GenerativeX

生成AIによるビジネスの変革を支援する

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

エグゼクティブ・サマリー

最速では公表から数十分で悪用される今、防衛は影響範囲の限定と停止の速さで設計すべきである。

現状

2026年9月末から国内で漏えいの公表が相次ぎ、10月の2週間だけで**101件**に達した。
原因を公表したものの**過半**は委託先と外部連携である。

背景

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。
Criticalの公表件数は、2026年は1〜9月だけで前年1年間の2倍に達した。

推奨

被害の大きさは、**入口の数・到達できる範囲・停止までの時間**の掛け算で表せる（概念モデル）。
入口・範囲・時間を小さくするため、次の3点に直ちに着手すべきである。

GenerativeXの提言

① 依存ライブラリは、悪用確認済み・外部公開の重大な脆弱性から**すべて更新**し、
公開Webサービスには緊急遮断の機構を導入すべきである。

② 利用中のSaaSには**SBOM（ソフトウェア部品表）または同等の透明性と、修正期限・通知SLA**を求め、
更新が間に合っていないものは段階的に縮小し、利用を停止すべきである。

③ Entra ID・Google Workspace等に連携済みのアプリは、**権限を最小化し、新規の連携を承認制**にすべきである。

3点は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日で実行する。

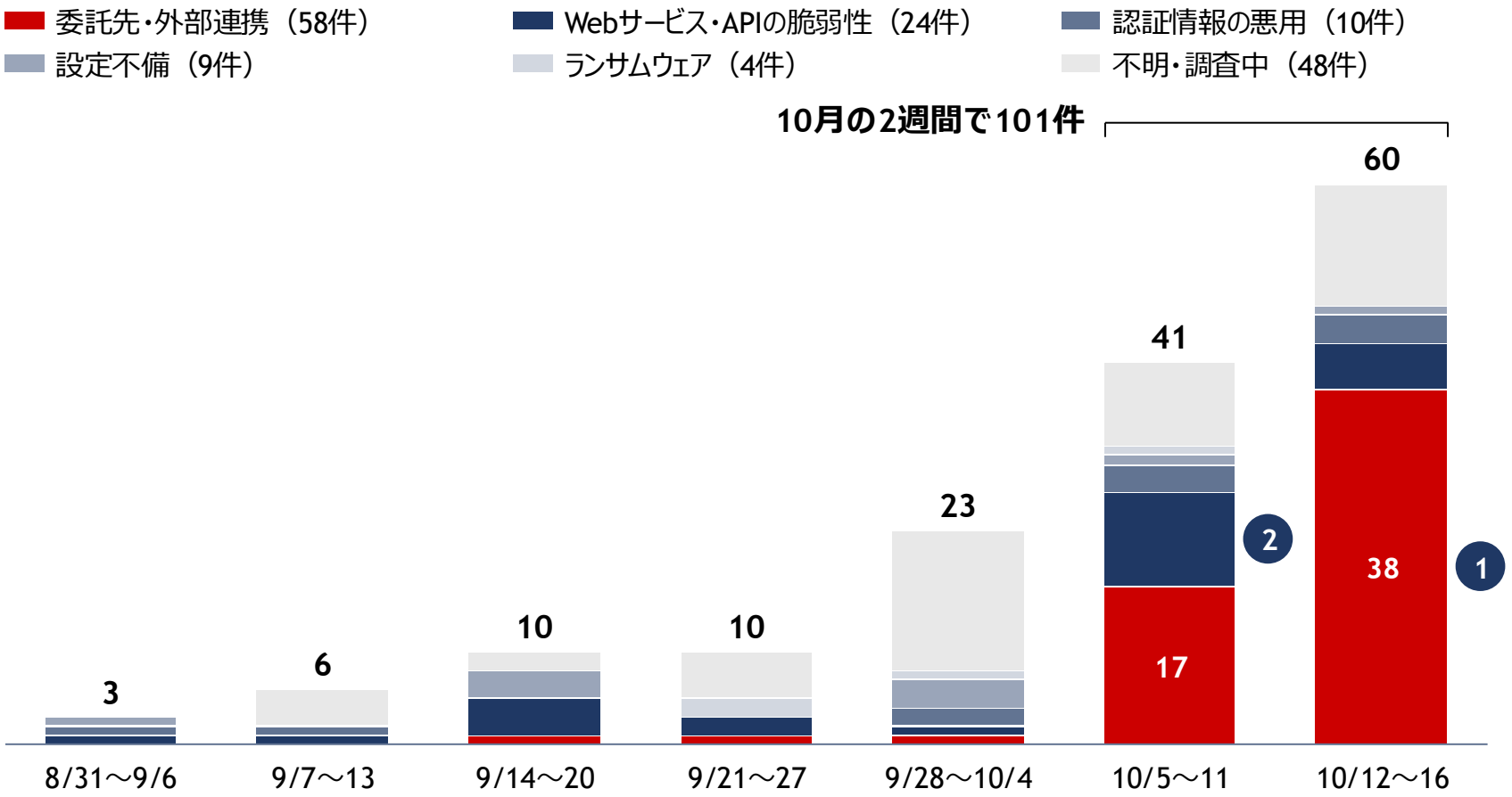
アジェンダ

1. **現状：攻撃の速さと被害の経路**
2. 提言：被害の3要素と緊急提言
3. 継続的な対策と経営の判断
4. 付録：AIエージェントと補足資料

国内の漏えいの公表は9月末から急増し、10月の2週間で101件に達した

原因を公表した105件の55%は委託先・外部連携で、うち約40件は5つの事業者の侵害を各委託元が公表したもの。

国内の組織による漏えい（可能性を含む）の公表件数（公表日の週別・原因別、計153件）；件



1. 対象は10月16日までの公表。1つの委託先の侵害を複数の委託元が公表した場合は、委託元ごとに数える
2. 原因は公表資料の記載で分類し、誤送信・紛失は対象外。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

資料：各組織の公表、報道をGenerativeXが収集（2026年10月19日時点）、JPCERT/CC (2026)

主なインサイト

1. **原因を公表した105件のうち58件（55%）が委託先・外部連携**
   うち**約40件**は、防災メール配信、クラウド事業者、Web制作、問い合わせ管理、コミュニティ基盤の**5事業者の侵害**を各委託元が公表

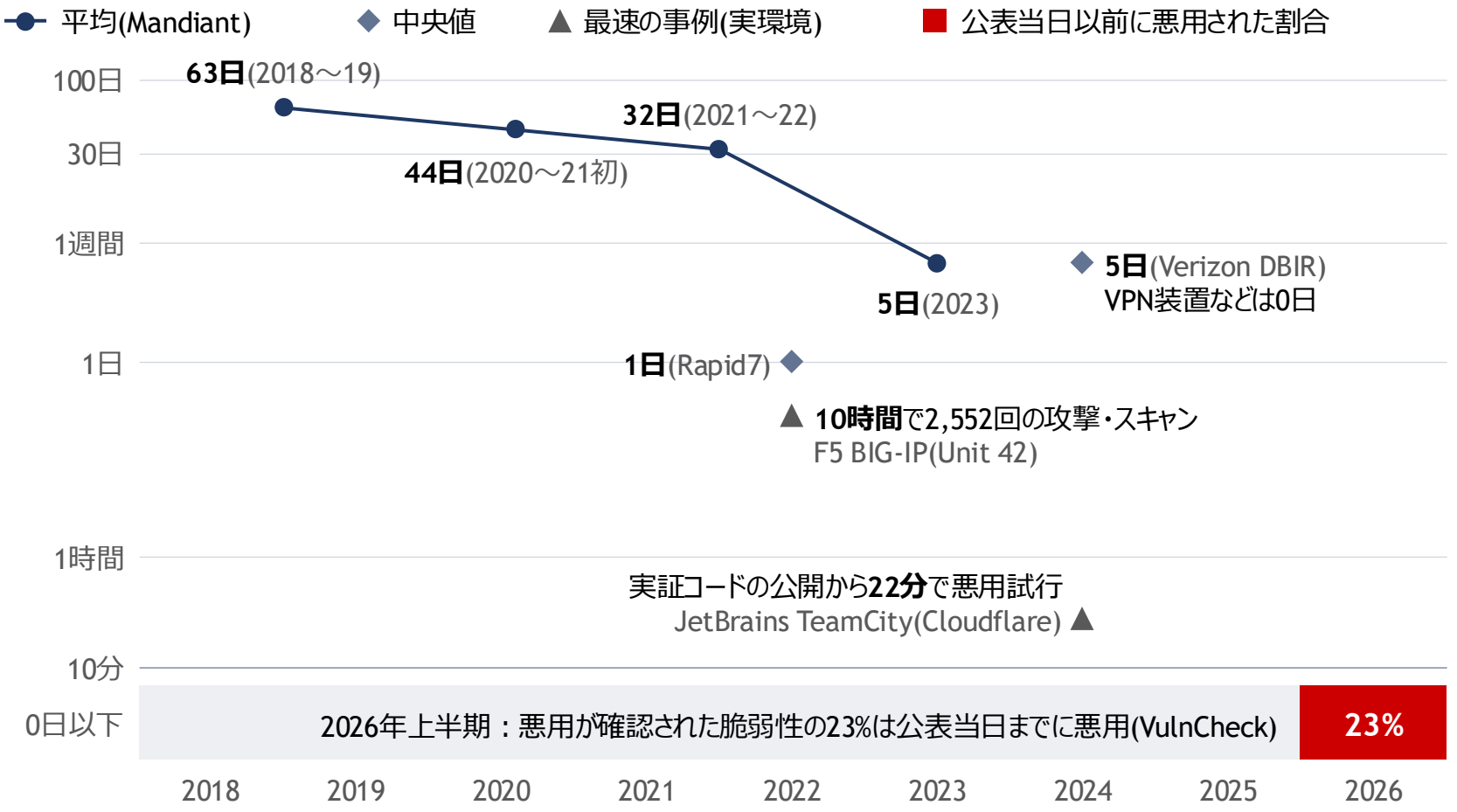
2. **Webサービス・APIの脆弱性が24件**
   APIの不正操作、SQLインジェクション、認可の不備。**JPCERT/CCの注意喚起**が挙げた点検項目と一致

