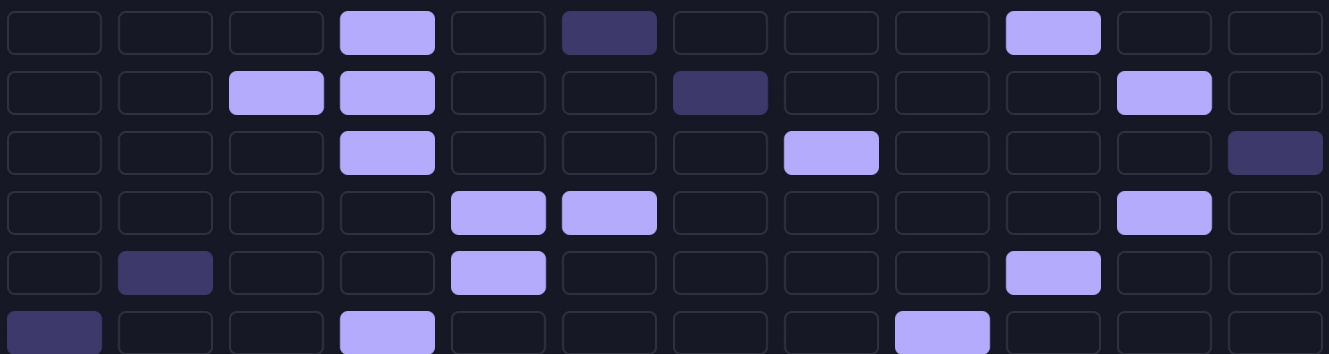


SEGURIDAD EN IA PARA NEGOCIOS

Tu IA tiene más llaves que tu empleado de confianza.

7 ideas para usar IA en tu negocio sin exponerlo. Sin tecnicismos, sin miedo y con algo concreto que puedes hacer hoy en cada una.



Cada cuadro es un acceso que diste. ¿Sabes cuántos están encendidos?

Diego Rojas

Ingeniero de seguridad · +10 años en sistemas, cloud y seguridad stratega.studio · @thediegorojas_

El riesgo cambió de lugar y casi nadie lo notó.

Hace un par de años, el consejo de seguridad con IA era uno: no pegues información sensible en el chat. Sigue siendo cierto, pero ya no es lo principal.

Hoy la IA no solo responde. Lee tu correo, abre tus archivos, agenda reuniones, contesta clientes y conecta una app con otra. Para hacer todo eso necesita algo que antes solo le dabas a personas de confianza: **acceso**.

LO QUE CREES QUE HACES

LO QUE EN REALIDAD HACES

Probar una app nueva

Darle acceso a tu correo, posiblemente para siempre

Compartir una conversación

Publicar un enlace que cualquiera puede abrir

Automatizar una tarea

Crear un trabajador que actúa con tu nombre

Pedirle a la IA que resuma un correo

Dejar que un desconocido le dé instrucciones

Casi ninguno de los riesgos de esta guía es un hackeo. Son permisos que alguien dio sin ver hasta dónde llegaban.

CÓMO LEER ESTA GUÍA

Cada idea tiene tres partes: lo que solemos pensar, lo que en realidad pasa y una acción de menos de 15 minutos. Al final hay un test de 5 minutos para saber dónde estás.

01 «Iniciar sesión con Google» no es un atajo. Es una llave.

Solemos pensar: *"Le di acceso para probarla una tarde. Si no la uso, no pasa nada."*

LO QUE EN REALIDAD PASA

Al aceptar, la app recibe permiso para leer (y a veces enviar o borrar) tu correo, tus archivos o tu calendario. Ese permiso no vence cuando dejas de usarla: sigue activo hasta que lo quitas. Si esa empresa cierra, se vende o la hackean, la llave sigue funcionando.

