Security for AI
AI for Security

Splunk で守る AI エージェントと、AI で加速する SOC

Splunk Services Japan
パートナー＆ソリューションアーキテクト技術本部　本部長
横田　聡 CISSP GMON

2026.07.10

Forward-looking statements

This presentation may contain forward-looking statements regarding future events, plans or the expected financial performance of our company, including our expectations regarding our products, technology, strategy, customers, markets, acquisitions and investments. These statements reflect management's current expectations, estimates and assumptions based on the information currently available to us. These forward-looking statements are not guarantees of future performance and involve significant risks, uncertainties and other factors that may cause our actual results, performance or achievements to be materially different from results, performance or achievements expressed or implied by the forward-looking statements contained in this presentation. For additional information about factors that could cause actual results to differ materially from those described in the forward-looking statements made in this presentation, please refer to Cisco's periodic reports its other filings with the SEC, including the risk factors identified in Cisco's most recent quarterly report on Form 10-Q and annual report on Form 10-K, copies of which may be obtained by visiting the Cisco Investor Relations website at investor.cisco.com or the SEC's website at www.sec.gov. Any projections in this presentation are based on information currently available to Cisco, which is subject to change. The forward-looking statements made in this presentation are made as of the time and date of this presentation. If reviewed after the initial presentation, even if made available by us, on our website or otherwise, it may not contain current or accurate information. We disclaim any obligation to update or revise any forward-looking statement based on new information, future events or otherwise, except as required by applicable law.

In addition, any information about our roadmap outlines or our general product direction is subject to change at any time without notice. It is for informational purposes only and shall not be incorporated into any contract or other commitment. We undertake no obligation either to develop the features or functionalities described, in alpha or beta or in preview (used interchangeably), or to include any such feature or functionality in a future release.

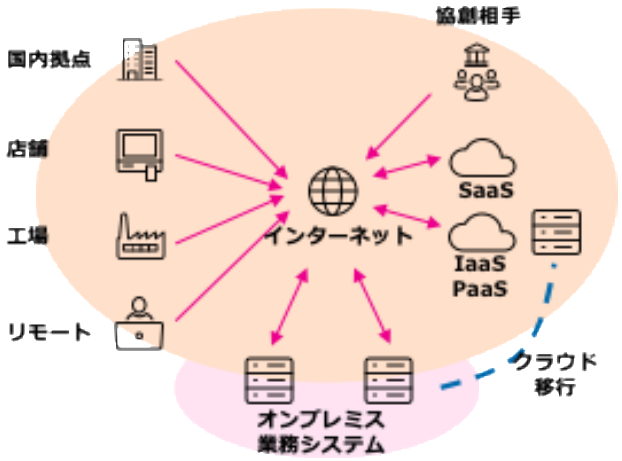
アジェンダ

- フロンティアAI時代の脅威と課題
- AI AgentでSOCを加速する
- 組織のAI活用を守る

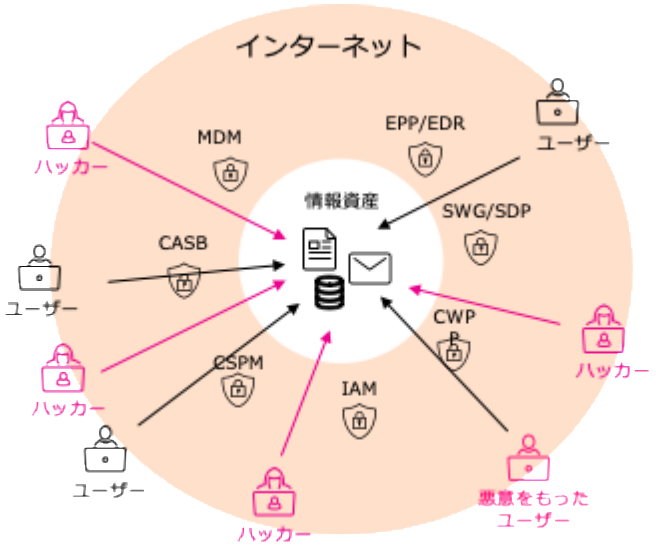
テクノロジーの変化によるリスクの変遷

2027年にかけてAIがサイバー攻撃の脅威を質・量ともに増大させると予測（英国国家サイバーセキュリティセンター（NCSC））。防御側においてはAI活用の有無によるデジタル格差が懸念されている。 出展：「情報セキュリティ10大脅威2026」解説書［組織編］

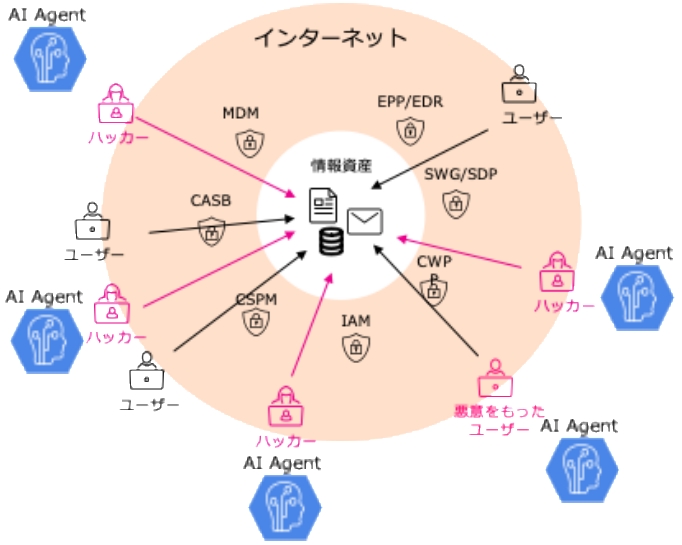
DX投資の加速 → **アタックサーフェス/リスクの拡大** → **フロンティアAIによるインシデントサイクルの加速**



