

# La Doble Cara de la IA: Promesa y Peligro para Nuestros Jóvenes

Navegando la era de la inteligencia artificial en la educación y la vida diaria.



La inteligencia artificial (IA) está transformando la manera en que aprendemos, creamos e interactuamos. Para los jóvenes, representa un universo de oportunidades sin precedentes. Sin embargo, esta misma tecnología introduce riesgos complejos y a menudo ocultos que amenazan su seguridad, privacidad y bienestar. Esta presentación explora ambos lados de la moneda, ofreciendo un mapa para guiar, no para prohibir, a la nueva generación en el ecosistema digital.



# La IA ya está aquí: El Ecosistema Digital de los Adolescentes

La IA no es una tecnología del futuro; es una realidad cotidiana para los jóvenes. Desde asistentes virtuales hasta plataformas educativas, su presencia es constante. Comprender su nivel de adopción es el primer paso para una orientación eficaz.

## Encuesta en EE. UU. (14-22 años)

**51%** ha utilizado herramientas de IA generativa.

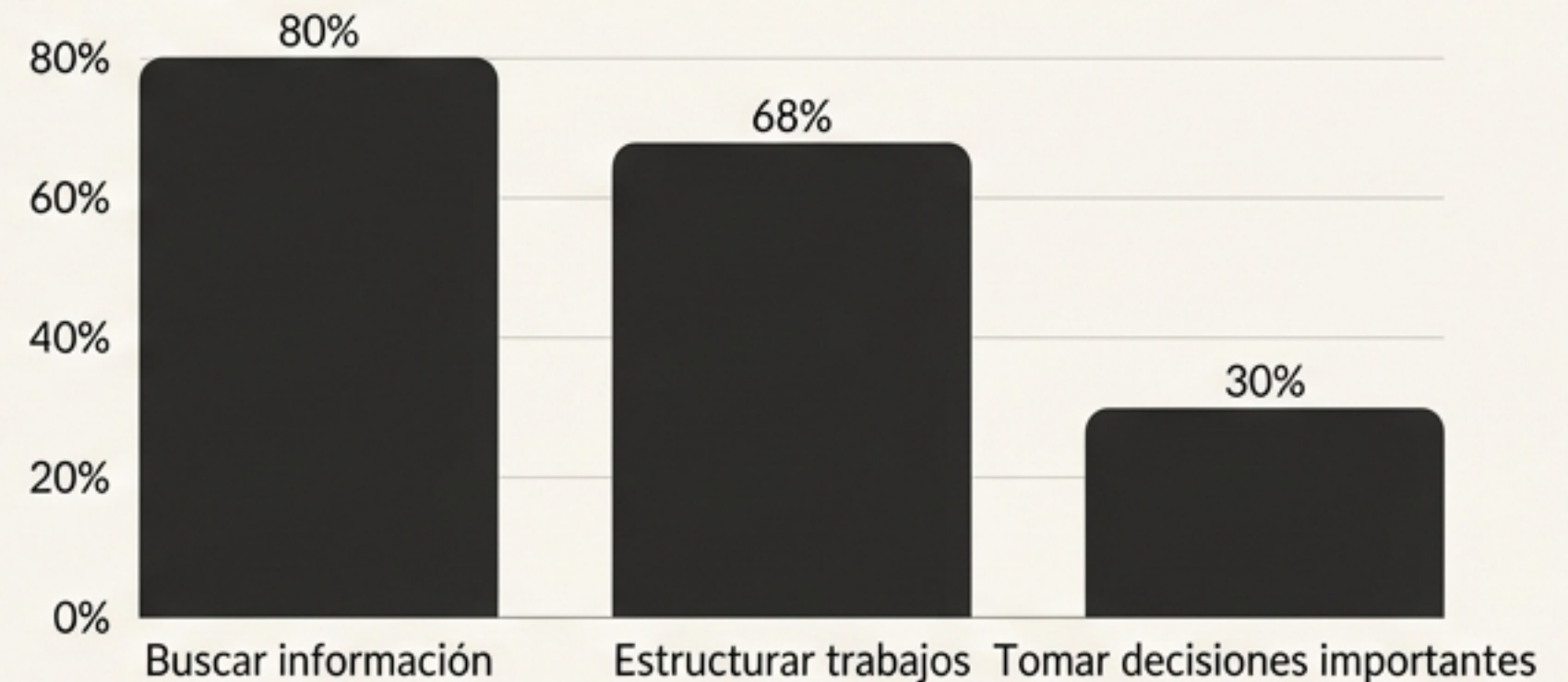
**4%** las usa a diario.

