INFORMACIÓN SOBRE EL PRODUCTO DE AKAMAI

# Firewall for AI

Akamai Firewall for AI es una solución de seguridad diseñada para proteger las aplicaciones basadas en IA, los modelos de lenguaje de gran tamaño (LLM) y las API basadas en IA frente a las ciberamenazas emergentes. Al proteger tanto las preguntas que recibe la IA como las respuestas que envía, el firewall elimina las brechas de seguridad que presenta la IA generativa.

Este firewall, dotado de detección en tiempo real, aplicación de reglas basada en políticas y medidas de seguridad adaptables, protege contra la inyección de instrucciones, las filtraciones de datos confidenciales, las vulnerabilidades frente a adversarios y los ataques de denegación de servicio (DoS) específicos de la IA.

Además de integrarse a la perfección con entornos en el Edge, la nube, híbridos y locales, Firewall for AI garantiza de forma sistemática la máxima seguridad, una buena gobernanza y la conformidad normativa sin sacrificar el rendimiento.

## Protección contra amenazas específicas de la IA

Firewall for AI proporciona una seguridad integral para las aplicaciones basadas en IA, ya que identifica y mitiga las vulnerabilidades específicas de la IA que las herramientas de seguridad tradicionales no logran abordar.

- **Defensa contra la inyección de instrucciones:** protege contra los atacantes que manipulan los modelos de IA mediante la introducción de datos engañosos.
- **Prevención de pérdida de datos (DLP):** detecta y bloquea las filtraciones de datos confidenciales en las respuestas generadas por IA, y protege contra la recepción de datos confidenciales en las solicitudes.
- **Filtrado de contenido tóxico e inapropiado:** detecta los discursos de odio, la información engañosa y el contenido ofensivo antes de su difusión.
- **Seguridad de la IA frente a los adversarios:** protege contra la ejecución remota de código, los ataques de puerta trasera y los ataques de envenenamiento de datos.
- **Mitigación de la denegación de servicio:** mitiga los ataques de denegación de servicio basados en IA, puesto que controla el abuso de consultas y la sobrecarga del modelo.

### Ventajas para su empresa

**Estrategia de seguridad unificada para la IA**
Seguridad estandarizada para la IA en entornos en el Edge, en la nube, híbridos y locales

**Detección automatizada de amenazas a la IA**
Protecciones específicas para la IA sin necesidad de ajustar las reglas de forma manual

**Integración perfecta con soluciones WAAP**
Amplía la protección de API y aplicaciones web (WAAP) con defensas para la IA

**Prevención del uso indebido de la IA y de los riesgos legales**
Bloquea las filtraciones de datos, los robos de propiedad intelectual y las infracciones de la normativa

**Seguridad simplificada para la IA**
No es necesario que los ingenieros internos apliquen manualmente las políticas de seguridad

**Flexibilidad en entornos multinube**
Protege las cargas de trabajo de IA en todos los entornos

**Protección de nivel empresarial para la IA**
Con el respaldo de la inteligencia ante amenazas globales de Akamai

## Opciones de implementación flexibles

Firewall for AI ofrece varios modelos de implementación adaptados a diferentes arquitecturas de IA y entornos de nube.

| Modelo de implementación | Descripción |
|---|---|
| Integración con el Edge de Akamai | Protege las aplicaciones de IA en línea en el Edge de Akamai con una aplicación de seguridad de baja latencia. |
| API basada en REST | Analiza las entradas y salidas de la IA mediante un sistema de detección y puntuación de riesgos basado en API. |
| Implementación del proxy inverso (capacidad de hoja de ruta) | Dirige el tráfico de la IA a través del proxy seguro de Akamai para llevar a cabo una inspección y filtrado en profundidad. |

Esta flexibilidad permite a las organizaciones proteger los LLM implementados en cualquier lugar, también en entornos multinube, híbridos y locales.

## Cómo funciona

### Análisis del tráfico de la IA

El firewall supervisa y analiza las interacciones de la IA, inspecciona las solicitudes de los usuarios y los resultados generados por la IA para detectar posibles amenazas antes de que lleguen al modelo o al usuario final. Mediante el análisis del ciclo de consulta y respuesta de la IA, el firewall previene de forma eficaz los riesgos de seguridad sin que afecte al rendimiento de las aplicaciones.

### Puntuación de riesgo y respuesta adaptativa ante las amenazas

Las interacciones de la IA se evalúan en función de múltiples indicadores de seguridad, entre los que se incluyen la inyección de instrucciones, la filtración de datos confidenciales y las vulnerabilidades frente a adversarios.

### Aplicación de medidas de seguridad

Firewall for AI aplica tres medidas de seguridad esenciales basadas en la puntuación de riesgo y la propensión al riesgo del cliente:

- **Supervisar:** registra las amenazas detectadas para analizarlas sin interferir en las consultas o respuestas de la IA.
- **Modificar:** ajusta en línea los resultados generados por la IA, lo que permite eliminar o alterar el contenido poco seguro a la vez que se mantiene el flujo natural de la conversación.
- **Denegar:** bloquea las entradas de mayor riesgo para que no lleguen al modelo de IA e impide que se devuelvan respuestas poco seguras a los usuarios.

## Confianza en el cumplimiento y la gobernanza

Firewall for AI ayuda a cumplir con las normas de seguridad y conformidad. A medida que las aplicaciones basadas en IA plantean nuevos retos normativos, es fundamental controlar la privacidad de los datos, la integridad de los modelos y los riesgos de seguridad.

### Alineación con la normativa

El firewall ayuda a las organizaciones a cumplir las directrices de privacidad, seguridad y protección. Mediante la aplicación de políticas de seguridad específicas para la IA, las empresas pueden mitigar los riesgos relacionados con las normativas de protección de datos, el uso ético de la IA y los requisitos de gobernanza corporativa.

## Análisis de seguridad y registros

Firewall for AI proporciona registros de auditoría detallados y análisis de seguridad en tiempo real, que aportan a los equipos de seguridad una visión detallada de los eventos de seguridad de la IA. Mediante la supervisión de los patrones de consulta, los indicadores de amenazas y los comportamientos de respuesta, las organizaciones pueden detectar anomalías de forma proactiva, aplicar controles de políticas y generar informes de cumplimiento.

## Protección de nivel empresarial para la IA

Con el apoyo de la inteligencia global ante amenazas de Akamai, el firewall se adapta constantemente a las amenazas emergentes contra la seguridad de la IA. Al aprovechar la información en tiempo real de la investigación sobre la seguridad de la IA y el modelado de amenazas, las organizaciones pueden adoptar una estrategia de seguridad resiliente y garantizar al mismo tiempo que sus aplicaciones de IA funcionen de forma segura y responsable.

Hable con un experto para obtener más información.

©2025 Akamai Technologies | Asistencia |
| Publicado en marzo de 2025