

ホワイトペーパー

侵害を前提としたサイバー防衛の設計

侵入の阻止から、影響範囲の限定と封じ込め速度へ

2026年10月19日

GenerativeX

目次

要旨	3
序章 午前2時47分	4
緊急提言	5
第1章 攻撃は速くなり、修正は遅れている	7
第2章 2026年秋、国内で起きたこと	12
第3章 漏えいは外部連携から起きる	15
第4章 GenerativeXの考え方	17
第5章 4つの柱	19
第6章 AIエージェント	21
第7章 経営が決めること	22
第8章 推奨する実行計画	24
結び	25
付録A 実装の詳細	26
付録B 本稿の前提と限界	30
付録C 既存の枠組みとの関係	31
付録D 図のデータと集計方法	32
付録E 出典	35

要旨

経営は、侵入されることを前提に、被害の大きさを左右する3つの要素、すなわち入口の数、1つの侵害で到達できる範囲、停止までの時間を小さくする判断を自ら下すべきである。

2026年9月末から、国内で不正アクセスによる漏えいの公表が相次いでいる。GenerativeXが集めた公表は10月の2週間で101件に達し、原因を公表したものの過半は委託先と外部連携だった。攻撃は修正より速い。悪用が確認された脆弱性の公表から大規模悪用までの中央値は5日、組織が修正を終えるまでの中央値は32～38日である。[\[R36\]](#)Criticalと評価された脆弱性の公表件数は、2026年は9か月で前年1年間の2倍に達し[\[R40\]](#)、修正版が出てても更新は遅れている。脆弱性の発見と悪用に長けたAIモデルも、一般に入手できる状況になりつつある。[\[R04\]](#)[\[R05\]](#)

GenerativeXは、次の3点に直ちに着手することを推奨する。

1. 依存ライブラリを、悪用が確認された脆弱性と外部公開システムの重大な脆弱性から解消し、全面的に更新する。続いて、公開Webサービスに緊急遮断の機構を導入する。
2. 利用中のSaaSに、SBOMまたは同等の透明性と、修正期限・通知SLAの短縮を求める。ライブラリの更新が間に合っていないものは、権限とデータを縮小し、最終的に利用を停止する。
3. Entra ID、Google Workspace等に連携済みのアプリの権限を最小化し、新規の連携を承認制にする。

3点はいずれも業務の停止を伴うため、経営が判断して全社に指示し、72時間、30日、90日の計画で実行する（[第8章](#)）。

序章 午前2時47分

2026年10月15日午前2時47分、Next.jsの開発チームが複数の脆弱性に対する緊急更新を公開した。[R08]

GenerativeXの監視は1分以内にこれを検知した。社内で管理する296リポジトリのうち148が該当し、22が本番環境で稼働していた。3時09分に修正のプルリクエストが自動生成され始め、3時13分、当番の担当者が公開中の57ホストを一括で遮断した。全更新を終えて遮断を解除したのは4時39分である。

修正の完了を待たずに遮断したのは、公表から悪用までの時間がすでに極めて短いからである。

Cloudflareは2024年、実証コードの公開から22分後に悪用の試行を観測している。[R33]しかも公表の時刻は配布元が決める。Next.jsの直近5回のセキュリティ更新は、いずれも日本時間の深夜0時から3時に公開された。[R44]

この時間で対応できたのは、公表された脆弱性を1分ごとに全リポジトリの依存関係と照合し、修正のプルリクエストを自動で作成し、公開中のサービスをWAFで一括して遮断できる仕組みを事前に用意していたからである（図1）。図1は一社の構成例である。各社に必要なのは、担当者の招集を待たずに、公表から停止までを進められる準備である。



図1 GenerativeXにおける脆弱性の自動検知・修正・遮断の仕組み

この夜、修正の大半は自動で進んだ。人が担ったのは、顧客向けサービスを止めるという判断である。深夜にその判断を誰がどの権限で下すかは、経営が事前に決めておくべき事項である。

緊急提言

3つの提言は、いずれも業務を止め、現場の負担を増やす。サービスの停止、SaaSの解約、連携の失効は、経営が判断して全社で進める。

提言① 依存ライブラリの全面的な更新と、公開Webサービスの緊急遮断の機構の導入

最優先は依存ライブラリの更新である。稼働中のシステムが使うライブラリ、フレームワーク、ランタイムと、VPN装置やファイアウォールなど外部との接続点にある機器のファームウェアを一覧にし、OSV、GitHub Advisory Database、JVN、配布元のセキュリティ情報と照合する。悪用が確認された脆弱性と、外部公開システムの重大な脆弱性を先に解消する。米CISAも2026年、悪用の危険が高いものから修正する方針を示した。[\[R42\]](#)残りも計画的にすべて更新し、最低でも既知の脆弱性を解消した版まで上げる。メジャーバージョンの更新は互換性を壊しうるため、サポート中の系列の最新の修正版を基本とし、回帰テストを経て段階的に適用する。使っていない機能に関わる脆弱性も対象に含める。[\[R08\]](#)

更新の経路そのものも攻撃される。2025年9月には、npmの500を超えるパッケージが、開発者の認証情報やクラウドのAPIキーを盗んで自己増殖するマルウェア「Shai-Hulud」に改ざんされた。[\[R55\]](#)依存関係はロックファイルで版を固定し、公開直後の版は安全を確認してから取り込む。開発者とCI/CDの認証情報は、フィッシングに強い多要素認証と有効期限の短いトークンで守る。

直ちに修正できないものは、WAFによる仮想パッチ、該当する機能の停止、アクセス元の制限、権限や鍵の失効の順にリスクを下げ、許容できる水準まで下げられない場合にサービス全体を停止する。

更新に続いて、公開Webサービスに緊急遮断の機構を導入する。更新しても次の脆弱性は公表され、悪用は修正より先に始まる。遮断の手段は、DNSの切り替え、CDNやロードバランサーでの遮断、メンテナンスページへの切り替えのいずれでもよい。夜間に担当者が一人でも数分で実行できること、決定者と実行者が事前に決まっていることが条件である。

根拠 修正版が出て更新は遅れており、Next.jsでは7週間前に修正された重大な脆弱性を抱えた版がダウンロードの38%を占める（図7）。大規模悪用は公表から0～5日で始まり、修正の完了には32～38日かかる（図6）。この提言は、入口の数と停止までの時間を小さくする。

提言② ライブラリの更新が間に合っていないSaaSは、利用を停止する

SaaSの脆弱性を直すのは事業者であり、事業者がライブラリを何日で更新するかが、そのまま自社のリスクになる。重要なデータと権限は、更新が間に合っているSaaSにだけ預ける。

契約、経費、各部門への確認から利用中のSaaSを一覧にし（無料利用や個人登録も含める）、預けたデータと付与した権限が大きいサービスから順に、事業者に次の3点を求める。

- 1. SBOMまたは同等の透明性：**サービスを構成するソフトウェア部品とその版の一覧（SBOM）、またはそれに代わる構成情報の開示を求める。経済産業省の手引きも、受発注時にSBOMの提供を取り決める契約の考え方を示している。[\[R41\]](#)SBOMに含まれる脆弱性が実際に影響するかは、事業者による影響判定（VEX）で確かめる。[\[R54\]](#)
- 2. 修正スケジュールと通知SLAの確認と短縮：**重大な脆弱性が公表されてから修正版を適用するまでの期限、対応状況の報告、利用者への通知の期限（通知SLA）を確認し、短縮を求める。米CISAは2026年、悪用が確認され、システムの全面的な制御につながる脆弱性の修正期限を3日とした。[\[R42\]](#)30日や90日の約束は、悪用の速さに対して遅すぎる。
- 3. 第三者による裏付け：**事業者の申告は、第三者による脆弱性診断の結果や第三者認証の報告書で裏付ける。チェックシートによる監査は、この3点で中身を確認して初めて意味を持つ。

大規模なSaaSで部品の一覧の開示が難しい場合は、影響判定、対応状況の報告、通知SLAで代える。求めに応じない事業者や更新の遅れが確認された事業者は、預けるデータと権限を縮小し、代替と移行の計画を用意したうえで、改善の期限を過ぎた時点で利用を停止する。

根拠 広く使われるライブラリでも、修正版の公開から7週間後に4割近くが古い版のまま使われている（図7）。SaaS事業者も同じライブラリを使っており、更新の遅れはそのまま利用企業の入口になる。外部の実業1社が侵害されると、100から700を超える組織が影響を受ける（図10）。この提言は、入口の数と到達できる範囲を小さくする。

提言③ 連携済みアプリの権限の最小化

SaaSの契約を見直しても、付与済みの権限は残る。Entra ID、Google Workspace、Slack、GitHubなどで連携済みの外部アプリとスコープを一覧にし、必要性や所有者が不明なもの、権限が広すぎるものを失効する。以後は新規の連携を管理者の承認制とし、承認に期限を設ける。[\[R16\]](#)サインインのための連携とデータへのアクセス権限は別物であり、前者を残したまま後者を絞ることができる。[\[R25\]](#)

根拠 2026年9月以降の国内の漏えいの公表で、原因を公表したものの過半（55%）は委託先・外部連携である（図8）。OAuth連携を起点とする被害は、1件で数百の組織に及ぶ（図10）。この提言は、到達できる範囲を小さくする。

第1章 攻撃は速くなり、修正は遅れている

AIによる攻撃の実例と能力評価

2025年11月、Anthropicは、同社のClaude Codeがサイバー諜報活動に使われ、偵察、認証情報の収集、横展開といった工程の多くをAIが実行していたと報告した。[\[R06\]](#)2026年9月には、米国NIST CAISIがZ.aiのモデルGLM-5.3の脆弱性発見・悪用能力を評価し[\[R04\]](#)、Anthropicは同等の能力が一般に入手できるモデルにも広がりつつあると報告した。[\[R05\]](#)いずれも管理された条件下での能力評価である。

それでも示唆ははっきりしている。脆弱性を探し、悪用手順を組み、標的ごとに調整する作業には、これまで熟練者の時間が必要で、それが攻撃の回数を抑えていた。AIがその一部を担えるようになり、この制約が緩みつつある。