境界型からオープンなネットワークへのシフト



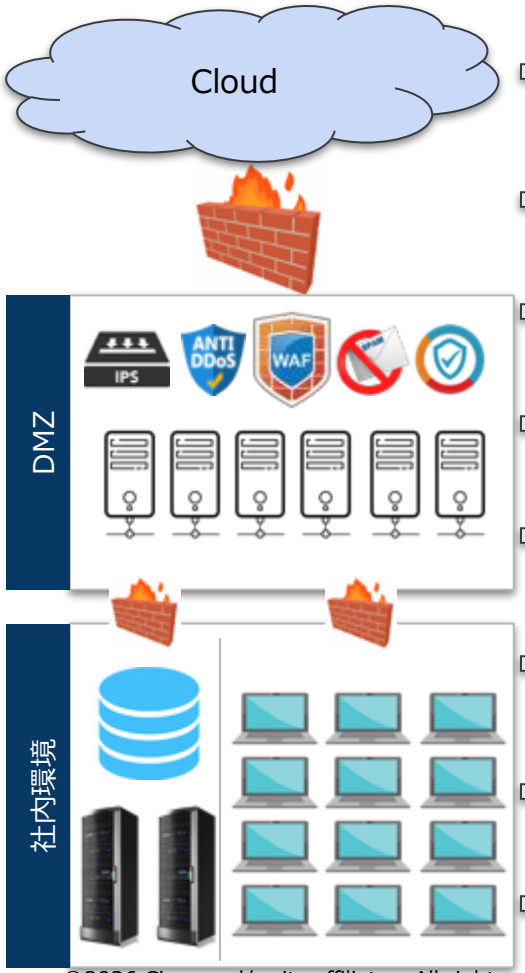
自社のみならずサプライチェーンのリスクマネジメントまで



- 攻撃者：脆弱性探索、マルウェア生成、侵入拡大の自動化
- 組織内：AI Agentによる意図しない情報流出・操作リスクの増大

基本のき：必要なログは対応するリスクによって異なる

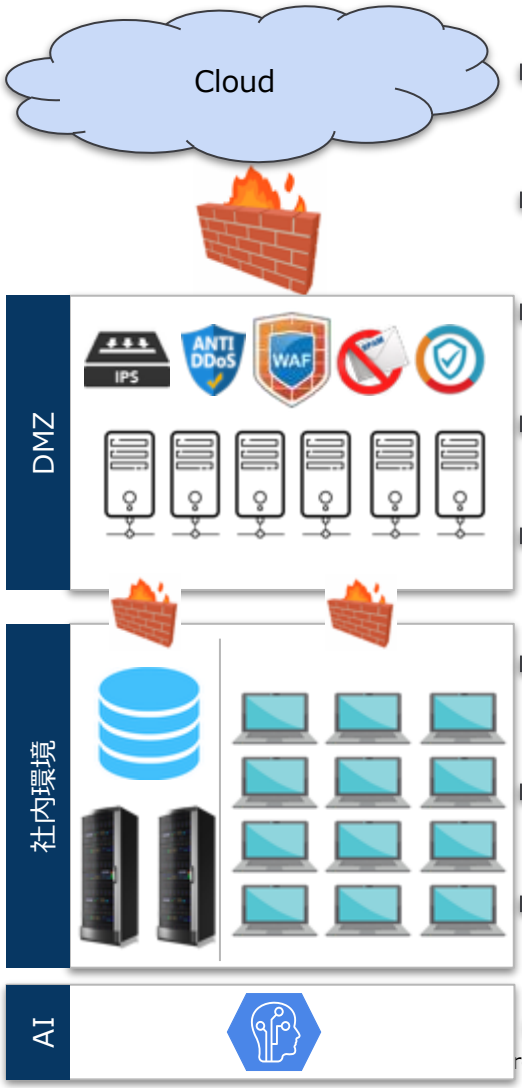
システムに潜在的に存在する脆弱性や脅威から考えられるリスクを理解し、それらのリスクがどこに存在しているかを把握しなければならない。このため、**該当するシステムのデータソースを活用**することで問題の事前・事後調査が可能になり、可視化をすることで**顕在化する前に食い止める**ことができる



存在するリスク	想定される被害	対象システム	必要データソース
クラウドリスク	情報漏えい、仮想環境悪用	public Cloud	SaaS,CASB,CSPM,Azure Audit,O365,AAD等,利用しているクラウドアクセスログ
サービス停止	Web停止による社会的影響	公開サーバ群	Neflow/Stream、サーバアクセスログ、 セキュリティデバイスログ
アカウント悪用	アカウント情報漏えい、特権アクセスによる情報漏えい	認証システム	AD,Web,SSH,FTP,SSO,MFA,VPN等の各種認証ログ
脆弱性悪用	サーバへの恒久的なのっとり	公開サーバ群	脆弱性診断結果ログ,Webアプリケーション脆弱性結果,パッチ適用有無情報
標的型攻撃	Ransomware等による金銭的な脅し、情報漏えい	従業員利用システム	EDRのログ及びアラート,Windows Eventlog,Sysmon,PowerShell等
情報漏えい	社会的信頼の損失、損害賠償の発生	サーバ、データベースシステム	EDR,DNS,サーバアクセスログ,データベースauditログ,Netflow/Stream,DLP等
内部不正	企業内機微情報の他社への移転	従業員利用システム	標的型攻撃と同等に加え、Printer,USB,入退館室ログ
コンプライアンス	コンプラ違反による罰金	対象領域	対応するコンプライアンス内容に準じる

AIエージェントの普及により再定義が必要なリスク・監視データ

AIエージェント攻撃の出現により、既存リスクの速度・規模が増大し、監視すべきデータソースが拡張される。
主たるリスクカテゴリは従来とほぼ変わらないが、新たにAIガバナンスリスクの可視化が求められる。