打ち手： 委託先と連携先を台帳で管理し、侵害時の停止手段を事前に定める（提言②③）

2026年上半期に悪用が確認された脆弱性の23%は公表当日までに悪用され、最速は実証コードの公開から22分

同じ方法で測ったMandiantの平均も63日から5日に短縮し、修正を待たずに停止する手段が必要である。

公表から悪用までの時間（印の種類ごとに出典と測定方法が異なる）；対数目盛



1. 平均は公表から最初の悪用観測まで、Verizon DBIRの中央値は公表から大規模悪用まで、TeamCityは実証コードの公開から
2. VPN装置などの0日は、公表当日以前に大規模悪用が始まっていたことを示す。対数目盛の外のため数値で表示

資料: Mandiant (2024)、Rapid7 (2023)、Palo Alto Networks Unit 42 (2022)、Cloudflare (2024)、Verizon DBIR (2025)、VulnCheck (2026)、Google Threat Intelligence Group (2026)、Anthropic (2026)

主なインサイト

1 悪用の過半は1週間以内

2022年に悪用された脆弱性のうち、**56%は公表から7日以内**に悪用(Rapid7)

2 悪用件数は月18件に増加

実環境で悪用された脆弱性は**2026年に月平均18件**。2025年は10.5件(Google Threat Intelligence Group)

3 悪用コードの作成は8時間

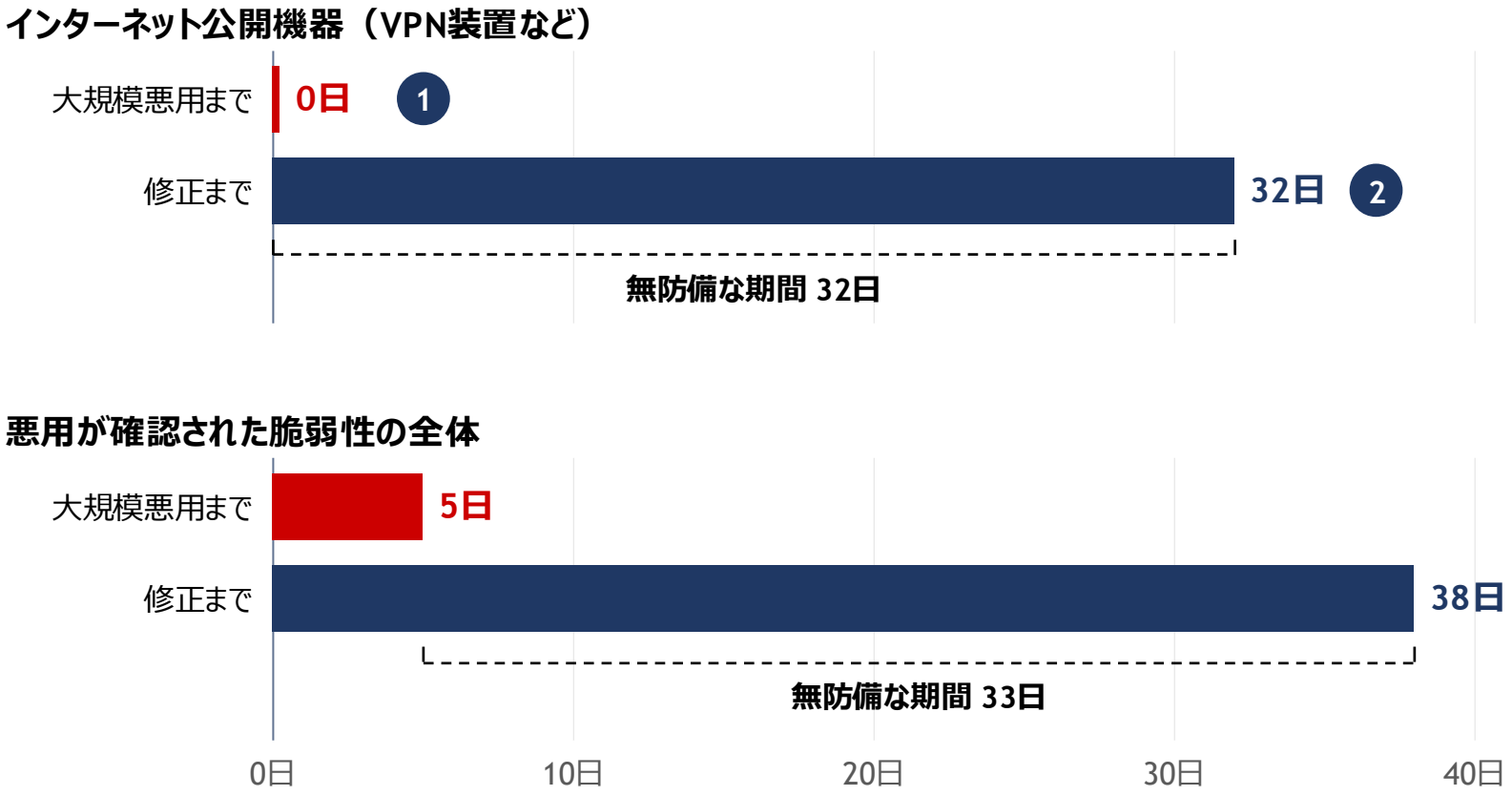
公表済みの脆弱性から動作する悪用コードを、**人の作業20分とモデルの処理8時間**で作成(Anthropic の評価)

打ち手: 公表を検知した時点で、修正を待たずに公開サービスを停止できるようにすべきである

大規模悪用は公表から0〜5日で始まり、修正には32〜38日かかる

大規模悪用の開始から修正の完了まで約1か月の差があり、修正を待たずに停止する仕組みが必要である。

公表から大規模悪用までと、修正を終えるまでの日数（中央値）；日



1. いずれも中央値。大規模悪用までは公表から、修正までは組織が修正を終えるまで。対象はKEV（米CISAの悪用確認済み脆弱性カタログ）の脆弱性
資料: Verizon「2025 Data Breach Investigations Report」

