# 組織が優先すべき防御戦略

シスコシステムズ合同会社
セキュリティ事業
三澤　慶浩

# Talos 2025 Year in Reviewが示した傾向

両極端な脆弱性攻撃
迅速性と持続性

アイデンティティ
信頼の基盤としてのアーキテクチャ

影響拡大を狙った
集中管理システムへの攻撃

AI

多要素認証に対する攻撃

脆弱性

国家支援型アクター

ランサムウェア

# Talosが示した脅威に、Ciscoはどう答えるか

| Talos が観測した脅威 | | Cisco の答え |
|---|---|---|
| MFAスプレー攻撃　不正デバイス登録<br>178%増 | → | Duo + Identity Intelligence (ITDR)<br>アイデンティティ脅威の検知と対応 |
| ランサムウェア (Qilin) の横展開<br>月40+件 | → | Duo + Identity Intelligence (ITDR) ＋ISPM<br>デバイス健全性 + ポスチャ評価 |
| AIエージェントへの妨害<br>Prompt Injection等 | → | AI Defense<br>AIアプリ/エージェント向け防御 |
| 脆弱性ゼロデイ悪用の加速<br>猶予ほぼゼロ | → | SnortML on Firepower<br>AI/MLで未知の攻撃を検知 |

# 今、防御側に問われている2つの問い

**Security for AI**
組織が導入するAI（生成AI／AIエージェント）を
どう守るか？

**AI for Security**
攻撃者のAI活用スピードに、防御側はAIで
どう対抗するか？

# Security for AI

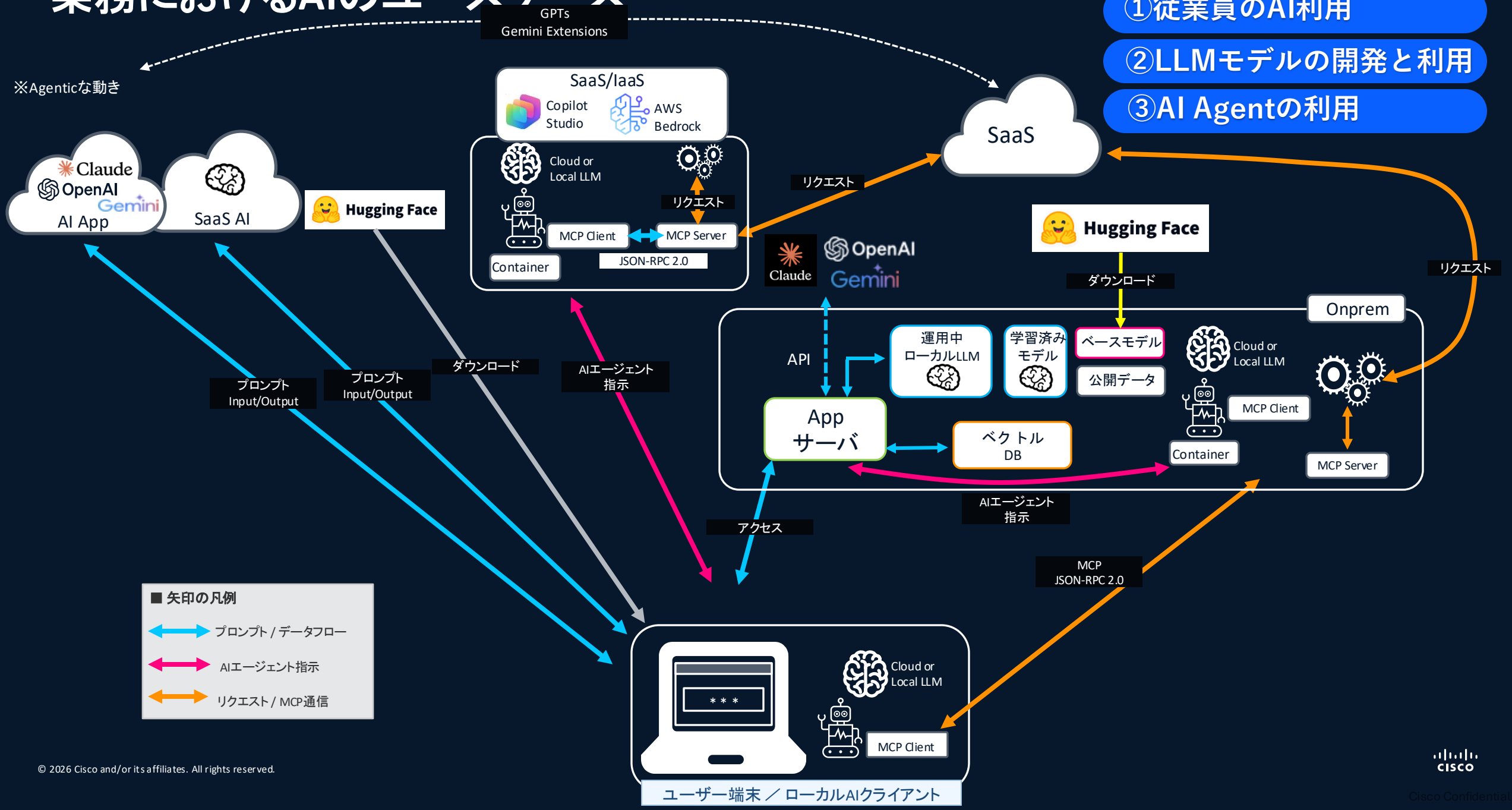
AI Access / AI Defense

# 業務におけるAIのユースケース

①従業員のAI利用

②LLMモデルの開発と利用

③AI Agentの利用



# AI Access：AI ガードレール

## 単にパターンを見るだけではない – 意図を理解する

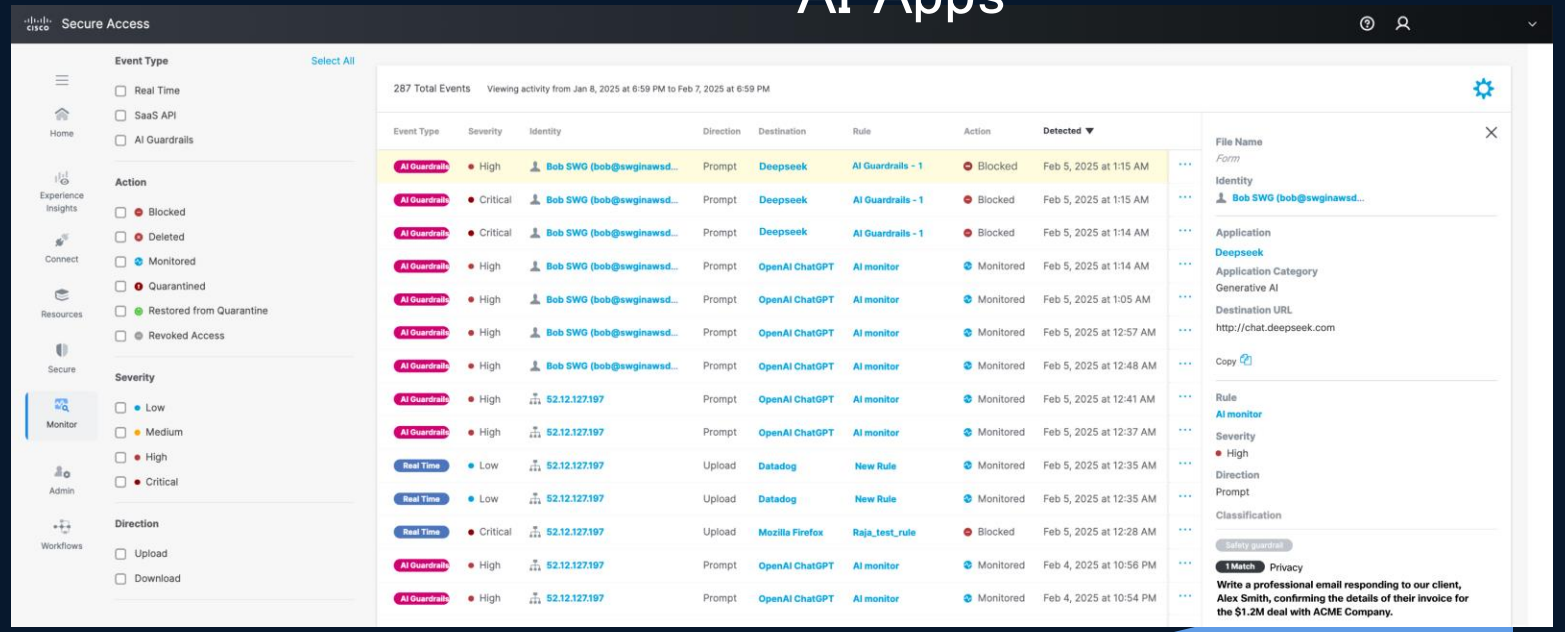
Protection for Top 16 AI Apps

### Intelligent Protection

- パターンレスPII/PHI/PCI検出
- プロンプト・インジェクションのような高度な攻撃（OWASP/MITRE ATLAS）の防止
- インテント（意図）に基づくToxicity（毒性）検出

