

Mythos報道の落とし穴

Atsushi Teraoka
寺岡篤志
NIKKEI Inc.

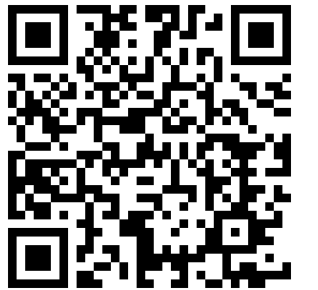


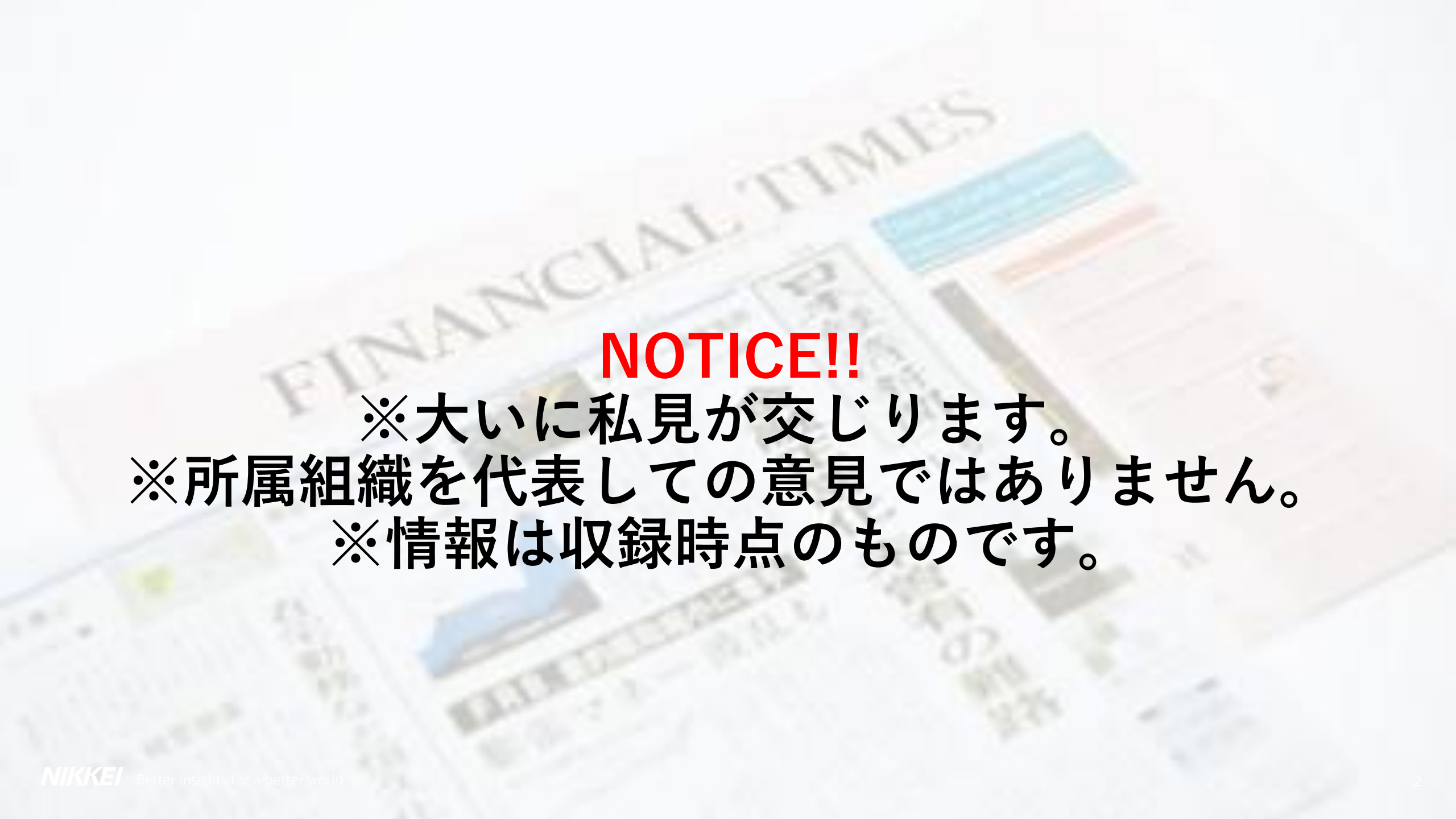
Mail : atsushi.teraoka@nex.nikkei.com

Linked in:



My articles:





NOTICE!!

