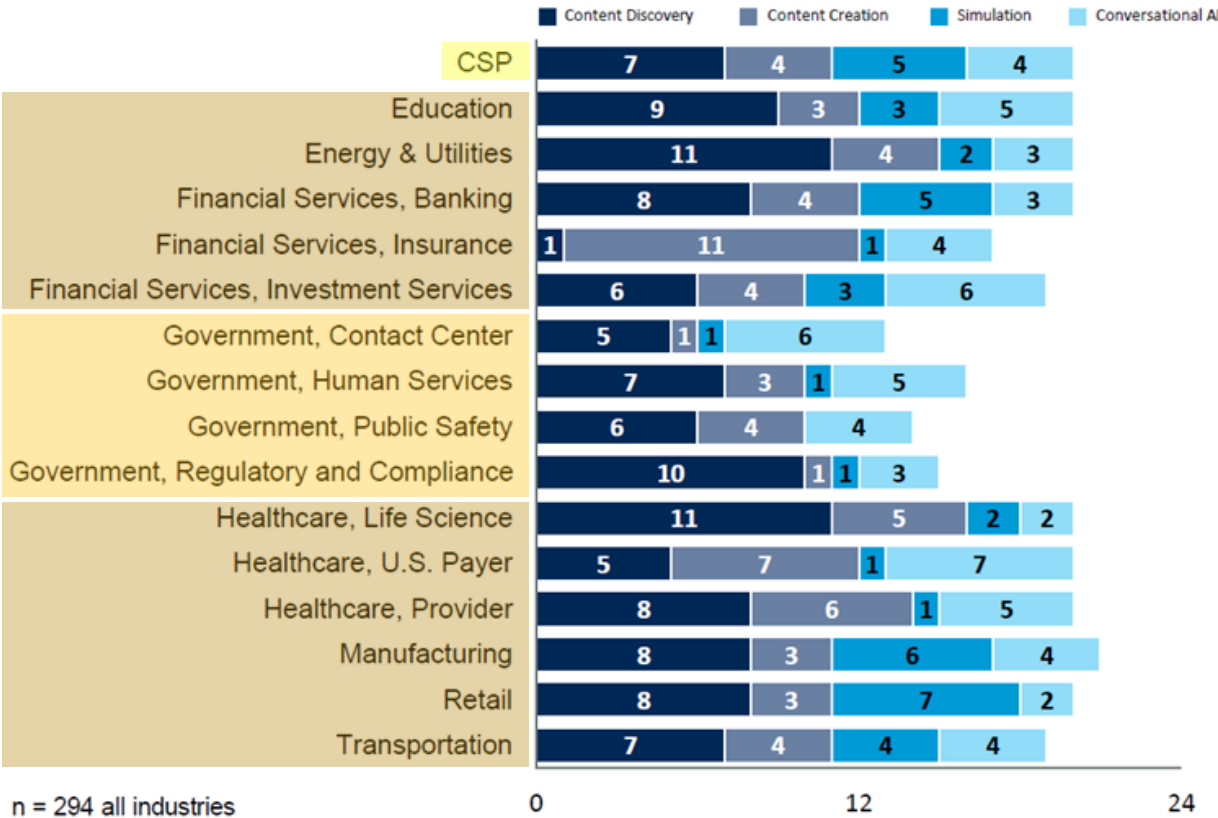
- ほぼ全ての企業がAIを活用した**戦略を策定中**
- **ユースケースを絞り込んだ上**、仮説を元にした**試験的な環境**を構築し、AIの取り組みを開始
- 主な検討事項には、**データの整理・管理、導入モデルの選択、既存のプロセスとの一貫性、効率の改善**など

