DELL Technologies

NVIDIA

生成AIのセキュリティに関する5つの重要な考慮事項

Dell AI Factory with NVIDIAを利用して、安全で拡張性のあるインフラストラクチャ基盤の導入を加速

生成AIに潜む変革の可能性

先見性のある人たちが想像し始めたばかりの方法で変革を実現する、生成AIに秘められた力。

76%

生成AIは組織に変革的な価値をもたらすと考えているITリーダーとビジネス リーダーの割合。[1]

AI

高度な分析と論理的な手法を使用して、イベントを解釈し、意思決定と行動をサポートし、自動化します。

生成AI

膨大な量のデータを活用して、自然言語プロンプトやその他のコード以外の非従来型の入力から新しいコンテンツを生成するテクノロジーと手法。

シミュレーション

- デジタル ツイン
- 合成データ
- 設計フレームワーク
- 予測

コンテンツの作成

- コーディング
- 数学
- 文章/音声
- 画像/ビデオ
- オーディオ

コンテンツの検出

- 自然言語検索
- 大規模なデータセット分析
- ナレッジ マネジメント
- パーソナライズされた教育とトレーニング

ユーザー エクスペリエンス

- 70以上の言語のリアルタイム翻訳
- 自然な表情とボディ ランゲージを使用したパーソナライズされた対話

1 デル・テクノロジーズ『Innovation Catalyst Study』（2024年2月）

可能性が高まる一方でリスクも増大

ビジネス リーダーは、データ、コンプライアンス、ガバナンス、その他のリスクに伴う影響を回避して、迅速に行動したいと考える傾向があります。しかし、セキュリティに関して、生成AIは諸刃の剣です。

メリット

- 脅威検出の向上
- 運用効率の向上
- パーソナライズされたセキュリティ意識向上トレーニング

欠点

- 攻撃の巧妙化
- 高度なソーシャル エンジニアリング
- シャドーAI

33%

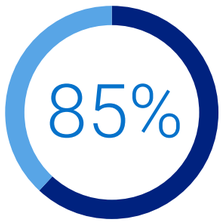
組織が軽減に取り組んでいる最も重要な生成AIリスクとしてサイバーセキュリティを挙げた回答者の割合。[2]

2 『McKinsey Global Survey on AI: The state of AI in early』（2024年5月）

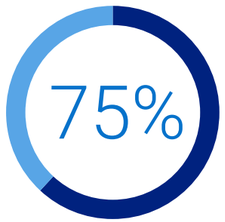
考慮事項1

新たな脅威ランドスケープ

生成AIには、将来性とともに厳しい現実があります。攻撃者は、従来の防御をすり抜けるさらに複雑な新しい攻撃を生み出しており、サイバーセキュリティ チームでは追いつくのが困難になっています。



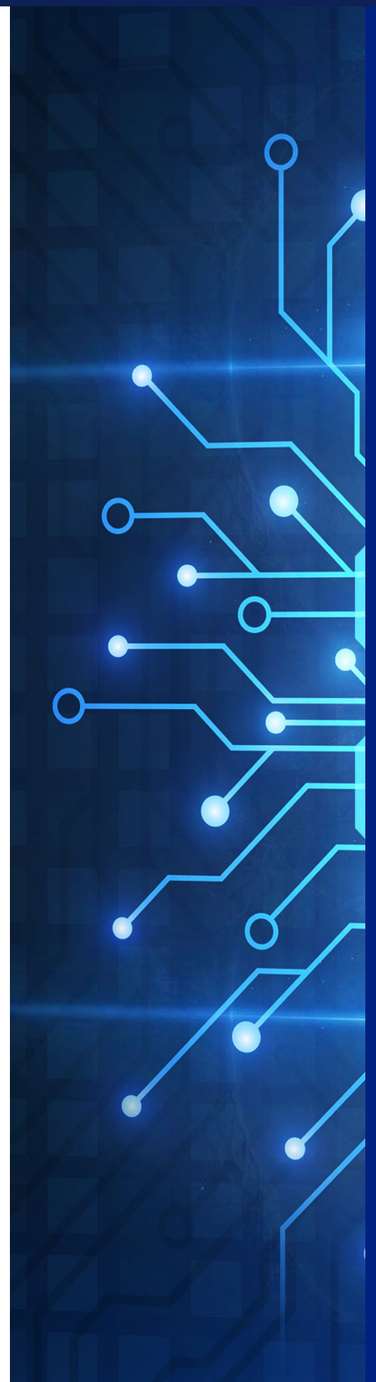
AIによってサイバーセキュリティ攻撃がより巧妙になったと考えている回答者の割合。[3]