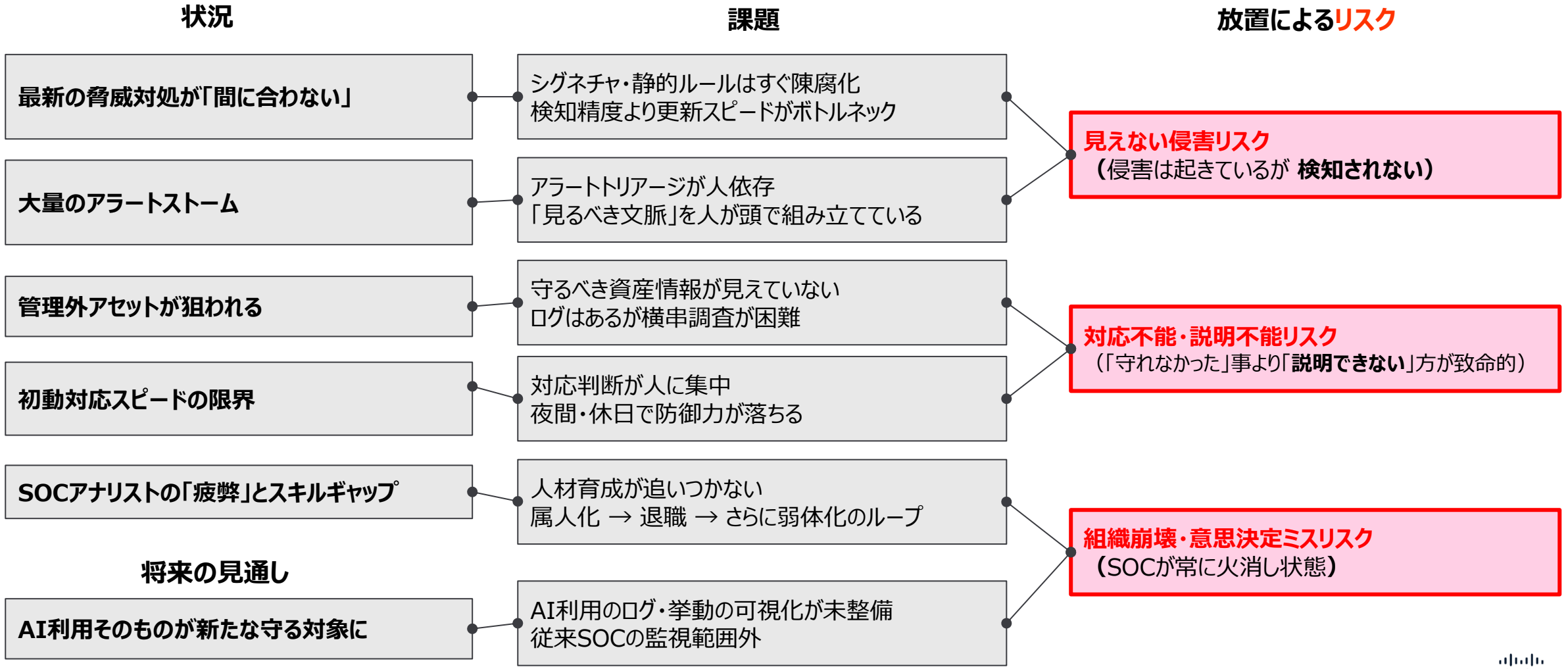


存在するリスク	想定される被害	対象システム	必要データソース
クラウドリスク	情報漏えい、仮想環境悪用	public Cloud SaaS AI サービス	SaaS,CASB,CSPM,Azure Audit,O365,AAD等, 利用しているクラウドアクセスログ SaaS AI利用ログ, AI DLP
サービス停止	Web停止による社会的影響	公開サーバ群	Neflow/Stream、サーバアクセスログ、 セキュリティデバイスログ
アカウント悪用	アカウント情報漏えい、特権アクセスによる情報漏えい	認証システム	AD,Web,SSH,FTP,SSO,MFA,VPN等の各種認証ログ、NHI(Non Human Identity)認証ログ
脆弱性悪用	サーバへの恒久的なのっとり	公開サーバ群	脆弱性診断結果ログ,Webアプリケーション脆弱性結果,パッチ適用有無情報
標的型攻撃	Ransomware等による金銭的な脅し、情報漏えい	従業員利用システム	EDRのログ及びアラート,Windows Eventlog,Sysmon,PowerShell等
情報漏えい	社会的信頼の損失、損害賠償の発生	サーバ、データベースシステム SaaS AIサービス	EDR,DNS,サーバアクセスログ,データベースauditログ,Netflow/Stream,DLP等
内部不正	企業内機微情報の他社への移転	従業員利用 システム	標的型攻撃と同等に加え、 Printer,USB,入退館室ログ
コンプライアンス	コンプラ違反による罰金	対象領域	対応するコンプライアンス内容に準じる AI意思決定記録、AI利用ポリシー準拠ログ
AIガバナンス リスク	シャドーAI、モデルポイズニング、AIエージェントの権限過剰/暴走リスク、AIドリフトなど	AI/LLMシステム SaaS AI全般	（上記赤字＋） AIモデル完全性検証ログ、エージェント行動監査

AI AgentでSOCを加速する

攻撃者がAIを用いる時代のSOC運用課題

攻撃側がAIを活用してくる事で、SOCの課題は量・速さ・質のすべてにおいて対応の難易度が増しています



SplunkのAgentic SOCによる解決

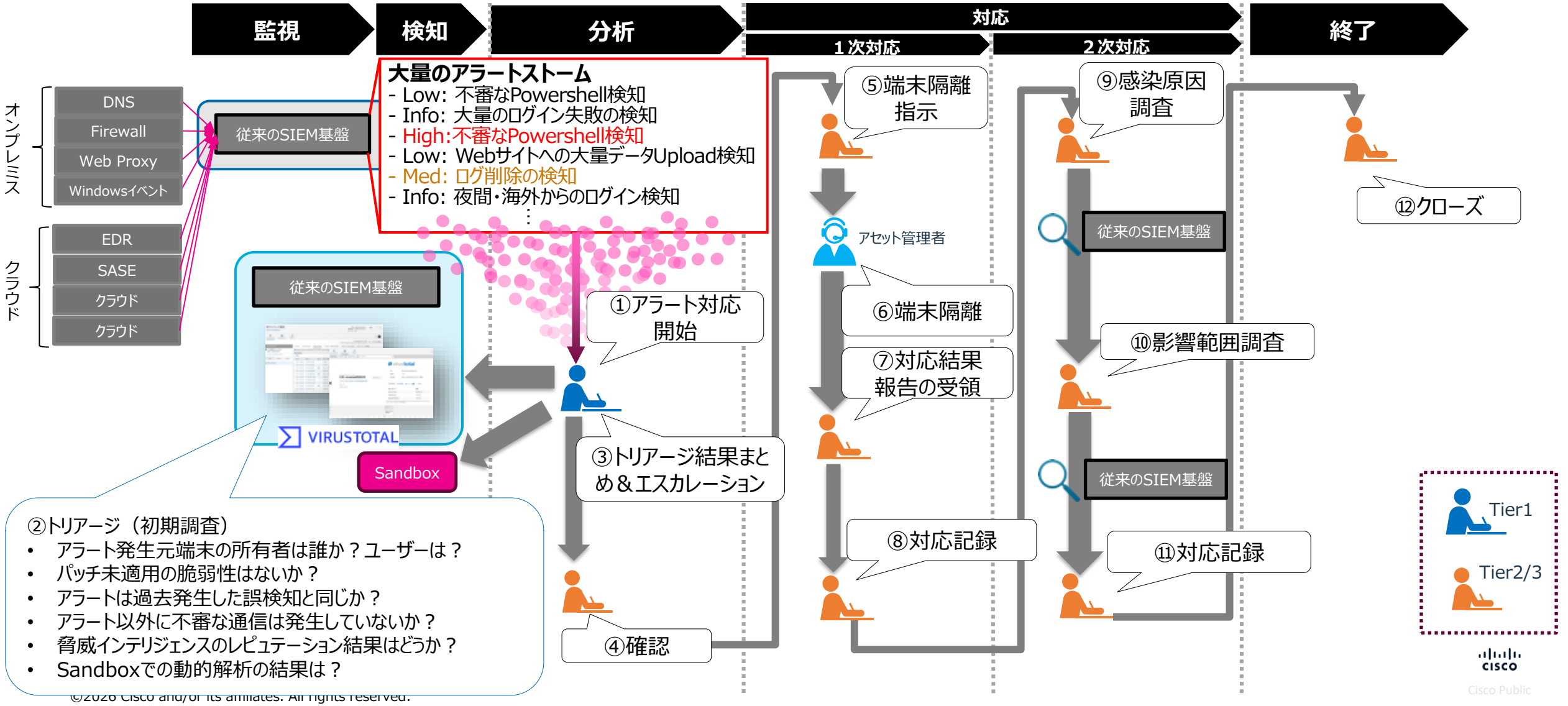
CiscoとSplunkが提供するHuman in the LoopのAIソリューション

状況	課題	ソリューション
最新の脅威対処が「間に合わない」	シグネチャ・静的ルールはすぐ陳腐化 検知精度より更新スピードがボトルネック	**AI**による検知ルールの推奨とチューニング提案
大量のアラートストーム	アラートトリアージが人依存 「見るべき文脈」を人が頭で組み立てている	**AI**によるアラートトリアージの自動化
管理外アセットが狙われる	守るべき資産情報が見えていない ログはあるが横串調査が困難	Splunkの柔軟な正規化とアセットリスク管理支援
初動対応スピードの限界	対応判断が人に集中 夜間・休日で防御力が落ちる	**AI**による推奨対応手順の提示、自動化支援
SOCアナリストの「疲弊」とスキルギャップ	人材育成が追いつかない 属人化 → 退職 → さらに弱体化のループ	Splunkによるセキュリティ人材育成コンテンツのデリバリー
将来の見通し		
AI利用そのものが新たな守る対象に	AI利用のログ・挙動の可視化が未整備 従来SOCの監視範囲外	Agent Observability(Galileo), Cisco AI Defense, SSE, Duo等との連携によるAIセキュリティガバナンス・監視の実現支援

