2026 年版 ThreatLabz:
フロンティア AI への対応状況レポート



目次



本書の要旨 3

数字で見る評価 5
評価対象グループ 6

脅威の見通し 7

調査方法 9

主な調査結果 10

新たな攻撃チェーン：
フロンティア AI が小さなギャップを自律的に
つなぎ合わせ一連の攻撃を成立させる仕組み 15

事例に基づく証拠：
進行中の 3 社 16

実装の現実性：
90 日間のスプリントで実際に必要となること 18

推奨 19

©2026 Zscaler, Inc.All rights reserved.

本書の要旨

昨年、ある有名な AI 企業から電話を受けました。ちょうどその企業が、自律的なセキュリティ テストを実行できるよう調整されたフロンティア AI モデルをリリースしたとの知らせを受けたところでした。その電話をきっかけに、こうした自律型攻撃に対する組織の対応状況を検証し、組織が備えられるよう支援する取り組みが始まりました。フロンティア AI モデルの能力に驚かされたのと同じくらい、調査結果にも驚かされました。

なお、私たちはフロンティア AI を組織の環境に適用したわけではありません。まずは自社のインフラ上でテストを実施し、そこで得られた知見をもとに、現在提供している評価プログラムを開発しました。

目的は、単にギャップを指摘することではなく、組織が新たな脅威の波に備えるなかで、セキュリティをどこで強化できるのか明らかにすることです。私たちは皆、ともに学んでいるところです。今回の調査結果は、単なる警告としてでなく、あらゆる組織のレジリエンスをどのように向上させられるかについて共有すべき教訓として役立つことを意図しています。

Zscaler は、ほとんどの組織が運用しているハイブリッド環境やレガシー環境とは異なるクラウドネイティブ アーキテクチャーを採用しています。そのため、実際のモデルを用いたフロンティア テストで使用されているのと同じ手法を用い、同等の有効性を持つモデルと TTP によって再現することで、こうした攻撃者が組織の環境内でどのように行動するかを模倣しました。これにより、自社独自のようなアーキテクチャーではなく、組織インフラの運用上の現実に即して評価プログラムを構築および調整することができました。

最大の発見は、AI に対する有効な対応策は単に AI を増やすことではなく、設計段階から悪用可能な経路を排除するアーキテクチャーであるということです。

世界中の組織は、フロンティア AI を悪用したサイバー攻撃に備えられていません。しかし、その理由は、多くの人が考えているものとは異なります。危険なのは、AI がこれまで発見されたことのない極めて特殊な脆弱性を見つけ出すことではありません。危険なのは、フロンティア モデルが明らかな脆弱性、つまりあらゆる組織がすでに存在すると把握しているギャップを高速で連鎖させ、完全な攻撃経路を構築できることです。ほとんどのアーキテクチャーは、そのような攻撃を想定して設計されていません。そして、人間であれツールであれ、AI を上回る速さで対応しようとするのは、勝ち目のない戦いを続けることになります。勝つための唯一の方法がアーキテクチャーです。幸いなことに、適切なゼロトラスト アーキテクチャーを導入していれば、ほとんどの組織は想像よりもレジリエンスの実現に近いところにいます。

フロンティア AI 対応状況評価プログラムでは、特定のタイプの攻撃者に対する組織のレジリエンスを評価します。それは高速で動作し、自律的な偵察、脆弱性の連鎖、ラテラル ムーブメントを実行できるフロンティア AI モデルです。私たちは、このタイプの攻撃者を表すために「フロンティア AI」という用語を使用しています。これは、従来のスキャナーや人間のアナリストでは同等の速度で再現できない方法で、攻撃対象領域全体を横断して推論できるためです。本番環境を対象とした並行テストでは、フロンティア モデルが、実際のコンテキストに基づいて高いシグナル対ノイズ比の推論を行えることが確認されました。これにより、攻撃チェーンの発見を大幅に加速させる機能が示されました。こうした能力は理論上のものではなく、すでに実証されています。

©2026 Zscaler, Inc.All rights reserved.

2026 年上半期に評価を実施した数百の組織を対象とした調査では、金融サービス、医療、製造、エネルギー、その他 8 業界の Fortune 500 企業が含まれ、フロンティア AI 対応状況スコアの平均は 100 点満点中 37 点でした。ほぼすべての組織 (91%) が公開可能な AI インフラを持っており、アイデンティティーベースのアクセス制御を導入していた組織は 1 社もありませんでした (100%)。さらに、セキュリティ スタックが設計どおりに機能するために必要な量の暗号化トラフィックを検査していた組織は、1 社もありませんでした。平均的な組織では自社ネットワークを通過するデータの 64% を事実上可視化できておらず、防御機能を保有していることとそれらを実際に運用することとの間にあるギャップを解消できている組織もありませんでした。

最も緊急性の高い新たなギャップは、上記で明らかになった構造的な問題をさらに深刻化させています。評価対象となったすべての組織で、AI ガバナンスはゼロからのスタートとなっています。コパイロット、推論 API、MCP エンドポイント、シャドー AI ゲートウェイなどの AI 接点は、セキュリティ プログラムが対応し、管理できる速度を上回るペースで増加しています。

これは、フロンティア AI の脅威への備えを向上させる大きな機会となります。調査の結果、約 3 分の 1 の組織では、重要なチョーク ポイントにデセプションを拡張することで防御を大幅に強化できることがわかりました。同様に、侵入防止を完全な施行モードに移行すること、アプリケーション層の異常検知を統合すること、セグメンテーション ポリシーを改良することは、SOC のアラート精度と全体的な防御態勢を向上させるための具体的かつ効果の大きい取り組みです。

セキュリティ リーダーは、このレポートを取締役会レベルでの議論を促すきっかけとして捉えてください。今回の調査対象グループの平均である 37 点から、73 ～ 98 点のフロンティア AI 対応状況スコアを目標とする水準まで引き上げることは、90 日以内に実現可能です。この 73 ～ 98 点という水準は、5 つすべての防御の柱が、フロンティア AI クラスの攻撃に対して運用上の有効性を発揮するための閾値です。そのためには、経営層による次の 3 つの取り組みが必要です。(1) フロンティア AI に対するレジリエンスを全社的な方針として確立し、取締役会への報告事項としてください。(2) 社内の経営層間で認識を統一し、集中的な取り組みのための専用のリソースを確保してください。(3) 二本立てのスプリントを実施します。つまり、休眠状態にある制御を有効化するとともに、すべての AI 接点にアクセス制御を適用してください。

すべての組織に共通する運用上の出発点は同じです。まず、すでにあるものから始めましょう。戦略上の出発点はガバナンスです。どの組織も、AI アクセスのセキュリティを先送りすることはできません。先送りすれば、現時点ですでに悪用可能な状態にさらされていることを受け入れることになります。



DEEPEN DESAI
Zscaler、サイバーセキュリティ担当 EVP



GEORGE MOSER
Zscaler、グローバル VP 兼 CXO インレジデンス責任者



©2026 Zscaler, Inc.All rights reserved.

数字で見る

100% で、AI に対するアイデンティティーベースの認証が導入されていません。

91% は、外部からの攻撃に対して完全に無防備な状態になっています。

100 点満点中 37 点が、フロンティア AI 対応状況スコアの平均的な開始時点の値です。最低値は 17 点、最高値は 60 点でした。

SSL インスペクションの率は **0.4 ～ 68.1% の範囲で、平均約 36%** でした。これは、インライン セキュリティ制御 (DLP、IPS、サンドボックス、AI エグレス) によって許容可能なセキュリティ カバレッジを確保するために必要と考えられる 70% の検査水準を大きく下回っています。

483,250 個のアプリケーションにアクセス可能でした。

代表的なワイルドカード固有の事例では、195 個のワイルドカード ドメインだけで **108,033** 個を超えるアプリケーションが公開されていました (ワイルドカードとは、「*」や先頭の「.」などの記号をポリシーで使用し、複数のドメインやパスを一括して表すものです。管理を効率化できる一方、適切に制御されていない場合には、攻撃対象領域を大幅に拡大させます)。

10 社中 9 社で、VPN アプライアンスが公開されていました。これらはそれぞれ、認証前の攻撃対象領域となっており、フロンティア AI モデルが高速で探索できる可能性があります。

750 以上の資産で、インターネットに接続されたインフラ上に CISA の既知の悪用された脆弱性が確認されました。これは積極的な調査によって確認されたもので、評価時点で実際の攻撃に使用されていました。

