# 2025年 Cyber IQ 調査

## ―生成AIの台頭とデジタル国境の形成に伴う サイバーリスクに企業はどう対応すべきか―

pwc

www.pwc.com/jp

# 目次



# はじめに

デジタル技術の急速な革新の中で、企業は新たなサイバーリスクに直面しています。地政学的な緊張や各国・地域の規制変化が企業環境に影響を及ぼす中、サイバー攻撃はますます高度化し、複雑さを増しています。

このような状況下で、PwC Japanグループは「Cyber IQ」という概念を提唱しています。Cyber IQは、過去の事例やベンチマークなど多様な情報を収集・分析し、組織のサイバーレジリエンスを強化するための資質を指します。

本レポートでは、法規制、生成AI、サプライチェーン、脆弱性管理、デジタルアイデンティティなどの急激な変化を考慮し、企業が取るべき対応策について考察しました。この知見が日本企業の皆さまのセキュリティ対策に役立つことを心より願っております。

# 1 地政学リスクにおけるサイバー脅威とAIリスクの潮流

## 日々進化するサイバー攻撃と生成AIがもたらす新たな脅威への対処

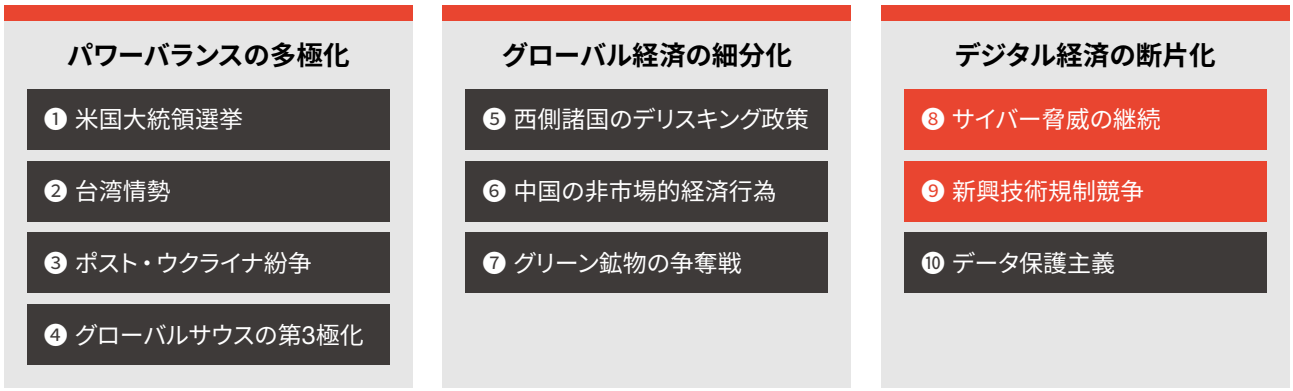
近年、日本企業を取り巻く国際情勢は劇的に変化しています。経営層には国際情勢を形成する潮流を理解し、従来以上に高度な方法で地政学リスクを管理することが求められるようになりました。

PwCは、2024年のビジネスに影響を及ぼす外部環境のうち、地政学上の重大なリスクを企業がどのように捉え、対応していくべきか検討する材料となるレポート「2024年地政学リスク展望」を公表しました。

2024年における主要な地政学リスクは、「パワーバランスの多極化」「グローバル経済の細分化」「デジタル経済の断片化」の3つが挙げられます（図表1）。

国境のないデジタル空間においても、国際情勢の不安定化を背景に国家間に障壁が生まれ、企業活動が制約されつつあります。海外で事業を展開する日本企業には、経済安全保障を考慮した企業戦略が必要となり、サイバーセキュリティ対応やAI規制法をはじめとする海外法規制への対応が求められます。こうした背景を鑑みて、第1章では地政学リスクのトレンドのひとつである「デジタル経済の断片化」について掘り下げて解説します。

**図表1：日本企業を取り巻く国際情勢―2024年の10大地政学リスク**




出所：PwC、「2024年地政学リスク展望」

### デジタル経済の断片化

かつて「国境のない空間」として期待されたデジタル領域においても、地政学的な緊張を背景に分断化が進んでいます。各国・地域は重要インフラの保護やデータ主権の確保を理由に、デジタル空間の管理を強化しており、新たな「デジタル国境」が形成されつつあります。特に安全保障の観点から、各国政府による規制は強まるばかりです。重要インフラに関わるシステムでは特定国の製品やサービスを排除する動きが広がり、重要データの国内保存を義務付けるデータローカライゼーション規制も世界的に拡大しています。

さらに、生成AIをはじめとする新興技術の台頭は、デジタル経済の複雑性を一層高めています。各国・地域で異なるAI規制への対応や、データの越境移転に関するルールの遵守など、企業はグローバルなデジタルビジネスモデルの抜本的な見直しを迫られています。

このような状況下で海外展開を進める日本企業は、経済安全保障の観点を考慮した慎重な戦略が不可欠です。特に日々進化するサイバー攻撃への対応と、急速に発展する生成AIがもたらす新たな脅威への対処は、喫緊の課題と言えるでしょう。次からは、これら2つの重要課題について詳しく見ていきます。

## 国家戦略の一翼を担うサイバー攻撃の実態

最初に企業が地政学リスクをどのように捉えているかを紹介します。PwCが実施した「企業の地政学リスク対応実態調査2024」[※1]によると、経営者が最も懸念している地政学リスクは「ロシア・中国・北朝鮮などによるサイバーアタック／サイバーテロ」であり、回答の40%を占め、3年連続で首位となりました。この背景には、2023年に日本へのサイバー攻撃関連通信の観測数が過去最高を記録し、企業活動に重大な支障をきたす事案が発生したことが大きく影響しています。

- 調査対象：海外展開する年間売上高100億円以上の企業の管理職400名
- 実施時期：2024年7月
- 調査方法：オンライン
- 対象業種：製造業、サービス業など産業全般
- 調査履歴：2019年3月、2021年8月、2022年8月、2023年8月に続く第5回目

こうした状況下、世界各国・地域ではデジタル分野に関する規制が急速に整備されています。PwCのデジタル規制を監視するリサーチ部門の調査によると、企業が対応すべきデジタル関連のガイドラインや法令は84件で、2020年の30件から3倍程度増加しました（図表2）。

これらの規制は、AI、IoT、電子商取引、サイバーセキュリティ、プライバシーなど多岐にわたります。特にデジタル分野の法規制は技術的な専門知識が必要となるため、従来の社内法務部門での対応が困難であり、事業部門やITセキュリティ部門との連携が不可欠となります。経営者は事業を展開する国や地域ごとの規制を把握し、自社への影響を分析した上で、適切な対応策を講じなければなりません。

※1　PwC、「企業の地政学リスク対応実態調査2024」
https://www.pwc.com/jp/ja/press-room/2024/geopolitics2408.html

## 図表2：グローバル企業が対応すべき法令・ガイドライン一例

### AI

- EU 欧州委員会：AI指針
- EU 欧州委員会：欧州AI法
- 米国 ホワイトハウス：AI権利章典
- 米国 国防総省：責任あるAIの指針の採用
- 米国 連邦取引委員会：ビジネス向けのAIアルゴリズム利用に係るガイダンス
- 米国 予算行政管理局：AI規制に係るガイダンス
- 米国 国立標準技術研究所（NIST）：AIリスクマネジメント枠組み
- 英国データ保護機関（ICO）：AIおよびデータ保護リスクツールキットv1.0
- 中国 国家インターネット情報弁公室（CAC）：インターネット情報サービスアルゴリズムレコメンデーション管理規定
- 中国 国家インターネット情報弁公室（CAC）：インターネット情報サービス深度合成アルゴリズム管理規定
- 中国 国家インターネット情報弁公室（CAC）：生成人工知能サービスの管理に関する暫定措置

### デジタルサービス

- EU 電子商取引指令
- EU デジタルサービス法
- EU デジタル市場法

### IoT／OT

- 米国 IoTサイバーセキュリティ改善法
- 米国 カリフォルニア州 IoTセキュリティ法
- 米国 オレゴン州 IoT法
- EU サイバーセキュリティ法
- EU サイバーレジリエンス法
- EU RED委任規制2022/30（RE指令）
- 英国 製品セキュリティと通信インフラに関する法
- ブラジル 通信機器サイバーセキュリティ要件（法律第77号）

### サイバーセキュリティ

- 日本 サイバーセキュリティ基本法
- 中国 サイバーセキュリティ法
- 米国 2022年重要インフラ向けサイバーインシデント報告法
- EU サイバーセキュリティ法
- EU 共通の高度サイバーセキュリティ措置に関する指令（NIS 2指令）
- EU サイバーレジリエンス法案
- EU デジタルオペレーションレジリエンス法
- 英国 ネットワーク・情報システム規則
- インド IT法2000／インフォメーション・テクノロジー・ルール2021
- インド 国家サイバーセキュリティポリシー
- インド CERT-Inサイバーセキュリティ指令2022
- 国連 UNECE規則サイバーセキュリティ（UN-R155）
- 国連 ソフトウェアアップデート（UN-R156）
- ISO ISO/SAE 21434

### プライバシー

- 日本 個人情報保護法
- 中国 個人情報保護法
- 中国 データセキュリティ法
- 香港 個人情報保護法
- 台湾 個人情報保護法
- 韓国 個人情報保護法
- インド 個人情報保護法
- インドネシア 個人データ保護法
- オーストラリア 個人情報保護法
- シンガポール 個人データ保護法
- タイ 個人情報保護法
- フィリピン 個人情報保護法
- ブルネイ 個人データ保護規定草案
- ベトナム 個人データ保護法
- ニュージーランド プライバシー法2020
- マレーシア 個人情報保護法
- スリランカ 個人データ保護法2022年第9号
- EU 一般データ保護規則
- EU eプライバシー規制
- ベラルーシ 個人情報保護法
- ドイツ 電気通信およびテレメディアにおけるデータ保護およびプライバシーの規制に関する連邦法
- 英国 データ保護法
- 英国 一般データ保護規則
- スイス データ保護に関する新連邦法
- イスラエル 個人情報保護法
- カタール データ保護規則2021
- モロッコ 個人データの処理に関する個人の保護に関する法律第09-08号
- チュニジア 基本法 第63-2004号
- トルコ 個人情報保護法
- ロシア 2006年7月27日付個人データに関する連邦法152-FZ号の改正法
- 南アフリカ 個人情報保護法
- アラブ首長国連邦 個人情報保護に関する2021年連邦政令第45号
- ケニア データ保護法
- ルワンダ 個人情報保護法とプライバシーに関する法律
- カナダ 個人情報保護および電子文書法
- カナダ 2022年デジタル憲章実施法
- 米国 カリフォルニア州消費者プライバシー法
- 米国 バージニア州消費者データ保護法
- 米国 コロラド州プライバシー法
- 米国 コネチカット州データ保護法
- 米国 ユタ州消費者プライバシー法
- チリ 私生活の保護に関する法律
- ブラジル データ保護法
- ペルー 個人情報保護法
- メキシコ 民間団体が保有する個人データの保護に関する連邦法
- パナマ データ保護法
- エクアドル 個人情報保護法
- アルゼンチン 個人情報保護法 25.326

出所：各国・地域の公的情報よりPwC作成

次に、日本企業が直面しているサイバー攻撃について見てみましょう。

PwCでは、サイバーインテリジェンス活動の一環として、グローバルで活動を続ける脅威アクター（攻撃者グループ）を常時監視しています。監視を担当するのは日本を含むグローバル専門チームで、日本独自の観点での分析結果も提供しています。同チームでは現在、362の脅威アクターを特定・追跡しており、そのうち35グループが、日本を攻撃対象としていることを把握しています。

脅威アクターの中でも特に注目すべきは、中国を拠点とする脅威アクターの動向です。PwCが独自に調査したところ、2024年9月時点で中国を拠点とする脅威アクターは75あることが確認されています。その中でも「Red Apollo」「Red Kelpie」「Red Dev 5」「Red Ladon」は日本を標的として活動することが確認されており、また米国司法省やその同盟国により図表3のとおり起訴されています。

**図表3：中国を拠点とする脅威アクターの事例**

| | 脅威アクター PwC呼称 | 別名例 | 起訴の概要 |
|---|---|---|---|
| **1** | Red Apollo | APT10,<br>Stone Panda,<br>MenuPass Team など | 2018年、商業スパイを可能にし、「データ、知的財産、ビジネスや技術の機密情報」を標的とした広範な一連のサイバー活動で起訴された |
| **2** | Red Kelpie | APT41,<br>WICKED PANDA,<br>BARIUM 等 | 2020年、「商業的利益を目的とした」高機密なビジネス情報の窃取を含むサイバー活動で起訴された。Red Kelpieは、Chengdu 404 Network Technologyとの関連が公表されている |
| **3** | Red Dev 5 | Dark Shadow,<br>Storm-0062,<br>BRONZE SPRING など | 2020年、特に商業スパイ行為と、技術設計、製造プロセス、テストメカニズムおよび結果、ソースコード、化学構造などの知的財産ならびに企業秘密の窃盗に関連する数多くのサイバースパイ行為で起訴された。Red Dev 5は中国の国家安全部（MSS）の広東省国家安全部（GDSS）に帰属すると公表されている |
| **4** | Red Ladon | APT40,<br>Leviathan,<br>GADOLINIUM,<br>Gingham Typhoon など | 2021年、知的財産や企業の機密情報を狙ったさまざまなサイバースパイ行為で起訴された。Red Ladonは、以下を含む広範な機密情報を窃取した。<br>- 潜水艇や自律走行車<br>- 特殊化学品の処方<br>- 独自の遺伝子配列技術<br>- 感染症研究に関する情報など<br>起訴状では、Red Ladonが「対象国内の国有企業（大規模な高速鉄道開発プロジェクトなど）の契約を確保しようとする中国の努力を支援するために外国の情報を盗んだ」と明確に言及されており、MSSの海南国家安全部（HSSD）に帰属すると公表されている |

これらの組織は、表面上は個別に活動しているように見えますが、実際には中国の第14次5カ年計画（2021～2025年）に沿う形で標的を選定し、連携して活動を展開していると推測されます。例えば、Red Kelpieが標的とする金融技術、Red Ladonが焦点を当てる先端技術は、いずれも中国が国家戦略として掲げる重点領域と合致するものです。

こうしたグループは、それぞれが得意とする領域で情報収集活動を行いながら、収集した情報を相互に共有・活用している可能性が高いと考えられます。各脅威アクターは個別組織として高度な技術力を保持しながら、国家戦略の一翼を担い、協調的に活動している実態が浮かび上がってきています。こうした観点からも現在検討が進められている第15次5カ年計画（2026～2030年）の動向について注視することが重要です。

## 脅威アクターに企業はどのように対峙すべきか

では、こうした脅威アクターに対し、企業はどのように対峙すべきなのでしょうか。

最も重要なのは、脅威アクターの攻撃手法を正確に理解することです。この理解を深めるために用いられるのが「MITRE ATT&CK」[※2]です。例えば、脅威アクターであるVolt Typhoonの攻撃手法をこのフレームワークに当てはめて分析すると、初期アクセスではスピアフィッシングや認証情報の窃取を行い、その後、LOTL（環境寄生型攻撃）を駆使してターゲットのシステム内をくまなく“調査”し、最終的にデータの窃取などを行います（図表4）。攻撃者の侵入を100%ブロックすることは不可能ですが、攻撃パターンを把握することで、「どの段階で」「どの攻撃を」「どのような手段で」ブロックするかという具体的な対策を検討できるのです。

**図表4：Volt Typhpooが用いる主要な戦術・技術・手順（TTPs）**

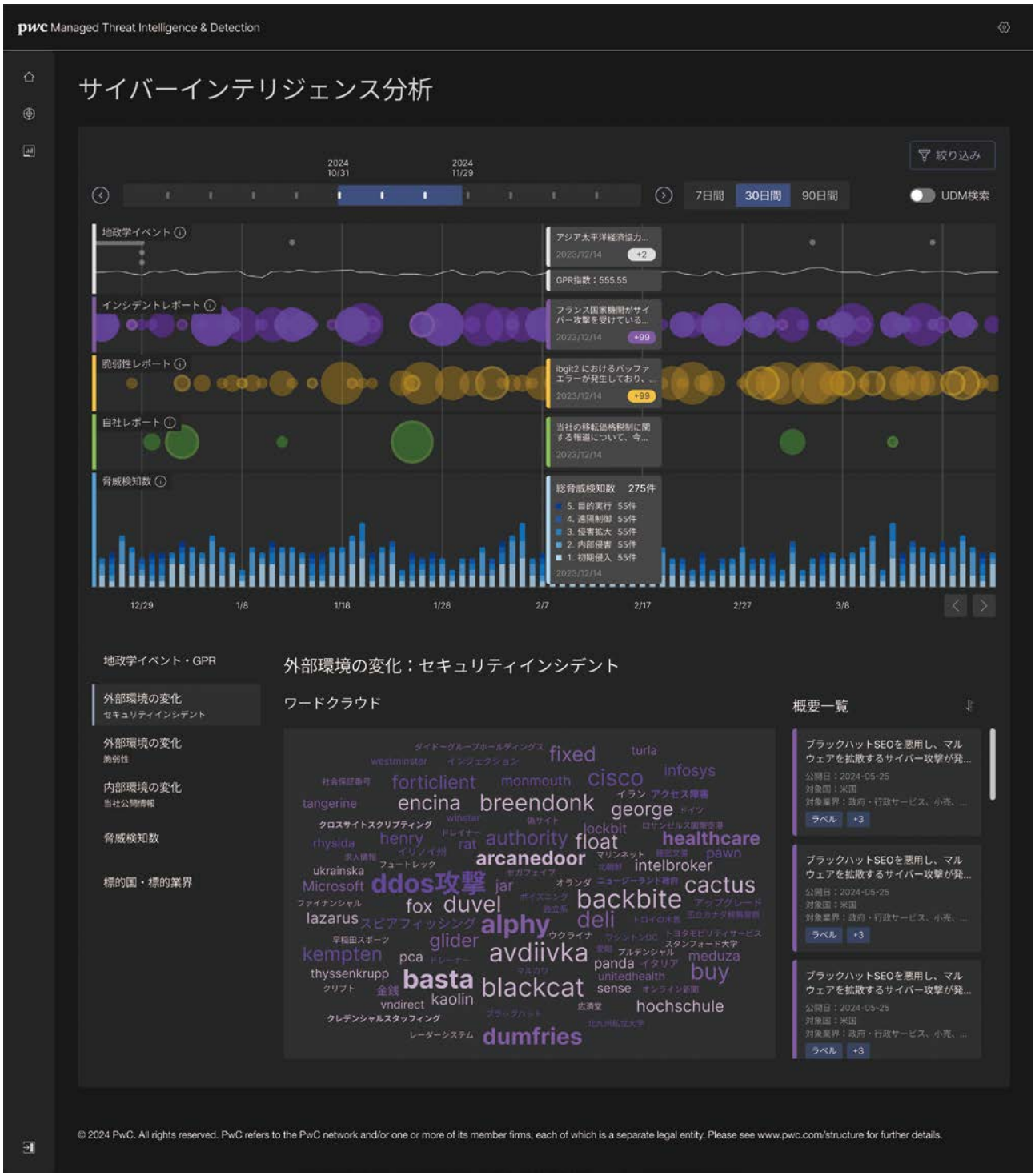
**MITRE ATT&CKに基づいた分析**

| 偵察、開発 | 初期アクセス、侵入 | 永続化、権限昇格、防御回避 | 探索、認証アクセス、横感染、収集 | 遠隔操作 | 情報持ち出し、インパクト |
|---|---|---|---|---|---|
| 被害者ホストの情報収集 | オンラインにさらされたアプリケーションへのエクスプロイト | サーバーソフトウェアコンポーネント | 権限を持つグループの探索 | プロキシ | C2チャンネルを介した持ち出し |
| 被害者の身元情報の収集 | 正当なアカウントの悪用 | 正当なアカウントの悪用 | システム情報の検出 | 侵入ツールの転送 | 別プロトコル経由での持ち出し |
| 被害者ネットワーク情報の収集 | コマンドおよびスクリプト言語・インタプリタ言語 | アーティファクトの隠ぺい | ネットワーク設定の検出 | 暗号通信路 | |
| 被害者の組織情報を収集 | WMIの悪用 | 痕跡消去 | ユーザーアカウントの探索 | | |
| オープンウェブサイトやドメインの検索 | リモートアクセスサービスの悪用 | システムバイナリプロキシ実行（検知回避技法） | 収集データのアーカイブ | | |
| 被害者が所有するウェブサイトの検索 | | ファイルまたは情報の難読化 | OS認証情報のダンプ | | |
| インフラの侵害 | | 特権昇格のための悪用 | エクスプロイトによる認証情報アクセス | | |
| 攻撃インフラの確立：ボットネット | | | 安全でない認証情報 | | |
| | | | リモートサービスのセッション乗っ取り | | |

※2　MITRE ATT&CK　サイバー攻撃者の戦術・技術・手順（TTPs）を体系的に整理したナレッジベース。企業や組織のセキュリティ対策の指針としてグローバルで活用されている。

