MENLO SECURITY

生成AIが セキュリティポスチャに 与える影響とは

ChatGPTなどの生成AIプラットフォームの利用が人々の働き方をどのように変化させ、組織にどのようなセキュリティリスクをもたらしているのかを、改めて考えてみましょう

レポート NO.2

生成AI：多様化と懸念 そして未解決の問題

2023年11月30日、OpenAIのChatGPTはリリースから1年を迎えました。状況は日々変わっていますが、生成AIの利用をめぐるセキュリティリスクは、依然としてセキュリティチームやITチームにとって大きな懸念事項となっています。これまでの12ヶ月間で、生成AIの使用に伴う多くのリスクが表面化しました。これらには、これらのプラットフォームの使用時に顧客データや営業秘密、機密情報、さらには知的財産などの重要データが流出することが含まれます。

組織は、生成AIによって生産性を向上させイノベーションを可能にしながら、独自のデータやその他の知的財産の潜在的な損失を防止しなければならず、そのための適切なバランスを模索しています。適切なバランスを見つけるためには、組織は市場の最新状況を理解しておく必要があります。

2023年7月から2023年12月までの6ヶ月間で、生成AIをめぐる変化には以下のようなものがありました：

- **多様化と専門化：**市場への資金の流入が増加したことで、新しいプラットフォームが次々に立ち上がり、プラットフォームの数が大幅に増加すると共に専門化が進みました。
- **データプライバシーへの懸念：**組織はエンドユーザーからデータが流出することを心配しているだけでなく、データプライバシーに関して、プラットフォーム自体に対しても懸念を持っています。2023年3月、OpenAIから約120万人の加入者のデータが流出しました。これはOpenAIにとって初めての文書化された侵害で、このデータには、ユーザーの氏名、メールアドレス、支払い先住所、クレジットカードの種類、クレジットカード番号の一部、有効期限などが含まれていた可能性があります。この流出により、プラットフォーム自身のセキュリティや、モデルのトレーニングに個人データがどのように使用されているのかについて、疑問の声が上がり始めました。

- **自社専用AI：**多くの組織で、自社の特定のニーズに合わせてアルゴリズムを訓練する傾向が強まっています。社内にチームを作って、自社用のAIモデルを構築して維持するためには膨大なコストとリソースが必要になりますが、その代わりに汎用のAIプラットフォームを利用して自社のデータでアルゴリズムを微調整することで、安価かつ迅速に自社専用のAIを構築することができます。その際には、学習データが社外向けのアルゴリズムで使用されないようにします。自社専用AIではデータ流出の懸念は減りますが、組織が認識しておくべき、プライバシーとコンプライアンスに関する懸念は依然として存在します。このような動きは、ビジネス環境内での汎用生成AIプラットフォームの採用率の変化に繋がる可能性があります。
- **成長率の低下：**ChatGPTは、わずか2ヶ月で1億人以上のユーザーを集め、史上最も急成長したプラットフォームのひとつとなりました。しかし、生成AIの指数関数的な利用拡大はもはや続いていません。成長は続いているものの、勢いは鈍化しています。

これらの変化にも関わらず、生成AIは今後も存在し続け、進化していくことは明らかです。つまり、セキュリティチームとITチームは、進化するテクノロジーとポリシーに確実に追従し、環境の変化に合わせた保護を提供し続けなければなりません。

