2026 年 11 月 10 日　全 6 頁

AI によるサイバー攻撃の嵐

深刻な現状のおさらいと乗り越える方法の考察

フロンティア研究開発センター　シニア・フェロー
プリンシパル・データサイエンティスト　坂本 博勝

はじめに

アンソロピック社の Claude Mythos Preview という AI モデルが出現し、サイバーセキュリティに関する大きな騒動が起きたのは 2026 年 4 月のことだった。

そして 2026 年 11 月現在、サイバーセキュリティ深刻化の潮流はさらに大きくなっている。毎日のように多くの企業が、サイバー攻撃を受け情報漏洩被害にあったことを公表している。

4 月に突如として現れた、AI の発達が深刻なサイバー被害を企業にもたらしかねないという懸念は、数ヵ月を経て猛スピードで現実リスクとなり始め、社会不安を急速に作り出している。一連の社会現象は関係しており、地続きだ。

今回は今話題の、企業に対する AI を使ったサイバー攻撃をテーマに、被害や対処方法の「現状」と「未来」について、論考していきたい。

日本における企業向けサイバー攻撃の現状

サイバー攻撃を受け情報漏洩被害にあった日本企業に関する被害企業の一覧整理については、他の品質の高い情報リソースに役目を譲ることとし、本稿で詳述はしない。[1]

マクニカの調査によれば、国内 Web 侵害の公表件数は 2026 年 4 月から増加傾向にあり、8 月から急増していると評価できる模様だ。[2]

日本企業の被害状況をまとめると、以下のような特徴がある。

- 9 月以降、10 万件以上の個人情報が漏洩した可能性のある大規模事案の公表が 150 件以上あり、前年と比べても非常に発生頻度が高い
- 情報漏洩事案の前年比急増率が突出しており、システムの停止や企業脅迫といった被害の

[1] 情報漏洩・流出・不正アクセスの企業一覧（随時更新）
https://smartscope.blog/blog/japan-data-breach-list-tier-2026-09-10/
[2] 相次ぐ WEB システムからの情報漏洩事案について
https://security.macnica.co.jp/blog/2026/10/web-incidents2026.html

このレポートは投資勧誘を意図して提供するものではありません。このレポートの掲載情報は信頼できると考えられる情報源から作成しておりますが、その正確性、完全性を保証するものではありません。また、記載された意見や予測等は作成時点のものであり今後予告なく変更されることがあります。㈱大和総研と大和証券㈱は、㈱大和証券グループ本社を親会社とする大和証券グループの会社です。内容に関する一切の権利は㈱大和総研にあります。無断での複製・転載・転送等はご遠慮ください。

報告件数における増加率を上回っている[3]

- サイバー被害を受けた企業/団体は、業種や民間/公共団体の区別を問わず、無差別的である

10 月中旬になると、実際のサイバー侵害ログを解析したうえでの調査レポートが公開され始め、サイバー攻撃に AI を活用したと思われる痕跡も確認されてきている。[4] [5]

中国関連のサイバー犯罪グループの関与を確認したという発表もある。[6]

強く断定するにはまだ情報が少ないものの、日本企業における情報漏洩事案の頻発について、AI による自律的な判断と行動を取り込んでスピード・効果・攻撃件数を向上させた、レベルアップされたサイバー攻撃が含まれている蓋然性が、高いと言えるだろう。誰もが漠然と予感していた印象が、評価として現実化してきている。

アンソロピックは 9 月に、中国製のオープンウェイトの AI モデルを名指しして、サイバーセキュリティ上の脅威であると評価するブログを公開している。[7]

※モデルの開発国・開発企業と、個別のサイバー犯罪活動の主体を直接結び付けるものではない。

ポイントは 3 点。

A)	該当 AI モデルのサイバー攻撃能力が、 4 月に話題になった Claude Mythos Preview に匹敵する高い水準だということ
B)	それにも関わらず、サイバー攻撃を依頼された際の AI の拒否態度が強固でなく、実行してしまうケースが少なくないこと
C)	さらにそのようなサイバー攻撃力の高さとガードレールの甘さを併せ持った AI が、オープンウェイトモデルであり誰でも入手と改変が可能であること