企業がこうした分析を進める上では、活動地域・業種といった企業の特性と脅威アクターの分析結果を組み合わせ、その企業に対して想定される戦術・技術・手順（TTPs）を特定することが重要です。具体的には、システムから収集された各種アラートやログ、最新の脅威情報やインシデント情報、業界特性に基づく脅威アクター情報や既存のセキュリティ対策の状況など、さまざまな要素をもとに分析を実施します。将来起こり得る攻撃の予測や、過去に発生した攻撃の検証をすることで、効果的な防御ができるとともに、経営層にも理解しやすい形でレポーティングをすることが可能となります（図表5）。

**図表5：サイバーインテリジェンス分析のイメージ（PwCの「Managed Threat Intelligence &Detection」）**




【参考URL】 https://www.pwc.com/jp/ja/services/digital-trust/cyber-intelligence/managed-threat-intelligence-detection.html

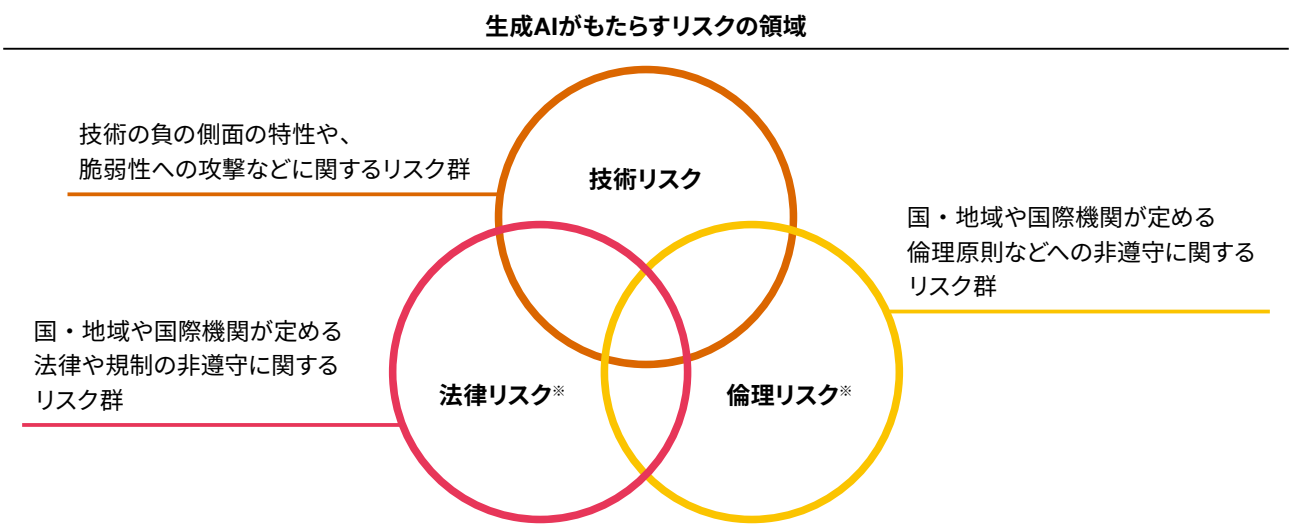
## AI特有の新たなリスクへの対応

次にAIを取り巻くリスクを見ていきましょう。

企業においてAIは着実に普及しています。PwCが実施した「生成AIに関する実態調査2024 春」では、91%の企業が既に生成AIを活用しているか、具体的な検討を始めていることが明らかになりました。2023年春の調査では同比率が22%であったことからも、企業における生成AI導入は急速に進んでいることが分かります。今後1年から2年の間に、多くの企業が生成AIを活用したビジネスを展開することは間違いありません。

しかし、生成AIのビジネス活用が拡大する一方で、AI特有の新たなリスクへの対応が課題となっています。ここではAIのリスクを、「技術」「法律」「倫理」の3つの領域に分類して考えます（図表6）。

**図表6：AIを取り巻くリスク**




※法律と倫理の境界は国・地域によって異なるため、リスクの境界も国・地域によって異なる

| 生成AIに関する代表的なリスク | 技術 | 法律 | 倫理 |
|---|---|---|---|
| ●：既存のAIも有するリスク<br>●：生成AIのもたらす新しいリスク | | | |
| ● AIに対するサイバー攻撃のリスク | ● | ● | |
| ● ハルシネーション（虚偽情報生成）のリスク | ● | ● | ● |
| ● AIが学習した機密情報が漏洩するリスク | ● | ● | |
| ● 利用者が入力した機密情報が漏洩するリスク | ● | ● | |
| ● 生成物の著作権／肖像権侵害のリスク | | ● | ● |
| ● 生成物に対する著作権の保護適用に関するリスク | | ● | |
| ● レピュテーションリスク | | | ● |
| ● 有害／差別的コンテンツ生成のリスク | ● | ● | ● |

技術面では、AIシステムへのサイバー攻撃が最も懸念されるリスクです。具体的には、AIモデル自体を標的とした攻撃です。実際、敵対的サンプルを用いてAIの判断を誤らせたり、モデルの内部パラメータを改ざんしたりする攻撃が確認されています。また、学習データを操作し、AIの性能を低下させる攻撃も報告されています。学習データにノイズを混入させたり、意図的に偏りのあるデータを注入したりすることで、AIモデルの判断精度を劣化させるのです。

加えて、大規模言語モデルでは、学習過程で取り込んだ機密情報が想定外の形で出力されるリスクも生じています。学習データに含まれる個人情報や企業の機密情報が、特定の入力パターンによって意図せず出力される可能性もあります。プロンプトインジェクション攻撃などにより、こうした情報漏洩を意図的に引き起こす手法も確認されています。

法律・倫理面では、技術の進化に法整備が追いついていないという課題があり、リスクも多岐にわたっています。中でも企業が最も留意すべきは、利用者が入力した機密情報の意図しない漏洩です。機密情報や個人情報を学習データに利用していなくても、これらを含むプロンプトが他者との会話履歴に残ったり、システムのログに保存されたりする可能性は十分に考えられます。

また、AIが学習データとして使用した既存の著作物と酷似したアウトプットを「AIの成果物」としてしまった場合は、訴訟に発展することもあります。一方でAIが生成した文章やコードなどの成果物に対する著作権保護のあり方も、十分に議論されていません。

さらに「技術」「法律」「倫理」の3つの領域に加えて、複数の領域にまたがる重要なリスクも存在します。AIが誤った情報を提示してしまう「ハルシネーション」は、技術的な限界に起因する問題でありながら、誤情報の拡散による社会的影響や法的責任などの問題もはらんでいます。その上、AIの不適切な利用や出力による企業の評価低下、いわゆるレピュテーションリスクは、技術的な不具合、倫理的な問題、法令違反など、さまざまな要因から生じる可能性があり、包括的な対策が求められます。

これらのリスクへの対応をさらに複雑にしているのが、国・地域による規制環境の違いです。例えば、データプライバシーに関する規制は、欧州連合（EU）の一般データ保護規則（GDPR）のように厳格な法的規制として確立している地域がある一方で、ガイドラインレベルにとどまる地域も存在します。加えて、AIモデルの学習に使用するデータの取り扱いについても、国・地域によって法的解釈が異なり、一元的なデータ管理は困難です。

このような規制環境の違いを背景に、各国・地域では独自の生成AI規制の整備が進んでいます。PwCが実施した調査によると、EU、米国、中国、日本など主要国において、それぞれの価値観や産業政策を反映した規制フレームワークが形成されつつあります（図表7）。グローバルに事業を展開する企業は、進出先の規制動向を注視し、各国・地域の要件に適合したリスク管理体制を構築する必要があります。

日本はこれまで「人間中心のAI社会原則」を基盤として、AI事業者ガイドラインを策定してきました。2024年度には新技術への対応や国際動向を踏まえた改定が予定されています。さらに金融や医療など、業界特性を考慮した個別のガイドライン整備も進められています。こうしたことから、企業には自社の事業領域に関連する規制の動向を把握し、計画的な対応を進めることが求められます。

## AIリスク対応のアプローチ―インテリジェンスの活用と実践

先述したとおり、生成AIの普及に伴い、2023年からAI関連インシデントは急増しています。発生したインシデントの多くは心理的影響、企業評判、経済的損失に関連するものですが、近年では物理的な被害も報告され始めています。

サイバーセキュリティ対策と同様、AIリスク対応では、インテリジェンスを活用したアプローチが有効です。ここで言うインテリジェンスは、技術的な情報収集にとどまりません。世界各国・地域で策定される規制やガイドラインの動向把握、同業他社で発生したインシデントの分析、さらには自社と同様のAI活用をしている企業の体制やインシデント対応事例など、幅広い視点での情報収集と分析を指します。特に技術領域では、AIを活用した攻撃手法の研究や、AIシステムを標的とした攻撃への対策など、最新の脅威に関する知見を継続的に蓄積することが重要です。

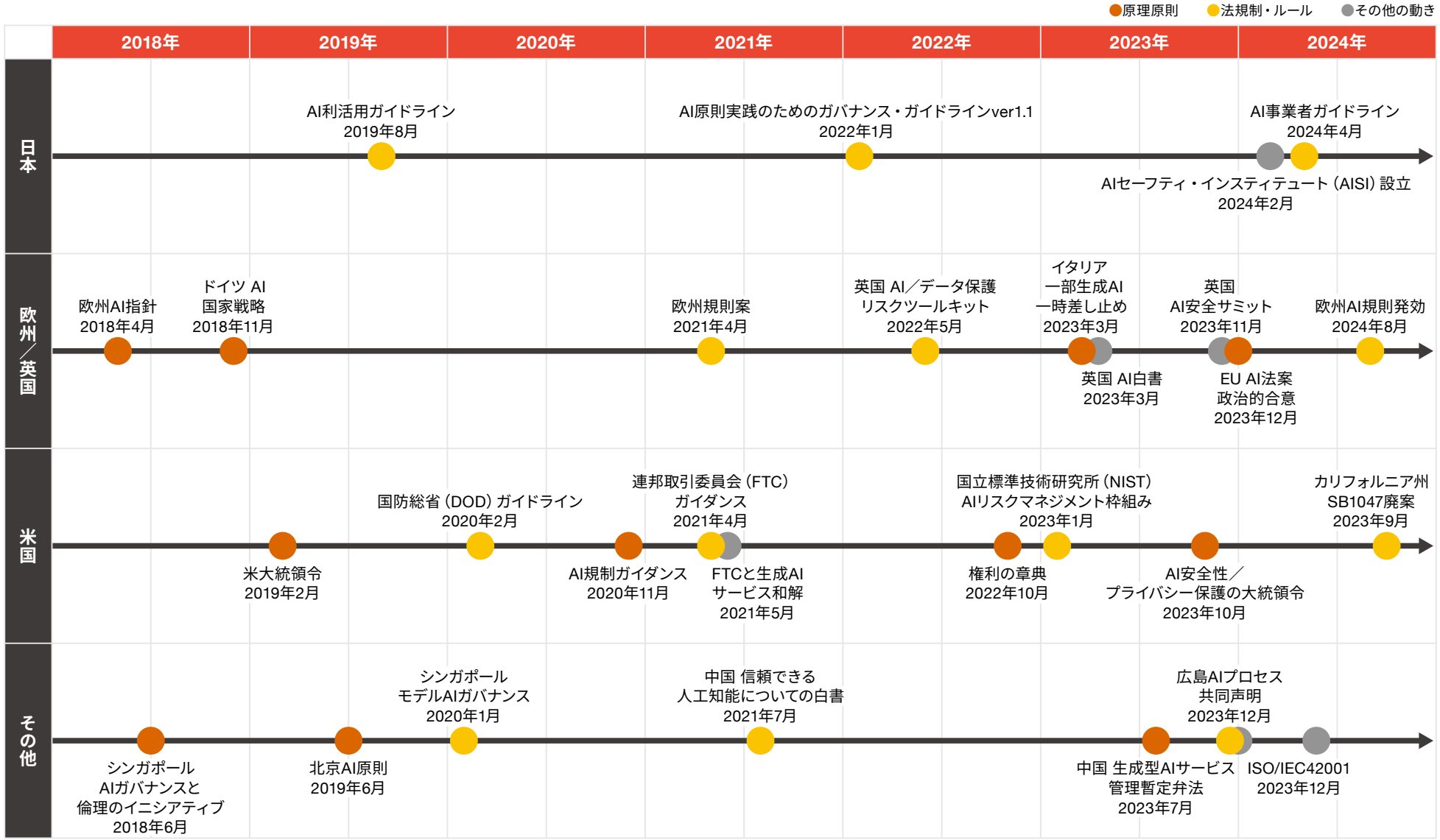
AIインシデントへの対応には、技術部門だけでなく、法務、コンプライアンス、広報など、組織横断的な協力体制の構築が不可欠です。各部門がそれぞれの専門性を活かしながら、包括的なリスク管理体制を整備することが求められています。

とりわけ重要なのは、AIの進化スピードを考慮した定期的なリスク評価の見直しと、新たなリスクへの迅速な対応体制の確立です。このためには、技術部門と法務部門の緊密な情報共有、広報部門を含めた危機管理体制の構築、そして経営層への適切な報告ラインの確保が必要となります。併せて、インシデント事例からの教訓を組織的に共有し、定期的な訓練と体制の見直しを行い、最新の脅威情報に基づいて対応手順を更新していくことも重要です。

このようなインテリジェンスを活用した実践的な取り組みとして、「AIレッドチーム」の活動があります。

第2章では「AIレッドチーム」を含む、企業が注目すべき事項について詳説します。

## 図表7：AI関連規制に関するグローバル動向




出所：2024年4月19日 AI事業者ガイドライン第1.0版の図にPwCの独自情報を追加して作成
https://www.meti.go.jp/press/2024/04/20240419004/20240419004.html

# 2-1 本社主導で進めるべきサイバーセキュリティ関連の法規制対応

## 1. サイバーセキュリティ関連の法規制を取り巻く状況

法規制は、各国・地域の政治経済的背景や経済安全保障を含む政策を実現する手段の一つのツールです。サイバーセキュリティを含むデジタル分野の法規制も同様で、現在進行形で各国・地域においてデジタル法規制が整備されつつありますが、その内容には類似点と相違点が存在します。

サイバーセキュリティ関連の法制の動向としては、多くの各国・地域でサイバーセキュリティに関する基本的な法制度（例：日本のサイバーセキュリティ基本法や欧州NIS2指令）が制定されました。それに加えて近年では、重要インフラのセキュリティ、製品セキュリティ、データの越境移転、サイバーインシデントの当局報告のための法規制の整備が各国・地域で進行しており、この傾向は今後も継続すると考えられます。

特に最近IoTに関する製品セキュリティに関する法規制が進んでいる背景の一つとして、政府や重要インフラでのIoTの普及が進み、製品セキュリティの基盤となる法規制や認証制度の必要性が高まっていることが挙げられます。製品セキュリティに関する代表的な法規制として、多くの企業にとって対応が課題になりつつあるのが、欧州のCRA（Cyber Resilience Act）と欧州データ法です。

CRAは、製品のセキュリティ対策に加えて、脆弱性対応も必須要件となっています。主要要件として、セキュリティ設計をはじめとする技術文書の10年間保管、サードパーティコンポーネントのセキュリティ保証、上市後最低5年間の脆弱性対応、インシデント時の24時間以内のENISA（あるいは加盟国CSIRT）への報告義務、SBOM対応など多くの要件があります。

欧州データ法では、IoT機器を使用することから生成されるデータが広く対象となっており、製品利用者はデータへのアクセス、第三者との共有などの権利を有すると定めています。そのため、従来のIoT製品の設計またはビジネスモデルの再考を促す内容になっています。

IT環境が複雑化・多様化するにつれて、新しい脅威が出現し、その脅威に対抗するために法規制が拡充されるという流れが今後も継続すると考えられ、企業の担当者にとっても、世界中で制定されるサイバーセキュリティ関連の法規制への対応が新たな重要課題となりつつあります。

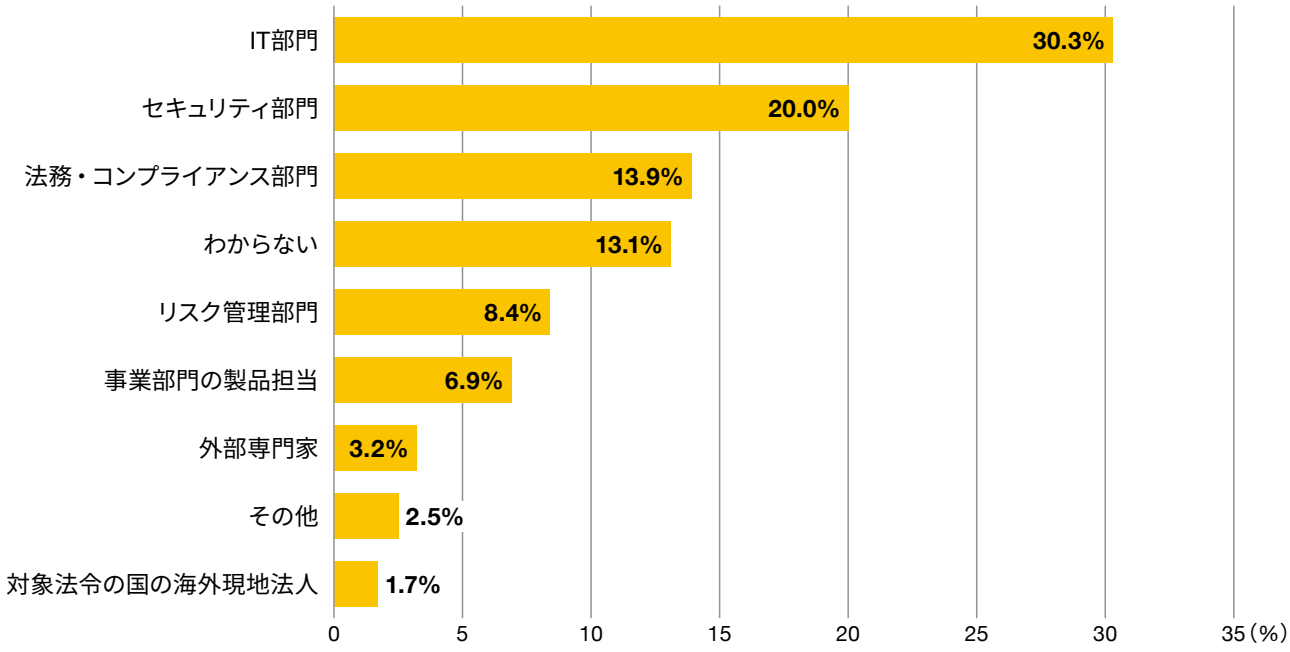
## 2. 多様なサイバーセキュリティ関連の法規制に対応する日本企業の課題

企業は進出国・地域のサイバーセキュリティ関連の法規制に、それぞれ対応せざるを得ません。それぞれの法規制において、特定時点までに遵守しなければならない義務は何か、遵守計画を作れば足りるものは何か、あるいは努力義務は何かを明確に把握した上で対応する必要があります。しかし、各国・地域の動向を継続的にモニタリングして、対応することは多くの企業にとっては高いハードルになります。

PwCの「2025年 Cyber IQ調査」では、日本企業においてサイバーセキュリティの意思決定や企画に携わる300人にサイバーセキュリティ法規制への対応を尋ねました。まずは、サイバーセキュリティ関連の法規制のモニタリングを行う部門について調査しました。（図表8）。

**図表8：サイバーセキュリティ関連の法規制※3に対応している部門**

**Q.** サイバーセキュリティ関連の法規制※3のモニタリングは誰が行っていますか。（複数選択可、n=300）




出所：PwC、「2025年 Cyber IQ調査」

回答からは、以下3点の傾向が確認できました。

1. グローバルなサイバーセキュリティ関連の法規制をモニタリングしている部門として、IT部門（30.3%）とセキュリティ部門が多い（20.0%）。これらの部門は法令モニタリングを専門としている部門ではないため、主担当の業務との兼務で法令モニタリングをしていると考えられます。

2. 法規制モニタリングの従来の主担当と想定される、法務・コンプライアンス部門（13.9%）でも、グローバルなサイバーセキュリティ関連の法規制のモニタリング関与度は高くありません。原因として、モニタリングリソースの不足、サイバーセキュリティに関するナレッジの不足が考えられます。

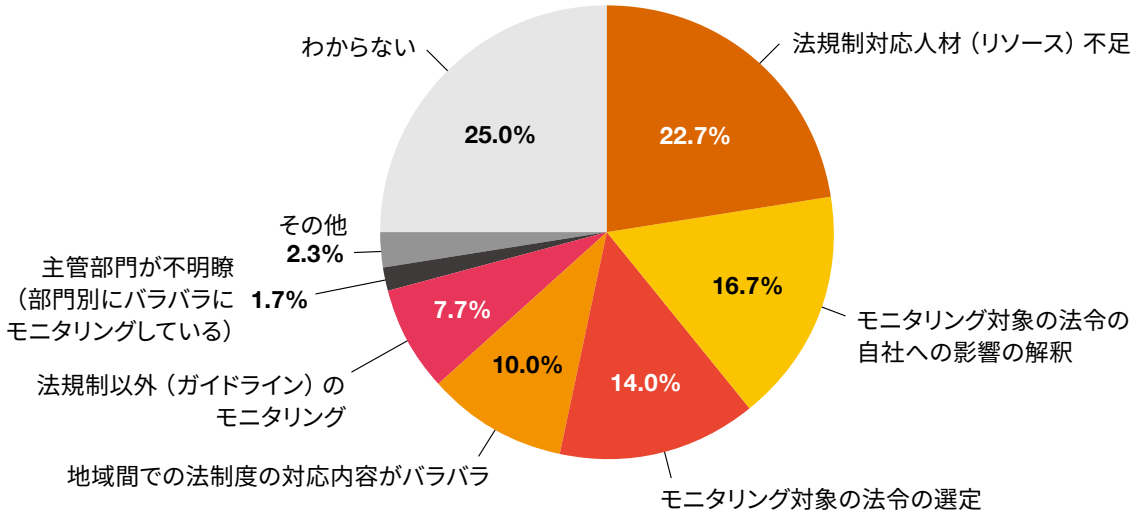
3. 外部専門家を使ったモニタリングは少数（3.2%）。専門家を活用した体系的な法規制モニタリングではなく、自社リソースでモニタリングを実施しているか、実態としてモニタリングができていない可能性が考えられます。

