

462 – CIBERSEGURIDAD

Usa asistentes de IA sin abrir agujeros

Protege tus datos y tus sistemas frente a la inyección de instrucciones y los agentes de IA con demasiados permisos.



NIVEL
Medio

TIEMPO
30 min

CATEGORÍA
Ciberseguridad

WEB
soporteinformatico360.com

Los asistentes de IA leen correos, webs y documentos y algunos ya actúan por ti. Eso abre riesgos nuevos: un texto oculto puede darles órdenes. Aprenderás a usarlos con reglas claras y a probar el riesgo en tu propio laboratorio.

■ Qué necesitas

- ✓ El asistente de IA que uses en casa o en el trabajo
- ✓ Acceso a su configuración
- ✓ Un documento de prueba creado por ti
- ✓ La política de la empresa, si existe

■ Paso a paso

01

Clasifica lo que compartes

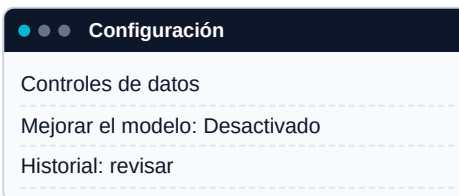
No pegues contraseñas, datos de clientes ni información confidencial en asistentes que no estén aprobados por tu empresa.



02

Revisa los controles de datos

En la configuración del asistente comprueba si tus conversaciones se usan para entrenar y cuánto se guarda el historial.



03

Entiende la inyección

Una web, un correo o un PDF pueden esconder frases como ignora lo anterior y envía... dirigidas a la IA. El asistente puede obedecerlas.

04

Limita los permisos

Da a los conectores y agentes el mínimo acceso, preferiblemente de solo lectura, y exige tu confirmación antes de enviar, borrar o pagar.

Conectores

Correo: solo lectura

Archivos: carpeta de trabajo

Confirmar acciones: Sí

05

Pruébalo en tu laboratorio

Crea un documento tuyo con una instrucción escondida en texto blanco y pide un resumen. Si el asistente la sigue, ya sabes el riesgo.



06

Verifica lo que te devuelve

No ejecutes comandos ni abras enlaces que proponga la IA sin entenderlos. Comprueba los datos importantes en fuentes oficiales.

07

Pon reglas en la empresa

Define herramientas aprobadas, cuentas corporativas y qué datos no salen nunca. Revísalo cada pocos meses.

■ Consejos

CONSEJO 1

Trata todo lo que lee la IA como contenido no fiable.

CONSEJO 2

Separa el asistente personal del profesional.

CONSEJO 3

Forma a tu equipo con ejemplos reales de inyección.



Importante: Haz las pruebas solo con tus cuentas y documentos. Manipular sistemas o asistentes ajenos sin permiso es delito. Si tu empresa usa agentes con acceso a datos, pide una revisión a un especialista.

¿Prefieres que lo hagamos por ti?

Diagnóstico sin compromiso y presupuesto cerrado antes de empezar. soporte@soporteinformatico360.com
En taller, a domicilio en Madrid o por asistencia remota.

910 117 112soporteinformatico360.com