INFINIDAT

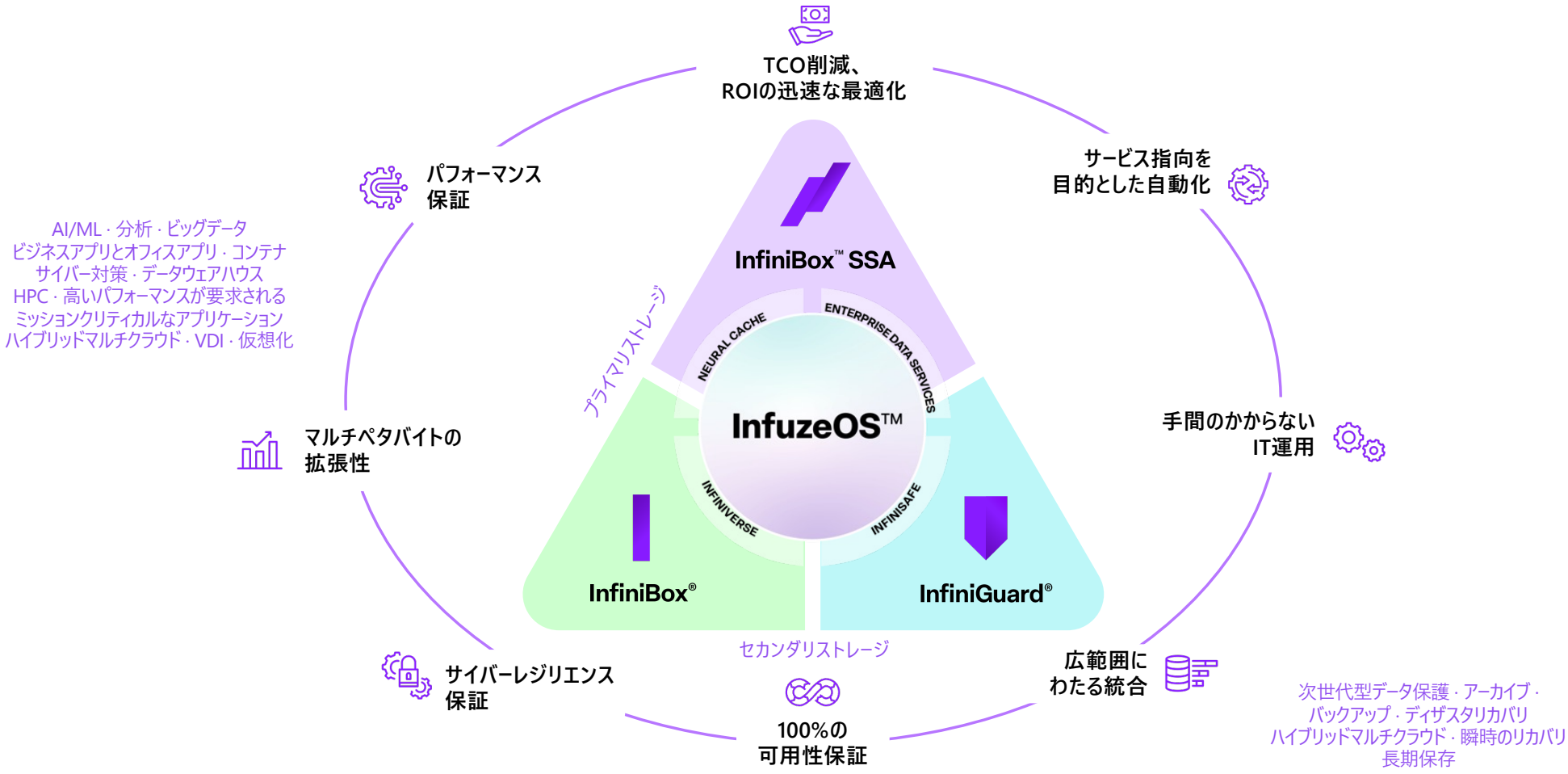
RAGワークフローの最適化：生成AIの応答速度が劇的に向上

安全性、迅速性、拡張性に優れた方法で、非公開の非構造化データをインテリジェンスへと変換

Eric Herzog – 最高マーケティング責任者（CMO）
eherzog@Infinidat.com

2025年3月5日

InfuzeOS - 1つのSDSアーキテクチャであらゆるプラットフォームに対応

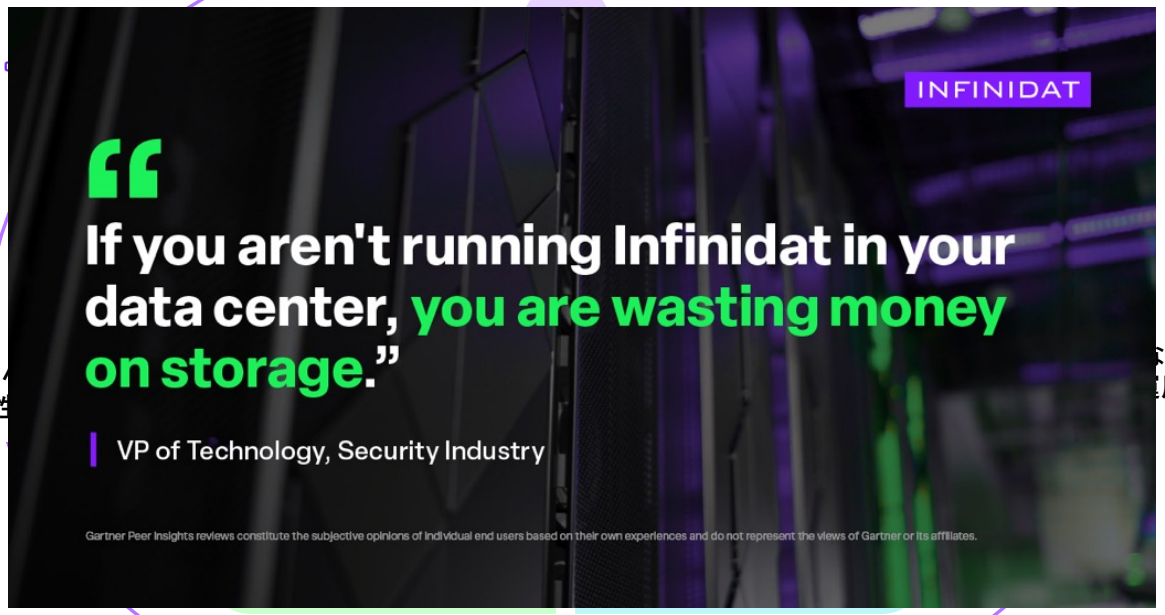


© 2025 Infinidat

InfuzeOS - 1つのSDSアーキテクチャであらゆるプラットフォームに対応

TCO削減、
ROIの迅速な最適化

AI/ML · 分析 · ビッグデータ
ビジネスアプリとオフィスアプリ · コンテナ
サイバー対策 · データウェアハウス
HPC · 高いパフォーマンスが要求される
ミッションクリティカルなアプリケーション
ハイブリッドマルチクラウド · VDI · 仮想化



マルチ
拡張性

ない
用

サイバーレジリエンス
保証

セカンダリストレージ

100%の
可用性保証

広範囲に
わたる統合

次世代型データ保護 · アーカイブ ·
バックアップ · ディザスタリカバリ
ハイブリッドマルチクラウド · 瞬時のリカバリ
長期保存

© 2025 Infinidat

業界での評価: 2024

大規模企業向け
プライマリストレージに関する
2024年レーダーレポート

Eric Herzog –
CMO Allianceにより
「CMO to Watch 2024」
に選出

Steve Sullivan –
2024 CRN®
Channel Chief

InfiniBox® G4 –
「最も革新的な
ハイパースケーラー実装」部門で
「Best of Show」賞を受賞

James 'JT' Lewis –
2024 CRN®
Channel Chief
(EMEAおよびAPAC)

InfiniBox™ SSA II –