また、サイバーセキュリティ関連の法規制のモニタリングおよび対応を行う上での最大の課題を尋ねました（図表9）。

※3　サイバーセキュリティ関連の法規制の例として、欧州NIS2指令、欧州データ法、欧州CRA、米国CIRCIA、米国Cyber Trust Mark、中国サイバーセキュリティ法、中国データセキュリティ法、日本サイバーセキュリティ基本法等が含まれる。

## 図表9：サイバーセキュリティ関連の法規制のモニタリングにおける難しさ

**Q.** サイバーセキュリティ関連の法規制のモニタリングを行う上での最大の難しさは何ですか。（単一選択、n=300）




※構成比は小数点以下第2位を四捨五入しているため、合計しても必ずしも100とはならない
出所：PwC、「2025年 Cyber IQ調査」

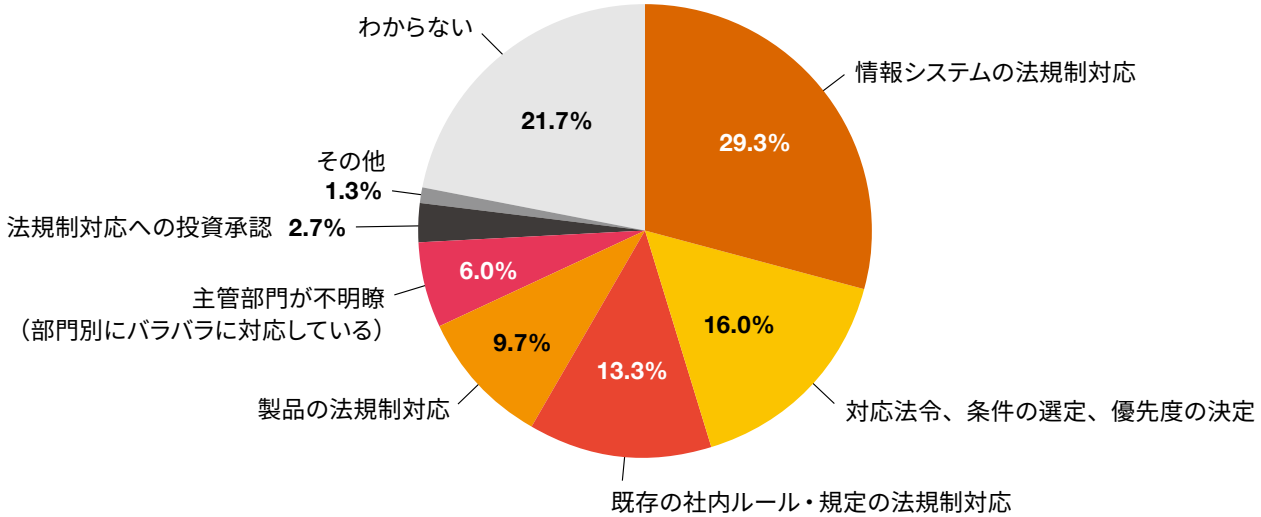
モニタリングにおける最大の難しさとしては、以下2点の示唆がありました。

1.「法規制対応人材（リソース）不足」（22.7%）が最も多く、モニタリングできる人材や部門が不足していることを示しています。

2.「モニタリング対象の法令の自社への影響の解釈」（16.7%）、「モニタリング対象の法令の選定」（14.0%）がリソース不足に次いで高いのは、法規制モニタリングできているとしても、その後の解釈で自社への影響を判断する点において課題があることを示しています。

## 図表10：サイバーセキュリティ法規制の対応における難しさ

**Q.** サイバーセキュリティ関連の法規制の対応を行う上での最大の難しさは何ですか。（単一選択、n=300）




出所：PwC、「2025年 Cyber IQ調査」

モニタリング後に想定されるサイバーセキュリティ関連の法規制への対応で最も大きな課題となるのは、「情報システムの法規制対応」（29.3%）でした（図表10）。これは社内の情報システムをサイバーセキュリティ要件に対応するために更新する（例：要件に合わせて追加のセキュリティ対策を導入するなど）ことの難しさを示しています。また、複数のサイバーセキュリティ関連の法規制に一度に対応することは難しいため、対応するサイバーセキュリティ関連の法規制の優先度を決めなければなりません。そのため、「対応法令、条件の選定、優先度の決定」（16.0%）も2番目に大きな課題として捉えられており、サイバーセキュリティ法規制の抽象的な要件に、多種多様な情報システムを対応させる難しさを示唆する結果となっています。

このような「グローバル」「サイバーセキュリティ」「法規制」全てに対応できる人材を社内で持ち、情報システムや製品をグローバルサイバーセキュリティ法規制に対応させることは、単独の企業で対応するのは非常に困難であることが調査から判明しました。事業を展開している国・地域のサイバーセキュリティ関連の法規制や政策動向に詳しい専門家を活用したモニタリング体制を整え、自社や製品・サービスに対する影響を分析し、専門家のサポートを得て対応することを推奨します。

また、サイバーセキュリティ関連の法規制への対応は、以下の3点から日本の本社が主体となって推進することを推奨します。

1. データの越境に関する規制の存在

近年、データが国境を越えることを制限する法規制が各国・地域で成立しています。欧州や英国のGDPR（一般データ保護規則）は、個人情報を域外に出すことに対して多くの制約を設けています。また中国の「データ越境移転安全評価弁法（数据出境安全评估办法）」のようなデータ越境に関する直接的な法規制なども存在します。しかしながら、データの利活用はビジネスにとってますます不可欠となっており、各国・地域のデータ移転に関する要件を理解しながら、データを集約して活用するには、ビジネス戦略に携わる日本本社が直接的に対応することが推奨されます。

2. 組織横断的な対応の必要性

サイバーセキュリティに関連する法規制は、国・地域で要件が共通する部分も多く、各拠点で個別に対応すると二重対応になる可能性があります。そのため、法規制の共通要件に本社主導で対応し、必要に応じて差分要件を各国拠点が対応する方が効率的です。特に製品セキュリティに関する法規制の中に機能要件が含まれる場合は、対象となるシステムや製品を所管し、製品仕様と法規制の要件を比較できる製品部門の対応が不可欠になり、通常そのような製品部門は本社である日本に存在します。製品部門と法務部門などによる組織横断的な対応は、本社での直接的なイニシアティブが必要になります。

3. グループ全体への罰則規定

欧州GDPRやNIS2指令のような、「全世界年間総売上高」に対する罰則規定は、日本企業のグローバル事業全体に大きな影響が及ぶリスクがあります。そのため、本社が直接的な対応方針を策定し、リスクがありそうな現地法人やシステムに適切に対応することが欠かせません。

以上のような海外子会社のリソースの制約や、サイバーセキュリティ関連の法規制の持つ性質からも、日本の本社がイニシアティブを取り、サイバーセキュリティ法規制への対応を一時的な「コスト」と捉えるのではなく、対象の国・地域でビジネスを長期的に拡大するための「投資」として対応を行うことを推奨します。

# 2-2 グローバルのガイドラインでも取り扱われるAIレッドチームの有効性

## AIサービスの急速な普及

多くの企業の経営計画を見ると、AIサービスの利活用について言及されており、AIサービスをいかに自社の事業に取り込むかが重要な経営アジェンダとなっていることが分かります。実際、グローバルでのAI市場は急速に拡大しており、今後も順調に成長することが見込まれています※4。

その大きなきっかけとなったのが、大規模言語モデルをはじめとした生成AIの登場です。これまでのAIは、AIに関する専門知識を持つ者が、予測や分析などの高度なタスクに対して利用する技術でしたが、生成AIにおいては自然言語による指示が可能であるため、専門的な知見を持たない者も利用できるようになり、活用の幅が専門的なタスクからより身近なタスクへと広がりました。

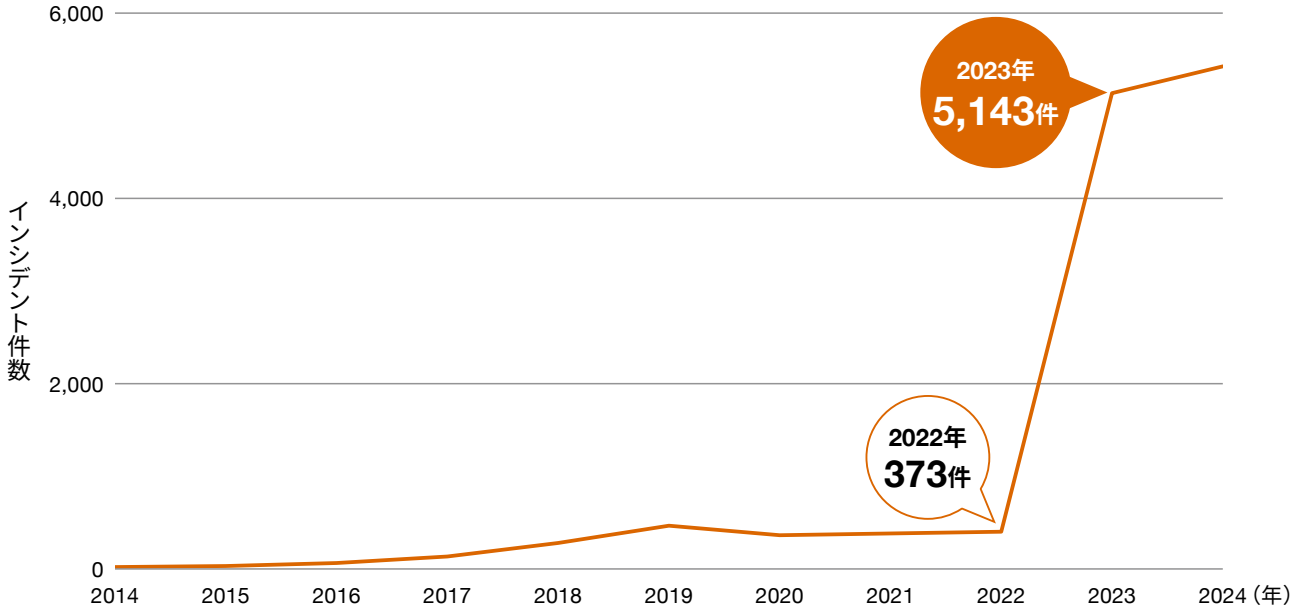
また、大量の学習データ、高い計算処理能力を持つ計算機などの必要なリソースもAIサービスの提供者側が用意するため、利用者はリソースを気にすることなく、スマートフォンなどの身近なデバイスで安価にAIサービスを利用できるようになりました。

## AIリスクの急増

AIの急速な普及に伴い、多くの企業はAIリスクに直面しており、AIに関するインシデントが急増しています。経済協力開発機構（OECD）が公開している「OECD AI Incidents Monitor」によると、生成AIサービスが普及した2023年からAIサービスに関連するインシデントが急増し、2022年においては400件に満たなかったインシデントが、2023年には5,000件を超えたことが報告されています（図表11）。

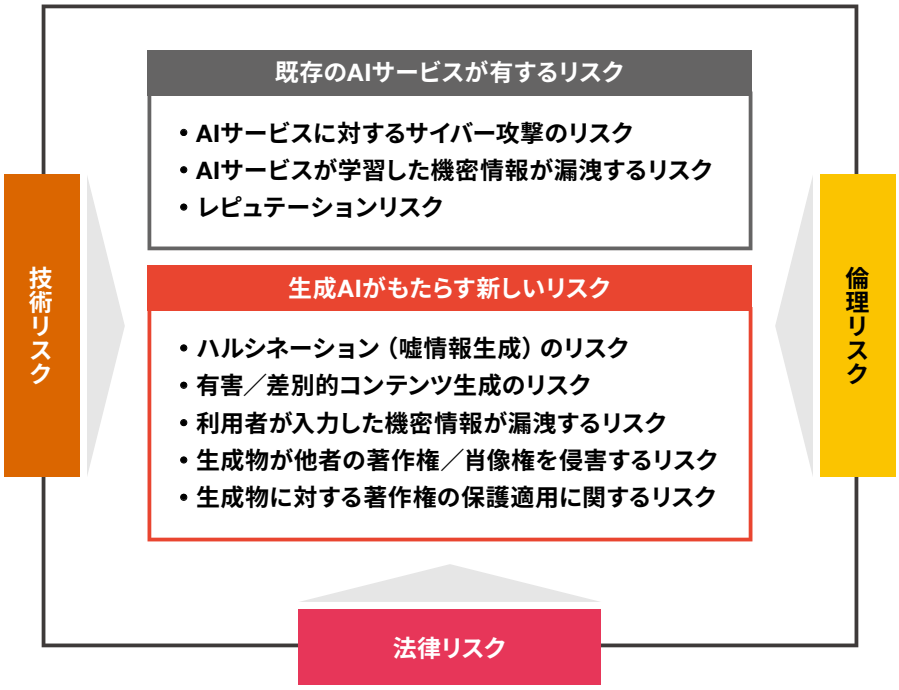
※4　総務省, 2024, 情報通信白書令和6年版, 2024/7/25閲覧, https://www.soumu.go.jp/johotsusintokei/whitepaper/r06.html

**図表11：AIに関するインシデント件数の推移**




では、AIインシデントを誘発するAIリスクとは、一体どのようなものなのでしょうか。従来のITシステムにおいては、サイバー攻撃などを起因とした機密性・完全性・可用性の侵害などが主たるリスクでしたが、AIリスクはその特有の性質から新たなリスクをはらんでいます。特に生成AIにおいては、利用者がアクセスしやすい点、人間らしいアウトプットが得られる点などが特徴として挙げられますが、これによってリスク領域は「技術リスク」にとどまることなく、「法律リスク」や「倫理リスク」と広範にわたります（図表12）。例えば、生成AIは事実と異なることをもっともらしく回答する「ハルシネーション」というリスクをもたらします。ハルシネーションを含んだ情報を利用することで、誤判断や誤情報の流布に発展し、結果としてビジネスに悪影響をもたらす可能性があります。

**図表12：代表的なAIリスク**



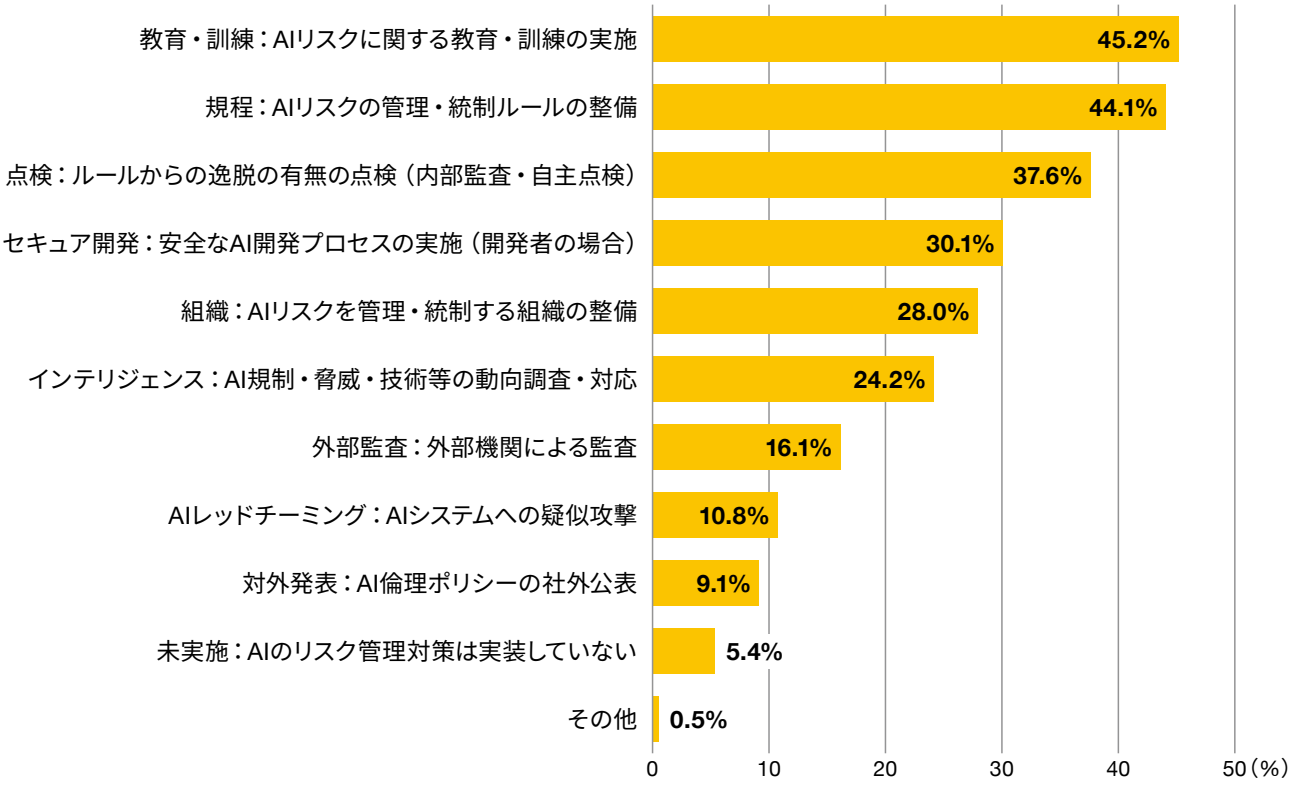
## AIリスクに対する企業の取り組み実態

このようなAIリスクに対して、各企業では具体的にどのような取り組みを実施しているのでしょうか。AIを開発、提供、または利用をしている企業（n=186）が具体的にどのような対策を講じているか整理した結果（図表13）を見ると、「未実施：AIのリスク管理対策は実施していない」と回答した企業は5%程度であり、多くの企業ではAIリスクに対して何らかの対策を講じていることが分かります。対策の中身を見ていくと、おおよそ半数の企業が規定を整備していることが分かります。2024年に総務省、経済産業省は「AI事業者ガイドライン」を公表しており、企業のAIリスク対策の拠り所としての活用が想定されるため、規定の整備はこれからもより多くの企業で進められていくと考えられます。また、おおよそ半数の企業が教育・訓練の実施を行い、4割程度の企業が内部監査や自主点検を実施しており、これらの活動を通して策定したルールの浸透や、ルールの履行に努めている状況が伺えます。また、3割程度の企業は組織を整備し、所管が曖昧になりがちなAIリスクの管理・統制を推進する役割と責任を明確化しています。このように、他のコーポレートガバナンスと同様に、AIというテーマについてもリスクを管理・統制するための組織・人、ルール、プロセスを整備し、AIガバナンスを構築している企業が一定数存在します。一方で見方を変えると、規定を整備している企業が5割に満たないなど、多くの企業においてはAIガバナンスの整備は道半ばともいえるでしょう。AIガバナンスが整備されていない企業においては、「AI事業者ガイドライン」といった外部ガイドラインなどを参照し、まずはAIガバナンスの整備を早急に進めることが望まれます。

さらに、「AIレッドチーミング：AIシステムへの疑似攻撃」を実施している企業は1割程度にとどまっています。ITセキュリティの領域においては、ペネトレーションテストやレッドチーミングといったシステムへ疑似攻撃を仕掛け、実際の脅威・リスクを明らかにし、効果的な対策を講じる取り組みが進んでいます。それに比べるとAIに対してはレッドチーミングはまだ十分に実施されていない状況が伺えます。このような企業においては、自社が実際に活用するAIサービスで想定されるリアルな脅威・リスクを特定し、効果的な対策を講じていくことが、次の一手として必要となってくるでしょう。その活動を支えるのが「AIレッドチーム」であり、今後より多くの企業で導入されていくと考えられます。

**図表13：AIリスクへの具体的な対策状況**

**Q.** 実施しているAIのリスク管理対策について教えてください。（いくつでも、n=186）



## 「AIレッドチーム」の重要性

従来のITサービスにおけるレッドチームの重要性が認知され普及したように、「AIレッドチーム」も今後必須要件として普及していくことが想定されます。従来のITシステムにおけるレッドチームの重要性の多くは、AIレッドチームにも当てはまります。例えば、従来のITサービスにおけるレッドチームでは、実際の攻撃者の視点からリスクを特定するため、ITサービスの性質に即した攻撃シナリオを再現します。これにより実際のサイバー攻撃において顕在化し得る脅威を特定、分析、評価できるため、有効なリスク特定手法と言えます。この点についてはAIサービスにおいても有効です。また、AIは技術の進歩は勿論ですが、それに伴い次々と脅威・攻撃手法が発見、報告されており、従来のITシステム以上に未知のリスクが発見される可能性も秘めています。その上、ルールベースで動作する従来のITサービスと異なり、AIは確率的な動作をすることからリスクを予見しづらく、繰り返し疑似攻撃を仕掛けてみないと発見できないリスクもあるでしょう。

