レジリエンスの再定義：ジェネレーティブAI（生成AI）時代におけるサイバーセキュリティ

生成AIとサイバーセキュリティにおいて、攻撃側と防御側のどちらが有利でしょうか？世界経済フォーラムの「2024年グローバルサイバーセキュリティ展望レポート」によると、経営幹部の56%が、今後2年間で攻撃側が防御側よりも優位に立つと考えており、生成AI時代に向けサイバーセキュリティ体制を再構築する必要があると強調しています。

生成AIが脅威に対して与える影響

多くの企業や政府機関と同様に、サイバー犯罪者も生成AIを積極的に活用しようとしており、その結果、生成AIを利用したサイバー攻撃が増加しています。たとえば、ランサムウェア攻撃は顕著に増加しており、2022年末のChatGPTのリリース以来76%も増加しています[2]。これらの攻撃は、生成AIを利用したフィッシングを通じて開始されることが多く、地方自治体、教育機関、製造業、医療機関などの分野に影響を及ぼしています。Fraud GPTやPentestGPTなどの悪意ある大規模言語モデル（LLM）は、サイバー攻撃を容易にするコンテンツを作成しています。これらLLMは、ダークウェブ上で月額わずか200ドルで購入できるのです。

ChatGPT[3]のリリース以来、フィッシング攻撃も1,265%という驚異的な増加を見せています。たとえば、経営幹部を装って不正に資金送金を承認させる音声ディープフェイクが急増しています。最近では、香港銀行が高度なディープフェイク詐欺により2,500万ドルの損失を被りました。詐欺集団は、同社の最高技術責任者と他の従業員をデジタルに再現し、電話会議にて同僚に送金を指示しました[4]。

注1：2024年グローバルサイバーセキュリティ展望レポート
注2：アクセンチュアサイバー脅威インテリジェンスレポート
注3：SlashNextサイバーフィッシングレポート2023
注4：ディープフェイク詐欺集団、史上初のAI強盗で2500万ドルを奪う

ハッカー集団Ghost Secなどの犯罪グループは、ダークLLMを使ってPythonベースのランサムウェアを作成する実験を行っています。このランサムウェアは高度に難読化され配布されており、攻撃が成功する可能性が高まっています。

アクセンチュアサイバーインテリジェンスの調査によると、金融サービス、政府、エネルギーなどの特定の業界は、生成AI攻撃の標的になりやすいことがわかっています。これらの業界はより高度な技術を使用する傾向があるため、洗練された攻撃に対してより脆弱となります。そのため、金融サービスや政府などの部門は、生成AI攻撃に対する防御を積極的に開発し、カスタマイズしています。

生成AIの脆弱性

生成AIによって、組織はこれまで以上に広範囲な脅威や、より高度な攻撃者、新しい攻撃の対象になりやすくなります。組織がテストケースや個別のユースケースから大規模な生成AI実装に移行すると、システム内のユーザーのデータ増加や統合増加など、導入の規模と複雑さが増すため、サイバーセキュリティのリスクが増大します。これらのリスク増大は、生成AIモデルの混乱やプロンプト・インジェクションから、トレーニングデータの露出、盗難、操作まで、あらゆるものに及びます。このような脆弱性は新しいものであり、ほとんどの組織は対処する準備ができていません。これらの新たなリスクを適切に軽減するためには、シャドーAIの検出、LLMプロンプトと応答のフィルタリング、特殊なAIワークロード統合テストなどの新機能が必要です。

AIを利用した攻撃に対抗する場合でも、自社のAI環境を保護する場合でも、組織はセキュリティ体制を迅速にアップデートする必要があります。優位に立つための鍵は、設計段階からセキュリティを組み込むことです。