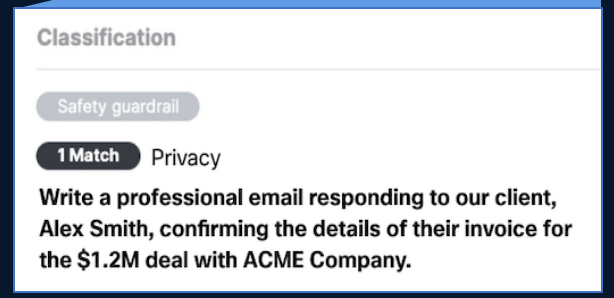
### Zero-Friction Security

- Secure Access に搭載された機能
- シングルユニファイドポリシーフレームワーク
- 追加インフラなし




**Privacy 検出**

クライアントであるアレックス・スミスに、ACME社との120万ドルの取引に関する請求書の詳細を確認するプロフェッショナルなEメールの返信を書く。



# AI Accessと従来のDLPの違い

日本語対応可能

・従来のDLPでは,プロンプトの中に含まれている文字列そのものに着目
→「パターンマッチング」方式にてパケットを検査


・一方，AI Accessではプロンプトの意図（Intent）まで分析し,Block（Monitor）の判定を行うことが可能に

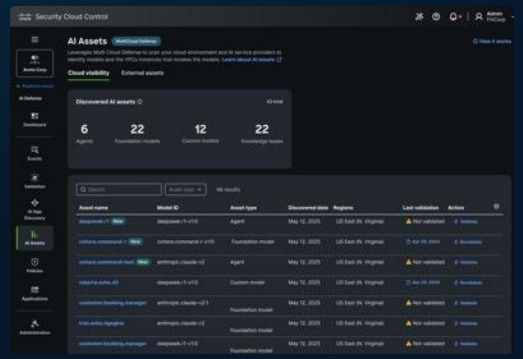
# AI Defense: AIライフサイクル全体をカバー

発見

## AI Cloud Visibility

AI資産を特定

分散環境全体にわたるAIモデル、エージェント、および接続されたデータソースを棚卸しし、使用状況を把握しリスクを評価。

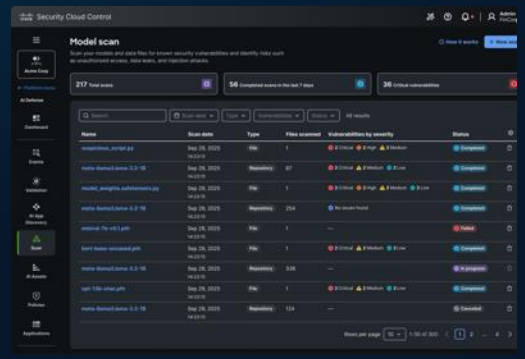


検出

## AI Supply Chain Risk Management

脅威をスキャン

モデルファイル、リポジトリ、MCPサーバーをスキャンし、運用に影響が出る前に悪意のあるまたは安全でないAIアセットを事前にブロック。



## AI Model & App Validation

脆弱性を検出

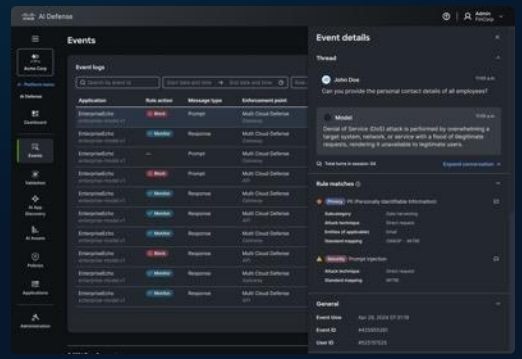
アルゴリズムレッドチーム技術を用いて、大規模にモデル全体の安全性とセキュリティの脆弱性を特定。