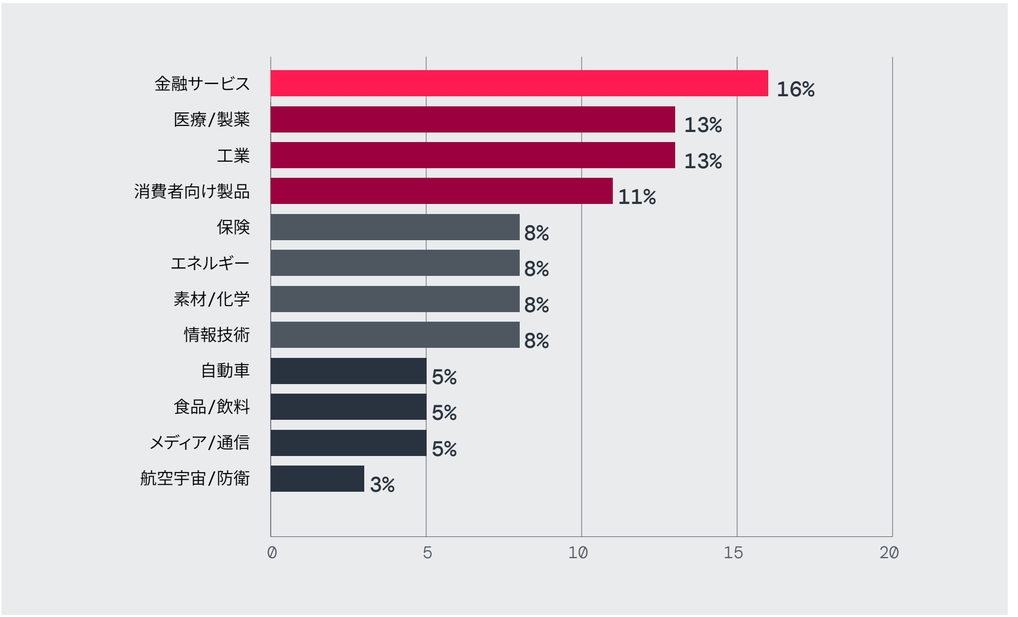


©2026 Zscaler, Inc. All rights reserved.



評価対象
グループ

このレポートの調査結果は、2026 年上半期に数百の組織を対象として実施したフロンティア AI 対応状況評価に基づいています。そのうち 32% は Fortune 500 企業です。調査対象グループは 12 業界にわたり、その構成は以下のとおりです。



評価対象となった組織は、主に大企業であり、インターネット上から発見可能な資産を数万単位で保有していました。インターネット上から発見可能な資産とは、認証なしでインターネットからアクセス可能なシステムやアプリケーション、エンドポイントなどを指します。

©2026 Zscaler, Inc.All rights reserved.



脅威の見通し

フロンティア AI モデルは高速で組織の脆弱性を特定し、それらを連鎖させて攻撃する能力を実証している

前述の調査結果は現状を示しています。つまり、2026 年上半期に調査対象グループで確認された状況です。ここから取り上げる脅威の見通しでは、なぜ対策を有効化するまでの時間軸が重要なのかを説明します。こうした状況を悪用する攻撃者の能力は成熟しており、一般的な組織の運用サイクルよりも速いペースで進化しています。以下の各トレンドは、評価で明らかになった 1 つ以上の調査結果を直接強調するものです。

Zscaler のフロンティア モデル テスト プログラムでは、フロンティア AI クラスのモデルを 3 つのモードで評価しました。内部情報を持たないブラックボックス、ソース コードやコンパイル済みバイナリー、静的ファイルを使用するホワイトボックス、そしてシステムに関する一部または完全なコンテキストを使用するグレーボックスです。これらのテストにより、実際の環境に即したコンテキストが与えられた場合、フロンティア AI モデルは高いシグナル対ノイズ比で組織環境を対象に推論し、人間のアナリストや従来のスキャナーでは見逃された複数段階の攻撃チェーンを一貫して特定できることが独立して確認されました。実際の本番アーキテクチャーを対象としたテストでは、実際の環境に即したコンテキストが攻撃経路の発見を大幅に加速させることが実証されました。

この独立した検証により、調査結果 1 ～ 6 で特定された設定上のギャップが、自律的に動作するフロンティア AI モデルによって悪用可能であること、また、現在のフロンティア モデルが評価対象グループで確認されたこれらの具体的な状況を悪用できることが確認されました。

AI を活用した脆弱性の発見により、2027 年にかけて新たなエクスプロイトが利用可能になるまでの期間が短縮される可能性が非常に高い

エクスプロイトまでの平均時間はマイナスに転じています。現在では、自動パッチ適用が導入されている場合でも、組織環境全体にパッチを適用する前に脆弱性が悪用されています。これは構造的な問題です。脆弱性が悪用されるまでの時間が、組織の変更管理が適応できる速度を上回るペースで短縮されているためです。

評価対象グループ全体で確認されたとおり、KEV が存在し、EPSS スコアが 1.0 に近い、または 1.0 に達しているインターネットに接続された資産では、脆弱性の公開から実際に悪用されるまでの時間が数時間にまで短縮されています。標準的なメンテナンス時間帯にパッチを適用する組織では、修正対応のサイクルが完了する前に既知の脆弱性が悪用される事態にほぼ確実に直面することになります。

©2026 Zscaler, Inc.All rights reserved.

AI 攻撃対象領域の拡大は、今後 12 ～ 18 か月にわたりガバナンス プログラムを上回るペースで進む可能性が高い

AI ガバナンスが普遍的に欠如している状況は、AI の導入速度とセキュリティ プログラムの適応との間に構造的な遅れがあることを示しています。開発部門、SaaS ベンダー、事業部門は、セキュリティ レビューを行わずにコパイロット、推論エンドポイント、MCP インターフェイスを導入しています。確認された普及率に基づくと、ガバナンスのない AI 攻撃対象領域は、ガバナンスが運用上のカバレッジを達成する前に、2 倍に拡大する可能性があります。対応を遅らせる組織は、評価期間中に悪用される可能性が高くなります。

アイデンティティーの侵害を介したラテラル ムーブメントは、ゼロトラスト環境に対する悪用後の攻撃手法として最も一般的な手法になる可能性が非常に高い

ワイルドカードの調査結果は、アイデンティティーの侵害が最も容易な侵入経路であることを示しています。フロンティア AI クラスの攻撃者は、従来のネットワーク悪用よりも認証情報の窃取を優先する可能性が非常に高いです。その見返りは桁違いに大きく、1 つのアイデンティティーから数千のアプリケーションにアクセスできます。セッションが期限切れにならず、デバイスの状態の検証も行っていない組織が、ほぼ確実に最初の標的となります。

暗号化トラフィックの死角が、AI を悪用したデータ窃取における主要な流出チャネルになる可能性が高い

フロンティア AI は、暗号化された API 呼び出し、モデルとのやり取り、クラウド トラフィックに機密データを埋め込むことができます。正規に見えるパターンを生成する AI ツールと、検査できないセキュリティ スタックが組み合わさることで、大容量かつ不可視された流出チャネルが生まれます。これはすでに悪用されている可能性が高く、今後ほぼ確実に深刻化するでしょう。

自律型エクスプロイト チェーンにより、初期アクセスから完全な侵害に至るまでの時間は数分に短縮される可能性が高い

露出した VPN アプライアンス、KEV 脆弱性、ワイルドカード セグメンテーション、休眠状態の検知機能が重なることで、AI を悪用する攻撃者が何の対応も引き起こすことなく、偵察から内部ネットワークへの広範なアクセスに至る可能性が高い環境が生まれています。このデータは、この一連の攻撃が現在すでに評価対象グループの大半に対して実行可能であることを示唆しています。エンドツーエンドで実行する自律型ツールは、今後 12 か月以内に高度な攻撃者が利用できるようになる可能性が高いでしょう。



©2026 Zscaler, Inc.All rights reserved.