エンタープライズストレージ向け
サイバー

Gartner® Peer Insights
's' Choice賞を
(全地域で4回、
で2回)

InfiniBox®とInfiniBox™ SSA –
「リーダー」の評価を獲得 –
プライマリストレージに関する
GigaOmレーダーレポートv.5

CRN®の「Storage Vendor 100」
に5年連続でランクイン

InfiniBox®とInfiniGuard® -
5 2PB+サイバーセキュア
バックアップターゲットとして
世界トップ5にランクイン

「Storage Awards」受賞 -
Capacity Storage
Vendor of the Year

Infinidat –
InfuzeOS™ Cloud Editionで
「2024 CRN® Tech Innovators
Awards」を受賞、InfiniSafe® Cyber
Detectionでファイナリストに入選

InfiniGuard
Editor's Ch
Product Award
ストレージのパフォ
バックアップ、最適化で
第2位を獲得

Eric Herzog –
「2024 CRN® Inclusive
Channel Leaders」受賞

SSA Express –
「Product of the Year」
のファイナリスト

Lynsy Marshall - チャネル
マーケティングリーダーとして
トップ100にランクイン -
Channel Insider

Infinidat –
マーケティングチームとして
世界トップ50にランクイン

受賞

Infinidat –「Infrastructure an
Storage Manufacturer
of the Year」のファイナリスト

Lynsy Marshall –
「CRN 2024 Women of
the Year」のファイナリスト

最もクールな
ソフトウェアデファインド
ストレージベンダー50社
(2024年) にランクイン

© 2023 Infinidat. All Rights Reserved.

