Guía - Mes Europeo de la Ciberseguridad

Usa la IA de forma segura

Seis claves para usar esta herramienta digital con confianza



vuela

Junta de Andalucía

Agencia Digital de Andalucía

La **Inteligencia Artificial (IA)** está cada vez más presente en nuestro día a día: en el trabajo, en las noticias, en las redes sociales…

Tal vez ya la usas para resolver dudas, o quizás todavía no te has animado a dar el salto. En cualquier caso, es **una herramienta digital con muchas utilidades**, y sobre todo, con mucho futuro. La IA ha llegado para quedarse, así que es el momento de aprender a usarla y aprovechar todo su potencial.

Como ocurre con todas las herramientas, es importante aprender a utilizar la IA de manera responsable, para evitar posibles riesgos e inconvenientes.

Para acompañarte en este proceso, en esta guía recogemos **seis claves que te ayudarán a sacarle partido a la IA de forma segura y con criterio**, incorporando esta herramienta a tu rutina digital.

Clave 1: Comunícate bien con la Inteligencia Artificial

Clave 2: No te fíes de todo lo que te responde la IA

Clave 3: Cuida los datos que compartes con la IA

Clave 4: Configura la IA para proteger tu privacidad

Clave 5: Detecta contenido falso generado con IA

Clave 6: Genera contenido de IA con responsabilidad



Clave 1: Comunícate bien con la Inteligencia Artificial

Para que herramientas de Inteligencia Artificial como ChatGPT, Gemini o Claude te den respuestas útiles, no basta con escribir la primera pregunta que se te ocurra. Cuanto más clara sea tu instrucción —lo que se conoce como ***"prompt"***— mejor será la respuesta que recibas.

Cinco elementos de un buen prompt:

Rol: dile a la IA en qué papel se debe meter. Ej.: "Actúa como un profesor de historia".

Contexto: explica tu situación. Ej.: "Quiero preparar un examen".

Tarea: qué quieres exactamente que te ofrezca. Ej.: "Redacta un resumen" o "Hazme un esquema".

Formato: cómo quieres recibirlo. Ej.: "En una lista de 5 puntos" o "En 200 palabras".

Restricciones: qué debe evitar en su respuesta. Ej.: "Sin tecnicismos" o "Sin usar palabras en inglés".

Dos reglas de oro:

Detalla tus necesidades: cuanta más información le des con estos cinco elementos, más útil será su respuesta.

Vuelve a intentarlo: si el resultado no te convence, no repitas la misma instrucción esperando algo distinto. Añade más contexto, concreta la tarea o especifica el formato que necesitas.

Prompt útil: **combina estos elementos en una sola instrucción, por ejemplo:**

"Actúa como [rol]. Mi situación es [contexto]. Quiero que [tarea]. Dame el resultado en [formato] y evita [restricción]."

Clave 2: No te fíes de todo lo que te responde la IA

La **Inteligencia Artificial** no "sabe" las cosas, sino que predice qué palabra es más probable que venga después de otra, a partir de toda la información con la que ha entrenado. Por eso **puede equivocarse**, y al usarla, puede que te encuentres las siguientes situaciones:

Alucinaciones



La IA puede **inventar información** que suena completamente real y presentarla con la misma seguridad con la que da una respuesta correcta. **La IA no valida si la información es veraz** y por eso puede inventar nombres, cifras, hechos o incluso documentos completos.

Sesgos



La IA **puede arrastrar prejuicios o discriminaciones** presentes en los datos con los que fue entrenada. Antes de dar por buena una respuesta, revisa si ha asumido, sin motivo, alguna desigualdad: el género de una profesión, si está discriminando por edad, procedencia o cualquier otra característica o si el lenguaje usado resulta poco inclusivo o injusto con algún colectivo.

Inconsistencia



Una misma IA **puede generar respuestas diferentes ante la misma pregunta** formulada en momentos distintos o con pequeñas variaciones. Por eso, no puedes asumir que una respuesta será siempre la misma ni utilizarla como referencia definitiva sin verificarla.

Desactualización



Los modelos de IA tienen una fecha de corte en sus datos de entrenamiento. Esto significa que **pueden generar información obsoleta sin advertirlo.** Un cambio normativo, un dato actualizado o un evento reciente pueden no estar reflejados en las respuestas.