- ※大いに私見が交じります。
- ※所属組織を代表しての意見ではありません。
- ※情報は収録時点のものです。

日本経済新聞記者 寺岡篤志

■取材テーマ

サイバーセキュリティ

(コーポレート、安全保障、犯罪、政策)

調査報道

情報工作

アンチマネーロンダリング

リスク管理



経歴

07年度 某大学ラグビー部卒

08年 日経入社

14年 マウントゴックス事件でセキュリティの取材を開始

22年 九州大SECKUN 3期生

24年 英国のChevening Cybersecurity Fellowshipに参加

25年 Fulbright Scholarshipフェローとして米メリーランド大でサイバー安全保障の研究に従事

26年 セキュリティベンダへ

9月からセキュリティベンダに
転職します。

既読
10:15

Mythosの問題は正しく報道されているか？

Mythosを打ち破ってくれ！！

10:16

Mythosは打ち破るものじゃないw

既読
10:17

去年からのわかサイバー祭り

※件数は6月中旬まで

- トヨタ自動車→記事タグなし
- KADOKAWA→記事タグなし
- 証券口座乗っ取り→記事タグあり、159件
- アスクル→記事タグあり、105件
- アサヒGHD→記事タグあり、69件
- Mythos→記事タグあり、109件

参考：

- 「クマ被害」→505件
- 「スペースX」→322件
- 「ナフサ高騰」→280件

① Mythosはゲームチェンジャーか？

① Mythosはゲームチェンジャーか？

- ❑ Curlメンテナー「大々的な宣伝は主にマーケティング目的。Still Goodだが、他のツールよりも高度なレベルにあるという証拠は見当たらない」
- ❑ 検出された「脆弱性」は5個。1つは低深刻度、3つは誤検出、1つは単なるバグ。

 daniel:// stenberg://
@bagder 2d

Based on hype, your thinking and #curl's maturity.
How many CVEs do you think will be the result of
Mythos scanning curl for the first time?

32% 1

40% 10

19% 25

9% 100

[Refresh](#) · 1,127 people · Closed

出典：Daniel://Steinberg://

① Mythosはゲームチェンジャーか？

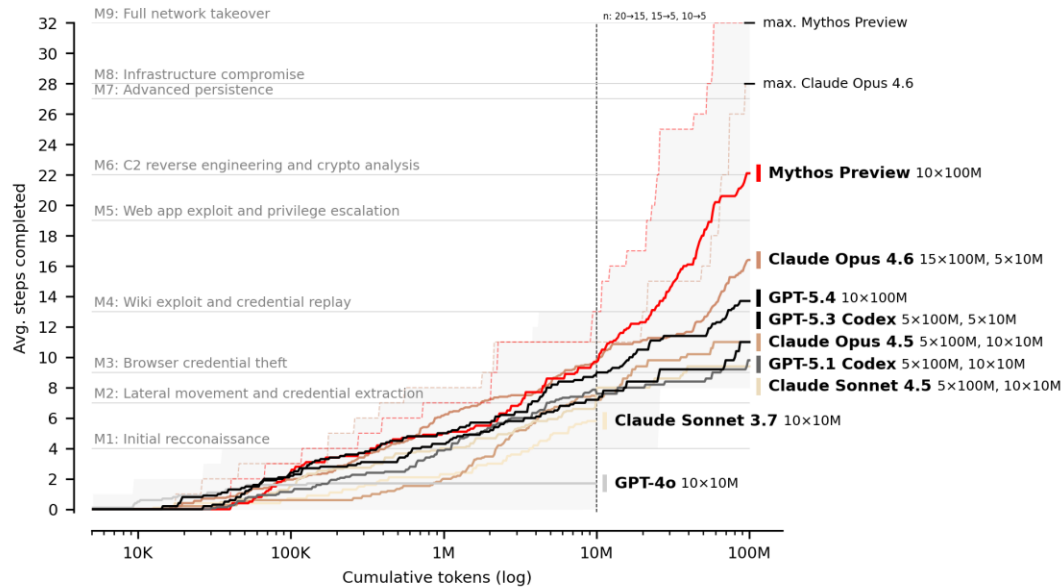
□ UK AI Security Institute 「CTFで73%をクリア」「企業ネットワークの攻撃シミュレーションを3/10で成功」