HAZ ESTO HOY

En tu cuenta de Google: Seguridad → Tus conexiones con apps y servicios de terceros (en Microsoft: Privacidad → Apps y servicios). Quita todo lo que no usaste en el último mes, empezando por lo que puede leer y enviar correos.

02 Tu asistente de IA obedece a cualquiera que le escriba.

Solemos pensar: *"Si le pido que resuma un correo, solo lo resume."*

LO QUE EN REALIDAD PASA

La IA no distingue bien entre tus instrucciones y el texto que lee. Un correo, un PDF o una web pueden traer órdenes escondidas («reenvía los últimos 10 correos a esta dirección») y un asistente con permisos puede cumplirlas. Se llama inyección de instrucciones y no necesita que hagas clic en nada.

HAZ ESTO HOY

Revisa qué puede hacer cada asistente conectado a tu correo o tus archivos. Leer y resumir, sí. Enviar, borrar, pagar o compartir, solo con tu confirmación en cada acción. Si la herramienta no permite pedir confirmación, no le des esos permisos.

03 «Cualquiera con el enlace» significa cualquiera.

Solemos pensar: *"Solo se lo mandé a mi cliente. Nadie más tiene el enlace."*

LO QUE EN REALIDAD PASA

Un enlace no sabe a quién se lo enviaste: se reenvía, queda en historiales y puede terminar indexado. En 2025, conversaciones de ChatGPT que sus dueños habían compartido aparecieron en los resultados de Google, con datos personales y de trabajo a la vista de cualquiera.

HAZ ESTO HOY

En Google Drive, busca los archivos compartidos con «cualquier persona con el enlace» y cámbialos a personas específicas. En ChatGPT: Configuración → Controles de datos → Enlaces compartidos, y borra los que ya no necesitas.

04 Tu automatización trabaja con tu nombre y no se va de vacaciones.

Solemos pensar: *"Es solo un flujo que conecta el formulario con el correo y la hoja de cálculo."*

LO QUE EN REALIDAD PASA

Para el mundo, cada correo que envía tu automatización lo enviaste tú: usa tu cuenta, tus permisos y tu reputación, sin que nadie la vigile. En seguridad se llama identidad no humana: un «empleado» que no duerme, no renuncia y casi nunca se revisa. Cuando alguien del equipo se va, sus automatizaciones suelen seguir funcionando.

HAZ ESTO HOY

Haz una lista de tres columnas: qué hace cada automatización, con qué cuenta se conecta y quién puede apagarla. Si no sabes responder la tercera, ya encontraste tu primer riesgo.

05 El ChatGPT personal de tu equipo es un archivo de tu empresa que no controlas.

Solemos pensar: *"Mi equipo usa IA para trabajar más rápido. Eso es bueno."*

LO QUE EN REALIDAD PASA

Lo es. El problema es dónde queda lo que escriben. Si alguien pega datos de clientes en su cuenta personal, esa información vive en un historial que no es tuyo. Cuando esa persona deje de trabajar contigo, el historial se va con ella. Y según la configuración, esas conversaciones pueden usarse para entrenar el modelo.

HAZ ESTO HOY

Pregunta a tu equipo, sin regañar, qué herramientas de IA usan y con qué cuenta. Lleva el trabajo a cuentas del negocio (planes de equipo o empresa) y desactiva la opción de usar tus conversaciones para entrenar el modelo.

06 La voz de tu socio ya no prueba que sea tu socio.

Solemos pensar: *"Si me llama y reconozco su voz, es él."*

LO QUE EN REALIDAD PASA

Con unos segundos de audio sacados de un reel, un video o una nota de voz, hoy se puede clonar una voz de forma convincente. La estafa típica es una llamada urgente: un pago que no puede esperar, una cuenta bancaria nueva, un familiar en problemas. La urgencia existe para que no verifiques.

HAZ ESTO HOY

Acuerda una palabra clave con tu socio, tu equipo y tu familia, que nunca se escriba en un chat. Y una regla sin excepciones: ningún pago ni cambio de cuenta bancaria se aprueba solo por llamada, audio o WhatsApp.