Qué hacer para no caer en estos errores de la IA:

- **Contrasta siempre:** verifica cualquier dato antes de darlo por válido con otras fuentes de confianza (medios de comunicación, instituciones, profesionales, etc.)
- **Revisa antes de compartir:** antes de publicar o enviar un contenido generado por IA, tienes que leerlo, corregirlo y asegurarte de que es adecuado y veraz.
- **Utiliza tu criterio:** no te bases solo en la IA para tomar decisiones importantes. Sus resultados deben utilizarse como apoyo, no para que decida por ti.



Clave 3: Cuida los datos que compartes con la IA

La Inteligencia Artificial puede utilizar tus conversaciones para seguir entrenándose, guardando todo lo que escribes. Esto puede suponer un riesgo si esa información llegara a manos equivocadas.

Por ejemplo, si compartes **datos sensibles o personales**, estos pueden aparecer, de forma directa o indirecta, en respuestas generadas para otras personas; y si la herramienta sufre un ataque, podrían acceder a lo que hayas guardado en ella, igual que con cualquier otra base de datos online.

Para reducir este riesgo, muchas plataformas incluyen en sus **ajustes de privacidad** la opción de **desactivar el uso de tu historial para entrenar el modelo.** Como primera pauta, te recomendamos buscar esta opción y activarla antes de usar la IA con normalidad.

Además de este ajuste, ten en cuenta estos hábitos:

Cero datos privados: no escribas tu DNI, tu domicilio, números de cuenta, contraseñas, historiales médicos ni datos confidenciales de tu trabajo en ningún asistente de IA.

Y si alguna vez necesitas la ayuda de la IA para resolver un caso real, "disfraza" la información: sustituye nombres por variables ficticias, por ejemplo "Cliente A" o "Empresa X".

Haz limpieza: elimina de vez en cuando el historial de tus conversaciones, sobre todo si has usado la IA para tratar algún asunto personal.

Lee la letra pequeña: antes de usar una nueva aplicación de IA, revisa su política de privacidad para saber qué hace con tu información y durante cuánto tiempo la conserva.

¿Y si ya he compartido mis datos personales con la IA?

- **No te alarmes.** Si crees que se te ha "escapado" información de tu teléfono o tu correo electrónico a través de una herramienta de IA, lo primero es borrar esa conversación o limpiar el historial de la aplicación.
- **A partir de ahí, basta con estar un poco más alerta:** si alguien llegara a usar esa información, podrías recibir algún SMS o correo extraño intentando engañarte, lo que se conoce como *phishing*. Ignora cualquier mensaje desconocido que te pida hacer algo con urgencia y no accedas a sus enlaces.

Clave 4: Configura la IA para proteger tu privacidad

Proteger tu privacidad también depende de cómo configures las herramientas de IA que usas a diario, no solo de cómo las usas.

Al igual que ocurre con cualquier aplicación que instalas en el móvil o usas en el ordenador, cada **asistente de IA pide permisos y recopila información sobre ti.**

Revisar esos ajustes es tan importante como saber usar la propia herramienta.

Aquí te recomendamos algunos **pasos para proteger tus datos al usar herramientas de Inteligencia Artificial** u otras aplicaciones que utilizan la IA:

Usa herramientas de confianza: desconfía de aplicaciones de IA poco conocidas, sin reseñas o que soliciten permisos excesivos nada más instalarlas.

Refuerza el acceso: usa contraseñas seguras y activa siempre la verificación en dos pasos en las plataformas de IA que lo permitan.

Política de privacidad: comprueba qué datos recopila la herramienta, cómo los almacena y si los utiliza para otros fines, como entrenar sus propios modelos.

Controla los permisos: desconfía de aplicaciones que pidan acceso a contactos, ubicación o archivos que no necesitan para funcionar. Una app de edición de fotos con IA, por ejemplo, no necesita acceso a tu micrófono.

Actualizaciones y soporte: usa herramientas que se actualicen con frecuencia y ofrezcan algún canal para resolver dudas o incidencias.

Vigila el uso de la IA en otras aplicaciones: en redes sociales como Facebook e Instagram, Meta puede utilizar las fotos y publicaciones públicas de tu perfil para entrenar sus propios modelos de IA, salvo que te opongas expresamente desde los ajustes de privacidad de cada red.

Para evitar que tus fotos se usen de esta forma, lo más recomendable es tener el perfil en privado; así, solo las personas que apruebes como seguidoras podrán ver tu contenido.

Revisa esta opción en la configuración de cada red social que utilices, ya que el proceso varía de una a otra.

Clave 5: Detecta contenido falso generado con IA

Los ciberdelincuentes también usan la Inteligencia Artificial para generar **mensajes, documentos y contenidos con un nivel de realismo muy convincente**, con los que intentan apropiarse de tu dinero o de tus datos.