調査方法

フロンティア AI の脅威検証

OpenAI GPT-5.5 と Anthropic Claude Opus 4.8 を、制御されたハーネス環境下で組織環境に即したコンテキストを付与して活用しました。「フロンティア AI クラス」という表現を用いるのは、特定のモデルよりも能力の閾値が重要であるためです。その閾値とは、自律的な多段階推論、ツールの使用、コンテキストに基づく情報検索、そして自律的なエクスプロイトの手順策定を指します。現在または将来のモデルがこの閾値以上の能力を備えていれば、いずれもこの評価の対象とする攻撃者クラスに該当します。

これと並行して、Zscaler は 3 つのモードとしてブラックボックス (内部情報なし)、ホワイトボックス (ソース コード、コンパイル済みバイナリー、静的ファイル)、グレーボックス (システム コンテキストの一部または全部) を用いて、フロンティア モデルの社内テストを実施しました。これらのテストにより、フロンティア AI モデルは、組織環境に即したコンテキストを与えられた場合、組織環境に対して高いシグナル対ノイズ比で推論できることが確認されました。

スコアリング フレームワーク

各組織について、フロンティア AI 脅威モデルに基づき、5 つの防御領域を対象とした加重 100 点満点の総合スコアで評価しました。その内訳は、攻撃からの不可視化 (25 点、25%)、影響範囲の制限 (25 点、25%)、AI 使用の管理 (20 点、20%)、データの保護 (20 点、20%)、デセプションとデコイ (10 点、10%) です。スコアリングでは、制御の所有権と運用の有効化状況の両方を評価しました。

データ ソース

評価には、外部アタック サーフェス管理 (EASM)、AI を活用した脆弱性の優先順位付け、AI 固有の攻撃対象領域の検出、プライベート アクセス設定のレビュー、さらに可能な場合には検査対象範囲やポリシーの施行、認証態勢を網羅するインライン テレメトリーなど、複数のカテゴリーのインプットを用いました。

評価モデル

包括的な EASM の可視性を基盤として、Opus エンジンで調査結果を検証および優先順位付けし、対応が必要な、悪用可能であることが確認された問題を特定しました。候補を平均 84%、最も複雑なケースでは 92% 絞り込みました。また、修復 ROI に基づいて CVE を再スコアリングし、CISA KEV と調査結果を照合するとともに、検証済みの攻撃対象領域全体にわたる多段階のエクスプロイト経路を特定しました。

この評価で対象となる内容 (と対象外の内容)

この評価は、外部攻撃対象領域の露出状況とセキュリティ制御の構成状態を評価します。SIEM のルールレベルにおける SOC の検知有効性、インシデント対応の準備状況、アイデンティティー プロバイダーのセキュリティ態勢、エンドポイント検知のカバレッジは評価しません。このスコアは、AI を悪用する攻撃者が有効な制御に遭遇することなく侵入、水平移動、流出を行える可能性を示すものであり、総合的なセキュリティ態勢を示すスコアではありません。

フロンティア AI クラスの攻撃は、偵察から情報流出までのキル チェーンを数分にまで短縮でき、成熟した SOC の対応閾値を下回り、人間による介入が行われる前に構成上の隙を突くことが可能である点が極めて重要です。

©2026 Zscaler, Inc.All rights reserved.

主な
調査結果

調査結果 1

組織はいくつかの簡単な手順で既存の AI 防御機能を最大限に活用できる

組織は、実践的なサイバー管理によって防御力を大幅に強化できます。

有効化の機会は広範に及びます。64% の組織では、ライセンス済みのディセプション プラットフォームが高精度の検知深度を追加するために展開可能な状態にあります。15% の組織では、侵入防止システムがモニタリング モードから完全な施行モードに移行できる準備が整っています。18% は、認証機能を追加のサイトに拡張可能です。構成の最適化だけでも、組織は 100 点中 73 ～ 98 点の目標スコアに到達できます。評価対象グループの平均開始スコアが 37 点であったことを考えると、これは個々のスコアが大幅に改善したことを意味します。

AI クラスの脅威に対してこれが重要である理由は、完全な施行モード以下で動作するセキュリティ ツールは、高速で行動する攻撃者に対して抵抗力が低下するためです。

調査結果 2

評価対象となったすべての組織において、その規模や制御範囲にかかわらず AI ガバナンスはまだ 0 年目の段階にある

この結果は普遍的です。評価対象となった組織の 100% が、自社の AI に対するアイデンティティーベース認証を備えておらず、91% が AI エンドポイントを外部に公開しています。AI ガバナンスの柱における平均スコアは 20 点満点中 2.1 点 (最大値の 10.5%) でした。評価対象グループの最高スコアは 20 点満点中 3 点でした。

この露出は重大です。インターネットに公開された MCP エンドポイントや、認証されていないチャットボットまたはコパイロットなど、公開されている AI エンドポイントは評価対象グループ全体で一般的に見られました。確認された最も深刻な事例には、極めて重大性の高い生成 AI ゲートウェイがインターネットにさらされているケースが含まれていました。

このギャップが生じている理由は、AI 接点が開発部門、事業部門、サードパーティー ベンダーによって、セキュリティ プログラムの適応速度を上回るペースで展開されているためです。AI ガバナンスは将来の取り組みではありません。評価対象となったすべての組織において現在進行形で悪用可能なギャップであり、管理されていない攻撃対象領域の中で最も急速に拡大しているカテゴリーです。

©2025 Zscaler, Inc. All rights reserved.

調査結果 3

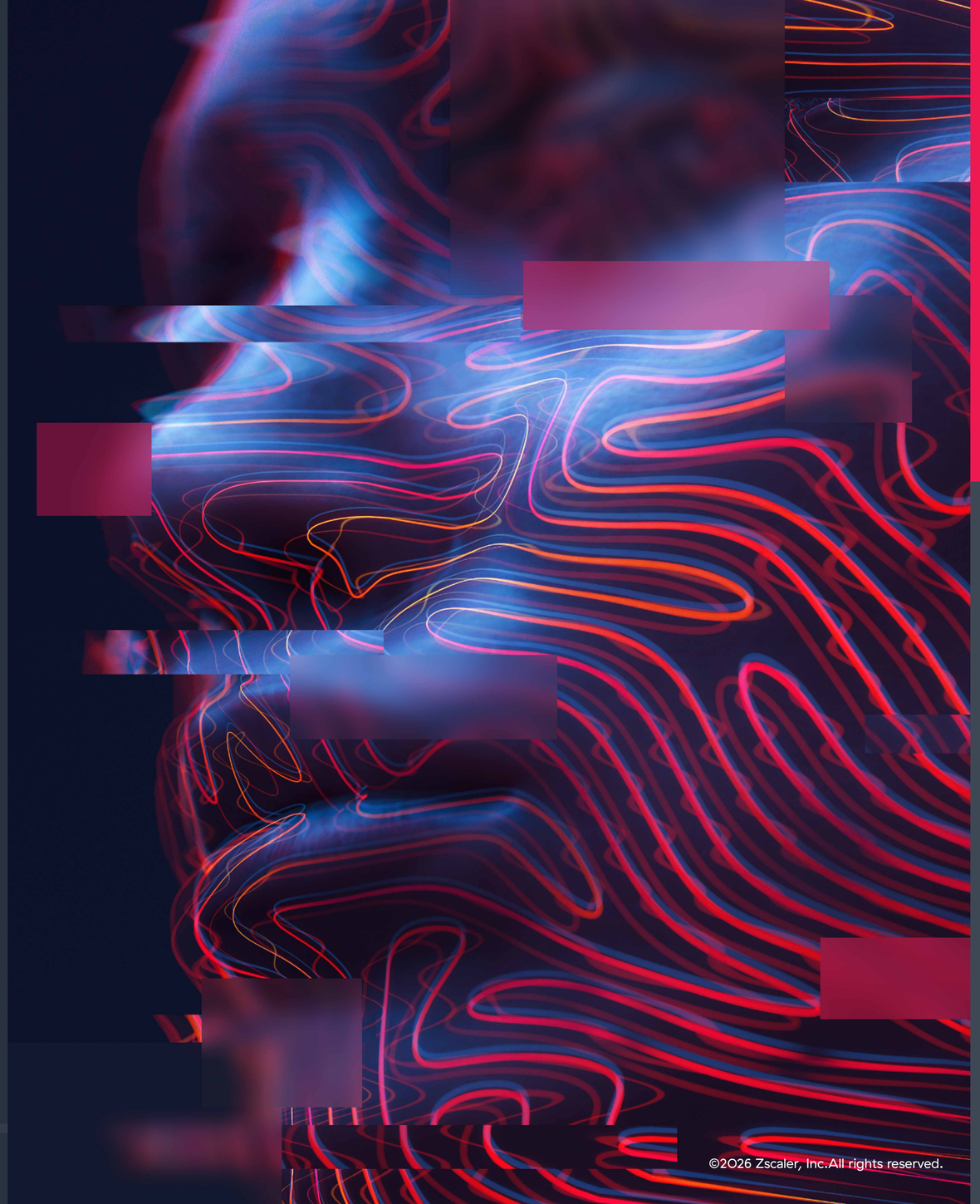
暗号化トラフィックの検査が限定的である場合、トラフィックの大半に対して導入済みのセキュリティ制御の有効性が低下する

暗号化トラフィックの検査は長年にわたりセキュリティ運用の優先事項とされてきました。AI を悪用した脅威環境では、包括的な検査範囲が極めて重要になります。流出を行うフロンティア AI モデルは、ほぼ確実に暗号化されたチャネルを悪用し、とりわけ AI の API 呼び出し、クラウド ストレージへのアップロード、正規のトラフィック パターンに紛れるモデル間通信が利用されます。検査されていないトラフィック内で活動する攻撃者は、インライン制御から見えません。このことから、暗号化トラフィックの検査は運用上のベストプラクティスから、AI クラスの脅威に対抗するための基盤的要件へと格上げされます。