## Encuesta en España (14-18 años)

**15%** la usa a diario.

**51%** la usa varias veces por semana.

## ¿Para qué usan la IA los adolescentes en España?



*“El rápido auge de la Inteligencia Artificial Generativa (IAG) está transformando la manera en la que la infancia y la adolescencia interactúa con los medios digitales”. - Adolescere SEMA*



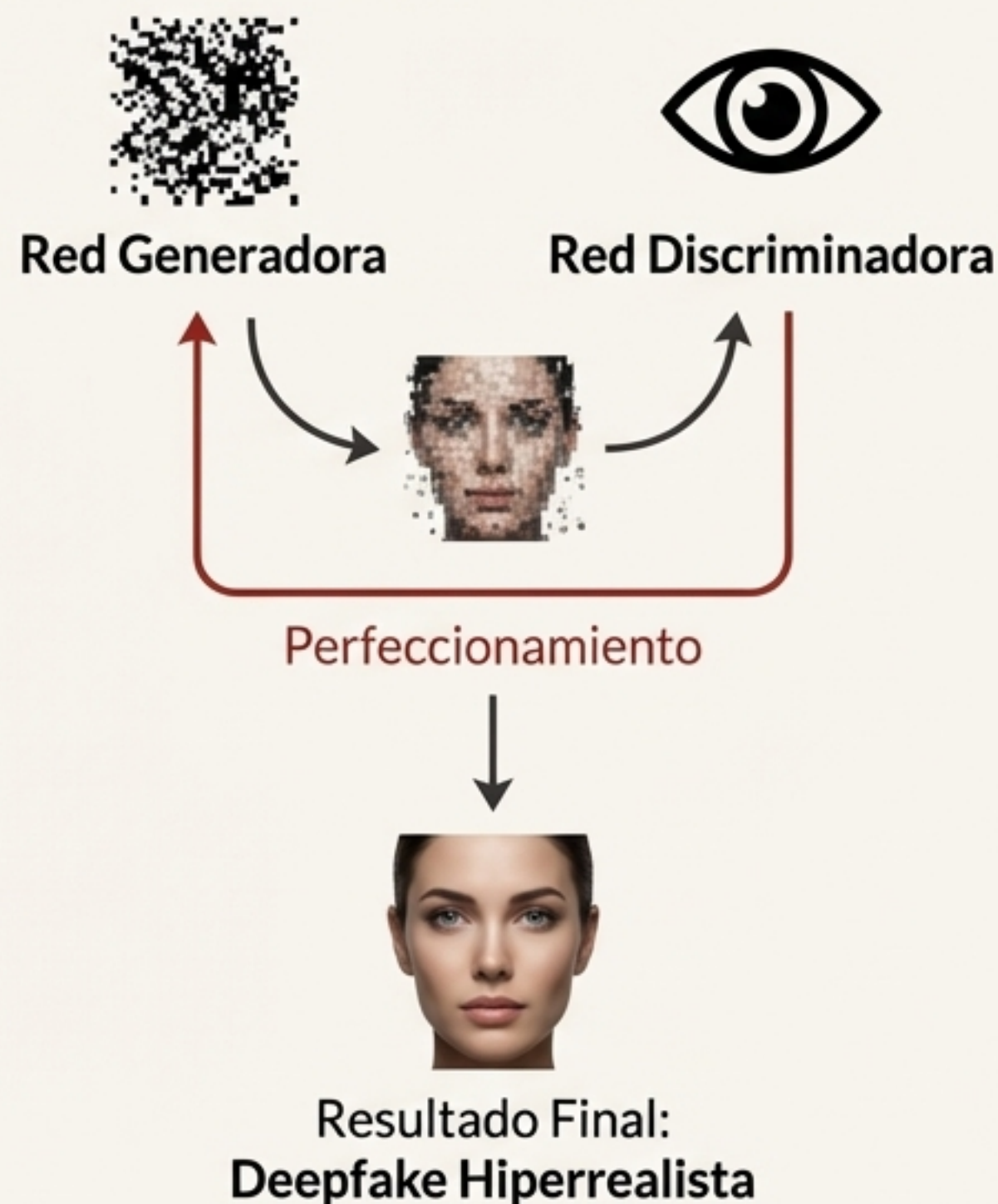
# La Crisis de la Confianza Visual: ¿Qué es un Deepfake?