過去12か月間に攻撃の増加を感じたと回答したセキュリティ プロフェッショナルの割合。[4]

このような新たな脅威から防御するために、企業は侵入テスト、モニタリング、監査などを通じて攻撃対象領域を最小限に抑えることに注力する必要があります。

3 『2024 Human Risk in Cybersecurity Survey』、EY（2024年5月）
4 Voice of SecOps Report『Generative AI and Cybersecurity: Bright Future or Business Battleground?』（2023年）



新たな攻撃ベクトル

高度なマルウェア
生成AIを利用して「自己進化」し続ける高度なマルウェアが増えており、シグネチャベースの検出といった既存のセキュリティでは検出されないように、コードを絶えず変更しています。

高度にパーソナライズされたフィッシング メールとキャンペーン
従来の詐欺の特徴がなく、本物に見える悪意のあるEメールの頻度が増加しています。

説得力のあるディープ フェイク データ
個人情報の盗難、金融詐欺、虚偽の情報が、文章、音声、画像、ビデオなどの人間の行動を模倣する能力によって容易になっています。

自動偵察
情報を収集し、潜在的なターゲットのネットワークやシステムの脆弱性と弱点を特定して、より標的を絞った攻撃を容易にします。

考慮事項2

導入と実装のリスク

生成AIの潜在的なメリットを活用したいと考えている組織には、大量の高品質なデータ、つまり、モデルが最良の成果を生み出すために使用できる入力データが必要です。ただし、データとリスクは密接に関連しています。情報を活用する前に、組織独自の要件、入力、リスクを慎重に評価し、考慮する必要があります。

大規模言語モデル(LLM)の脆弱性

生成AIサービスは、プロンプト インジェクション攻撃に対して脆弱です。この攻撃では、攻撃者が出力を操作してセキュリティ対策を回避したり、モデルの改良に使用された可能性のあるファイルに不正アクセスしたりします。

データ ポイズニング

攻撃者は、トレーニング段階で、改ざんされたデータをLLMに対して意図的に供給することができます。この結果、データに埋め込まれたバックドアを介した攻撃に対して、モデルが脆弱になるおそれがあります。実例として、スパムメールでスパム フィルターのトレーニングを実施し、それを攻撃および悪用することが挙げられます。

規制の複雑さ

世界中の規制当局が、生成AIの理解、制御、安全性の保証に向けて急いでいます。生成AIモデルは、データの保存、処理、使用の方法を規定する現在のデータ主権規則の対象となっていますが、規制当局は依然として知的財産権および著作権情報の監視を定義しています。規制の遵守には費用がかかる可能性がありますが、既存の規制および新たな規制に違反すると、罰金などの罰則が科せられるおそれがあります。

考慮事項3

シャドーAI

現在、多くの従業員はすでに、ChatGPTのようなテキスト、画像、ビデオの公開された生成プログラムを使用して、日常の作業を補強しています。しかし、これらのツールが適切なガバナンスなしに使用されると、企業の知的財産やデータを保護しようとする組織にとって重大な脅威となります。このような生成AIの不正使用は、シャドーAIとして知られています。