平均的な組織が検査している暗号化トラフィックは約 36% にとどまります。この評価で定義された適格な検査閾値である 70%、すなわち、暗号化トラフィックのうち、アクティブなインライン セキュリティ制御 (DLP、IPS、サンドボックス、AI エグレス) を通過し、それらの制御が最適なセキュリティ カバレッジを達成できる割合を満たす組織は 1 社もありません。最悪の場合、暗号化トラフィック全体の 99.6% が、セキュリティ インフラを通過しながら完全に未検査のままです。

暗号化トラフィックが未検査のまま残されると、DLP、IPS、サンドボックス、AI エグレスといったインライン セキュリティ制御は機能不全となり、組織はデータ流出やマルウェア配信といった重大な脅威を把握できなくなります。サイバー脅威のうち 86% を超えており、AI モデルとのすべてのやり取りが暗号化チャネルを使用していることを踏まえると、包括的なトラフィック検査は AI クラスの攻撃に対抗するための不可欠な前提条件です。

根本原因は技術ではなく運用です。大多数の組織はオプトインの検査モデル (初期設定が「検査しない」) を採用しています。調査対象グループ全体で、組織はしばしば「他のポリシーを回避する」や「検査しない」といった SSL ルールに依存しており、これらは特定のトラフィック カテゴリー全体に対して検査を無効化します。影響を受けた組織では、こうした回避ルールの数は 2 〜 51 件に及んでいました。



©2026 Zscaler, Inc.All rights reserved.

調査結果 4

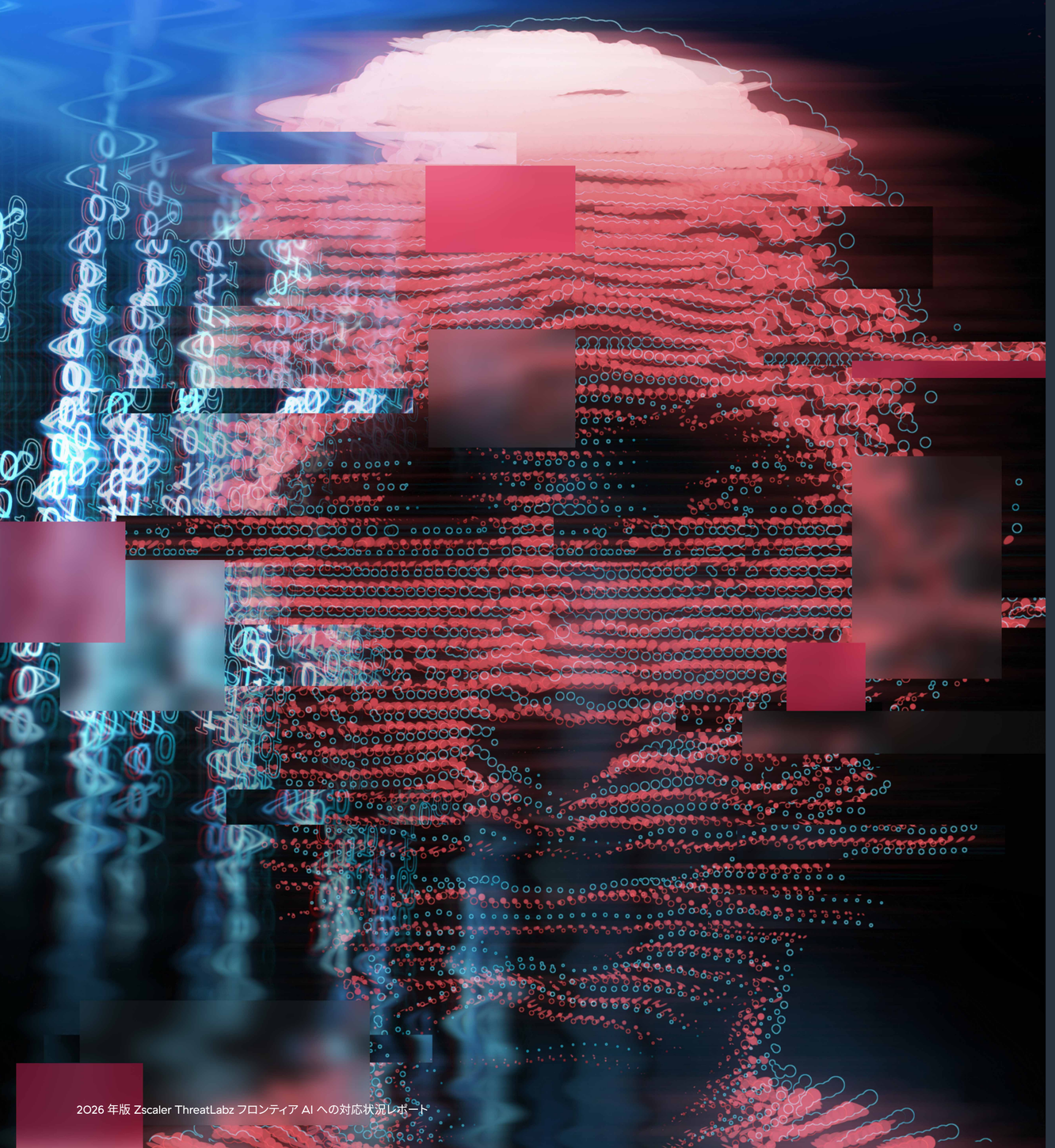
評価対象の環境は、AI を悪用した攻撃を必要な速度で検知および防御できるようまだ構成されていない

調査結果 4 は、重大な検知ギャップを浮き彫りにしています。これは、デセプション、IPS、異常検知、セッション監視にわたる有効化済みのセキュリティ制御が、AI クラスの攻撃に対応するよう構成を最適化されていないために機能していないというものです。その結果、組織は阻止できるだけの対策を備えているにもかかわらず、脅威を特定できず、脅威にさらされています。したがって、フロンティア AI クラスの攻撃は SOC から見えないままとなり、デセプション、IPS、振る舞いの異常、セッション タイムアウトのいずれについてもトリガーが発生しません。そのため、侵害の最初の兆候は、内部での検知ではなく、外部からの通知となる可能性が高くなります。

- **デセプション (早期警戒システム):** 組織の 64% では、デセプション プラットフォームのカバレッジが重要なセグメントにまで拡張されていません。攻撃者がラテラル ムーブメントを行ったとしても、現在の構成ではトリップ ワイヤーに遭遇しません。デセプションは、高精度のデコイを使用してラテラル ムーブメントを瞬時に検知し、自動的に封じ込めることで、AI を悪用した攻撃に対してシグネチャーに依存しない効果的な検知を実現します。
- **侵入防止 (インラインでのブロック):** 組織の 15% では、侵入防止システムを完全な施行モードに移行することで、防御を強化できる明確な機会があります。有効化されている環境でも、「他のポリシーを回避する」ルールにより、ほとんどのトラフィックが評価されていませんでした。暗号化されたチャネルを通過する既知のエクスプロイトは、検査もブロックもされないまま通過してしまいます。
- **アプリケーションレイヤーの異常検知 (挙動シグナル):** 複数の組織では、インライン ルールがまだ設定されていませんでした。ワイルドカード トラバーサルの通信量に対する異常検知も行われていませんでした。1 つのアイデンティティーが数千ものアプリケーションに順番にアクセスしたとしても、アラートが生成される可能性は非常に低い状況でした。
- **セッションとアイデンティティーのモニタリング (封じ込めのトリガー):** 複数の組織では、有効期限を設定せずに認証タイムアウトを構成していたため、侵害されたセッションが無期限に有効な状態で残る可能性があります。ある組織では、18 件のアクセス ルールのいずれにもデバイスの状態チェックが設定されていませんでした。盗まれた認証情報が、再認証を求められることなく数週間にわたって悪用される可能性がありました。
- **フロンティア AI クラスの攻撃の SOC からの見え方：**SOC は何も検知しません。デセプションのトリガーも、IPS アラートも、挙動異常の検知も、セッション タイムアウトもありません。侵害を示す最初の兆候は、内部での検知ではなく、外部からの通知となる可能性が非常に高いでしょう。



©2026 Zscaler, Inc.All rights reserved.