主なインサイト

1 大規模悪用の開始は公表から0〜5日

全体で公表から**5日**、VPN装置などインターネットに直接公開している機器では**公表当日以前**に大規模悪用が始まる

2 修正の完了は公表から32〜38日

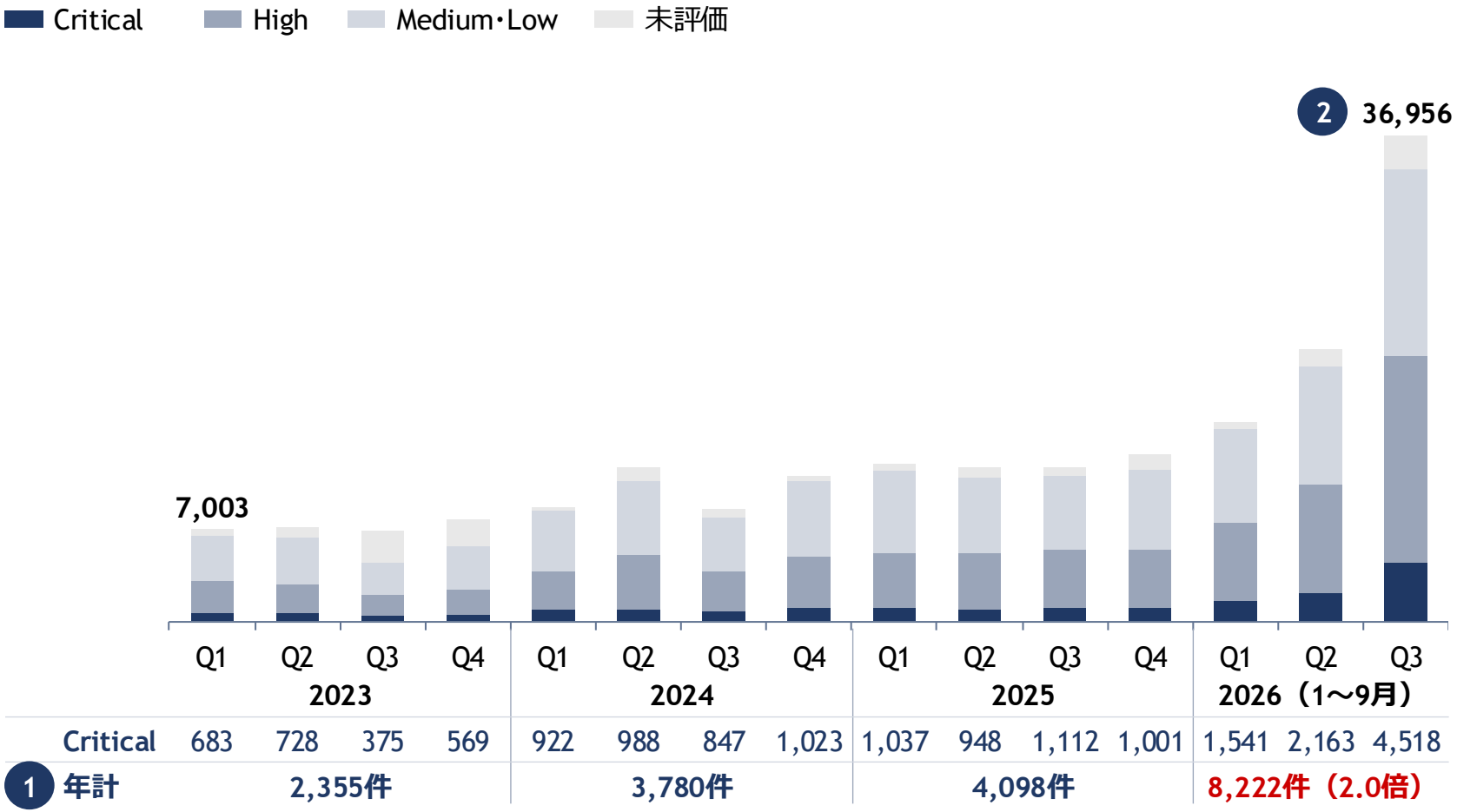
組織が修正を終えるまで全体で**38日**、公開機器でも**32日**。この差を埋めるには、修正の迅速化と並行して停止の手段が必要

打ち手: 修正を待たずに停止できる手段を用意し、SaaS事業者には修正期限の短縮を求める

Criticalと評価された脆弱性の公表件数は、2026年に前年の2倍を超えた

2026年は1～9月だけでCriticalが8,222件公表され、2025年1年間の4,098件の2倍に達した。

四半期ごとのCVE公表件数（CVSSの重大度別）；件



		2023				2024				2025				2026（1～9月）		
		Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3
	Critical	683	728	375	569	922	988	847	1,023	1,037	948	1,112	1,001	1,541	2,163	4,518
1	年計	2,355件				3,780件				4,098件				8,222件（2.0倍）		

1. 重大度はCVEを公表した組織（CNA）が付与したCVSSの評価（なければCISAの評価）。Criticalは9.0以上。CNAの評価がそろう2023年以降を表示
資料: CVE Program cvelistV5（2026-10-18時点）をGenerativeXが集計、Google Threat Intelligence Group (2026)

主なインサイト

1 CNAの評価だけでも2.3倍

評価の付与方法の変化の影響を除くため、CVEを公表した組織（CNA）の評価だけで数えても、**2026年1～9月は2025年の2.3倍**

2 CVE全体も5倍超

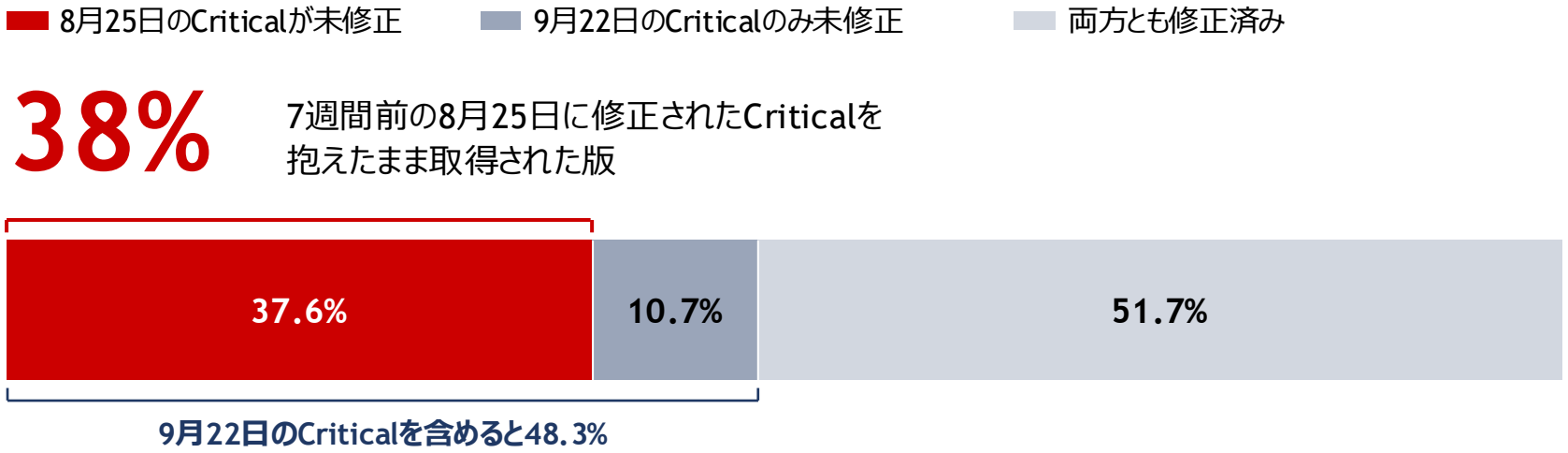
四半期の公表件数は**7,003件から36,956件**に増加。Google Threat Intelligence Groupも、2026年1～8月に月あたりの公表件数の倍増を報告

打ち手: 稼働中の版の照合、影響の判定、修正の作成を自動化し、人は停止と承認の判断に集中する

Next.jsのダウンロードの38%は、7週間前に修正された重大な脆弱性を抱えた版

修正版の公開後も更新は遅れ、curlのように広く使われる部品では影響を受けうる環境が200億を超える。

Next.jsの週間ダウンロードの版別内訳（2026年10月10日～16日、正式版4,077万件）；%



修正版の公開から集計週までの経過



1. 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別の指標
資料: npm registry（版別ダウンロード、2026-10-10～16）、Next.jsセキュリティ勧告（GitHub）、curl (2026)