公表から悪用までの時間

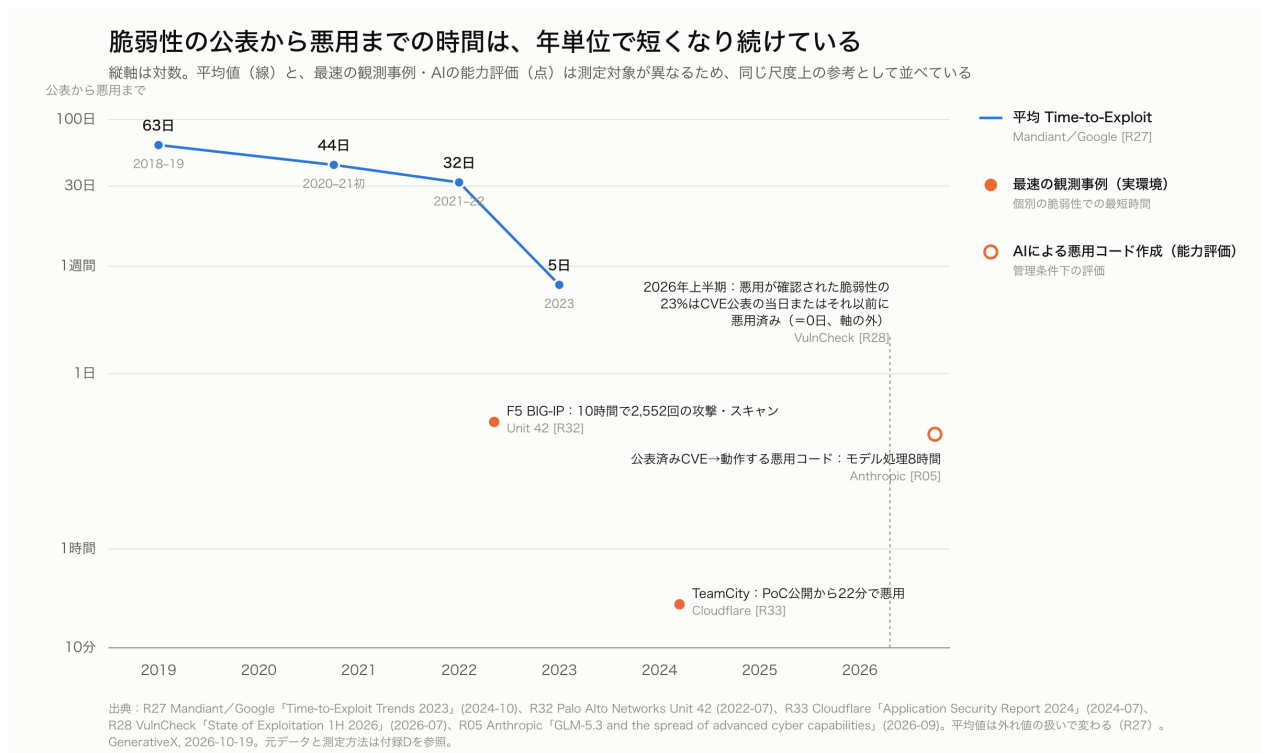


図2 脆弱性の公表から悪用までの時間

図2の数値は測定方法がそれぞれ異なる（元データと出典は付録D）。ただし傾向は一致しており、公表から悪用までの時間は短くなり続け、2026年には実環境で悪用される脆弱性の数も増えた。Google Threat Intelligence Groupは、増加の主因を、攻撃者がLLMで修正差分や公表情報を解析し、既知の脆

弱性を短期間で武器化していることに求めている。[R29]CVEの説明文を与えられたAIが公表済み脆弱性の多くを悪用できたという研究もある。[R30]

重大な脆弱性の公表件数

悪用までの時間が縮むのと同時に、対応が必要な脆弱性の数も増えている。CVEを公表した組織（CNA）が付与したCVSSで集計すると、Critical（CVSS 9.0以上）と評価された脆弱性は、2023年の2,355件から2025年には4,098件に増え、2026年は1～9月だけで8,222件に達した（図3）。[R40]評価の付与方法の変化の影響を除くため、CNAの評価だけで数えても、2026年1～9月は2025年1年間の2.3倍である（付録D）。公表件数の増加は、そのまま被害の増加を意味するものではないが、確認と修正の作業量の増加を意味する。

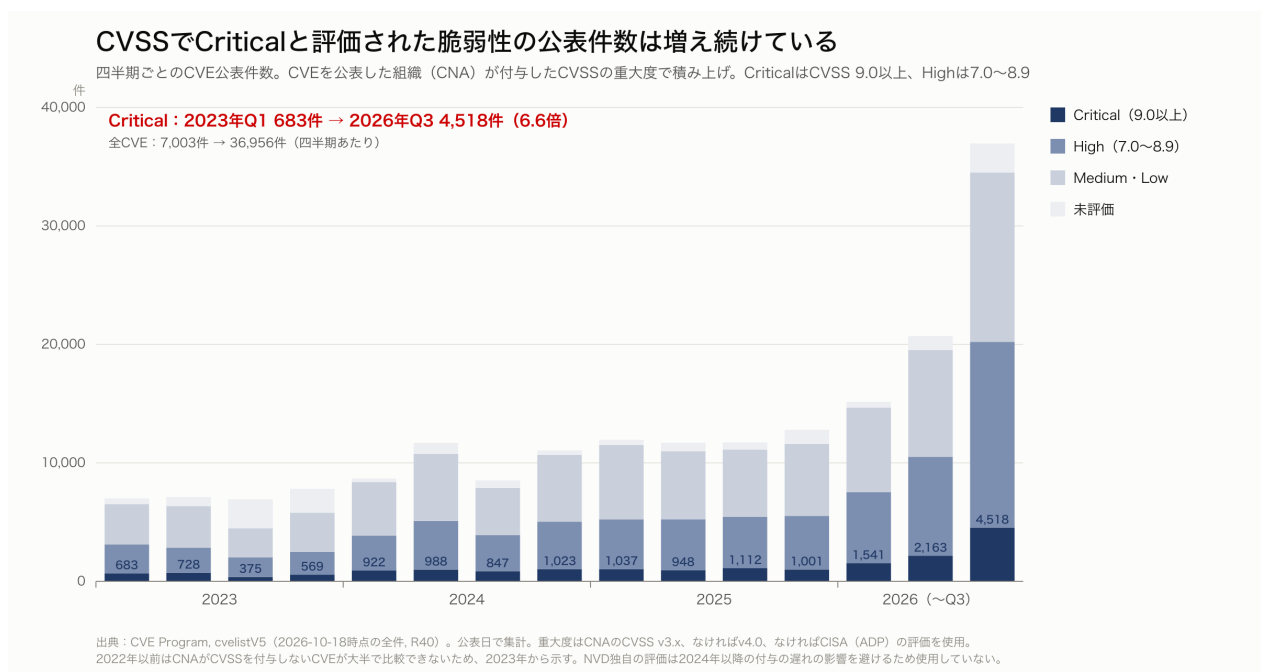


図3 CVSSの重大度別にみたCVE公表件数（四半期）

この件数に対しては、一件ずつ人が確認して修正する運用は遅すぎる。稼働中の版の照合、影響の判定、修正の作成は仕組みで自動化し、人は停止と承認の判断に集中する体制が必要である。

Criticalの中身を影響の種類で分けると、リモートからのコード実行につながるものが2023～2025年は37～39%を占め、2026年も3割を超える（図4）。コード実行につながるCriticalは、2023年の922件から2025年には1,563件になり、2026年は1～9月で2,599件に達した。認証を回避できるCriticalも、2023年の188件から2026年1～9月の933件へ約5倍に増えた。[R40]

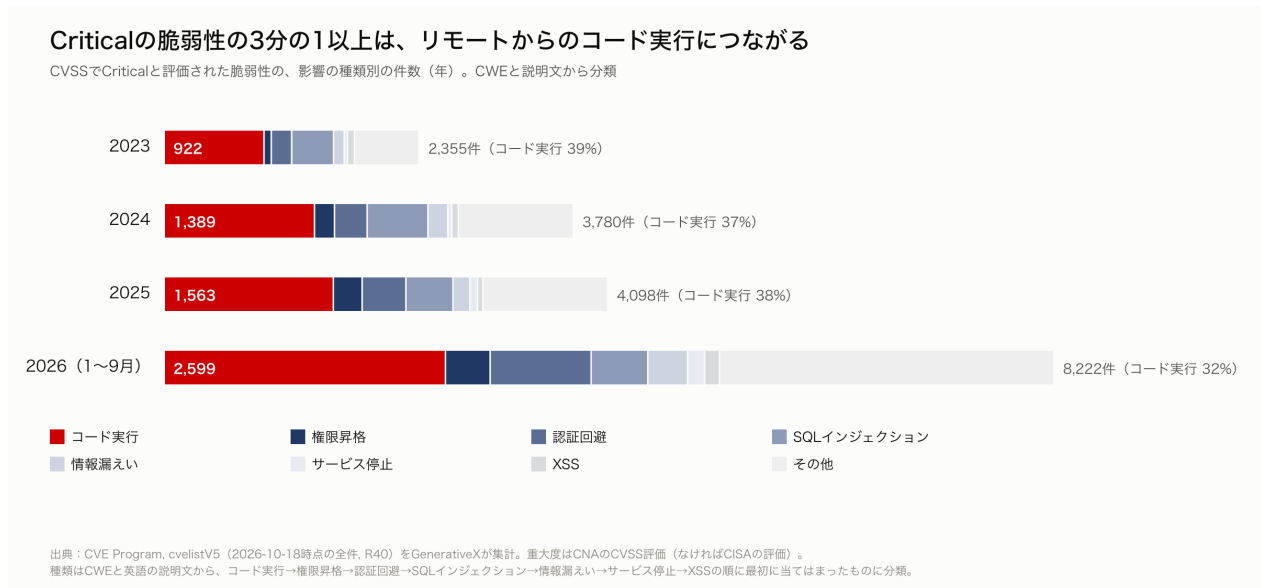


図4 Criticalの脆弱性の影響の種類別件数（年）

公表は日本の深夜に集中する

脆弱性の公表は、日本の業務時間を待たない。2025年1月から2026年9月に公表されたCritical 12,320件の公表時刻を日本時間に換算すると、32.2%が深夜0時から6時に集中する（図5）。9時台の2,015件は協定世界時0時の一括公表を多く含み、10時から18時の公表は15.2%にとどまる。[R40]序章のNext.jsの緊急更新が深夜に公開されたのも、この傾向のとおりである。公表を人が翌朝に確認する運用では、最初の数時間を毎回失う。