調査結果 5

広範なセグメンテーションにより、プライベート アクセス プラットフォームがラテラル ムーブメントの高速道路と化している

今回の分析では、ZTNA を利用している組織の 64% で、意図した以上のアプリケーションを公開する可能性のある、ワイルドカードを多用した構成が確認されました。あるケースでは、わずか 195 個のワイルドカード ドメインによって 10 万 8,000 を超えるアプリケーションが到達可能となっており、1 つのアイデンティティーから環境の非常に広範な部分にアクセスできる可能性がありました。こうした構成は、オンボーディングを簡素化するために導入時や移行時に設定されることが多くあります。しかし、環境の拡大に伴ってそのまま残り、当初の意図を超えてアクセス範囲を広げてしまう可能性があります。定期的なレビューと見直しを行わなければ、ワイルドカードベースのポリシーによって最小特権の目標が損なわれ、フィッシングやセッションの窃取などによってアイデンティティーが侵害された場合の影響が拡大する可能性があります。

複数の組織では、認証タイムアウトを「無期限」に設定することで、この問題をさらに悪化させています。侵害されたセッションは無期限に有効な状態で残ります。影響範囲の制限で最高スコア (25 点満点中 14.4 点) を獲得した組織でさえ、ワイルドカード構成が維持されていました。スコアがより低かった組織では、実質的にセグメンテーションがまったく行われていませんでした。

このパターンは、プライベート アクセス環境には、ゼロトラスト アーキテクチャーが施行するために設計された最小特権セグメンテーションがまだ反映されていないことを意味します。大規模なアイデンティティー侵害が可能なフロンティア AI クラスの攻撃者を相手にする場合、ワイルドカードを中心とした構成は、評価対象となった環境においてラテラル ムーブメントにつながる最も大きな攻撃経路となります。

©2026 Zscaler, Inc.All rights reserved.

調査結果 6

3 分の 1 を超える組織で、インターネットに公開されたインフラに既知の悪用された脆弱性が残されている

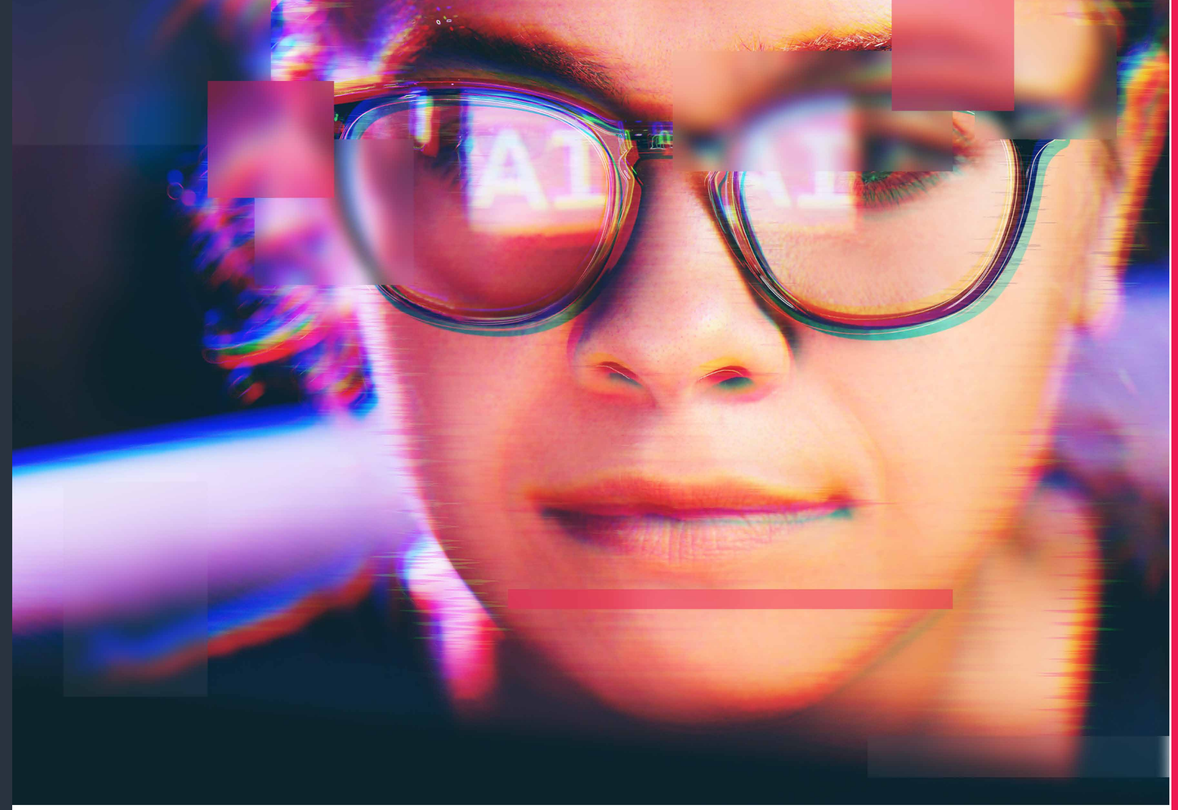
こうした脆弱性が残り続けるのは、必ずしも脆弱性が把握されていないからではありません。メンテナンス期間、変更管理要件、インフラ間の依存関係など、現実の運用上の制約と修復を両立させる必要があるためです。調査対象グループの半数以上 (53%) が、インターネットに公開された資産に存在する CISA の既知の悪用された脆弱性 (KEV) について、修復の推奨を受けていました。積極的に調査されている間、3 分の 1 を超える組織で、こうした実際に悪用可能な KEV が境界上に残されていました。

その結果、防御側が修復できるスピードと攻撃者が脆弱性を特定して悪用できるスピードとの間にますます乖離が生じています。攻撃者が AI を悪用して脆弱性の発見と悪用を加速させるなかで、この脆弱性が続く時間は、今後さらに短くなる可能性が高いでしょう。

3 つの CVE が、調査対象グループの大半で確認されました。いずれも CISA の既知の悪用された脆弱性カタログに掲載されています。すべてアクティブなプロービングによって確認されました。

- **CVE-2024-38475 (Apache mod_rewrite):** EPSS スコアは 0.9988 ～ 0.9997
- **CVE-2021-40438 (Apache mod_proxy SSRF):** EPSS スコアは 0.9999 ～ 1.0000
- **CVE-2023-44487 (HTTP/2 Rapid Reset):** EPSS スコアは 0.999

高度な組織でも、これらの脆弱性が残り続けていることは、運用上の現実を反映しています。Web 層のインフラ、特に Apache や F5/NGINX は、標準的なパッチ自動化の対象外となることが多く、変更期間に合わせた手動での調整が必要です。



現在の露出は既知の脆弱性によるものですが、境界上に既知の KEV が残っている組織は、既存の修復が完了する前に、同じシステムに存在するまだ公開されていない脆弱性が悪用されるリスクにも直面しています。

調査結果 5 は、このリスクをさらに増加させます。インターネットに公開されたパッチ未適用の各資産は、認証前の侵入口となります。また、1 つのアイデンティティーから数十万ものアプリケーションにアクセスできる環境では、1 件の CVE が悪用されるだけで、影響を受けたサーバーをはるかに超えてアクセス範囲が拡大する可能性があります。

優先すべき対策は、外部アタック サーフェス管理によって資産の可視性を継続的に維持すること、仲介型アクセスによって発見可能な攻撃対象領域を削減すること、悪用可能性に基づく優先順位付けによって業務上の正当な理由から公開されている資産について検知、トリアージ、パッチ適用のサイクルを短縮すること、そして修復期間中は自律型の仮想パッチを利用することです。

©2026 Zscaler, Inc.All rights reserved.

新たな攻撃チェーン：フロンティア AI が小さなギャップを自律的につなぎ合わせ一連の攻撃を成立させる仕組み

これらの調査結果は、個別に存在しているわけではありません。より広範な攻撃チェーンを構成する要素が相互に作用しています。それぞれ単独では対応可能なギャップですが、組み合わさることで、どの 1 つの調査結果から想定されるものよりも、はるかに大きなエクスポージャーを生み出します。

コンテキストに基づくフロンティア AI モデルは、パッチ未適用の CVE、まだ施行されていない IPS、ワイルドカード構成、管理されていない AI エンドポイント、未検査のエグレス経路といった条件を自律的に特定して組み合わせ、人間が関与することなく高速で継続的な悪用経路を構築できます。従来のスキャナーやペネトレーション テストでは、このように複合的な攻撃経路を同等のスピードで明らかにすることはできません。次の一連の経路が、評価対象となった環境全体で検証されました。

ステップ 1: インターネット スキャンによって公開されている VPN アプライアンス (500 台以上のうちの 1 台) を特定します。Web 層で CVE-2024-38475 (Apache mod_rewrite、EPSS 0.9997) を特定します。

ステップ 2: 認証前の脆弱性を悪用します。認証情報なしで初期アクセスを取得します。IPS は無効になっており、アラートは生成されません。

ステップ 3: 侵害された VPN インフラから認証情報を収集します。1 つの有効な組織のアイデンティティーを取得します。

ステップ 4: プライベート アクセス プラットフォームに認証します。ワイルドカード構成により、17,000 以上のアプリケーションへのアクセスが許可されます。タイムアウトは「無期限」に設定されています。デバイス状態のチェックはありません。