Un **deepfake** es contenido (video, audio o imagen) falso pero hiperrealista, generado por algoritmos de aprendizaje profundo que suplantando la apariencia o voz de una persona. La tecnología analiza grandes cantidades de datos de la víctima para crear una imitación convincente.

## ¿Cómo se crea?

- Dos redes neuronales compiten entre sí:
  1. Una **red generadora** que produce las imágenes sintéticas.
  2. Una **red discriminadora** que intenta diferenciar el contenido real del falso.
- Este ciclo se repite millones de veces, perfeccionando la falsificación hasta que se vuelve indistinguible de la realidad.

Los deepfakes rompen la confianza en el mundo de la realidad. Los deepfakes rompen la confianza fundamental en lo que vemos y oímos. Si el video y el audio ya no son pruebas fiables, se erosiona la base de la comunicación y la verdad.





# El Impacto Real de los Deepfakes: Fraude, Acoso y Desinformación

Los *deepfakes* ya no son un experimento de laboratorio. Son herramientas utilizadas activamente para cometer delitos con consecuencias devastadoras.



## Fraude Financiero Corporativo (Ataque Dirigido)

**Caso:** Un empleado de la multinacional Arup fue engañado para transferir 25 millones de dólares tras participar en una videollamada donde el director financiero y otros ejecutivos eran recreaciones deepfake.

**Impacto:** Pérdida financiera masiva, vulnerabilidad crítica para la seguridad empresarial.



## Acoso y Daño Reputacional (Ataque Masivo)

**Caso:** Un director de atletismo en una escuela de EE. UU. creó un audio deepfake racista y antisemita suplantando la voz del director del colegio para desacreditarlo. El audio se volvió viral.

**Impacto:** Amenazas de muerte, división en la comunidad y despido del profesor suplantado.



## Manipulación Política y Desinformación

**Caso:** Se difundió una llamada automática con una voz deepfake de un político estilizado (evitando realismo fotográfico) para disuadir a los votantes de participar en las primarias de New Hampshire.

**Impacto:** Fraude electoral, erosión de la confianza en los procesos democráticos.

## Riesgo para los Jóvenes

Creación de deepfakes con fines de acoso, ciberacoso potenciado y explotación sexual.



# El Espejo Roto: Cuando la IA Aprende Nuestros Prejuicios

**El sesgo algorítmico** describe errores sistemáticos y repetidos que crean resultados injustos, privilegiando a un grupo de usuarios sobre otros. No es un error matemático, sino la digitalización de los sesgos sociales e históricos preexistentes.



**La Causa Principal:** La calidad y representatividad de los datos de entrenamiento.

Si los datos de entrenamiento tergiversan la demografía real o si ciertos grupos están insuficientemente representados, el algoritmo no podrá tomar decisiones precisas o justas para esos grupos.