Before

従来のSOC運用イメージ

大量のセキュリティアラートをSIEMを使って1件ずつ対応する従来のSOC運用イメージ



②トリアージ（初期調査）
- アラート発生元端末の所有者は誰か？ユーザーは？
- パッチ未適用の脆弱性はないか？
- アラートは過去発生した誤検知と同じか？
- アラート以外に不審な通信は発生していないか？
- 脅威インテリジェンスのレピュテーション結果はどうか？
- Sandboxでの動的解析の結果は？

After

RSAC2026後

SplunkのAgentic SOC運用イメージ

Tier1オペレーターの反復作業をAgentic AIがサポートし、スキルギャップを埋めて工数負担を下げる

監視 | 検知 | 分析 | 対応（1次対応 / 2次対応） | 終了

オンプレミス: DNS / Firewall / Web Proxy / Windowsイベント

クラウド: EDR / SASE / クラウド / クラウド

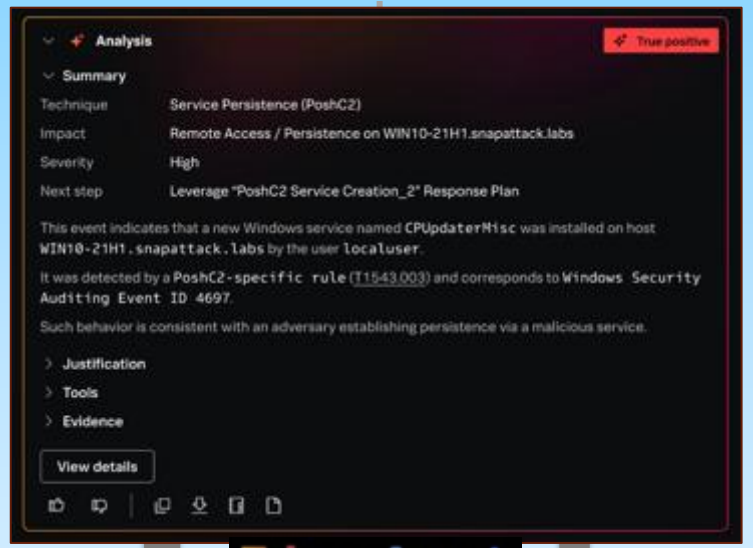
Splunk Enterprise Security™

TALOS

ML AIによる振る舞い検知

①マシン/ユーザーの通常と異なる挙動を検知

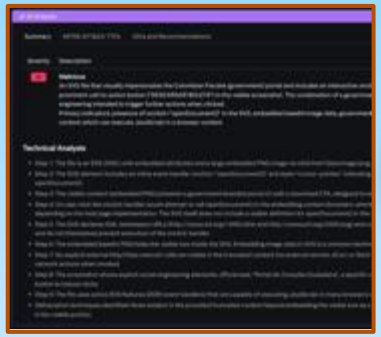
トリアージAIによる自動調査
推奨手順の提示



Triage Agent

②大量アラートから要注意アラートのみトリアージ
③アラートサマリーと推奨アクションを生成

AIによるURL/検体の
動作解析サマリーの提示



Malware Threat Reversing Agent

AIによる
調査レポート作成

AI Assistant in Enterprise Security

Tier1

Tier2/3

④トリアージ結果の確認

⑤推奨手順に
従った追加調査と対応

⑥分析レポートを元に
影響範囲の調査

⑦AIの作成した調査レ
ポートを確認しクローズ

Before

SIEM検知エンジニアリングの継続的改善活動イメージ

一度に大量の検知ルールを実装するのではなく、特定の脅威シナリオに基づく検知ルールを段階的に導入することを繰り返す

1. 自組織に関連する脅威シナリオの調査

- 組織に対して発生し得る具体的な攻撃の例として、脅威シナリオを調査する
- 脅威シナリオの調査例
  ・自組織で利用しているSaaSや、製品に関連する攻撃の情報を入手
  参考：Splunk Blogリンク
  ・セキュリティベンダーの調査レポートを入手
  参考：Kaspersky ガイドリンク
  ・戦術的・運用上の脅威インテリジェンスの入手
  ・セキュリティコンテンツのアップデート情報の入手
  参考：ESCU Release notesリンク

2. 実装対象の検知ルールの検討・選定

- 脅威シナリオの調査結果に基づき、実装を行う検知ルールを検討する
- Splunkセキュリティコンテンツの SIEM検知ルールを選定
  ・Splunk SSE 、またはSplunk ESの使用事例ライブラリを利用
  参考：Qiita記事リンク
- 検知ルールの実装に必要なデータソースも合わせて検討
- セキュリティコンテンツの利用が難しい場合はカスタムでのサーチ作成

3. 検知ルールの設定作成

- Splunkセキュリティコンテンツは、CIM化されたデータモデルのフィールド名でサーチ文が作成されている。そのため、実環境のデータに合わせてサーチ文を変更
- その他、検知ルール設定項目を作成
- Splunk ESのRBAを利用する場合は、検知ロジックを作成
- アラート発生後の調査手順が決まっている場合は関連設定を事前に実装できるが、後のフェーズでも検知ルールに追加設定が可能

4. 検知ルール設定後のチューニング

- サーチ文の更新や、ルックアップテーブルファイルを作成し、過検知・誤検知が発生しているイベントの抑止
- アラート発生状況を確認し、サーチ文中の閾値を調整
- 実環境でのスケジュールサーチの実行時間を確認し、サーチ実行間隔の調整
- Splunk ESのインシデントレビューのNotableイベント表示項目の追加

5. アラート発生後の調査関連設定の整備

- Splunk ESのNotableイベントで、アラート発生後の調査を行うためにドリルダウンサーチを追加
- Splunk ESの次のステップの箇所に、Notableイベントの種類ごとの対応手順を追加
- Splunk ESの調査管理ワークベンチへのNotableイベント連携に関連する設定追加
- Notableイベント発生後の適応応答アクションを追加
- 必要があれば、ダッシュボードを作成