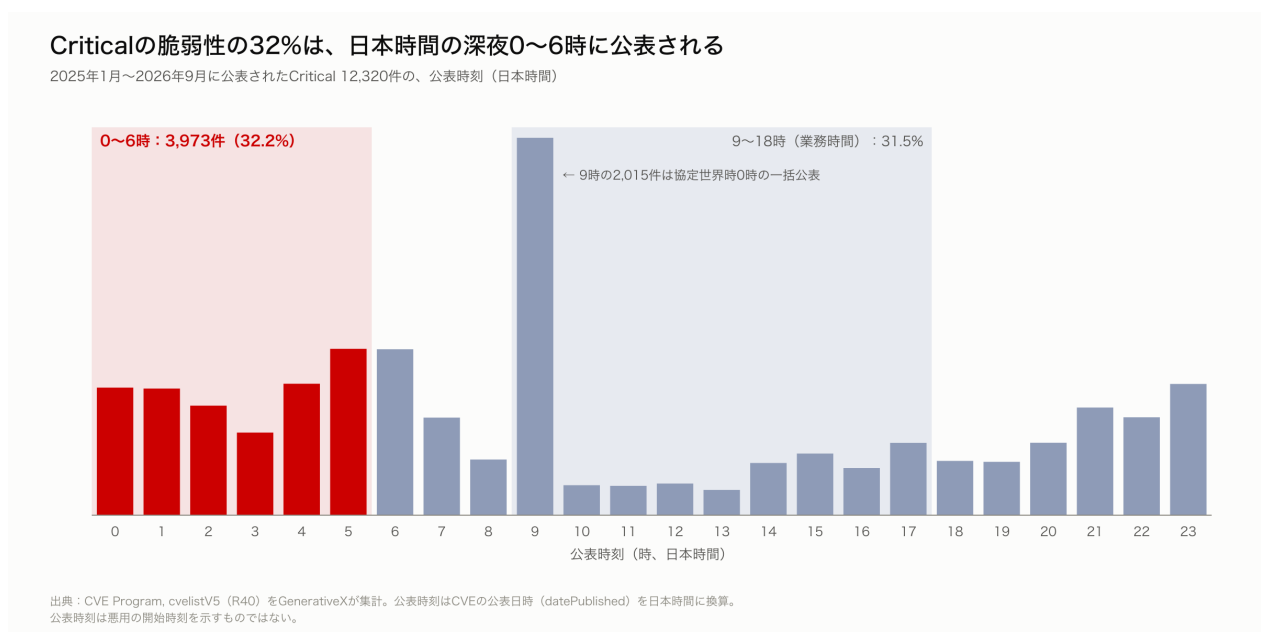


図5 Criticalの脆弱性の公表時刻（日本時間）

悪用は修正より先に始まる

攻撃側と防御側の時間を並べると、差ははっきりする。Verizonの2025年の調査では、悪用が確認された脆弱性の公表から大規模悪用までの中央値は5日、VPN装置やファイアウォールなどインターネットに直接公開している機器（エッジ機器）では0日だった。0日は、半数以上の脆弱性で公表の当日か、それ以前から大規模な悪用が始まっていたことを示す。一方、組織が修正を終えるまでの中央値は、全体で38日、エッジ機器で32日だった（図6）。[R36]攻撃が始まってから修正が終わるまでの約1か月、システムは既知の脆弱性を抱えたまま公開されている。この差を埋めるには、修正を速めると同時に、修正を待たずに停止できる仕組みが必要である。

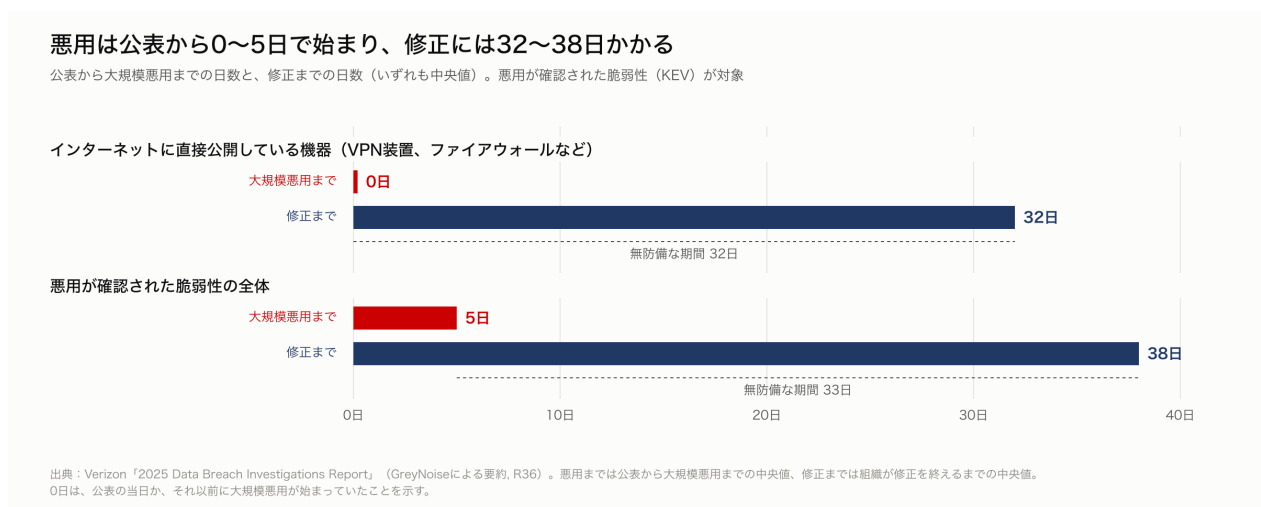


図6 公表から大規模悪用までの日数と、修正までの日数

更新は修正版の公開から遅れている

修正版が出て、利用者の更新は遅れている。Next.jsの2026年10月10日から16日までのnpmでのダウンロード4,077万件を版ごとに分けると、37.6%が、7週間前の8月25日に修正版が出たCriticalの脆弱性を抱えた版だった。9月22日に修正されたCriticalも含めると48.3%になる（図7）。[R43][R44]ダウンロード数は開発環境やCIでの取得を含み、稼働中のシステム数とは一致しないが、重大な脆弱性の修正から2か月近くたっても、古い版が広く使われ続けていることを示している。

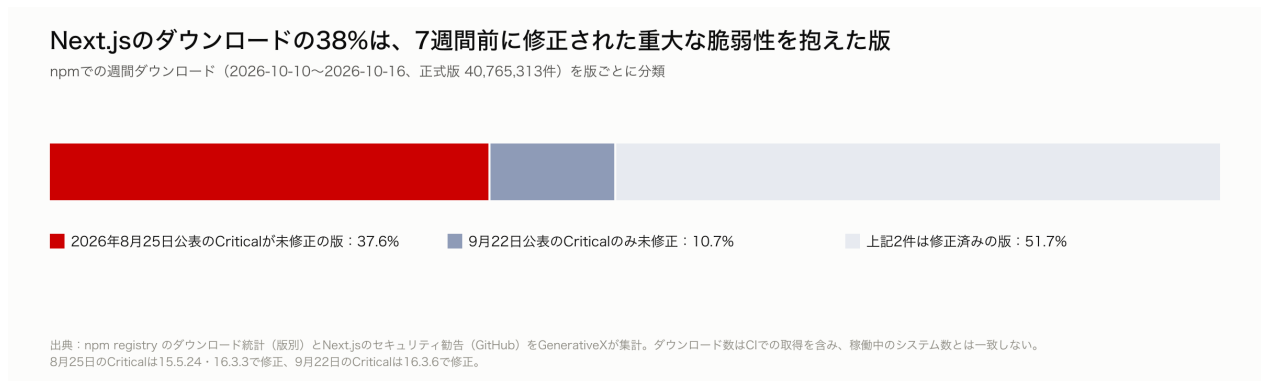


図7 Next.jsの週間ダウンロードの版別の内訳

影響を受けるシステムの規模も大きい。2026年10月14日に修正されたcurlの脆弱性は、悪意あるサーバーへの接続で、特定の条件のもとでコード実行に転用できることが実証されている。[\[R39\]](#)curlは「200億を超える」環境で使われているとされ[\[R45\]](#)、サーバー、端末、組み込み機器の多くが、利用者が意識しないまま影響を受けうる。

未公表の脆弱性

攻撃者は未公表の脆弱性も使う。企業が使うフレームワークやライブラリの多くはオープンソースで、コードは誰でも読める。未知の脆弱性を探す作業はこれまで一部の組織に限られていたが、CAISIとAnthropicが評価したのはまさにこの能力である。[\[R04\]](#)[\[R05\]](#)2026年10月には、セキュリティ企業XBOWが、公開されているOpenAIのモデルを使い、Chromeの未知の脆弱性3件を6日間で発見・連結したと公表した。[\[R31\]](#)

稼働中のソフトウェアには、まだ公表されていない脆弱性があると考えべきである。公表前の攻撃に備えるには、修正の速さと同じ重みで、侵入された場合の影響範囲を限定し、速やかに封じ込める準備が必要である。

第2章 2026年秋、国内で起きたこと

公表は9月末から急増した

GenerativeXが各組織の公表と報道から収集したところ、2026年8月31日から10月16日までに、国内の組織による不正アクセス等の漏えい（可能性を含む）の公表は153件あった。週ごとの公表は9月半ばまで10件以下だったが、9月28日の週に23件、10月5日の週に41件、10月12日の週に60件と増えた（図8）。原因を公表した105件のうち、最も多いのは委託先・外部連携の58件（55%）である。このうち約40件は、災害情報のメール配信サービスやクラウド事業者など5つの事業者の侵害を、それぞれの委託元が公表したものである。次いでWebサービス・APIの脆弱性が24件ある。10月8日のJPCERT/CCの注意喚起の後には、公表と報道そのものが増えた可能性もある。個人情報保護委員会が処理した漏えい等報告は、急増の前の2026年4～6月に6,101件で、前年同期の5,652件から8%増えた。この件数には誤送信や紛失も含まれる。[R52]