影響を受けうる環境の規模（curlの例）

200億超

curlが使われている環境の数（curlの公表）

2026年10月14日

curlが、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た脆弱性を修正

利用者が意識しないまま影響を受けうる

サーバー、端末、組み込み機器の多くが対象

打ち手: 修正版を適用できない公開サービスは停止し（提言①）、更新が遅れているSaaSは利用を停止する（提言②）

閉域網もSSOも、侵入の入口を残している

閉域網は外部との接続点の機器から侵入され、悪用された脆弱性の68%は認証なしで攻撃に使える。

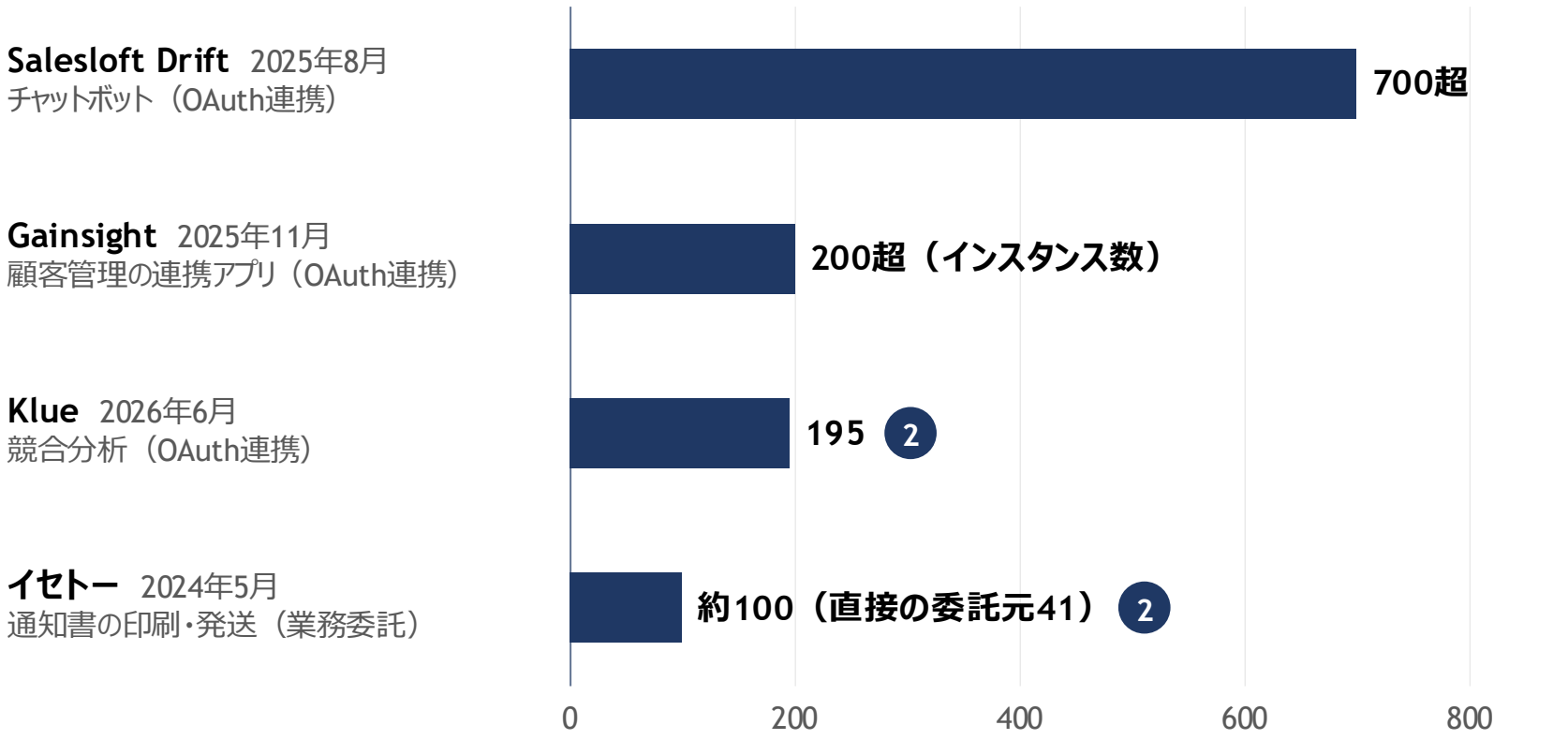
	「閉域網に置いているから大丈夫」	「SSOと多要素認証を入れているから大丈夫」
よくある判断		
事例	**大阪急性期・総合医療センター（2022年10月）**	**Next.jsとcurlの修正（2026年10月14〜15日）**
	給食事業者 **VPN機器** 侵入の入口 → 閉域網 → 病院 **給食サーバー** 踏み台 → **基幹システム** 暗号化	**利用者** → **SSO・多要素認証** → **公開サービス** コードを実行 **攻撃者** → 認証なしで脆弱性を悪用 → **公開サービス**
	電子カルテを含む基幹システムの再稼働まで**43日**、診療システム全体の復旧まで**73日**。逸失利益は**十数億円以上**（見込み）	**Next.js** 認証なしでリモートからコードを実行できる脆弱性2件（Critical） **curl** 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
統計	2026年1〜8月に悪用された脆弱性のうち、VPN装置やファイアウォールなどの割合	2026年上半期に悪用された脆弱性215件のうち、認証なしで悪用できた件数
	14%	**146件（68%）**
		リモートでコードを実行できる82件のうち、60件が認証なしで悪用可能
打ち手	**閉域網の内側も操作ごとに認可し、外部との接続点の機器を最優先で更新する**	**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**

資料: 大阪急性期・総合医療センター調査委員会 (2023)、Google Threat Intelligence Group (2026)、Recorded Future (2026)、Next.js (2026)、curl (2026)

1つの事業者の侵害で影響を受けた（可能性がある）組織・インスタンスは、100から700超

侵害されたのは権限やデータを預けた先であり、データ侵害に第三者が関与する割合は30%に倍増した。

事業者1社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; 組織・インスタンス



1. いずれも公表時点の値。Drift・Gainsightは影響を受けた可能性がある数で、Gainsight自身は影響を受けた顧客を少数と説明
2. イセトーは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（漏えいした個人データは約308万人分）

資料: Google Threat Intelligence Group（CyberScoop 2025、The Register 2025）、Cloud Security Alliance (2026)、Klue (2026)、個人情報保護委員会 (2025)、Verizon「2025 Data Breach Investigations Report」

主なインサイト

1 **第三者が関与した侵害は30%**
Verizonの調査で、データ侵害のうち第三者が関与した割合は**前年から倍増**

2 **起点は休眠した認証情報と旧版のVPN機器**
Klueは**2022年に作成され休眠していた認証情報**、イセトーは**2021年4月の版のまま稼働したVPN機器**が侵入口

打ち手: 連携済みアプリの権限を最小化し、預けるデータを絞る（提言③）

アジェンダ

1.現状：攻撃の速さと被害の経路

2.提言：被害の3要素と緊急提言

3.継続的な対策と経営の判断

4.付録：AIエージェントと補足資料

被害の大きさは、入口の数・到達できる範囲・停止までの時間の掛け算で整理できる

掛け算は被害を左右する要素を整理した概念モデルであり、緊急提言と4つの柱はいずれかの要素を小さくする。

	入口の数	×	到達できる範囲	×	停止までの時間	=	被害の大きさ
具体例	悪用されうる入口（攻撃面） 公開API・管理画面・ID・ 外部連携・VPN装置		侵害1件から、データ・ クラウド管理基盤・CI/CD・ サービスアカウントまで		修正の完了まで32～38日 → 緊急遮断まで数分		**すべて 小さくすれば 被害は大きく縮む**
緊急提言	**① 全ライブラリの更新**		**③ 連携権限の最小化**		**① 緊急遮断の機構の導入**		
	② 更新が遅いSaaSの利用停止						
4つの柱	**② 削減する**		**③ 影響範囲を限定する**		**④ 停止までの時間を短縮する**		
	① 把握する（他の3つの柱の前提）						

1. 32～38日は、悪用が確認された脆弱性を組織が修正し終えるまでの中央値（インターネット公開機器32日、全体38日）
資料: Verizon「2025 Data Breach Investigations Report」

緊急提言は、全ライブラリの更新と緊急遮断・更新が遅いSaaSの利用停止・連携権限の最小化の3点

3点はいずれも業務の停止を伴うため、経営の判断で全社に実施すべきである。

提言	実施すること	完了の基準
1 全ライブラリを更新し、緊急遮断の機構を導入する **根拠** Next.jsは修正の7週間後も38%が古い版	• 悪用確認済み・外部公開の重大な脆弱性から解消し、全ライブラリを計画的に更新 • 公開Webサービスに緊急遮断の機構を導入	**悪用確認済みの脆弱性を解消済みで、夜間でも数分で遮断できる**
2 更新が間に合っていないSaaSは利用を停止する **根拠** 1社の侵害で100～700超の組織に影響	• 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 • 申告は第三者の脆弱性診断と認証の報告書で裏付ける	**主要なSaaSの部品と修正期限を把握し、応じない事業者の利用を縮小・停止している**
3 連携アプリの権限を、必要最小限に絞る **根拠** 原因を公表した漏えいの55%が委託先・外部連携	• Entra ID・Google Workspace等の連携アプリとスコープを一覧にする • 所有者が不明なもの、権限が過大なものを失効する	**新規の連携を管理者が審査し、許可に期限を設けている**

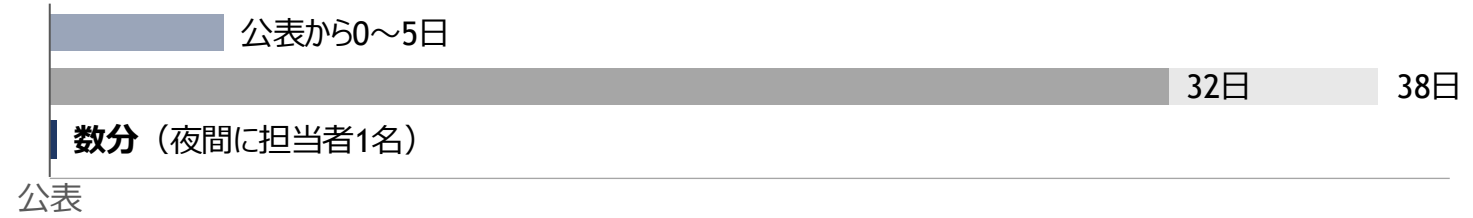
資料: npm・Next.js (2026)、Google Threat Intelligence Group ほか、GenerativeX集計

提言①：依存ライブラリを全面的に更新し、公開Webサービスに緊急遮断の機構を導入する

最優先はライブラリの更新であり、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すべきである。

実施の順序：リスクの高い脆弱性から解消して全ライブラリを更新し、続いて公開Webサービスに緊急遮断の機構を導入する
責任者：セキュリティ責任者、サービス・システム責任者 ／ **期限：**72時間で着手と悪用確認済みの解消、30日で全ライブラリの更新と緊急遮断の機構

実施の項目	実施すること	完了の基準
① **全ライブラリの更新**	VPN装置等のファームウェアを含めて一覧にし、悪用確認済み・外部公開の重大な脆弱性から解消、残りもサポート中の系列の最新の修正版へ更新する 更新経路も攻撃される（2025年9月、npmの500超のパッケージが改ざん）。版はロックファイルで確定する	悪用確認済みの脆弱性を解消済み。全ライブラリを回帰テストを経て更新済み
② **緊急遮断の機構**	DNS・CDN・ロードバランサーで遮断する機構を導入し、遮断の決定者と実行者を事前に指名する	夜間に担当者1名で、数分で遮断できる
③ **直ちに修正できない場合**	WAFによる仮想パッチ、該当機能の停止、アクセス元の制限、権限・アクセスキーの失効、サービス全体の停止の順に検討する	残るリスクを、許容できる水準まで下げている