## 「AIレッドチーム」を巡る世界の動向

ここまで「AIレッドチーム」の重要性について、リスクの側面から迫りました。ここではAIを巡る世界の規制強化などの側面からその重要性に迫ります。

AIリスクが拡大する状況に鑑み、国際的にAIの信頼に関する議論が活発化し、各地域・国・団体において法令・ガイドラインの整備が進められています。EUでは「EU AI規制法」が成立し（今後、段階的に適用予定）、規制強化が進んでいます。世界各国・地域でも今後AIに関する法令・ガイドラインの整備による規制強化が加速していくことが予想されます。

既に発行されている法令・ガイドラインを見てみると、「AIリスクを特定する」という趣旨の要件が推奨事項として言及されているケースが多くあることが分かります。そしてその具体的な取り組みとして「AIレッドチーム」が例示・紹介されることが増えてきました（図表14）。

このように、世界の規制強化の側面からも、「AIレッドチーム」の重要性は高まってきていると考えられます。

実際に、米国の先進テクノロジー企業を中心に、法律・経済など、多様なドメインの専門家を外部から募集してAIレッドチームを立ち上げる企業、SMEs（内容領域専門家／Subject Matter Expert）からAIレッドチームを組成し安全保障の専門家へのインタビューを通してリスク領域を特定する企業など、実際にAIレッドチームを組成する企業も出始めてきております。

日本国内においては、独立行政法人情報処理推進機構（IPA）が、2024年2月に「AIセーフティ・インスティテュート（AISI）」を設立しています。この新しい機関は、AIの安全性に対する国際的な関心の高まりを受けて、AIの安全性評価手法の検討や基準の策定を行うことを目的としています。AISIは、2024年9月に、AIセーフティ評価の観点や評価項目例、評価に関する手法の概要などを提示した「AIセーフティに関する評価観点ガイド」や「AIセーフティに関するレッドチーミング手法ガイド」を公開しております。このような状況からも日本国内においても「AIレッドチーム」の重要性の認知は加速していくと考えられます。

**図表14：世界の法令・ガイドラインでの「AIレッドチーム」に対する言及**

| 法規制・ガイドライン | 発行組織 | 時期 | 概要 |
|---|---|---|---|
| • 全AI関係者向けの広島プロセス国際指針※5<br>• 高度なAIシステムを開発する組織向けの広島プロセス国際行動規範※6 | 主要国首脳会議（G7サミット） | 2023年10月 | 高度な AI システムの設計、開発、導入において遵守すべき原則「1. AIライフサイクル全体にわたるリスクを特定、評価、軽減するために、高度な AIシス テムの開発全体を通じて、その導入前及び市場投入前も含め、適切な措置を講じる」の具体的な手法として「レッドチーム」が例示されている。 |
| 安全なAIシステム開発のガイドライン※7 | 英国国家サイバーセキュリティセンター（NCSC）、米国サイバーセキュリティ・インフラセキュリティ庁（CISA）、および各国の機関（内閣サイバーセキュリティセンター＜NISC＞を含む） | 2023年11月 | AIを使用するシステムの提供者向けガイドラインで、AIシステム開発のライフサイクルの4つの主要な領域（安全な設計、安全な開発、安全な展開、安全な運用と保守）に分けて、組織のAIシステム開発プロセスに対する全体的なリスクを軽減するための考慮事項と対策が記載されている。システムをリリースする際には、ベンチマークやレッドチームなどを行うことが推奨されている。 |
| AI事業者ガイドライン※8 | 総務省・経済産業省 | 2024年3月 | 高度なAIシステムに関係する事業者の共通指針「I）AIライフサイクル全体にわたるリスクを特定、評価、軽減するために、高度なAIシステムの開発全体を通じて、その導入前および市場投入前も含め、適切な措置を講じる」の具体的な手法として「レッドチーム」が例示されている。 |
| The Draft NIST Assessing Risks and Impacts of AI (ARIA) Pilot Evaluation Plan※9 | 米国立標準技術研究所（NIST） | 2024年6月 | ARIAはNIST「AI RISK MANAGEMENT FRAMEWORK※10」を拡張し、AIリスクと影響を分析・監視するためのリスク測定機能の実用化を支援するための評価プログラムである。AI技術のリスクと影響を評価するために、「レッドチーミング」を含む3つのレベルのテストとして「モデルテスト」「レッドチーミング」「フィールドテスト」が提示されている。 |
| • AIセーフティに関する評価観点ガイド※11<br>• AIセーフティに関するレッドチーミング手法ガイド※12 | AIセーフティ・インスティテュート（AISI） | 2024年9月 | 「AIセーフティに関する評価観点ガイド」では、AIシステムの開発や提供に携わる者がAIセーフティ評価を実施する際に参照できる基本的な考え方を提示している。<br>「AIセーフティに関するレッドチーミング手法ガイド」では、AIシステムの開発や提供に携わる者が、対象のAIシステムに施したリスクへの対策を、攻撃者の視点から評価するためのレッドチーミング手法に関する基本的な考慮事項を示している。 |

※5　主要国首脳会議（G7サミット）, 2023, 全AI関係者向けの広島プロセス国際指針, 2024/8/6閲覧, https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document04.pdf

※6　主要国首脳会議（G7サミット）, 2023, 高度な AI システムを開発する組織向けの広島プロセス国際行動規範, 2024/8/6閲覧, https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document05.pdf

※7　2023, Guidelines for secure AI system development, 2024/8/6閲覧, https://www.ncsc.gov.uk/collection/guidelines-secure-ai-system-development

※8　総務省・経済産業省, 2024, AI事業者ガイドライン（第1.0版）, 2024/8/6閲覧, https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20240419_1.pdf

※9　米国立標準技術研究所（NIST）, 2024, The Draft NIST Assessing Risks and Impacts of AI (ARIA) Pilot Evaluation Plan, 2024/8/6閲覧, https://ai-challenges.nist.gov/aria/docs/evaluation_plan.pdf

※10　米国立標準技術研究所（NIST）, 2024, Artificial Intelligence RISK MANAGEMENT FRAMEWORK, 2024/8/6閲覧, https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf

※11　Japan AI Safety Institute（AISI）, 2024, AIセーフティに関する評価観点ガイド https://aisi.go.jp/assets/pdf/ai_safety_eval_v1.01_ja.pdf

※12　Japan AI Safety Institute（AISI）, 2024, AIセーフティに関するレッドチーミング手法ガイド https://aisi.go.jp/assets/pdf/ai_safety_RT_v1.00_ja.pdf

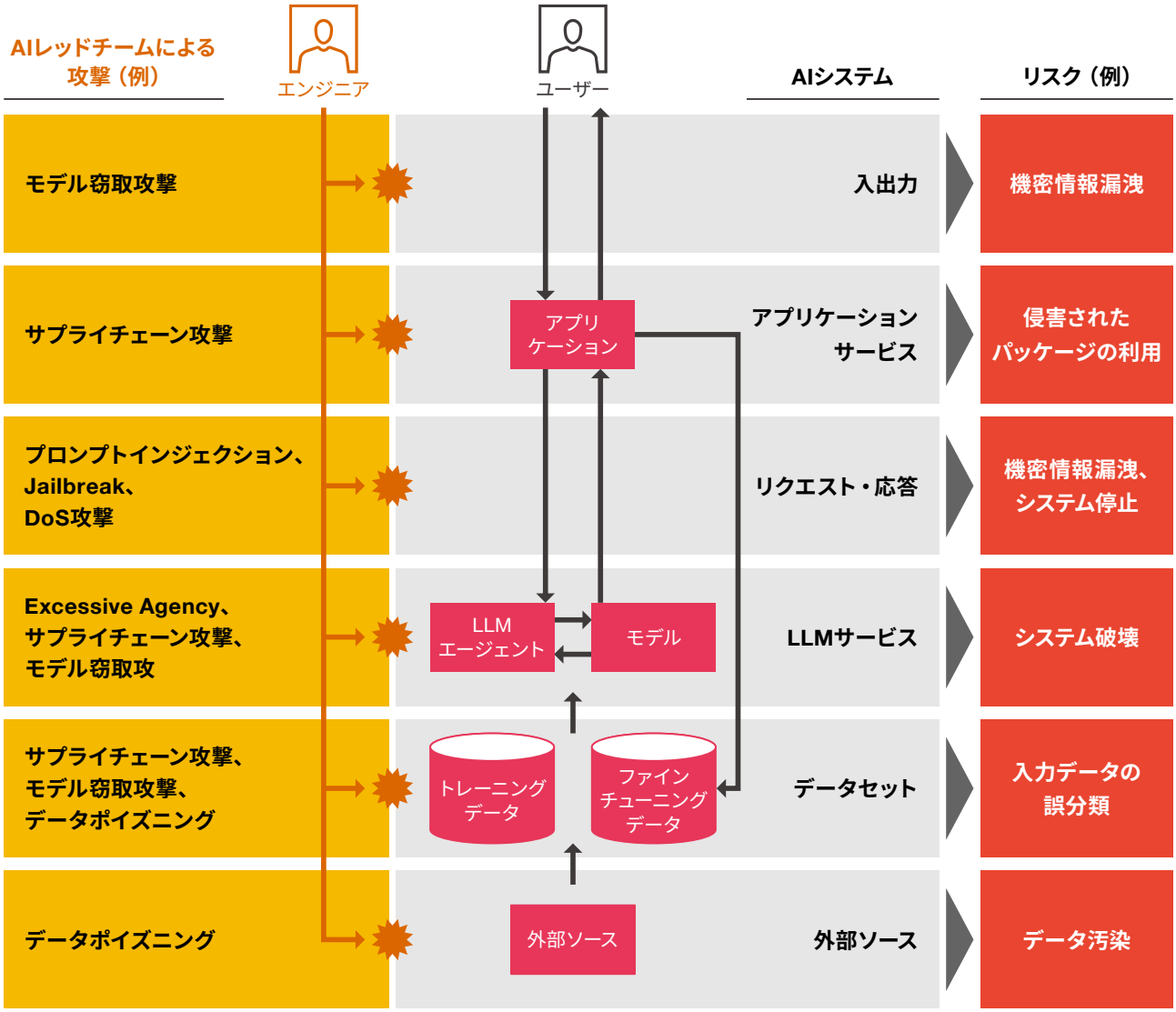
## 「AIレッドチーム」とは

ここでは、「AIレッドチーム」が具体的にどのような取り組みであるかを紹介します。

「AIレッドチーム」とは、AIを利用したサービスに対して、リスク起因者（サイバー攻撃者、犯罪者、愉快犯など）の立場からエンジニアが高度な疑似攻撃を行うことで、脆弱性とそれに伴うビジネスリスクを特定する組織や取り組みを指します（図表15）。AIサービスのインシデントを未然に防ぐため、リリース前や運用中に脅威を評価し、脆弱性・リスクを特定し、効果的なセキュリティ対策を講じることを目的として実施します。

「AIレッドチーム」の代表的なアプローチは、「1. テスト計画」「2. テスト（疑似攻撃）の実施」「3. 対策定義」から成るものです。「1. テスト計画」では、テスト対象となるAIサービスや、当該サービスの特性やシステムなどを踏まえ蓋然性のあるリスクシナリオを特定し、テスト計画を立案します。「2. テスト（疑似攻撃）の実施」においては、立案したテスト計画に基づき、エンジニアがテスト（疑似攻撃）を実施し、脅威が成立し得るか、AIサービスに脆弱性やリスクがあるかを特定します。そして「3. 対策定義」では、テストで検出したAIサービスの脆弱性やリスクを踏まえ、講じるべき対策を定義します。

**図表15：AIレッドチームによるテスト（疑似攻撃の実施）イメージ**



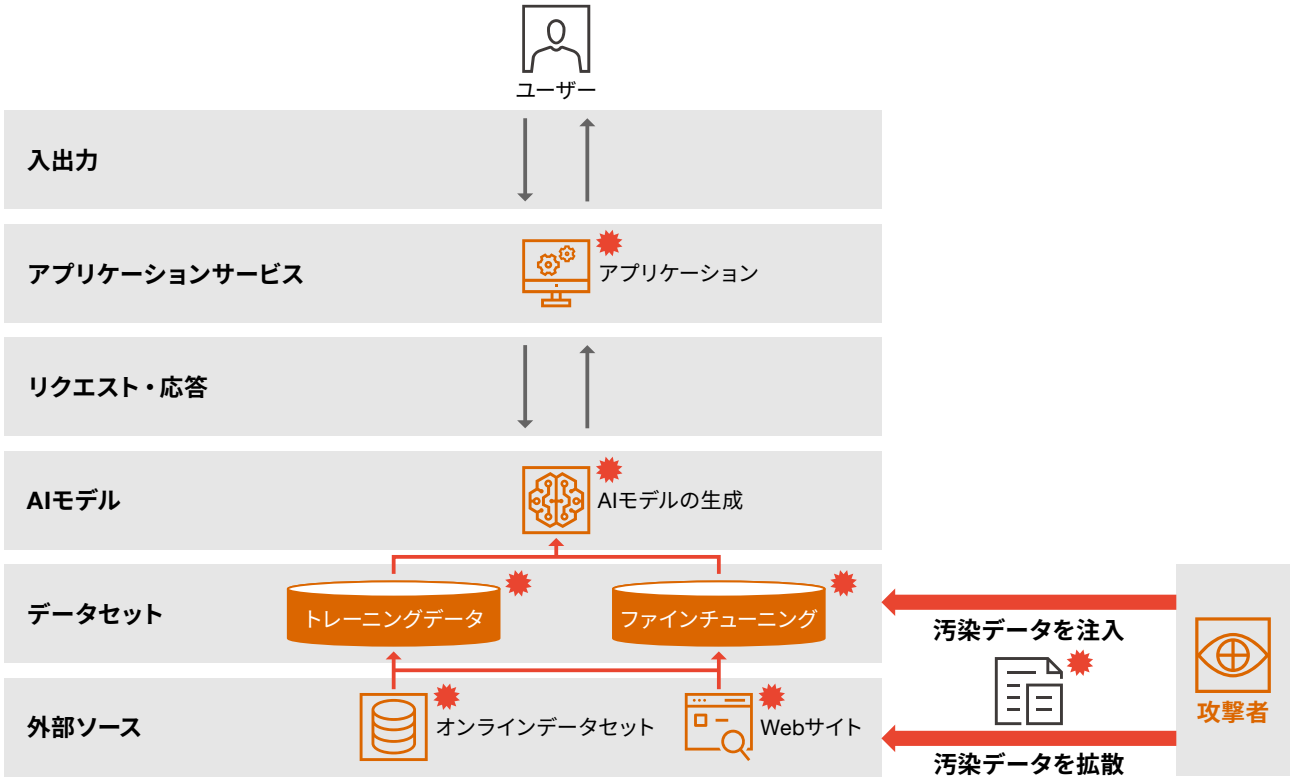
「AIレッドチーム」でテストする疑似攻撃として「Data poisoning attack（データポイズニング攻撃）」「Membership inference attacks（メンバーシップ推論攻撃）」を紹介します。

## Data poisoning attack（データポイズニング攻撃）

Data poisoning attack（データポイズニング攻撃）（図表16）とは、攻撃者が不正なデータを学習データへ直接または間接的に注入することで、学習したAIの出力結果にバイアス、エラーを誘発したり、AIモデルの意思決定や予測機能の精度に負の影響を与えたりする攻撃を指します。学習データへの不正なデータの注入は、AIの学習データへ直接アクセスして注入するケースや、AIモデルを生成する際のデータセットとなる外部ソースのオンラインデータセットならびにウェブサイトに不正なデータを注入しておくケースなどが存在します。

例えば、自動運転システムで用いられるAIに対してこの攻撃が成立した場合、汚染されたデータで学習することによりAIのモデル精度が低下し、車両の制御に悪影響をもたらして深刻な事故の発生につながる可能性があります。この例は安全性に関するリスク例ですが、モデル精度の低下は、AIサービスによっては倫理リスクや法律リスクを招く可能性もあります。

**図表16：Data poisoning attackのイメージ**

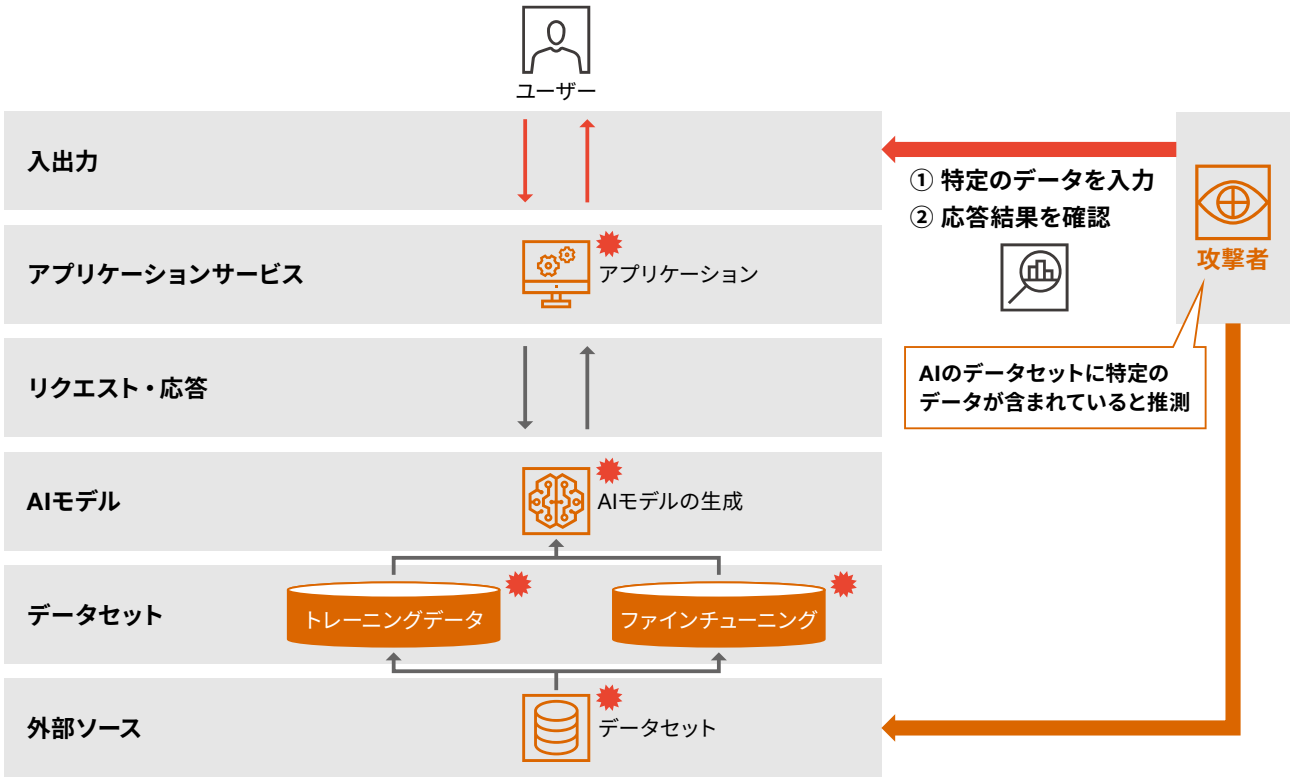



## Membership inference attacks（メンバーシップ推論攻撃）

学習データを過学習すると、AIに入力したデータが学習データに含まれる場合とそうでない場合に、AIの応答結果の特徴が異なる（信頼度スコアの偏り）ことが知られています。Membership inference attacks（メンバーシップ推論攻撃）（図表17）とは、攻撃者がこの特徴を悪用して、入力データが学習データに含まれているか否かを推測する攻撃です。

例えば、医療の研究などの目的で患者の病歴などのデータを学習させているAIサービスに対し、この攻撃が成立した場合、特定の患者が学習データに含まれることが推測され、病歴などが攻撃者に知られてプライバシー侵害事案に発展する可能性があります。

**図表17：Membership inference attacksのイメージ**




次に「AIレッドチーム」で効果的な成果を上げ、成功に導くために重要なポイントを紹介します。

1. AIサービスに関わるビジネスリスクを特定できるか

AIリスクは「技術リスク」にとどまることなく、「法律リスク」や「倫理リスク」と多岐にわたります。そのため、AIシステムの「技術リスク」のみに視点を当ててテスト計画を立案すると、「法律リスク」「倫理リスク」などを想起できず、テスト計画から蓋然性のあるリスクシナリオが漏れてしまう懸念が生じます。

AIの不正利用に伴って考慮すべきリスクとインパクトは、そのビジネスユースケースに依存するため、AIシステムのみに視点を当てるのではなく「ビジネスユースケース」から想起されるリスクシナリオを特定できるかが重要です。