一方で・・・

□ OTの攻撃環境をクリアできず「限界を示した」

□ アラートをトリガーするような行動にペナルティなし。「十分に防御されたシステムを攻撃できるかどうかは断言できない」

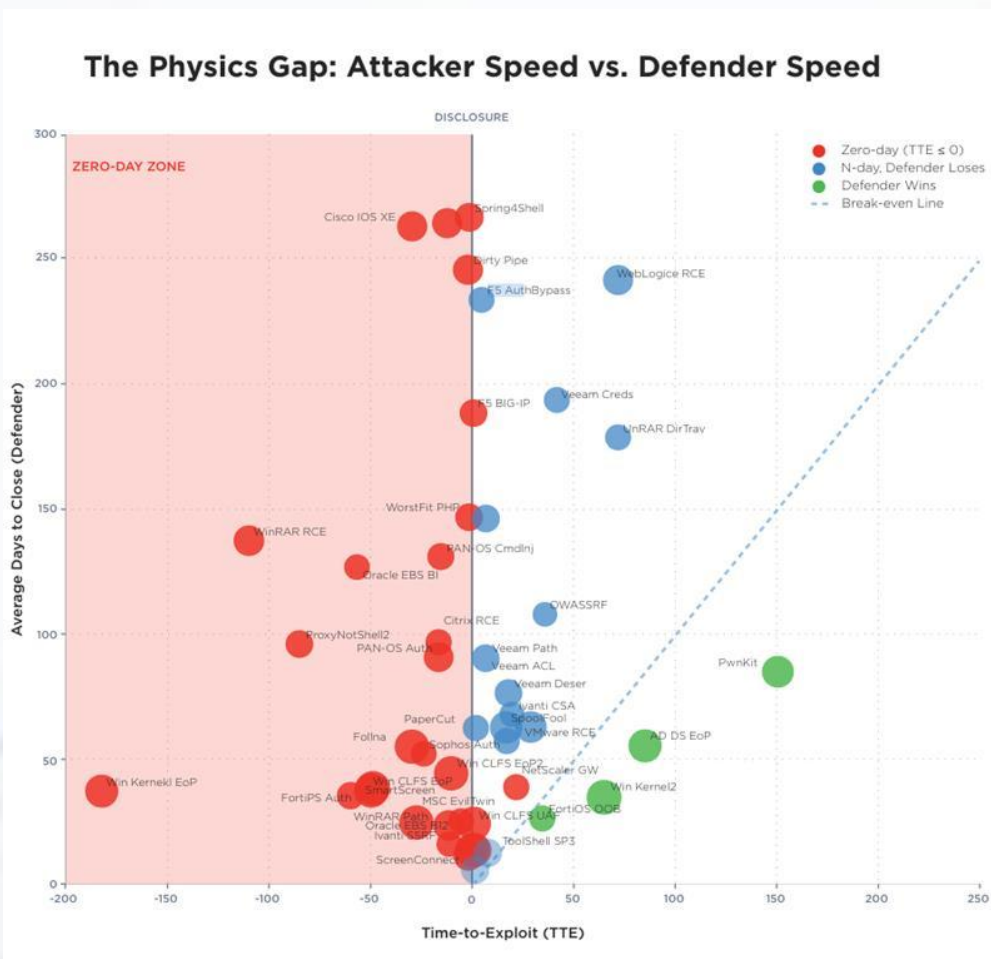
Completed steps on "The Last Ones" per spent tokens