**Resultado Sesgado**

“Si los humanos son imperfectos, y si los datos que creamos y usamos para entrenar a las máquinas también son imperfectos, entonces la IA que aprende de ellos no será neutra ni infalible”. – Guía Curricular Didáctica Avanzada



# El Costo Humano del Sesgo: Justicia y Salud en la Era Algorítmica

Las decisiones algorítmicas sesgadas ya están afectando áreas críticas de la sociedad, perpetuando la desigualdad bajo una apariencia de objetividad matemática.



## Justicia Penal y Predicción de Reincidencia

- **Sistema:** El software **COMPAS**, utilizado en el sistema judicial de EE. UU. para predecir la probabilidad de que un acusado reincida.
- **El Sesgo:** Una investigación de ProPublica reveló que el algoritmo etiquetó incorrectamente a los acusados negros como de “alto riesgo” casi el doble de veces que a los acusados blancos (45% vs. 23%).
- **Consecuencia:** Refuerzo de disparidades raciales en las sentencias y audiencias de libertad condicional.



## Sistemas de Salud y Asignación de Cuidados

- **Sistema:** Un algoritmo comercial utilizado en hospitales de EE. UU. para identificar pacientes que necesitan programas de “gestión de cuidados de alto riesgo”.
- **El Sesgo:** Utilizaba el gasto sanitario pasado como indicador de necesidad futura. Debido a desigualdades sistémicas, los pacientes negros gastan menos que los blancos con las mismas condiciones, por lo que el algoritmo subestimaba drásticamente sus necesidades de salud.
- **Consecuencia:** Discriminación en el acceso a cuidados médicos, perpetuando desigualdades en la salud.



# La Huella Digital Permanente: Privacidad y Seguridad en el Aula



“¿Sabemos realmente qué datos están siendo recolectados por las herramientas que usamos en nuestras aulas?” - *\*Podcast ‘Después de Clase’\**”

## Los Riesgos

- **Recopilación de Datos Sensibles:** Muchas plataformas de IA recopilan información personal como calificaciones, comportamientos en clase e incluso datos sobre el bienestar emocional.
- **Políticas de Datos Poco Claras:** Herramientas diseñadas para adultos a menudo ingresan a las escuelas sin políticas de tratamiento de datos adecuadas para menores.
- **Consentimiento Involuntario:** Los jóvenes pueden ser presionados por la IA para revelar información personal o íntima en interacciones que parecen seguras.
- **Uso de Datos para Entrenamiento:** Existe una preocupación creciente por el uso de imágenes o datos personales de menores sin el consentimiento adecuado para entrenar futuros modelos de IA.

**\*\*El Impacto para los Jóvenes\*\*:** La creación de una “huella digital” que los acompaña de por vida, exponiéndolos a riesgos de manipulación, filtraciones de datos y vigilancia.



La Solución: Hacia una Tutoría Digital Consciente

# Nuestro Rol es Guiar, no Prohibir

Ante la complejidad de la IA, la prohibición no es una solución sostenible. Nuestra responsabilidad como educadores, padres y tutores es convertirnos en **“tutores digitales”**: guías proactivos que empoderan a los jóvenes con las herramientas para navegar el entorno digital de manera segura, ética y crítica.

## Los Tres Pilares de la Tutoría Digital



### 1. Educar

Fomentar el pensamiento crítico y la alfabetización mediática.



### 2. Proteger

Implementar salvaguardas técnicas y hábitos seguros.



### 3. Colaborar

Construir una cultura de confianza y comunicación abierta.



# Pilar 1: Educar para un Pensamiento Crítico

La primera línea de defensa no es un filtro de software, sino una mente crítica. Debemos enseñar a los jóvenes a cuestionar, verificar y comprender la tecnología que utilizan.