ステップ 5: アプリケーションを高速で横断します。経路上にデセプション デコイはなく、SOC にはアラートが表示されません。

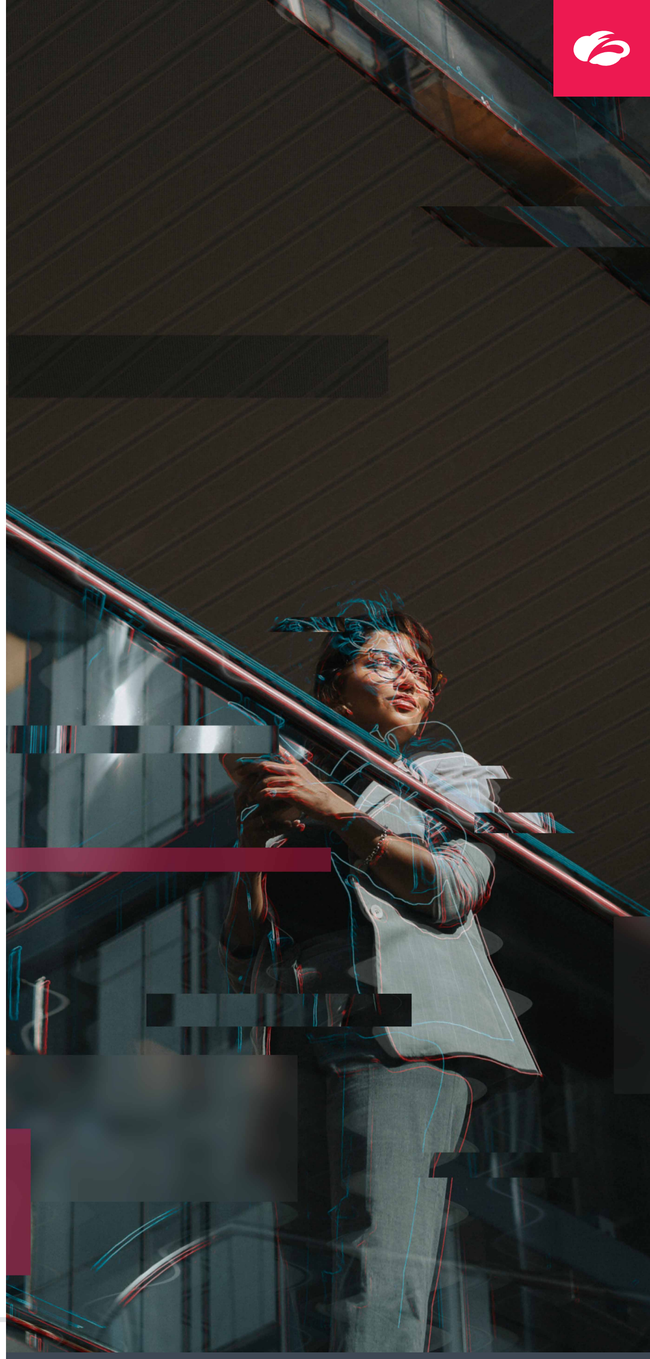
ステップ 6: 管理されていない AI 推論エンドポイント (300 以上のうちの 1 つ) に到達します。追加の認証なしでアクセスできます。モデルのコンテキスト ウィンドウを介して独自データを照会します。

ステップ 7: 暗号化されたチャネルを介してデータを持ち出します。トラフィックの 3 分の 2 は検査されていません。DLP ではエグレスを可視化できず、アラートは生成されません。

初期スキャンからデータ持ち出しまでの合計時間：内部での攻撃テストでは、組織の AI システムが最初の重大な障害に至るまでの時間の中央値は、わずか 16 分でした。

Zscaler のフロンティア モデル テスト プログラム (フロンティア AI クラスと GPT 5.5 を内部情報なし、ソースコードへのアクセス、システム コンテキストへの部分的 / 完全なアクセスという 3 つのモードで評価) では、この一連の推論能力が検証されました。コンテキストに基づくテスト環境により、フロンティア モデルは、従来のスキャナーや人間のアナリストが同等のスピードでは明らかにできない、多段階の悪用経路を特定できます。

すべての構成要素が、2026 年上半期に組織環境で確認されました。このチェーンは、フロンティア AI 対応状況スコアの評価基準となる攻撃モデルを示しています。



©2026 Zscaler, Inc.All rights reserved.

事例に基づく証拠：進行中の 3 社

以下の匿名化された事例研究では、修復のさまざまな段階にある組織を取り上げています。評価で特定されたギャップを解消するための、導入時の状態と実際の運用手法を紹介します。

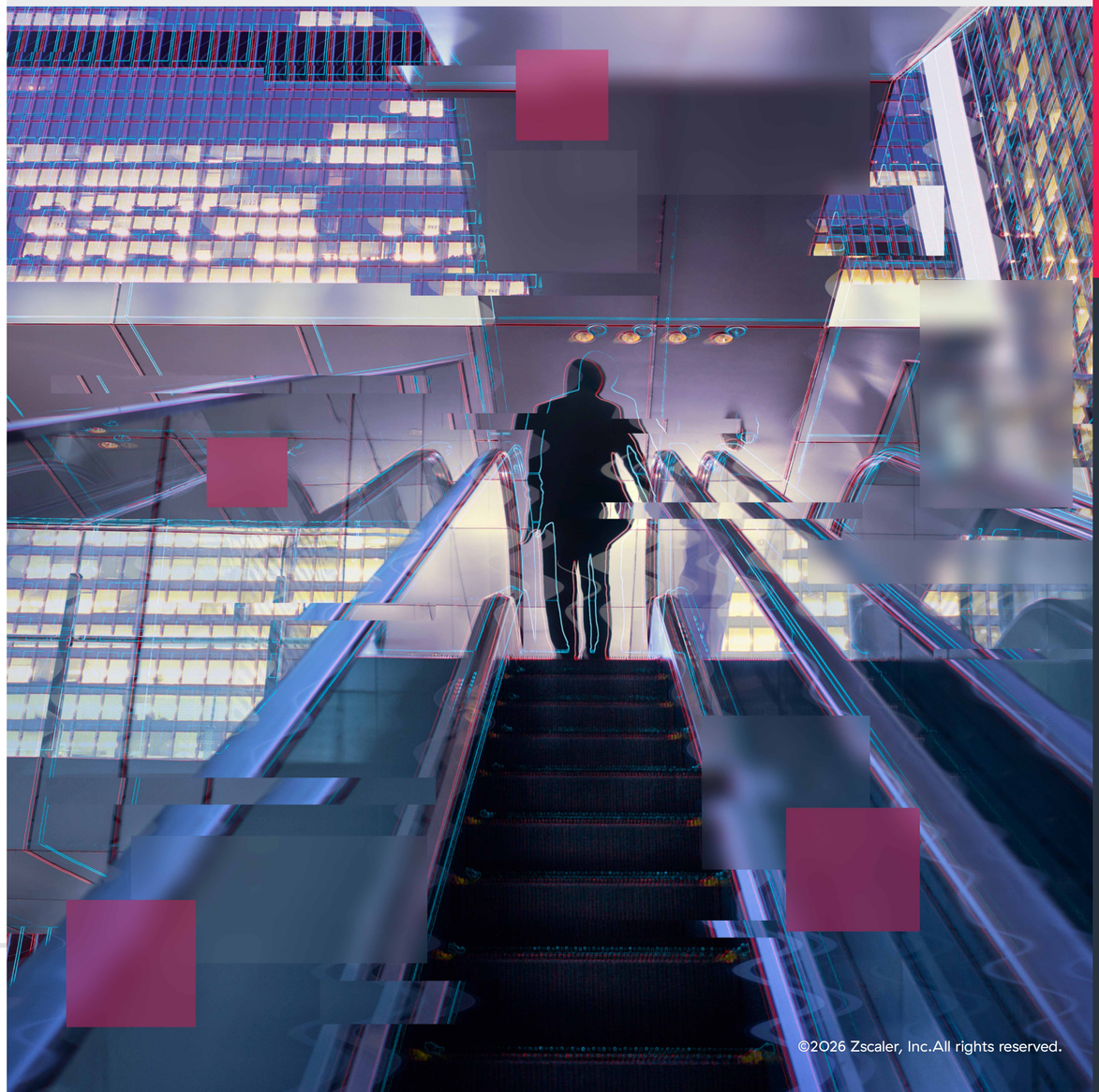
ケース 1

世界的な金融サービス企業

導入時の状態：インターネットに公開された資産が 10,000 件以上。検証済みの CVE が 4,000 件以上 (固有の CVE が 900 件以上、うち 165 件が重大)。調査委対象グループで最も多く確認された 3 つの CVE (CVE-2024-38475、CVE-2021-40438、CVE-2023-44487) が公開されたインフラ上で確認されました。

評価結果：この組織の広大なインターネット公開環境には、複数世代にわたるインフラの構築と更新に伴って蓄積された、パッチ未適用の重大な脆弱性が存在していました。SSL インスペクションも、トラフィックの大半をカバーするまでには至っていませんでした。デセプション機能は権限が付与されていたものの、導入には至っていませんでした。AI の利用環境も、ガバナンスの境界外にありました。

修復の開始：重大な CVE 165 件について、優先順位を付けてパッチを適用しました。SSL インスペクションをオプトインからオプトアウトに変更しました。重要な金融セグメントでデセプションを有効化しました。AI 接点に対するアクセス制御を開始しました。



©2026 Zscaler, Inc.All rights reserved.