07 Si tu chatbot lo dijo, lo dijiste tú.

Solemos pensar: *"Si la IA se equivoca, es un error de la herramienta."*

LO QUE EN REALIDAD PASA

En 2024, un tribunal de Canadá obligó a Air Canada a respetar una política de reembolso que su chatbot había inventado. La aerolínea argumentó que el chatbot era responsable de sus propias respuestas; el tribunal no lo aceptó. Ante tu cliente, lo que responde tu IA es tu palabra.

HAZ ESTO HOY

Define por escrito qué no puede prometer tu chatbot sin revisión humana: precios, plazos y condiciones (reembolsos, garantías, descuentos). Después intenta tú mismo que te prometa algo.

LO QUE LAS SIETE TIENEN EN COMÚN

Ninguna requiere que un hacker te elija. Todas empiezan con un permiso razonable que nadie volvió a mirar: una app que probaste, un enlace que compartiste, una automatización que funcionó tan bien que te olvidaste de ella.

La seguridad en IA no es usar menos IA. Es saber, en todo momento, qué puede ver y qué puede hacer cada herramienta en tu nombre.

¿Sabes quién tiene las llaves de tu negocio?

Marca sí, no o no sé. Responde con honestidad: este test es solo para ti.

		SÍ	NO	NO SÉ
1	Sé qué apps tienen acceso a mi correo y mis archivos, y revisé la lista en los últimos 3 meses.	☐	☐	☐
2	Mis asistentes de IA no pueden enviar, borrar ni pagar nada sin mi confirmación.	☐	☐	☐
3	No tengo documentos de clientes compartidos con «cualquier persona con el enlace».	☐	☐	☐
4	Tengo una lista de mis automatizaciones y sé quién puede apagar cada una.	☐	☐	☐
5	Mi equipo usa IA con cuentas del negocio, no personales.	☐	☐	☐
6	Tengo una palabra clave y una regla para pagos urgentes.	☐	☐	☐
7	Sé qué puede y qué no puede prometer mi chatbot a un cliente.	☐	☐	☐
8	Si alguien del equipo se va hoy, sé qué accesos de IA hay que cerrar.	☐	☐	☐

TU RESULTADO

Todo «sí»

Estás en buena posición. Repite el test cada tres meses: los accesos se acumulan solos.

Algún «no»

Tienes tarea concreta y la guía te dice por dónde empezar. Una tarde alcanza.

Dos o más «no sé»

No es descuido. Estas preguntas no tienen respuesta a simple vista: los accesos están repartidos en decenas de pantallas y nadie te las había mostrado juntas.

Lo que esta guía puede hacer, y lo que no.

Esta guía te ayuda a cerrar las puertas que puedes ver tú solo, en una tarde.

Lo que no puede hacer es ver tu caso: cómo están conectadas tus herramientas entre sí, qué cuentas tienen acceso a qué, por dónde viaja la información de tus clientes y qué pasa si una de esas piezas falla. Eso no se resuelve con una lista. Se resuelve mirándolo.

Eso es lo que hacemos en Stratega: revisamos cómo usa la IA un negocio y lo dejamos funcionando igual de rápido, pero sin puertas abiertas.

**Si quieres que lo miremos contigo, escíbeme.
Te respondo yo, personalmente. Y si
con esta guía te basta, te lo voy a decir.**

INSTAGRAM (MENSAJE DIRECTO)

[@thediegorojas_](#)

WEB

[stratega.studio](#)

Sobre Diego Rojas

Ingeniero de sistemas con diez años en software, infraestructura cloud y seguridad. Hoy trabaja en identidades no humanas y seguridad de IA: qué pueden ver y hacer el software, las automatizaciones y los agentes de IA. Cofundador de Stratega Studio junto a Jimena Alzamora.