生成AIのリスクとメリット

誤った応答や不正確な応答がもたらすリスク

- 法的およびコンプライアンス上の問題
- ブランドの評判やイメージの低下
- 顧客満足度の低下
- ビジネス機会損失による競争力の低下

生成AIによる正確なデータ生成のメリット

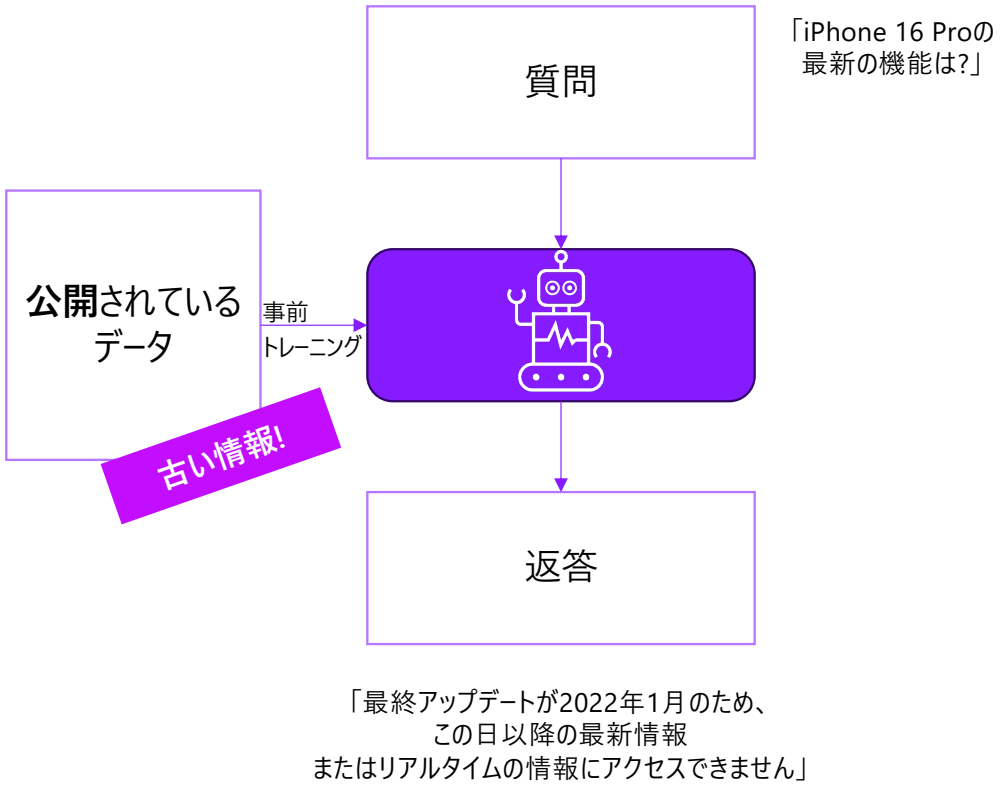
- 優れた顧客体験
- 生産性と業務効率の向上
- 顧客満足度の向上
- イノベーションの促進と迅速なコンテンツ作成
- コスト削減と利益拡大

**「Amazon（AMZN）では、
質問に返す答えが不正確という問題の
解決に時間を要しているため、
AlexaデバイスのAI対応アップグレードが
遅れる見込みです」**

https://markets.businessinsider.com/news/stocks/amazon-amzn-fails-to-provide-answers-in-ongoing-alexa-ai-delay-1034374823

