

Ampliación del Proyecto “Cegados por la IA” – Acta de Debate y Respuestas

1. Introducción

Tras el debate mantenido en clase, hemos revisado y reforzado los puntos críticos del proyecto con el objetivo de resolver las dudas y objeciones planteadas, garantizando la solidez técnica, ética y regulatoria del sistema.

Este documento amplía la especificación del Sprint 1 y actúa como plan de mejora para futuras iteraciones.

2. Principales cuestiones debatidas y respuestas reforzadas

2.1. Privacidad ilusoria: ¿Tratar pero no almacenar?

Crítica recibida:

“Según algunos artículos, procesar sin consentimiento sigue sin justificar la grabación. ¿Cómo garantizáis la privacidad real de terceros?”

Respuesta reforzada:

El sistema implementa un proceso de anonimización previo al análisis. En cuanto la cámara capta una imagen:

Se aplica difuminado automático de rostros y matrículas antes de cualquier inferencia.

No existe transmisión ni almacenamiento, ya que el procesamiento es local (Edge AI).

Las imágenes son descartadas tras la generación del audio.

La base jurídica es el art. 9.2.g del GDPR, que permite el tratamiento de datos sensibles con fines de accesibilidad. Además, se añade una opción de modo privacidad reforzada, donde la cámara se apaga automáticamente en entornos sensibles (baños, vestuarios, hospitalares).

2.2. Procesamiento “on the edge” sin GPU ni nube

Crítica recibida:

“¿Cómo se puede procesar vídeo en tiempo real sin conexión a la nube ni GPU? ¿Es técnicamente viable?”

Respuesta reforzada:

Para 2029 se espera que la miniaturización de hardware y la computación en el borde (Edge AI) estén completamente maduras.

El sistema se basará en:

Chips neuromórficos optimizados para IA multimodal, capaces de ejecutar modelos compactos sin disipar calor excesivo.

Modelos ligeros preentrenados externamente y luego desplegados en local (similar al enfoque de Gemini Nano o Apple Neural Engine).

Procesamiento distribuido: el chip solo ejecuta inferencia (no entrenamiento).

De esta forma, se garantiza el procesamiento sin conexión y sin requerir GPU dedicada.

2.3. Reconocimiento fallido o incierto de objetos

Crítica recibida:

“¿Qué hace el sistema si no identifica correctamente un objeto o lo confunde?”

Respuesta reforzada:

El modelo aplica un umbral de confianza.

Si la probabilidad de acierto es baja, el sistema genera una descripción genérica (“objeto desconocido delante”).

No se inventan respuestas ni se asume información incierta.

En casos de duda, el sistema verbaliza su incertidumbre (“no estoy seguro de qué es eso”).

Además, se incluirá una fase de test de validación en entorno real con personas ciegas y profesionales de accesibilidad visual para afinar el comportamiento ante ambigüedades.

2.4. Sustitución de perros guía y dependencia tecnológica

Crítica recibida:

“¿No puede sustituir a los perros guía y generar dependencia del usuario?”

Respuesta reforzada:

El sistema se concibe como herramienta complementaria, no sustitutiva.

Se impartirá un curso gratuito de formación y concienciación (en colaboración con la ONCE) sobre el uso responsable.

Las gafas fomentan la autonomía consciente, manteniendo el bastón o el perro guía como soporte principal.

En caso de fallo, el usuario está preparado para interpretar y reaccionar adecuadamente.

2.5. Accesibilidad económica y riesgo de discriminación

Crítica recibida:

“¿No se generará una discriminación entre personas con recursos y sin ellos?”

Respuesta reforzada:

El equipo prevé establecer acuerdos con entidades como la ONCE, fundaciones o programas públicos para subvencionar el coste del dispositivo.

Además, se publicará el software bajo una licencia restrictiva de uso social, permitiendo su adopción por organizaciones sin ánimo de lucro.

2.6. Mejora del modelo sin recopilar datos

Crítica recibida:

“Si no se almacenan datos, ¿cómo se reentrena o mejora el modelo?”

Respuesta reforzada:

La mejora se realiza a través de reportes voluntarios, no mediante datos reales.

Los usuarios pueden reportar situaciones problemáticas (“no detectó un semáforo”) sin adjuntar imágenes.

Esos reportes se transforman en casos sintéticos de prueba para mejorar el modelo en entornos controlados.

De este modo, se garantiza la retroalimentación sin vulnerar la privacidad.

2.7. Supervisión humana y fiabilidad del sistema

Crítica recibida:

“¿No es falsa supervisión humana si el usuario confía ciegamente en las gafas?”

Respuesta reforzada:

La supervisión humana efectiva es requisito central del proyecto:

El usuario tiene control total: puede activar/desactivar cámara, elegir nivel de detalle y activar modo privado.

Se realizarán pruebas de validación previas para comprobar que el usuario sabe interpretar correctamente las descripciones.

El sistema informará al usuario cuando su nivel de confianza sea bajo (“no estoy seguro”).

Además, los reportes de error se revisarán manualmente por el equipo de desarrollo.

2.8. Ciclo de vida de los datos

Crítica recibida:

“No está claramente definido el ciclo de vida de los datos.”

Respuesta reforzada:

El modo de privacidad reforzada borra toda la información inmediatamente durante la sesión.

El ciclo completo se define ahora como sigue:

Fase	Dato tratado	Duración	Medidas aplicadas
Captura	Imagen de vídeo	Instantáneo	Difuminado previo, sin almacenamiento
Procesamiento	Frames de vídeo	<1 seg	Procesamiento local
Generación de audio	Descripción sintetizada	Volátil	Audio en memoria RAM
Fin de sesión	Configuración temporal	Hasta cierre sesión	Borrado automático y cifrado temporal

2.9. Reentrenamiento y nuevos objetos en la sociedad

Crítica recibida:

“¿Qué pasa si aparecen nuevos objetos o contextos que el modelo no reconoce?”

Respuesta reforzada:

El mantenimiento del sistema incluirá un plan de actualización anual:

Nuevas versiones del modelo entrenadas en entornos controlados con datos públicos y sintéticos.

Comunicación transparente a los usuarios sobre cada nueva actualización y sus mejoras.

Garantía de compatibilidad retroactiva del hardware (vía firmware).

Si un modelo deja de cumplir estándares de precisión, las gafas no se retirarán del mercado, pero recibirán una actualización que amplíe su base de conocimiento.

3. Conclusiones

Las observaciones del debate han permitido fortalecer la base ética, técnica y legal del proyecto.

“Cegados por la IA” se consolida como una solución de accesibilidad tecnológica responsable, basada en:

Privacidad por diseño y procesamiento local.

Supervisión humana y control del usuario.

Actualizaciones seguras sin comprometer la privacidad.

Colaboración con organizaciones sociales para garantizar la accesibilidad económica.

Con estas mejoras, el grupo considera que el proyecto está preparado para avanzar al siguiente sprint, centrado en la validación técnica y la evaluación de impacto ético.