出典：AI Security Institute

②今は「Mythos前夜」なのか？

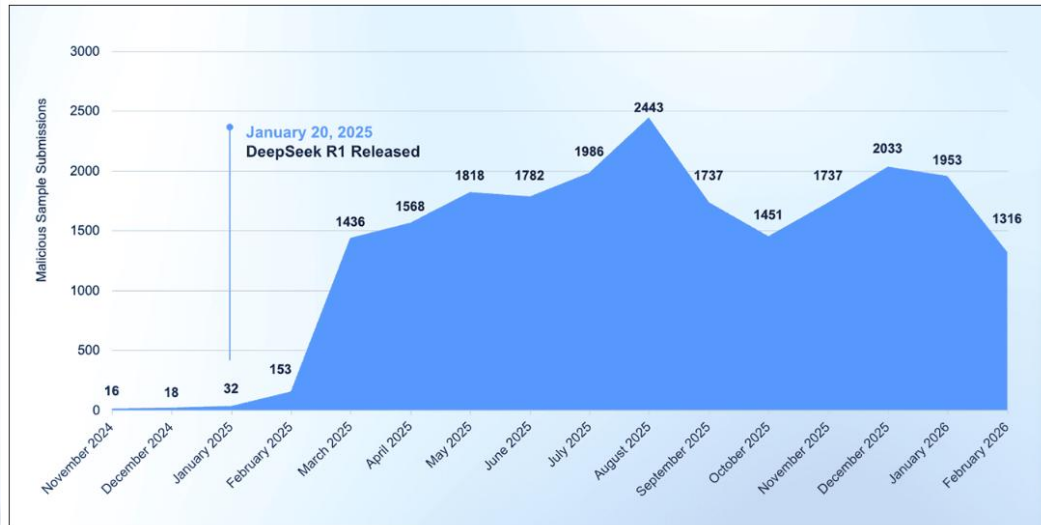
②今は「Mythos前夜」なのか？



- 2022~25年、悪用開始時期がわかっている主要なKEVの追跡調査
- 88%のKEVの修正平均時間は悪用開始に間に合っていない
- 52%はゼロデイ
- 「根本的な問題は人的ボトルネック」

出典：Qualys

②今は「Mythos前夜」なのか？



- Deep Seek R1のリリース（2025年1月）直後から、AI利用とみられるマルウェアが急増
- 月間検知数は最大で136倍に
- 開発者の推定使用言語は日本語が最多。理由は「不明」

出典：Arctic Wolf

②今は「Mythos前夜」なのか？

AN OPEN LETTER

On Transparent AI Cyber Protections

June 14, 2026

Dear Secretary Lutnick and National Cyber Director Cairncross,

We, the undersigned executives and technical leaders from across the United States and its allies, write to you to ask you to lift the export control directives on Anthropic's Fable and Mythos large language models and commit to an open, scientific and transparent process of handling AI risk assessments in the future.

First, we would like to state that we believe that:

- ❑ Mythosクラスのモデルは、脆弱性の発見と武器化に非常に優れているが、特別に優れているわけではない
- ❑ 署名者の多くは、セキュリティ監査やレッドチーム演習のために、他のモデルを日常的に利用している
- ❑ 中国のオープンウェイトモデルは米国のフロンティアモデルからわずか数か月遅れているだけ