生成AI - 課題は「正確な回答を得る」こと

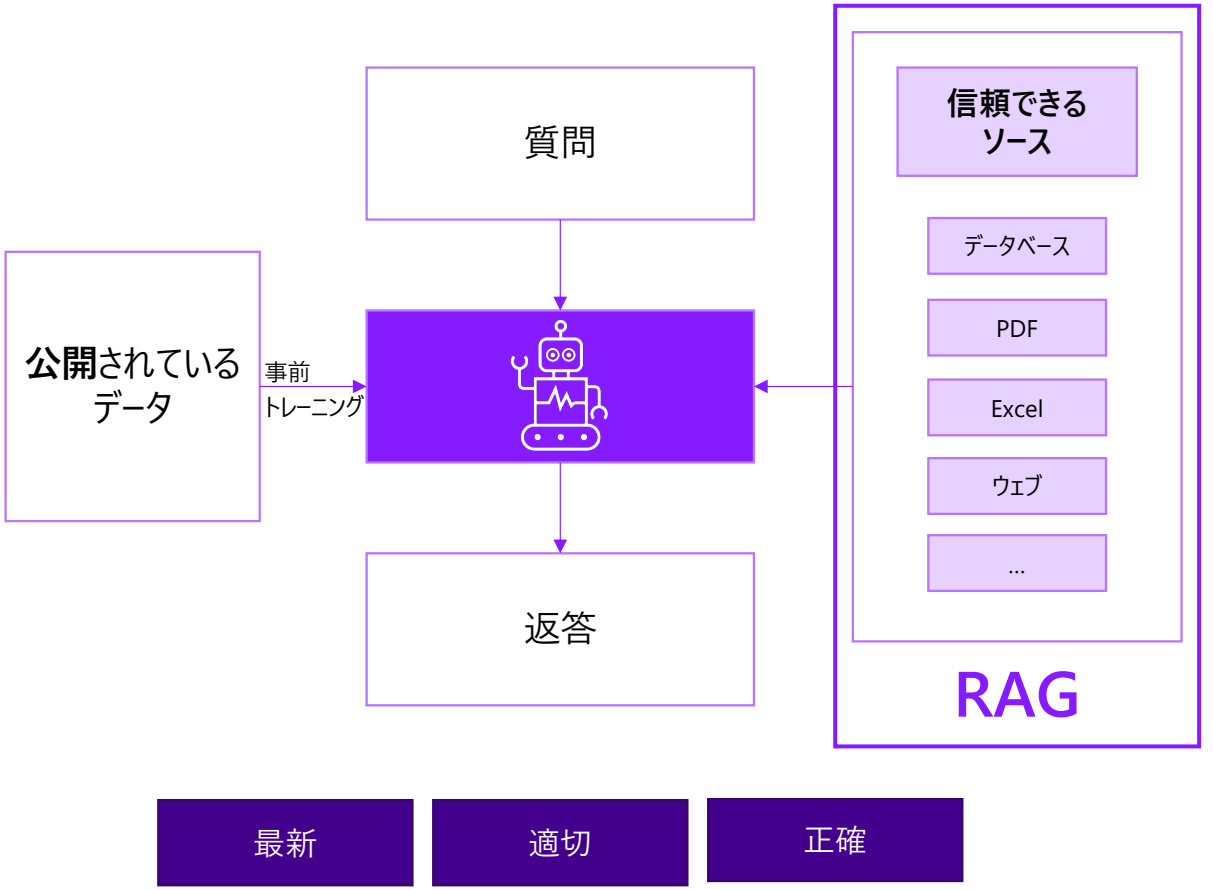
- 通常、LLMとSLMは公開データに基づいて
  トレーニングされているため、特定分野の情報や
  企業独自のデータを学習する機会が不足している
- その結果、「ハルシネーション」と呼ばれる誤情報や
  信頼できない回答が生成される
- もっともらしい回答に見えるが、実際は誤情報や
  古い情報
- 不正確な情報に基づいてビジネス上の意思決定を
  行ったり、顧客に誤った情報を提供したりすることに
  なり、企業の評判を損なう可能性がある
- 生成AIの信頼性を高めて、それを維持するには、
  信頼できる最新の正確なデータでAIを継続的に
  強化する必要がある



