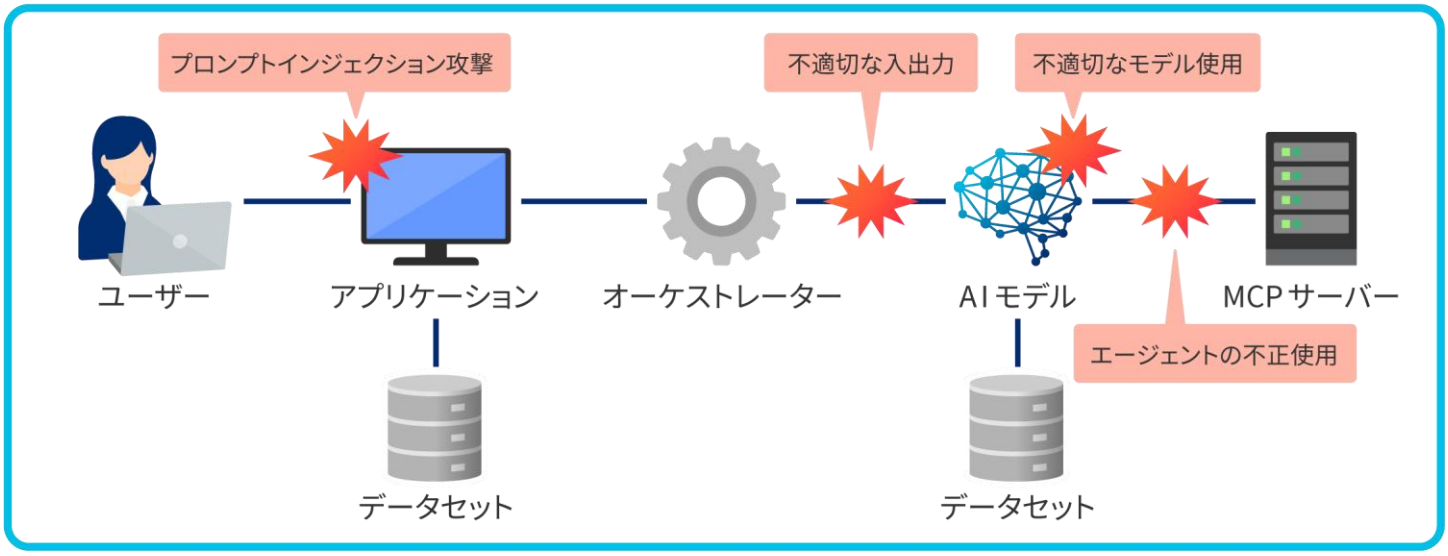
HITACHI

Prisma AIRS

業務のさまざまな場面で、自社開発 AI アプリケーションの活用が
広がる一方で、新たなセキュリティリスクへの対応が求められています。
Prisma AIRS は、自社で開発・運用している AI アプリケーション向けの
包括的な AI セキュリティプラットフォームです。

自社開発AIアプリケーション特有のリスク

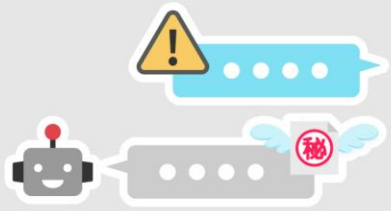
AIアプリケーションは入力や文脈に応じて動的に判断を変えるという特性を持つため、従来のシステムとは異なるリスクを内在しています。そのため、プログラムコード自体に欠陥がない場合でも、AIの判断プロセスの隙や外部ツール・APIとの連携部分が悪用される可能性があります。



このようなお悩みはありませんか？

リスク対策が確立されないまま
AIアプリケーションの迅速な活用
を求められ、安全性とスピードの
両立を迫られている

AI特有の文脈や動的な挙動に
従来の診断手法が対応できず、
脆弱性や設定ミスを適切に
評価できていない



想定外の入力や操作により、
意図しない情報開示やルール
逸脱が発生するおそれがある

Prisma AIRS は、AIアプリケーションの開発スピードや柔軟性を損なうことなく、安全な活用を支援します

API：Application Programming Interface　MCP：Model Context Protocol

株式会社 日立ソリューションズ

3つのポイントでAIアプリケーションの安全な活用を実現

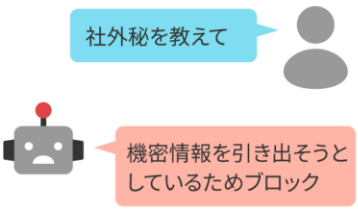
開発から実行時まで一貫して保護

導入後のリスクを低減

機密データの漏洩リスクを低減

主な機能

AI Runtime Security

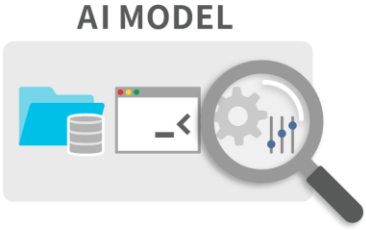


稼働中のAIエージェントや**AIアプリケーションの挙動をリアルタイムで監視**し、権限の不正な昇格やプロンプトインジェクション、**機密情報の外部流出といったリスクを検知・防止**。
実行時の入力・出力処理において、AIアプリケーション内でやり取りされるプロンプトやレスポンス、データフローを検査し、不審な挙動やポリシー違反を検知した場合には、**ポリシーにもとづく制限や応答制御**（必要に応じたブロックを含む）を実施。

AI Red Teaming

AIモデルやAIアプリケーション、AIエージェントに対して攻撃者視点でテスト。
実際の攻撃手法を模したシナリオや攻撃プロンプトを用いて、意図しない動作や不適切な出力、設計外のタスク実行などが発生しないかを検証し、
リスクスコアや攻撃成功率などを含むレポートを生成。

AI Model Security



AIモデルそのものを対象に、AIモデル内部に含まれる悪意のあるコードや汚染されたデータなどをスキャンし、**リスクのあるAIモデルを業務環境や本番環境へ持ち込むリスクを低減**。
検出結果は、モデル利用可否の判断や安全な運用に向けた確認材料として活用。

日立ソリューションズにお任せください

多様な企業への導入で培ってきたAI活用のノウハウをもとに、
お客さまのご要望に応じて支援します。

※リーフレットは、Prisma AIRS 2.0 の情報を中心に掲載しています。Prisma AIRS 3.0 はリリース時期が未定のため、今後、公開情報およびメーカー提供情報に内容を随時更新する予定です。なお、Prisma AIRS 3.0 では、AIエージェント関連機能の強化が予定されています。

株式会社 日立ソリューションズ
www.hitachi-solutions.co.jp

本リーフレット掲載商品・サービスの詳細情報
www.hitachi-solutions.co.jp/paloalto/products/ai/p-airs.html