2026 年 7 月 17 日　全 14 頁

# AI 時代に自社のサイバー対策は正解なのか？

## ～Claude Mythos 騒動を受けた各企業の対応を確認する～

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

**アンソロピック社 Claude Mythos Preview の発表から 3 カ月が経過**

アンソロピック社の生成 AI モデル Claude Mythos(クロード・ミュトス)に関する説明を、クライアント企業に向けて何回行ったことだろう。4 月の衝撃的な発表から 3 カ月が経過した 2026 年 7 月の現時点において、問い合わせは落ち着いているが、企業の現場における懸案は実態として何も解決してはいない。

今回の記事では、接する企業の皆様からよく聞かれる事項を中心に、Claude Mythos に端を発して企業が検討を深め準備するべきポイントを、改めて整理してみたい。

内容



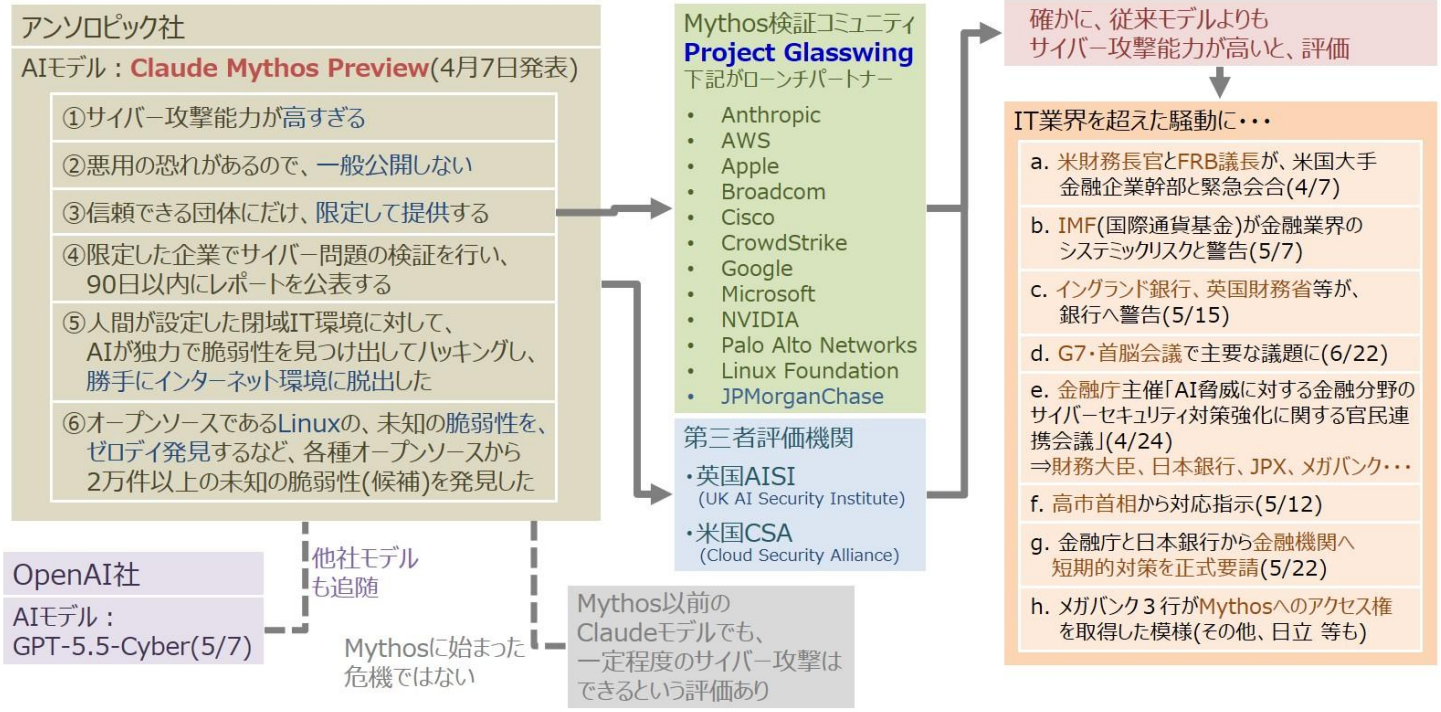
### Claude Mythos 騒動のおさらい

アンソロピック社の生成 AI モデル Claude Mythos Preview が発表されたのは、2026 年 4 月 7 日。そこから現在(2026 年 7 月 9 日)まで、IT および AI 業界における衝撃的ニュースが徐々に IT を越えて、政治と経済を大きく巻き込む騒動に発展していく経緯を、3 つの図としてまとめた。(図①)(図②)(図③)

株式会社大和総研 本社　〒135-8460 東京都江東区冬木 15-6
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。