注5：アクセンチュアサイバー脅威インテリジェンスレポート

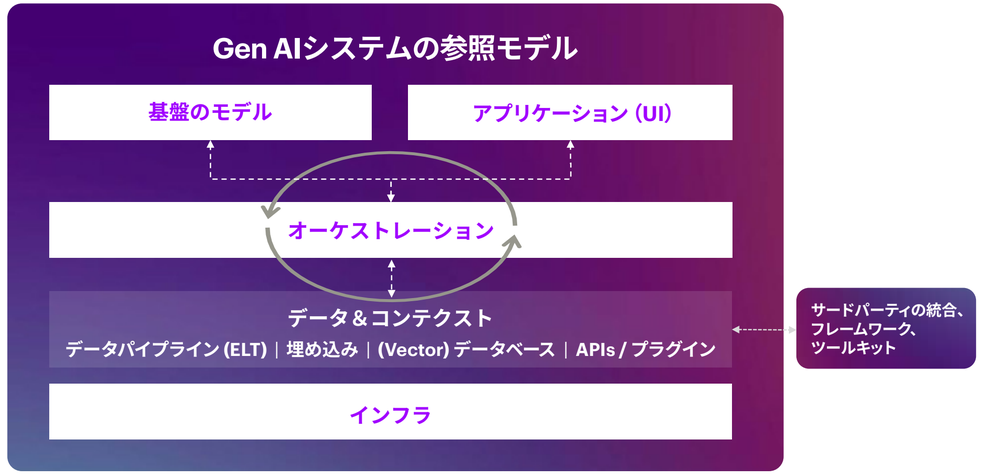
安全な生成AIの導入プロセスを加速

成功している企業は、セキュリティがビジネスを遅らせるものではなく、むしろ生成AIの成功を加速させる鍵であることを認識しています。生成AIの大規模な導入を加速させ、企業の生成AI環境を効果的に保護するためには、企業は以下の推奨事項を活用する必要があります。

ガバナンス、リスク、コンプライアンス（GRC）にAIセキュリティを組み込む：生成AIセキュリティはGRCの不可欠な部分であり、明確なガバナンスフレームワーク、ポリシー、プロセスを確立する必要があります。組織は、進化する規制にも常に対応する必要があります。たとえば、欧州連合AI法は、AIシステムがセキュリティを考慮して発展および展開されることを目的としており、バイデン政権の大統領令はAIの安全な開発・使用の基盤を築きました。規制当局と民間および公共のパートナーシップを結ぶことで、企業は将来の規制に影響を与えることができます。

生成AIのセキュリティリスクのレベルを評価する：最新のサイバーインテリジェンスに基づく包括的なセキュリティ評価を実施し、生成AI環境内の現在のセキュリティ成熟度を把握します。生成AIアーキテクチャを評価し、業界のベストプラクティスとの整合性を確保します。アクセンチュアの生成AIセキュリティ診断は、生成AIを安全に導入するために改善すべきポイントについてのアドバイスを提供します。

あらゆる層で生成AI環境を保護：組織は、データ層、基盤モデル、生成AIアプリケーション、IDアクセスとコントロールを含む生成AIスタック全体のセキュリティ保護に重点を置く必要があります。従来のセキュリティ対策を再現することもできますが、生成AI環境特有の脆弱性に対処するには、AI固有のソリューションも検討する必要があります。攻撃者と同じ生成AIを利用した攻撃を実施し、新たな脅威に対するテストを行い、組織のレジエンスを確保します。



生成AIでサイバーレジリエンスを改革

生成AIが、サイバーディフェンスとサイバーセキュリティに改革の機会をもたらすことは朗報です。生成AIを最大限に活用することで、組織は潜在的な攻撃者に対抗し、サイバーディフェンス能力を強化することができます。

従来のセキュリティソリューションだけでは、AIによるリスクに対抗するには不十分です。組織はAIを活用した防御テクノロジーを採用し、攻撃者が組織に対して使用する可能性のある生成AIテクノロジーを使用してテストをする必要があります。その例として、AIを活用したレッドチーミングやペネトレーションテストがあります。これらは、生成AIの規制が進化するにつれ、組織にとって必須となるでしょう。

多くのプラットフォーム企業やハイパースケーラーは、自社の環境やより広範な利用のためにAIセキュリティ機能をリリースしています。アクセンチュアのマネージド検出および対応（MxDR）サービスは、Google Cloudのセキュリティ固有の生成AIインテリジェンスを活用しており、一般的なセキュリティ環境や他のクラウドと統合できるように設計されています。この分野には、環境を保護するために生成AIセキュリティ固有のソリューションをゼロから作成した新しいプレーヤーも参入しています。

セキュリティベンダーを統合すると、複雑さが軽減され、全体的なセキュリティ体制が強化されることもあります。現在、多くの組織では40〜50種類のセキュリティツールを導入していますが、コストがかかり、安全ではありません。

クライアントストーリー

レンドリース：レンドリース社では、変化する世界的なリスクに適応できる柔軟性を必要としていました。そこで、アクセンチュアとGoogle Cloudが協力し、AIを活用したセキュリティユースケース向けに設計された特定の大規模言語モデルを搭載した次世代の検出および対応機能を作成しました。その結果、インシデントの検出、対応アクション、コミュニケーション、修復において改善が見られました。

ナショナルオーストラリア銀行：アクセンチュアはナショナルオーストラリア銀行と協力し、セキュリティを重視した生成AIを搭載したエコシステムへと同行の環境を変革しました。データレイク、生成AIアプリケーション、デジタルIDを保護するために、特定のデータ保護およびセキュリティ対策が実施されました。セキュリティは、生成AIの責任ある導入を可能にし、銀行のビジネス目標の達成を加速する上で重要な役割を果たしました。

問い合わせ

生成AIは、企業の変革、次世代のセキュリティ機能、そしてサイバーレジリエンスを促進させます。アクセンチュアが生成AI時代のサイバーセキュリティの改革にどのように貢献できるかを体験するには、当社のスタジオにぜひお越しください。

accenture