© 2025 Infinidat

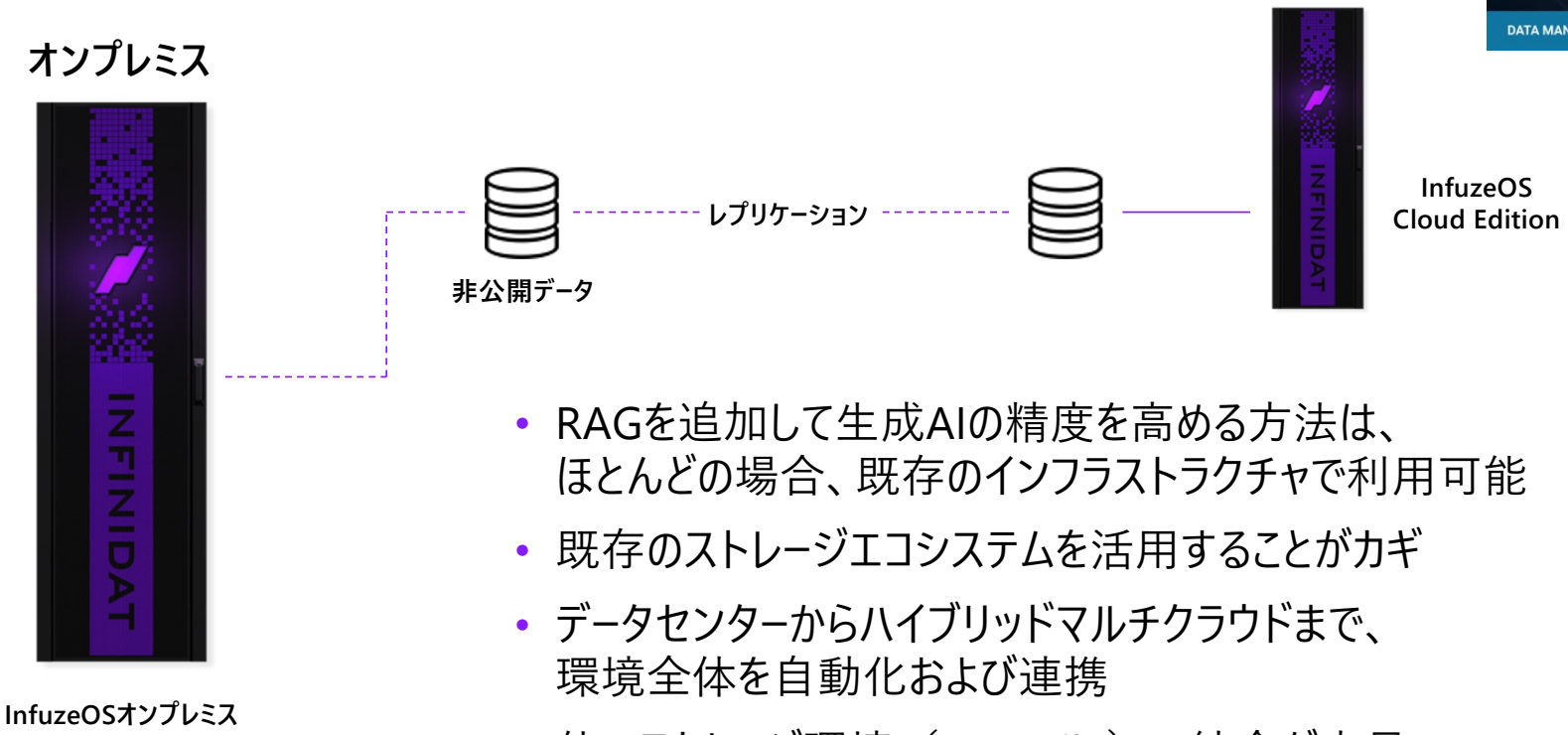
解決手段は検索拡張生成（Retrieval-Augmented Generation: RAG）

- 検索拡張生成（RAG）は、
  生成AIのパワーと企業の非公開データを
  組み合わせて、新たなインサイトの取得と
  ワークフローの効率化を可能にする
- 外部ソースからの情報をLLMとSLMに
  統合することで、より正確で目的に
  特化した成果が得られる
- エージェント型