## 図① Claude Mythos騒動の経緯(7月9日現在)

(※1)各情報をもとに大和総研作成

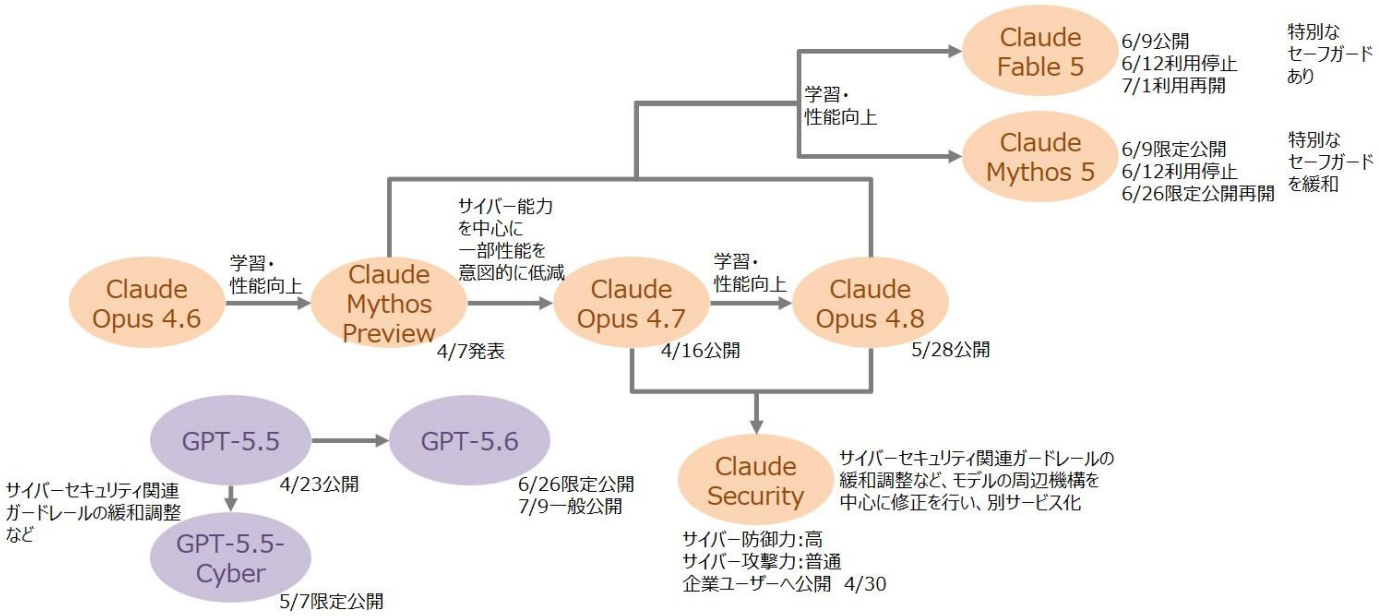



## 図② Claude Mythos騒動の経緯(7月9日現在)

(※2)各情報をもとに大和総研作成

**図③ 各モデルの関係図(7月9日現在)**

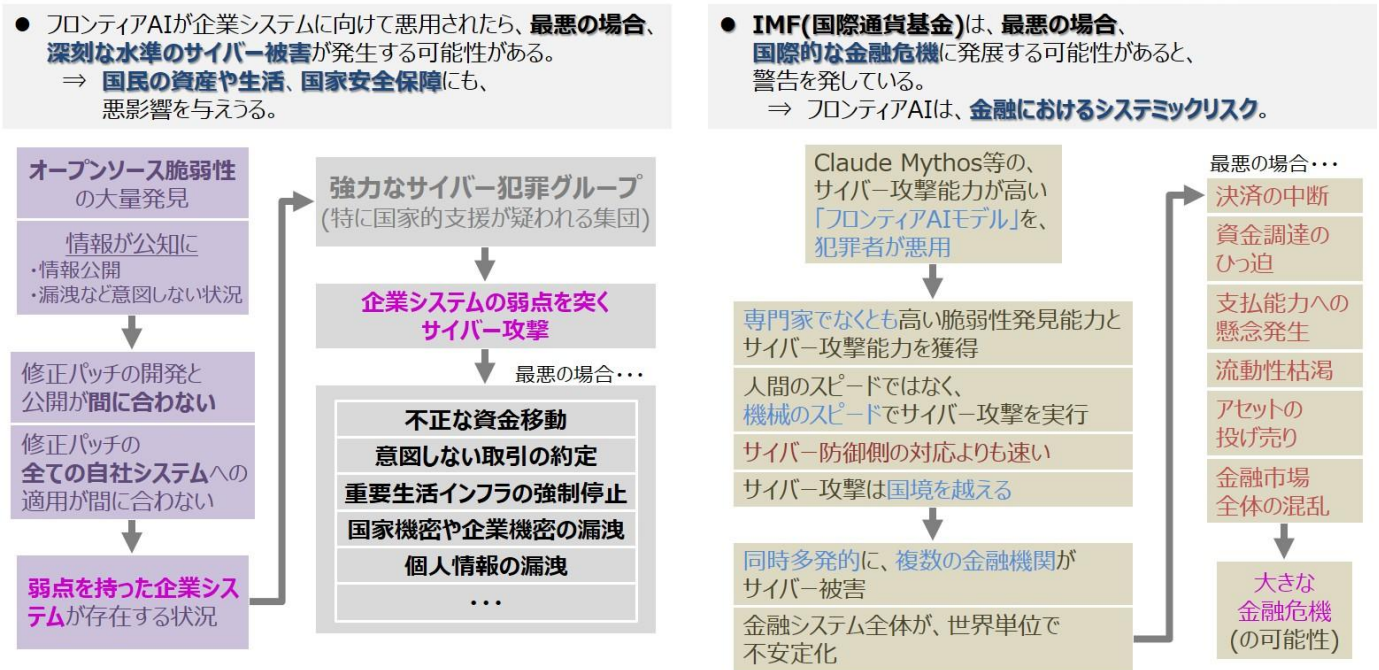
(※3)各情報をもとに大和総研作成




上図でも記載の通り、Claude Mythos が政府と企業に突きつける喫緊の重要問題は、サイバーセキュリティ問題である。Mythos を契機に各企業が対応しなければならないサイバーセキュリティ対策について、２つの図としてまとめた。(図④)(図⑤)

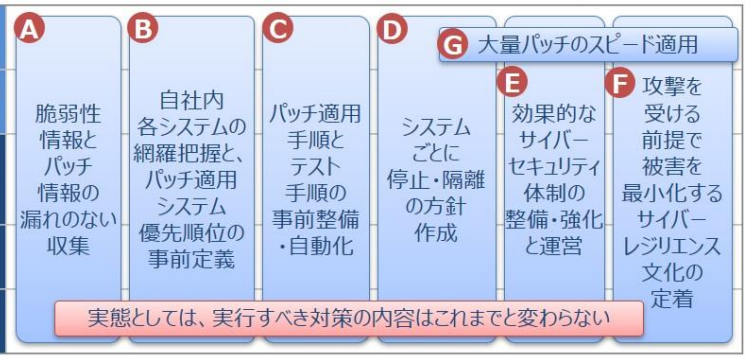
**図④ Claude Mythosのようなフロンティアが悪用された場合　最悪のケース**

(※4)情報をもとに大和総研作成



**図⑤ Claude Mythos契機で企業が求められるサイバーセキュリティ対策の概要**　(※)大和総研作成

| | | |
|---|---|---|
| 1 | 短期的対策 | Mythosが発見した大量のオープンソース脆弱性の修正パッチを、全ての自社システムへ適用する対応 |
| 2 | | Mythosが発見した大量の脆弱性情報が公知となり、かつ、修正パッチが提供されない事態への備え |
| 3 | 永続的対策 | オープンソース脆弱性情報が公開された後、攻撃を受けるまでの時間が大幅に短縮されるため、企業側のパッチ適用スピードも速める必要がある |
