



Guía para el uso
seguro de la
**Inteligencia
Artificial**

Índice

01. Introducción.....	3
1. 1. Contexto.....	3
1. 2. Objetivo del documento	4
1. 3. Aclaraciones	4
02. Medidas para un uso seguro de la IA	5
2. 1. Paso 1: Regula el uso de la IA	6
2. 2. Paso 2: Inventaría los casos de uso de la IA	7
2. 4. Paso 4: Controla el uso de la IA.....	9
2. 5. Paso 5: Evalúa los resultados... y vuelve a empezar.....	10
03. Conclusiones	11

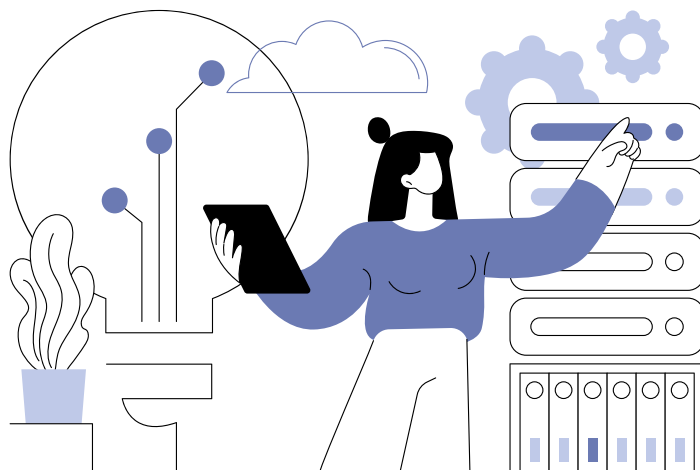


01.

Introducción

La inteligencia artificial se ha incorporado con rapidez a la vida cotidiana, al trabajo y a la relación de la ciudadanía con los servicios digitales. Herramientas capaces de generar textos, imágenes, código, resúmenes, recomendaciones o análisis de grandes volúmenes de información ofrecen oportunidades importantes para mejorar la productividad, la accesibilidad y la innovación. Sin embargo, su uso también introduce riesgos que deben conocerse y gestionarse de forma responsable.

Este documento tiene como finalidad ofrecer una referencia clara y práctica para utilizar la IA de forma segura. No pretende frenar su adopción ni presentar la tecnología como una amenaza inevitable, sino ayudar a distinguir entre usos adecuados e inadecuados, promover decisiones informadas y reducir los riesgos asociados a la privacidad, la seguridad de la información, la fiabilidad de los resultados, la propiedad intelectual, la dependencia tecnológica y el cumplimiento normativo.



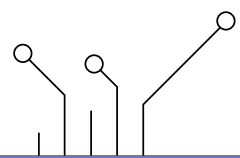
La seguridad en el uso de la IA no depende solo de la tecnología utilizada. Depende también de cómo se seleccionan las herramientas, qué datos se introducen, quién valida los resultados, qué controles existen, cómo se informa a las personas usuarias y cómo se revisan los riesgos a lo largo del tiempo. Por ello, el enfoque recomendado debe combinar medidas organizativas, técnicas y formativas, siempre con un lenguaje comprensible para cualquier persona usuaria.

1.1.

Contexto

La IA generativa y otros sistemas de IA permiten automatizar tareas, apoyar la toma de decisiones, mejorar la atención a personas usuarias, detectar patrones y acelerar procesos de análisis. En el ámbito público, empresarial y ciudadano, estas capacidades pueden generar valor si se aplican con proporcionalidad, transparencia y supervisión humana.

Al mismo tiempo, el uso de IA puede producir efectos indeseados si se emplea sin criterios claros. Entre los riesgos más relevantes destacan la introducción de datos personales o confidenciales en herramientas no autorizadas, la generación de respuestas incorrectas o inventadas, la pérdida de control sobre la información, la exposición a sesgos, la utilización de contenidos protegidos por derechos de terceros, el uso de servicios ubicados en jurisdicciones no adecuadas y la integración insegura de soluciones de IA en sistemas corporativos.