リスクを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

資料: CISA BOD 26-04 (2026)、CISA (2025)、Next.js (2026)

提言①の実践例：GenerativeXは公表から22分で修正を自動作成し、26分で57ホストを遮断

検知と修正の作成は自動で進み、57ホストの遮断は当番の担当者が修正の完了前に判断した。

GenerativeXの対応事例　2026年10月15日（日本時間）Next.jsの緊急更新

	02:47 公表	02:48 検知	03:09 修正	03:13 遮断	04:39 解除
公表から	起点	+1分	+22分	**+26分**	+112分
出来事	Next.jsが緊急更新を公開	296リポジトリを自動照合 148が該当、22が本番	AIエージェントが修正のプルリクエストを自動作成	**WAFで公開中の57ホストを一括遮断**	全更新を終えて遮断を解除
担い手	Next.js開発チーム	GenerativeX 自動化	GenerativeX 自動化	GenerativeX **当番の判断**	GenerativeX 担当者

公表の時刻　Next.jsの直近5回の緊急更新は、いずれも日本時間の深夜0～3時に公開

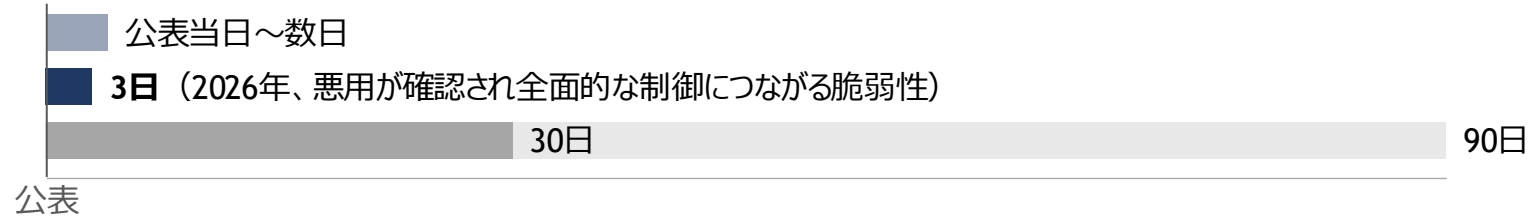
資料: GenerativeXの対応記録（2026年10月15日）、Next.js (2026)

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るため、事業者の更新の速さを開示と期限で確かめるべきである。

実施の順序：利用中のSaaSを一覧にし、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
責任者：調達・法務、セキュリティ責任者 ／ **期限：**30日以内に主要SaaSへ要求

実施の項目	実施すること	完了の基準
① **SBOMまたは同等の透明性**	部品とその版の一覧（SBOM）か同等の構成情報の開示を求め、影響の有無はVEXで確認する VEX：事業者による影響判定。大規模なSaaSは、影響判定・対応状況の報告・通知SLAで代替できる	主要なSaaSの部品と版、影響の有無を自社で把握している
② **修正スケジュールと通知SLAの確認と短縮**	公表から修正適用までの期限、対応状況の報告、利用者への通知の期限を確認し、短縮を求める	修正の期限と通知SLAを、契約で合意している
③ **第三者による裏付け**	事業者の申告を、第三者の脆弱性診断の結果と第三者認証の報告書で確認する	主要なSaaSの申告を、第三者の資料で確認済み



応じない事業者は、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、改善の期限を過ぎた時点で利用を停止する

資料：経済産業省 (2024)、CISA BOD 26-04 (2026)、CISA VEX (2023)

提言③：Entra ID・Google Workspace等に連携済みのアプリは、権限を最小化し、新規の連携を承認制にする

原因を公表した国内の漏えいの過半は委託先・外部連携であり、連携アプリに付与した権限がそのまま被害の範囲になる。

実施の順序：Entra ID、Google Workspace、Slack、GitHub等の連携アプリを、メールやファイル全体への権限を持つものから見直す
責任者：情報システム（IdP管理者） ／ **期限：**72時間で一覧化と失効に着手、30日で承認制

実施の項目	実施すること	完了の基準
1 一覧化	連携済みの外部アプリについて、 スコープと所有者を一覧にする スコープ：メールの読み取りなど、アプリに許可した操作の範囲	すべての連携に、所有者と用途が 記録されている
2 失効	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 サインイン用の連携を残し、データへのアクセスだけを絞れる	所有者が不明な連携が0件
3 承認制	新規の連携を管理者の承認制とし、 承認に期限を設ける Google Workspaceは、アプリごとにデータへのアクセスを制限できる	期限を過ぎた承認が、 自動で失効する

複製済みのデータは失効後も連携先の管理下にあるため、連携の前に提供するデータの項目と量を絞る

資料：Google Workspace、Google

アジェンダ

1.現状：攻撃の速さと被害の経路

2.提言：被害の3要素と緊急提言

3.継続的な対策と経営の判断

4.付録：AIエージェントと補足資料

対策は、把握・削減・影響範囲の限定・停止までの時間の短縮の4つの柱で実行する

4つの柱を同じ資産と連携に同時に適用し、侵害されても被害を限定し、すぐに停止できる状態をつくる。

	攻撃される対象を減らす		侵害されても被害を抑える	
	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**
答える問い	何を保有し、何に接続し、どの権限を付与しているか	そのうち不要なものは何か	侵害されたとき、影響はどこまで及ぶか	危険な状態を何分で終わらせられるか
最初に実施すること	公開されている入口、恒常的な特権、OAuth連携を洗い出す	所有者が答えられないOAuth連携を失効する	クラウド管理基盤・CI/CD・サービスアカウントの権限を通常業務から分離し、長期有効な認証情報を短期化する	修正版の反映までの停止範囲と判断者を事前に指名する
効果の確認	**所有者不明の項目を未解決として計上**	**重要データへの不要な経路の減少数**	**主体1つの侵害を想定した検証での、API・データベース・ストレージの拒否記録**	**稼働環境で修正版の動作を確認するまでの実測時間**

経営の確認事項を、侵入の有無から影響範囲と停止までの時間に切り替える

経営は侵害を前提に影響範囲と停止までの時間を確認し、停止の判断と報告の体制を平時に整える。

	これまでの確認事項	これから確認すべきこと
侵入	侵入されない対策はできているか	端末・SaaS・AIのいずれかが侵害されたとき、 **影響はどこまで及ぶか**
修正	パッチは適用したか	公表から修正の反映まで**何分か**。 反映までの間、**誰が何を停止するか**
外部サービス	そのSaaSは安全か	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	AIは安全に使えているか	悪意ある指示を読み込んだとき、 **どの操作まで実行できるか**

停止の判断に備え、重要サービスの順位、シナリオ別の対応手順、判断者、再開の確認事項を平時に決定する

停止の判断は侵害の発生前に準備するものであり、経営は平時にシナリオごとの停止と再開の基準を決定すべきである。

1 **重要なサービスとデータ**　サービスとデータを事業への重要度で順位づけ、常時稼働が必要なものと一時停止できるものを区別する

2 **シナリオごとの停止条件**　シナリオ別の対応手順として、停止の基準と単位、停止中の業務の続け方、顧客への連絡を決定する

シナリオ 想定する事象	停止の基準 何が判明したら停止するか	停止の単位 全体・機能・アプリ・トークン・顧客	停止中の業務 手作業への切り替えなど	顧客への連絡 いつ、何を知らせるか
重大な脆弱性の公表 記入例	稼働中の版が対象と判明	該当する公開Webサービス	メンテナンス画面に切り替える	停止と再開を告知する
トークンの侵害 記入例	不正利用を確認	該当するトークンと連携	連携なしの手順に切り替える	影響する顧客に個別連絡
連携先の外部サービスの侵害 記入例	連携先が侵害を公表	侵害された連携先との連携	手作業に切り替える	預けたデータの範囲を確認して連絡する
AIエージェントの承認範囲外の動作 記入例	承認範囲外の操作を検知	該当するタスクとトークン	人の手順に切り替える	外部送信があれば連絡する

3 **判断者と実行者**　夜間・休日を含め、停止を判断する者と実行する者を指名する

4 **再開の条件**　修正版の反映、権限の失効、残存アクセスの解消を確認してから再開する

注：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

経営は5つの指標を月次で報告させ、被害の大きさの推移を確認する

5つの指標は入口の数・到達できる範囲・停止までの時間のいずれかを測り、月次の推移で対策の効果を示す。

	測る要素				
指標	入口	範囲	時間	経営が知るべきこと	目標の例
外部公開資産の把握率	●			未把握の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**