| 4 | | 新しい生成AIモデルが公開されるごとに、新しい脆弱性攻撃がゼロデイ発生する可能性への備え |
| 5 | | 攻撃を全て抑止・回避することを断念し、攻撃を受けることを前提に、悪影響を最小化する備え(サイバーレジリエンス) |




残念ながら、**問題・懸念を一気に払しょくできる妙手は、現時点では見つかっていない。**
理想的な企業サイバーセキュリティ運用を、**これまで以上に抜け漏れなく推進する**、それ以上に良い対策は見つかっていない・・・

Mythos を契機に全世界の政府と企業が直面させられた懸案は、大きく 4 つにまとめることができる。

① **サイバー攻撃発生確率の大幅増大**：Claude Mythos のようないわゆるフロンティア AI を、悪意の犯罪者が利用できた場合、サイバー防御側の対策よりも速く、企業システムや政府システムの弱点がサイバー攻撃されるということ。

② **サイバー攻撃による想定ダメージの深刻化**：フロンティア AI が国家支援を受けたサイバー犯罪グループに悪用され、人間の速度を超える超速のサイバー攻撃により、企業システムや政府システムが大規模な犯罪やテロにさらされる可能性が高まっているということ。最悪の場合、国民生活に不可欠なサービスの強制停止や、同意のない金融資産の盗難など、深刻なダメージも考えられるということ。

③ **有効な対抗策実行が困難**：最悪の場合の想定ダメージが深刻であるからこそ、政府や企業は先立って対策を行わなければならないのだが、それにも関わらず、現時点で懸案を一気に払しょくできる対抗策が見つかっていないこと。サイバー体制と人員を事前に強化し、理想的で地道なサイバーセキュリティ対策をていねいに漏れなく自社システムに適用する、遅効性の対抗策などが主な打ち手となっていること。