2. AI特有のリスクを特定するためのテスト（疑似攻撃）を実施できるか

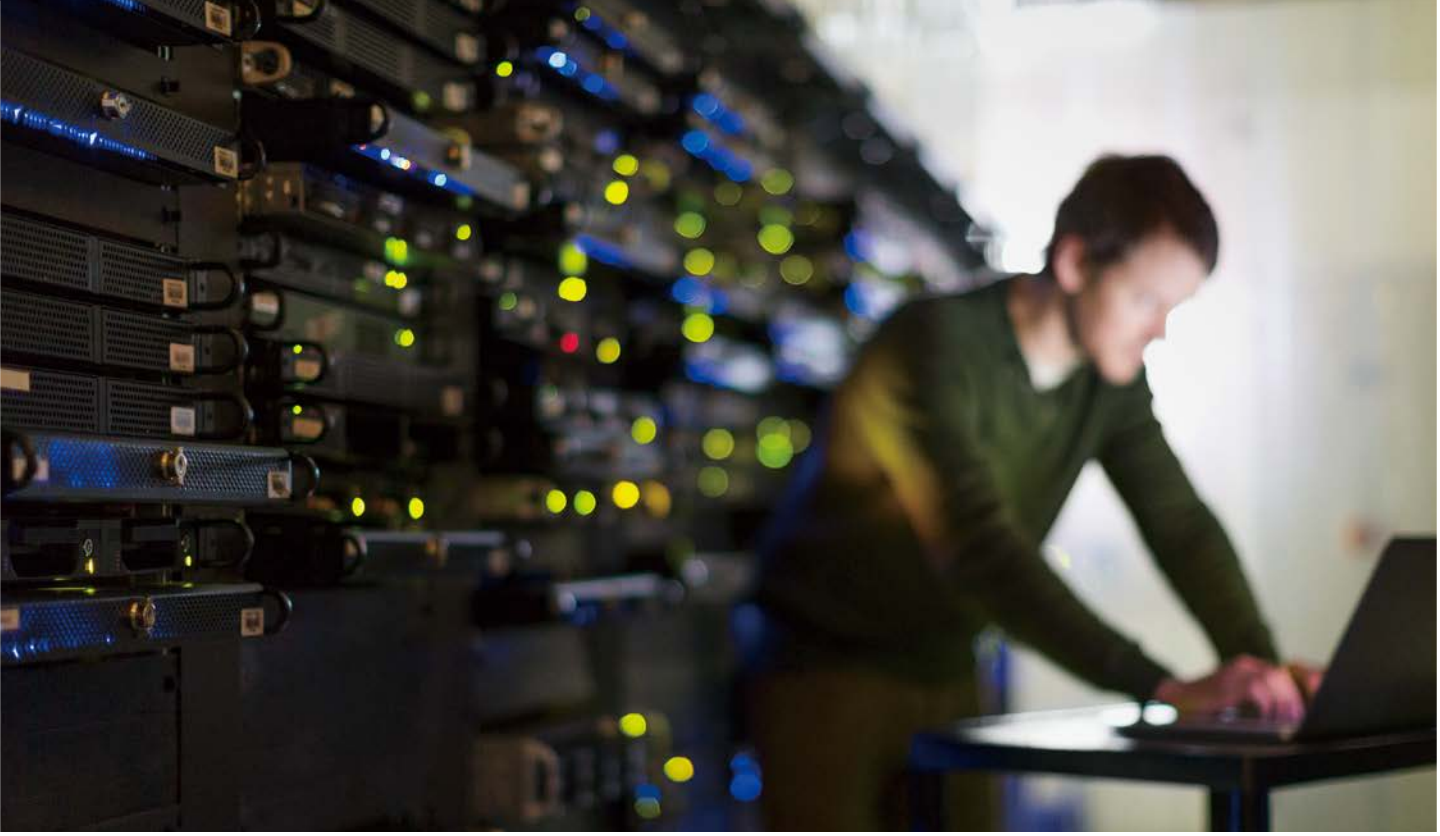
AIは確率的プロセスに基づいて動作するなど、仕組み・挙動・脆弱性を生むポイントなどが従来のITサービスとは異なります。また、AIに関する新たな攻撃手法は日々発見・報告されています。このためAIサービスのリスク特定に求められる専門性やその手法も従来のITサービスとは異なります。AI独自のリスクを効果的に特定するためには、AIに精通するエンジニアによるテストが必要不可欠と言えるでしょう。

3. 多角的な対策（サービス設計、実装、運用）を定義できるか

検出された課題に対して対症療法となる対策のみを講じるのではなく、将来的にリスクを生みづらい管理態勢の整備や全体最適を踏まえたROIの高い対策などを検討することも重要です。例えば、課題の表面だけを見て直接的な実装レベルの対策（例：入出力のフィルター）のみを行うのではなく、AIガバナンスの整備・改善、MLOps（Machine Learning Operations）の改善・高度化も含めたサービス設計、実装、運用と多角的に対策を検討していくことが重要です。

## 「AIレッドチーム」の今後の展望

AIリスクの増大、今後の世界の規制強化、日本国内での重要性の認知の加速に伴い、「AIレッドチーム」はより一般的なものとなり、米国などの先進テクノロジー企業だけでなく、日本国内の多くの企業でも実施する企業が増加していくことが予想されます。AIガバナンス態勢を構築した企業には、是非「AIレッドチーム」の実施を推奨します。AIサービス固有のリスクを特定し、有効な対策を講じることはもちろんですが、今後のビジネスドライバーのひとつである「AI」についての理解をより一層深めるきっかけとなるでしょう。



# 2-3 サプライチェーンを脅かすデジタルリスクに企業はどう立ち向かうべきか

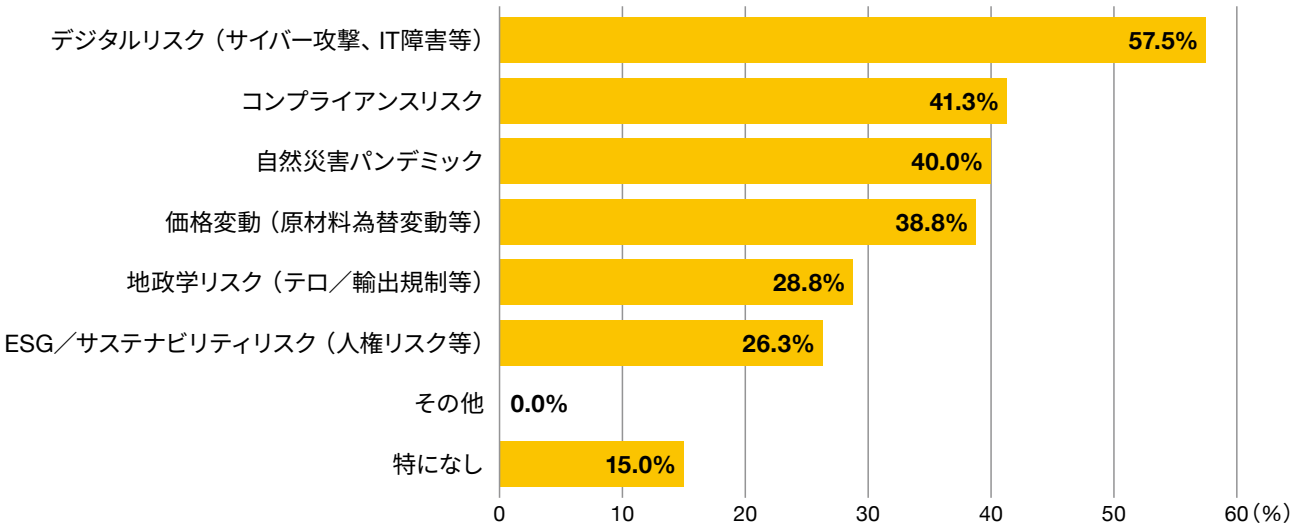
## サプライチェーンを脅かすデジタルリスク動向

自然災害、政治・地政学リスク、環境問題、人権問題などによりサプライチェーンを取り巻く外的環境が不安定になるなか、サイバー攻撃をはじめとしたデジタルリスクが顕著に高まっています。企業のサプライチェーンは、調達先（ハードウェア／ソフトウェア）やサービスプロバイダー、オープンソースソフトウェア（OSS）、業務委託先など複数のサードパーティにより複雑に構成されているため、自社のみならずサプライチェーン全体のデジタルレジリエンスの確保が重要になっています。また、それを受けて各国政府や国際機関はデジタル領域における規制やガイドラインを強化しており、企業はこれらの規制への対応が求められています。

サプライチェーンリスクに対する企業の対応状況を把握するため、PwCはサプライチェーン（調達先や業務委託先）を統制する立場にある経営層80名を対象に、具体的にどのようなリスクを懸念しているかアンケート調査を行いました。調査の結果、サプライチェーン上のリスクのうち「デジタルリスク」を懸念していると回答した経営層が最も多く57.5%、次いで「コンプライアンスリスク」が41.3%、「自然災害・パンデミック」が40.0%となりました。この結果を踏まえても、サイバー攻撃をはじめとするデジタルリスクへの懸念がここ数年で顕著に高まっていると考察できます（図表18）。

**図表18：経営層が懸念するサプライチェーン上のリスク**

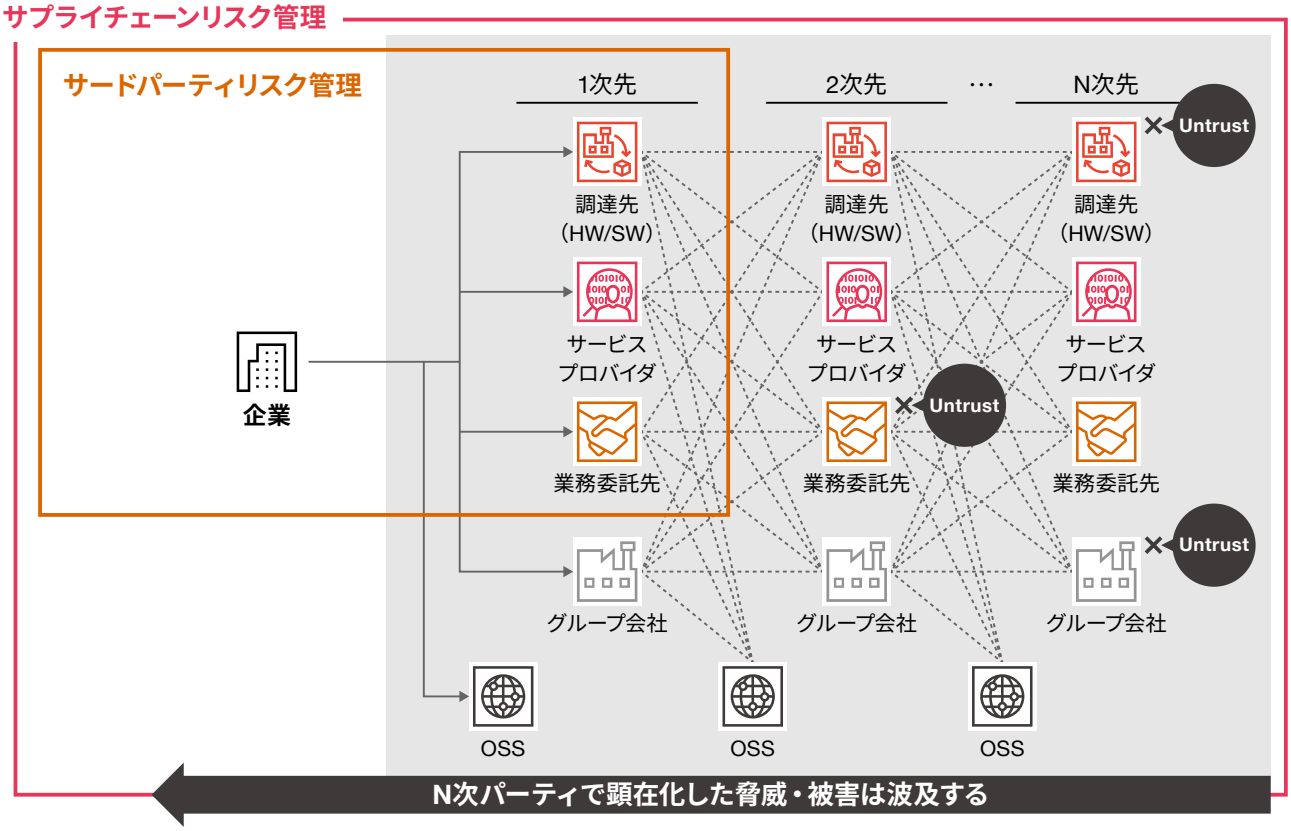
Q．貴社（経営者）が懸念するサプライチェーン上のリスクとして、あてはまるものを全てお選びください。（いくつでも、n=80）




## サプライチェーンリスクの特性

自社や直接の取引先が信頼できたとしても、信頼できない（Untrsutな）取引先が1社でもあると、チェーンを通じて脅威・被害が波及することがあります（図表19）。自社と直接取引のないフォースパーティ以降の調達元・委託先のリスクや、自社が購入したソフトウェア／ハードウェア製品に含まれている脆弱な要素への統制は手薄になりがちです。サプライチェーンのガバナンスを「どの範囲」で「どのレベル」まで徹底していくかが、対策の検討の大きな論点になっています。

**図表19：サプライチェーンリスクの特性**



## サプライチェーンデジタルリスクに関する国内外の法規制・制度

グローバルにおけるサプライチェーンリスクマネジメントの重要性は今後ますます高まり、ガイドライン主流（soft law）から法令主流（hard law）へのシフトが加速していくでしょう。特に、重要インフラ事業者に対するサプライチェーン要求は一層厳重となり、違反時には膨大な制裁金が科されることが予想されます。

具体的なデジタル法規制の例を図表20に示します。欧州では「ネットワーク・情報システムの安全に関する指令（NIS指令）」の修正であるNIS2指令が制定され、2024年10月から適用されました。NIS2指令では重要インフラ事業者に該当する法人に対して製品やサービスを供給するサプライヤー経由でインシデントが発生することを防ぐために、サプライチェーンを含めたセキュリティリスク管理が求められています。

日本においても、経済安全保障の強化を目的に2022年5月、経済安全保障推進法が成立しました。電力やガス、水道や石油の他、鉄道や電気通信、金融など、社会生活を支える基盤を持つ業種を対象に特定重要設備を導入・更新したり、維持管理を外部委託したりする際に所管官庁への事前の届け出が必要になります。

また、諸外国ではサイバー対策の可視化やレーティング（格付け）の動きが活発化しており、さまざまな制度が確立されています。国内においても、IPAが2024年9月末にIoT製品に対する「セキュリティ適合性評価及びラベリング制度（JC-STAR）」を公開し、同制度は2025年春より運用が開始されます。その他、経済産業省にて「サプライチェーン強化に向けたセキュリティ対策評価制度」に関する議論が開始しており、このような格付け制度に対して「企業がどう向き合うか」「どう活用するか」が今後の重要な論点になりつつあります。

**図表20：サプライチェーンデジタルリスクに関連するデジタル法規制の例**

| | | 欧州<br>ネットワークおよび情報システムに関する指令（NIS2指令） | 中国<br>サイバーセキュリティ法に基づくインフラ保護規則／インシデント報告弁法 | 日本<br>経済安保推進法に基づく基幹インフラ役務の安定的な提供の確保に関する制度 |
|---|---|---|---|---|
| 重要インフラ事業者に該当する法人の義務 | セキュリティ対策 | ✓ | ✓ | ✓ |
| | サプライチェーン | ✓ | ✓ | ✓ |
| | インシデント報告 | ✓ | ✓ | 義務ではなく推奨 |
| | 主要な罰則 | ■ Important Entity（製造業含む）<br>• 罰金として700万ユーロ、あるいは適用対象の事業体が属するグループの全世界での年間売上の1.4%のいずれか高い方（ICTサービスマネジメントなどEssential Entityは1,000万ユーロか全世界売上の2%） | • 当局からの是正措置に応じない場合、停止や罰金最大100万元（責任者は最大10万元）<br>• 安全でない機器などの調達については、調達金額の10倍以下（責任者には最大10万元） | • 情報開示要求、立入検査<br>• 調達先の見直し勧告、調達停止命令 |
| サプライヤーに要求される可能性がある協力事項 | | • 契約上の要件（情報収集・共有、下請含むセキュリティ管理）<br>• 製品・サービス提供時の認証取得（特にICT製品） | • 契約上の要件（情報収集・共有、下請含むセキュリティ管理）<br>• 納品前のセキュリティ審査 | • 契約上の要件（情報収集・共有、下請含むセキュリティ管理）<br>• 納品前のサプライチェーンなどの審査 |

## サプライチェーンリスク対応のリアルな実態

サプライチェーンリスクに対する取り組みが求められる一方で、各企業は「サプライチェーン可視化の幅（どこまでをスコープとして管理すればよいか）」や「統制の強度・深さ（どの程度のレベルで統制を効かせるか）」「対策検討・推進にかかるマンパワー、運用工数の負荷」などに難しさを感じています。

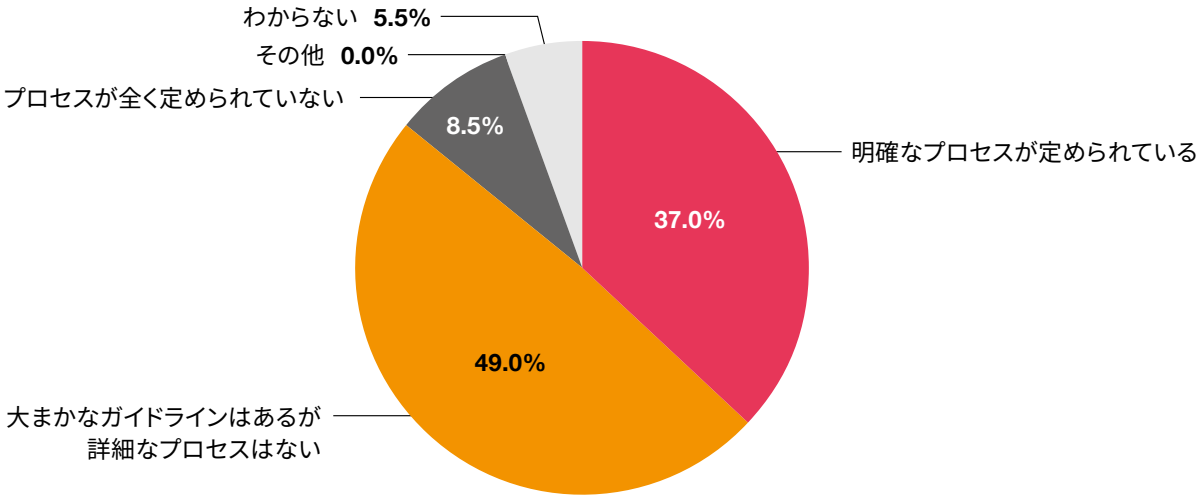
具体的に現場担当者から聞こえてくる「生の声」の例として、以下が挙げられます。

- 全ての取引先を対象とすべきか、再委託先までを対象とすべきか、適切な対象設定が難しい
- チェックシートを導入しているものの自己申告のため効果が不透明で、信頼性にも疑問が残る
- 規模や業種の異なる取引先に対して統一の基準を設けることが難しい
- 下請法などの遵守も含め、どの程度までセキュリティ対策を強制してよいのか分からない
- 数千～数万社の取引先を管理する負荷が大きい。対策を行うにも、多くの部署を巻き込む必要があり、膨大な工数がかかる

上記のような難しさが障壁となり、日本企業はサプライチェーンリスク管理に多くの問題を抱えています。サプライチェーン（調達先や業務委託先）を統制する立場にある日本企業200社に対し、取引先企業でインシデントが発生した際の対応プロセスがどの程度整備されているかアンケート調査しました。「大まかなガイドラインはあるが詳細なプロセスはない」「プロセスが全く定められていない」「わからない」と回答した企業が6割を超えており、サプライチェーン全体を意識したリスクガバナンスは、未だ発展途上であることが読み取れます（図表21）。

**図表21：取引先企業でインシデントが発生した際の対応プロセス整備状況**

Q. 貴社では、取引先企業のインシデント発生時の対応プロセスは現在どのように定められていますか。（ひとつだけ、n=200）




## 企業はどう立ち向かうべきか？解決の方向性

それでは、サプライチェーンに対して統制する立場にある企業は、どのようにしてリスク強化を図ればよいのでしょうか。

「サプライチェーンデジタルリスクに課題がある」と回答した日本企業の経営層56名に対し、どのような打ち手が必要と考えているかアンケート調査しました。その結果「リスクに応じて濃淡をつけた統制（ガバナンス）強度の設定」が必要と回答した経営層が53.6%と最も多い結果となりました。総花的なリスク対応ではなく、優先リスクへの実効的な対応を多くの企業が重視していることがこの結果から読み取れます（図表22）。

濃淡をつけた統制（ガバナンス）強度の設定を重視する背景には、前述した課題である「サプライチェーン可視化の幅（どこまでをスコープとして管理すればよいか）」があると推察されます。自社を取り巻くサプライチェーン全てに一律の対策を講じることは非常に困難です。サプライチェーンを統制する企業は、サプライチェーンを分類し、取り扱う情報やサービスの重要性（例．保有情報や事業規模）などを捉えた上で、濃淡のあるリスク対応を行うことがファーストステップです。