Para evitarlo, te presentamos las formas más habituales, y te damos claves para saber cómo reaccionar ante cada una:

Correos fraudulentos (phishing hiperpersonalizado)



Son correos muy personalizados, generados con Inteligencia Artificial con un estilo natural y sin errores gramaticales, que incluyen un enlace a una web falsa o te piden información personal.

Ante la duda, no uses el enlace del mensaje para acceder al contenido: llama primero al teléfono oficial de la entidad que supuestamente te ha contactado para comprobar si el mensaje es real.

Facturas falsificadas



Son documentos casi idénticos a los originales, generados con IA para parecer legítimos. Revisa los detalles con calma y llama a la entidad que figura como emisora de la factura para confirmar que el documento es auténtico antes de hacer ningún pago.

Chatbots fraudulentos



Puedes tener conversaciones con chatbots muy creíbles, con las que intentan obtener información sensible o dinero. Mantente alerta y confirma siempre por otra vía antes de compartir tus datos.

Páginas webs falsas



Existen en Internet tiendas online muy realistas que venden productos que nunca llegan, con el objetivo de quedarse con tu dinero y tus datos. Antes de comprar, busca el nombre de la tienda en otras fuentes y, si no encuentras información ni opiniones de otros compradores, desconfía.

Audios y vídeos (deepfakes)

Otro riesgo son los audios y vídeos manipulados con IA para suplantar la identidad de una persona, conocidos como deepfakes. Si algo de lo que ves o escuchas te genera dudas —una videollamada, una nota de voz, una noticia con imágenes llamativas— fíjate en algunas señales antes de darlo por real:

- Bordes de la cara extraños o parpadeos irregulares.
- Cambios de iluminación raros o cortes bruscos en el audio.
- Desfase entre el movimiento de los labios y lo que se escucha.
- Un tono de urgencia injustificada, típico de los intentos de engaño.

También puede ocurrir lo contrario: un contenido no te genera ninguna duda porque mucha gente lo comparte y coincide con lo que ya piensas, así que lo das por real sin cuestionarlo. Para no caer en esto:

- **No te fíes de tus impulsos:** si un contenido te provoca una reacción emocional fuerte —odio, indignación, miedo— y además te incita a compartirlo con urgencia, tómate un segundo antes de hacerlo.
- **No lo compartas de inmediato:** analízalo con calma. Este tipo de contenido suele estar diseñado precisamente para influir en la opinión pública o generar alarma.



Clave 6: Genera contenido con responsabilidad

Los **generadores de imágenes con IA** te permiten crear ilustraciones, logotipos o fotografías realistas en cuestión de segundos con solo escribir una descripción.

Es una posibilidad muy útil, pero conviene usarla con responsabilidad:

Uso ético: aprovecha estas herramientas para potenciar tu creatividad, ahorrar tiempo o mejorar tu trabajo, nunca para engañar, manipular o suplantar el trabajo de otras personas.

Frena la desinformación: evita generar o difundir imágenes falsas que puedan manipular opiniones o crear confusión, especialmente en redes sociales.

Pide permiso antes de usar la foto de otra persona: cada vez es más frecuente compartir con una herramienta de IA una foto para que la retoque de algún modo. Por ejemplo, convirtiéndote en un personaje de dibujos animados, o colocándote junto a una persona famosa. Si vas a subir una foto en la que aparecen otras personas —familiares, amistades, compañeros—, pídeles permiso antes de hacerlo.

Retos virales en redes sociales de generar imágenes con IA



Estos retos parecen inofensivos, pero tienen un coste que no siempre es visible. Es lo que ocurrió con el reto de convertir fotos personales en ilustraciones al estilo de Studio Ghibli o un selfie tuyo si hubieras vivido en los 80's: millones de personas subieron su rostro a herramientas de IA sin plantearse qué pasaba con esas imágenes después.

Al compartir tu foto, o la de alguien conocido, con una herramienta de IA gratuita, estás entregando algo más que una imagen: le das acceso a la biometría de tu cara, un dato que, a diferencia de una contraseña, no puedes cambiar si se ve comprometido.

Además, pierdes el control sobre esa imagen una vez subida, que podría acabar usándose para entrenar otros modelos o, en el peor de los casos, en manos de ciberdelincuentes que la utilicen para crear contenido falso.

Antes de sumarte a uno de estos retos, comprueba en la política de privacidad de la herramienta si tus fotos se usan para entrenar sus modelos de IA generativa, y valora si realmente compensa el riesgo.