En este contexto, Cyberzaintza promueve un enfoque de uso seguro, responsable y proporcionado de la IA, orientado a que las personas y organizaciones puedan aprovechar sus beneficios sin comprometer la seguridad, la privacidad, la confianza ni los derechos de las personas afectadas. Este enfoque está basado en el modelo de ciberseguridad CyberOrratz, que establece los controles de seguridad aplicables para cada nivel de implantación de este, considerando los controles específicos para la Inteligencia Artificial.

1. 2.

Objetivo del documento

El objetivo de este documento es establecer criterios y medidas prácticas para orientar el uso seguro de la inteligencia artificial. Para ello, se describen las medidas mínimas que deberían tenerse en cuenta antes y durante el uso de herramientas o servicios basados en IA, tanto desde una perspectiva individual como organizativa.

El documento busca que cualquier persona pueda comprender qué precauciones debe adoptar al usar una herramienta de IA, qué criterios debe considerar para determinar si su uso es seguro y qué mecanismos debe adoptar para su control. Además, ofrece una base común para que las organizaciones puedan ordenar sus políticas internas, inventariar sus soluciones de IA, analizar riesgos, controlar usos no autorizados y gobernar su seguridad de manera continua.



1. 3.

Aclaraciones

Este documento ha sido elaborado parcialmente con la asistencia de herramientas de inteligencia artificial y revisado por Cyberzaintza.



02. Medidas para un uso seguro de la IA

El uso seguro de la IA exige pasar de recomendaciones genéricas a medidas concretas, verificables y adaptadas al contexto de cada organización. Las medidas descritas en este capítulo se orientan a construir un marco práctico: definir reglas de uso, conocer qué soluciones se utilizan, analizar sus riesgos, aplicar controles proporcionados y capacitar a las personas usuarias.

Estas medidas no deben entenderse como actuaciones aisladas. Funcionan de manera complementaria: una política de uso de IA resulta difícil de aplicar si no existe un inventario de herramientas; un inventario pierde valor si no se analizan los riesgos; y los controles técnicos son insuficientes si las personas usuarias no comprenden qué pueden hacer y qué deben evitar. El objetivo es crear un ciclo de gobierno y mejora continua del uso de la IA.

Las medidas aquí propuestas son aplicables al uso de sistemas, servicios, herramientas, modelos y funcionalidades de inteligencia artificial utilizados por una organización o por sus personas usuarias en el desarrollo de sus actividades. Incluye tanto soluciones propias como soluciones de terceros, herramientas gratuitas o de pago, servicios en la nube, funcionalidades integradas en aplicaciones corporativas y usos puntuales realizados desde equipos o redes de la organización.

El alcance comprende, como mínimo, los usos de IA que traten información corporativa, datos personales, documentación interna, código, credenciales, información de clientes, información de proveedores, datos sujetos a obligaciones legales o cualquier contenido cuyo uso indebido pueda impactar en la seguridad, privacidad, continuidad, reputación o cumplimiento normativo de la organización.

Estas medidas se resumen, esquemáticamente, de la siguiente forma:



1

REGULA EL USO
DE LA IA

2. 1. Paso 1: Regula el uso de la IA



La finalidad de este primer paso es establecer reglas claras sobre qué usos de la inteligencia artificial están permitidos, cuáles requieren autorización y cuáles deben quedar prohibidos. Una política de uso de IA permite que las personas usuarias sepan cómo actuar, reduce decisiones improvisadas y proporciona un marco común para proteger la información, cumplir la normativa aplicable y evitar daños derivados de un uso inadecuado.



La responsabilidad principal debería recaer en la persona o función responsable de la seguridad de la información, en coordinación con las áreas responsables del negocio, de la protección de datos personales, del cumplimiento normativo y de los sistemas de información. La alta dirección debe aprobar la política y garantizar que se comunica adecuadamente, se aplica en la práctica y se revisa cuando cambien los riesgos, las herramientas o el marco regulatorio.