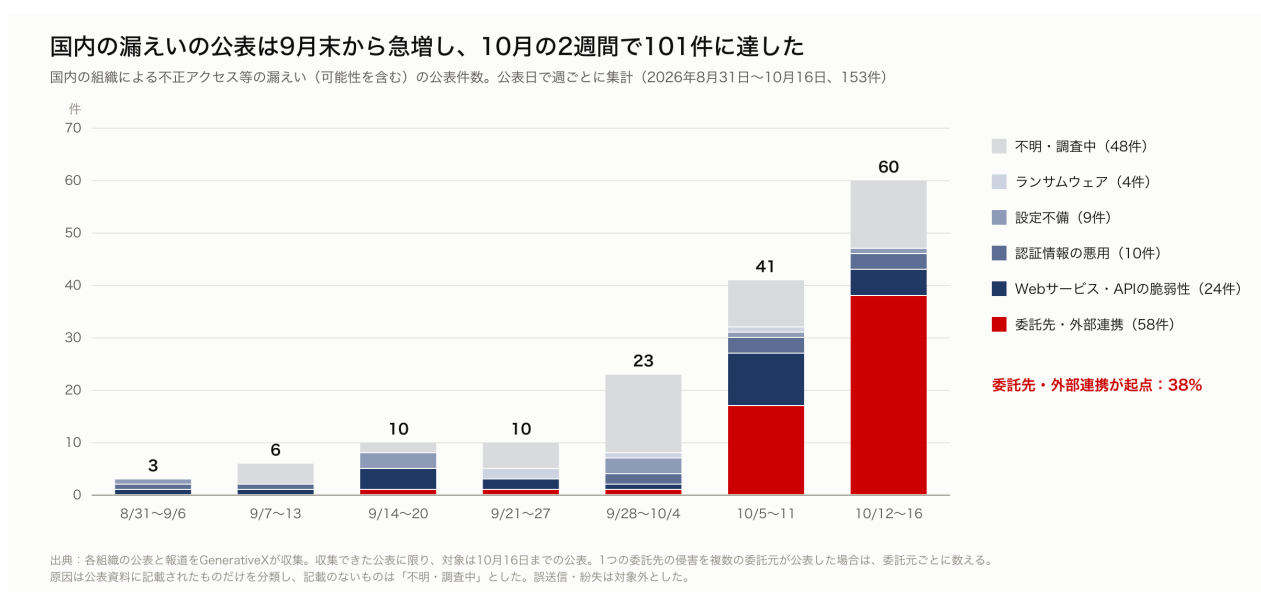


図8 国内の漏えいの公表件数（週別・原因別）

ランサムウェアの犯行グループがリークサイトに載せた日本の組織の数は、9月が15件で、2月以降の月9～15件と同じ水準にある（図9）。9月以降の公表の急増は、委託先とWebサービス・APIを起点とするものが中心である。リークサイトの掲載件数そのものも増えており、2025年は前年の2.3倍の115件、2026年は9月までで107件に達した。[R51]

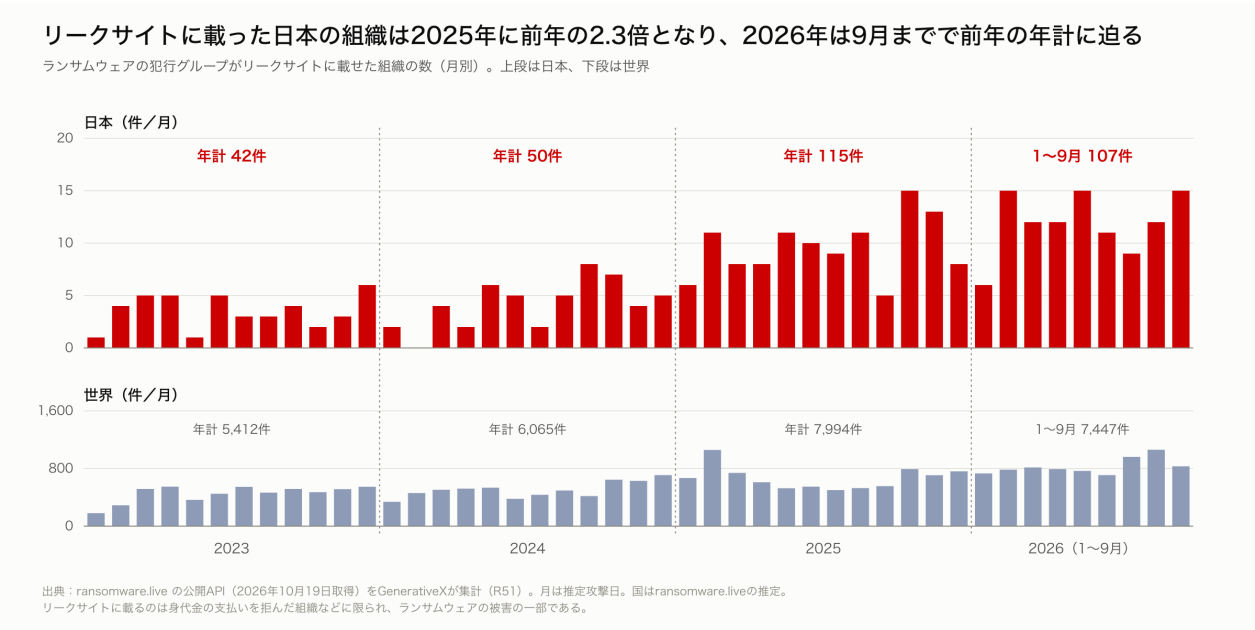


図9 ランサムウェアのリークサイトに載った組織の数（月別）

注意喚起と緊急更新

日付	発信元	内容	自社にとっての意味
10/8 (10/9更新)	JPCERT/CC	相次ぐ不正アクセスについて注意喚起。既知の脆弱性に加え、APIの不正操作、設定・管理の不備を点検項目に挙げた [R24]	パッチに加え、APIの認可と設定を見直す
10/13	Gartner	日本企業向けの緊急提言。公開資産、認証認可、第三者に加え、経営による停止・改修の判断を求めた [R02]	停止と改修の判断を経営が担う体制を整える
10/13	トレンドマイクロ	公開資産の更新、サプライヤーの点検、AI利用の把握を推奨。国内事案とAI利用の関連は確認していないと明記 [R03]	原因の特定を待たずに点検を進める
10/13	NICT CYNEX	公開アプリを解析してサーバーを攻撃する手口に注意喚起 [R15]	クライアント側に秘密鍵や認可の判断を置いていないか確認する
10/15（日本時間）	Next.js	複数の脆弱性に対する緊急セキュリティ更新 [R08]	使っていない機能の脆弱性も含めて直ちに更新する
10/15	JPCERT/CC	調査した攻撃インフラで、AIを用いた既知脆弱性のスキャンとAPI経由の攻撃を確認 [R01]	AIを使う攻撃者を前提に、照合と遮断を自動化する

攻撃の対象は公開Webアプリの脆弱性に限られず、APIの認可や設定の不備にも及んでいた。[R24][R15]公開画面の診断に加え、その背後のAPIの認可も点検する必要がある。AIを用いたスキャンは国内でも観測

された。[\[R01\]](#)個々の事案の原因がAIかどうかにかかわらず、企業はAIを使う攻撃者を前提に備えるべきである。

攻撃者が到達し得る対象は、従業員の端末、クラウドの管理基盤、外部SaaS、SaaS間の連携トークン、取引先に渡したデータ、社内のAIエージェントにまで広がっている。次章では、このうち最も見落とされやすい外部連携を取り上げる。

第3章 漏えいは外部連携から起きる

第2章でみたとおり、国内の漏えいの原因で最も多いのは委託先と外部連携である（図8）。海外でも同じ構造の被害が続いている。

2025年8月、Googleの脅威インテリジェンスグループ（Mandiant）は、Salesloft社のチャットボット製品「Drift」を起点とする大規模なデータ窃取を報告した。^[R07]攻撃者はDriftを侵害し、Driftが保管していた各社Salesforce向けのOAuthトークンを使って顧客データを取得した。Drift Emailと連携していた一部のGoogle Workspaceアカウントへのアクセスも確認されている。影響を受けた可能性がある組織は700を超える。^[R47]被害企業の多くで、侵害されたのは自社のSalesforceやGoogle Workspaceではなく、権限を付与していた第三者だった。

外部の事業者1社の侵害で、その事業者に権限やデータを預けていた100から700を超える組織が影響を受けた事例が、2024年以降も相次いでいる（図10）。2025年11月には、顧客管理の連携アプリを提供するGainsightを起点に、200を超えるSalesforceのインスタンスが影響を受けた可能性がある^[R48]とGoogleが報告した。2026年6月には、競合分析サービスのKlueで、2022年に作られ使われていなかった認証情報から各社のOAuthトークンが盗まれ、195の組織が影響を受けた。^{[R49][R53]}国内では2024年5月、通知書の印刷と発送を受託していたイセトーが、2021年4月から更新していなかったVPN機器を経由して侵入された。個人情報保護委員会によれば、直接の委託元41団体、再委託元を含め約100団体の、約308万人分の個人データが漏えいした。^[R50]Verizonの2025年の調査では、データ侵害のうち第三者が関与したものの割合は前年から倍増し、30%に達した。^[R46]

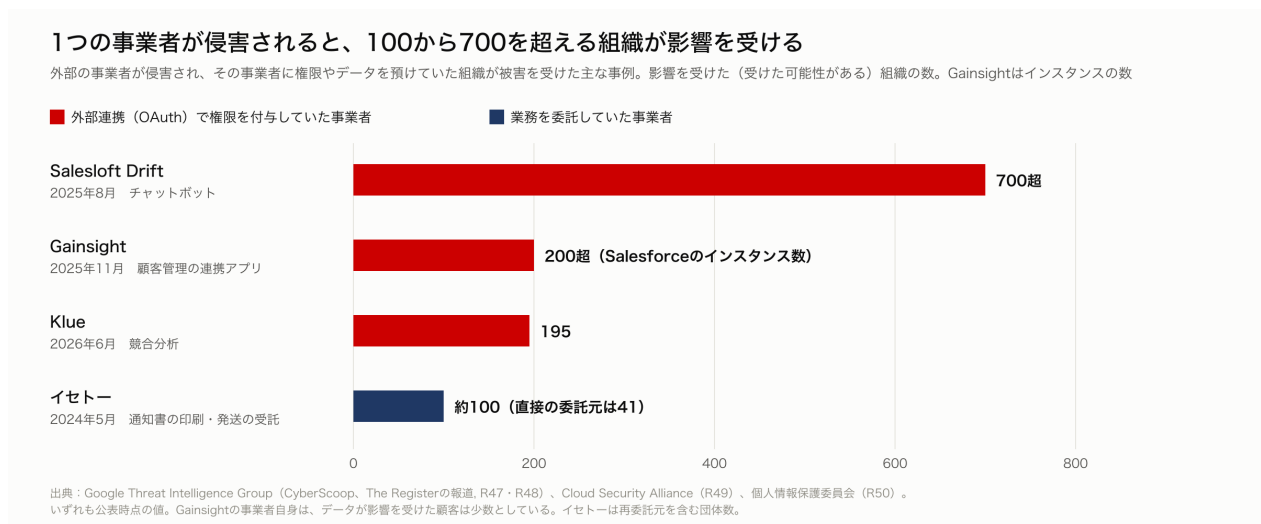


図10 1つの事業者の侵害で影響を受けた（可能性がある）組織・インスタンスの数

これらの事例を踏まえると、外部サービスについて確認すべきことは、「そのSaaSは安全か」に加えて「そのSaaSが侵害されたら何を失うか」である。どの権限を付与し、その権限でどのデータまで到達でき、何分で失効できるか。