After

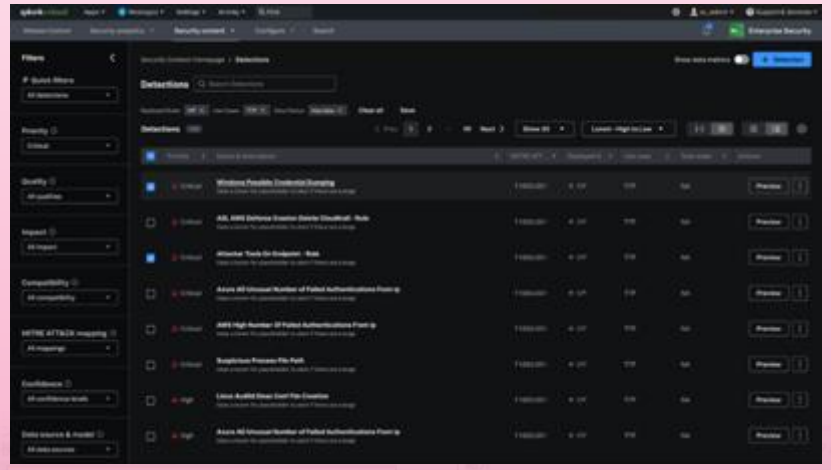
RSAC2026後

AIがサポートする検知エンジニアリングのイメージ

AIがサポートする事で、新しい脅威に対する検知ルールのアップデートの負担を下げて、常に最新の脅威に対応し続ける事が可能に

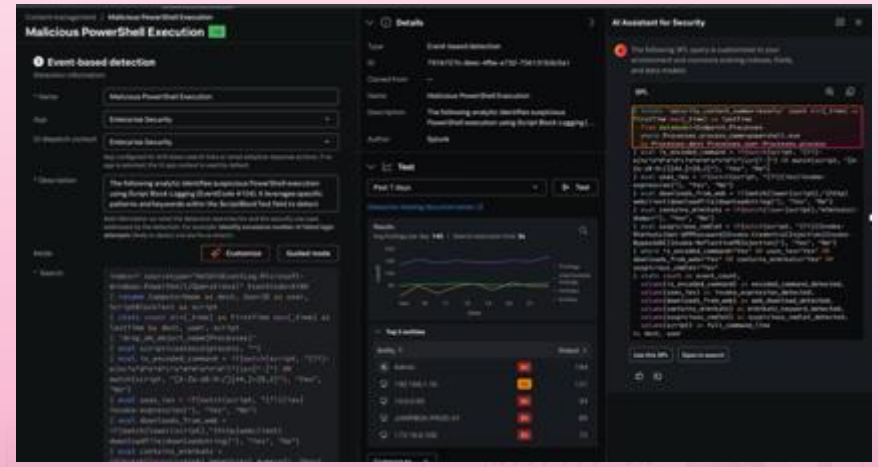
1. 自組織に関連する脅威シナリオの調査
2. 実装対象の検知ルールの検討・選定
3. 検知ルールの設定作成
4. 検知ルール設定後のチューニング
5. アラート発生後の調査関連設定の整備

自社のデータに合わせて、検知ルールをレコメンデーション



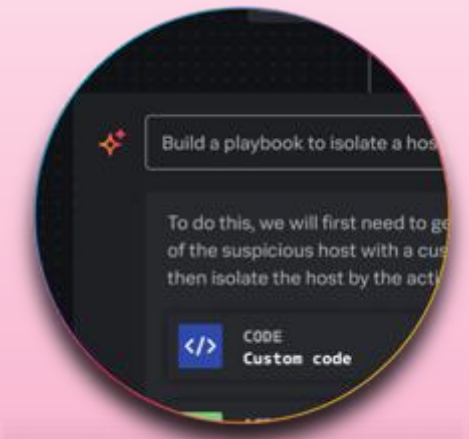
Detection Studio

検知ルール(SPL)の最適化をAIがサポート



Detection Builder Agent

自然言語を用いてアラート対応の手順化をサポート



Automation Builder Agent

Demo
Triage Agent

組織のAI活用を守る

AI活用における課題

5つの立場、同じ3つの悩み

CISO
「全社でどこにAIリスクがあるのか見えない。野良AI、AIアプリ、ガードレイル違反がバラバラに検知され、経営に説明できる全体像がない」

CFO
「AIのトークン/API課金や利用量が部門・アプリ単位で見えず、コスト増の原因や投資対効果を説明できない」

リスク・コンプライアンス企画
「EU AI Actや社内AIポリシーへの準拠を求められるが、誰が・いつ・どのデータをAIに渡したかの証跡が残っていない」

SOC
「野良AI利用やガードレイル違反のアラートは来るが、重大度・対応先・業務影響を判断できない」

AI App開発
「本番AIエージェントがなぜ誤回答・暴走・危険なツール実行をしたのか追えない。セキュアに作れと言われても、検証と運用監視が手作業になる」

見えない
Shadow AI・ 品質/コスト

守れない
インジェクション・データ漏洩・過剰権限

証明できない
監査証跡・規制対応

Agentic AI セキュリティフレームワーク (フロー視点)