De este modo, la organización debería definir, aprobar y difundir a todos los afectados una Política de Uso de la Inteligencia Artificial que contemple los siguientes aspectos:

- **Definir usos permitidos, condicionados y prohibidos:**

Los usos permitidos son aquellos que pueden realizarse sin autorización adicional porque no introducen riesgos relevantes, por ejemplo, pedir a una IA cuyo uso esté preautorizado por la organización que ayude a mejorar la redacción de un texto público o a generar ideas generales sin aportar información confidencial.

Los usos condicionados son aquellos que pueden ser útiles, pero requieren controles previos, como utilizar IA para revisar documentación interna, apoyar análisis técnicos, resumir actas o generar código. En estos casos puede exigirse autorización, uso de herramientas aprobadas, anonimización de datos o revisión humana reforzada.

Los usos prohibidos deben incluir aquellos que supongan introducir secretos, credenciales, datos personales innecesarios, información clasificada, documentación confidencial o decisiones automatizadas no autorizadas. Por ejemplo, no debería introducirse en una herramienta pública de IA un listado de personas con datos identificativos, expedientes, contraseñas o información estratégica no publicada.

- **Determinar qué información puede introducirse en herramientas de IA:** La política debe explicar de forma sencilla qué tipos de información pueden usarse y cuáles no. Una regla práctica es permitir información pública o expresamente autorizada y prohibir información confidencial, datos personales no necesarios, secretos técnicos, claves, tokens, contratos no publicados o documentación sometida a restricciones.

Cuando sea posible, debe promoverse el uso de datos anonimizados o seudonimizados. Por ejemplo, en lugar de introducir nombres reales de personas, puede utilizarse "persona usuaria A" o "departamento B", siempre que la finalidad pueda alcanzarse sin identificar a nadie.

- **Establecer qué herramientas de IA pueden utilizarse:** No todas las herramientas ofrecen las mismas garantías. La política debe indicar si existe una lista de herramientas autorizadas, qué requisitos deben cumplir y en qué

condiciones pueden utilizarse. Entre los criterios relevantes se encuentran la ubicación y tratamiento de los datos, las condiciones contractuales, la posibilidad de que los datos se usen para entrenar modelos, las garantías de seguridad y la trazabilidad disponible.

- **Definir cómo deben tratarse los resultados generados por IA:** Los resultados de la IA deben considerarse borradores o apoyos, no verdades automáticas. Deben revisarse antes de incorporarse a documentos, decisiones, comunicaciones o sistemas. Esto es especialmente importante en textos jurídicos, técnicos, médicos, financieros, de seguridad o con impacto en derechos de las personas. Así mismo, es clave determinar si se va a permitir a la IA llevar a cabo actuaciones de manera autónoma (mediante agentes) o sólo generar información de apoyo.

Un ejemplo sencillo: si una IA genera una recomendación de seguridad, la persona responsable debe comprobar que la recomendación es técnicamente correcta, proporcional y aplicable al entorno real antes de difundirla o implantarla. Así mismo, se debe considerar si se permite a la IA ejecutar autónomamente dicha recomendación.

- **Asignar responsabilidades de uso, autorización y supervisión:** La política debe identificar quién puede autorizar nuevos usos, quién valida resultados, quién resuelve dudas, quién gestiona excepciones y quién revisa el cumplimiento. Esto evita que el uso de IA dependa exclusivamente del criterio individual de cada persona usuaria.
- **Incluir requisitos de privacidad, seguridad, propiedad intelectual y cumplimiento:** Deben incorporarse advertencias claras sobre protección de datos personales, confidencialidad, secreto profesional, derechos de autor, uso de contenidos de terceros, trazabilidad, conservación de evidencias y cumplimiento del marco normativo aplicable.
- **Revisar la política periódicamente:** La IA evoluciona rápidamente. Por ello, la política debe revisarse de forma periódica y siempre que aparezcan nuevas herramientas, nuevos casos de uso, incidentes, cambios normativos o riesgos relevantes.