業界別のユースケース分布

GenAIは
どのようなエリアに
展開されようと
しているのか

投資分野

- コンテンツディスカバリー
- コンテンツ作成
- シミュレーション
- 会話型AI



膨大な数のユースケース - 何処から始めるのか？

ビジネス戦略と運用管理	製品のイノベーションと研究開発	製造とサプライチェーン	マーケティング、営業、カスタマーサービス	IT、人事、ファイナンス
<ul style="list-style-type: none"> プロジェクト管理リソース/タスクのスケジューリング、予算編成と最適化、アラート リソース利用と持続可能性の最適化 販売パートナーへのインサイト、最適化指標、契約コンプライアンス状況/推奨事項 規制の変更と影響を特定して理解する 数値としての結果と成果を予測し、ワークロードと需要を予測 ファシリティマネジメントの最適化(スマートビルディング) 会議の自動文字起こし、要約されたメモ、フォローアップの割り当て 会議スケジュールの自動化(ボット) 従業員と顧客の感染症スクリーニング(安全衛生監視) 予測可能で適応性のある環境安全性スクリーニング 	<ul style="list-style-type: none"> デザインアイデアのブレインストーミングと画像生成 科学データと文献の検索と要約 障害を起こす可能性があるものをスクリーニング 満たされていないニーズ、十分なサービスを受けていない顧客、意図しない結果を特定する 新しい製品/サービスの機会を特定する 医療/画像診断から導かれるパーソナライズされた治療 人口/公衆衛生分析とキャンペーン 	<ul style="list-style-type: none"> 在庫とサプライチェーンのロジスティクス、ルート選択、スケジューリング、予算の最適化、アラート 交通の管理と最適化、緊急時対応 製造プロセスの品質管理とアラート 材料の無駄を削減(材料と機材の監視・追跡、利用状況管理と予測、最適化) 売り場/ディスプレイ/場所の最適化 ビデオ/センサーによる監視、リスク評価と軽減 音/光/大気/水質汚濁のコンプライアンスと最小化 運搬機材、生産設備、機器の活用状況管理/生産性の最適化 運搬機材と機器の予知保全 	<ul style="list-style-type: none"> コピーライティングの下書き(キャンペーンメッセージング、顧客パーソナライゼーション、翻訳) 感情分析、顧客スコアリング、ネクストベストオファー 検索エンジンやキーワードの最適化/自動化 セールストレーニング/コーチング カスタマーサポートのコンテンツの審査、学習、検索、要約 カスタマーサービスチャットボット/デジタルヒューマン 顧客レビューの自動応答 非構造化データソースから導き出された市場動向 リードの特定、評価、ランク付け、ターゲティング 顧客離れリスクの評価とアラート 顧客/患者データの検証と保護 本人確認 	<ul style="list-style-type: none"> AIによるコーディングおよび文書化アシスタント コードQAとテスト、合成データ作成 ログスクリーニングと早期警告システム ネットワークエンジニアリングによるレジリエンス、最適化、自己修復 メタデータの抽出と合成(言語プログラムやAPI用クエリ言語などで読み取り可能なタグ付け) インテリジェント検索 ネットワーク、コンピューティング、ストレージハードウェアの予知保全 自動侵入テスト 候補者の履歴書選考 従業員のパフォーマンススクリーニング パーソナライズされたトレーニング(コンテンツと配信) 従業員の離職リスク評価とアラート 詐欺/浪費/乱用の検出と防止 人的資源と財源の予測、および予算編成

ほとんどの企業・組織には200を超えるユースケースがあると想定されます

企業戦略に沿った導入と適用領域の拡大まで AIジャーニーを簡素化するための道のり

戦略の策定

ビジネス上の利害関係者
との共同で…ユースケー
スとガバナンスを定義

戦略の実行・推進

策定したユースケースのため
に、導入モデルに適合できる
ようにデータを準備

適用

ユースケースへの効果を最
大化するための生成AIの導
入モデルを適用して実行

評価

実際の規模での運用を行
うことで、生成AIのため
のスキルを習得し、運用
可能な状態にする

道程

何から始めるのかを決定



ビジョンの策定とソリュー
ションのアーキテクチャ



ユースケースの決定と
それを実現するモデルの選択

必要なデータの準備



データマネジメン
トスキームの構築



データエンジニアリ
ング

AIプラットフォームと テクノロジーの導入



生成AIソフトウェア



生成AIのための
ハードウェア



PC

目的としたユースケース に適合するよう、 導入モデルのテスト



言語モデル



AIユースケースの実行

運用と評価



AIとデータのガバナンス



生成AIの評価と効果を
最大化するための対策

お客様の目標

病院機構

- 医師のがん診断における治験、創薬について生成AIが支援
- 医師と看護師の文書のレビューと要約
- 従業員向け生成AIインスタンス

コスト削減

クラウドコストの削減
データセンター設置面積の削減
消費電力削減



マイグレーション

データが増大してもハイパフォーマンスを維持できるための仕組み

セキュリティ

HIPAA規制により、
高度に制御され
カスタマイズされた環境



パフォーマンス

GCPではネイティブファイルで12ジョブしか処理できないため、ビジネスユニットごとに複数のGCPクラスターを動かしてボリュームを処理していた
(PowerScaleでは1日平均1,000ジョブを処理)

お客様の目標

大手私立大学・高等学校

セキュリティ

学生情報の保護と
内部の脅威に対する保護に
重点を置く



パフォーマンス

HPC、AI、生成AIワークロード向けに
将来を見据えた環境が望まれる
(1つのデータレイクに)



サポート

包括的なサポートとサービス



運用効率

学生向けのより効率的な
教育システムを構築したい
学生向けの生成AIイノベーションラボを構築



ソリューション

4x PowerScale
F600

4x PowerScale
A300

1x PowerEdge
XE8640

2x Dell S5232F
Switches



採用に至る主な提案機能

- パフォーマンス – NFS over RDMA、GPU Direct
- 拡張性 – 生成AIラボとHPCワークロードの基盤を単一のソリューションに構築
- セキュリティ – SmartLock機能、Ransomwareの検出/防御機能
- 古いデータをアーカイブノードへマイグレート