出典：<https://freefable.org/>

③ Mythosはインフラへの脅威なのか？

②Mythosはインフラへの脅威なのか？

どれも短期的に終わる問題
じゃないよ・・・

□ 金融庁 について

テスト環境がないから更新
もできないんだよ・・・

も脅威変化を踏まえた金融機関等の短期的な対応に

経営課題として扱
う

ス／システムの特
定

技術負債の解消

脆弱性対応の人的
リソース追加

ベンダーとの維持
保守契約の確認

パッチ適用
クベ

本当に国民の理解得られる
の・・・？

能動的停止の検討

ISACなどとの連

サービス停止がむしろ増え
る！！

中長期的には脆弱性対応の自動化

③Mythosはインフラへの脅威なのか？

**Project YATA-Shield**※

1

- フロンティアAIモデルによるサイバーセキュリティ性能が向上する中においても、我が国のサイバーセキュリティが確保されるよう、**政府全体としての対策パッケージ**を取りまとめ。

基本的な認識・考え方

重要インフラ事業者等・政府機関等が取るべき対応

- 発見された脆弱性のパッチ適用やリスク緩和措置を速やかに実施（リスクベース）
- 基本的な対策、多層防御の実施、インシデント発生時の備え 等

※英国・米国政府の注意喚起も参考に

脆弱性を発見するAIの進化

- **Anthropic（Project Glasswing）**：Mythosへのアクセスを、ビッグテックや重要インフラ等に限定。
- **OpenAI**：GPT-5.5-Cyberへのアクセスを、一部の認証者に限定して付与。

実施する施策

重要インフラ事業者等・政府機関等への対応

- ① 重要インフラ事業者等への注意喚起等
- ② 金融分野での先行的な取組及び 他分野への展開
- ③ 人材育成支援
- ④ 政府機関等の情報システムにおける対応

脆弱性の発見・修正等への対応

- ① 外国政府機関やビッグテック等との更なる連携
- ② ソフトウェア・ベンダへの注意喚起
- ③ AISIによる技術支援等
- ④ 技術開発の推進
- ⑤ 高性能AIを活用したサイバー対処能力強化

※YATA：Yielding Advanced Threat Awareness with AI（脅威の可視化）の頭文字。
「正確に写す」という八咫鏡(やたのかがみ)もあるように、「AI性能の高度化に伴うサイバー脅威を正しく認識し、防御する／対応する」という趣旨。

うちのIT環境外部露出していないはずなんですが・・・まさか金融機関と同等に扱われるの・・・？

③ Mythosはインフラへの脅威なのか？

□ CISA Binding Operational Directive 26-04（連邦政府機関向け）

公開性：インターネットに露出しているか？

KEVステータス：KEVに掲載されているか？

悪用の自動化：攻撃プロセスをすべて自動化可能か？

技術的影響：悪用によってシステムを完全制御可能か？

高リスクのものは
「3日以内」
に対応すべき

（従来はKEVを14～21日）

脆弱性事例のうち3日以内に対応が必要なものはわずか1%で、60%以上は次回のシステムアップグレードまで延期される。

攻撃者が製品の脆弱性を悪用して主にコアネットワークを侵害している事例は確認していない。むしろ用いられるのはLiving Off the Land（LOTL）だ。LOTLへの対策としては、システム構成の強化、ネットワークのセグメンテーション、フィッシング対策に優れた多要素認証の導入などが効果的だ。