（脚注 7 の情報をもとに大和総研作成）

サイバー犯罪者が、強力なサイバー攻撃能力を持つ AI を活用し、サイバー被害を深刻化させる近い将来の事態に警鐘を鳴らしていたのだが、残念ながら、今私たちが直面している状況に非常に近い。警告が早くも現実化してしまっていると考えられる。

※なお、中国製のオープンウェイト AI モデルが効果的なサイバー攻撃に応用可能であるという

---

[3] 2026 年 1 月～9 月の国内セキュリティインシデントを振り返る
https://www.trendmicro.com/ja_jp/jp-security/26/j/securitytrend-20261016-01.html

[4] 相次ぐ不正アクセスの攻撃元の多くが AI であることを確認
https://gmo-cybersecurity.com/blog/20261019/

[5] AI エージェントによるウェブサイトへの攻撃の注意喚起
https://www.scutum.jp/information/waf_tech_blog/2026/10/waf-blog-085.html

[6] AI-driven, large-scale nation-state campaign using ARTEX enables supply chain compromise
https://www.microsoft.com/en-us/security/blog/2026/10/21/storm-0940-uses-artex-agentic-ai-connectwise-supply-chain-compromise/

[7] GLM-5.3 and the spread of advanced cyber capabilities
https://www.anthropic.com/research/glm-5-3-and-the-spread-of-advanced-cyber-capabilities

話と、中国関連のサイバー犯罪グループがそれらを実行しているという話は、無条件に結びつく内容ではない。

日本だけでなく、韓国をはじめ世界各国でサイバー攻撃が報告されている。私たちは今、サイバー攻撃と被害が AI によって 1 段階深刻化する、歴史的転換点をリアルタイムで経験している。

サイバー攻撃と被害の次の展開予想

それにしてもなぜ、情報漏洩事案が急増しており、システムの停止や企業脅迫といった種類の被害事案の増加率を大きく上回っているのだろうか。理由の考察については、データが少ないこともあり、現時点ではまだ共通理解の形成まで至っていないと考えている。複数の専門家によって多様な意見が交わされている段階という印象だ。

＜情報漏洩事案が多い理由　仮説ベース＞

①	個人情報を大量窃盗して、闇マーケットで売ろうとしている
②	個人情報を大量窃盗して、複数の企業のデータを結合し、フィッシング詐欺や特殊詐欺の精度を上げようとしている
③	個人情報を大量窃盗して、後にデータ公開で企業を脅迫し、身代金を得ようとしている
④	データ窃盗は、企業システムの Web 接点から少ない手数の攻撃で実現できるケースがあるので、相対的に手数が多くかかる種類の他の攻撃よりも、優先して実行されている
⑤	実はランサムウェア攻撃の件数も同じくらい増えているのだが、多くの企業がまだ気付いておらず潜在しているので、情報漏洩事案が多く見えている
⑥	サイバー犯罪者が、データ窃盗活動を通じて企業システムの状況を探索し、次の段階のより深刻なサイバー攻撃の準備をしている

（各社公表資料等を基に大和総研作成）

どの考え方も、もっともらしい内容を含み、かつ、同時に十分な根拠がない。確固とした証拠というよりも推測に基づいた議論が、継続中の状況だ。

私も同じで強い根拠を示すことは難しいのだが、個人的には、【⑤】も可能性があると評価している。

理由は、データ窃盗というゴールが、AI を使う先駆的なサイバー犯罪者が目指しそうな経済的利得の規模に見合わないのではないかと思うからだ。個人データを売るとか、詐欺の精度を上げて儲けを増やすとか、個人データの身代金を得るとか、そのような行為から得られる犯罪成果は、サイバー攻撃の先駆的な AI 化という試みに対して不釣り合いなのではないかと評価している。攻撃者の意図は裏付けられてはいないが、もっと大きな規模の犯罪成果を企図する可能性も否定できないだろう。

