Lista de verificación para la aplicación de principios de uso responsable de la IA en los servicios de seguridad

La siguiente lista de verificación y los estudios de caso posteriores se ofrecen como apoyo para comprender el potencial de la IA y gestionar cuestiones éticas relacionadas con su implementación. Son herramientas complementarias a la *Declaración sobre el uso responsable de la IA en los servicios de seguridad* de la International Security Ligue.

Gestión de Riesgos

❑ Garantizar el cumplimiento de la normativa aplicable durante todo el ciclo de vida del sistema de IA.
❑ Identificar y analizar los riesgos previsibles para la salud, la seguridad, los derechos fundamentales y la seguridad del cliente, incluidos los riesgos de uso indebido.
❑ Implementar medidas específicas para reducir los riesgos a un nivel aceptable.
❑ Establecer procedimientos de respaldo y medidas de mitigación para los riesgos que no se puedan eliminar.

Gobernanza de Datos

❑ Asegurar el cumplimiento total del RGPD y garantizar la seguridad cibernética y física de los conjuntos de datos.
❑ Utilizar datos fiables, sólidos y de alta calidad, garantizando la ausencia de sesgos en los datos de entrada.
❑ Asegurar la trazabilidad, explicabilidad, auditabilidad y rendición de cuentas en el procesamiento de datos y resultados de IA.

Supervisión Humana

❑ Ofrecer formación al personal para garantizar conocimientos adecuados sobre IA.
❑ Asignar responsabilidades proporcionales al caso de uso específico.
❑ Garantizar que la supervisión humana esté alineada con normas éticas.

Resiliencia

❑ Garantizar la precisión, robustez y ciberseguridad del sistema de IA a lo largo de su ciclo de vida.
❑ Abordar riesgos físicos (p. ej., control de acceso, eliminación segura, resiliencia de centros de datos).
❑ Mitigar riesgos cibernéticos (p. ej., evitar manipulación de datos, seguir normas de resiliencia digital).
❑ Establecer procedimientos de contingencia y planes de emergencia ante incidentes.

Registro de Datos

❑ Mantener registros automáticos durante al menos seis meses.
❑ Asegurar la trazabilidad, explicabilidad y rendición de cuentas del funcionamiento del sistema de IA.
❑ Documentar claramente las decisiones tomadas por la IA y asignar responsabilidades.

Transparencia y Explicabilidad

❑ Informar a las personas afectadas sobre interacciones con IA y ofrecer mecanismos de reclamación.
❑ Poner la documentación de IA a disposición de las autoridades para generar confianza pública.
❑ Ofrecer explicaciones claras sobre las decisiones de la IA y el uso previsto de los datos.

Evaluación del Impacto sobre los Derechos Fundamentales

❑ Realizar evaluaciones en contextos de uso por autoridades públicas, cubriendo riesgos, supervisión y gobernanza.
❑ Actualizar las evaluaciones si se producen cambios significativos o si los elementos quedan obsoletos.

Diligencia Debida

❑ Verificar que los sistemas de IA estén entrenados con datos de alta calidad, diversos y representativos.
❑ Garantizar el cumplimiento de leyes y reglamentos sobre ciberseguridad.
❑ Usar sistemas de IA transparentes con instrucciones claras sobre su finalidad, supervisión y métricas de riesgo.

Participación de los Trabajadores

❑ Informar e implicar a los trabajadores sobre el uso y finalidad de los sistemas de IA.
❑ Proporcionar formación y recursos para asegurar la comprensión de sus capacidades y limitaciones.
❑ Establecer canales protegidos por legislación laboral para plantear preocupaciones éticas.

Estudios de caso sobre implementaciones de IA en servicios de seguridad

A continuación, se presentan tres ejemplos reales (anonimizados) a nivel global que demuestran integraciones exitosas de IA en seguridad privada, así como los dilemas éticos que pueden surgir durante su implementación.

#1. Seguridad en Eventos **EUROPA**

Contexto: Un gran evento deportivo internacional en Europa enfrentaba el reto de gestionar multitudes, garantizar la seguridad y responder a amenazas dinámicas en múltiples sedes.

Uso de IA: La empresa de seguridad contratada desplegó analítica de video con IA integrada a la infraestructura de CCTV para detectar comportamientos anómalos en la multitud (p. ej., patrones de estampida, violaciones de perímetro, objetos abandonados).

Resultados:

- Se mejoraron los tiempos de respuesta en un 40%, permitiendo la intervención temprana en dos incidentes de aglomeración.
- El sistema emitió alertas en tiempo real al personal en tierra, evitando lesiones y perturbaciones.
- La anonimización de datos y la gobernanza ética garantizaron el cumplimiento del RGPD y la aceptación pública.

Consideración ética: El sistema fue configurado para evitar el reconocimiento biométrico, optando por análisis conductual no intrusivo — demostrando que seguridad y privacidad pueden coexistir.

#2. Priorización de Alarmas con IA en Seguridad Industrial **ASIA**

Contexto: Una empresa logística multinacional con más de 150 almacenes en Asia-Pacífico sufría frecuentes falsas alarmas, lo que causaba desensibilización y retrasos en la revisión manual.

Uso de IA: Se introdujo una IA para priorizar alarmas, analizando patrones de movimiento, anomalías por horario e integrándolos con datos operativos y meteorológicos (p. ej., horarios de montacargas).

Resultados:

- Las falsas alarmas se redujeron en un 85%, disminuyendo desplazamientos innecesarios y fatiga del personal.
- La detección de amenazas reales se agilizó (p. ej., un intento de intrusión fue frustrado en 4 minutos).
- El sistema utilizó pocos datos personales y se entrenó con parámetros operativos, reduciendo los riesgos éticos.

Beneficio para trabajadores: El personal de seguridad se centra ahora en la resolución activa en lugar de en la vigilancia repetitiva, mejorando la satisfacción y la calidad de las decisiones.

#3. Dilema Ético en el Uso de Reconocimiento Facial **LATINOAMÉRICA**

Contexto: Una promotora inmobiliaria privada contrató una agencia de seguridad para probar el reconocimiento facial como control de acceso en una urbanización cerrada en Sudamérica.

Uso de IA: El sistema se entrenó con imágenes extraídas de redes sociales para crear una base de datos de residentes, con el objetivo de evitar suplantaciones y entradas no autorizadas.

Dilema ético:

- Los residentes se quejaron al saber que sus imágenes se usaron sin consentimiento.
- El sistema cometió errores con visitantes de minorías étnicas, negando el acceso en tres ocasiones injustamente.
- Las autoridades abrieron una investigación preliminar por violación de la ley de protección de datos.

Resultado:

El despliegue fue suspendido y se reemplazó por un sistema híbrido de credencial + IA menos intrusivo. El caso resaltó la importancia del consentimiento informado, la mitigación del sesgo y la transparencia en el uso de biometría con IA.

Este documento tiene como único objetivo ayudar a los proveedores de servicios de seguridad a aplicar principios éticos en el uso de IA. No constituye asesoramiento legal y no debe considerarse como base para cumplir ninguna normativa específica. Ha sido elaborado por el grupo de trabajo *Tecnología y Sociedad* de la Ligue, compuesto por representantes de empresas miembros. Con autorización, incorpora elementos de la "Carta sobre el uso ético y responsable de la inteligencia artificial en los servicios de seguridad privada europeos" de CoESS. Es un documento complementario a la *Declaración sobre el uso responsable de la IA en los servicios de seguridad* de la International Security Ligue.