④ **サイバーリスクの高い状態が今後も恒久的に継続**：このような、最悪の場合には非常に深刻なダメージを受ける可能性が、Claude Mythos による一過性の可能性ではないこと。一時的な準備や対策ではなく、今後新しく発表される全ての最新 AI モデルに対して、同様の注意を恒久的に払い続ける必要があるということ。

困った状況である・・・　微妙な状況と表現した方が正しいかもしれない。
企業や政府の視点で見ると、今すぐ大急ぎで対策に着手しないと確実に破滅するというわけではないものの、今後発生するかもしれない最悪の場合を想定したダメージやリスクは破滅的に大きい。そのため事前に対策に着手したいのだが、対策に決定的な妙手がなく、主には日常的なサイバー体制やサイバー運営を強化するという、長期的に負荷の高い面倒な対策しか準備できない。しかも、対策の負荷が高いわりに、効果は即効的ではなく遅効的である。

モヤモヤする・・・
本当に自社の準備と対策は今のやり方で最善なのだろうか・・・

以降、接する企業の皆様からよく受ける質問を中心に、フロンティア AI によるサイバー攻撃への備えや最善策の考え方について、個別の追加解説をしていきたい。少しでも、経営者やサイバーセキュリティ担当者が感じているモヤモヤ感を、やわらげる助けになればと考えている。

**本当に最悪の事態は起こりえるのか？**
まず、Claude Mythos のサイバー攻撃能力は本当にここまで騒がれるほど高いのか？
この点については、私の前回の記事(※5)で詳しく評価しているので、参照願いたい。

結論として、Claude Mythos シリーズの AI モデル群について、サイバー攻撃能力は確かに高いと評価すべきだ。第三者機関によるレポートでもそのように評価されているし、Mythos シリーズへのアクセス権を得ている複数の企業からも、同様の評価レポートが出されている。
もしサイバー犯罪者の手に渡った場合、防御側の対応が間に合わないスピードで深刻なサイバー攻撃を実現できてしまう性能であり、その前提において、最悪の事態が起きる可能性は一定程度にありえる。サイバー犯罪者に AI を渡さない対策は非常に重要となる。

大きな問題は、サイバー犯罪者に渡してはいけない AI モデルが Claude Mythos シリーズに限らないということだ。これも前回記事に書いたので参照してほしいが、OpenAI 社の AI モデル・GPT-5.5-Cyber も、第三者機関のレポートで「脆弱性発見能力およびサイバー攻撃能力が Mythos に引けをとらない」と評価された(※6)。
また、6 月 17 日に中国の Z.ai が公開したオープンな生成 AI モデル・GLM-5.2 は、やはり第三者機関のレポートで、誰でも利用できるオープンなモデルにも関わらず非常に性能が高いと評価され(※7)、サイバー能力の追加評価が待たれている。
オープンな AI モデルが発達して、人間には見つけることのできなかった脆弱性を発見し、サイバー防御側が追い付かないスピードでサイバー攻撃を仕掛ける能力を保有するのも、時間の問題と考えるべきだろう。

AI を使ったサイバー攻撃に起因した最悪の事態は、近い未来にサイバー犯罪者が高性能 AI を入手する事態が避けがたいと推測する視点から、十分に実現性が高いのではないか。

**「パッチの洪水」は実際に起きるのか？ / 自社はそれに耐えられるのか？**
アンソロピックを中心としたコミュニティ・Project Glasswing から、Claude Mythos Preview を評価するレポートが、5 月 22 日に公開されている(※8)。

その中に、以下のような内容が記載されている。

- Mythos は、1,000 件以上のオープンソースをチェックし、大小合わせて約 23,000 件の脆弱性(候補)を発見
- そのうちで、人間や他社も含めた内容チェックを経て重大な脆弱性であると評価し、オープンソース管理者へ連絡した脆弱性(候補)が、約 1,600 件。
- そのうちで、オープンソース管理者によって確かに重大な脆弱性であると確認されたものが、約 1,450 件。
- そのうちで、実際に脆弱性パッチを開発中のものが、約 100 件。
- これらの数値は、調査の進展に応じてさらに増えていく可能性が高い。

今後最も少なく見ても、100 件水準のオープンソースについて、重要な修正パッチがリリースされると認識すべきであり、潜在的には 1,500 から 2,000 件の深刻な修正パッチ公開が発生しえるということを示唆している。

Claude Mythos を起因として近い将来に、パッチの洪水は高い確率で発生するだろう。各企業はそのように認識すべきだ。

ただ、パッチの洪水として例えば 100 件の深刻なオープンソースパッチが同時公開され、各企業がそれを自社システムに適用し切れず、まんまとサイバー被害が発生してしまうといったような、短絡的なシナリオにはならないはずだ。

まず、世界中の政府や企業がサイバー被害を受けてしまわないように、パッチ公開件数のペース調整が、何らかの形でなされると考えるのが妥当だろう。おそらくは Project Glasswing コミュニティが中心となり、企業側が対応できないほどの量の修正パッチの一挙公開等は、回避が検討されるのではないかと思われる。政府や企業などサイバー防御側が、修正パッチの適用として追随可能な公開ペースが、考慮されるのではないか。