濃淡のあるリスク対応の実現に向け、例えば以下のようなアプローチが有効です。

1. 全体像の可視化・スコープ定義

   サプライチェーンの全体像を整理し、自社における重要なサプライチェーン（例．重要業務、重要な委託先、重要な部品など）を定義する

2. 事業環境を踏まえたリスク分析

   ガイドライン（モノサシ）へのFit&Gapだけでなく、事業環境（取り扱う情報、システム環境など）を勘案したリスク分析・深掘りを行う

3. 課題の抽出・優先課題の設定

   可視化したリスクに対して課題を抽出し、課題の優先順位を設定する

4. 優先課題への取り組み

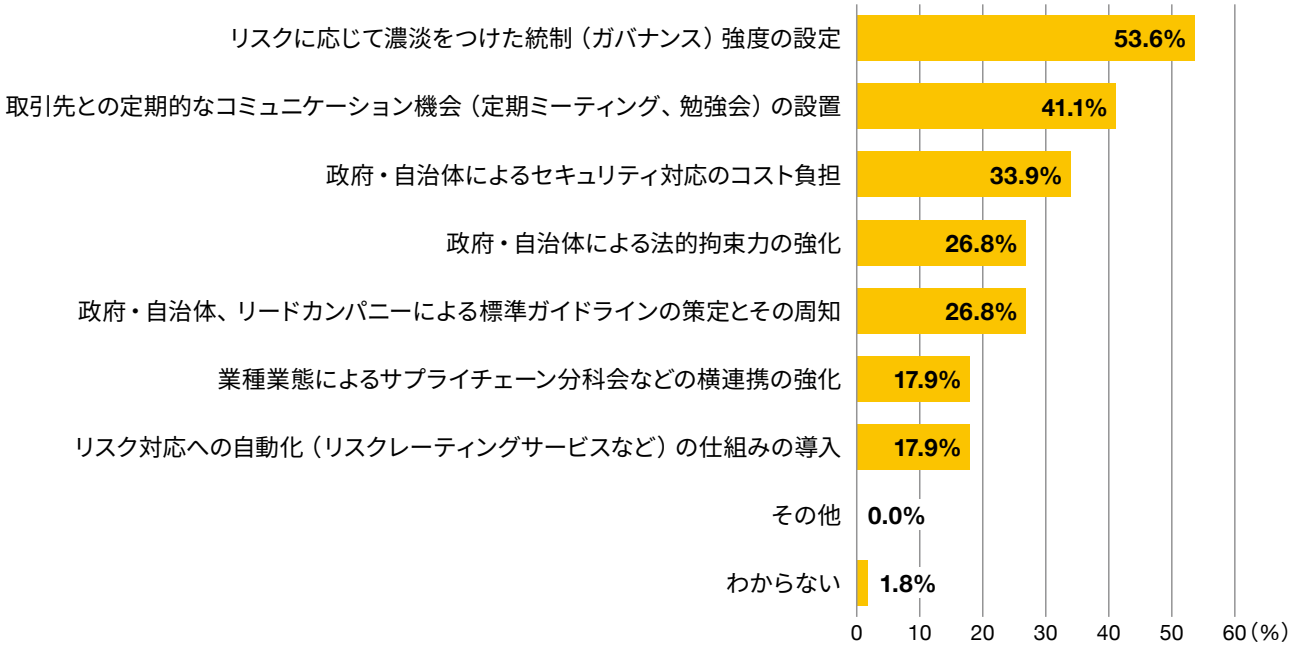
   サプライチェーンを脅かすクリティカルなリスク・課題に対し、優先的に経営リソース（ヒト・モノ・カネ）を投入して推進する

その他にも「取引先との定期的なコミュニケーション機会の（定期ミーティング、勉強会）設置」が41.1%で次点となっています。従来のような「一律的なセキュリティ要求リストを用意し、あとは取引先の選定および契約で縛る」といった役割分担論主体のガバナンスではなく、実務的な連携を通じた共助体制の構築が重視されていることが、この結果から分かります。

取引先とのコミュニケーションにおいては、対策にかかるコストや取引先の体力を理解し、関係者における合意を得るための工夫や配慮が大切です。例えば「シンプルで分かりやすいルール設計」を意識し、取引先に合理的な範囲を超えた要求を行うことがないよう方針を展開したり、「自律的統制を促すような教育と研修」を意識して実行に向けた助言を行っていく、などの工夫が有効です。

**図表22：経営層が考える課題解決に必要な打ち手**

**Q．**課題や障壁を乗り越えるために、必要と考えている打ち手をお選びください。（いくつでも、n=56）



## サプライチェーンリスク対応に挑む企業の先進事例

サプライチェーンリスク管理の体制やプロセスが一定レベルに成熟してきた企業は、「効率化」や「自動化」に取り組み始めています。そこで新たな武器として注目を集めているのが、セキュリティレーティングサービスです。レーティングサービスは、攻撃者と同じ目線でサプライヤーのリスクを評価・スコアリングするサービスであり、サプライチェーンマネジメント市場の中で高い成長が期待されています。

このサービスは、タイムリーかつ客観的、人手によらない自動評価という多くのメリットがある一方で、下記に挙げるような難所を理解し、適切に使いこなすことが重要です。

1. 公開範囲によっては評価も限定的になる

一般的にIPアドレスやドメイン情報をもとに評価することから、子会社と親会社が同じドメインを利用していると、セキュリティレーティングサービスから見て判別が難しくなります。そのため公開システムや公開Web・ドメインの有無や関係によって、当該サービスによる評価結果が限定的になったり、詳細な結果が得られなかったりする可能性があります。

2. 評価が困難な領域が存在する

客観的に評価可能なサービスであるものの、公開情報より評価を行うことから、その企業における統制状況や体制の規模感、業務や規定の整備状況、人員の持つ知見・対応レベルといった内部的な面の評価は困難なことを理解しておく必要があります。

3. 評価結果を読み解くために専門的なスキルが必要

セキュリティレーティングサービスは、それぞれのサービスごとに独自の算出ロジックによって企業の危険度レベルなどを分かりやすく評価する一方で、算出ロジックの一部として公開されるシステムの脆弱性などの技術的な側面を取り扱うことから、評価結果の詳細を理解するには、一定度のセキュリティ知見を保有していることが求められます。

一部の先進企業では、レーティングサービスの活用を進めており、例えば「クリティカルな問題を持つ取引先のスクリーニング」をレーティングサービスで実現している事例があります。さまざまな事業を展開している企業の取引先数は多い場合で数万を超えます。これらの企業では数万ある取引先の中で、できるだけ工数をかけずに、クリティカルな問題を持つ企業を特定したいという課題がありました。そこで、生産影響のある企業に対象を絞る（1次スクリーニング）、レーティングサービスで一定の閾値以下の企業に絞る（2次スクリーニング）、そこからさらに取引先チェックシートで詳細を点検する（3次スクリーニング）という運用を構築し、膨大なセキュリティ評価業務を簡略化に成功しました。

レーティングサービスは決して万能ではありません。適用する範囲とそれに期待する目的を明確にすることで、よりスピーディーかつ効果的に取引先のセキュリティレベルを向上させることが期待できます。従来は会社独自でチェックシートを作成し、取引先の評価を行うことが一般的でしたが、レーティングサービスの台頭により取引先のチェック運用を専門サービスに任せる流れが一気に加速しました。今後、取引先評価業務において、レーティングサービスが当たり前のように活用される未来もそう遠くないかもしれません。

## まとめ

サプライチェーンを脅かすデジタルリスクが高まっています。取引先の1社が攻撃を受けるだけで自社の操業停止に陥ってしまうケースもあり、サプライチェーン全体を意識したリスクガバナンスは経営課題の一つになっています。課題は多岐にわたるため、総花的なリスク対応ではなく、クリティカルなリスクを見極め、より実効性のあるリスク対応を意識することが必要です。取引先とのコミュニケーションにおいては、対策にかかるコストや取引先の体力を理解し、合意を形成するための工夫や配慮を行うなど、今まで以上に踏み込んだガバナンスが求められています。

サプライチェーンリスクマネジメントは、長い道のりになることを覚悟しなければなりません。検討・推進に2、3年のロードマップを掲げる企業も多く見られます。工数も時間も多大にかかるからこそ「納得感の醸成」と「実効性の担保」を意識し、初めから100点を目指すのではなくスモールスタートで、一歩一歩、成果を出していくことが重要です。

# 2-4 脆弱性対応プロセス改善への取り組み

―意思決定機能の成長を実現するためのAcuity Rampというコンセプト―

## 脆弱性管理の現状

「2024年 Cyber IQ調査：第1章-4 脆弱性対応プロセスのトレンドと変化の必要性」でもお伝えしたとおり、脆弱性の悪用へ対応するための脆弱性対応プロセス整備は組織にとって重要な課題です。新たな評価方式であるSSVC（Stakeholder-Specific Vulnerability Categorization）を採用している組織や、移行を検討している組織も確認され始めています。

このような状況の中、SSVCの提案元である米カーネギーメロン大学ソフトウェア工学研究所からAcuity Rampというコンセプトが公開されました。これは、SSVCを導入した後、組織のニーズやリソース、能力が変化する中で、意思決定機能をよりステークホルダーの望む形へ成長させるための考え方です。

本節では、SSVCの現状と課題に言及し、Acuity Rampを解説するとともに、その有用性について考察します。

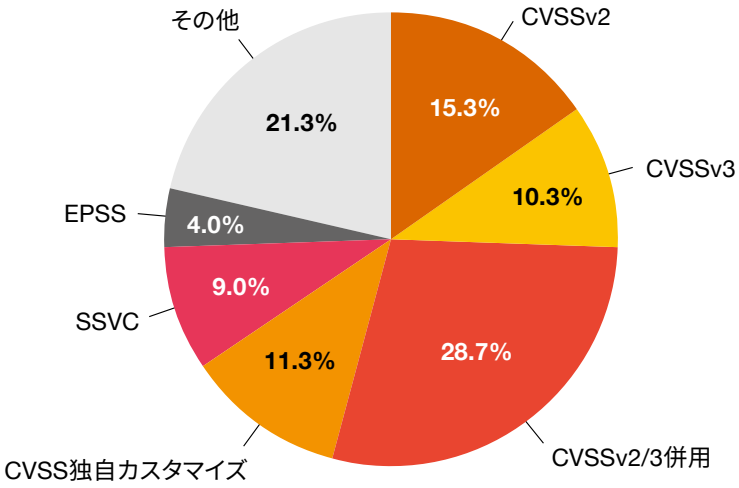
## SSVCの現状

SSVCは、米カーネギーメロン大学ソフトウェア工学研究所によって提唱された脆弱性評価指標です。脆弱性のリスク値といったスコアではなくステークホルダーごとの具体的なアクションを決定木に基づいて導出する方法論となっており、従来のデファクトスタンダードであるCVSS（Common Vulnerability Scoring System）の課題へ対応するために提案されました。

今回の調査で脆弱性対応プロセスに採用している評価方式を確認した結果、9.0%がSSVCを採用していることが確認されました（図表23）。

## 図表23：SSVCの採用状況

**Q.** 脆弱性対応プロセスに現在採用している評価方式（1つだけ、n=300）



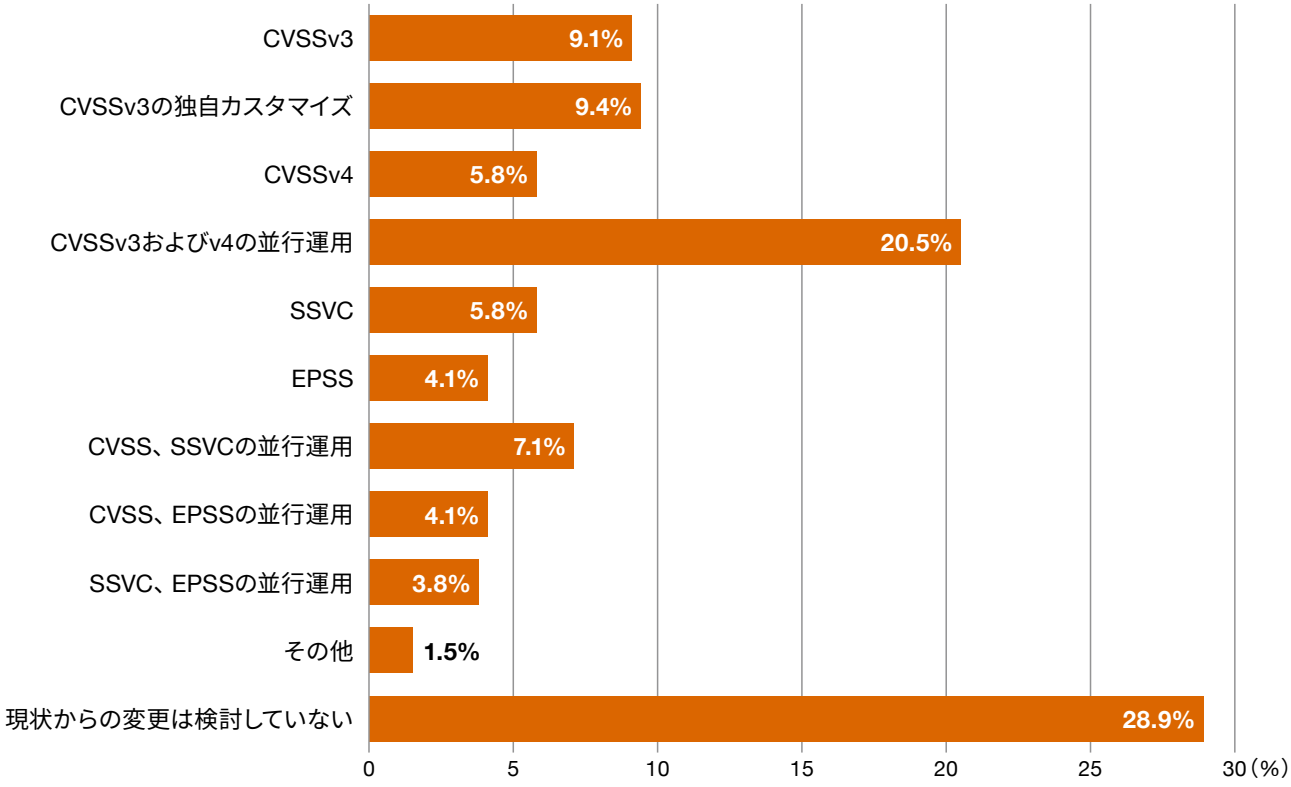

※構成比は小数点以下第2位を四捨五入しているため、合計しても必ずしも100とはならない

また、CVSS（v2、v3、カスタマイズ含む）を採用している回答の割合は、65.6%となり、依然としてCVSSが主流であることが分かります。

一方で、今後採用を検討している評価方式に関して、SSVCの導入を検討している回答の割合は16.7%（複数選択の総計におけるSSVCを含む回答の割合より算出）となり、潜在的なニーズが伺えます（図表24）。

## 図表24：今後採用を検討している評価方式

**Q.** 今後採用を検討している評価方式（いくつでも、n=395）

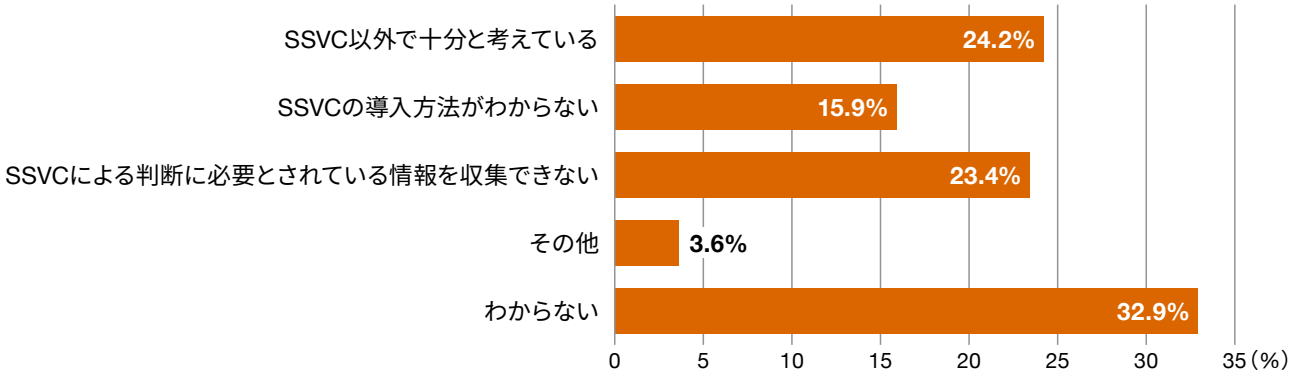


SSVCの導入を検討・採用していない理由として「SSVC以外で十分と考えている」が24.2%と最も多いものの、「SSVCによる判断に必要とされている情報を収集できない」（23.4%）、「SSVCの導入方法がわからない」（15.9%）と導入が困難であるとする回答が一定数を占めています（図表25）。

SSVCを検討している際に生じている課題として「必要な情報（System Exposure：システムの外部露出）が不足」（15.5%）が内部情報として最も多い回答です。また、「必要な情報（Exploitation：脆弱性の悪用状況）が不足」（14.5%）が外部情報として最も多い回答となっています。これらがSSVCの導入ネックとなっているように見えます（図表26）。

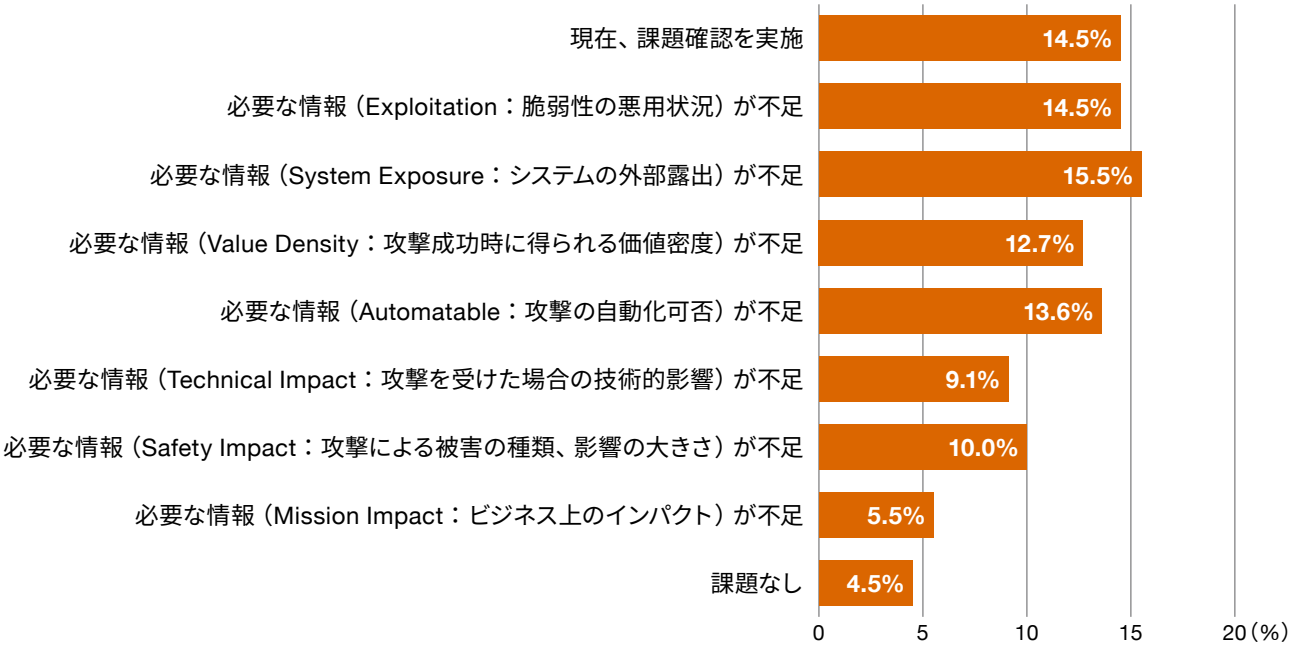
**図表25：SSVCを検討・採用しない理由**

**Q.** SSVCを検討・採用しない理由（いくつでも、n=252）




**図表26：SSVCを検討している中で生じている課題**

**Q.** SSVCを検討している中で生じている課題（いくつでも、n=110）



## 意思決定機能の成長を実現するためのAcuity Rampというコンセプト

前述のとおりSSVCの提案元である米カーネギーメロン大学ソフトウェア工学研究所からAcuity Rampというコンセプトが公開されています。これはSSVCを導入した後、組織のニーズやリソース、能力が変化する中で、意思決定機能をよりステークホルダーの望む形へ成長させるための考え方です。