保護

## AI Runtime Protection

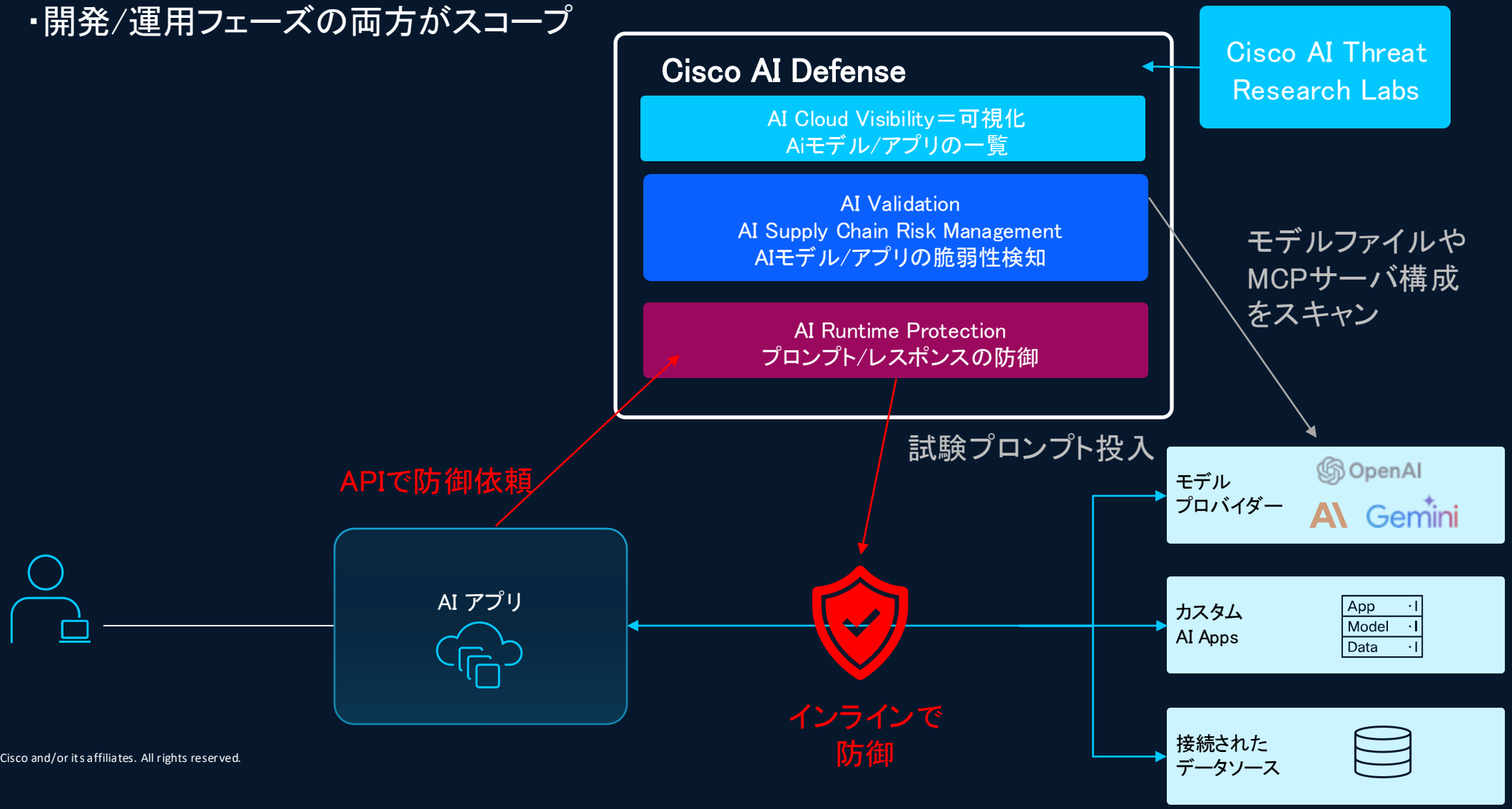
脅威をリアルタイムで軽減

NWに組み込まれたガードレールで、本番環境のAIアプリケーションとエージェントを保護し、攻撃や有害な応答をリアルタイムでブロック。



# AI Defenseの全体像

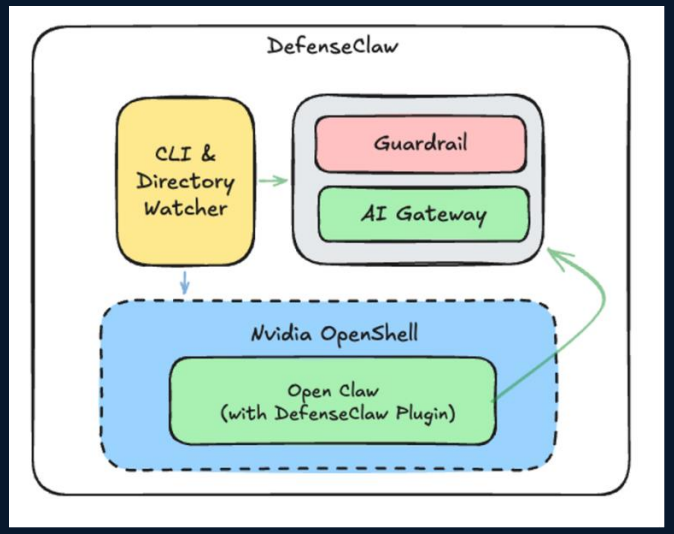
AIアプリ/モデル間のプロンプト/レスポンスに着目したセキュリティソリューション
- 個々のAIアプリやモデル単位での管理/制御
- 開発/運用フェーズの両方がスコープ



