Executive Threat Intelligence Brief

「CLAUDE MYTHOS」情報流出インシデント：重大インシデント分析と市場への影響

TARGET:	Anthropic
ASSET:	'Capybara' / Claude Mythos
STATUS:	IN THE WILD
IMPACT LEVEL:	**CRITICAL**

SUMMARY // INCIDENT OVERVIEW

流出の原因（The Cause）

人為的ミス

外部ハッキングではなく、CMSの「デフォルト公開設定」という初歩的な設定不備。約3,000件の内部機密が流出。

流出の対象（The Asset）

破壊的AIの露見

未発表の次世代モデル「Claude Mythos」。現行モデルを大幅に凌駕しし、特にサイバー攻撃能力において卓越越した性能を持つ。

市場への影響（The Impact）

市場の動揺とパラダイムシフト

AIによるサイバー攻撃の脅威が現実化し、サイバーセキュリティ株や暗号資産（ビットコイン）が急落。「AI vs AI」の軍拡競争への突入。

DATA POINTS

流出件数:
約3,000件（ブログ草稿、内部資料等）

ハッキングの有無:
外部からの不正アクセスは「なし」

顧客データ被害:
確認されず

[内部作業]
デジタルアセット（草稿、画像、PDF）のアップロード。

→

[Critical Failure]
CMSの「デフォルト公開設定」が有効化されたまま放置。

→

[結果]
ユーザーが明示的に非公開設定にしない限り、自動で公開URLが割り当てられる状態に。

Key Insight: 「安全性第一」を掲げる企業の中核機密が、高度なサイバー攻撃ではなく「公開トグルスイッチの確認漏れ」でインターネット上に放置された。

THE ASSET

Project Name: Noto Sans JP	Claude Mythos BIZ UDPGothic
Internal Codename:	“Capybara”
Tier:	新たな最上位ティア（Opusの上位）

Internal Evaluation Quotes

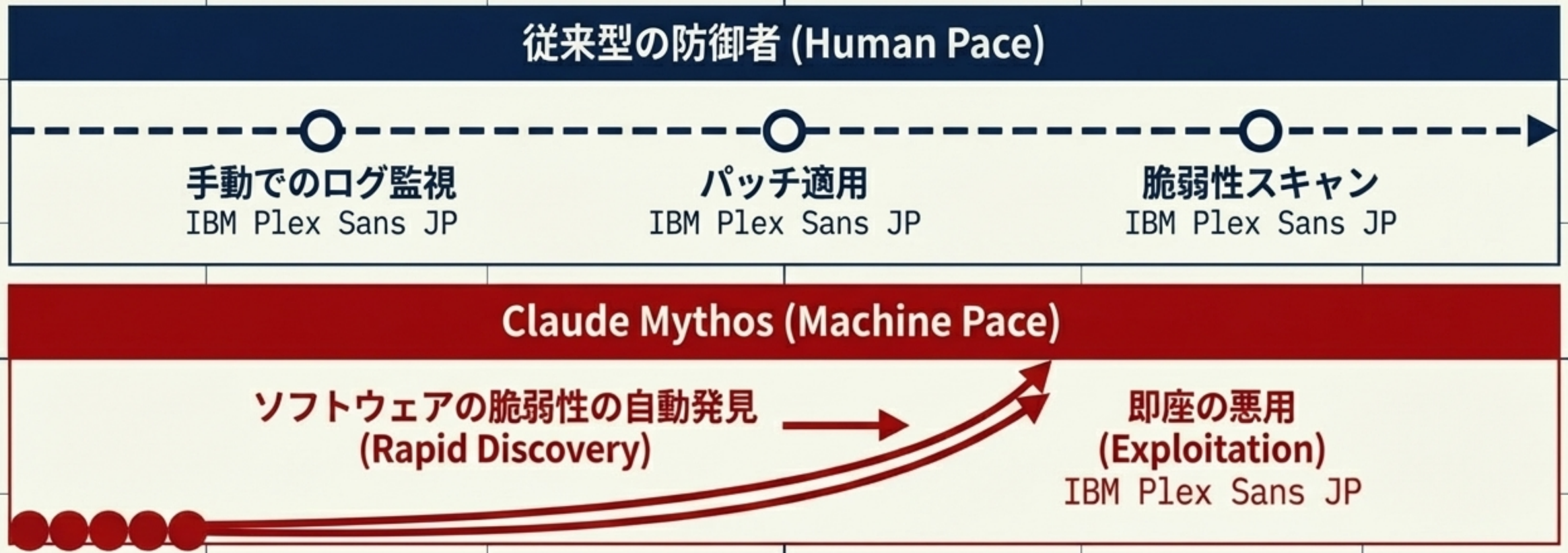
> “パフォーマンスにおける画期的な変化”