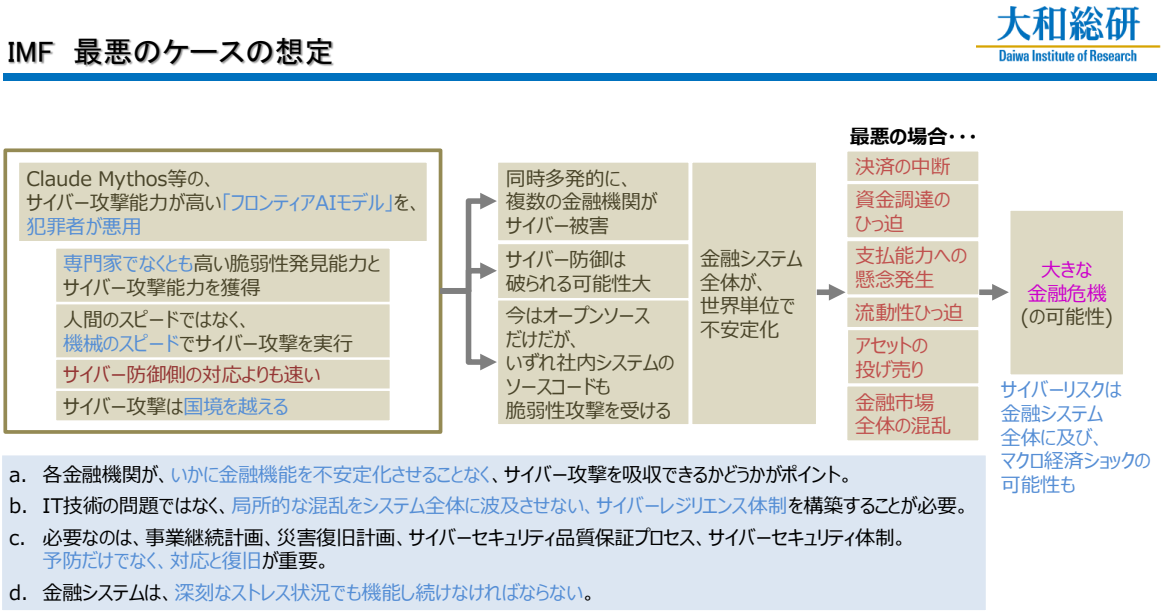
そしてその延長として、【⑥】の検討が避けられない。

すでに、他人の金融口座の乗っ取りや違法な資金操作の被害が発生している。[8] [9]

サイバー攻撃者の意図が明確に裏付けられている現状ではないが、より被害深刻度の高い犯罪行為がサイバー攻撃者の発展的な狙いで、現時点での活動は前段の準備に過ぎない可能性も、疑うべきなのではないか。継続して注視し、警戒していく必要がある。

IMF(国際通貨基金)は 2026 年 5 月、「サイバー攻撃能力が高い AI の出現は、世界に金融システム不安と金融危機をもたらす可能性がある」というブログを公開している。[10]

5 月当時は、最悪の事態を想定したレアケースのリスクを評価していると解釈し読み流していたのだが、AI サイバー攻撃によって金融口座乗っ取りが行われ犯罪者が口座資産を盗難する事態がもし頻発するならば、IMF が警告した金融パニックが本当に現実化しかねない。



Copyright(C) Daiwa Institute of Research All rights reserved.

(脚注 10 の情報をもとに大和総研作成)

今後必要な対応

いずれにしても、企業の Web システムを標的としたサイバー攻撃は今後も収束せず、攻撃件数の多い状況がまだ続くだろう。しかも、AI による自律的な判断と行動を取り込んでスピー

[8] 連結子会社の提供サービスへの不正アクセスに関するお知らせ
https://ir.group.gmo/pdf/irlibrary/gmo_disclose_info20261023.pdf

[9] 外部委託先への不正アクセスに伴うお客様情報の漏洩について(第 2 報)
https://www.iwaicosmo.co.jp/info/2026/10201710.html

[10] Financial Stability Risks Mount as Artificial Intelligence Fuels Cyberattacks
https://www.imf.org/en/blogs/articles/2026/05/07/financial-stability-risks-mount-as-artificial-intelligence-fuels-cyberattacks

ド・効果・攻撃件数を向上させた、レベルアップされたサイバー攻撃を受け続ける状況を、各企業は覚悟すべきだ。

AI を使ったサイバー攻撃において、AI は自律的に、そして個社別の IT 設定に応じて柔軟に、人間が見過ごしているシステムの穴を見つけ侵入して来る恐れがある。単体では影響の小さいシステム不備を複数組み合わせて侵入する場合もあれば、今まで誰も気付かなかったシステム不備に未知の攻撃を仕掛けて侵入する場合もあるだろう。情報漏洩報道の第一波、第二波に名前が挙がらなかった企業でも、次の攻撃の時には、プロンプトの中身を変えた AI による変化した攻撃を防御できないかもしれない。