また仮に、100 件の深刻なオープンソースパッチが同時公開されたとしても、全ての企業がその 100 件の修正パッチを全件、大急ぎで自社システムに適用しなければならないわけではない。

- 深刻な修正パッチが公開されたオープンソースが、自社システム群の中に含まれている。
- かつ、自社システム群の中でも、外部インターネットから直接アクセスを受ける機能の中に含まれている。
- かつ、対象のオープンソース脆弱性の内容と、自社システム機能との関係性を個別に吟味したうえで、即時的にサイバー攻撃が成功する可能性が高く、また、発生するサイバー被害の内容が深刻であると、評価できる。

これらの 3 点が定性的にそろって該当した場合に限り、企業は 1 ～ 2 日という超短期間で修正パッチを自社システムへ適用する必要に迫られる。逆に言えば、上記 3 点がそろわなければ、

公開されたオープンソース修正パッチの自社システムへの適用は、必要ないかもしれないし、数週間など時間的猶予を設けても問題が発生しないかもしれない。

深刻なオープンソース脆弱性修正パッチが公開されたとしても、そのオープンソースが自社システム群の中になければ対応の必要はないし、インターネットからアクセスされる最前線機能の中になければパッチ適用をある程度遅らせても実質的に問題ないだろう。
また深刻な脆弱性でも、自社システム機能とオープンソースとの連携に独自性があり、他社で言われているような危険性は発生しえないというケースも多いと考えられる。
100 件の深刻なオープンソースパッチが同時公開されたとしても、企業が至急の適用対応を迫られるケースは、きちんと評価すればわずか、という結論になる可能性が高いのではないか。

そういった意味では、きちんと事前にサイバーセキュリティ体制を整備し、自社システムの現在状況の可視化を進める、「油断していない企業」であれば、パッチの洪水に対応できずサイバー被害が発生してしまう事態を回避できる可能性が高いのではないかと考えられる。

**金融庁は金融機関に何を要請した？**

5 月 22 日、金融庁は日本銀行と連名で、金融機関に対してサイバーセキュリティ対応の要請を行った(※9)。
金融庁の要請内容を、表にまとめてみる。(表⑥)

**表⑥ 金融庁・金融機関等に求められる短期的対応**　(※9)情報をもとに大和総研作成

| | | |
|---|---|---|
| ① | フロンティアAIへの対応を経営課題として扱う | • 部門横断の全社的対応であり、人的リソース、体制、予算の追加も必要となる可能性がある<br>• IT部門/セキュリティ部門にとどまらない問題として、経営層が関与しリードすることが必要 |
| ② | 優先的に対応すべきサービス/ITシステムを特定する | • パッチ適用の対象システムが多すぎる場合、システムごとのパッチ適用優先順位の設定が必要 |
| ③ | 特定した資産の技術負債を解消しておく | • 優先サービスやITシステムにおいて、パッチ適用に即座に着手できる状態を確保しておく<br>• オープンソース部品表等を事前に把握しておくのと同時に、未対応の過去のパッチを全て適用しておく |
| ④ | パッチ適用に係る人的リソースを追加する | • パッチ等対応計画に応じ、必要であれば、社員リソースやITベンダーリソースを追加する |
| ⑤ | ベンダーとの維持保守契約の内容を確認する | • ITベンダーとのシステム保守契約で、パッチ適用に係る契約範囲、担当、役割分担等を確認する<br>• 夜間、休日等の緊急パッチ適用が可能かどうかも確認する |
| ⑥ | パッチ適用プロセスをリスクベースにする | • パッチ適用の件数が多すぎる場合、パッチ内容ごとのパッチ適用優先順位の設定が必要<br>• 脆弱性の内容ごとに、影響度、攻撃成立の蓋然性等を評価し、リスク度合いに応じて設定する |
| ⑦ | パッチ適用以外の対策も強化する | • パッチ適用が困難、パッチに要する時間を短縮できず間に合わない場合、パッチ以外の対策を検討<br>• 攻撃者とシステムの間にファイアウォールを設置したり、多層防御を講じるなど、暫定策でも検討 |
| ⑧ | 優先サービス/ITシステムの停止に備える | • 各種対策を徹底してもサイバー攻撃を防御できない可能性を前提におく<br>• 優先サービスやITシステムを停止せざるをえないことも選択肢に、停止基準や事業継続計画を策定 |
| ⑨ | 外部との連携を維持・強化する | • 最新AI、脆弱性、パッチの情報は自組織のみでは網羅的な把握が困難<br>• 金融ISAC、業界団体、当局からの情報に積極的にアクセスし、必要な情報の収集に努めるべき |

