AI主導のセキュリティ

VECTRA

AIの兵器化の進化

このタイムラインは、AIが加速段階から真の兵器化へと移行した転換点を示しています。

2023年11月

ChatGPT のリリース
民主化の始まり

2023年

+1,265%
フィッシング攻撃の急増
大規模な実験段階

-31%
AIフィッシング vs 人間
依然としてAIは劣る

フェーズ 1
実現
AIが攻撃者のワークフローに組み込まれる

2024年

AIを使った
5つの国家レベルの脅威
運用上の正当性が評価される

長文コンテキストモデル
GPT-4 Turbo (128k) 及びClaude 3 (200k)
メモリが優位性となる

536.5B：AI/MLの
トランザクション
犯罪者による悪用が急増

フェーズ 2
加速
スピードと規模が重要になり始める

285件：GitHub上の
ジェイルブレイクリポジトリ
ガードレールが迂回され、一般化

64 – 92%
ジェイルブレイクの成功率
マルチターン攻撃が成熟

フェーズ 3
兵器化
制約が崩壊し始める

2025年3月

+24%
フィッシング攻撃 vs 人間

転換点
2年間で55%の増加
品質と規模が交差

2022年から2025年にかけて、AIを悪用した攻撃は実験段階から大規模で部分的に自律的な運用へと移行しました。そして、速度、規模、現実性を制約していた人間の限界を超える結果を残しています。

2025年：自律化

人間のペースを超えた攻撃

速度

100万トークン規模のコンテキストモデル

準備時間が数時間から数秒へ短縮

規模

80〜90%が自律運用

1人のオペレーターが複数のキャンペーンを管理

制御

42以上のAPTグループを確認

攻撃のペースが人間の限界に左右されない

人間の能力を超えた攻撃

AIの兵器化は新しい動きですが、AIを活用した防御はすでに存在します

Vectra AIは2011年より、防御者が攻撃者の振る舞いを検知する支援を目的として構築された機械学習を活用してきました。

Vectra AIが現代のネットワークを最新の攻撃からどのように保護するかをご覧ください

Vectra AIとは

Vectra AIは、現代のネットワークを最新の攻撃から保護するAI主導のサイバーセキュリティを提供しています。高度なサイバー攻撃が既存の制御を回避し、検知を逃れて顧客のデータセンター、キャンパス、リモートワーク、アイデンティティ、クラウド、IoT/OT環境にアクセスした場合、Vectra AI プラットフォームは攻撃のあらゆる動きを監視し、リアルタイムで点と点を結び付けて、侵入を阻止します。また、当社はAIセキュリティに関する35件の特許を取得し、MITRE DEFENDで最も多くのベンダーリファレンスを誇ります。他のツールでは検知できない攻撃を見つけ、阻止するために世界中の組織がVectra AIを活用しています。詳細については、https://ja.vectra.ai/ をご参照ください。.