AI エージェントを安全に運用するための 6+1 ステップ

発見・認証

1 エージェント検出

AGENT DISCOVERY

リポジトリ／エンドポイント／3rd パーティ

2 ポスチャーチェック

POSTURE CHECK

スキルスキャナー・モデル来歴の確認

3 認証

AUTHENTICATE

エージェントに新しい ID を付与する

アクション制御 Action Control

4 認可

AUTHORIZATION

スコープを縮小（自己減衰トークン）

5 セマンティック検査

SEMANTIC INSPECTION

プロンプトや推論内容を検査する

6 コンテンツ検査

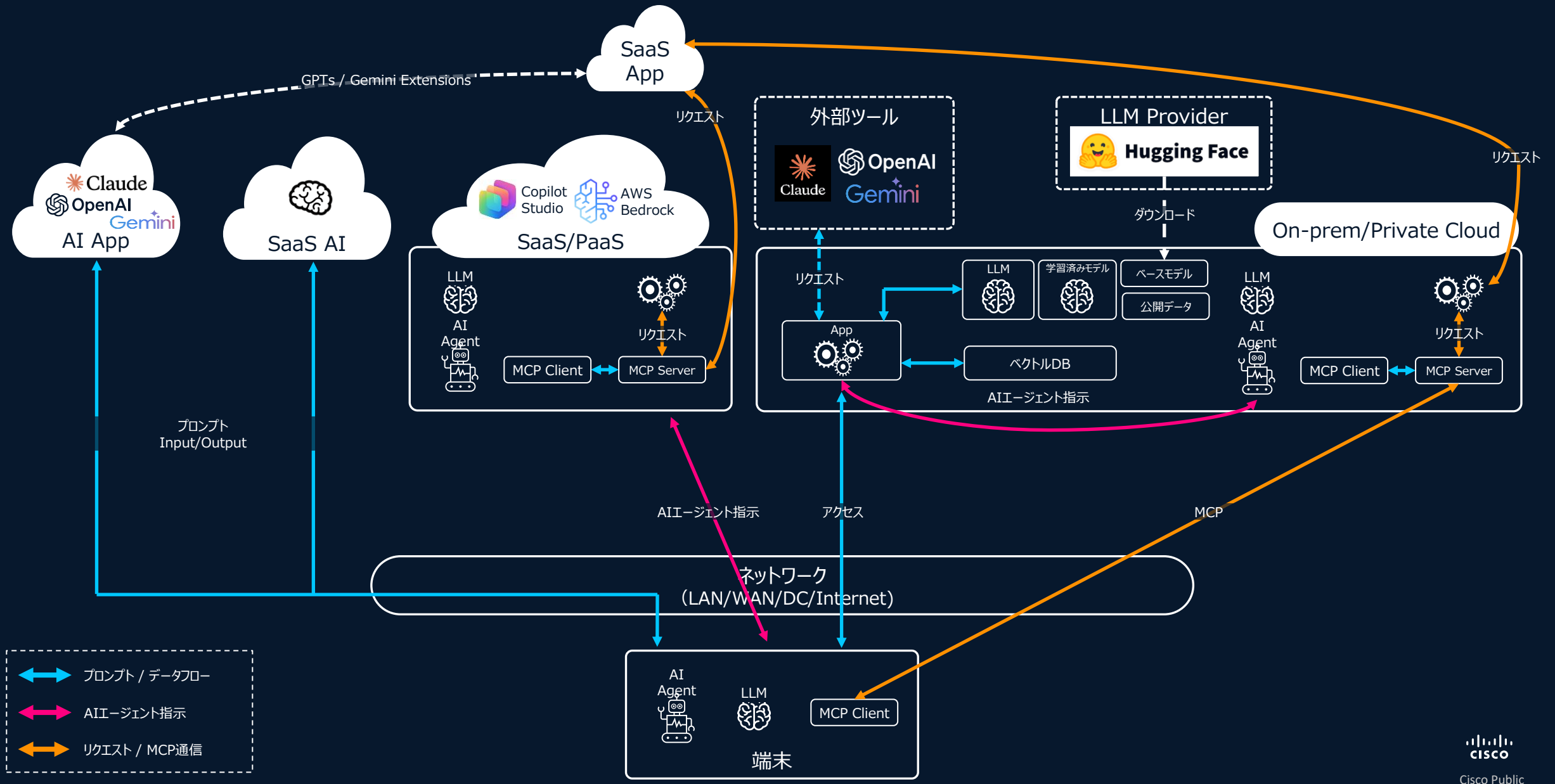
CONTENT INSPECTION

DLP・マルウェアの検査

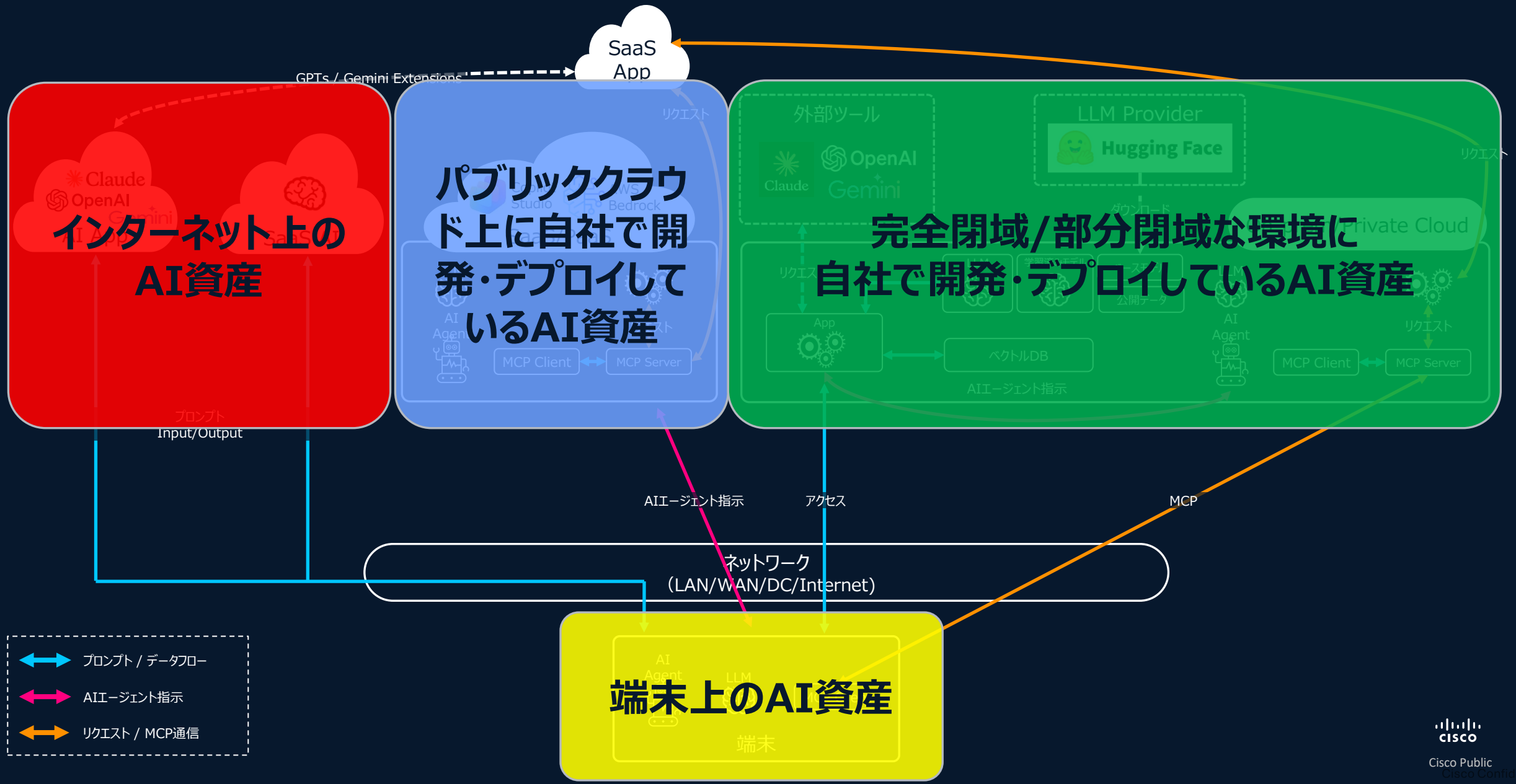
+1 エージェント評価 AGENT EVAL / TRACING

全工程を横断して監視・追跡（トレーシング）

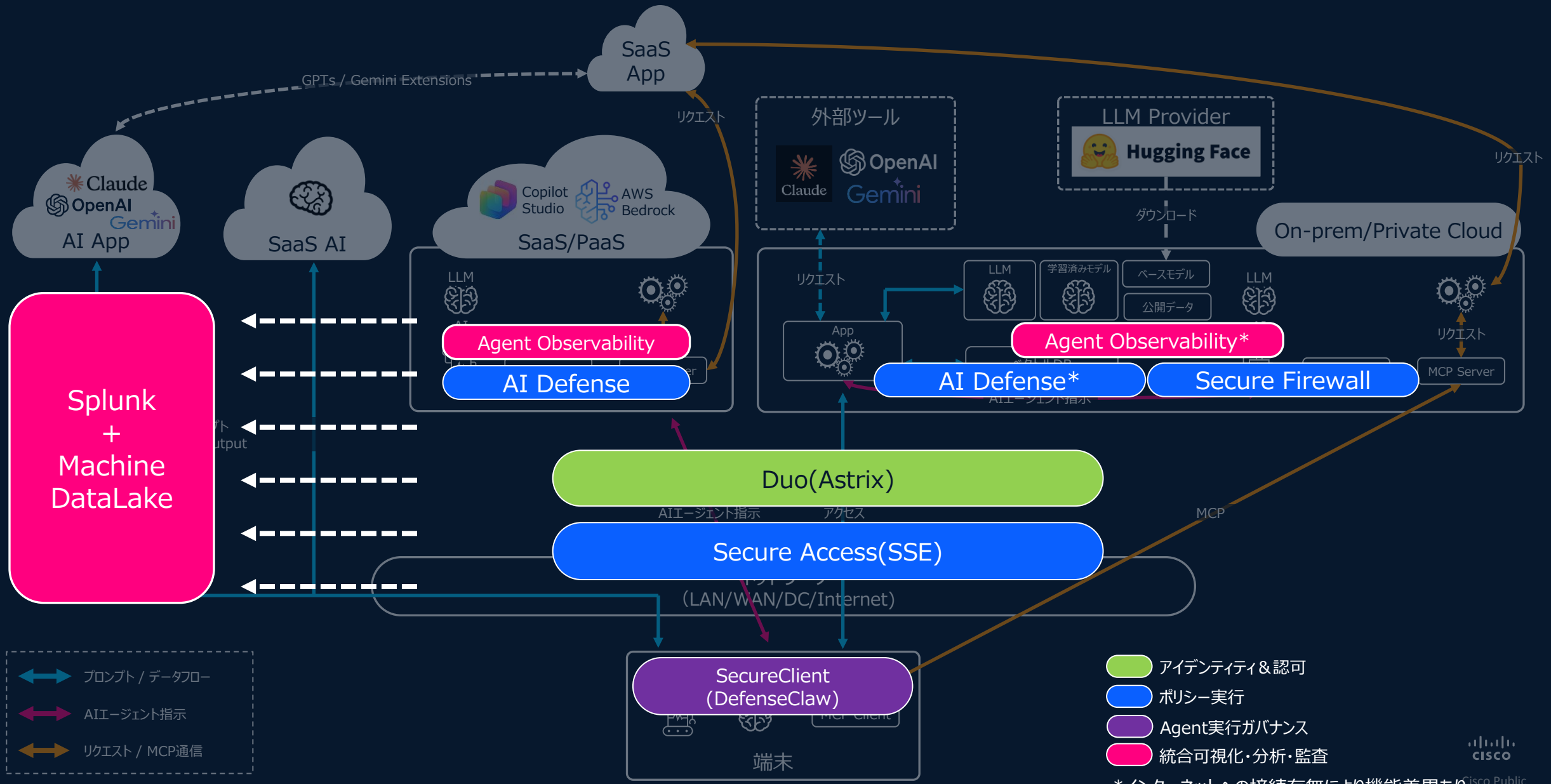
AI利活用ユースケース（全体像）



AI利活用ユースケース（全体像）



AI 利活用ユースケースに対する Ciscoのアプローチ



*インターネットへの接続有無により機能差異あり

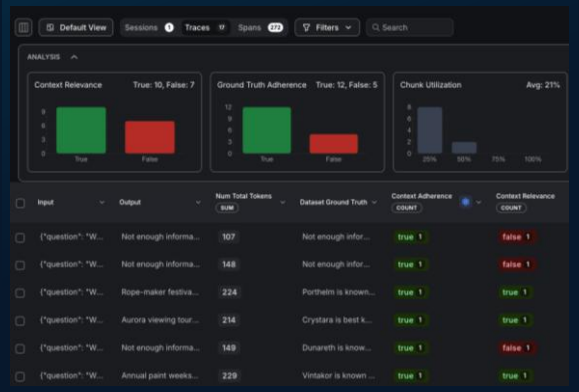
Agent Observability powered by Galileo

①評価