© 2025 Infinidat

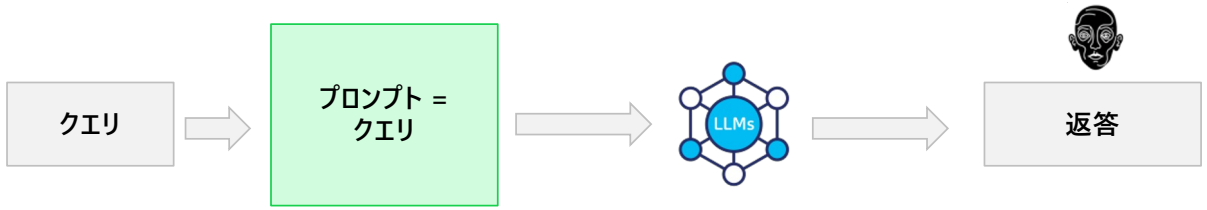
Infinidatの生成AIソリューション



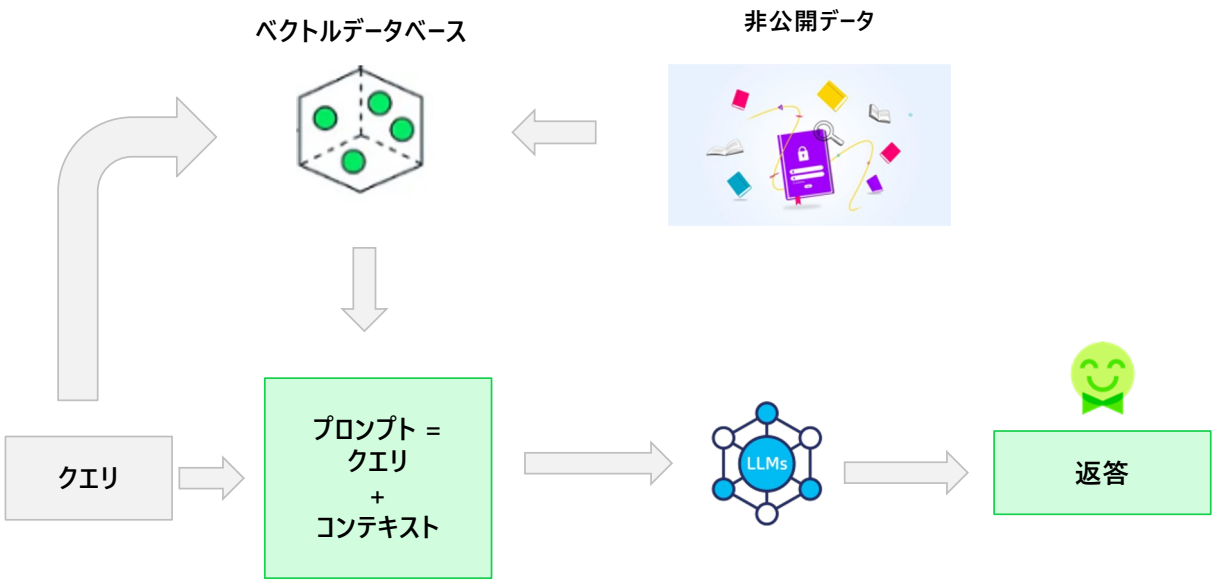
- RAGを追加して生成AIの精度を高める方法は、ほとんどの場合、既存のインフラストラクチャで利用可能
- 既存のストレージエコシステムを活用することがカギ
- データセンターからハイブリッドマルチクラウドまで、環境全体を自動化および連携
- 他のストレージ環境（NFS/File）の統合が容易
- エージェント型RAGにより、柔軟性と拡張性をさらに向上