注：入口 = 入口の数、範囲 = 到達できる範囲、時間 = 停止までの時間。目標は例であり、自社の重要サービスと許容できる停止時間に合わせて設定する

推奨する実行計画：経営のコミットメントのもと、トップダウンで72時間・30日・90日の順に実行する

緊急提言はサービスの停止やSaaSの解約を伴うため、経営が判断し、トップダウンで全社に展開すべきである。

	最初の72時間 **更新の着手と悪用確認済みの解消**	30日以内 **更新の完了と緊急遮断**	90日以内 **構成の変更**
経営のコミットメント	**セキュリティ総点検を指示する**	トップダウンで全社に展開する	
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	セキュリティ責任者、調達・法務	サービス・システム責任者、データ責任者
実施すること	• 最優先で、悪用が確認された脆弱性と公開システムから更新に着手する • 更新が困難なものは停止するか公開範囲を限定する • SaaSと連携アプリを一覧にし、所有者不明のものを停止・失効する • 不要な管理画面と休眠中の特権IDを停止する	• 全ライブラリの更新を完了する • 公開Webサービスに緊急遮断の機構を導入する • クラウド・SaaS・契約などを照合し、台帳にまとめる • 重要サービスの実行権限を縮小し、外部共有とOAuth連携を再審査する • SaaS事業者にSBOMの開示と修正期限の短縮を求める	• 顧客・環境ごとの分離など、構成を変更する • 取引先へのデータ提供条件を整える • 影響範囲を主体ごとに検証し、封じ込めの所要時間を実測する • 復旧演習で、バックアップの独立性を確認する • 契約・権限付与・デプロイの手順に台帳の更新を組み込む
成果物	**✓経営の判断記録** **✓ライブラリの更新記録** **✓サービス停止と権限失効の記録**	**✓緊急遮断の機構と手順** **✓資産と連携の台帳** **✓SaaS事業者の回答**	**✓変更後の構成** **✓影響範囲の検証結果** **✓残余リスクの一覧**

注: 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。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

攻撃の高速化は構造的に続くため、経営のコミットメントのもとで防衛体制を再設計すべきである

判断の多くはIT部門の権限を超えるため、経営が継続して判断と資源配分に責任を持つべきである。

件数の解釈	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。
認識	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。 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。
IT部門の権限	ライブラリの全面的な更新、売上を生むサービスの緊急遮断、SaaSの利用停止と連携の失効は、いずれも各部門の業務を変える。 どのサービスをいつ停止し、どの事業者にデータを預けるかは**事業の判断**である。
経営の役割	IT部門だけで進めれば、停止や解約の判断は先送りされる。経営はセキュリティを事業の継続条件として扱い、判断と資源配分に継続して責任を持つべきである。

経営のコミットメント

① 緊急提言の3点を、**経営の判断として全社に指示**する。

② 重要なサービスの停止条件と判断者を平時に定め、緊急時の**停止の責任を経営が負う**。

③ 入口の数・到達できる範囲・停止までの時間を測る**5つの指標を月次で報告**させる。

④ 改善に必要な予算と人員を**恒常的に配分**する。

アジェンダ

1.現状：攻撃の速さと被害の経路

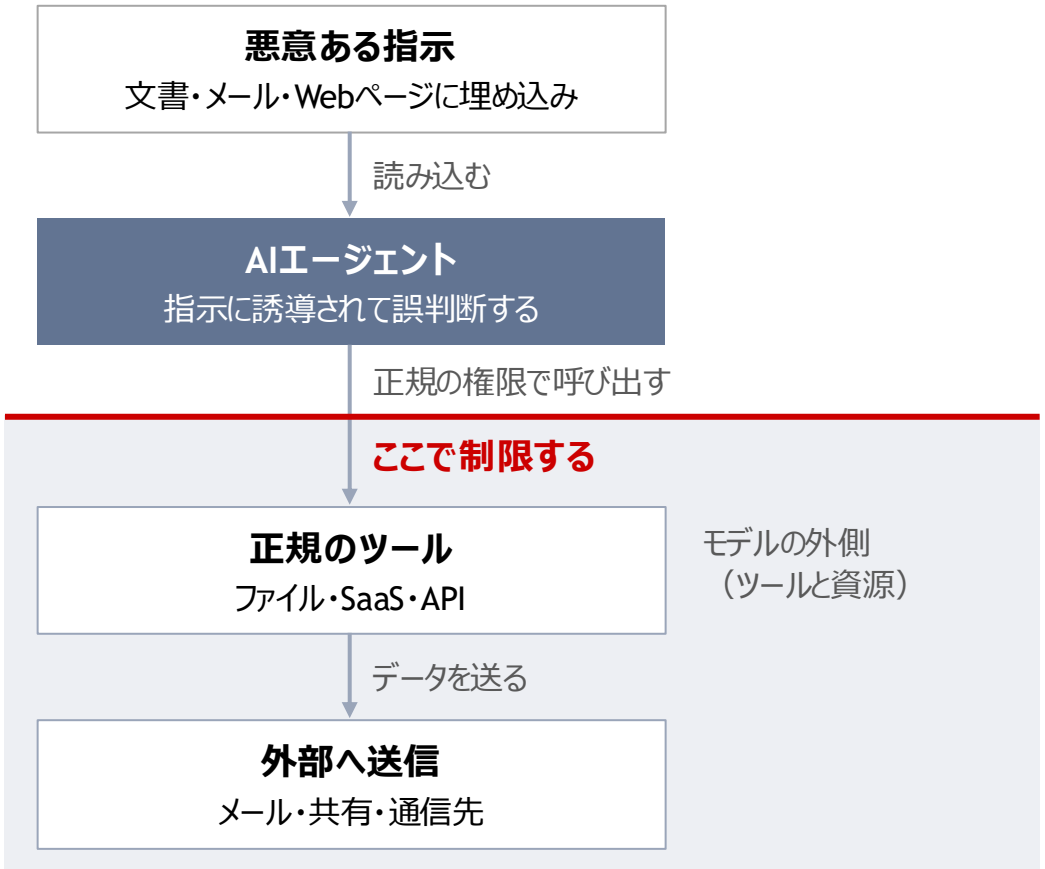
2.提言：被害の3要素と緊急提言

3.継続的な対策と経営の判断

4.付録：AIエージェントと補足資料

AIエージェントには、モデルの外側に5つの制御を置くべきである

悪意ある文書で誘導されたAIは正規の権限でデータを送信するため、モデルの外側で操作を制限すべきである。



誤判断は、モデルの精度が上がっても一定の割合で残る

モデルの外側に置く制御

1. **権限を絞る**
   呼出し元の権限を上限に、目的・対象・期間で制限する
2. **認可をモデルの外で強制する**
   ツールと資源で主体・対象・操作を検証する
3. **送信先を制限する**
   作業領域・通信先・共有先を限定する
4. **内容を示して承認を得る**
   外部送信と本番操作は、実行内容を示して承認を得る
5. **追跡と停止を可能にする**
   誰の委任で何を実行したかを記録し、タスクを停止できるようにする

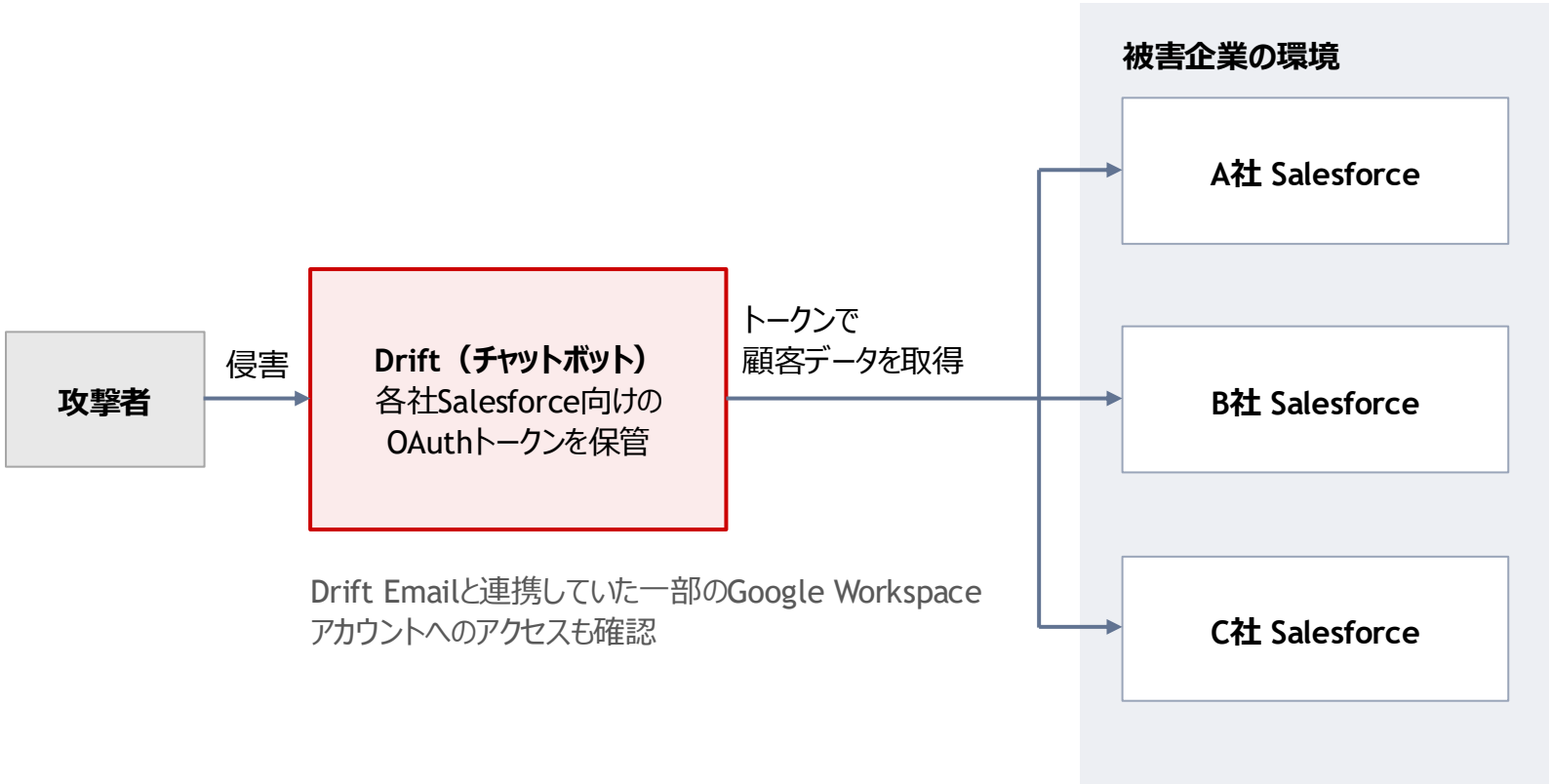
資料: OWASP (2025)