失効で止められるのは、その後のアクセスだけである。Googleは、第三者アプリが利用者のデータを自社のサーバーへ複製し得ること、複製後の安全性や削除はGoogleだけでは管理できないことを説明している。[\[R23\]](#)取引先に送った資料、委託先に預けた顧客リスト、従業員の端末に同期されたファイルも同様である。データは提供する前に、項目、量、保持期間を絞る必要がある。

1つのSaaS契約には、少なくとも4つの管理対象がある。サービス自体、付与した権限、提供したデータ、そして停止や連携解除をしたときの業務への影響である。サービス一覧、OAuth連携一覧、データ一覧、契約一覧は、互いに照合して管理する。

第4章 GenerativeXの考え方

第1章から第3章を踏まえ、GenerativeXは次の3つを防衛の前提に置く。

侵入は起きるものとして扱う。稼働中のソフトウェアには未公表の脆弱性があり、それをAIで探す攻撃者も増える。脆弱性診断とパッチは引き続き必要だが、経営が管理すべきは侵入された場合の被害の大きさである。

外部に付与した権限と外部に提供したデータは、自社の管理対象とする。自社のサーバーが安全でも、権限やデータを渡した第三者が侵害されれば情報は漏えいする。

AIエージェントは、誤った判断をする前提で設計する。誤った判断をしても実行できる操作が限られるよう、認可と実行制御をモデルの外側で強制する。

被害の大きさを左右する3つの要素

侵入を前提にすると、被害の大きさは、入口の数、1つの侵害で到達できる範囲、停止までの時間の掛け算で表せる。これは厳密な計算式ではなく、被害を左右する要素を整理するための概念モデルである。入口の数は、公開API、管理画面、ID、外部連携、VPN装置などのうち、悪用されうるものの広さ（攻撃面）を指す。到達できる範囲は、侵害された1つの主体から、データ、クラウドの管理基盤、CI/CD、サービスアカウントや長期有効な認証情報を経て、どこまで広がるかを指す。どれか1つを小さくすれば被害は小さくなり、3つすべてを小さくすれば大きく縮む。緊急提言の①は入口の数と停止までの時間を、②は入口の数と到達できる範囲を、③は到達できる範囲を小さくする。

この前提に立つと、経営がセキュリティ部門に確認すべきことも変わる。

これまでの確認事項	これから確認すべきこと
侵入されない対策はできているか	端末、SaaS、AIのいずれかが1つが侵害されたとき、影響はどこまで及ぶか
パッチは適用したか	公表から修正の反映まで何分かかるか。間に合わないとき、誰が何を止めるか
そのSaaSは安全か	そのSaaSが侵害されたら何を失うか。提供したデータはどうなるか
AIは安全に使えているか	AIが悪意ある指示を読み込んだとき、どの操作まで実行できてしまうか

閉域網もSSOも、侵入の入口を残している

「閉域網に置いているから」「SSOと多要素認証を入れているから」という理由で、自社は安全だと判断する企業は多い。いずれも有効な対策だが、侵入の入口は残る。

閉域網の安全性は、閉域網に接続しているすべての事業者と機器の安全性で決まる。2022年10月、大阪急性期・総合医療センターでは、攻撃者が給食事業者のデータセンターに、リモート保守用に設置されたVPN機器の脆弱性を使って侵入した。攻撃者はそこで窃取した認証情報を使い、閉域網でつながった病院の給食サーバーを踏み台にして、電子カルテを含む基幹システムをランサムウェアで暗号化した。基幹システムの再稼働に43日、診療システム全体の復旧に73日を要し、逸失利益は十数億円以上と見込まれている。[\[R37\]](#)保守用のVPN機器のようなエッジ機器は、実環境で悪用される脆弱性の主要な入口でもある。Google Threat Intelligence Groupの集計では、2026年1～8月に悪用された脆弱性の14%がエッジ機器とセキュリティ機器だった。[\[R29\]](#)閉域網の内側も外部と同じ基準で操作ごとに認可し、外部との接続点にある機器を最優先で更新する。

SSOと多要素認証は、人のログインを守る。一方、アプリケーションやライブラリの脆弱性を突く攻撃は、ログインを経ずにサーバー上でコードを実行する。Recorded Futureの集計では、2026年上半期に実環境で悪用された215件の脆弱性のうち68%にあたる146件が、事前の認証なしに悪用できた。リモートでコードを実行できる脆弱性は82件あり、そのうち60件はネットワーク越しに認証なしで悪用できた。[\[R38\]](#)2026年10月15日（日本時間）にはNext.jsが、認証なしでリモートからコードを実行できる重大な脆弱性2件を修正し[\[R08\]](#)、前日の10月14日にはcurlも、特定の条件のもとでコード実行に転用できることが実証された脆弱性を修正した。[\[R39\]](#)AIはこの種の脆弱性の発見にも使われており、AIが発見した脆弱性の50%はリモートでのコード実行につながる。[\[R29\]](#)認証基盤の整備と並行して、公開サービスの更新と停止の手段を用意する。

ゼロトラストとの関係

これらはゼロトラスト（NIST SP 800-207）の延長にある。[\[R11\]](#)ゼロトラストは、ネットワーク上の位置を信頼の根拠にせず、アクセスごとに主体と資源を検証する原則である。本稿はこれを3つの点で広げる。第三者に付与した権限と提供したデータ（自社の認可が及ばない）、AIエージェント（SP 800-207の公表時点では想定されていない主体）、そして正規の権限が侵害された後の影響範囲と停止の判断である。NIST CSF 2.0やサプライチェーンリスク管理など、既存の枠組みの上に積み増して使う。[\[R10\]](#)[\[R13\]](#)

第5章 4つの柱

緊急提言の後も、3つの要素は継続して小さくし続ける必要がある。GenerativeXは、そのための対策を4つの柱で整理する。①把握は他の柱の前提となり、②は入口の数を、③は到達できる範囲を、④は停止までの時間を小さくする（図11）。4つの柱は、同じ資産や連携に並行して適用する。



図11 被害の大きさを左右する3つの要素と4つの柱

柱	答える問い	効果を確認する材料
① 把握する	何を保有し、何に接続し、何を委任・提供しているか	責任者、実効権限、データの所在、最終確認日
② 削減する	そのうち不要なものは何か	廃止・失効・削除の結果と、残す理由
③ 影響範囲を限定する	残したものが侵害されたとき、影響はどこまで及ぶか	1つの主体の侵害を仮定した検証結果
④ 停止までの時間を短縮する	危険な状態を何分で終わらせられるか	暫定緩和、恒久修正、失効、業務復旧の実測時間

把握する

多くの資産一覧は名称の一覧にとどまり、誰がどの権限でどのデータにアクセスでき、誰がそれを止められるかの記録を欠いている。棚卸しは、各項目を止められる状態で記録できて初めて完了とする（記録項目は付録A）。まずは重要データにつながる経路のうち、公開されている入口、管理者権限と恒常的な特権、外部サービスへのOAuth連携の三種類を洗い出す。被害が大きく、一覧化しやすいからである。

削減する

未使用のAPI、旧環境の管理画面、使われていないOAuth連携、休眠アカウント、過去案件のデータ、不要なライブラリ。使われていなくても、保守と防御の責任は残る。まずはOAuth連携のうち所有者が答えられないものを失効する。認証、監査、バックアップといった防御機能は維持し、業務に必要な正規の手段も残して、個人アカウントへの迂回を防ぐ。効果は、重要データへの不要な経路がいくつ減ったかで測る。

影響範囲を限定する

顧客、案件、環境ごとにデータ、実行ID、鍵、管理権限を分ける。ネットワークの到達制限だけに頼らず、API、データベース、ストレージの側で操作を認可する。管理基盤と通常業務、バックアップと本番の権限も分ける。クラウドの管理基盤、CI/CD、サービスアカウントの権限は通常業務の権限から切り離し、長期有効な認証情報は有効期限の短いトークンに置き換えて、侵入後の横展開を抑える。社内や同じグループ内からのアクセスも、操作ごとに認可する。[R11]

検証は、侵害される主体を1つ選んで実際に試す。担当者一人のセッションが盗まれたら他社のデータを見られるか。アプリ1つが侵害されたら他環境の秘密情報に到達できるか。本番の管理者が侵害されてもバックアップを守れるか。結果は、資源側でアクセスが拒否された記録で確認する。なお、読み取り専用の権限が防ぐのは改ざんであり、漏えいには別の対策が要る。

停止までの時間を短縮する

公表から悪用までが数十分から数時間であれば、プルリクエストの作成は途中経過にすぎない。影響の特定、暫定緩和（序章の遮断がこれにあたる）、恒久修正、権限の失効、業務の復旧について、それぞれ時間を計る。完了の定義は、稼働環境で修正版が動いていることの確認とする。マージやデプロイの成功は途中経過である。WAFやIP制限は暫定防御として有効だが、正規権限の悪用やデータの持ち出しには権限の失効で対応する。

第6章 AIエージェント

序章の夜、修正のプルリクエストを作成したのはAIエージェントである。数時間以内の対応には、AIによる自動化が有効である。一方でAIエージェントは、権限を持ち、データを読み、外部に送信する主体でもある。

AIエージェントには、読み込んだ文書に書かれた指示に従ってしまうという固有の弱点がある。OWASPが「間接プロンプトインジェクション」と呼ぶリスクで、悪意あるメールやWebページに埋め込まれた指示が、エージェントの操作を誘導する。[\[R18\]](#)誘導されたエージェントは、外部からの侵入を伴わずに、正規の権限でデータを外部に送る。

モデルの精度が上がっても誤りは残る。対策は、誤った判断をしても許容範囲を超えて実行できない仕組みをモデルの外側に置くことである。[\[R19\]](#)[\[R20\]](#)GenerativeXは、AIエージェントに次の5つを求めている。

- **権限を絞る**：呼出し元の権限を上限とし、タスクの目的、対象、期間でさらに制限する。読取、変更、削除、外部送信、管理操作を区別する。
- **認可をモデルの外で強制する**：ツールの実行側と資源側で、主体、対象、操作を検証する。権限の根拠は、認可基盤が発行した情報に限る。
- **実行環境と送信先を制限する**：作業領域、参照ファイル、実行時間、通信先を限定する。同じSaaSの中でも送信先のテナントや共有先を区別する。
- **高リスクな操作は内容を示して承認を取る**：外部送信、破壊的な変更、本番操作では、何をどこへ実行するかを示して承認を得る。承認は操作ごとにする。
- **追跡と停止ができるようにする**：誰の委任で何を実行したかを記録し、実行中のタスク、保持するトークン、外部ジョブを止められるようにする。