ケース 2

世界的な食品メーカー

導入時の状態：VPN アプライアンス 13 台が公開されていました。ガバナンスの境界外にある公開された AI 環境が 40 件ありました (調査対象グループの規模に対する AI の露出密度はグループ内で最大)。公開されたインフラ上で、検証済みの CVE が 2 件確認されました。

評価結果：組織の急速なデジタル トランスフォーメーションに、セキュリティ プログラムによる新たな攻撃対象領域への対応が追いついていませんでした。AI エンドポイントは業務上の取り組みの一環として各事業部門によって構築されていましたが、セキュリティ アーキテクチャーはまだ経由していませんでした。プライベート アクセス環境のセグメンテーション ポリシーも確立されていませんでした。VPN インフラは、移行が完了するまでインターネットに公開された状態でした。

修復の開始：VPN をプライベート アクセスの背後に配置して保護しました (即時対応)。40 件のエンドポイントに対する AI 利用環境のアクセスを制御し、セグメンテーション ポリシーをゼロから策定するとともに、本番セグメントにデセプションを導入しました。

ケース 3

世界的な高級品企業グループ

導入時の状態：対象範囲は 487 件の公開対象名であり、1 か月あたり 922TB の暗号化トラフィックが検査されておらず、定義済みアプリケーションが 10,000 件ありました。

評価結果：同社の検査アーキテクチャーは、保守的なオプトイン モデルで導入されていたため、暗号化トラフィックの大半が検査されていませんでした。10,000 件のアプリケーションが存在する一方、セグメンテーションはほぼ行われておらず、環境の拡大にアクセス ガバナンスが追いついていませんでした。

修復の開始：SSL インスペクションへの移行を最優先で実施し、10,000 件のアプリケーション全体でワイルドカードを置き換え、デセプションを有効化し、AI ガバナンスのベースラインを確立しました。

©2026 Zscaler, Inc.All rights reserved.

実装の現実性：90 日間のスプリントで実際に必要となること

レジリエンスを実現する道筋は、技術的には 90 日以内に達成可能です。多くの場合、制御はすでに導入済みであり、依存関係も十分に把握されているためです。しかし、容易に実現できるわけではないため、経営層による指示が必要です。フロンティア AI への対応状況のギャップを解消することは、取締役会レベルでの可視性、CISO と CEO の認識共有、そして専任の実行体制から恩恵を受ける全社的な取り組みです。これは、進化する脅威環境において認識を共有する機会でもあります。セキュリティ リーダーは、以下のことを実行すべきです。

1. 現在の対応状況について取締役会に報告し、SEC、DORA、NIS2 の規制要件を踏まえて説明します。
2. 明確な責任者を定め、経営層への定期的な報告を行う 90 日間のスプリントについて、経営層からの支援を確保します。
3. 小規模な専任チームを編成して実装を進めます。検査対象範囲の拡大やワイルドカードの置き換えに伴うアクセス設定の見直しによって生じる摩擦は、導入プロセスにおいて想定されるものとして取り組みます。

SSL インスペクションには、部門横断的な連携が必要です。組織は、IT 部門とセキュリティ部門が連携して現在の検査対象範囲を完全に把握し、検査が必要なトラフィック ストリームを特定するとともに、検査が運用上不可能な領域 (証明書ピンニングを使用するアプリケーションや従来の統合など) を判断する必要があります。検査を適用できない場合は、ブラウザー分離、エンドポイント DLP、CASB ベースの API 検査、DNS レイヤーのフィルタリングなどの代替制御によって、インライン復号を必要とせずに補完的な可視性と脅威防御を実現できます。

ワイルドカードの置き換えには、アプリケーション所有者との連携が必要です。200 件を超えるワイルドカードを、指定した FQDN に置き換える作業は、セキュリティ部門だけで完結するものではありません。それぞれのワイルドカードは、事業部門、開発部門、ベンダーが所有するアプリケーションに解決されるため、アクセスを継続できることを各担当者が検証する必要があります。セキュリティ部門が目標とするアーキテクチャーを定義しますが、実行には部門横断的な連携が必要です。

このタイムラインでは、専任の実行体制を前提としています。専任のスプリント チーム (エンジニア 2 ～ 4 人) を割り当てた組織は、予定どおりに対応を進めることができました。既存の業務と並行して修復を進めた組織では、取り組みの勢いを維持するために追加の支援が有効でした。成功の鍵となったのは、経営層による定期的な確認、専任チームに確保された実行力、そして目標スコアに対する週次のマイルストーン報告など、経営層からの支援です。

大規模なセキュリティ更新によって一時的に運用上の変更が生じる可能性はありますが、こうした取り組みは、長期的によりレジリエンスの高い防御を構築するために不可欠です。今回の評価では、初期スコアが 37 点といったギャップに予防的に対応することが、現代のデジタル脅威を管理するうえで重要なステップであることが明らかになりました。こうした短期的な調整と、強固な保護の必要性とのバランスを取ることが、組織全体のセキュリティ態勢を強化するうえで重要です。



©2026 Zscaler, Inc.All rights reserved.

推奨事項

1 セキュリティ制御を施行モードに切り替える

評価対象となったすべての組織で、影響の大きいセキュリティ制御 (IPS、認証、悪意のある IP アドレスのブロック、ドメイン フロンティング対策) が、設定変更によって完全な施行モードに移行できる状態にありました。セキュリティ運用部門は、自社で保有するすべての機能について有効化状況を確認し、必要に応じてベンダーによる設定監査や有効化支援を活用したうえで、14 日間のスプリントで施行モードを有効にしてください。

すべての SOC 部門が懸念する点は理解しています。IPS をモニターからブロックに移行すると、誤検知が発生し、トラフィックが中断される恐れがあります。適切なアプローチは段階的な導入です。まずは重要度の低いセグメントで開始し、微調整と検証を行ったうえで、明確なロールバック基準を設定して本番環境に展開します。

2 AI 接点すべてに対して直ちにアイデンティティー認識型のアクセス制御を適用する

インターネットからアクセス可能なすべての AI エンドポイントは、組織のアイデンティティー認証の背後に配置する必要があります。この対策は、評価対象となったすべての組織に推奨されます。アイデンティティー認識型のアクセス制御を経路上に導入するまで、AI インフラはインターネットに接続できるあらゆる攻撃者によって悪用される可能性があります。

3 SSL/TLS インスペクションの対象範囲を 90 日以内に最低 70% の適切な範囲まで引き上げる

検査モデルをオプトインからオプトアウトに変更し、デフォルトですべてのトラフィックを検査したうえで、対象を限定した除外を設定します。「他のポリシーを回避する」ルールは排除します。検査しない例外を必要最小限まで減らします。アプリケーションの動作に影響が生じる期間として、2 ～ 4 週間を想定します。ヘルプデスクの問い合わせ増加に対応できる体制を確保します。この検査範囲がなければ、DLP、IPS、サンドボックス、AI エグレス制御は、トラフィックの大半に対して機能しません。

SSL インスペクションは普遍的に適用するのではなく、リスクベースで実施する必要があります。まず最も重要なアプリケーションとリスクの高いトラフィックから開始し、その後、対象を選別しながら範囲を拡大しましょう。証明書ピンニングや規制上の制約などにより例外が必要となるアプリケーションやサービスもあるため、組織はあらかじめ例外適用の経路と迅速な対応手順を定義しておく必要があります。検査が実行できない場合は、代替制御によって可視性のギャップを縮小します。

4 重要なセグメントのワイルドカード ドメインを指定した FQDN に置き換える

すべてのプライベート アクセス構成について、ワイルドカード エントリーを監査します。最重要セグメントである金融システム、IP リポジトリー、アイデンティティー インフラを優先しましょう。認証タイムアウトを設定し、すべてのアクセス ルールでデバイスの状態の検証を実施するとともに、各アイデンティティーを必要なアプリケーションのみに限定するアプリケーションレベルのセグメンテーションを実装します。アプリケーション所有者との部門横断的な連携を計画してください。



©2026 Zscaler, Inc.All rights reserved.