知的財産の喪失

すでに企業は、公開されている生成AIツールで従業員が機密情報を共有することにより知的財産が失われる状況に対処しています。

ソース コード データの漏洩

開発者がChatGPTを使用してソース コードを最適化しようとして、データ漏洩が発生しています。

シャドーAIの課題に対処するためには、安全なAIガバナンスに関して意思決定を行う権限を持つ、全社的な評議会や取締役会を設置する必要があります。

データはどこにあるのか？ ワークロードはどこに配置すべきか？

AIは、データがどこにあっても、それと組み合わされたときに最大限の効果を発揮します。インフラストラクチャとLLMを完全に制御すれば、知的財産の喪失やソース コード データ漏洩のリスクはありません。

コスト

オンプレミスの実装を活用することで、TCOを3年間で最大75%削減できます。[5]

セキュリティとプライバシー

オンプレミスのワークフローと運用により、組織全体に安全なAI/生成AI環境を構築します。特に機密データを扱う業界では、データ セキュリティとコンプライアンス規制の遵守を厳格に管理します。

5 Dellの委託によりEnterprise Strategy Groupが実施した、オンプレミスのDellインフラストラクチャとネイティブのパブリッククラウドインフラストラクチャ アズ ア サービスを比較した調査（2024年4月）に基づきます。分析されたモデルでは、ユーザー数5,000人の組織でのRAGを活用した70億パラメーターのLLMは最大38%コスト効率が高く、ユーザー数50,000人の組織でのRAGを活用した700億パラメーターのLLMは最大75%コスト効率が高いことが示されています。実際の結果は状況によって異なります。『Economic Summary』

考慮事項4

評価基準

過去1年間、AIコミュニティーは、責任ある開発と導入、影響の評価、リスクの軽減という3つの重要な課題にますます焦点を当てるようになりました。企業が生成AIモデルを評価する際には、いくつかの重要な注意点を考慮する必要があります。

一貫した報告要件がない

トップクラスの開発者は、さまざまな信頼できるAIベンチマークに照らしてモデルをテストします。報告の標準化が不十分なため、主要なAIモデルのリスクと限界を体系的に比較することは困難です。

出力に含まれる著作権保護された素材

一般的なLLMの出力には、著作権保護された素材が含まれている可能性があります。これにより法律に違反し、その素材を使用する企業が罰則のリスクにさらされるおそれがあります。

脆弱性の複雑化

研究者たちは、モデルに無限にランダムな単語を繰り返させるなど、あまり明白はないものの、LLMが有害な行動を示す原因となる戦略を発見しています。

開発者の透明性の欠如

多くの場合、AI開発者はトレーニングデータや方法論について積極的に情報を公開していません。これにより、AIシステムの堅牢性と安全性の理解を深める取り組みが妨げられています。

考慮事項5

セキュリティ上のメリット

生成AIによってもたらされるものは、セキュリティ リスクだけではありません。潜在的なセキュリティ上のメリットもあります。生成AIは、今やサイバーセキュリティに欠かせない要素となりつつあり、新たな保護手段をもたらしています。

より豊富なインサイトへの迅速なアクセスと自動脅威検出を備えた拡張性のあるセキュリティ運用の構築を開始できるようになり、効率を高め、人手不足のセキュリティ チームを補完することができます。

脅威検知および対応

生成AIは、履歴データを分析し、パターンと異常を特定することで、進化し続ける新たな脅威をリアルタイムで認識できます。ネットワーク トラフィック、システム ログ、ユーザーの行動を継続的に監視して、セキュリティの脅威の可能性を示す不規則なアクティビティを迅速に特定できます。

その結果、適応性の高い脅威検出が可能となり、変化する攻撃ベクトルに迅速に対応して、新たなサイバー脅威に対するプロアクティブな防御メカニズムを提供します。

脅威のシミュレーションとトレーニング

企業は、生成AIを使用して、制御された環境でさまざまなサイバーセキュリティの脅威と攻撃シナリオをシミュレートできます。その結果、時間が重要になる状況で、チームはサイバー脅威を特定、対応、軽減するための準備を整えることができます。