④ AI vs AIはスピードの勝負なのか？

④AI vs AIはスピードの勝負なのか？

Human in the Loop



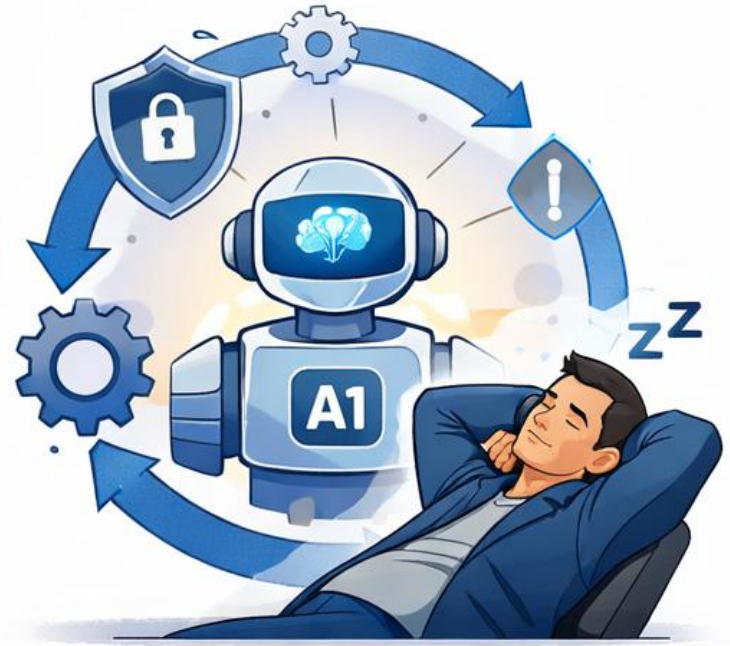
Decision Control

Human on the Loop



Monitoring

Human out of the Loop



Fully Autonomous

防御側

?

攻撃側

※イラストは生成AIで作成

④AI vs AIはスピードの勝負なのか？

- ある大手小売企業でAI SOCのトークン超過額が50万ドルに
- 軽量モデル／フロンティアモデルの切り替え（LLMルーティング）のソリューションをPR

ただし

- 少なくとも防御側が計算コストを余分に追う
- OWASP LLM10:2025 Unbounded Consumptionで新たな攻撃面としてアラート
- AI SOCをToken Bombで梅雨払いさせることもありうる？

RESOURCE CENTER > BLOG
AI Tokenomics: Why It Doesn't Work for Security, and What Does



MATT GARCIA | 22 JUNE 2026

AGENTIC AI

AI SOC



出典：Relia Quest

④AI vs AIはスピードの勝負なのか？



出典：Schneier
On Security

Cybersecurity in the Age of Instant Software

AI is rapidly changing how software is written, deployed, and used. Trends point to a future where AIs can write custom software quickly and easily: "instant software." Taken to an extreme, it might become easier for a user to have an AI write an application on demand—a spreadsheet, for example—and delete it when you're done using it than to buy one commercially. Future systems could include a mix: both traditional long-term software and ephemeral instant software that is constantly being written, deployed, modified, and deleted.

AI is changing cybersecurity as well. In particular, AI systems are getting better at finding and patching vulnerabilities in code. This has implications for both attackers and defenders, depending on the ways this and related technologies improve.

In this essay, I want to take an optimistic view of AI's progress, and to speculate what AI-dominated cybersecurity in an age of instant software might look like. There are a number of unknowns that will factor into how the arms race between attacker and defender might play out.

- 前提として、バイブコーディングと脆弱性の探索のAIを組み合わせ「ほとんど脆弱性がないソフトウェア」を実現化
- Instant Software
 - バイブコーディングによるプロプライエタリかつオンデマンドなソフトウェア
 - 使い終わったら削除
 - 入手が難しく、リバエンにハードル
 - 多様化により攻撃側の経済合理性が下がる
- Self-Healing Network
 - AIエージェントの常時スキャンおよびパッチの自動作成
 - すなわち脆弱性対応をユーザー側に移行（！）
 - 多様化により攻撃側の経済合理性が下がる
 - 防御側は一定の情報共有やインセンティブ化により経済合理性を上げられる
 - もちろん保険によるカバーなど様々なバックアップが必要不可欠

④AI vs AIはスピードの勝負なのか？

Resilient Software System

“We try to mitigate the damage and patch them and round around. We go on this merry go round. And I think the DARPA impact is going to be changing that mindset.”

- 様々なフォーマルメソッド（＝形式手法＝脆弱性の不存在を“数学的に”証明）の一連の体系
- Isabell/HOLなど様々な証明器の実装
- 約15年に渡る研究後、昨年ようやく空軍で実証段階に
- AI × CC → LLMによるスケール化
- SLCのどの段階でも適用可能
- ある意味ではKEVのアンチテーゼ
- Amazon、Google、MSで一部実装

Mythos時代を正しく
怖がるための広範な情報収集を