© 2025 Infinidat

非公開データからインテリジェンスを得るためのベストツール



- 古いデータ
- ハルシネーション



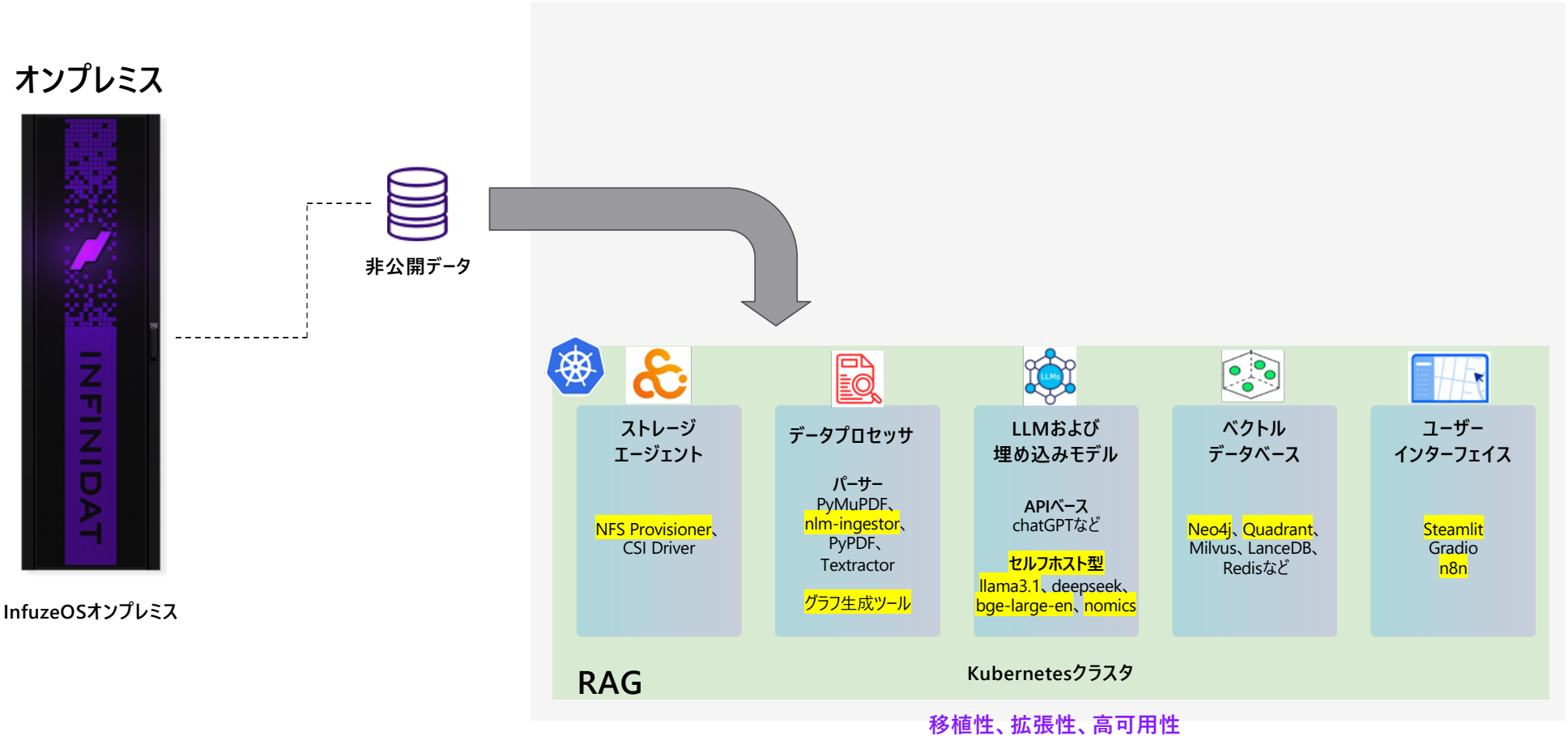
RAG

重要な課題:

- セキュリティ
- 拡張性
- 複雑な運用

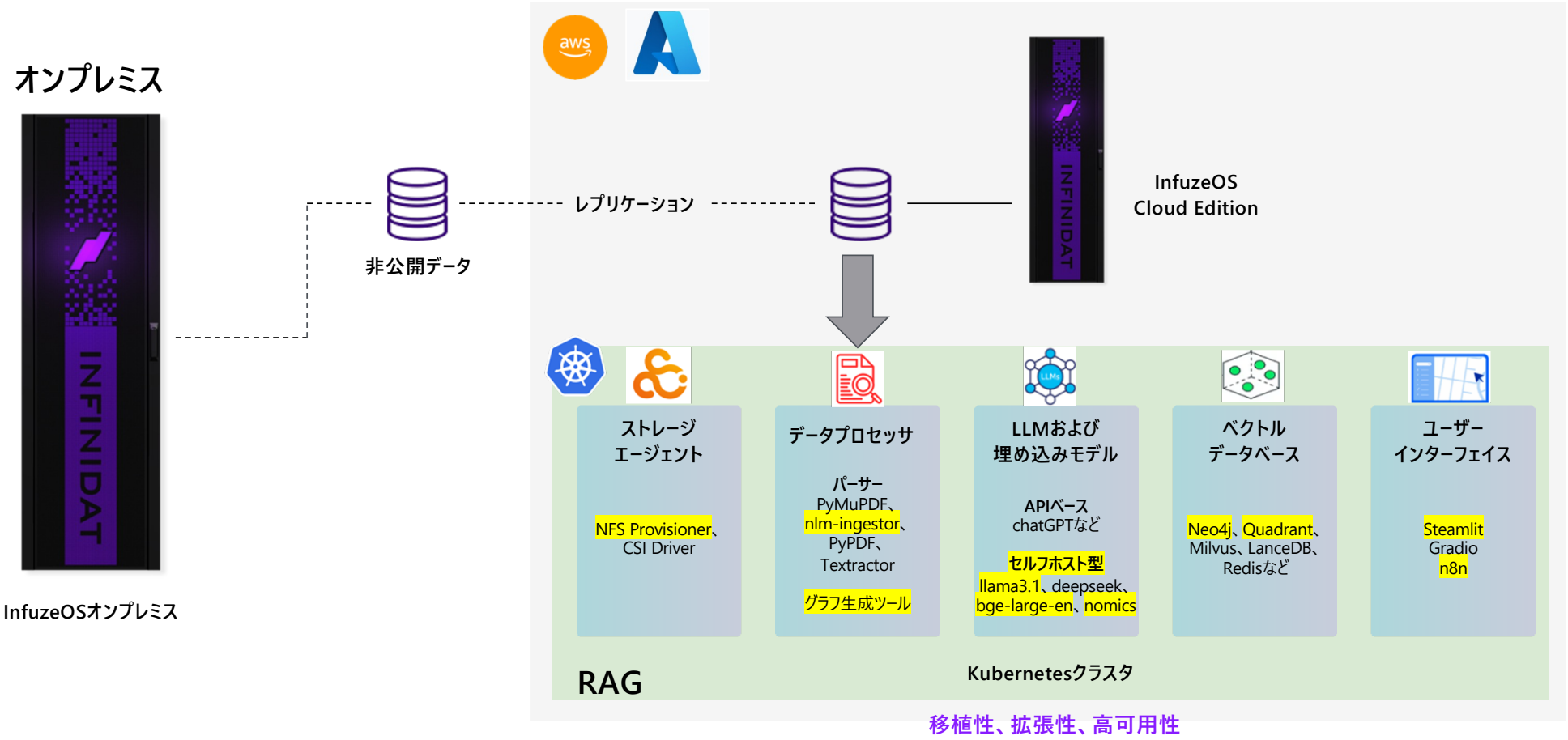
© 2025 Infinidat

安全かつ制限のないインテリジェンスを実現するフレームワーク（オンプレミス）



© 2025 Infinidat

安全かつ制限のないインテリジェンスを実現するフレームワーク（クラウド）



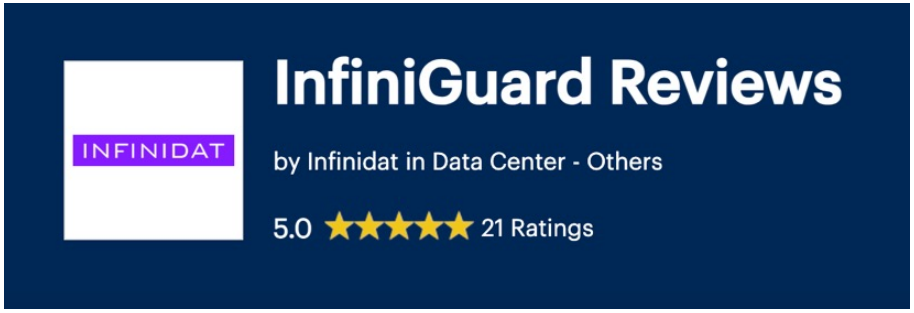
© 2025 Infinidat

RAGにおけるInfiniBoxの強み

- 現在、当社のお客様の多くは、ベクトルデータの格納と取得をサポートするデータベースを利用し、さらにInfiniBoxを使用してミッションクリティカルなアプリケーションを運用しているため、RAGをスムーズに導入可能
- InfiniBoxは、ニューラルキャッシュ（Neural Cache）を活用してきわめて低いレイテンシで超高速パフォーマンスを実現するため、埋め込み段階でRAGワークロードプロセスが停滞しない
- データを迅速に取得できるストレージシステム「InfiniBox」を大規模言語モデル（LLM）と組み合わせることで、一貫性があり、文脈を考慮した応答生成が可能
- オンプレミスと同様のストレージ運用をクラウドベースで実現するInfuzeOS Cloud Editionにより、ストレージ管理者は、強力なInfuzeOSをクラウドで利用して、データセットをAIリソースに迅速にレプリケーションできる
- エージェント型RAGアーキテクチャ活用により高い柔軟性を実現

© 2025 Infinidat

エンドユーザーの評価



2025年2月25日現在

© 2023 Infinidat. All Rights Reserved.

エンドユーザーの評価

2025年2月25日現在

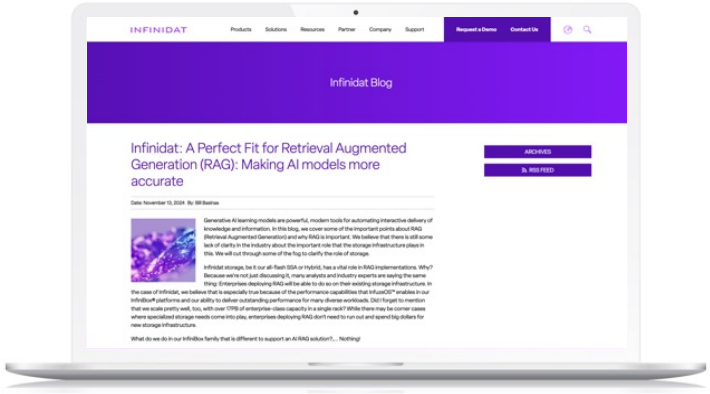
© 2023 Infinidat. All Rights Reserved.

AIソリューションに関するウェブサイト

https://www.infinidat.com/ja/artificial-intelligence

© 2023 Infinidat. All Rights Reserved.

AIソリューションの関連情報



ブログ:

Infinidat: 検索拡張生成（RAG）に最適なソリューション: AIモデルの精度を向上

https://www.infinidat.com/en/blog/infinidat-perfect-fit-retrieval-augmented-generation-rag-making-ai-models-more-accurate

アナリストによるホワイトペーパー:

https://www.infinidat.com/en/resource-pdfs/infinidat-cures-rag-response-time.pdf

アナリストによるビデオ:

https://blocksandfiles.com/2025/02/13/how-to-reduce-ai-hallucinations/

ソリューションに関するウェブページ:

https://www.infinidat.com/ja/artificial-intelligence

© 2025 Infinidat

ご視聴ありがとうございました

INFINIDAT