5 重要なネットワーク セグメントにデセプションを導入する

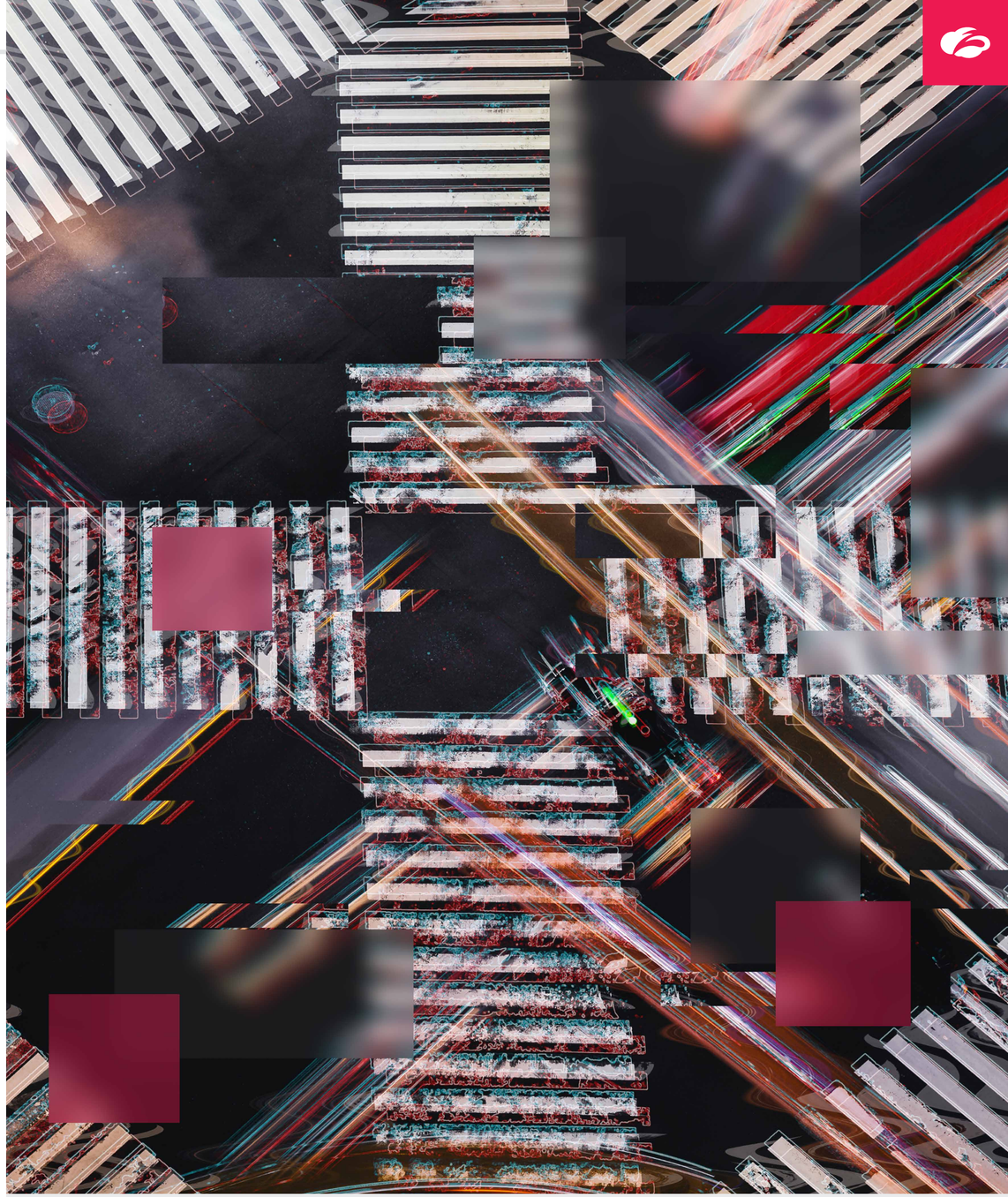
組織は最重要アプリケーションの近くにデコイを配置することで、防御を強化できます。これにより、高精度のアラートを生成する独立した検知レイヤーを構築し、SOC 部門がより迅速に対応できるようになります。デコイがトリガーされると、デセプションはアナリストの承認を必要とすることなく ZPA を通じて、サインイン中のすべてのデバイスで侵害されたユーザーのプライベート アプリケーションへのアクセスを自動的に取り消します。これにより、攻撃者が別のシステムに侵入を拡大する前にアクセスを遮断し、SOC の対応キューの速度ではなく攻撃の速度でラテラル ムーブメントを阻止できます。

6 インターネットに公開された VPN アプライアンスをゼロトラストの仲介型アクセスの背後に移行する

発見可能なすべての VPN アプライアンスは、認証前の攻撃対象領域となります。30 日以内に、すべての VPN インフラを仲介型アクセスの背後に配置します。90 日以内に、公開されたアプライアンスを廃止します。これにより、フロンティア AI モデルが最初に調査する可能性のある侵入口を排除します。

7 インターネットに公開された資産上の CISA-KEV 脆弱性に対して評価から 72 時間以内にパッチを適用する

CVE-2024-38475、CVE-2021-40438、CVE-2023-44487 は、EPSS スコアが 1.0 またはそれに近い値であり、実際に悪用されていることが確認されています。直ちにパッチを適用するか、仮想パッチを適用してください。メンテナンス期間が必要な場合は、その間、エクスプロイトの試みをブロックするために自律型の仮想パッチを導入します。



©2026 Zscaler, Inc.All rights reserved.

規制上の
リスクに関する注記

スコアが 40 点未満の組織では、重大なサイバー インシデントの発生可能性と潜在的な影響の双方を高める制御上の弱点を抱えている可能性があります。組織には、リスクへの露出軽減とレジリエンスの強化に向けてこのレポートの推奨を確認し、対応することをお勧めします。場合によっては、こうした弱点に関連するインシデントが、SEC Item 1.05 (Form 8-K) に基づく開示に関する検討事項を生じさせることもあります。

DORA または NIS2 の対象となる組織において、基本的な制御が不足していることは、ICT リスク管理と全体的な運用上のレジリエンスを強化する余地があることを示している可能性もあります。

次の
ステップ

フロンティア AI 対応状況評価では、各組織に対し、定量化した開始時スコア、優先順位を付けた修復手順、そして主に既存の機能を拡張または微調整することで達成可能な 90 日間の目標を提示します。このレポートのデータは、業界のベースラインを示すものです。すべてのセキュリティ リーダーに問われているのは、自社がこのベースラインに対してどの位置にあるのかということです。

調査結果をご覧いただきました。次は、自社のリスクをご確認ください。

Zscaler の専門家にご相談いただき、お客様の環境にどのような AI リスクが存在する可能性があるのか、そして何から対応すべきなのかをご確認ください。

リスクを把握し、迅速に行動しながら、安全を確保しましょう。
詳細はこちらをご覧ください：zscaler.com/jp/campaign/zscaler-ai-security-demo

ThreatLabz
について

ThreatLabz は、Zscaler が誇る世界トップクラスのセキュリティ調査部門であり、Zscaler のプラットフォームを使用する世界中の組織が常に保護された状態にあることを保証する責任を担います。ThreatLabz のメンバーは、マルウェアの調査や振る舞い分析に加え、Zscaler のプラットフォームの高度な脅威対策を実現するための新しいプロトタイプ モジュールの研究開発も進めています。また、定期的に社内のセキュリティ監査を実施して、Zscaler の製品とインフラがセキュリティ コンプライアンス基準を満たしていることを確認します。ThreatLabz は、新たな脅威に関する詳細な分析を定期的にポータル (research.zscaler.com) で公開しています。

フォローはこちら：X @ThreatLabz | ThreatLabz セキュリティ リサーチ ブログ

©2026 Zscaler, Inc.All rights reserved.

zscaler™ | Act Fast. Stay Secure.

Zscaler について

Zscaler (NASDAQ: ZS) は、より効率的で、俊敏性や回復性に優れたセキュアなデジタル トランスフォーメーションを加速しています。Zscaler Zero Trust Exchange™ プラットフォームは、ユーザー、デバイス、アプリケーションをどこからでも安全に接続させることで、数多くのお客様をサイバー攻撃や情報漏洩から保護しています。世界 150 拠点以上のデータ センターに分散された SSE ベースの Zero Trust Exchange™ は、世界最大のインライン型クラウド セキュリティ プラットフォームです。詳細は、**zscaler.com/jp** をご覧いただくか、Twitter で **@zscaler** をフォローしてください。

© 2026 Zscaler, Inc. All rights reserved. Zscaler™ および **zscaler.com/jp/legal/trademarks** に記載されたその他の商標は、米国および/または各国の Zscaler, Inc. における (i) 登録商標またはサービス マーク、または (ii) 商標またはサービス マークです。その他の商標はすべて、それぞれの所有者に帰属します。

+1 408.533.0288 **Zscaler, Inc. (HQ) • 120 Holger Way • San Jose, CA 95134** **zscaler.com/jp**