Aprende más

En Andalucía Vuela te acompañamos a aprender a usar cualquier herramienta del mundo digital de forma responsable y segura:

Aprende Volando

Cursos gratuitos de 10 horas para mejorar tu nivel digital sin importar tu punto de partida o tu edad.

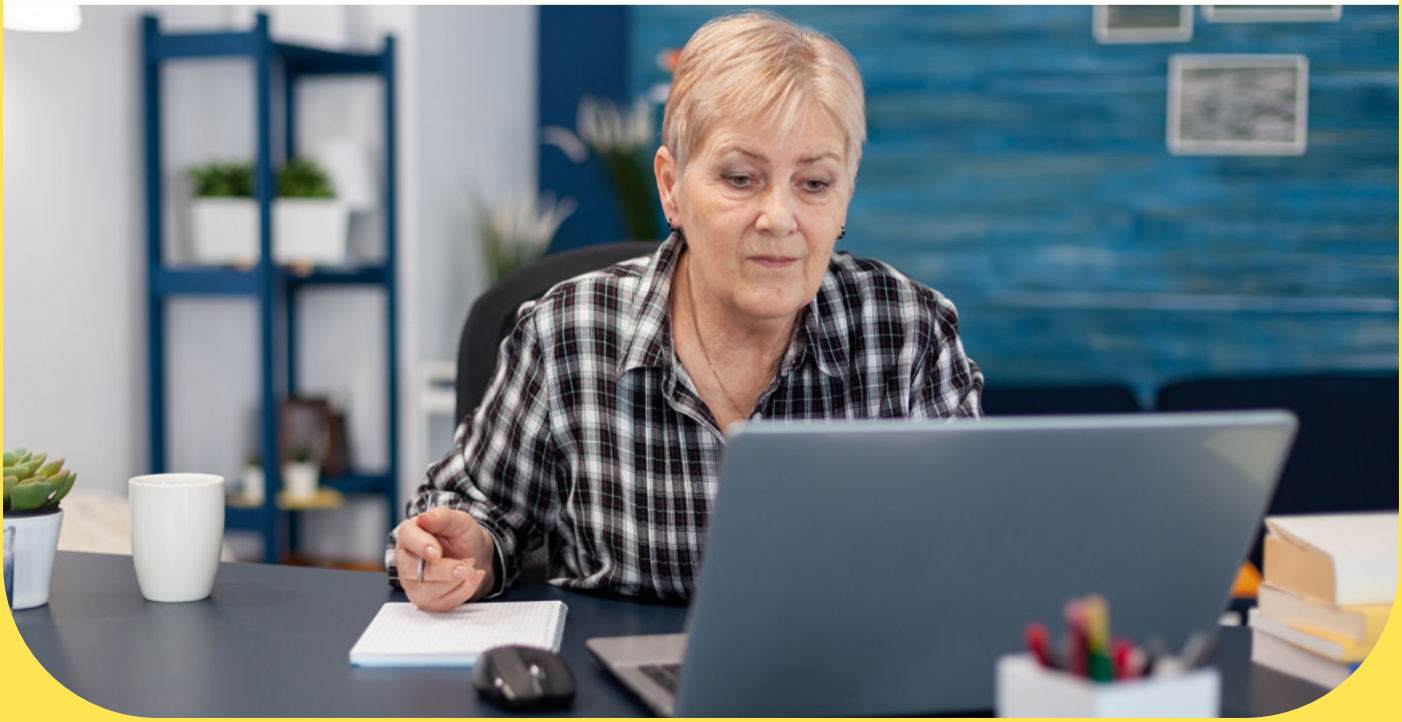
Más información aquí

Puntos Vuela

Más de 760 centros abiertos a todas las edades y repartidos por toda Andalucía para explorar el mundo digital y aprender competencias digitales con apoyo personalizado.

Más información aquí

**Guía generada a partir de la guía de ciberseguridad "E-learning. Inteligencia artificial y ciberseguridad" del Centro de Ciberseguridad de Andalucía. Este documento forma parte del Proyecto RED ARGOS, integrado en el programa de Redes Inteligentes de Especialización Tecnológica (RETECH) y está desarrollado en el marco del Plan de Recuperación, Transformación y Resiliencia, y financiadas por la Unión Europea NextGeneration-EU, a través de INCIBE*



vuela

Más cerca, más digitales, más oportunidades

andaluciavuela.es



Junta de Andalucía

Agencia Digital de Andalucía