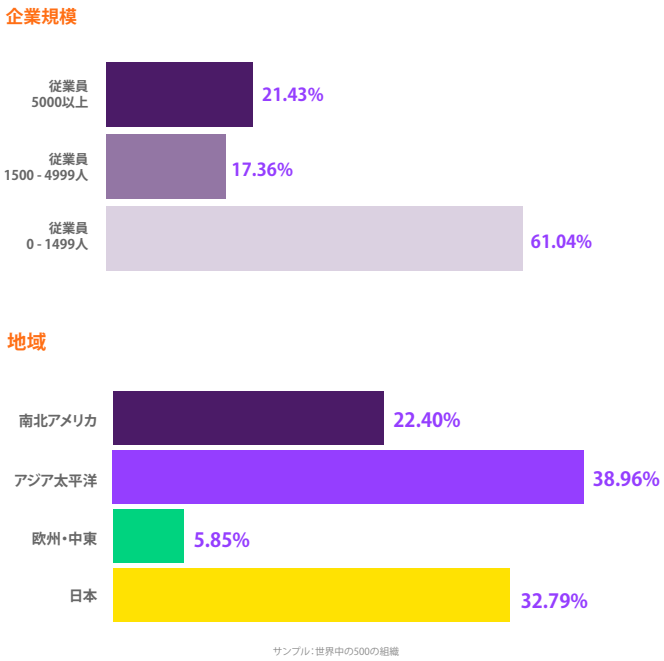
AIがフィッシング詐欺に与える影響：

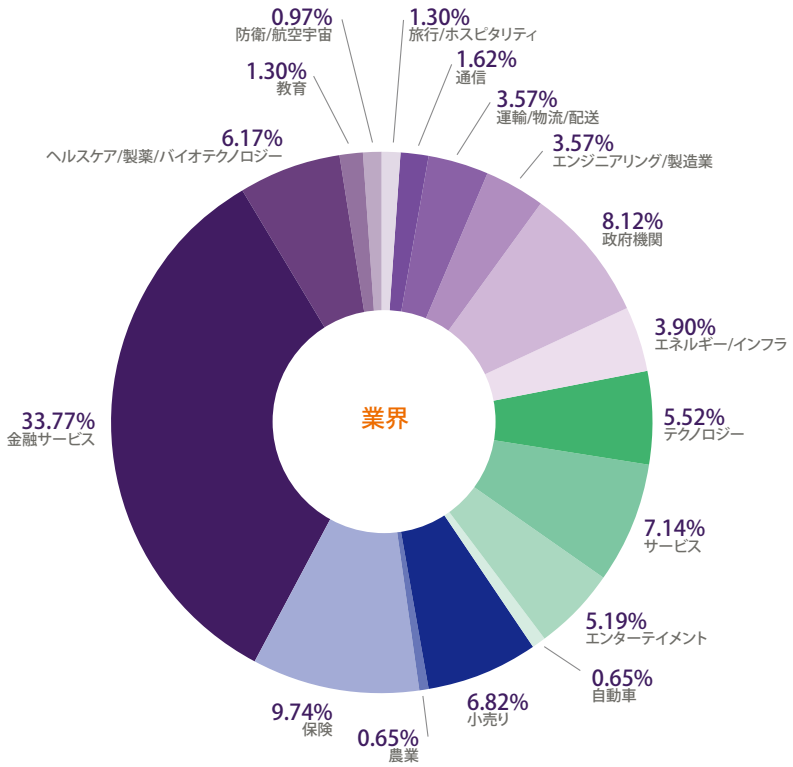
本レポートはデータ流出に注目していますが、AIがフィッシング詐欺を生成することについての懸念も高まっています。何百万人ものユーザーが日々の生活を改善するためにこれらのプラットフォームを利用しているのと同様に、悪意のある攻撃者もこれらのプラットフォームを利用してフィッシングキャンペーンを強化している可能性があります。
これまでのフィッシング詐欺ではスペルミスや文法の誤り、ぎこちない言い回しが目立ちましたが、ChatGPTのようなプラットフォームはこれらの問題をすべて解決してくれます。さらに、ハッキングコードを生成するためにChatGPTが使われる可能性もあります。悪意のある攻撃者がどのようにして生成AIを使用し、組織を標的にするかについての様々な方法を理解することは重要ですが、もっと重要なことは、この種の脅威を阻止するための防御策を講じておくことです。

メソドロジー

2023年6月、Menlo Securityは、世界中の500の組織をサンプルとして生成AIとのインタラクションを分析しました。このレポートNo.2では、従業員による生成AIの使用によってサイバーセキュリティがどのような影響を受けたかについての変化に注目します。

正確な比較を行うため、特に断りのない限り、データはNo.1と同じ6つの生成AI領域を比較します。また、本レポートでは、生成AIドメインをより広いカテゴリとして検討します。

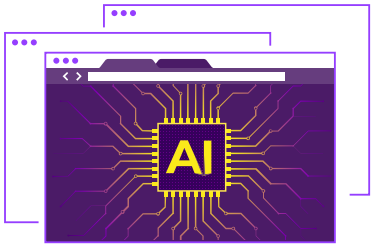




洞察1：

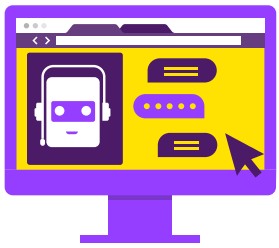
企業からの生成AIサイトへの訪問者数とパワーユーザーの数は一貫して増加している

前項で生成AI利用者の成長率が低下していることに触れましたが、企業内での利用者数は引き続き増加しています。この洞察は、ビジネスでの使用と個人での使用の違いに関連している可能性があります。ビジネス環境では、生成AIは新しいアイデアの創出、メール文の改善、コンテンツの作成、スペルや文法の間違いのチェックに役立っています。



