

# Desafíos éticos

En este documento vamos a profundizar más en el tema de los desafíos éticos de Cegados por la IA.

<b>Desafíos éticos</b>	<b>1</b>
Profundizar riesgos	1
Posturas “macro”	2
Personalización	3
Referencias	4

## Profundizar riesgos

Vamos a exponer algunos riesgos posibles que potencialmente podría tener nuestro sistema si fuera usado de forma masiva en todo el mundo.

### 1. Errores sistemáticos de percepción (millones de descripciones incorrectas)

Incluso un 1% de error (que sería un resultado bueno para visión artificial) implicaría millones de descripciones mal generadas cada día.

Problemas que puede generar:

- **Confundir objetos** peligrosos, p. ej., no detectar un coche o describirlo demasiado tarde.
- **Describir mal** señales, precios, medicamentos, alimentos, causando daño en la vida cotidiana.
- **Sesgos** sistemáticos: reconocimiento peor de rostros, objetos o entornos propios de ciertas culturas o regiones.
- **Falsas sensaciones de confianza**, llevando a la persona usuaria a arriesgarse más de lo que debería.

Colectivos afectados:

- Personas con discapacidad visual: impacto directo en seguridad y autonomía.
- Aseguradoras y sistemas sanitarios: aumento de accidentes en usuarios.

### 2. Dependencia excesiva del sistema

Si cientos de millones de personas dependen de Cegados por la IA, un fallo global podría paralizar a toda una población.

Problemas que puede generar:

- **Caídas del servidor**: millones de personas pierden su “percepción asistida”.
- **Actualización defectuosa**: “ceguera digital” simultánea.
- **Ataques** que incapaciten el servicio.

### 3. Riesgos de privacidad masivos

Aunque solo capture vídeo, la escala multiplica el problema: millones de cámaras andando por el mundo, grabando espacios públicos y privados.

Problemas que puede generar:

- **Exposición accidental** de escenas íntimas de terceros.
- **Reconocimiento involuntario** de personas: riesgo de trazabilidad masiva.
- Datos extremadamente sensibles si alguien consigue acceso (hacking, **filtraciones**).

Colectivos afectados:

- Ciudadanía general, incluso quienes no usan el producto.
- Negocios, escuelas, hospitales: espacios donde la grabación constante es un riesgo.

### 4. Problemas legales y regulatorios globales

Hasta ahora hemos especificado nuestro sistema teniendo en cuenta las leyes españolas y europeas, pero si se expande su uso a nivel mundial puede haber otras leyes que no hemos tenido en cuenta.

Problemas que puede generar:

- **Legislaciones diferentes** (UE, EE.UU., India, etc.) entrarían en choque.
- Obligación de borrar datos, exigir consentimiento, informar a terceros de grabación...

## Posturas “macro”

Postura de cada miembro del grupo:

- Diego Alonso Arceiz - Team Advocate: **safetyist**
- Carmen Fernández González - Product Owner: **skeptic**
- Ignacio Gutiérrez Sánchez - Scrum Master: **skeptic**
- Bryan Xavier Quilumba Farinango - Experto Git: **accelerationist**

En nuestro grupo convivimos perfiles más **accelerationist**, que ven en *Cegados por la IA* una herramienta de autonomía y progreso social, y otros más **safetyist**, preocupados por los riesgos de sesgo, dependencia y abuso a gran escala; también hay **skeptic** que rechazan teorías apocalípticas y recalcan que el impacto real no será existencial sino socioeconómico -fraudes, impacto sobre colectivos vulnerables, y medioambiental-.

Como empresa adoptamos una postura **intermedia entre escépticos y safetyistas**, priorizando beneficios concretos para personas con discapacidad visual, pero con fuertes límites técnicos, controles de uso y diseño centrado en la privacidad y neutralidad. Esto se traduce en decisiones como **evitar personalización ideológica, no inferir atributos sensibles, garantizar robustez frente a mal uso y mantener un proceso continuo de auditoría y evaluación ética**.

# Personalización

En Cegados por la IA una personalización excesiva podría llevar a:

- Interpretar la realidad de forma sesgada.
- Reforzar valores problemáticos (prejuicios).
- Aumentar los riesgos de dependencia o manipulación.
- Crear desigualdad entre usuarios con outputs radicalmente diferentes.

Si se permite **personalización profunda**, el sistema corre el **riesgo de convertirse en un filtro ideológico y comportamental** de la realidad, amplificando sesgos, dependencias y riesgos sociales. Si se **limita** demasiado, puede ser **percibido como rígido o poco útil**.

El equilibrio razonable para “Cegados por la IA” es:

- Personalización funcional → SÍ.
- Personalización interpretativa o ideológica → NO.

La personalización de nuestro sistema se hará a través de la **app móvil**, por medio de un menú de ajustes que estará definido por nosotros, por tanto las opciones de personalización estarán limitadas.

## **Límite 1:** Personalización solo superficial

- El sistema NUNCA adapta la interpretación de la realidad.
- Solo adapta:
  - Velocidad
  - Nivel de detalle
  - Orden de la información
  - Preferencias contextuales

## **Límite 2:** Prohibir personalización ideológica, moral, identitaria o discriminatoria.

- El usuario no puede pedir:
  - Describir a las personas de manera distinta según rasgos sensibles
  - Ocultar información que pueda ser relevante para su seguridad
  - Enfatizar prejuicios o valoraciones subjetivas

## **Límite 3:** Modelos que no guardan historial sensible

- No se almacena:
  - Comportamiento en espacios públicos
  - Personas que encuentra habitualmente
  - Rutas de desplazamiento
  - Objetos que posee o usa

## **Límite 4:** El modelo no hace juicios de valor

- Nunca debe decir cosas como:
  - “Esa persona parece peligrosa / poco confiable”
  - “Este lugar parece de mala categoría”
  - “Esa persona parece triste / nerviosa / agresiva”

## **Límite 5:** Política explícita de “seguridad del tercero no usuario”

La personalización nunca puede perjudicar a terceros grabados involuntariamente.

## Referencias

Conversación IA - <https://chatgpt.com/share/691b5e80-cad4-8002-a235-abd0b71905f7>

[AI Triad: A Dialogue Across Differences](#)

[The Age of AI Anxiety — and the Hope of Democratic Resistance](#)