2

INVENTARÍA
LOS CASOS DE
USO DE LA IA

2.2. Paso 2: Inventaría los casos de uso de la IA



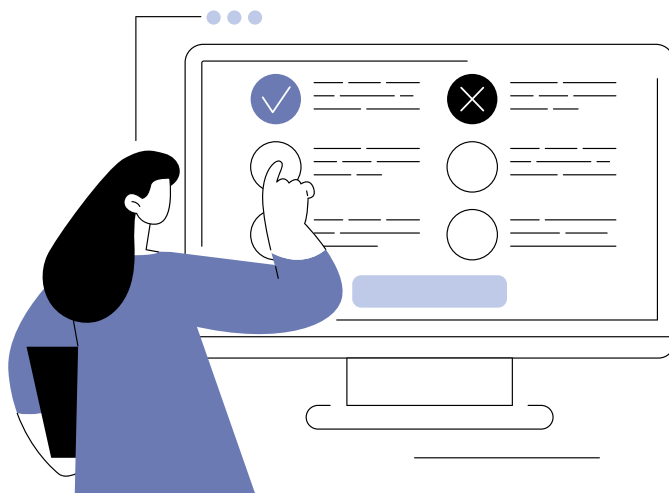
La finalidad de este segundo paso es conocer qué herramientas, modelos, servicios y casos de uso de IA existen dentro de la organización. No se puede proteger ni gobernar adecuadamente aquello que no se conoce. Un inventario permite identificar usos autorizados y no autorizados, clasificar riesgos, asignar responsables, controlar integraciones y decidir qué medidas de seguridad deben aplicarse.



El responsable principal debería ser la función responsable de los sistemas de información, con participación de seguridad de la información, protección de datos, compras, cumplimiento y las áreas usuarias. La responsabilidad no se limita a crear una lista inicial: incluye mantener el inventario actualizado, revisar altas y bajas, y coordinarlo con la gestión de riesgos y de cambios.

De ese modo, La organización debería mantener un inventario actualizado de las soluciones de Inteligencia Artificial utilizadas que contemple los siguientes aspectos:

- **Identificar herramientas, modelos, servicios y componentes de IA utilizados:** El inventario debe recoger tanto soluciones corporativas autorizadas como herramientas detectadas o declaradas por las personas usuarias. Deben incluirse motores LLM, agentes de IA, generadores de texto o imagen, herramientas de análisis documental, copilotos de productividad, motores de recomendación, chatbots, modelos integrados en aplicaciones y APIs de IA.
- **Documentar el propósito o caso de uso:** Cada elemento inventariado debe tener una finalidad descrita en lenguaje claro. Por ejemplo: "apoyo a la redacción de comunicaciones públicas", "clasificación preliminar de tickets", "resumen de documentos internos" o "detección de patrones anómalos en eventos de seguridad".
- **Identificar proveedor, propietario interno y área usuaria:** El inventario debe indicar quién presta el servicio, quién lo administra internamente y qué área lo utiliza. Esto facilita la gestión de incidencias, revisiones contractuales, bajas de servicio y análisis de responsabilidades.
- **Clasificar criticidad y riesgo asociado:** No todos los usos de IA tienen el mismo nivel de riesgo. Debe valorarse si la herramienta trata datos personales, información confidencial, procesos críticos, resultados con impacto en personas o decisiones relevantes. Un chatbot informativo con contenido público tiene un riesgo muy distinto al de una IA que apoya decisiones administrativas o analiza información sensible.
- **Identificar integraciones con otros sistemas:** Debe documentarse si la solución se conecta con aplicaciones corporativas, repositorios documentales, correo, sistemas de identidad, bases de datos, herramientas de seguridad o plataformas en la nube. Las integraciones amplían el impacto potencial de un fallo o uso inadecuado.
- **Indicar si procesa información sensible, datos personales o información confidencial:** Este punto es esencial para aplicar medidas de privacidad y seguridad. Si la herramienta trata datos personales, debe evaluarse la base de legitimación, la minimización, la información a personas afectadas, las garantías del proveedor y, cuando proceda, la necesidad de una evaluación de impacto.
- **Registrar el tratamiento de los resultados generados:** El inventario debe reflejar si los resultados se usan como borrador, recomendación, evidencia, contenido público, entrada para una decisión o salida automatizada. Cuanto mayor sea el impacto del resultado, mayor debe ser la supervisión humana.
- **Mantener trazabilidad de altas, modificaciones y bajas:** Toda incorporación, cambio relevante o retirada de una solución de IA debe registrarse. Por ejemplo, si una herramienta empieza a conectarse a un repositorio documental interno, el inventario debe actualizarse porque el riesgo cambia de forma material.