エージェントの挙動を改善する

エージェントとモデルを評価する

幻覚、バイアス、関連性、安全性に関するテストの提示文と回答例

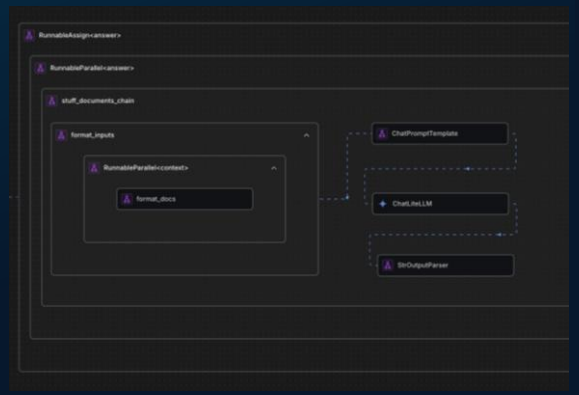


②観測

AIスタック全体を監視する

障害やエラーの診断

AIスタック全体にわたって、依存関係を追跡し、ワークフローを分析し、レイテンシやエラーなどのSLOを監視する



トークンの使用とコストを最適化する

*AI*コストの管理

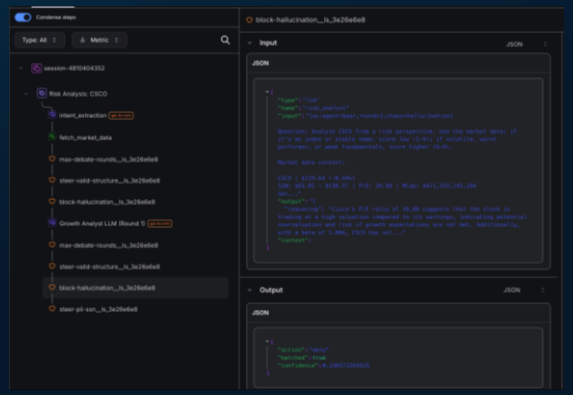
GPUやトークンの活用方法を理解し、コストを最適化する

③ガードレール

有害かつ不正確な出力をブロックする

有害な出力を防止する

リアルタイムのガードレール機能により、不正確または有害な出力を防止します



Splunk + Machine Data Lake

AI関連テレメトリを分析要件に合わせて適切なストレージにルーティング

大量のAI関連テレメトリ

- AI Defense
- Duo(Astrix)
- Secure Access(SSE)
- SecureClient (DefenseClaw)
- Agent Observability

AI関連テレメトリの
フィルタ、マスク、ルーティング

分析要件に合わせたデータ保管

AI Agent

MCP Client

ユーザー/MCP/エージェントによるログ検索

Splunk

・高速検索
・定期アラート監視
・ダッシュボード可視化 等

Splunk Index (Hot)