## Estrategias Clave

- **Verificar la Información Siempre:** Enseñar que la información generada por IA no es infalible. La recomendación clave de la Guardia Nacional es **verificar la información con otras fuentes** antes de aceptarla o compartirla. No usarla como sustituto, sino como complemento del aprendizaje.
- **Educar sobre Privacidad Digital:** Incluir sesiones sobre cómo identificar señales de peligro en línea, evitar compartir información sensible y configurar la privacidad en redes sociales.
- **Fomentar la Alfabetización Mediática:** Desarrollar la habilidad de distinguir entre contenido auténtico y manipulado (como los *deepfakes*). Enseñar a analizar inconsistencias. Enseñar a analizar inconsistencias en videos (parpadeo, reflejos, sombras) y a ser escépticos ante contenido demasiado perfecto o provocador.
- **Enseñar sobre Ética y Consentimiento:** Normalizar la discusión sobre el consentimiento en interacciones digitales, el plagio y el uso responsable de la IA para no dañar a otros.

# 50%

de los adolescentes en España siente que su escuela no les prepara para usar la IA de forma ética.

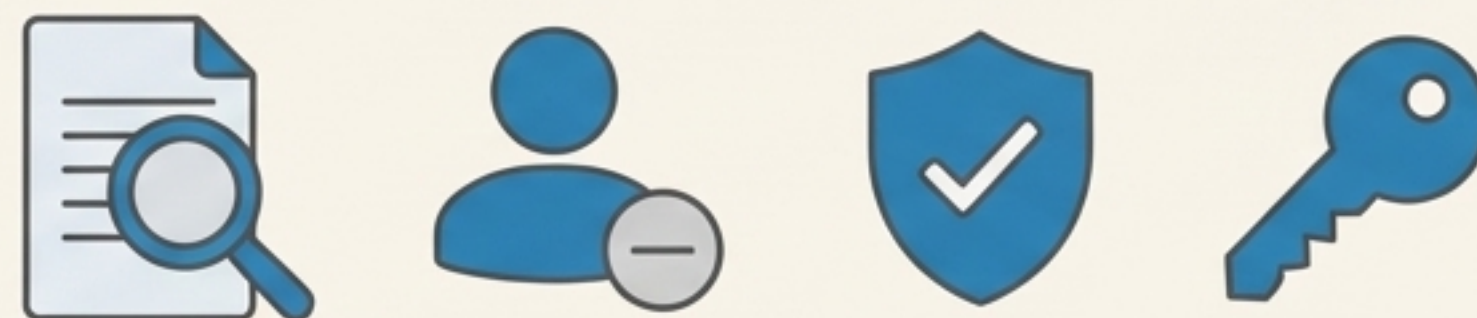


# Pilar 2: Proteger con Hábitos y Herramientas Seguras

La educación se complementa con **acciones concretas** para minimizar la exposición a riesgos y proteger la información personal de los estudiantes.

## Estrategias Clave

- **Leer y Entender las Políticas de Privacidad:** Antes de usar cualquier herramienta de IA, es crucial revisar sus términos.
- **Minimizar la Cantidad de Datos Compartidos:** Limitar la información que se comparte. Por ejemplo, crear cuentas de usuario anónimas para los estudiantes en lugar de usar sus nombres reales.
- **Utilizar Herramientas Seguras y Controles Parentales:** Elegir plataformas que usen encriptación y ofrezcan autenticación en dos pasos.
- **Generar Contraseñas Seguras:** Enseñar la importancia de crear contraseñas robustas (símbolos, mayúsculas, minúsculas, números) y cambiarlas periódicamente (cada 3-6 meses).



## Ejemplo Práctico

OpenAI recomienda el uso de ChatGPT a mayores de 18 años, o a partir de 13 con consentimiento parental. Usarlo en el colegio sin autorización envía un mensaje equivocado.

Utilizar herramientas como **Google Family Link** para gestionar el tiempo de uso, las aplicaciones permitidas y restringir contenido.

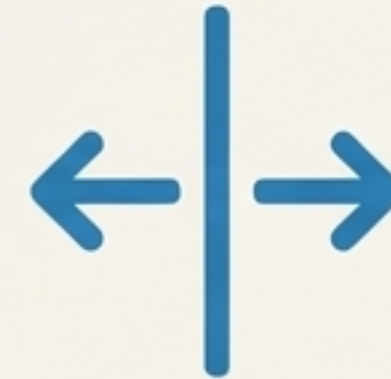


# Pilar 3: Colaborar para Construir una Cultura de Confianza

La tecnología no debe ser un campo de batalla entre generaciones. La protección más eficaz nace de la colaboración, la transparencia y una comunicación abierta y honesta.