3

ANALIZA LOS
RIESGOS DE
USO DE LA IA

2.3. Paso 3: Analiza los riesgos de uso de la IA



La finalidad de este tercer paso es identificar, valorar y tratar los riesgos derivados de cada caso de uso de IA antes de su implantación y durante su funcionamiento. El análisis de riesgos permite decidir si un uso es aceptable, qué controles deben aplicarse, qué riesgos residuales se asumen y qué usos deben limitarse o descartarse.



La responsabilidad debería recaer en la persona responsable de seguridad de la información, con la colaboración del área propietaria del caso de uso, los responsables de protección de datos, cumplimiento y compras y el responsable de los sistemas de información. Cuando el uso pueda afectar a derechos de personas o decisiones relevantes, la participación jurídica y de privacidad resulta especialmente importante.

Por tanto, la organización debería llevar a cabo un análisis de los riesgos específicos de seguridad asociados al uso de soluciones de Inteligencia Artificial, en el que se contemplen los siguientes aspectos:

- **Identificar el caso de uso y su valor para la organización:** Antes de analizar riesgos, debe describirse qué se quiere hacer, con qué herramienta, qué datos se utilizarán, quién usará el sistema y qué resultado (información, actuaciones, etc.) se espera obtener. También debe valorarse la importancia del caso de uso: no es lo mismo una ayuda para redactar un texto divulgativo que una herramienta que prioriza expedientes, clasifica incidentes o apoya decisiones con impacto en personas.
- **Determinar el catálogo de riesgos a considerar:** La organización deberá establecer el catálogo de riesgos específicos a considerar para cada caso de uso, pudiendo ser éstos de muy diversa naturaleza, tal y como se indica a continuación:
- **Riesgos inherentes al propio caso de uso:** Debe evaluarse la jurisdicción aplicable, la normativa de protección de datos, el posible encaje en el Reglamento Europeo de Inteligencia Artificial (sistemas de IA de riesgo inaceptable, alto, limitado o mínimo), las obligaciones sectoriales, los requisitos contractuales y cualquier limitación derivada de confidencialidad, propiedad intelectual o secreto profesional, así como las implicaciones asociadas a su integración con otros sistemas de información, tratamientos de datos o procesos corporativos.
- **Riesgos éticos y de confianza:** La IA puede generar sesgos, resultados discriminatorios, explicaciones insuficientes o recomendaciones difíciles de entender. Por ello, debe analizarse si el caso de uso requiere transparencia, explicabilidad, trazabilidad o mecanismos para que una persona pueda revisar y corregir, si fuese necesario, el resultado. La integridad y trazabilidad de la actividad de la IA, en su sentido más amplio, son factores clave.
- **Riesgos asociados al uso indebido de la IA:** También deben considerarse posibles usos no autorizados, uso con datos no permitidos o introducción de instrucciones maliciosas. Por ejemplo, una persona usuaria podría copiar información confidencial en una herramienta externa para resumirla, aunque esa práctica no esté autorizada.
- **Riesgos sobre los datos proporcionados a la IA:** Entre los riesgos más relevantes están la fuga de datos, el uso de información para entrenar modelos sin autorización, el almacenamiento no controlado, la inferencia de información sensible y el envenenamiento de datos. Debe prestarse especial atención a datos personales, credenciales, secretos, documentación interna y datos de terceros.
- **Riesgos sobre el propio modelo de IA:** Si la organización desarrolla o ajusta modelos propios, debe valorar riesgos de manipulación del modelo, extracción del modelo, ataques adversariales, pérdida de integridad, dependencia de componentes de terceros y exposición de parámetros o datos de entrenamiento. Si la organización utiliza modelos de terceros, debe valorar los riesgos asociados a las condiciones de uso de dichos modelos (incluyendo el modelo de costes), sus posibles cambios, la disponibilidad de dichos modelos, los niveles de servicio garantizados, etc.
- **Riesgos sobre los resultados generados por la IA:** Los resultados pueden contener errores, alucinaciones, referencias inexistentes, sesgos, información no autorizada o conclusiones no justificadas. Por ejemplo, una IA puede presentar como cierta una norma derogada, inventar una fuente o recomendar una actuación inapropiada para el entorno real. Evaluar estos riesgos es clave, ya que la responsabilidad (incluso legal) sobre los resultados (información o acciones) de la IA siempre es de la organización.
- **Valorar probabilidad, impacto y riesgo aceptable:** El análisis debe estimar la probabilidad de ocurrencia y el impacto de cada riesgo, considerando el valor del caso de uso y las medidas existentes para su mitigación. Así mismo, deberá determinarse cuál es el umbral de riesgo que la organización considera aceptable para el uso de la IA.
- **Definir acciones de tratamiento y aceptación formal:** Para los riesgos no aceptables deben definirse acciones concretas: limitar datos, cambiar de proveedor, aplicar anonimización, incorporar revisión humana, restringir perfiles, bloquear herramientas no autorizadas, reforzar registros, realizar auditorías o descartar el caso de uso. El riesgo residual debe ser aceptado formalmente por la organización.