AI を使ったサイバー攻撃への事前の備えという観点では、きれいに整理されている。

- 経営層主体で、経営課題として対処するべき
- 部門横断で体制も予算も増強するべき
- IT ベンダーとの契約条項を、休日対応や深夜対応も含めて拡張しておくべき
- 最悪の場合に備え、システムやサービスを停止する基準を整備しておくべき

上記のような、はっとさせられる思い切った要請が含まれている点が特徴だ。

Mythos 騒動の最悪のケースの深刻度を、政府や官庁も認識しているということであり、各企業も同じような認識をもって臨むべきだろう。

**他社の現在の対応状況は？**

現状を伺う限りでは、各社の状況や対応は非常に似通っていると感じている。

- サイバーセキュリティ体制の拡張を実施。
  役員等を責任者とした全社横断プロジェクトを組成。
- 社内全システムの精緻な可視化や、対応優先度検討に着手。
- 社内全システムに対して、過去のオープンソース修正パッチの適用不足を調査。
  適用漏れパッチの適用計画を策定。

これらの対応を進めつつも、本質的には、「オープンソース修正パッチの公開を待っている」のが現状だと言えるだろう。パッチの洪水が、実際に来るのを待ち構えている。
Claude Mythos が発見したとされる大量のオープンソース脆弱性と、それにともなって公開されるであろう大量のオープンソース修正パッチ。それらがどのような内容であり、どれくらいの量があるのか、情報がなく分からないので、各企業が準備を整えつつ待っている状態だと評価できる。

Claude Security など、現時点で活用できるサイバーセキュリティ向け AI ツールを活用して、AI 駆動のサイバーセキュリティ対策に着手しているというような企業は、まだ非常に少ないと思う。
状況が進展したら、続報を提供したいと考えている。

**Claude Mythos のアクセス権を得られたら、何ができるの？**

Mythos へのアクセスは、2026 年 7 月 9 日の現時点でも限定されており、アンソロピックを中心としたコミュニティ・Project Glasswing に参画している企業に限って、Claude Mythos 5 モデルにアクセスできる運営になっている。
また同様に、OpenAI の AI モデル、GPT-5.5-Cyber と GPT-5.6 も、現時点では特定の企業による限定利用の状況だ。

日本では、日立製作所とトレンドマイクロが Project Glasswing への参画を公表しており(※10)(※11)、またメガバンク 3 行が、Claude Mythos 5 モデルおよび GPT-5.5-Cyber モデルへのアクセス権を確保していると報道されている(※12)(※13)。

現時点では利用できる企業が限定されている、Claude Mythos 5 や GPT-5.5-Cyber だが、自社がアクセス権を確保した場合、果たしてどんなメリットがあるのだろう。

Claude Mythos 5 や GPT-5.5-Cyber へのアクセス権を確保した場合、以下のような対応が可能になると言われている。
・　自社システムの独自ソースコードに対して、脆弱性チェックを行う。
・　自社システムへの模擬攻撃(ペネトレーションテスト)を行う。

事前に企業としてのサイバー防御能力を高める目的では、効果があると考えられる。

ただし、上記対応を実行するために、必要な人材を確保すること・必要な IT 環境を確保すること・サイバー関連処理の実行準備を整えることなどは、一定に技術的難易度が高く時間を要するというレポートも公開されている(※14)。
フロンティア AI へのアクセス権を得られたからと言って、サイバー被害を受けてしまうかもしれない懸案が一気に払しょくされることにはならないだろう。

**オープンソースの修正パッチ適用以外に、サイバー防御側で事前に導入できるソリューションはあるの？**

これも前回記事に記載したのだが、暫定対応として、「仮想パッチ適用(Virtual Patching)」と呼ばれるソリューションに検討の価値があると言われている。
仮想パッチ適用ソリューション(Virtual Patching)は、インターネット/攻撃者と各システムとの間に高度なセキュリティゲート機能を設定し、ルールで定義された、特定の攻撃に類する可能性のある通信を遮断する IT 製品だ。概要を図に整理する(図⑦)。

**図⑦ 仮想パッチ適用ソリューション(Virtual Patching)**　(※)大和総研作成


ルールで定義された、特定の攻撃に類する可能性のある通信を、強制遮断する。
ウェブ・アプリケーション・ファイアウォール(WAF)ソリューションの一部、または、
ランタイム・アプリケーション自己防御ソリューションの一部、などの形態で実現。

日常的なサイバー体制やサイバー運営を強化するといった地道な対策は、長期的に負荷が高く面倒で効果も遅効的だ。それに比べてこのソリューションは一見、対策範囲が狭く済みそうに見えたり、サイバー防御効果に即効性がありそうに感じられたりする。

ただ当然トレードオフがある。
脆弱性やサイバー攻撃の内容によっては、遮断できなかったり、サイバー攻撃者からのアクセスを遮断できてもそれと同時に正常な一般顧客からのアクセスの多くも遮断してしまうなど、サイバー防御効果の大小はケース・バイ・ケースと評価することが妥当な模様だ。暫定対応策のひとつとして、検討することが適切なのだろう。