サンドボックスによる隔離は有効だが、漏えいを防ぐには権限の制限も要る。[\[R21\]](#)隔離環境の中に広い読取権限と外部送信の手段があれば、許可された操作だけで情報は外に出る。実行の隔離、データとツールの認可、送信制御を組み合わせる。

防衛にAIを使う場合も同じ原則を適用する。防衛用のエージェントに与える管理権限は、タスクごとに必要な範囲に限る。序章の夜、GenerativeXが自動化していたのは検知と修正の作成までで、遮断は人が判断した。次の段階として、重大な脆弱性では通知より先に遮断する自動化を進めているが、その場合も遮断の範囲、再開は人が判断すること、誤遮断時の復旧手順を先に決めている。

第7章 経営が決めること

公表から悪用までの時間が短くなるほど、サービスを停止するかどうかは事業の判断になる。修正の自動化が効果を持つのは、誰がどの権限で判断するかが決まっているときである。NIST CSF 2.0はガバナンスを独立した機能として明示し[R10]、Gartnerの提言も停止の判断を経営に求めている。[R02]

役割

役割	主な決定事項
経営・事業責任者	守る業務、許容できない被害、停止時の事業判断、残余リスクの受容
データ責任者	利用目的、外部提供、保持・削除
サービス・システム責任者	稼働資産、依存関係、更新、技術的な復旧
セキュリティ責任者	優先度、横断的な調査、緊急時の封じ込めの実行
調達・法務	委託・再委託、事故時の連絡、契約終了時のデータの扱い

役職は兼務でよい。ただし決定者と実行者は特定しておく。

事前に決めておく4つの事項

停止の判断は、侵害が起きる前に準備しておく。次の4つを平時に決めておく。

- 1. 重要なサービスとデータ：**どのサービスとデータが事業にとって重要かを順位づける。止められないもの、一時的に止めてもよいものを区別し、停止の判断はこの順位を基準に行う。
- 2. シナリオごとの停止条件：**想定するシナリオごとに、どの条件のときに、どの単位で停止するかを決める（シナリオ別の対応手順）。シナリオの例は、重大な脆弱性の公表、トークンの侵害、連携先の外部サービスの侵害、AIエージェントの承認範囲外の動作である。停止の単位は、サービス全体だけでなく機能、アプリ、トークン、顧客ごとに選べるようにする。あわせて、停止している間に業務をどう続けるか（手作業への切り替えなど）と、顧客にいつ何を知らせるかも、シナリオごとに用意しておく。
- 3. 判断者と実行者：**夜間・休日を含め、誰が停止を判断し、誰が実行するかを決める。
- 4. 再開の条件：**何が確認できたらサービスを再開するかを決める。修正版の反映、権限の失効、残存するアクセスがないことの確認などである。

停止そのものが重大な被害を生む業務では、一律の自動遮断は避ける。

指標

経営は、3つの要素の推移を、次の5つの指標で月次に報告させる。

指標	経営が知るべきこと	目標の例
外部公開資産の把握率	把握できていない入口がどれだけあるか	100%
外部連携の棚卸し率	データと権限を預けた先を把握しているか	所有者不明の連携0件
重要システムの侵害影響範囲	1つの認証情報でどこまで到達できるか	侵害を想定した検証で越境0件
重大な脆弱性への対応時間	公表から更新・遮断・暫定緩和でリスクを下げるまで何時間か	悪用が確認されたものは24時間以内
緊急遮断の実測時間	判断から遮断の完了まで何分か	夜間でも数分

第8章 推奨する実行計画

経営が対応を宣言し、トップダウンで全社に展開する。期間は目安であり、緊急のリスクには期間を待たずに対応する。

最初の72時間

責任者：経営が指示し、セキュリティ責任者とサービス・システム責任者が実行する。

経営が全社にセキュリティ総点検を指示する。稼働中のライブラリの版を脆弱性データベースと照合し、悪用が確認された脆弱性と公開システムから更新を始める。更新できないものは停止するか公開範囲を限定する。72時間で求めるのは、全ライブラリの更新への着手と、悪用が確認された脆弱性の解消である。あわせて、利用中のSaaSと連携済みアプリの一覧化に着手し、所有者や必要性が不明なもの、不要な管理画面、休眠中の特権IDを停止・失効する。

30日以内

責任者：セキュリティ責任者、調達・法務。

全ライブラリの更新を完了し、公開Webサービスに緊急遮断の機構を導入して、遮断の決定者と実行者を指名する。クラウド、IdP、SaaS、データ、契約の一覧を照合して台帳にまとめ、重要サービスの実行権限を絞り、外部共有とOAuth連携を再審査する。主要なSaaSの事業者に、SBOMまたは同等の透明性と、修正期限・通知SLAの短縮を求める。

90日以内

責任者：サービス・システム責任者、データ責任者。

顧客や環境ごとの分離、不要なサーバー側機能の廃止など、構成を変更する。主要な主体について侵害された場合の影響範囲を検証し、封じ込めの演習で所要時間を実測する。取引先へのデータ提供の条件（目的、項目、保持期間、削除の確認）を整え、AIエージェントを[第6章](#)の5つの要件で点検し、復旧演習でバックアップの独立性を確かめる。新規契約、権限付与、デプロイの手順に台帳の更新を組み込む。

結び

2026年9月末からの国内の漏えいの公表の急増は、攻撃の前提が変わったことを示している。公表件数の増加にはCNAの増加など複数の要因があり、国内の急増にもAI以外の要因が含まれる。それでも、AIによる攻撃能力の普及を踏まえれば、攻撃の高速化は構造的に続き、後戻りしにくい変化とみるべきである。企業は、攻撃の高速化が続くシナリオを前提に、防衛体制を再設計すべきである。

この脅威への対応は、IT部門の権限の範囲を超える。ライブラリの全面的な更新は開発の計画を変え、緊急遮断は売上を生むサービスを止め、SaaSの利用停止や連携の失効は各部門の業務を変える。どのサービスをどの条件で止め、どの事業者にどのデータを預けるかは、事業の判断である。IT部門だけで進めれば、これらの判断は先送りされ、入口の数、到達できる範囲、停止までの時間は大きいまま残る。

経営に求められるのは、セキュリティを事業の継続条件として扱い、継続してコミットすることである。緊急提言の3点を経営の判断として全社に指示し、重要なサービスの停止条件と判断者を平時に決め、緊急時の停止の責任を負う。[第7章](#)の5つの指標を月次で報告させ、その改善に予算と人員を恒常的に配分する。

侵入は起きる。被害の大きさは、経営の判断と準備で小さくできる。

付録A 実装の詳細

A.1 記録の対象

対策は、次の4つの要素で記録する。既存の枠組みにおける資産中心の認可、データフロー、外部依存の考え方に基づく。[R11][R13][R22]

- **保護対象**：情報（顧客情報、営業秘密、ソースコード、認証情報）、機能（受発注、決済）、管理権限（IdP、クラウドの管理基盤、CI/CD、DNS、暗号鍵）、復旧手段（バックアップ、復旧用ID）
- **アクセス主体**：人（従業員、委託先、取引先）と、人以外（アプリ、CI/CD、サービスアカウント、AIエージェント）。AIエージェントには委任者、実行ID、使えるツールも記録する
- **関係**：接続（公開Web、API、Webhook、AIが読む文書）、認可・委任（IAM、OAuth、共有権限）、データの移転（SaaSへの保存、取引先への送付、モデルへの入力）、依存（OSS、AIモデル、MCPサーバー、IdP、クラウド）
- **管理主体**：誰がその関係に制約をかけられるか。自社で制御できない場合は、相手の証跡、契約、提供量の制限で補う

OAuthによるデータ取得のように、1つの連携が接続、認可、データ移転の複数に該当することがある。

A.2 台帳の記録項目

記録するのは、秘密情報の保管場所や識別子である。

- 対象業務／保護対象
- 主体／実行ID／委任者
- 関係の種類／起点と終点／管理主体
- 許可する操作とデータ／許可しない範囲
- 目的／有効期間／最終利用日／最終確認日
- 責任者／承認者
- 侵害時の影響範囲
- 停止・失効の方法／外部に渡したデータの扱い／復旧手順
- 未確認事項／次の対応期限

A.3 適用例：問い合わせメールに回答案を作るAI連携

架空の例である。

関係	把握する	削減する	影響範囲を限定する	停止までの時間を短縮する
メールAPIの呼出し	使うAPIと実行元	不要なAPIを外す	呼出し元と頻度を制限	連携停止の手順を確認
メール参照権限の委任	実効権限とトークン	全メール・恒久権限を避ける	案件用の主体と対象に限定	認可とトークンを失効
本文の外部処理	どこに何が渡るか	必要な部分だけ渡す	保持期間を制限	処理停止と保存データの確認
返信の送信	送信先と承認経路	当初は送信せず回答案の作成に限定	送信時は内容を示して承認	誤送信の停止と影響確認

AIに「他案件のメールを読まない」と指示することと、API側で読めなくすることは別である。最小権限は、API側で案件単位の制限を実装して初めて成立する。

A.4 公開サービスの評価単位

脆弱性の影響は、稼働している成果物、有効な機能、到達可能な入口、実行権限で判断する。たとえばNext.jsの静的エクスポートでは、Server Componentsはビルド時に実行されるため、ソースにあってもリクエスト時には動かない。[\[R09\]](#)一方、2026年10月15日（日本時間）の更新は、類似コードの有無で判断せず直ちに更新するよう求めている。[\[R08\]](#)

業務の条件	検討する構成	注意点
利用実態がない	廃止	データ、認証情報、DNSも処理する
表示を静的に生成できる	静的配信と、範囲を絞ったAPI	APIの認可、成果物の公開範囲
画面がブラウザ側で完結する	SPAとAPIの分離	トークンの保管、XSS、資源単位の認可
リクエスト時の描画が必要	SSRを残し、機能と権限を限定	更新責任と監視対象の明確化

構成の見直しを待たずに、セキュリティ更新は先に適用する。

A.5 認証と資格情報

- ・特権の利用を中心に多要素認証を整備し、フィッシング耐性のある方式を優先する。[\[R26\]](#)
- ・管理されていない端末から重要資産を使えないようにする。
- ・人以外の主体には用途別のIDと短期の資格情報を使い、鍵はサービスごとに分ける。

- ・特権を持つ秘密鍵はサーバー側だけで保持する。ブラウザーや配布アプリに置いた情報は、難読化しても読まれる前提で扱う。[R01][R15]
- ・バックアップは削除・改ざんから保護し、復旧用の権限と鍵を本番から切り離す。