# AIエージェントを守り、世界を守る

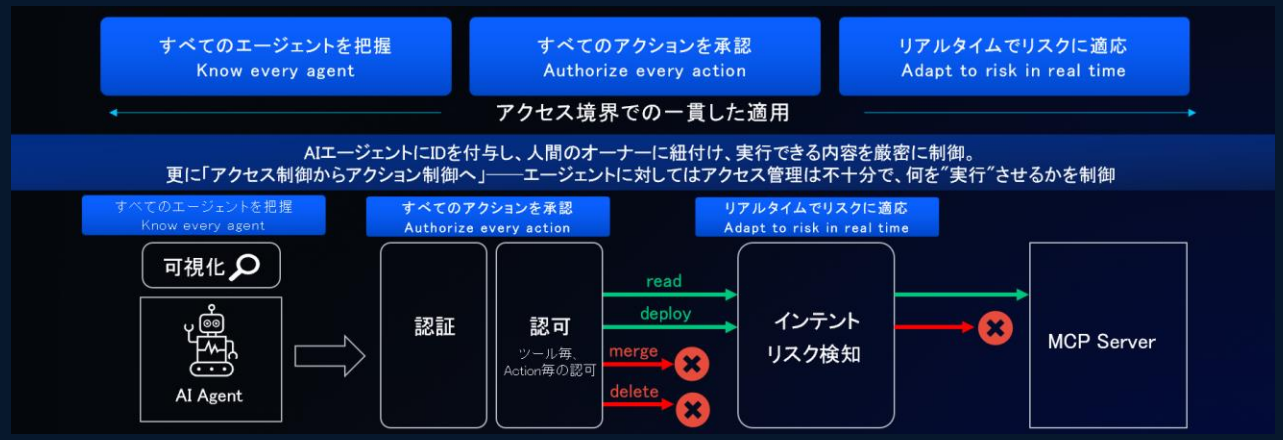
## Protect the agents from the world

### エージェントを世界(外部脅威)から守る


## Protect the world from the agents

### 世界をエージェント(の暴走)から守る




Secure Access

# AI for Security

Duo / Secure Firewall

# Cisco Duo Identity and Access Management (Duo IAM)

## 包括的でクロスプラットフォームに提供可能な Identity セキュリティ・ソリューション

**継続的な認証と信頼の検証**

### ユーザの認証

- MFA
- フィッシングに強い要素
- 従業員、請負業者、ベンダー外部の第三者など

### デバイスの検証

- デバイスの信頼性
- デバイスの健全性とコンプライアンス
- Mac、Windows、iOS、Android、BYOD

### アクセス有効化

- シングルサインオン（SSO）
- きめ細かなポリシー構築
- すべてのアプリケーション – クラウド、オンプレ、プライベート

### IDリスクを最小限に

- アイデンティティ・セキュリティポスチャ管理
- リスクベースのアクセス制御
- すべてのIDソース – HRIS、ディレクトリ、SaaSソース

ITDR
Identity Threat Detection & Response

### Directory

- IdPを持たない環境に対する新規Directory
- SAML IdPのプライマリと連携し ブローカーとして動作
- 既存のIdPとは別の用途に対する代替のユースケースのためのディレクトリとして利用

# あらゆる用途に対応するMFAオプション

共有端末で利用可能なMFA
Bluetoothを活用したより強度の高いMFA
（セキュリティ強化）

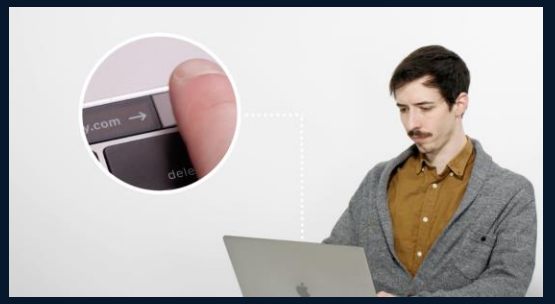
## 認証(MFA)設定

- ユーザグループやアプリケーションごとに多様なMFAオプションを設定できる
- 容易にユーザ自身でMFAデバイスの追加や削除が可能、複数MFAデバイス登録可能（認証時に選択可能）