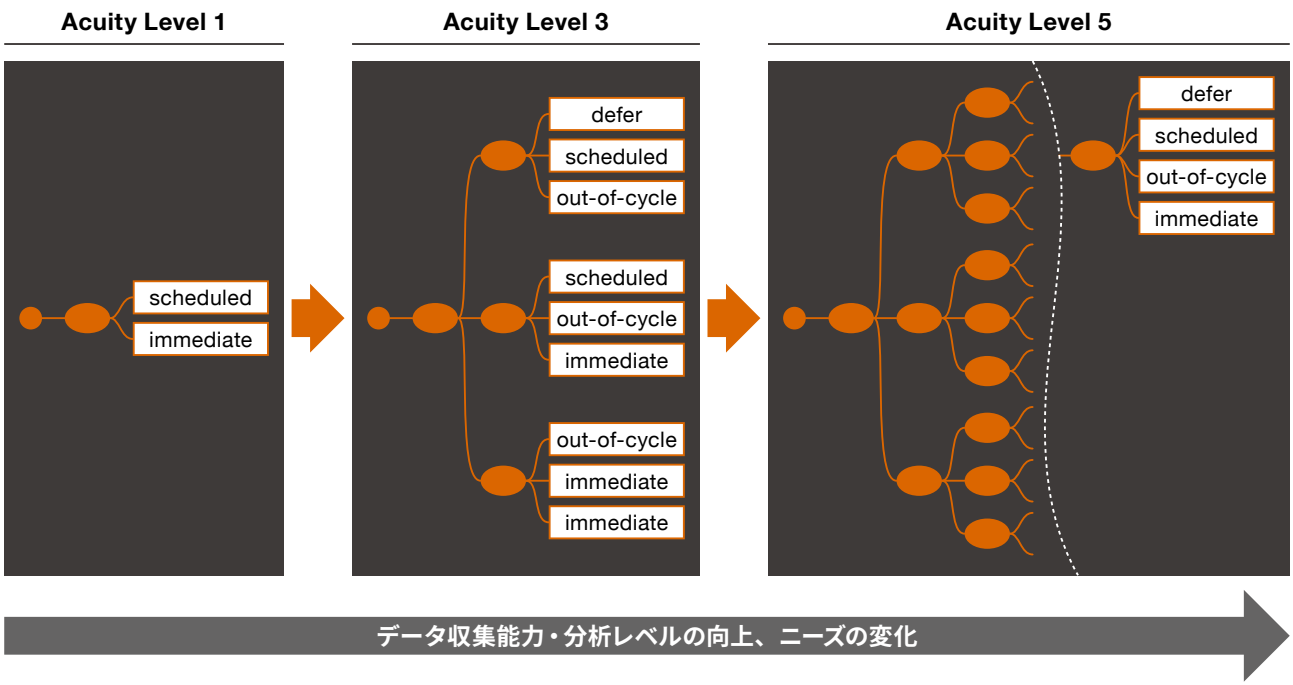
Acuity Rampの考え方で組織のデータ収集・分析能力レベルごとに想定される決定木の例を示します（図表27）。

Acuity Rampでは、このレベルのことをAcuity Levelと呼んでいます。例として記載しているレベル数とその内容は、SSVCとして定義されているものではなく、個々の組織で決める必要があります。

この例では組織のデータ収集、分析能力が低い場合にはシンプルな決定木を用います。その後、組織のデータ収集、分析能力が向上するにつれて、より詳細な決定木を用いることができます。

**図表27：ニーズ・リソース・能力変化によるSSVC決定木の成長例**

| Acuity Level | 決定木の構成要素 |
|---|---|
| 1 | KEV（Known Exploited Vulnerabilities Catalog） |
| 2 | Exploitation：脆弱性の悪用状況 |
| 3 | Exploitation：脆弱性の悪用状況、System Exposure：システムの外部露出 |
| 4 | Exploitation：脆弱性の悪用状況、System Exposure：システムの外部露出、Automatable：攻撃の自動化可否 |
| 5 | Exploitation：脆弱性の悪用状況、System Exposure：システムの外部露出、Automatable：攻撃の自動化可否、Mission Impact：ビジネス上のインパクト、Safety Impact：攻撃による被害の種類、影響の大きさ |



Acuity Level 1の状況で、組織がデータを収集・分析できるようになれば脆弱性に関して独自でデータを収集でき、KEV（Known Exploited Vulnerabilities Catalog）による判定だけでなく、Acuity Level 2のように決定木へExploitationを組み込むことで、より詳細な判断モデルに移行できます（図表28）。この場合、有償の脆弱性情報提供サービスなどを意思決定モデルに組み込むことが考えられるでしょう。

その後、ASM（Attack Surface Management）の実施などにより社内の資産管理能力が向上することで、インターネット上に公開されているシステムを特定でき、Acuity Level 3のように決定木へSystem Exposureを組み込むことができるようになります（図表29）。

さらに、時間の経過とともに組織の分析能力が向上することで、Acuity Leve4のようにAutomatableの判断ポイントを意思決定モデルに組み込むことが可能です（図表30）。

そして、プロセスの運用を進めていく中で、脆弱性が業務継続性や安全性に与える影響がより理解できるようになり、Mission Impact、Safety Impactの判断ポイントを意思決定モデルに取り入れられるようになります（図表31、32）。

**図表28：Exploitation：脆弱性の悪用状況の定義**

| 値 | 定義 |
|---|---|
| None | 実際に悪用されている証拠がない、また攻撃検証コードが公開されていない。 |
| PoC | ダークウェブ、ディープウェブなどのアンダーグラウンドで攻撃検証コードが売買されている、攻撃検証コードがインターネット上の脆弱性データベースやソースコードリポジトリで公開されている、または攻撃の再現手法が公知となっている。 |
| Active | 公的機関、セキュリティベンダー、ソーシャルメディアなどで脆弱性を悪用した攻撃の発生が確認、報告されている。 |

**図表29：System Exposure：システムの外部露出の定義**

| 値 | 定義 |
|---|---|
| Small | ネットワークに接続されていないサービスやプログラム、またはファイアウォールなどで厳格にアクセスが制御されたネットワーク。 |
| Controlled | 何らかのアクセス制限や緩和措置が施されているネットワーク。 |
| Open | アクセス制限を十分に実施できないインターネット、または広くアクセス可能なネットワーク（DNSサーバー、ウェブサーバー、VOIPサーバー、メールサーバーなどが該当する）。 |

**図表30：Automatable：攻撃の自動化可否の定義**

| 値 | 定義 |
|---|---|
| No | 攻撃者は、以下の理由により、キルチェーンのステップ1から4（1：reconnaissance, 2：weaponization, 3：delivery, 4：exploitation）を確実に自動化できない。<br>• ネットワーク経由での脆弱性の検出・列挙が不可能<br>• 武器化にあたり、人間による個別の関与が必要<br>• 配信にあたり、広く普及しているネットワークセキュリティ対策の回避が必要<br>• ASLR（Address Space Layout Randomization）などの既定で適用される脆弱性攻撃緩和策により、エクスプロイトの信頼性が低い |
| Yes | 攻撃者はキルチェーンのステップ1から4を確実に自動化できる。 |

**図表31：Mission Impact：ビジネス上のインパクトの定義**

| 値 | 定義 |
|---|---|
| Degraded | 重要ではない機能の劣化、またはほとんど影響がない状態。慢性的な劣化はミッション達成に必要な機能に影響がでる可能性がある。 |
| MEF※ Support Crippled | ミッション達成を直接的に支える活動が機能不全となるが、ミッション達成に必要な機能は一時的に継続する。 |
| MEF※ Failure | ミッション達成に不可欠な機能のいずれかにおいて、許容範囲を超えて長期間にわたり障害が発生する。 |
| Mission Failure | ミッション達成に不可欠な機能の複数または全てに障害が発生し、回復能力が低下する。回復能力の低下により組織のミッションが果たせなくなる。 |

※MEF：Mission Essential Function

**図表32：Safety Impact：攻撃による被害の種類、影響の大きさの定義**

| 値 | 定義 |
|---|---|
| None | 全ての観点でMinorを下回る影響 |
| Minor | 以下のいずれかに該当する場合<br>Physical Harm：システム利用者の不快感、軽微な労働安全衛生上の危険、物理的なシステムの安全マージンの減少<br>Environment：軽微な物的損害や環境的な損害<br>Financial：複数の人に対する金銭的な損失<br>Psychological：複数の人に対する精神的、または心理的な損害 |
| Major | 以下のいずれかに該当する場合<br>Physical Harm：システム利用者の身体的苦痛、負傷、重大な労働安全上の危険、物理的なシステムの障害<br>Environment：大きな物的損害や環境的損害<br>Financial：複数の人に対する破産につながるような大きな金銭的な損失<br>Psychological：集団に対する精神的、または心理的な損害 |
| Hazardous | 以下のいずれかに該当する場合<br>Physical Harm：重傷、または致命傷につながる危険、安全な運用を支援するサイバーフィジカルシステムの一部故障<br>Environment：生命への脅威、広範囲の環境破壊、公衆衛生リスクの可能性など<br>Financial：選挙や金融などの社会を支えるシステムが不安定化や安全ではない状態に陥る |
| Catastrophic | 以下のいずれかに該当する場合<br>Physical Harm：多数の死者の発生<br>Environment：公衆衛生への脅威、小規模な生態系の破壊につながる環境破壊など<br>Financial：選挙や金融などの社会を支えるシステムの崩壊 |

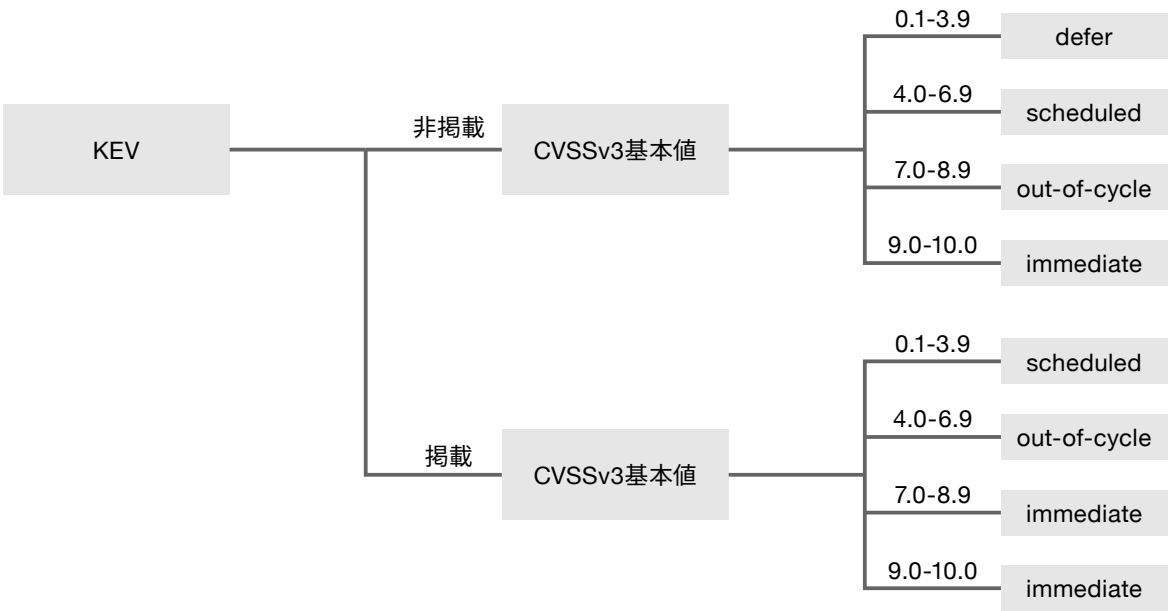
Acuity Rampでは、情報収集のコストと意思決定の質のトレードオフを行う方法を示しています。ある時点においてコストが高く取得できない情報がある場合でも、意思決定を行うためのモデルを作ることができる一方で、意思決定の質が低下するリスクがあります。意思決定の質が低い場合、緊急対応が不要な脆弱性に対しても緊急対応をしてしまうケースや、逆に緊急対応が必要にもかかわらず見逃してしまうケースが考えられます。例えば、コストやリソースの制約からAcuity Level 1を用いて意思決定を行った結果、KEVに載っていない脆弱性への対応を保留するという判断がなされた場合、まだKEVには載っていない危険性の高い脆弱性への対応が遅れることになります。Acuity Rampの考え方を活用する場合、こうしたトレードオフで犠牲になる部分を受け入れ可能か判断する必要があります。既にCVSSを運用している場合、それにSSVCを組み合わせることで、脆弱性対応プロセスの劣化を避けつつ、SSVCを導入可能です。また、Acuity Rampの考え方では、構築した意思決定モデルの判断ポイントを他へ置き換えることや、時間の経過とともに新しい判断ポイントを追加・カスタマイズすることで、組織の意思決定機能をシンプルなモデルから、より効率的で安全なモデルへ成長させることもできます。

## Acuity Rampの活用

今回の調査では、「SSVCの導入方法がわからない」「SSVCで必要な情報が不足している」という回答が多く見受けられました。しかし、Acuity Rampの考え方を用いることで、SSVCをスモールスタートし、徐々に脆弱性対応プロセスを成長させることができます。

脆弱性対応プロセスにCVSSを採用している組織がスモールスタートする場合、CVSSv3基本値にKEVを加える形で決定木を作成することが望ましいと考えます（図表33）。この方式の場合、KEVを加えることでCVSSv3基本値のみの判定よりも悪用状況を考慮した判定が可能になります。SSVC導入後は、前述したようにSystem Exposure要素の追加などによる段階的なプロセスの成長を検討していくことが可能です。

**図表33：CVSSとKEVを用いたSSVCの決定木例**




## まとめ

毎年多くの脆弱性が発見されている中、それに対応するための脆弱性対応プロセスの整備および改善は組織において重要な課題です。Acuity Rampの考え方を参考にすることで、脆弱性対応プロセスとしてSSVCを採用していない組織はSSVCをスモールスタートできます。また、既にSSVCを採用している組織においても、プロセスを成長させる上での指針として利用できると考えられます。

# 2-5 エコシステムを支えるデジタルアイデンティティ・トラストフレームワーク

## エコシステムにおける信頼の鍵

デジタル社会が進展する現代において、個人や企業は多くの情報をデジタル領域で共有、取引することが求められています。このような環境において、デジタルアイデンティティは、信頼性と安全性を確保するための基盤としてますます重要な役割を果たしています。しかし、デジタルアイデンティティの運用における課題は多岐にわたり、特に異なるシステムやサービス間での相互運用性の確保が求められています。

そのためには、デジタルアイデンティティ・トラストフレームワークと呼ばれる多様なシステムやサービス間での信頼性を確保し、安全な情報交換を実現するための包括的な枠組みの構築が不可欠です。本節では、このトラストフレームワークの意義、重要性、そしてその構成要素について詳しく解説し、未来のデジタルエコシステムにおける信頼性の向上に向けた道筋を探ります。

## デジタルアイデンティティ・トラストフレームワーク整備の重要性

デジタルアイデンティティ・トラストフレームワークは、人やデバイスのデータを扱うデジタルエコシステムにおいて、信頼できる取引を可能にするために策定されるルールやガイドラインをまとめた包括的な仕組みです。

このトラストフレームワークは、単なる概念にとどまらず、国・地域やISOなどの国際的な標準化団体、専門機関による具体的なガイドラインやプロトコルが含まれます。

その主な機能として、まず取引する相手が信頼できることを担保することが挙げられます。その実現に向けて、デジタルアイデンティティを扱うシステムやサービス同士が相互に信頼できる状態をつくるためのルールを定めなければなりません。さらに、主要な機能の一つとして、異なるステークホルダーをつなげる仕組みを提供する相互運用性の促進があります。これにより、ユーザー、サービスプロバイダー、デバイス、組織間で安全かつ円滑に相互作用が行える環境を構築します。

## なぜデジタルアイデンティティ・トラストフレームワークが重要なのか

### エコシステム形成とサプライチェーンにおける重要性

昨今、多くのビジネスやサービスは単独ではなく、多くのプレイヤーが連携して提供されることが一般的となってきました。そのような中、人やデバイスのデータは、企業、政府機関、個人を含む多くのエコシステムで活用されます。これらのデータの安全性を担保するため、サプライチェーン全体で信頼できるプロセスに則ってデータを扱う適切なID管理を行うことが重要です。

このフレームワークを整備するメリットを以下に記載します。

● ビジネス価値の創造

トラストフレームワークは、新規ビジネスやサービス開発における基盤として活用できます。

標準化された認証基盤を活用することで、新サービスの開発期間を短縮し、セキュリティ品質を確保しやすくなります。また、ユーザー認証の信頼性向上により、オンラインサービスにおける顧客体験が改善され、競争力の強化につながります。特に、異業種またはグローバル展開を検討する企業にとって、またはグローバル展開を検討する企業にとって、相互運用性の確保は重要な価値となります。例えばグローバル展開を行う場合、標準化された技術を使用しておけば、海外市場への参入障壁を低減できます。

● 企業のリスク管理における価値

デジタルアイデンティティのトラストフレームワークは、企業のリスク管理において実践的な解決策を提供します。サプライチェーンセキュリティにおいては、取引先との信頼関係をシステム的な形で構築できます。例えば、取引の相手方の真正性や実在性の確保のための技術やプロセスが標準化され、信頼性の評価基準が明確になることで、なりすましや取引詐欺のリスクを抑えられます。これにより、政府機関や企業が管理するID管理システムを連携させれば、アクセス権限の管理や監査の負担が軽減され、セキュリティリスクの可視化が容易なセキュアなデジタルアイデンティティインフラを形成することが可能となります。

● コスト最適化の実現

トラストフレームワークの導入は、長期的な視点でコスト削減に貢献します。個別に構築・維持していた認証システムを標準化された枠組みに統合することで、システム開発・運用コストを削減できます。また、セキュリティ運用の効率が向上し、人的リソースの最適配分が可能となります。さらに、インシデント対応コストの観点でも大きな効果が期待できます。統一された対応手順と役割および責任の明確化により、インシデント発生時の対応時間短縮と影響範囲を最小限に抑えられるでしょう。

● コンプライアンスの強化

個人情報保護法やGDPRなど国内外の個人データ保護の対応において、プライバシー侵害や目的外利用を防止し、透明性と信頼性を確保するため、法規範に伴う国際基準の準拠がますます重要となっています。トラストフレームワークでは、これらに関する明確な基準の提供も必要とされます。統一された基準に基づく管理により、法令遵守の確認が容易になり、監査対応の工数を削減できます。また、グローバルに連携する場合、GDPRなどのグローバル規制にも整合的な枠組みとなっているため、国際的なコンプライアンス要件への対応も効率化されます。

一方で、トラストフレームワークを整備しないままにエコシステムを作る場合、以下のような問題が発生する可能性があります。

● ユーザーエクスペリエンスの一貫性欠如

場当たりな対応の氾濫は、ユーザーの意思決定プロセスを妨げる要因となります。単一のトラストフレームワークの下で統一されたユーザーエクスペリエンスを提供することで、ユーザーの理解と信頼を促進できます。

● セキュリティリスク

トラストフレームワークを整備せずガバナンスやリスク評価を場当たりに実施している場合、アイデンティティシステムの運用は各組織の裁量に委ねられ、一貫性のない管理体制となります。これにより、ユーザーデータの安全性やシステム全体の信頼性が低下し、セキュリティリスクが高まる可能性があります。特に、プライバシー、セキュリティに関する概念的な課題は、統一された基準なしには適切に対処できません。

これらのことを考えると、トラストフレームワークは企業にとって、ビジネス機会創造、リスク管理、コスト最適化、コンプライアンス強化など、多面的な価値を提供する実践的な解決方法となり得るでしょう。

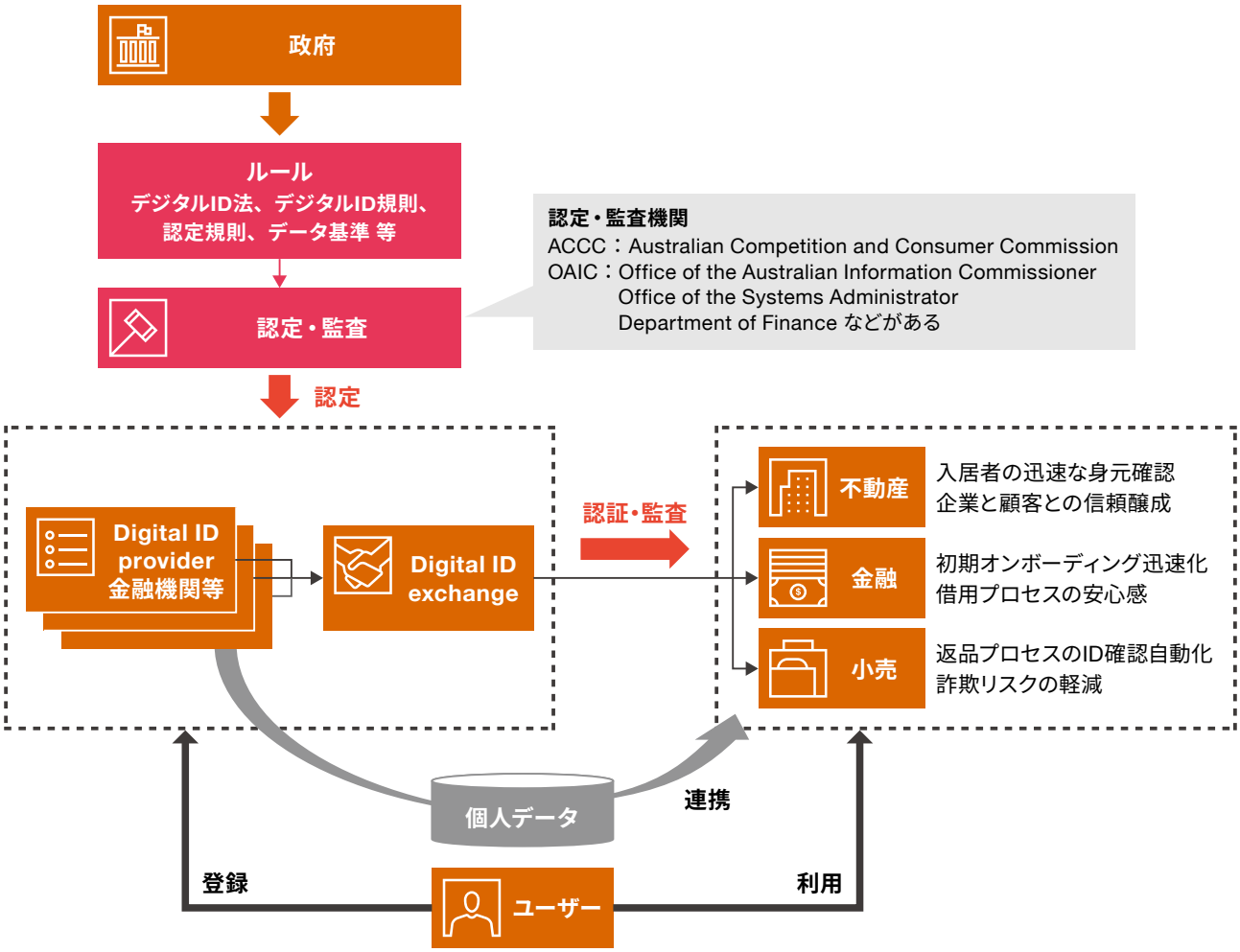
## 事例：オーストラリアのAGDIS

デジタルアイデンティティ・トラストフレームワークの考え方は、英国、EU、シンガポール、インドなど多くの国・地域で導入が進められていますが、ここではオーストラリアの事例を紹介します。

