AKAMAI 製品概要

# Firewall for AI

Akamai Firewall for AI は、AI ベースのアプリケーション、大規模言語モデル（LLM）、AI 主導の API を新たなサイバー脅威から保護するように設計された専用のセキュリティソリューションです。インバウンド AI クエリーとアウトバウンド AI 応答のセキュリティを確保することで、生成 AI がもたらすセキュリティギャップをファイアウォールが解消します。

このファイアウォールは、リアルタイム検知、ポリシーベースの適用、および適応型セキュリティ対策により、プロンプトインジェクション、機微な情報の漏えい、敵対的攻撃、AI 固有のサービス妨害（DoS）攻撃から保護します。

エッジ、クラウド、ハイブリッド、オンプレミス環境と統合できる Firewall for AI は、パフォーマンスを損なわずに、一貫したセキュリティ、安全性、ガバナンス、コンプライアンスを実現します。

## AI 固有の脅威防御

Firewall for AI は、従来のセキュリティツールでは対応できなかった AI 固有の脆弱性を特定して緩和することで、AI 主導のアプリケーションに包括的なセキュリティを提供します。

- **プロンプトインジェクションの防御** — 不正な入力によって AI モデルを操作する攻撃者から保護します。
- **データ損失防止（DLP）** — AI によって生成された応答で機微な情報の漏えいを検知して阻止し、要求内の機微な情報の受信を防止します。
- **有害コンテンツフィルタリング** — 配信前に、ヘイトスピーチ、誤情報、不快なコンテンツへのフラグ付けを行います。
- **攻撃的な AI セキュリティ** — リモートコード実行、モデルバックドア、データポイズニング攻撃から保護します。
- **サービス拒否攻撃の緩和** — クエリーの過剰な使用とモデルの過負荷を制御することで、AI 主導型の DoS 攻撃を緩和します。

### ビジネス上のメリット

**統一されたセキュリティポスチャ**
エッジ、クラウド、ハイブリッド、オンプレミス、全てに対応する標準化された AI セキュリティ

**自動化された AI 脅威検知**
手動のルール調整不要の AI 特化型の保護

**WAAP とのシームレスな統合**
AI 防御によって Web アプリケーションと API 保護（WAAP）を拡張

**AI の誤用や法的リスクを防止**
データ漏えい、IP 盗難、規制違反を阻止

**AI セキュリティのシンプル化**
社内エンジニアがセキュリティポリシーを手動で適用する必要なし

**マルチクラウドの柔軟性**
環境全体で AI ワークロードを保護

**エンタープライズクラスの AI 保護**
Akamai のグローバルな脅威インテリジェンスがサポート

## 柔軟性に富む展開オプション

Firewall for AI は、さまざまな AI アーキテクチャとクラウド環境に合わせた複数の導入モデルを提供します。

| 導入モデル | 説明 |
| --- | --- |
| Akamai エッジ統合 | 低レイテンシーのセキュリティ適用により、Akamai エッジにおいてインラインで AI アプリケーションを保護します。 |
| REST API | API ベースのリスク検知とスコアリングを介して AI の入出力をスキャンします。 |
| リバースプロキシの導入（ロードマップ機能） | AI トラフィックを Akamai の安全なプロキシ経由でルーティングし、詳細な検査とフィルタリングを行います。 |

この柔軟性により、マルチクラウド、ハイブリッド、オンプレミス環境など、場所を問わずに導入された LLM のセキュリティを確保できます。

## 仕組み

### AI トラフィックの分析

ファイアウォールは AI インタラクションを監視および分析し、受信ユーザープロンプトと AI 生成出力を検査して、モデルやエンドユーザーに到達する前に潜在的な脅威を検知します。AI クエリー応答サイクルを分析することで、ファイアウォールはアプリケーションのパフォーマンスを維持しながら、セキュリティリスクを効果的に防止します。

### リスクスコアリングと適応型脅威への対応

AI インタラクションは、プロンプトインジェクション、機微な情報の露出、攻撃の脅威など、複数のセキュリティ指標に対して評価されます。

### セキュリティ強制アクション

Firewall for AI は、リスクスコアと顧客のリスク選好に基づいて、次の 3 つの重要なセキュリティ対策を実施します。

- **監視：**AI クエリーや応答を妨害することなく、検知された脅威を分析用に記録します。
- **変更：**自然な会話の流れを維持しながら、AI によって生成された出力をインラインで調整し、安全でないコンテンツを削除または変更します。
- **拒否：**リスクの高い入力が AI モデルに到達するのをブロックし、安全でない応答がユーザーに返されることを防ぎます。

## コンプライアンスとガバナンスの信頼性

Firewall for AI は、セキュリティおよびコンプライアンスの基準を満たすのに役立ちます。AI を活用したアプリケーションが新たな規制上の課題をもたらす中、データプライバシー、モデルの整合性、セキュリティリスクの監視を維持することは非常に重要です。

### 規制の整合

Firewall for AI は、プライバシー、安全性、およびセキュリティのガイドライン順守を支援します。AI 固有のセキュリティポリシーを適用することで、企業はデータ保護規制、倫理的な AI の使用、および企業ガバナンスの要件に関連するリスクを緩和できます。

## セキュリティ分析とロギング

Firewall for AI は、詳細な監査ログとリアルタイムのセキュリティ分析を提供し、セキュリティチームが AI セキュリティイベントを可視化できるようにします。クエリーパターン、脅威インジケーター、および応答のふるまいを監視することで、組織は異常を事前に検知し、ポリシー制御を実施し、コンプライアンスレポートを生成できます。

## エンタープライズクラスの AI 保護

Akamai のグローバルな脅威インテリジェンスに支えられたファイアウォールは、新たに出現する AI セキュリティの脅威に継続的に適応します。AI セキュリティ調査と脅威モデリングから得られたリアルタイムのデータ知見を活用することで、AI アプリケーションを安全かつ責任を持って動作させると同時に、回復力の高いセキュリティポスチャを維持できます。

**エキスパートによるコンサルティングで詳細をご確認ください。**