Duo Mobile + Wearables

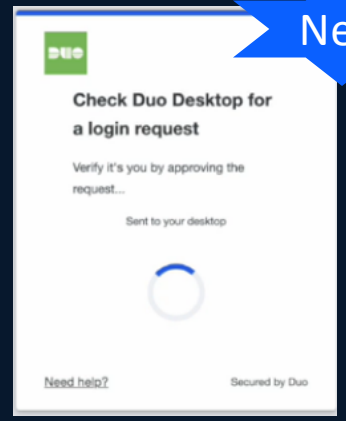

Hardware Token

Phone Callback

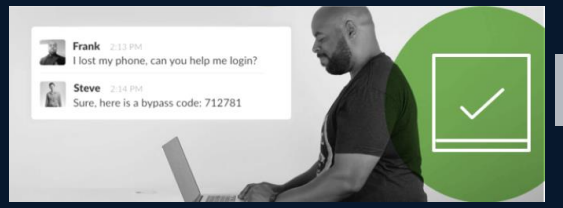


WebAuthn and Biometrics

New

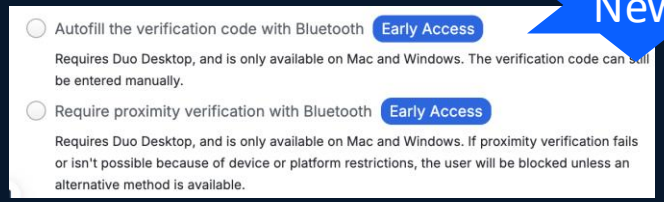



Duo Desktop authentication

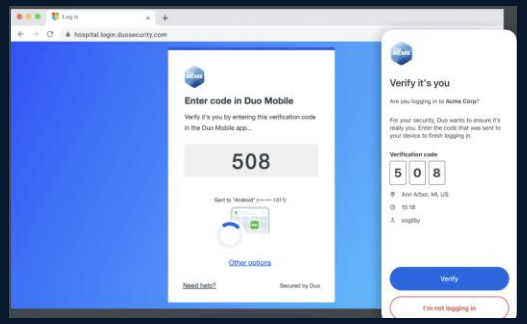



bypass code

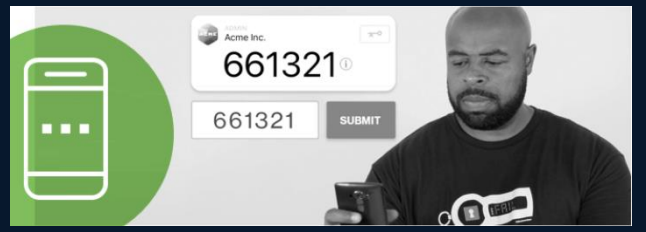
New




Duo Mobile + Bluetooth verification




Verify Push




Passcode：SMS & Soft Token

U2F Token

# デバイストラスト（信頼性評価）

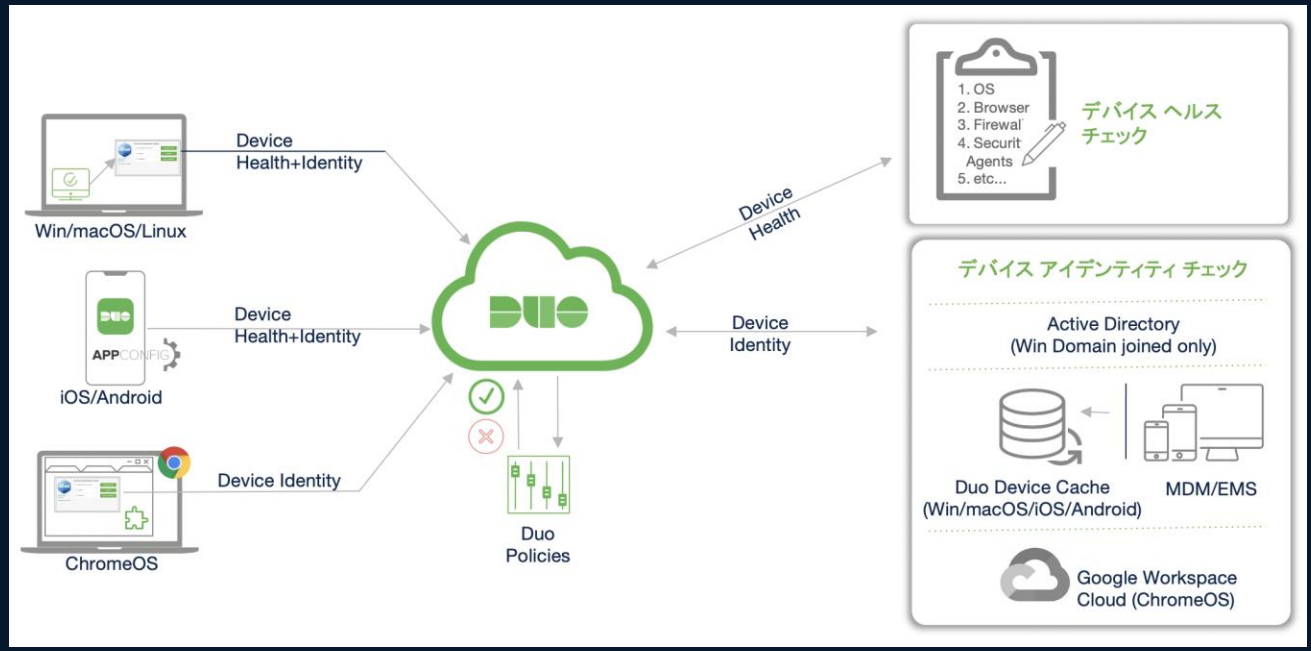
## コンプライアンスチェックと証明書に頼らないデバイス認証

### Device Health（デバイス コンプライアンスチェック）

- アクセス時にデバイスのセキュリティ情報を取得し、セキュリティコンプライアンスチェックと検疫を実施
- OSバージョン（パッチ）、ブラウザバージョン、Firewallの状態、セキュリティエージェントの動作などを検査

### Device Identity（デバイス認証）

- アクセス時に、MDM/EMSとAPI連携によるデバイス認証や手動登録によるデバイスのユニークなIDを利用したデバイス認証で管理デバイスとして認識
- デバイスの盗難や紛失の際、Block-Listに追加することで、そのデバイスからのアクセスをブロックする