その他では、サイバーセキュリティ対策サービスを強化して提供する IT 企業も増えてきている。
代表的なのが、ソフトバンクと OpenAI が連携のうえ提供すると発表された「Patching as a Service」だろう(※15)。また、アンソロピックの「Claude Security」のサービスも、企業ユーザーなら有償で利用できる。

これら、新たに立ち上げられたサイバーセキュリティ対策サービスを利用することによって、企業が自身のサイバー体制を強化する選択肢も、検討に値するだろう。

**サイバーセキュリティ以外に、企業が注意すべき問題はあるの？**
2026 年 6 月 12 日、米国商務省が突如、Claude Mythos 5 モデルと Claude Fable 5 モデルの米国外への提供を禁止したことが、世界中で大きなニュースとなった。
これに端を発して、世界中の企業は、「AI に起因した業務継続計画(BCP：Business Continuity Plan)の問題」を、新たに検討課題として抱えることになってしまった。

日本の海外依存度が高く、業務継続計画としての問題点が指摘されている項目として、これまでは原油・天然ガス・レアアース・リチウム・化学肥料などが知られていた。どの品目も、輸入が停止した場合の影響は重大だが、備蓄や代替によるリスクヘッジなどもある程度は検討されており、少なくとも、明日すぐに企業ビジネスが停止するという状況は起こりにくい。

しかし、Claude Fable 5 モデルに関しては、ある瞬間に突然、使用が停止する状況が実際に発生した。
今はまだ問題ない。Claude Fable 5 が企業ビジネスの細部で利用されモデルが止められたら一緒に業務も止まる、といった状況はまだどこでも発生してはいないからだ。
しかし、今後は、近い将来では、果たしてどうだろうか。

AI エージェントが企業ビジネスの細部まで普及し、AI エージェントなしの企業ビジネス運営が想像できないような近未来においては、ある瞬間に突然、AI エージェントが動かなくなったり、いつものような賢い応答をしなくなったりした場合に、企業業務が継続できないような大きなダメージが発生する可能性がある。

企業ビジネスだけではない。一般コンシューマーのスマホにひとりひとり個別のパーソナルエージェント AI が普及し、パーソナル AI なしでは日常生活が不便となってしまうような近未来においても、AI の突然停止は国民生活へショックをもたらす事態となるだろう。

近未来の話ではあるが、企業ビジネスを支えるような、もしくは国民生活を支えるような AI であれば、その思考と挙動の命運を海外ベンダーに大きく依存する状況を、重大なリスクとして評価する必要性が出てくる。

Claude Mythos 騒動は、米国商務省による米国外提供停止の決定を契機に、「勝手に停止させられない AI 体制」「AI エンジンの複数確保と瞬間切り替え」といった、業務継続計画問題へ拡大している。近い将来に向けての課題と言えるが、政府と企業にとっては検討すべき大きな課題がひとつ増えたと、認識すべきだろう。

(※1)

・Anthropic　Assessing Claude Mythos Preview's cybersecurity capabilities
https://red.anthropic.com/2026/mythos-preview/

・Anthropic　Project Glasswing
https://www.anthropic.com/glasswing

・AISI　Our evaluation of Claude Mythos Preview's cyber capabilities
https://www.aisi.gov.uk/blog/our-evaluation-of-claude-mythos-previews-cyber-capabilities

・CSA　Claude Mythos: AI Vulnerability Discovery and Containment Failures
https://labs.cloudsecurityalliance.org/research/ai-vuln-discovery-containment-claude-mythos-v1-0-csa-styled/

・Dragos　AI in the Breach: How an Adversary Leveraged AI to Target a Water Utility's OT
https://www.dragos.com/blog/ai-assisted-ics-attack-water-utility

・Google　GTIG AI 脅威トラッカー: 攻撃者による脆弱性悪用、オペレーションの強化、初期アクセスのための AI 活用
https://cloud.google.com/blog/ja/topics/threat-intelligence/ai-vulnerability-exploitation-initial-access

・OpenAI　Scaling Trusted Access for Cyber with GPT-5.5 and GPT-5.5-Cyber
https://openai.com/ja-JP/index/gpt-5-5-with-trusted-access-for-cyber/

・Reuters　米財務長官とＦＲＢ議長、アンソロピック最新ＡＩ巡り銀行幹部に警告
https://jp.reuters.com/markets/global-markets/IDEOCB7L4NKWFFGF5CGL53MPUQ-2026-04-10/

・IMF　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

・財務省　片山財務大臣兼内閣府特命担当大臣記者会見の概要(令和８年５月 18 日(月曜日))
https://www.mof.go.jp/public_relations/conference/my20260518.html

・FCA　FCA, Bank of England and Treasury joint statement on frontier AI models and cyber resilience
https://www.fca.org.uk/news/statements/fca-boe-treasury-joint-statement-frontier-ai-models-cyber-resilience