4

CONTROLA
EL USO
DE LA IA

2. 4. Paso 4: Controla el uso de la IA



La finalidad de este cuarto paso es disponer de mecanismos técnicos y organizativos que permitan controlar el uso de soluciones y herramientas de IA, aplicar restricciones proporcionadas y detectar usos inseguros o no autorizados. Estos controles permiten que la política de uso de IA no sea solo una declaración, sino una práctica verificable.



La responsabilidad principal debería corresponder al responsable de los sistemas de información, en coordinación con seguridad de la información y con las áreas de cumplimiento, privacidad y relación con las personas usuarias. La aplicación de controles debe ser proporcionada, respetar la normativa aplicable y comunicarse de forma clara a las personas afectadas.

De este modo, la organización debería desplegar herramientas y mecanismos de control del uso de herramientas de Inteligencia Artificial que sean capaces de:

- **Identificar el uso de herramientas de IA desde entornos corporativos:** La organización debe poder conocer, en la medida técnicamente viable, qué servicios de IA se utilizan desde sus redes, equipos o identidades corporativas. Esto puede realizarse mediante herramientas de seguridad web, CASB, SASE, registros de navegación, inventarios de aplicaciones o controles de endpoints.
- **Aplicar restricciones o bloqueos sobre herramientas no autorizadas:** Cuando una herramienta no ofrezca garantías suficientes o incumpla la política corporativa, debe poder bloquearse o limitarse. Por ejemplo, puede permitirse una solución corporativa de IA y bloquear herramientas públicas que utilicen los datos introducidos para entrenamiento no controlado.
- **Definir políticas diferenciadas por perfiles, áreas o tipo de información:** No todas las personas necesitan las mismas capacidades. Puede permitirse el uso de IA para redacción general a toda la plantilla, reservar herramientas de análisis documental para perfiles autorizados y exigir controles reforzados a quienes traten información sensible o datos personales.
- **Detectar y bloquear intentos de introducir información sensible:** Cuando sea viable y proporcionado, deben integrarse mecanismos de prevención de fuga de datos para advertir o impedir que se introduzcan datos personales, información confidencial, credenciales, claves, código sensible o documentación clasificada en herramientas no autorizadas.
- **Mantener registros y evidencias de uso:** La trazabilidad es necesaria para detectar abusos, investigar incidentes y demostrar cumplimiento. Los registros deben ser adecuados al riesgo y respetar la privacidad de las personas usuarias. No se trata de vigilar de forma indiscriminada, sino de disponer de evidencias suficientes para gestionar la seguridad.