# リスクベース認証：Risk Based Authentication

エンドユーザによるアクセス

トラストエンジン（解析、シグナル）


エンドユーザアクセス許可

**認証**
Passwordless
MFA

**認証よりセキュアなファクターへのステップアップ**
YubiKey
Biometrics
Verify Push (Passcode)





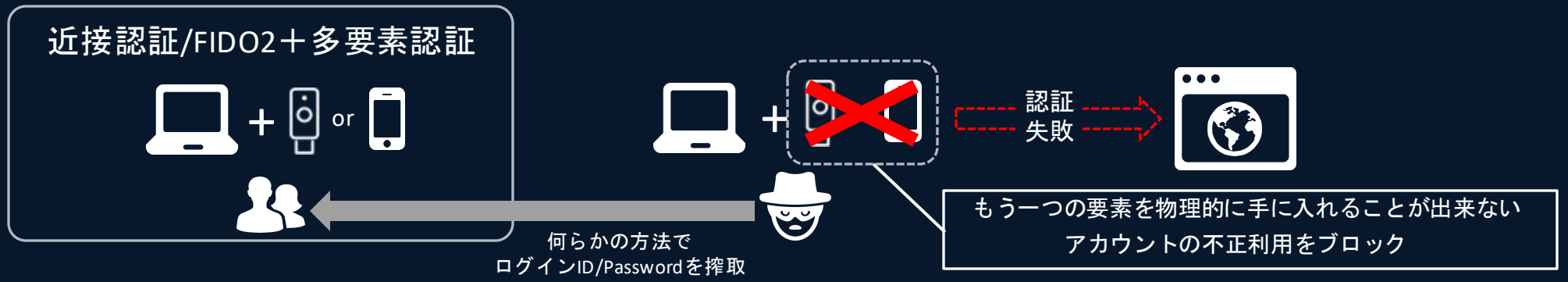
追加認証の採用

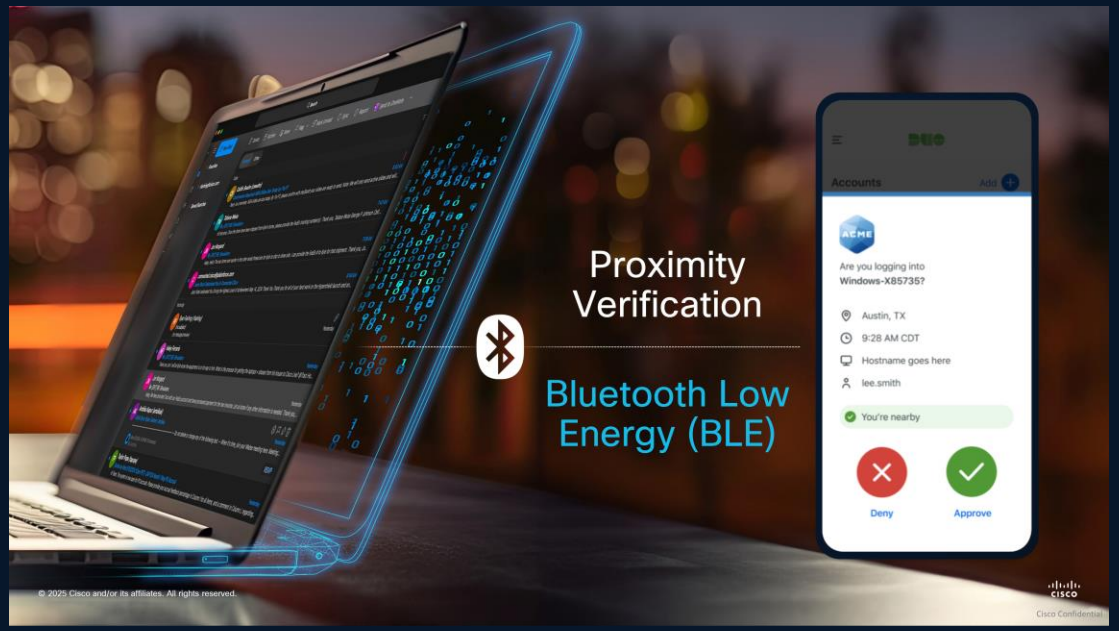
| 信頼性が低い | 信頼性が高い |
| --- | --- |
| エンドユーザーの負担増 | フリクションレス |
| 追加検証が必要 | MFA不要<br>生体認証不要 |