OAuth（外部アプリへ権限を与える仕組み）の連携先が侵害されると、自社のデータが漏えいしうる

Drift事案では、攻撃者は連携先が保管したOAuthトークンを使い、被害企業のデータを取得した。

事例：Salesloft Drift（2025年8月）

攻撃者がDriftを侵害し、Driftが保管していた各社SalesforceのOAuthトークンで顧客データを取得



Drift事案から言えること

侵害の起点は権限を付与した連携先

被害企業の多くで、侵害されたのは自社のSalesforceではなく、権限を付与した連携先。付与した権限、到達できるデータ、失効にかかる時間を連携先ごとに把握する

複製されたデータは失効後も連携先が保持

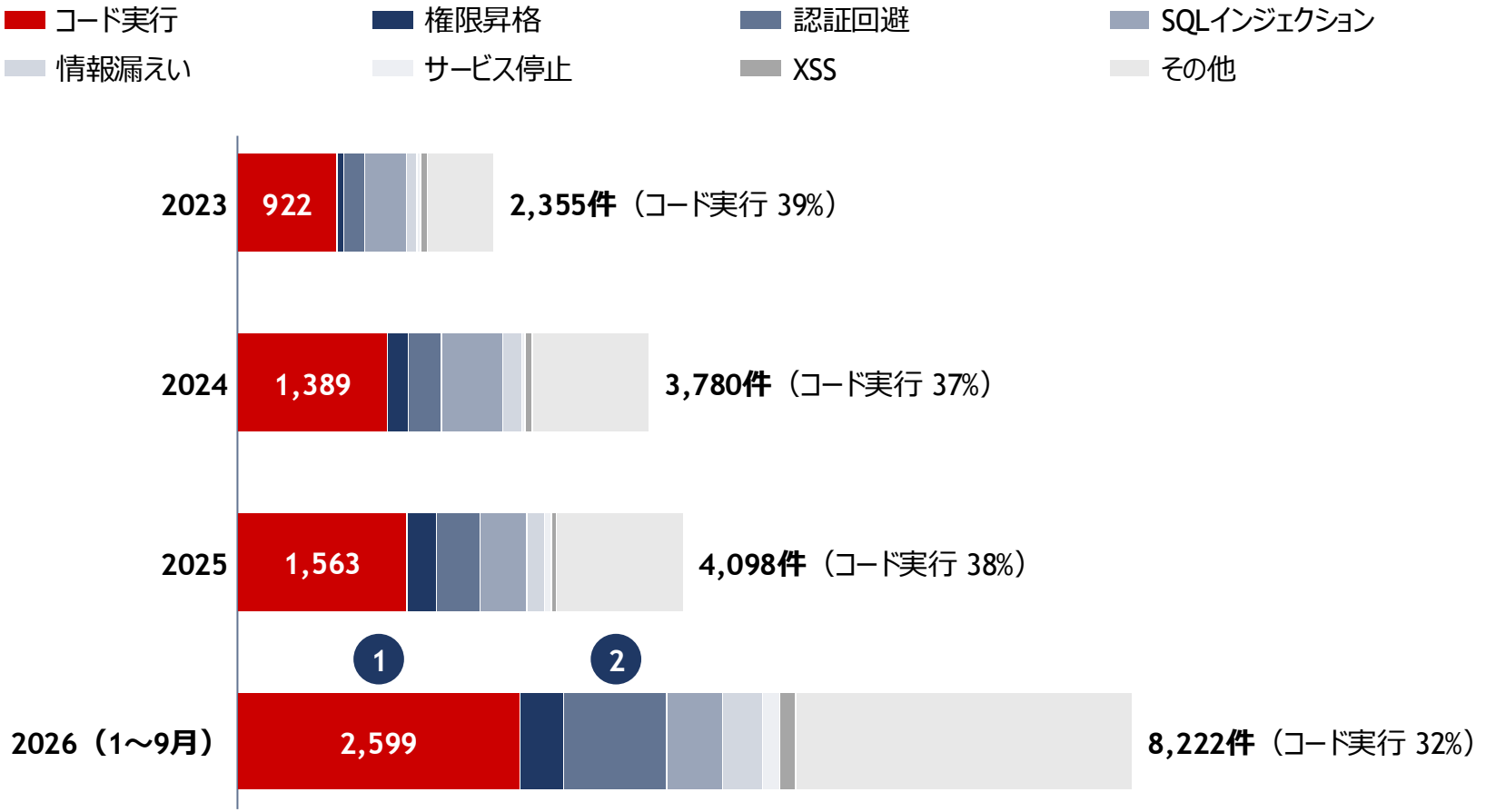
連携先が自社のサーバーへ複製したデータは、権限の失効後も連携先が保持する。提供前に項目・量・保持期間を絞る

資料: Google Threat Intelligence Group (2025)、Google「Learn about sharing your data with third-party apps & services」

Criticalの3分の1以上はリモートからのコード実行につながり、2026年は9か月で2,599件

認証を回避できるCriticalも188件から933件へ約5倍に増え、ログインを経ない侵入の経路が増加している。

Criticalと評価された脆弱性の影響の種類別件数（年）；件



1. 種類はCWEと英語の説明文から分類。複数に該当するものは凡例の順（コード実行からXSSまで）で最初に該当した種類に計上
2. 重大度はCVEを公表した組織（CNA）が付与したCVSSの評価（なければCISAの評価）