100%以上

生成AIサイトへの訪問者数の増加率



64%以上

生成AIサイトを訪問するユーザーの増加率

ユーザーはどのようなサイトを訪問しているのか？

一部の生成AIプラットフォームの人気に変化が見え始めています。OpenAIのChatGPTがトラフィックの大半を占めていることに変わりはありませんが、その一方で特定の機能を持つ生成AIプラットフォームへのアクセスや利用が増加しています。たとえば、ライティング支援ツールのQuillBotAIや、その他の文法に重点を置いたAIプラットフォームが人気を集め始めています。

洞察2:

組織は生成AIサイトにセキュリティ重視の技術ポリシーを適用しているものの、その多くはグループベースではなくドメインベースである

生成AIトラフィックに対するセキュリティポリシーは26%増加しました。生成AIが注目を集める中、セキュリティおよびITリーダーは、組織内での生成AIの使用に関してセキュリティに重点を置いたポリシーを適用する必要性を感じています。しかし、大多数の組織における現在のアプローチは、拡張性に欠けるものです。

ドメインベースでセキュリティポリシーを導入している組織を見ると、次のことがわかります:

92%

生成AIの使用に関してセキュリティに重点を置いたポリシーを導入している

8%

生成AIの使用を無制限に許可している

グループ単位でセキュリティポリシーを導入している組織を見ると、次のことがわかります:

79%

生成AIの使用に関してセキュリティに重点を置いたポリシーを導入している

21%

生成AIの使用を無制限に許可している

本レポートの前半で述べたように、生成AIの状況は絶えず進化しており、特に新しいプラットフォームや新しい機能が生まれる際の変化が顕著です。ドメインごとにポリシーを適用するセキュリティおよびITチームの場合は、ユーザーが未知のプラットフォームにアクセスして機密データを流出させることがないように、リストを頻繁に確認する必要があります。このプロセスには時間がかかり、結局のところ拡張性はありません。組織は、生成AIのグループレベルでポリシー管理を可能にし、広範に生成AIサイトを横断して保護することができるセキュリティ技術を採用する必要があります。

* 組織では、生成AIの使用を保護またはブロックするために、他の技術を導入している場合があることに注意してください。

洞察3：

ユーザーはさまざまな方法でデータを入力している

ほとんどのユーザーはキーボードを打つことで質問を入力しますが、一般的にデータ流出の原因はこの他に2つあります。それは、ファイルのアップロードとコピー＆ペーストです。興味深いことに、ファイルのアップロードによるデータ流出の発生率は増加傾向にあります。
以前は、ほとんどのソリューションではネイティブでのファイルのアップロードが許可されていませんでしたが、最近の生成AIプラットフォームの新しいバージョンでは、ファイルアップロード機能のような新しい機能が追加されています。

コピー＆ペーストやファイルのアップロードは、大量のデータを簡単に入力できるため、データ流出に非常に大きな影響を与える可能性があります。例えば、以下のような例があります：

- ソースコード、顧客リスト、ロードマップ計画のコピー＆ペースト
- 数百の列を持つスプレッドシートのアップロード

ドメイン単位でセキュリティポリシーを導入している組織を見ると：

6%減少

生成AIサイトへのコピー/貼り付けイベントがブロックされた件数が6%減少

80%増加

AIサイトへのファイルアップロードの試行件数が80%増加

トラフィックの大部分は6つの主要なサイトに向けられていますが、生成AIをカテゴリとして見ると、ファイルのアップロードは70%増加しています。これは、ドメインごとではなくグループレベルでセキュリティポリシーを有効にすることの重要性を示しています。

洞察4：

ユーザーは依然として機密データを生成AIに入力しようとしている

生成AIに関するデータ流出の影響は十分に文書化されており、多くの組織が企業ポリシーを設定し、生成AIの責任ある使用に関するポリシーをユーザーに配布していますが、生成AIプラットフォームでは今でもデータ流出イベントが発生しています。これは、ユーザーが相変わらず、故意にまたは無意識に、これらのプラットフォームに機密情報を入力していることを示しています。これにより、適切なサイバーセキュリティ技術で補完することの必要性が注目されることになります。