- **Monitorizar usos indebidos o comportamientos anómalos:** Deben definirse alertas ante patrones de riesgo, como uso masivo de herramientas no autorizadas, acceso desde ubicaciones inusuales, cargas repetidas de documentos sensibles o intentos de eludir controles. Estas alertas deben tratarse como eventos de seguridad.
- **Facilitar la aplicación de la política de uso de IA:** Los controles deben ayudar a las personas, no sustituir la formación. Por ejemplo, una advertencia automática antes de enviar información sensible puede evitar errores y reforzar el aprendizaje práctico.
- **Formar a las personas en el uso seguro de la IA:** La organización debe asegurarse de que las personas usuarias conocen los usos permitidos y prohibidos de la IA, así como los errores habituales que se pueden producir en torno a su uso. Las personas usuarias deben disponer de ejemplos sobre qué se puede hacer y qué no, saber que la IA puede inventar datos, reproducir sesgos o generar respuestas convincentes pero incorrectas, y conocer que un uso más seguro de la IA requiere formular instrucciones claras, limitar la información proporcionada, pedir explicaciones verificables, solicitar advertencias sobre incertidumbre y contrastar los resultados con fuentes fiables.



5

EVALÚA LOS
RESULTADOS...
Y VUELVE
A EMPEZAR

2. 5. Paso 5: Evalúa los resultados... y vuelve a empezar



La finalidad de este último paso es comprobar de forma periódica y estructurada si el uso de la inteligencia artificial genera los resultados esperados de manera segura, fiable, proporcionada y conforme a las normas internas, evaluando no solo su funcionamiento técnico, sino el valor real aportado, la calidad de los resultados, el control de riesgos y la adecuación de las medidas adoptadas. Esto debe entenderse como un proceso cíclico en el que, ante cambios o tras periodos de uso, la organización evalúa resultados, aprende y revisa continuamente las fases previas, dado que las herramientas, los modelos, los usos y las amenazas evolucionan.



La responsabilidad principal debe recaer en la función de seguridad de la información, en coordinación con el responsable del sistema de información y las áreas usuarias y, cuando proceda, con las funciones de protección de datos, cumplimiento, auditoría o gestión de riesgos. El área propietaria del caso de uso debe participar activamente evaluando la utilidad y fiabilidad de la IA en el contexto real, mientras que las funciones de seguridad, privacidad y cumplimiento deben determinar si los riesgos siguen siendo aceptables o requieren ajustes en los controles.

Por ello, la organización debería desarrollar un proceso de gestión de las medidas desplegadas para el uso de la Inteligencia Artificial, en el que se contemplen los siguientes aspectos:

- **Evaluar si el caso de uso cumple la finalidad prevista:** Debe verificarse si la herramienta de IA aporta el beneficio esperado en términos de eficiencia, reducción de tiempos, calidad de resultados, mejora del servicio o apoyo a tareas, comprobando en casos concretos que estos sean comprensibles, completos y útiles.
- **Revisar la calidad, exactitud y fiabilidad de los resultados:** Los resultados de la IA deben evaluarse periódicamente mediante revisiones y comprobaciones que aseguren su uso adecuado, especialmente cuando impactan en decisiones, documentos o análisis sensibles, analizando errores, omisiones, incoherencias, sesgos o referencias no verificables, y verificando, por ejemplo, que las recomendaciones técnicas sean correctas, proporcionales, compatibles con el contexto y alineadas con las políticas internas y requisitos regulatorios.
- **Comprobar que los riesgos siguen siendo aceptables:** El análisis de riesgos debe revisarse periódicamente o ante cambios relevantes (como proveedor, funcionalidades, usuarios, datos o integraciones), evaluando aspectos como confidencialidad, protección de datos, trazabilidad, dependencia del proveedor o fiabilidad de resultados, y, si el riesgo residual dejara de ser aceptable, deberían adoptarse medidas como limitar el uso, reforzar la supervisión humana, anonimizar datos, ajustar configuraciones, restringir accesos, actualizar contratos o suspender el caso de uso.
- **Analizar incidencias, desviaciones y usos no previstos:** La evaluación debe considerar incidencias detectadas durante el uso de la IA (errores relevantes, dudas recurrentes, usos no autorizados o desviaciones) para identificar debilidades en controles o políticas y adoptar medidas adecuadas. Por ejemplo, si se detecta uso de herramientas públicas para tratar información interna, será necesario analizar la causa, ofrecer alternativas seguras y reforzar la comunicación y las normas de uso.
- **Determinar acciones de mejora y volver a empezar el ciclo:** La evaluación de todos los aspectos anteriores debe conducir a acciones concretas, manteniendo o ampliando usos si los resultados son positivos y los riesgos están controlados, o revisando política, inventario, análisis de riesgos, controles, responsabilidades e instrucciones si se detectan problemas, en un enfoque de mejora continua que asume la evolución constante de la IA y la necesidad de revisión, aprendizaje y adaptación permanentes.



03.

Conclusiones

La inteligencia artificial representa una de las transformaciones tecnológicas más relevantes de los últimos años. **Su capacidad para apoyar tareas, acelerar análisis, generar contenidos y facilitar nuevos servicios puede aportar un valor significativo** a personas, organizaciones y administraciones públicas. Pero ese valor solo será sostenible si su uso se realiza con seguridad, responsabilidad y confianza.



El principal reto no consiste únicamente en elegir una herramienta concreta, sino en **gobernar adecuadamente su uso**. Una IA segura requiere **reglas claras, conocimiento de los casos de uso, análisis de riesgos, controles proporcionados, supervisión humana y formación continua**. Sin estos elementos, incluso una herramienta técnicamente avanzada puede generar riesgos graves: exposición de información, decisiones incorrectas, pérdida de confianza, incumplimientos normativos o dependencia de servicios que la organización no controla suficientemente.



Por ello, la aproximación recomendada debe ser práctica y equilibrada. No se trata de prohibir la IA de forma general, ni de permitirla sin límites. Se trata de identificar para qué puede utilizarse, con qué datos, por quién, en qué condiciones y con qué mecanismos de revisión. **La clave está en permitir usos beneficiosos y de bajo riesgo, reforzar los usos sensibles y evitar aquellos que puedan comprometer derechos, seguridad o confianza pública.**



Para la ciudadanía y el público general, el mensaje esencial es sencillo: **la IA puede ser una ayuda, pero no sustituye el criterio humano**. Conviene no introducir información personal o confidencial en herramientas no autorizadas, revisar siempre los resultados, desconfiar de respuestas que no puedan verificarse y utilizar la tecnología de forma consciente. La seguridad empieza por decisiones cotidianas bien informadas.



Para las organizaciones, el reto es más amplio: **deben construir un marco de gobierno que permita innovar sin perder control**. La política de uso de IA, el inventario de soluciones, el análisis de riesgos, las herramientas de control y la formación deben funcionar como un sistema integrado. Solo así puede demostrarse que la IA se utiliza de manera proporcional, trazable, revisable y alineada con los valores de seguridad, privacidad, transparencia y servicio a las personas.

En definitiva, un uso seguro de la IA no es un obstáculo para la innovación, sino una condición para que la innovación sea fiable. Cyberzaintza, desde su papel de impulso y coordinación de la ciberseguridad, puede contribuir a que la adopción de la IA se realice con criterio, con garantías y con una visión orientada tanto a la protección como al aprovechamiento responsable de sus beneficios.



Sarean ere, auzolana.
Por ti hacemos red.

Información general:
945 236 636
info@cyberzaintza.eus



www.ciberseguridad.euskadi.eus

EUSKO JAURLARITZA



GOBIERNO VASCO