A.6 用語

用語	意味
IdP	利用者の認証とID連携を担う基盤
OAuth	アプリにアクセス権を委任する仕組み。本人認証とは別
スコープ	アプリが要求するアクセス権の範囲
WAF	Webアプリへの通信を検査・遮断する仕組み
CVE	公表された脆弱性に付与される識別番号
CNA	CVEを採番し公表する組織（ソフトウェアの開発元など）
CVSS	脆弱性の深刻度を0～10で評価する共通の基準。9.0以上がCritical
CWE	脆弱性の種類の分類
SBOM	ソフトウェアを構成する部品とその版の一覧
VEX	SBOMに含まれる脆弱性が、その製品に実際に影響するかを示す情報
仮想パッチ	ソフトウェアを修正する前に、WAFなどで攻撃の通信を遮断すること
攻撃面	攻撃者が悪用を試みることのできる入口の広さ
サービスアカウント	人ではなくプログラムが使うアカウント
サプライチェーン攻撃	利用しているソフトウェアや事業者を侵害し、その利用者に被害を広げる攻撃
ロックファイル	依存するパッケージの版を固定して記録するファイル
SSO	1回の認証で複数のサービスを利用できる仕組み
CI	コードの変更ごとに自動でビルドとテストを行う仕組み
武器化	公表された脆弱性を、実際に悪用できる攻撃コードにすること
横展開	侵入した攻撃者が、社内の他のシステムへ範囲を広げること
KEV	実際に悪用が確認された脆弱性の一覧（米CISA等）
MCP	AIアプリと外部ツールを接続するプロトコル
間接プロンプトインジェクション	AIが読み込む文書に埋め込まれた指示で、AIの操作を誘導する攻撃

付録B 本稿の前提と限界

- 国内の事案とAIの関係は、JPCERT/CCが調査した攻撃インフラの観測に限って扱う。[R01]トレンドマイクロは、国内事案とAI利用の関連を確認していないと明記している。[R03]
- 第1章の能力評価は管理された条件下の結果であり、能力の目安として扱う。[R04][R05]
- 図2（付録D）の数値は測定方法が異なり、平均値、中央値、最速の事例を含む。図2は傾向を示すものとして扱う。
- 図8の国内事案は、GenerativeXが収集できた公表に限る。10月8日の注意喚起の後は、公表と報道そのものが増えた可能性がある。
- CVEの公表件数は、CVEを採番する組織（CNA）の増加にも影響される。
- SaaSの安全性は、国籍や企業規模ではなく、更新の速さと付与した権限で判断する。
- 安全の根拠には、単一の対策ではなく、停止までの時間と影響範囲の検証結果を用いる。
- 本稿の施策は被害の大きさを小さくするためのものである。物理的な盗難、内部不正、業界固有の規制は対象外とする。

付録C 既存の枠組みとの関係

本稿の要素	参照した枠組み
経営の責任と全体の管理	NIST CSF 2.0 [R10]
資源単位の認可	NIST SP 800-207（ゼロトラスト）[R11]
修正の反映と開発工程	NIST SP 800-218（SSDF）[R12]
外部サービスと取引先	NIST SP 800-161 [R13]
検知、対応、復旧	NIST SP 800-61 [R14]
認証	NIST SP 800-63B [R26]
OAuth	RFC 9700 [R17]
AIエージェント	OWASP [R18][R19][R20]、Microsoft [R22]

付録D 図のデータと集計方法

図2 脆弱性の公表から悪用までの時間

脆弱性の公表から悪用までの時間。測定対象と方法は出典ごとに異なり、平均値、中央値、最速の事例、管理条件下の能力評価を含む。

年	数値	測定対象	出典
2018～ 2019	平均 63日	公表から最初の悪 用観測まで	Mandiant [R27]
2020～ 2021初	平均 44日	同上	Mandiant [R27]
2021～ 2022	平均 32日	同上	Mandiant [R27]
2022	中央値 1日。56%が7日以内に悪用	同社が追跡した悪 用済み脆弱性	Rapid7 [R34]
2022	平均 19.5日で武器化。修正は平均30.6日	武器化された脆弱 性	Qualys [R35]
2022	公表から15分以内にスキャン開始。F5 BIG-IPで は10時間で2,552回の攻撃・スキャン	同社の観測	Palo Alto Networks Unit 42 [R32]
2023	平均 5日	公表から最初の悪 用観測まで	Mandiant [R27]
2024	実証コードの公開から22分で悪用試行 (JetBrains TeamCity)	個別事例	Cloudflare [R33]
2024	中央値 0日（エッジ機器）、5日（全体）	公表から大規模悪 用まで	Verizon DBIR 2025 [R36]
2026上半 期	23.43%が公表当日またはそれ以前に悪用	新たに悪用が確認 された495件	VulnCheck [R28]
2026	実環境での悪用 月平均18件（2025年は10.5 件）	同グループの観測	Google Threat Intelligence Group [R29]
2026	公表済み脆弱性から動作する悪用コードまで、人 の作業20分とモデルの処理8時間	能力評価（管理条 件下）	Anthropic [R05]

図2では、Mandiantの平均値を線、実環境での最速事例を点、能力評価を白抜きの点で示した。VulnCheckの「公表当日またはそれ以前」は0日にあたり、対数軸に載らないため注記とした。

図3 CVSSの重大度別にみたCVE公表件数

CVE Programの公式データ（cvelistV5、2026年10月18日時点の全件）から、公表日（datePublished）の四半期ごとに、公表済み（PUBLISHED）のCVEを数えた。[\[R40\]](#)重大度は、CVEを公表した組織（CNA）が付与したCVSS v3.x、なければv4.0、なければCISA（ADP）が付与したCVSS v3.xの評価を使った。NVDが独自に付与する評価は、2024年以降の付与の遅れによって直近の件数が少なく出るため除外した。2022年以前はCVSSの付与がないCVEが大半を占めるため、図3は年をまたいで比較できる2023年から示す。集計の手順と四半期ごとの値は、求めに応じてGenerativeXが提供する。

年	CVE全体	Critical	High	Criticalの前年比
2023	28,847	2,355	8,134	
2024	39,924	3,780	14,107	1.6倍
2025	48,152	4,098	17,348	1.1倍
2026（1～9月）	72,805	8,222	30,033	9か月で2025年の2.0倍

CISA（ADP）の評価は2024年以降に付与が広がったため、Criticalの増加がその影響によるものではないかを、CNAの評価だけで数えて確かめた。CNAの評価に限ったCriticalは、2023年1,350件、2024年2,168件、2025年2,851件、2026年1～9月6,449件で、増加の傾向は同じである（2026年は9か月で2025年の2.3倍）。

図4・図5 Criticalの脆弱性の影響の種類と公表時刻

図3と同じデータとCriticalの判定を使った。影響の種類は、CNAが付与したCWEと英語の説明文から、コード実行、権限昇格、認証回避、SQLインジェクション、情報漏えい、サービス停止、XSSの順に、最初に当てはまったものに分類した。いずれにも当てはまらないものは「その他」とした。公表時刻は、2025年1月から2026年9月に公表されたCritical 12,320件の公表日時（datePublished）を日本時間に換算した。公表時刻は悪用の開始時刻とは一致しない。

図6 悪用と修正の時間差

Verizon DBIR 2025（GreyNoiseによる要約）の中央値をそのまま用いた。[\[R36\]](#)対象は悪用が確認された脆弱性（CISAのKEVカタログ）で、悪用までは公表から大規模悪用まで、修正までは組織が修正を終えるまでの日数である。

図7 Next.jsの週間ダウンロードの版別の内訳

npmが公開する版別のダウンロード数（2026年10月10日～16日、正式版4,077万件）を、Next.jsのセキュリティ勧告の影響範囲と照合した。[\[R43\]](#)[\[R44\]](#)2026年8月25日公表のCritical（GHSA-2xp9-vwfh-vxw4）は10.0.0以上15.5.24未満と16.0.0以上16.3.3未満、9月22日公表のCritical（GHSA-vcvr-r3jv-pc5j）は16.2.0以上16.3.6未満に影響を受ける版とした。ダウンロード数は開発環境やCIでの取得を含み、稼働中のシステムの数とは一致しない。

図8 国内の漏えいの公表件数

2026年8月31日から10月16日までに国内の組織が公表した、不正アクセス等による漏えい（可能性を含む）を、各組織の公表と報道からGenerativeXが収集した（153件）。誤送信と紛失は対象外とした。1つの委託先の侵害を複数の委託元が公表した場合は、委託元ごとに数えた。原因は公表資料に記載されたものだけを分類し、記載のないものは「不明・調査中」とした。収集できた公表に限るため、実際の件数はこれを上回る可能性がある。

図9 ランサムウェアのリークサイトに載った組織の数

ransomware.liveの公開API（2026年10月19日取得）から、月別・国別の掲載件数を集計した。[\[R51\]](#)月はransomware.liveが推定した攻撃日、国はransomware.liveの推定による。リークサイトに載るのは身代金の支払いを拒んだ組織などに限られ、ランサムウェアの被害の一部である。

図10 1つの事業者の侵害で影響を受けた組織の数

各事案の公表時点の値を用いた。[\[R47\]](#)[\[R48\]](#)[\[R49\]](#)[\[R50\]](#)Salesloft DriftとGainsightは、Googleが影響を受けた可能性があるとした数である。Gainsight自身は、データが影響を受けた顧客は少数と説明している。イセトーは、個人情報保護委員会の資料にある再委託元を含む団体数である。