7～14日間保持

Machine Data Lake（MDL）

アラート後の
定型的な調査 等

MDL Lakehouse

90～180日間保持

多少検索に時間が
かかっても良いアドホックな調査

- IOCハンティング
- 監査 等

プロモート

プロモート

MDL Raw Storage
（低コストストレージ層によるRawデータの長期保存）

Splunk + MachineDataLake

Splunkを使ったAI活用状況の可視化

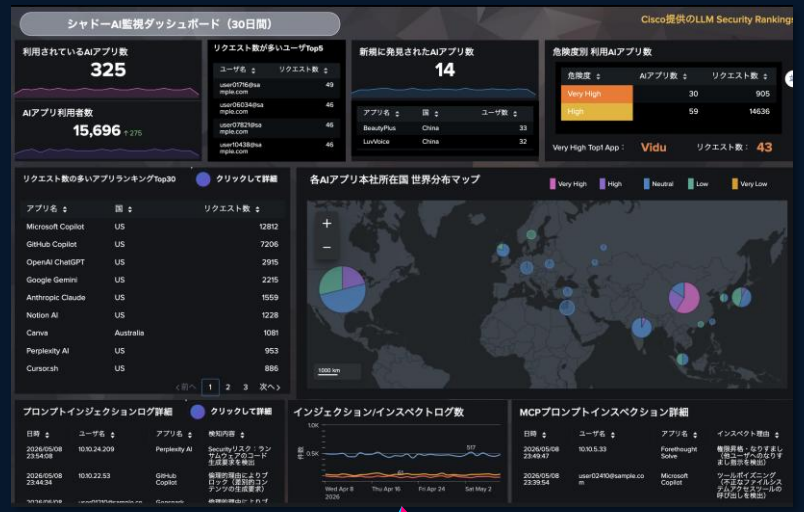
AI製品が多岐にわたる混迷期にこそオススメするSplunkによる可視化

CISO向け

用途：AIセキュリティ全体概要

セキュリティ部隊向け

用途：シャドーAI監視



DX推進部隊向け

用途：自社AI監視

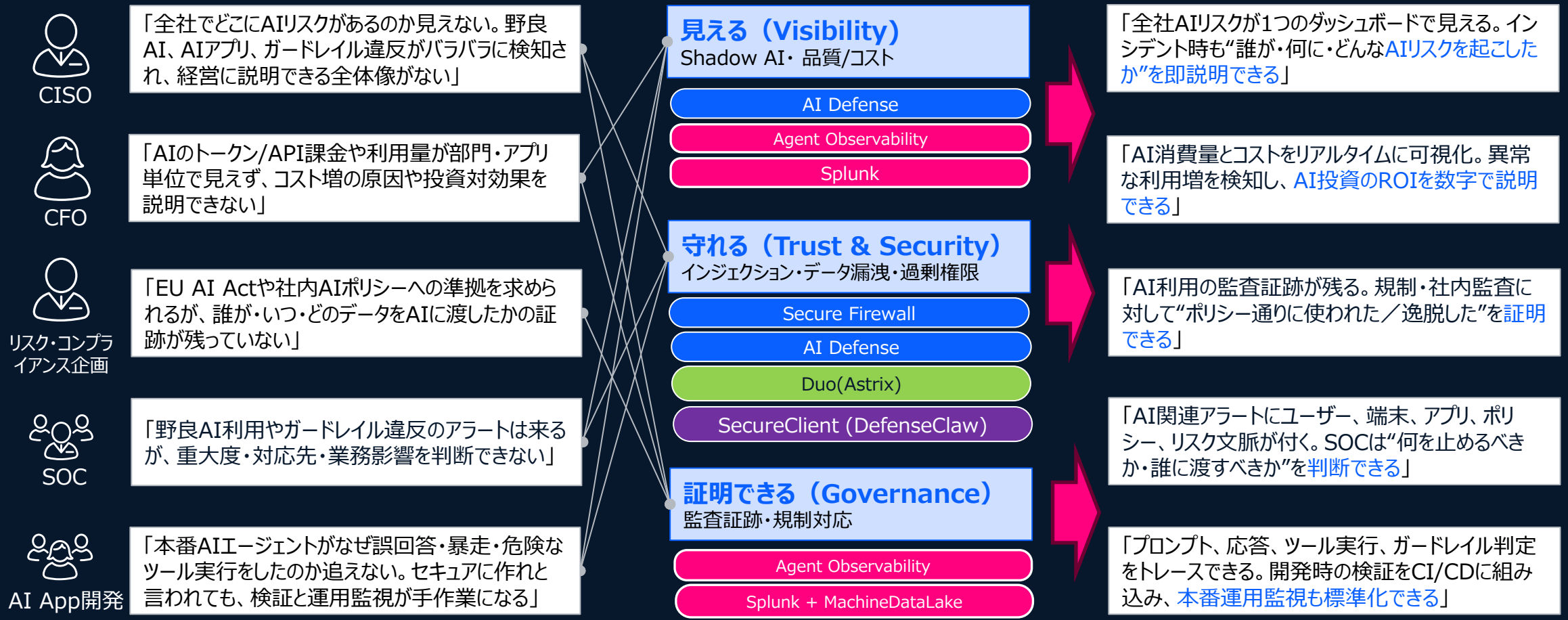
AI Defense

Secure Access(SSE)

Agent Observability

AI活用のコントロールが可能に

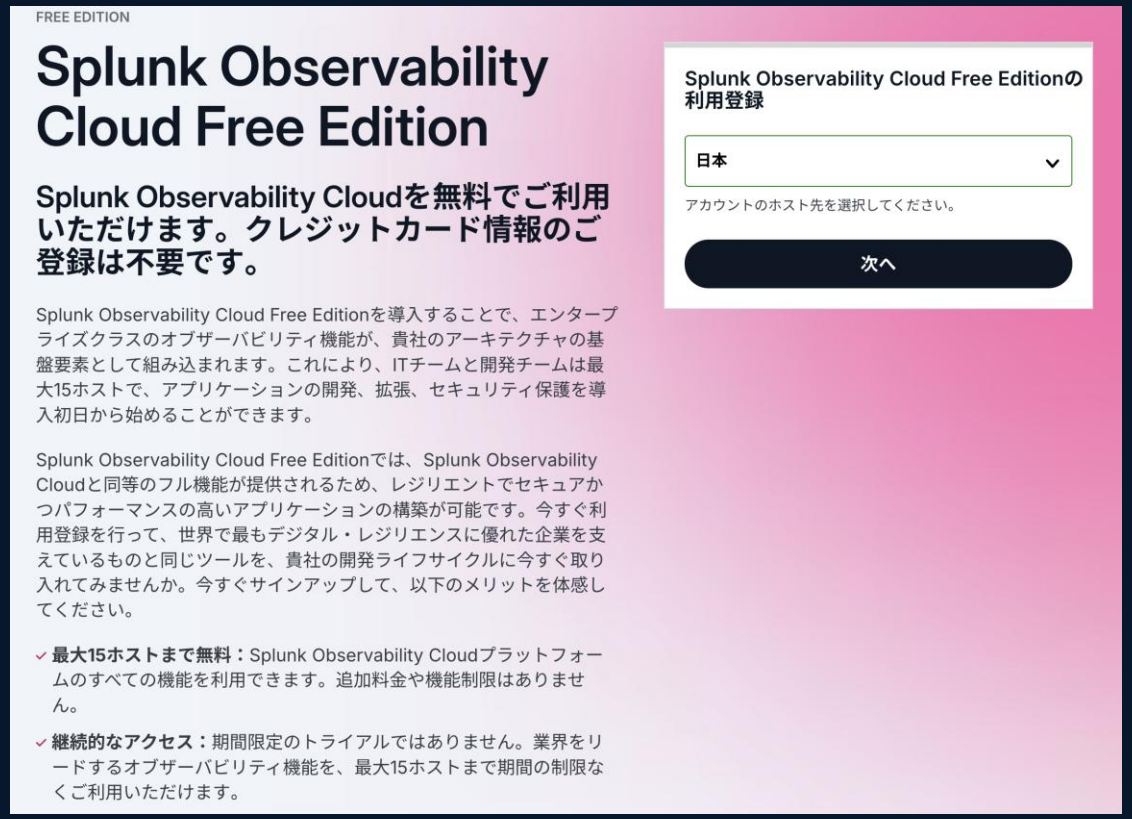
Cisco + Splunkにより、AIの利用・リスク・コスト・ガバナンスを横断的に可視化し、運用できる状態にする



Next Step

Agent Observability最初の一歩

① O11y Cloudの無期限トライアルで検証



https://www.splunk.com/ja_jp/download/observability-cloud-free-edition.html

② ZennのBlogコンテストに投稿してAmazonギフトをゲット！

https://zenn.dev/contests/splunk-opentelemetry-2026?tab=articles


CISCO