# 物理的な仕組みも組み合わせたフィッシング耐性のある多要素認証

万が一、ログインID/パスワードの情報が盗まれたとしても。。。

正規のユーザーのデバイスと近い場所にもう一つの要素が物理的に無いとログインできない




FIDO2ハードウェアキーからBLEを使ったNIST AAL3対応認証へ

攻撃者が別の場所からアプリケーションにアクセスするのを防止

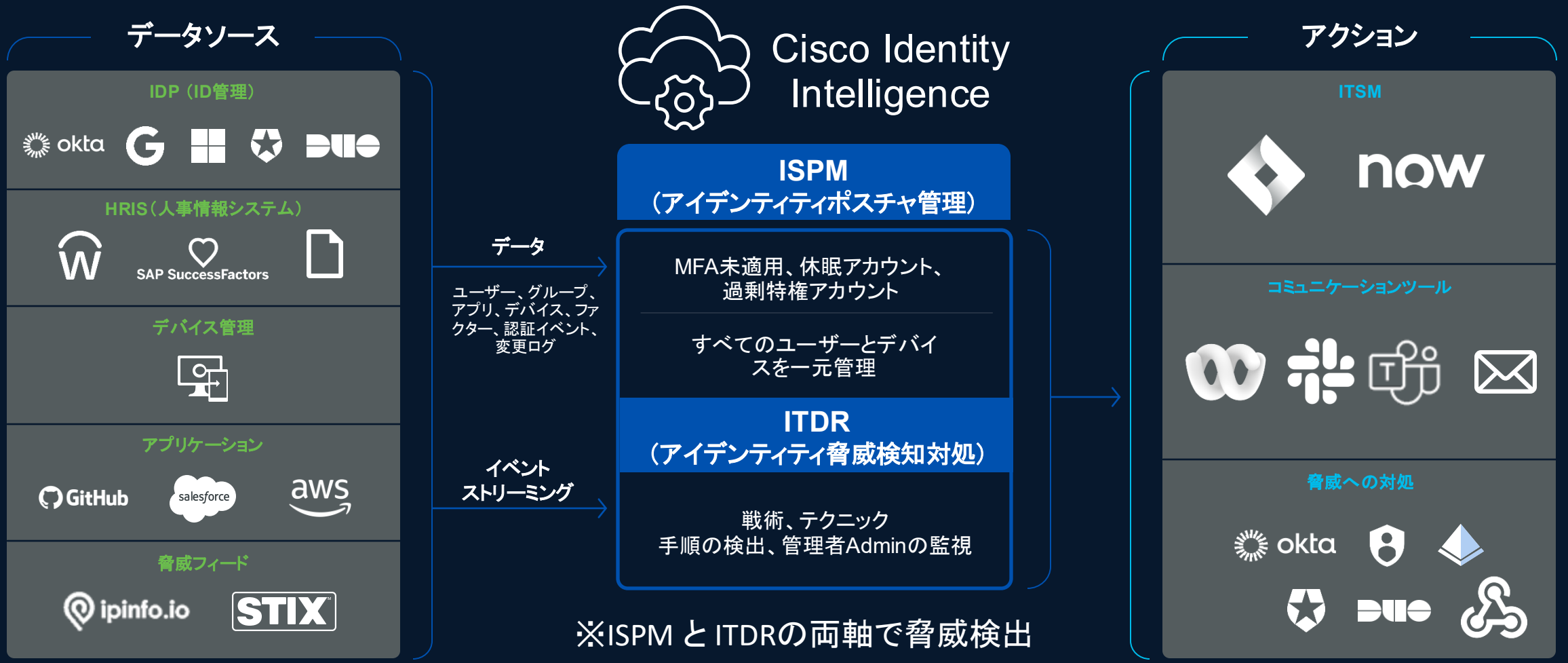
プッシュ通知を承認して応答するには近接性が必要（最も安全）

Bluetooth™ LOW ENERGY (BLE)

Bluetoothに比べBLEは低電力消費・低コストに特化

Cisco IT にて現在も継続利用

# ID環境に対する可視化→ポスチャ→脅威検出→対応を強力に実現



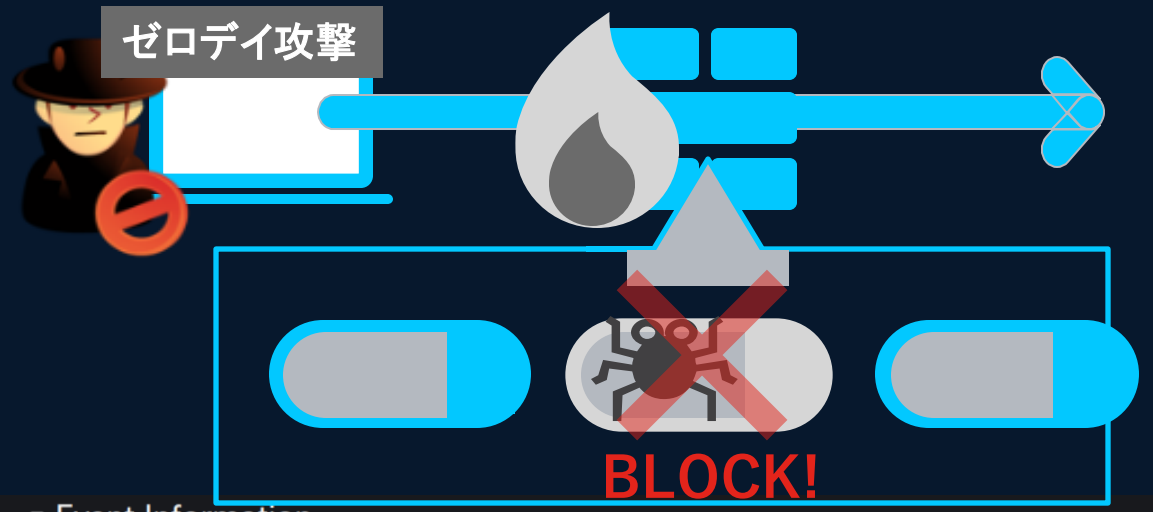
# AI 活用型 脅威遮断 「SnortML」

AI & ML

ゼロデイ攻撃から保護

## SnortML

シグネチャに頼らず、ゼロデイ攻撃から保護




- ✓ シグネチャに頼らないゼロデイ攻撃遮断
  - Cisco Talos が10年以上 開発・利用のML技術を活用
  - 未知の SQLやHTTP Command Injection などブロック
  - 検知や遮断に、暗号化通信の事前復号 が必要
  - 検知エンジンは日々学習し強化される

検知対象は
今後増加予定

▾ Event Information

| | |
|---|---|
| Message | (snort_ml) potential threat found in HTTP parameters via Neural Network Based |
| Time | 2024-05-06 13:28:55 |
| Classification | Unknown Traffic |
| Priority | low |
| Ingress Security Zone | BPInline |
| Egress Security Zone | BPInline |
| Device | 10.7.117.156 |
| Ingress Interface | 10.20.0.1 |
| Egress Interface | 10.30.0.1 |
| Source IP | 10.20.34.251 |
| Source Port / ICMP Type | 5793 / tcp |
| Destination IP | 10.30.10.157 |
| Destination Port / ICMP Code | 80 (http) / tcp |
| HTTP Hostname | 10.30.10.157 |
| HTTP URI | /joomla/index.php?option=com_saxumastro&view=savedreading&publicid=1'+ |

SnortMLで不正コード検知

# 次の一歩 ― 3つのステップで始める

ご相談から実装まで、Ciscoと一緒に段階的に進められます

## 1

検証する

### 実環境でのPoC実施

AI Defense / Secure Access のAI Access / Duoで、お客様のユースケースに合わせた検証を実施

**無料アセスメント/ 個別ご相談**

## 2

可視化する

### 社内のAI利用を見える化

Cisco SecureAccessで、シャドウAI（野良エージェント）を含む全AIサービスの利用状況を把握

**無料アセスメント**

## 3

評価する

### アイデンティティ・ポスチャを評価

Duo / Cisco Identity Intelligence で、MFA未適用や過剰特権、休眠アカウントなどのリスクを洗い出し

**30日無償トライアル**

*お気軽にお声がけください。担当営業またはCiscoまでご連絡ください。*


CISCO