詳細な分析と要約

生成AIにより、チームはさまざまなソースやモジュールから収集したデータを調査できるようになり、従来は時間がかかり面倒であったデータ分析を、より迅速かつ正確に実行することが可能です。チームは、インシデントと脅威評価の要約を自然言語で作成することもでき、効率が向上し、チームの成果が増大します。

パーソナライズされたセキュリティ意識向上トレーニング

生成AIの上に対話型AIを組み込み、AIアバターをユーザー インターフェイスに取り入れることで、組織は自然な表情とボディ ランゲージを使用して、パーソナライズされた対話を提供できます（24時間365日規模で利用可能）。これをセキュリティのトレーニングと教育に使用すると、より自然でカスタマイズされたインタラクティブな学習体験や自動評価などを提供できます。

Dell AI Factory with NVIDIA

業界初の包括的なターンキーAIソリューションが、AIの導入をスピードアップし、データをインサイトへ安全に変換します。Dell AI Factory with NVIDIAは、AIと生成AIの活用を模索する企業の複雑なニーズに対処します。最先端のインフラストラクチャとサービスをNVIDIA AIソフトウェアとともに使用すると、開発と導入が簡素化され、プロジェクトのタイムトゥバリューを短縮できます。

- ルート オブ トラストなど、主要機能を含むイントリンシック（内在的）セキュリティを特長とするインフラストラクチャが、セキュリティ侵害のリスクを軽減します。
- お客様が管理するオンプレミスのAIソリューションを使用して、知的財産の喪失につながる可能性のあるデータ漏洩からデータを保護します。
- 安全なアクセスでAIをデータに導入することで、厳格なコンプライアンス要件とデータ主権要件を満たします。
- データへのアクセス場所とアクセス権を制御して、ステークホルダーのプライバシーを保護します。

Dell AI Factory with NVIDIA

業界初のエンドツーエンドのエンタープライズAIソリューション

データはAIファクトリーとユース ケースを動かす燃料

最も有用なデータはオンプレミスとエッジにあります。デル・テクノロジーズは、最も有用なデータをAIで活用できるよう支援しており、そのデータの保存、保護、管理におけるリーダーです。

ユース ケースから成果へ

AIファクトリーが生み出すビジネス上の成果は、お客様が最優先するユース ケースによって推進されます。デル・テクノロジーズは、検証済みのソリューションとカスタマイズされたサービスを提供して、お客様の最も重要なAIユース ケースの導入をシンプルにします

セキュリティ リスクによってイノベーションが阻害される事態を回避

AIと生成AIの世界をナビゲートし、メリットを得られるようにお手伝いします。

戦略計画の策定

無料のAccelerator Workshop for GenAI

- 成功をもたらす戦略の策定に向けた取り組みを始める
- 課題とギャップに対処し、目標に優先順位を付け、機会を特定する
- 準備状況の評価を受けて、インフラストラクチャ要件、AIモデル、運用の統合などを詳細に把握する

技術的な準備

すぐに使用できるモバイル ラボ

成功に向けた取り組みを迅速に開始しましょう。NVIDIA GPU搭載のDell Precisionモバイル ワークステーション5690/7780と、開始を支援する2日間のコンサルティング サービスが含まれています。

- 生成AIのテストとデモンストレーション用の移植可能なサンドボックス環境
- 開発者向けにNVIDIA AI Workbenchプラットフォームで事前検証済み
- お客様のデータを使用して実装された最初のチャットボット ユース ケース
- 生成AIスキルを実験して構築するためのコスト パフォーマンスに優れた低リスクのアプローチ

NVIDIA GPU搭載のDELL PRECISIONモバイル ワークステーション5690/7780

今すぐ開始してください

DELL Technologies AI Factory WITH NVIDIA