オーストラリアは、デジタルアイデンティティ管理の分野で先進的な取り組みを進めています。その中で使用されているトラストフレームワークが「AGDIS（Australian Government Digital ID System）」です（図表34）。

AGDISは、安全なデジタルアイデンティティの利用を促進し、オンライン業務における信頼性を向上させることを目的としています。

**図表34：Australian Government Digital ID Systemの認定を受けた民間セクターのインフラ概要**




Digital ID provider：デジタルIDの管理を行う
Digital ID exchange：サービスプロバイダとDigital ID provider間の情報を連携する

## AGDISの目的と法制度

AGDISが目指すのは、安全なデジタルアイデンティティの利用を促進し、オンライン業務における信頼性を向上させることです。このフレームワークに基づき、個人は複数のデジタルサービスで認証を行えます。一方、各サービス提供者は厳格な認定基準に基づき、データセキュリティやプライバシー保護の規範を遵守することが求められます。体制としてはオーストラリア競争消費者委員会（ACCC）がDigital ID Regulatorとしてプロバイダーの認定と監査を行い、オーストラリア情報コミッショナーオフィス（OAIC）がプライバシー保護の役割を担う他、データ標準なども定められ運営されています。この枠組みにより、法的および運用上の透明性が確保されています。

### AGDISの活用事例

AGDISは、公共セクターと民間セクターの両方が認定を受けることができ、相互運用可能なデジタルアイデンティティエコシステムの構築を可能にしています。オーストラリアでは、AGDISの認定を受けた民間セクターのインフラが稼働しており、このインフラが提供するサービスを民間企業が活用することで、セキュアなエコシステムの実現に寄与しています。

例えば、個人情報を直接共有せず、安全かつ効率的にアイデンティティを証明するデジタル認証システムがあります。信頼性の高い本人確認を行っている金融機関などの既存の事業者（Digital ID Provider）から得た情報を活用して個人を認証することで、各業界のサービスプロバイダーに次のような具体的な仕組みとメリットをもたらしています。

● 不動産業界

住居の賃貸契約では、書類確認や個人情報の多重的な入力が必要になる場合が多く、不必要なデータ共有や管理上の負担が企業にのしかかることが一般的です。しかし、トラストフレームワークの仕組みの整備により、顧客は銀行やその他信頼できる情報元からアイデンティティ情報で即座に認証できます。これにより、従来のような手作業による運転免許証やパスポート情報の直接共有が不要になります。

実現効果は、以下のとおりです。

- データ共有の範囲を最小限に抑え、情報漏洩リスクを減少
- 検証プロセスをリアルタイムに実行し、書類準備に時間を取られることなく迅速な契約が可能
- テナントとオーナーの双方にとって、信頼性の高い取引環境を提供

結果として、日本で見られる紙ベースや手入力中心の賃貸契約の手続きを変革し、簡便で安全な管理を実現しています。この仕組みは、不動産業務の効率化と信頼性向上に役立てられています。

● 金融業界

ローン申請手続きでは、顧客が膨大な個人データを入力する時間的負担が課されるとともに、データの正確性を確保し続ける必要があります。こうした課題に対応するために、AGDISを活用した仕組みの整備により、顧客は信頼するデータ元（銀行など）の情報を活用して申請フォームを自動で部分的に入力できます。これにより、入力ミスを防ぎ、プロセス全体の効率化が促進されます。

実現効果は、以下のとおりです。

- 顧客データの自動入力により、入力エラーや名前の不一致を回避
- オンボーディングの迅速化で、顧客体験がスムーズに
- 金融機関にとっては、データ収集プロセスの簡素化と法規制遵守の一貫性確保に寄与

この仕組みは、銀行やローン提供業者が広範囲な手動プロセスを見直し、顧客満足度を高めつつ効率的なサービス提供を実現する方法として活用されています。

● 小売業界

オンラインショッピングにおける返品プロセスでは、顧客からの情報収集や本人確認が必要になりますが、非効率なプロセスが不正リスクを高め、顧客満足度の低下を招く懸念があります。AGDISを活用した認証により、顧客は銀行などの信頼できる情報元を利用して返品手続きを迅速化できます。これにより、店舗側が直接的に顧客情報を管理・収集する必要はありません。

実現効果は、以下のとおりです。

- 詐欺行為を防止し、プラットフォームおよび顧客双方の信頼を向上
- 便利な顧客体験を提供し、リピーターを確保
- 顧客情報を保護しつつ、迅速に返品を完了するプロセスを実現

このプロセスは、EC市場や店舗においても、消費者の課題を解決するための効果的かつ安心できる返品対応モデルを提供しています。

このように、AGDISの仕組みはさまざまな業界で活用されており、企業にも一般の市民にも以下の共通メリットがあります。一般の市民は、既に信頼する情報元を介してスムーズに本人確認を行えるため、手続きを簡潔化できます。企業にとっても、不必要なデータ収集や手作業の削減によりコストが削減され、同時に顧客との信頼関係を強化します。

**今後の発展**

オーストラリアでは2024年12月1日から「Digital ID Act2024」が施行されました。

「Digital ID Act 2024」は、2026年までに地方政府やより多くの民間組織を巻き込む形で拡大していく計画です。この発展により、デジタルIDエコシステムがさらに充実し、セキュリティ、利便性、そしてプライバシーの全てを強化する包括的な基盤が整います。

AGDISは、今後もオーストラリアのデジタル経済を支える重要なインフラとして進化し続けると予測されます。

## デジタルアイデンティティ・トラストフレームワークの構成要素

ここからは、デジタルアイデンティティ・トラストフレームワークについて解説していきます。

デジタルアイデンティティ・トラストフレームワークの主な構成要素として、次の要素があります。

● エコシステム全体が守るべき原則

この要素は、デジタルアイデンティティの運用において、倫理的および法的な基盤を提供します。具体的には、ユーザーを中心に考えることや、透明性を保証することが含まれます。これにより、全ての関係者が共通の価値観に基づいて行動することが求められます。

● ステークホルダーの義務と権利

フレームワークに参加する各ステークホルダーは、それぞれの役割に応じた義務と権利を持ちます。これを明文化することで、各ステークホルダーが果たすべき義務と享受する権利が明確になり、協力関係が円滑に進むようになります。

● 技術的要素とセキュリティ基準

デジタルアイデンティティの安全性を確保するために、必要な技術的要素とセキュリティ基準が設定されます。これには、セキュリティ基準の遵守状況を確認するためのガバナンスプロセスが含まれ、システム全体の信頼性と技術的な相互運用性を高めます。

● トラストマーク

トラストマークは、信頼性を可視化するためのシンボルです。これにより、ユーザーはサービスや事業者の信頼性を容易に確認でき、安心して利用できます。

● 運用上の役割とガバナンス

フレームワークを効果的に運用するためには、開発と維持、必要に応じた改訂を担当する組織が必要です。この組織は、フレームワークの運用に関するコンセンサスを形成し、継続的な改善を図ります。

これらの構成要素は、相互に関連し合いながら、一貫した信頼の枠組みを形成します。これにより、デジタルアイデンティティの安全で信頼性の高い運用が可能となります。

このトラストフレームワークは、これまでもさまざまな国・地域で検討されてきました。

ただし、そのトラストフレームワーク自体の種類が多くなるという弊害もあり、現在、規模や特性に応じたトラストフレームワークが検討されています。

この活動はSIDI Hubといい、次頁で紹介します。

## SIDI Hubの紹介

### SIDI Hubの目的と役割（国際連携の推進）

SIDI Hubは、デジタルアイデンティティの国境を越えた相互運用性を実現するというミッションを持って2023年に設立されました。その特徴は、多様な規制環境や文化的背景を持つ世界において、相互運用性の複雑な課題に取り組むためのステークホルダーの結集にあります。SIDI Hubは、人間中心設計の原則を適用し、国連の人権に関する基本文書に基づいて活動を展開しています。国・地域の政策や設計上の選択が価値観や文化によって形作られることを尊重しながら、多様なソリューションの相互運用を目指しています。

### SIDI Hubの活動概要

SIDI Hubは、25以上の非営利団体や多国籍組織によって組織・支援されており、25カ国以上のデジタルアイデンティティ専門家が参加しています。活動は図表35の5つのワークストリームを通じて展開されています。

これらのワークストリームは、2024年に開催された5つのサミット（ケープタウン、ベルリン、ワシントンD.C.、東京、ブラジル）を通じて進められ、2025年の戦略策定に向けた知見を蓄積しています。特に東京でのサミットでは、災害対応や教育資格、オープンバンキングなどのユースケースが議論され、デジタルIDの相互運用性の重要性が強調されました。

国・地域や業界といった巨大なエコシステム向けの検討のみでは、個別の企業には縁遠いものになりかねません。しかし、SIDI Hubのように具体的なユースケースや最小限の要件を想定したトラストフレームワークは、どの規模の企業や団体にとっても有用であると考えられます。

SIDI Hubで行われている議論を踏まえても、相互運用性は業界標準への準拠やエコシステムへの参画、技術革新への追随において重要な役割を果たします。また、それが欠如した場合は、金融サービスやエコシステム参加の機会損失、災害時の支援制限などのリスクをもたらします。さらに、相互運用性の導入はシステムの個別開発・運用の削減、不正リスクの低減、メンテナンスコストの削減につながるでしょう。

このように、相互運用性は単なるコストではなく、将来の競争力維持のための重要な投資として捉えるべきです。中長期的な視点での早期の取り組みが競争優位性の確保につながると考えられます。国・地域や企業は、自らの持つサービスを、将来的な外部との連携や緊急時の対応を見据え、可能な限り設計段階から相互運用性を備えた仕組みを構築することが重要です。

これは、ビジネス環境の変化や予期せぬ事態に迅速かつ柔軟に対応できる体制を整えるために不可欠です。

**図表35：SIDI Hubのワークストリーム**

| ワークストリーム名称 | 概要 | 主な活動内容 |
|---|---|---|
| Champion Use Cases<br>（チャンピオンユースケース） | 具体的な活用事例の特定と分析 | 実践的な活用事例の特定と分析を行う。難民支援や銀行口座開設などの具体的なケースを通じて、デジタルアイデンティティの要件を検討する |
| Trust Framework Mapping<br>（トラストフレームワークのマッピング） | 異なるフレームワーク間の比較分析 | 異なるトラストフレームワーク間の比較分析と、フレームワーク間の整合や国際標準との調和を図る |
| Minimum Requirements for Interoperability<br>（相互運用性の最小要件） | 相互運用に必要な基本要件の定義 | 相互運用性を確保するための最小限必要な技術要件を定義する |
| Metrics of Success<br>（成功指標） | 相互運用性達成度の評価基準設定 | 相互運用性の達成度を評価するための指標を設定する。実装の成功度を測定する基準を策定し、継続的な改善のための評価方法を確立する |
| Governance<br>（ガバナンス） | 全体の運営管理体制の検討 | フレームワーク全体の運営管理体制を検討する。SIDI Hub自体はガバナンス権限を持たず、政府および分野横断的な参加者によるガバナンスを確保と意思決定プロセスの在り方を検討する |

## 相互運用性における日本の課題

これまでに述べたように、企業や国境を越えたデジタルアイデンティティの相互運用性は、グローバル化が進む現代社会において重要性を増しています。しかし、相互運用性の実現にはさまざまな背景を持つ参加者の同意が必要であり、そのためには、ステークホルダー全体の価値観や文化を尊重しながら、多様なソリューションの相互運用を実現することが求められます。

日本では、個人情報保護法や電子署名および認証業務に関する法律など、個人情報の利活用や事業者間接続のための制度は個別に整備されています。しかし、総合的なものはなく、エコシステムを構築するために事業者はこれらを組み合わせて使用する必要があります。

また、各ステークホルダーの役割と責任が明確に定義されておらず、相互運用性を確保するための技術要件やそれを支えるガバナンス体制の確立も十分とは言えない状況です。

ただし今後、この状況は見直されていく可能性があります。

2024年10月に行われたG7 デジタル・技術大臣会合では、DPI（Digital Public Infrastructure／デジタル公共インフラ）についての議論が初めて実施されました。

日本政府は、以下の点を発信しました[※13]。

- DPIの議論を進めていくにあたっては、公的分野におけるAIの活用やデジタルアイデンティティなどの個々の要素を連結させ、総体としてのデジタル政府を実現することが必要であること
- その要素のすべてに共通するデータの共有とデータガバナンスが要になること
- 従ってその相互運用性に係る議論を進める際には、信頼性ある自由なデータ流通（Data Free Flow with Trust）推進を通じて、データ共有に関するさまざまな懸念や障壁に適切に対処していくべきであること

またこの会合の後、「デジタル・アイデンティティ・マッピングエクササイズ（Mapping Exercise of Digital Identity Approaches）」と呼ばれる、デジタルアイデンティティについてG7各国における共通概念や定義、国際的な技術標準の活用、保証レベルへのアプローチなどの現在のギャップと、将来の相互運用性を支援するための共通アプローチを探る文書が公表されました。

これらの活動からも今後、このトラストフレームワークが整備されていくことが期待されます。

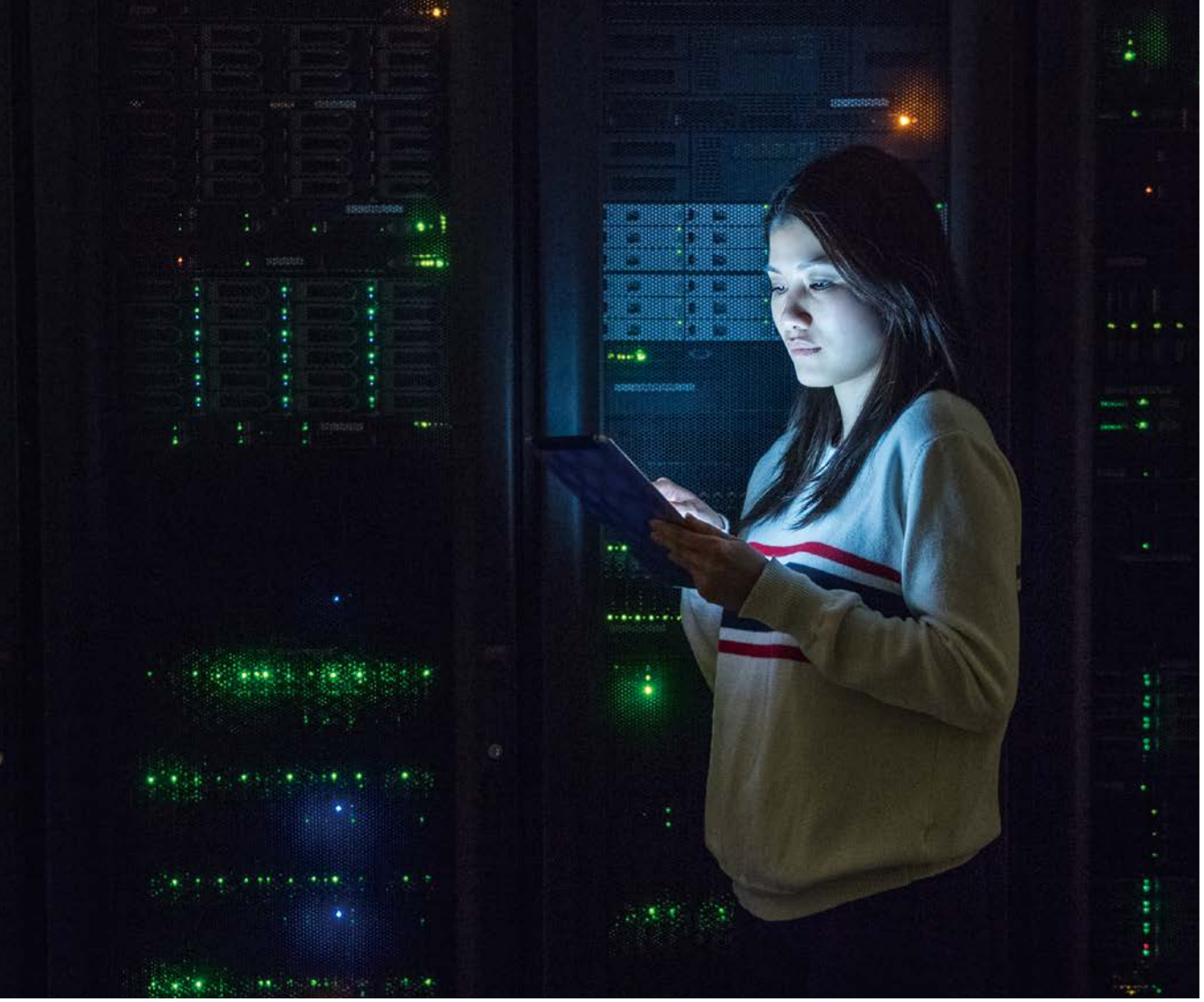
## まとめ

デジタルアイデンティティ・トラストフレームワークの整備は、相互運用性の実現において重要な役割を果たします。これにより、サプライチェーンやエコシステム形成など、さまざまな企業、国境を越えたデジタルアイデンティティの円滑な運用が実現し、ビジネス展開が適切かつ迅速に行え、競合劣後を回避できます。

整備を適切に進めるためには、専門家が参加するSIDI Hubのような国際標準化活動の知見を活用し、ステークホルダー間で相互運用性を確保しつつ、自社固有の要件にも対応できる柔軟な枠組みを構築することが重要です。

そして、継続的な維持管理、新たな課題への対応、定期的な見直しと更新のプロセスまで整備することで、競争力の強化と、安全で信頼性の高いデジタルサービスの提供が可能となるでしょう。

※13　出所：デジタル庁「デジタル・技術大臣会合の開催結果」 https://www.digital.go.jp/news/53ed2e40-a8be-4249-869d-e94f4f9a28fa



# おわりに

サイバー攻撃がなくなることはありません。サイバー攻撃の防御は終わりのない対応といえるでしょう。そして、その対応は単に過去の対策の繰り返しでは終わりません。なぜならば、AIに代表されるように技術は常に変化し、国際情勢に応じて攻撃者のターゲットや攻撃の手法、強度が変化し、技術や国際情勢に応じて法律などが変わるためです。そして、何よりも、企業のビジネスも変化します。

このように変化が激しい事業環境の中で、迅速かつ効果的に防御をするためには、情報を収集し、分析することが重要となります。

そのためには自社のガバナンスを強化し、自社の状況をよく理解できるようにするとともに、業界、取引先、政府などともコミュニケーションをとり、適切な対応ができる環境を維持することが重要です。

機先を制し、対策をすることによりサイバー攻撃の影響を最小化し、事業のレジリエンスを保つことが、企業に求められます。

# お問い合わせ先

**PwC Japanグループ**
https://www.pwc.com/jp/ja/contact.html

**www.pwc.com/jp**

PwC Japanグループは、日本におけるPwCグローバルネットワークのメンバーファームおよびそれらの関連会社（PwC Japan有限責任監査法人、PwCコンサルティング合同会社、PwCアドバイザリー合同会社、PwC税理士法人、PwC弁護士法人を含む）の総称です。各法人は独立した別法人として事業を行っています。
複雑化・多様化する企業の経営課題に対し、PwC Japanグループでは、監査およびブローダーアシュアランスサービス、コンサルティング、ディールアドバイザリー、税務、そして法務における卓越した専門性を結集し、それらを有機的に協働させる体制を整えています。また、公認会計士、税理士、弁護士、その他専門スタッフ約12,700人を擁するプロフェッショナル・サービス・ネットワークとして、クライアントニーズにより的確に対応したサービスの提供に努めています。
PwCは、社会における信頼を構築し、重要な課題を解決することをPurpose（存在意義）としています。私たちは、世界149カ国に及ぶグローバルネットワークに370,000人以上のスタッフを擁し、高品質な監査、税務、アドバイザリーサービスを提供しています。詳細はwww.pwc.comをご覧ください。

発刊年月：2025年3月　　管理番号：I202408-11