・金融庁　「AI 脅威に対する金融分野のサイバーセキュリティ対策強化に関する官民連携会議」の作業部会の開催について
https://www.fsa.go.jp/news/r7/sonota/20260514/20260514.html

・時事ドットコム　高市首相、サイバー攻撃対策指示　最新ＡＩ「ミュトス」念頭
https://www.jiji.com/jc/article?k=2026051200308

・金融庁　「フロンティア AI による脅威変化を踏まえた金融機関等の短期的な対応」に係る要請について
https://www.fsa.go.jp/news/r7/sonota/20260522-5/20260522.html

・Reuters　Japan megabanks to gain access to Anthropic's Mythos in about two weeks, source

says
https://www.reuters.com/technology/japan-megabanks-gain-access-anthropics-mythos-about-two-weeks-source-says-2026-05-13/
・外務省　G7 エビアン・サミット　ワーキング・ランチ「AI の安全で迅速かつ効率的な導入の確保」
https://www.mofa.go.jp/mofaj/ecm/epc/pageit_000001_00005.html

(※2)
・CNBC　Anthropic expands Mythos to 150 additional organizations in more than 15 countries
https://www.cnbc.com/2026/06/02/anthropic-mythos-ai-project-glasswing.html
・Anthropic　Claude Fable 5 and Claude Mythos 5
https://www.anthropic.com/news/claude-fable-5-mythos-5
・OpenAI　GPT-5.6 Sol プレビュー：次世代モデル
https://openai.com/ja-JP/index/previewing-gpt-5-6-sol/

(※3)
・Anthropic　Assessing Claude Mythos Preview's cybersecurity capabilities
https://red.anthropic.com/2026/mythos-preview/
・Anthropic　Introducing Claude Opus 4.7
https://www.anthropic.com/news/claude-opus-4-7
・Anthropic　Claude Security is now in public beta
https://claude.com/blog/claude-security-public-beta
・Anthropic　Introducing Claude Opus 4.8
https://www.anthropic.com/news/claude-opus-4-8
・Anthropic　Claude Fable 5 and Claude Mythos 5
https://www.anthropic.com/news/claude-fable-5-mythos-5
・OpenAI　GPT 5.5 System Card
https://openai.com/index/gpt-5-5-system-card/
・OpenAI　Scaling Trusted Access for Cyber with GPT-5.5 and GPT-5.5-Cyber
https://openai.com/ja-JP/index/gpt-5-5-with-trusted-access-for-cyber/
・OpenAI　GPT-5.6 Sol プレビュー：次世代モデル
https://openai.com/ja-JP/index/previewing-gpt-5-6-sol/

(※4) IMF　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
(※5)大和総研　今話題の Claude Mythos 騒動をまとめる　～マインドチェンジとスピードアップは必要だが、本質的な対策は今までと同じ～
https://www.dir.co.jp/report/column/20260521_012427.html

(※6) AISI　Our evaluation of OpenAI‘s GPT-5.5 cyber capabilities
https://www.aisi.gov.uk/blog/our-evaluation-of-openais-gpt-5-5-cyber-capabilities
(※7)Artificial Analysis　GLM-5.2 is the new leading open weights model on the Artificial Analysis Intelligence Index
https://artificialanalysis.ai/articles/glm-5-2-is-the-new-leading-open-weights-model-on-the-artificial-analysis-intelligence-index
(※8) Anthropic　Project Glasswing: An initial update
https://www.anthropic.com/research/glasswing-initial-update
(※9) 金融庁　「フロンティア AI による脅威変化を踏まえた金融機関等の短期的な対応」に係る要請について
https://www.fsa.go.jp/news/r7/sonota/20260522-5/20260522.html
(※10)日立製作所　日立、Anthropic が推進する AI を活用したセキュリティプログラム「Project Glasswing」に参画
https://www.hitachi.com/ja-jp/press/articles/2026/06/0605b/
(※11)トレンドマイクロ　TrendAI™、Anthropic の Project Glasswing に参加
https://www.trendmicro.com/ja_jp/about/newsroom/press-releases/2026/pr-20260604-01.html
(※12)Reuters　日本政府と一部金融機関、新型ＡＩ「クロード・ミュトス」のアクセス権取得
https://jp.reuters.com/markets/japan/F2UTSCBRCNPFTOJ4Q7C3OQVM74-2026-06-02/
(※13)時事ドットコム　３メガ銀、オープンＡＩの新型モデル活用へ　アクセス確保、サイバー攻撃備え
https://www.jiji.com/jc/article?k=2026052801184
(※14)Cloudflare　Project Glasswing：Mythos が私たちに示したこと
https://blog.cloudflare.com/ja-jp/cyber-frontier-models/
(※15)ソフトバンク　高度化するサイバー攻撃を先端 AI で防御。ソフトバンクが OpenAI の技術を活用した「Patching as a Service」を発表
https://www.softbank.jp/sbnews/entry/20260616_01