従業員が生成AIに入力しているデータの種類		
PCI	0.03%	私たちは、ユーザーがどれ位の頻度で機密情報を生成AIプラットフォームに入力しようとしているかを分析しました。過去30日間に、以下のカテゴリに関連するデータのDLPイベントがありました： 最も頻繁に流出する可能性のある情報は、個人を特定できる情報でした。 これらの組織ではMenlo Securityを導入し、これらのインスタンスをブロックしています。
PII	55.11%	
機密文書	39.86%	
医療情報	0.90%	
制限された情報	3.44%	
その他	0.67%	

連邦政府機関における生成AIの活用

生成型人工知能（AI）は世界中で何百万もの人々の注目を集めていますが、政府機関では、生成AIをさまざまな業務でどのように活用するかを考える初期の段階にあります。

AIへの注目が高まったことで、「Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence（人工知能の安全で安心かつ信頼できる開発と使用）」に関する大統領令が発令され、AIの安全で責任ある使用を最優先課題としました。この大統領令は人工知能全般に関するものですが、政府機関が考慮しなければならない生成AI特有の重要な側面もあります。同大統領令には次のように記されています：

> 「政府機関は、生成AIの使用を広範かつ一般的に禁止したりブロックしたりすることは推奨されない」が、代わりに「少なくとも米国民の権利に影響を及ぼすリスクが低い実験や日常業務の目的で」生成AIを利用するための適切な安全策を講じることを強く求めています。

政府機関は、生成AIが国民、企業、政府に与えるプラスの影響と、潜在的なセキュリティリスクを阻止することのバランスをとる必要があります。

ソリューション

ユーザーエクスペリエンスに影響を与えずに生成AIを保護できる技術を採用する

組織は、生成AIを安全に使用できるように、適切なテクノロジーを採用する必要があります。生成AIプラットフォームの使用を完全に禁止することは現実的ではなく、政府機関でも推奨されていません。組織には階層化されたアプローチが必要で、データ漏洩防止（DLP）のような単一の技術ではなく、従業員がこれらのプラットフォームを使用するさまざまな方法に対応できる機能が必要です。

例えば、ユーザーは生成AIプラットフォームに大量のデータをコピー&ペーストしたり、ファイルをアップロードしたりします。これらの操作を制限することで、組織は従業員による生成AIツールの使用を妨げない方法で彼らを保護することができます。解決策は、文字数制限によるコピー&ペーストの制御です。文字数を制限したり、既知のコードをブロックしたりして、入力フィールドに貼り付けられる内容を制限することで、大量のデータ流出を防ぐことができます。何千行ものソースコードを手動で入力するユーザーはいないので、貼り付け機能を制限することで、このタイプのデータ流出を効果的に防ぐことができます。また、ユーザーは入力しようとしている情報についてよく考えるようになります。

組織は生成AIについてもっと知る必要があります。組織は、イベントログの記録やセッション監視を開始するような追加の制御をトリガーするセキュリティポリシーを適用して、問題の解決とイベント後の分析を支援することもできます。また、内部関係者による侵害の調査では、それが意図的だったかどうかの証拠を得る必要があることを忘れてはなりません。イベントやブラウジングセッションを記録することで、ユーザーが悪意を持っていたか、単に不注意だったかを把握し、洞察を得ることができます。

最後に、組織は生成AIグループレベルでセキュリティ制御を可能にする技術を採用する必要があります。これまで見てきたように、新しい生成AIサイトの利用が広がっており、データ流出の可能性が高まっています。ポリシーがドメインごとに適用される場合、組織はリストを継続的に更新するか、従業員が使用している生成AIサイトに対する保護にギャップが生じるリスクを負うことになります。

www.Menlo Security.com

Menlo Securityについて

Menlo Securityは、Menlo Secure Cloud Browserによって高度に回避的な脅威を排除し、組織の生産性を維持します。Menlo Securityは、クラウドベースのセキュリティが目指す、導入展開が容易なゼロトラストアクセスを実現します。Menlo Secure Cloud Browserは、エンドユーザーがオンラインで業務を行う間、ユーザーからは見えない形でサイバー攻撃から防御し、同時にセキュリティチームの運用負担を軽減します。

Menlo Securityは、ユーザーを保護してアプリケーションへのアクセスを確保し、完全なエンタープライズブラウザソリューションを提供します。Menlo Securityなら、ワンクリックでブラウザセキュリティポリシーを導入することができ、SaaSやプライベートアプリケーションへのアクセスを保護して、ラストワンマイルまで企業データを守ります。信頼と実績のあるサイバー防御により、あらゆるブラウザでデジタルトランスフォーメーションを安全に実現します。Menlo Securityと共に、安心してビジネスを前進させましょう。

データ損失を防ぎながら生成AIを安全に有効にする方法について、今すぐお問い合わせ下さい。