付録E 出典

閲覧日は2026年10月18日（R44、R45～R55は10月19日）。

- [R01] JPCERT/CC 「連続不正アクセス事案の攻撃手法に関する調査」 2026-10-15
https://blogs.jpccert.or.jp/ja/2026/10/web_api_attacks.html
- [R02] Gartner 「多発する日本企業の不正アクセス事案に関する緊急的な提言」 2026-10-13
<https://www.gartner.co.jp/ja/newsroom/press-releases/2026-10-13-gartner-issues-urgent-recommendations-on-unauthorized-access-incidents-at-japanese-enterprises>
- [R03] トレンドマイクロ 「2026年の日本でのインシデント急増を受け、法人組織が行うべきガイダンスを公開」 2026-10-13
https://www.trendmicro.com/ja_jp/about/newsroom/press-releases/2026/pr-20261013-01.html
- [R04] NIST CAISI 「CAISI's Assessment of Z.ai's GLM-5.3 Cyber Capabilities」 2026-09-17
<https://www.nist.gov/news-events/news/2026/09/caisis-assessment-zais-glm-53-cyber-capabilities>
- [R05] Anthropic 「GLM-5.3 and the spread of advanced cyber capabilities」 2026-09-29
<https://www.anthropic.com/research/glm-5-3-and-the-spread-of-advanced-cyber-capabilities>
- [R06] Anthropic 「Disrupting the first reported AI-orchestrated cyber espionage campaign」 2025-11-13
<https://www.anthropic.com/research/disrupting-AI-espionage>
- [R07] Google Threat Intelligence Group 「Widespread Data Theft Targets Salesforce Instances via Salesloft Drift」 2025-08-26
<https://cloud.google.com/blog/topics/threat-intelligence/data-theft-salesforce-instances-via-salesloft-drift/>
- [R08] Next.js 「Next.js Security Update for Upstream Vulnerabilities」 2026-10-14
<https://nextjs.org/blog/nextjs-security-update-october-14-2026>
- [R09] Next.js 「How to create a static export of your Next.js application」
<https://nextjs.org/docs/app/guides/static-exports>
- [R10] NIST 「The NIST Cybersecurity Framework (CSF) 2.0」 2024-02-26
<https://csrc.nist.gov/pubs/cswp/29/the-nist-cybersecurity-framework-csf-20/final>
- [R11] NIST 「SP 800-207: Zero Trust Architecture」 2020-08
<https://csrc.nist.gov/pubs/sp/800/207/final>
- [R12] NIST 「SP 800-218: Secure Software Development Framework v1.1」 2022-02
<https://csrc.nist.gov/pubs/sp/800/218/final>
- [R13] NIST 「SP 800-161 Rev. 1: Cybersecurity Supply Chain Risk Management Practices」 2022-05
<https://csrc.nist.gov/pubs/sp/800/161/r1/upd1/final>
- [R14] NIST 「SP 800-61 Rev. 3: Incident Response Recommendations」 2025-04-03
<https://csrc.nist.gov/pubs/sp/800/61/r3/final>
- [R15] NICT CYNEX 「相次ぐ情報漏洩、その原因の1つは「アプリ解析からのサーバ攻撃」」 2026-10-13
<https://cynex.nict.go.jp/library/261013-01.html>
- [R16] Google Workspace 「Control which third-party & internal apps access Google Workspace data」
<https://knowledge.workspace.google.com/admin/apps/control-which-third-party-and-internal-apps-access-google-workspace-data>

- [R17] IETF 「RFC 9700: Best Current Practice for OAuth 2.0 Security」 2025-01
<https://www.rfc-editor.org/rfc/rfc9700.html>
- [R18] OWASP 「LLM01:2025 Prompt Injection」
<https://genai.owasp.org/llmrisk/llm01-prompt-injection/>
- [R19] OWASP 「LLM06:2025 Excessive Agency」
<https://genai.owasp.org/llmrisk/llm062025-excessive-agency/>
- [R20] OWASP 「AI Agent Security Cheat Sheet」
https://cheatsheetseries.owasp.org/cheatsheets/AI_Agent_Security_Cheat_Sheet.html
- [R21] Anthropic 「Beyond permission prompts: making Claude Code more secure and autonomous」 2025-10-20
<https://www.anthropic.com/engineering/claude-code-sandboxing>
- [R22] Microsoft 「エージェントの脅威モデル」
<https://learn.microsoft.com/ja-jp/agents/architecture/threat-models>
- [R23] Google 「Learn about sharing your data with third-party apps & services」
<https://support.google.com/accounts/answer/14013600>
- [R24] JPCERT/CC 「直近で相次いでいる国内組織における不正アクセスに関する注意喚起」 2026-10-08
<https://www.jpcert.or.jp/at/2026/at260030.html>
- [R25] Google 「Manage connections between your Google Account & third parties」
<https://support.google.com/accounts/answer/13533235?hl=en>
- [R26] NIST 「SP 800-63B-4: Authentication and Authenticator Management」 2025-07-31
<https://csrc.nist.gov/pubs/sp/800/63/b/4/final>
- [R27] Mandiant 「How Low Can You Go? An Analysis of 2023 Time-to-Exploit Trends」 2024-10-16
<https://cloud.google.com/blog/topics/threat-intelligence/time-to-exploit-trends-2023>
- [R28] VulnCheck 「State of Exploitation 1H 2026」 2026-07-28
<https://www.vulncheck.com/blog/state-of-exploitation-1h-2026>
- [R29] Google Threat Intelligence Group 「Vulnerability Discovery and Exploitation Trends in the AI Era」 2026-10-01
<https://cloud.google.com/blog/topics/threat-intelligence/vulnerability-discovery-and-exploitation-trends-in-the-ai-era>
- [R30] R. Fang ほか 「LLM Agents can Autonomously Exploit One-day Vulnerabilities」 arXiv, 2024-04
<https://arxiv.org/abs/2404.08144>
- [R31] The Register 「Security researchers spent 6 days doing their Chromework, built 1B user exploit」 2026-10-14 (XBOWの公表に関する報道)
<https://www.theregister.com/security/2026/10/14/security-researchers-spent-6-days-doing-their-chromework-built-1b-user-exploit/5303134>
- [R32] The Register 「Palo Alto Networks: Attackers scan for vulnerabilities within 15 minutes of disclosure」 2022-07-27 (Unit 42の報告に関する報道)
https://www.theregister.com/2022/07/27/palo_alto_unit_42/
- [R33] Cloudflare 「Application Security Report: 2024 update」 2024-07-11
<https://blog.cloudflare.com/application-security-report-2024-update/>
- [R34] Rapid7 「2022 Vulnerability Intelligence Report」 2023-02-28
<https://investors.rapid7.com/news/news-details/2023/Rapid7-Vulnerability-Intelligence-Report-Shows-Attackers-Developing-and-Deploying-Exploits-Faster-Than-Ever/>

- [R35] Qualys 「2023 TruRisk Research Report」 2023
<https://www.qualys.com/2023-trurisk-research-report>
- [R36] GreyNoise 「Verizon DBIR 2025: Edge KEVs Increasingly Left Unpatched, Exploited」 2025-05-06 (Verizon DBIR 2025の要約)
<https://www.greynoise.io/blog/verizon-dbir-2025-edge-kevs-increasingly-left-unpatched-exploited>
- [R37] 大阪急性期・総合医療センター 情報セキュリティインシデント調査委員会 「情報セキュリティインシデント調査報告書 概要」 2023-03-28
https://www.gh.opho.jp/pdf/reportgaiyo_v01.pdf
- [R38] Recorded Future (Insikt Group) 「H1 2026 Malware and Vulnerability Trends」 2026
<https://www.recordedfuture.com/research/h1-2026-malware-vulnerability-trends>
- [R39] curl 「CVE-2026-92392: HTTP/2 upload Use-After-Free」 2026-10-14
<https://curl.se/docs/CVE-2026-92392.html>
- [R40] CVE Program 「cvelistV5」 2026-10-18時点の全件スナップショット (GenerativeXが公表日とCNAのCVSS評価で集計)
<https://github.com/CVEProject/cvelistV5>
- [R41] 経済産業省 「ソフトウェア管理に向けたSBOM (Software Bill of Materials) の導入に関する手引 ver2.0」 2024-08-29
<https://www.meti.go.jp/policy/netsecurity/wg1/SBOMv2.pdf>
- [R42] CISA 「BOD 26-04: Prioritizing Security Updates Based on Risk」 2026-06-10
<https://www.cisa.gov/news-events/directives/bod-26-04-prioritizing-security-updates-based-risk>
- [R43] npm 「Download counts by version: next (last-week)」 2026-10-10～10-16 (GenerativeXが版ごとに集計)
<https://api.npmjs.org/versions/next/last-week>
- [R44] Vercel 「Next.js Security Advisories」 2026-10-18閲覧 (各脆弱性の影響範囲と修正版)
<https://github.com/vercel/next.js/security/advisories>
- [R45] curl 「curl」 2026-10-19閲覧 (「over twenty billion installations」)
<https://curl.se/>
- [R46] Verizon 「2025 Data Breach Investigations Report」 2025-04-23 (第三者が関与した侵害の割合)
<https://www.verizon.com/about/news/2025-data-breach-investigations-report>
- [R47] CyberScoop 「Salesloft Drift compromise scope expands」 2025-08-28 (影響を受けた可能性がある組織は700超とするGoogleの見解)
<https://cyberscoop.com/salesloft-drift-compromise-scope-expands/>
- [R48] The Register 「Gainsight CEO downplays breach, says only a 'handful' of customers had data stolen」 2025-11-26 (Googleは200超のインスタンス、Gainsightは少数と説明)
https://www.theregister.com/2025/11/26/gainsight_ceos_handful_customers_data_stolen/
- [R49] Cloud Security Alliance 「ShinyHunters' OAuth Pivot: A Year of SaaS Supply-Chain Breaches」 2026-07-16 (Clueが顧客に通知した影響組織数195)
<https://labs.cloudsecurityalliance.org/research/csa-research-note-shinyhunters-oauth-saas-abuse-20260716-csa/>
- [R50] 個人情報保護委員会 「株式会社イセトーに対する個人情報の保護に関する法律に基づく行政上の対応について」 2025-03-19
https://www.ppc.go.jp/files/pdf/250319_shiryou-2.pdf
- [R51] ransomware.live 「API v2」 2026-10-19取得 (国別・月別のリークサイト掲載件数をGenerativeXが集計)
<https://api.ransomware.live/v2>

- [R52]** 個人情報保護委員会「令和8年度第1四半期の漏えい等報告の処理件数」2026-09-02
https://www.ppc.go.jp/files/pdf/260902quarter-report_roueihoukoku.pdf
- [R53]** Klue「An update on recent Klue security incident」2026-06-18
<https://klue.com/blog/an-update-on-recent-klue-security-incident>
- [R54]** CISA「Minimum Requirements for Vulnerability Exploitability eXchange (VEX)」2023-04-21
<https://www.cisa.gov/resources-tools/resources/minimum-requirements-vulnerability-exploitability-exchange-vex>
- [R55]** CISA「Widespread Supply Chain Compromise Impacting npm Ecosystem」2025-09-23
<https://www.cisa.gov/news-events/alerts/2025/09/23/widespread-supply-chain-compromise-impacting-npm-ecosystem>

侵害を前提としたサイバー防衛の設計

侵入の阻止から、影響範囲の限定と封じ込め速度へ

発行日 2026年10月19日
調査基準日 2026年10月18日
発行 GenerativeX
Web <https://gen-x.co.jp>

本書の内容は調査基準日時点の公開情報に基づく。引用した資料の内容と範囲は付録Eを参照のこと。