そのような困難な状況の中で、各企業には今後、どのような対応が求められるのだろう。

当然ながら、公表済みの深刻な脆弱性への対応は、最低限の対応として完了させておかなければならない。自社システム群のどこにどのような脆弱性が残っているか把握し、深刻なものへは補正対応を済ませる。そうしなければ、もういつ、AI によってサイバー攻撃されてもおかしくない状況だ。実際に推進すると簡単な対応ではないのだが、それでもこのラインが最低限の対策となる。

また、これも基本的な対応策となるが、いわゆる多層防御を、自社システムに対して適切に設定する必要がある。自社システムの設定不備をひとつ突破されたとしても、次の壁で被害に至る事態を食い止める。

＜多層防御の例＞

● 自社システムと Web との境界線に、ウェブアプリケーションファイアウォールやネットワークファイアウォールを設置し、通信の入口段階で適切な防御対応を行う
● ランタイムアプリケーション自己防御等の製品を適切に導入・設定し、自社システムのアプリケーション実行局面においても適切な防御機構を機能させる
● 自社データベース周辺には、データ保護に特化したセキュリティ製品を適切に導入・設定し、アプリケーションやネットワークとは独立した個別の防御を適切に行う
● ユーザーの成りすまし対策には多要素認証の設定を徹底し、ユーザー情報が犯罪者に漏れた場合でも成りすましが成功しづらい状況を確保する

（大和総研作成）

次に、AI を使った自社システムへの疑似サイバー攻撃テストを実施し、事前に AI サイバー攻撃を想定した自社システムの弱点を把握する方法がある。いわゆる、AI ペネトレーションテストといわれるサービスを、提供するツールや事業者が複数存在する。思いもよらないシステム不備が発見されるかもしれないので、見つけ次第補正する。

ここまで述べた企業の防御対応には、一定の人員と体制が必要だ。サイバーセキュリティ組織と運営の強化を目的に、採用・人員増強・体制整備・人材育成・専門コミュニティとの窓口

強化など、覚悟を決めて人的投資と増強を行わなければ、AI 時代の適切なサイバー防御を実現することはできないだろう。

さらに総括するとどの企業でも、システム投資も人的投資も、従来よりも増額することが避けられないだろうと評価している。昨今、AI への投資がどの企業でも盛んであり並行して AI 人材への投資の必要性もよく語られているが、営業支援や業務効率化を実現する AI への投資だけではなく、サイバーセキュリティを支援する AI と人材への投資も、両方同じくらい大切だという議論が、2027 年頃には普通になっているのではないかと予想する。

最後に、共同防衛の戦略が検討される可能性がある。

サイバーセキュリティはもうすでに、非競争領域になっている。同業他社同士でも協力して、お互いがサイバー攻撃にあわない共通の目標を目指す戦略が、受け入れられやすい状況と評価できる。

業界のトップ企業や業界団体、業界内でシェアの高い IT ベンダーなどが主導した、共同サイバー防衛機構が、検討され実現していくことも近い将来に起こりえるだろう。

自民党の国家サイバーセキュリティ戦略本部は、政府へのサイバーセキュリティ対策提言を取りまとめたと発表し、各省庁をまたいだ体制を構築し国を挙げてサイバー防御能力を高める必要性を訴えている。[11]

「サイバー対処能力強化法」「サイバー対処能力強化法整備法」も 2026 年 10 月に施行され、法定の手続きのもとで自衛隊もしくは警察庁がサイバー攻撃元サーバーへアクセスし、無害化措置を行うことも可能となっている。[12]

サイバーセキュリティの共同防衛に、将来的に国が関わる構成となる可能性も考えられないことではなく、むしろ各企業は今から、サイバー防御を個社ではなく共同で実現するための活動を、検討していくことが望ましいかもしれない。

[11] 現在は「サイバー有事前夜」　サイバーセキュリティ対策へ提言取りまとめ
https://www.jimin.jp/news/information/214417.html

[12] サイバー対処能力強化法等の施行について
https://www.cyber.go.jp/pdf/press/siryou_20260930.pdf