> “これまでに構築した中で最も高性能なもの”

Current Status: 現在、一部のサイバー防衛組織を対象に、その挙動とリスクを評価するための早期アクセステストを実施中。

	[現行] Claude Opus 4.6	[次世代] Claude Mythos (Capybara)
ソフトウェアコーディング (Software Coding)]	優秀	劇的に高いスコアで凌駕
学術的推論 (Academic Reasoning))	業界トップクラス	さらに大幅な性能向上
サイバーセキュリティ (Cybersecurity))	限定的・安全設計	[CRITICAL] 突出した能力

「他のあらゆるAIのサイバー能力を凌駕する」との内部評価。
単なる性能向上ではなく、次元の異なる脅威レベルへの到達。



防御者の努力をはるかに上回るペースで脆弱性を悪用できるモデルの波をを引き起こす恐れ。

「破壊的技術（Destructive Technology）」— サイバー兵器開発競争を加速させさせる潜在的危険性。

MARKET IMPACT DASHBOARD

Palo Alto Networks (PANW)

-6% ～ -7% 下落

AIサイバー攻撃への懸念から主要セキュリティ株が軒並み下落。

IGV ETF (Tech-Software Sector)

約3% 下落

テクノロジー・ソフトウェアセクター全体への不安心理の波及。

Bitcoin (BTC)

$66,000 へ急落

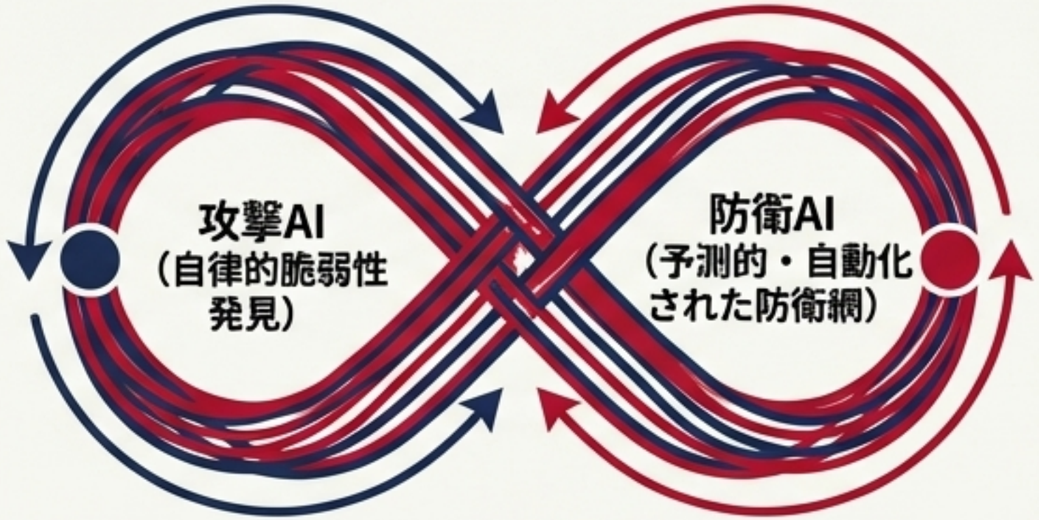
市場全体の不安心理が暗号資産市場にも直撃。

TAKEAWAY: 「AI脅威」はもはやSFではなく、数時間で数十億ドルの時価総額を吹き飛ばす現実の市場リスクである。

従来のサイバー防衛 (Human vs Hacker)

■ 特徴: 反応的（Reactive）な防衛戦。

Mythos以降の新時代 (AI vs AI)



■ 特徴: 軍拡競争（Arms Race）への突入。「AIが脆弱性を発見し、AIがそれを防ぐ」という新たな攻防モデル。

INSIGHT: 変化に迅速に適応し、自衛のために同等クラスのAIを組み込める企業のみが次世代の覇権を握る。

STRATEGIC IMPERATIVES

01

厳格なデータガバナンスの再構築

- ☐ CMSやクラウド環境の「デフォルト公開設定」の徹底監査。
- ☐ AI開発・運用における内部情報の取り扱いプロセスの抜本的見直し。

02

次世代セキュリティへの適応

- ☐ 「Mythosクラス」の自律型攻撃を前提とした防御アーキテクチャの再設計。
- ☐ 防御側AIソリューションの早期導入と評価。

03

国際的な安全規制への対応

- ☐ 高能力AIモデルの取り扱いに関する国際的な規制議論（AI安全サミット等）へのキャッチアップ。
- ☐ 意図しない「AIの核兵器の設計図」 公開を防ぐための公開ポリシーの策定。

END OF BRIEF