## Estrategias Clave

- **Fomentar una Cultura de Consentimiento Informado:** Antes de implementar una nueva tecnología de IA en clase, organizar reuniones con los padres para explicar sus beneficios y riesgos, asegurando que todos estén de acuerdo.
- **Crear Canales de Comunicación Abiertos:** Los jóvenes deben saber que pueden reportar cualquier cosa que comprometa su integridad o seguridad sin temor a represalias.
- **Promover el Aprendizaje Intergeneracional:** Fomentar un diálogo donde los adultos no solo enseñan, sino que también aprenden de las experiencias digitales de los jóvenes.



Un estudio reveló que el 70% de los profesores pone límites al uso de la IA, pero 6 de cada 10 alumnos reconocen haberlos ignorado. Esto evidencia una brecha que solo la confianza puede cerrar.



El 45% de los padres que ha hablado sobre IA con sus hijos reconoce que **le ha servido para aprender más.**



# Un Vistazo Profundo a los Riesgos de la IA Generativa para los Jóvenes

Un análisis empírico de interacciones de jóvenes con IA Generativa identificó 84 riesgos específicos, clasificados en seis áreas clave del desarrollo individual.

| Bienestar Mental                                                                                                                                                                                                       | Desarrollo Conductual y Social                                                                                                                                                                                                              | Interacción Tóxica con la IA                                                                                                                                        | Mal Uso (Interpersonal)                                                                                                                                                                               | Sesgos y Discriminación                                                                                                                                      | Privacidad                                                                                                                                                                                              |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"><li>• Dependencia emocional de la IAG.</li><li>• Aumento de vulnerabilidades (depresión, ansiedad).</li><li>• Dificultad para diferenciar experiencias virtuales y reales.</li></ul> | <ul style="list-style-type: none"><li>• Refuerzo de conductas negativas (manipulación, agresión).</li><li>• Aislamiento social progresivo.</li><li>• Normalización de la violación del consentimiento en interacciones simuladas.</li></ul> | <ul style="list-style-type: none"><li>• Generación de insultos o lenguaje violento sin provocación.</li><li>• Producción de contenido sexual inapropiado.</li></ul> | <ul style="list-style-type: none"><li>• Creación de deepfakes para acoso o explotación.</li><li>• Difusión de desinformación.</li><li>• Ciberacoso potenciado por herramientas generativas.</li></ul> | <ul style="list-style-type: none"><li>• Refuerzo de estereotipos negativos (género, raza).</li><li>• Generación de contenido con discurso de odio.</li></ul> | <ul style="list-style-type: none"><li>• Uso de datos personales de menores sin consentimiento.</li><li>• Presión de la IA para revelar información sensible.</li><li>• Filtraciones de datos.</li></ul> |



# La Alfabetización Ética: Nuestra Mayor Defensa

Las vulnerabilidades más significativas de la IA no son **fallos técnicos**, sino manipulaciones de la verdad y fallas en la equidad. Por lo tanto, nuestra principal línea de defensa no es la tecnología misma, sino la **alfabetización ética** de **los usuarios**. El **objetivo** es transformar a los jóvenes de meros consumidores de tecnología a creadores de las reglas y estándares éticos que gobernarán el futuro digital.

## Llamado a la Acción

- **Concientizar, Dialogar, Guiar:** Fomente conversaciones sobre los riesgos y beneficios de la IA en casa y en el aula.
- **Reportar:** Si usted o alguien que conoce ha sido víctima de un delito digital, no está solo.



**Recurso Clave: Guardia Nacional de México**

Línea de Atención Ciudadana: **088**

Correo de Reporte: **cnac@gn.gob.mx**

**Recomendación:** Guarde siempre capturas de pantalla y evidencia digital antes de reportar.