お客様の目標

通信事業者におけるGenAI導入

- 数百ある生成AIのユースケースを念頭に、将来のデータ戦略を検討するために導入
- 現在はAzure上でSnowflakeのインスタンスを使用
- ChatGPTを利用してChatbotを構築して利用中、次に内部データを使用してRAGを構築し、十万人規模の従業員へのリリースを計画
- カスタマーサービスのための数十万件の録音されたお客様との会話を分析および分類し、アップセルの機会を特定

セキュリティ

お客様とサービスを利用する企業
にとっての重要なコンポーネント
静止データ暗号化、データセット
スナップショット



パフォーマンス

パフォーマンス テストで大量のデータの
処理を実施
要件に応じたスケーリング



コスト削減

Azure の利用に掛かるコストを精査
消費電力を重視



運用の合理化

使い慣れたテクノロジーと拡張性、
コストの整合性によって効率を向上



ソリューション

5x PowerScale
F710

2x PowerScale
Z9264 BE
Switches

6x PowerEdge
XE9680

2x PowerEdge
XE8640

2x PowerEdge
R760XD

2x
PowerSwitch
S5232F



採用のポイントとなった機能

- セキュリティ
- エンドツーエンドのソリューションとサポート
- As a Serviceの価格設定で、拡張ニーズに合わせてソリューションを適切なサイズに調整可能

AIと分析を使用して生産を管理する例

予測可能な品質

部品の故障に原因を特定する

ラインを早い段階で生産調整し、
不良部品を排除



X線による亀裂の検出
金属組成、温度、湿度、圧力
振動、電力変動からの異常検出
部品サプライヤーの情報



光学的な異常検出
ニッケル/スズ比、はんだ温度
組立速度の管理、
フラックスメーカーの情報



焦げ、壊れ、未調理などの
光学異常検出
油の温度、調理時間の管理
原材料の供給源・種類・世代の
管理

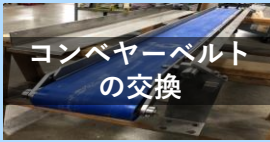




自動車、半導体

食品、医療、
軍事 / 航空宇宙、製薬

その他
多くの業界

Smart Factory:

AI/GenAI と Analytics を使用した生産管理の例

<p>予知保全</p>	<p>生産に影響が及ぶ前に 機器のメンテナンスの問題 を特定</p>	 <p>コンベヤーベルト の交換</p>	<p>光学検出(ビデオ、X線) ローラー振動 ベアリングノイズ(オーディオセンサー) ベルトの経年</p>
<p>インベントリ予測</p>	<p>在庫コストと在庫不足の リスクを軽減するための 高度な購買プロセス</p>	 <p>ジャストインタイム</p>	<p>主要な購入者履歴 季節性、コスト/調達し易さの傾向 サプライヤーに影響を与える 世界的なサプライの状況</p>
<p>安全性の担保</p>	<p>ルールの徹底 違反/ショートカットの検出</p>	 <p>衝突防止</p>	<p>ビデオ監視 顔認識 ロボット管理 手続きの徹底</p>
<p>仮想現実/拡張現実</p>	<p>製造現場での リアルタイムアシスト</p>	 <p>検査補助</p>	<p>メンテナンス手順 ライン機器の故障解析 製造不良の検出と修理</p>
<p>セキュリティ/ サイバーセキュリティ</p>	<p>工場およびIT/OTへの 物理的および リモートアクセスの保護</p>	 <p>データ保護</p>	<p>ビデオ監視 顔認識 エアギャップ ユーザーの振る舞い/プロセスの異常</p>

Dell Technologiesの優位性: Edge to Core to Cloud

データの発生から活用まで、データ管理の一貫性を実現



AIはデータから始まる

- ▶ このような業種のお客様では、自社のAI活用にあたって、活用可能な大量のデータを保有しています
- ▶ そのような場合、AIは学習やカスタマイズに利用するデータによって出来栄が決まります
- ▶ 先ずは、データの格納方法を整えることから始める場合もあります
- ▶ Dell PowerScaleやECSは、セキュリティにも配慮して設計されており、大量の非構造化データを扱うことができます
- ▶ PowerScaleは、単一のネームスペース、NVMe オールフラッシュの利用、GPU Direct機能が備わっており、優れた拡張性とパフォーマンスを提供します
- ▶ ECSは、スケーラブルで障害からの回復力が高く、コスト効率の高いオブジェクトストレージプラットフォームで、膨大な量のデータを効率的に保存・管理し、分析用の基盤として活用頂けます



ありがとうございました
宜しくお願い致します