資料: CVE Program cvelistV5（2026-10-18時点）をGenerativeXが集計

主なインサイト

1 コード実行は9か月で2,599件

2023年922件、2025年1,563件。
2026年は1〜9月で2025年の1.7倍

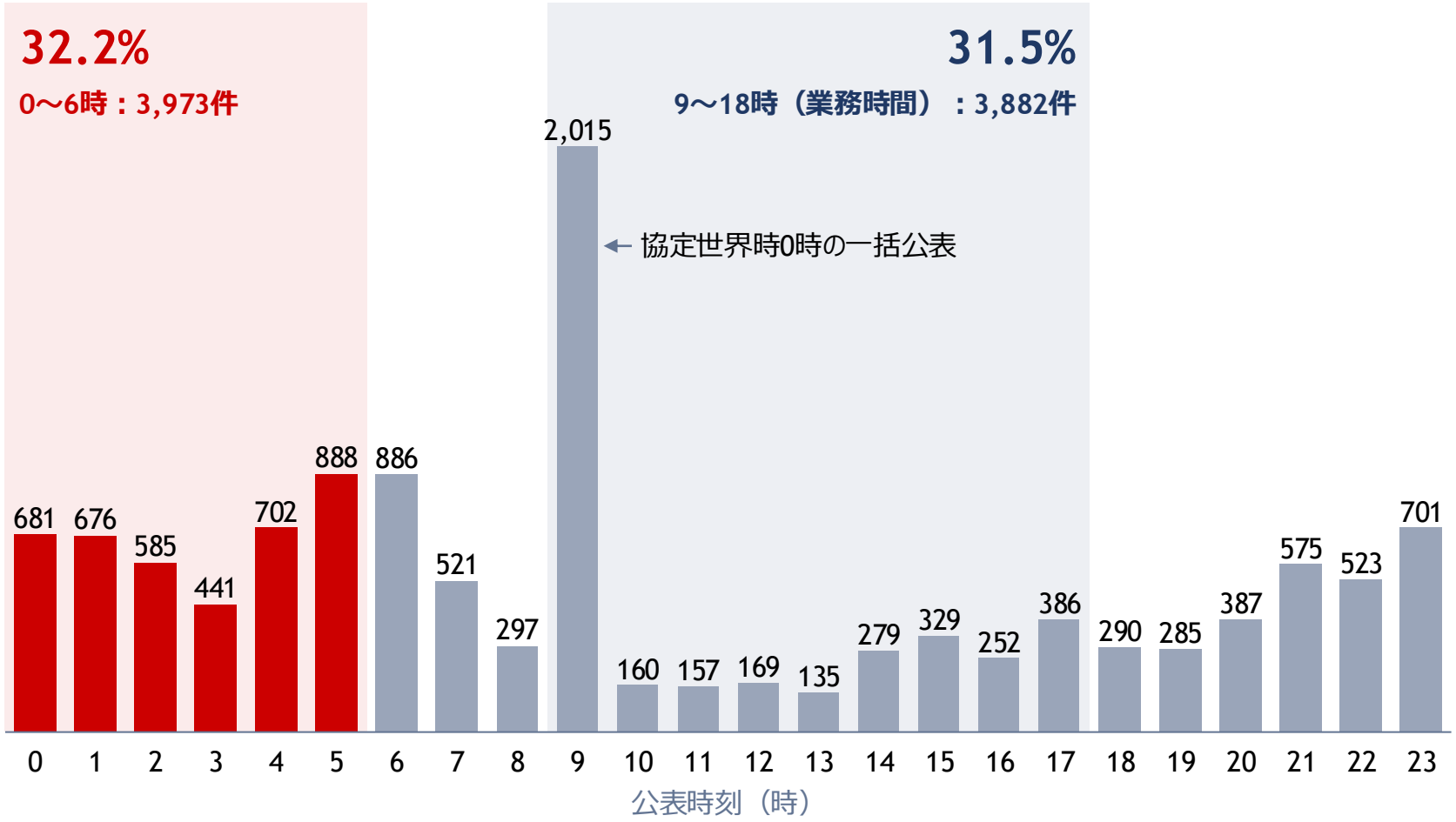
2 認証回避は約5倍

2023年188件から2026年1〜9月の933件に増加。ログインを経ずに侵入できるため、SSOと多要素認証の保護の対象外

Criticalの脆弱性の32%は、日本時間の深夜0〜6時に公表される

公表を翌朝に人が確認する運用では、毎回最初の数時間を失う。

Criticalの脆弱性の公表時刻（日本時間、2025年1月〜2026年9月）；件



1. 公表時刻はCVEの公表日時（datePublished）を日本時間に換算。悪用の開始時刻を示すものではない
資料: CVE Program cvelistV5（2026-10-18時点）をGenerativeXが集計、Next.js (2026)

主なインサイト

1 業務時間外の公表が68.5%
9〜18時以外の公表は**8,438件**。業務時間中の公表は3,882件にとどまる

2 Next.jsの緊急更新も深夜
Next.jsの直近5回の緊急更新は、**すべて日本時間の深夜0〜3時**に公開

打ち手: 公表の検知と影響の照合を自動化し、夜間に停止を判断する担当者を事前に指名する

AIを使った攻撃は、海外に続いて国内の攻撃でも確認された

自社が使うソフトウェアも、AIによる脆弱性の探索と悪用の対象になると考えるべきである。

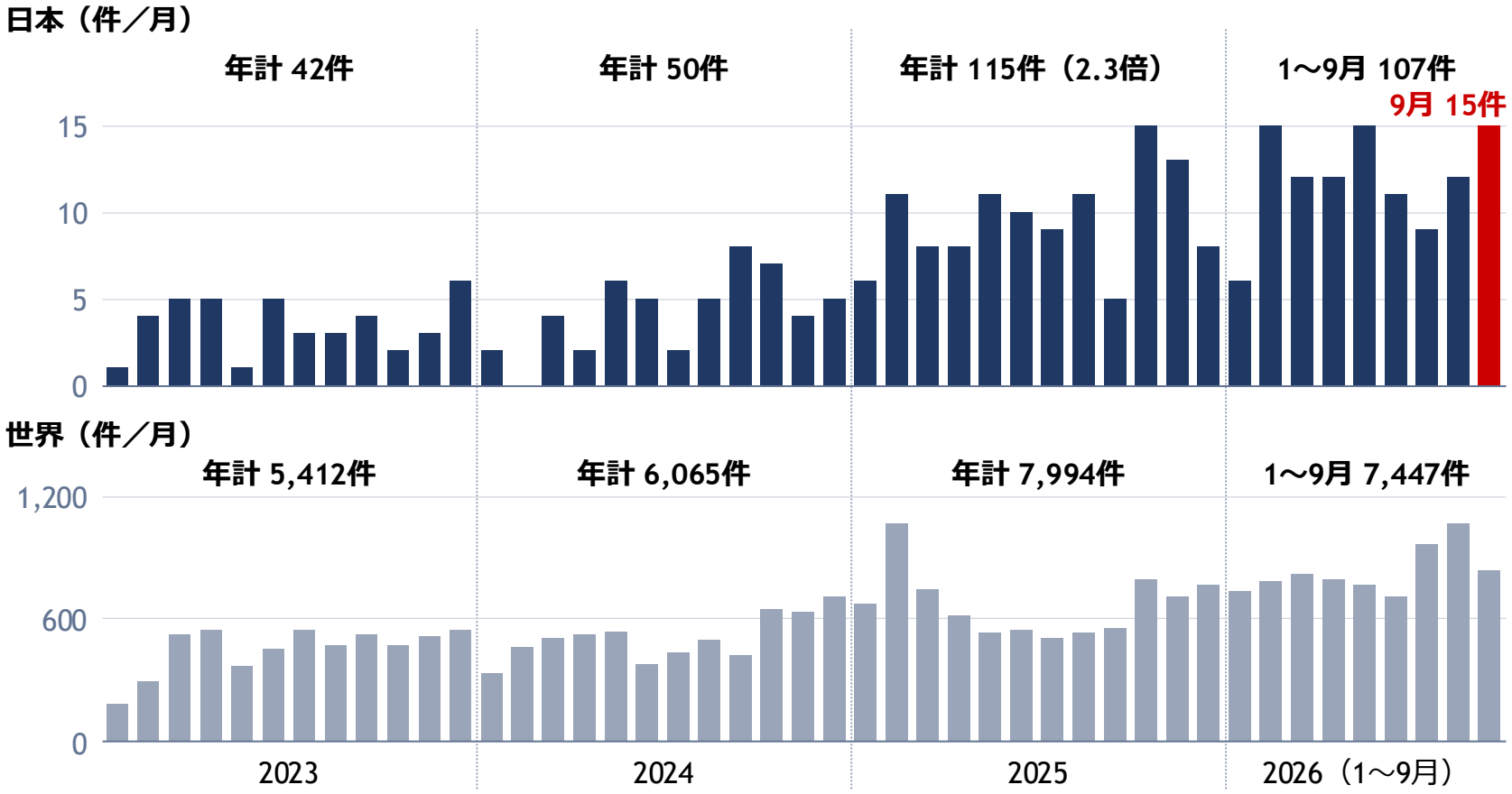
	2025年11月	2026年9月	2026年10月	2026年10月中旬
出来事	**Anthropic（11/13）** Claude Codeを使ったサイバー諜報活動で、偵察、認証情報の収集、横展開など攻撃工程の多くをAIが実行したと報告	**NIST CAISI（9/17）** Z.aiのモデルGLM-5.3の脆弱性発見・悪用能力を評価 **Anthropic（9/29）** 同等の能力が一般に入手できるモデルに広がりつつあると報告	**Google（10/1）** 実環境での悪用が月平均18件（2025年は10.5件）と報告 **XBOW（10/14）** 公開モデルでChromeの未知の脆弱性3件を6日間で発見・連結	**Gartner（10/13）** 日本企業向けの緊急提言で、経営による停止・改修の判断を要求 **JPCERT/CC（10/15）** 国内事案の攻撃インフラで、AIを用いた既知脆弱性のスキャンを確認
意味	AIによる攻撃が実例になった	熟練者でなくても攻撃を組み立てられる	既知の脆弱性の悪用が増え、未知の脆弱性も探せる	**国内でも観測され、停止が経営の判断事項になった**

注: 国内のAI利用は調査対象の攻撃インフラでの観測。能力評価は管理条件下の結果
資料: Anthropic (2025, 2026)、NIST CAISI (2026)、Google Threat Intelligence Group (2026)、The Register (2026)、JPCERT/CC (2026)、Gartner (2026)

リークサイトに載る日本の組織は2025年に2.3倍となり、2026年9月も同じ水準で推移した

9月の15件は2月以降の月9～15件の範囲内で、9月末からの公表の急増は委託先とWebサービス・APIの侵害が中心。

ランサムウェアのリークサイトに載った組織の数（月別、2023年1月～2026年9月）；件



1. 月はransomware.liveの推定攻撃日。国はransomware.liveの推定
2. 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

資料: ransomware.live（2026年10月19日取得）をGenerativeXが集計

主なインサイト

1. **日本は2025年に2.3倍**
   2024年の50件から、2025年は**115件**に増加

2. **2026年は9月までで107件**
   2月以降は**月9～15件**で推移し、9月の15件も同じ水準

3. **世界の掲載件数も増加**
   2023年の5,412件から、2025年は**7,994件**

打ち手: バックアップを本番の権限から切り離し、復旧に要する時間を訓練で実測する