

Estrategia Gallega  
de Inteligencia Artificial 2030

# Guía práctica para la gestión de la inteligencia artificial en las administraciones públicas



XUNTA  
DE GALICIA



AXENCIA PARA A  
MODERNIZACIÓN  
TECNOLÓXICA DE GALICIA

**Edita.** Xunta de Galicia. Axencia para a Modernización Tecnolóxica de Galicia – Amtega

Santiago de Compostela. Año 2023

*Los textos de este documento se distribuyen bajo licencia Creative Commons Atribución-Compartir Igual 4.0 Internacional (CC BY-SA 4.0). Disponible en: <https://creativecommons.org/licenses/by-sa/4.0/deed.es>. Las imágenes fueron obtenidas de "Deloitte Brand Space", no está permitida su extracción o redistribución como imágenes independientes, separadas del documento original.*



AXENCIA PARA A  
MODERNIZACIÓN  
TECNOLÓXICA DE GALICIA





Julián Cerviño Iglesia

## Director de la Agencia para la Modernización Tecnológica de Galicia

El desarrollo de la Inteligencia artificial está llamado a marcar un antes y un después en la trayectoria de la Humanidad. Por primera vez, estamos creando máquinas con la capacidad de procesar con criterios inteligentes volúmenes ingentes de datos con capacidad para hacer muchas cosas. Y eso debe ser objeto de un saber, donde hemos de tener claro lo que estamos haciendo y a dónde queremos llegar con ello. Porque, el reto de la sociedades modernas y humanistas es no perder la referencia; las máquinas las hacen personas.

Ahora estamos ante un horizonte totalmente nuevo y, con enormes incertidumbres. Debemos pensar que el desarrollo de la inteligencia artificial supone una enorme oportunidad para procesar rápidamente y resolver dilemas que se están cuestionando sobre nuestro derecho a la intimidad. Nuestro continente fue pionero al proponer la primera legislación de IA del mundo, que se publicó en 2021, basado en la importancia de una gestión adecuada de los riesgos asociados a esta tecnología.







En el marco del sector público, la inteligencia artificial permitirá transformar la relación entre la administración pública y la ciudadanía, impulsando la evolución hacia una gobernanza más inteligente y centrada en las personas. Al permitir análisis de datos más rápidos y precisos, esta tecnología creará una administración más ágil y con capacidad para tomar mejores decisiones, al tiempo que impulsará una mayor transparencia y adaptabilidad a las necesidades cambiantes de la ciudadanía.

Consciente de este reto, la Xunta de Galicia asumió ya en 2021 un compromiso estratégico por garantizar el uso eficaz, eficiente, ético y seguro de la inteligencia artificial en la administración autonómica. Así se plasma en la Estrategia Gallega de Inteligencia Artificial 2030, hoja de ruta con la que la Xunta de Galicia aspira a liderar la creación de las condiciones propicias para el desarrollo y la implantación de una inteligencia artificial ética y fiable, que respete los valores y normas de la región, refuerce su presencia en las cadenas de valor de la industria y, ante todo, sirva a la ciudadanía. Para ello, la Xunta de Galicia ha propuesto un texto en lo que será la primera legislación de inteligencia artificial de una región europea con el anteproyecto de ley para el desarrollo e impulso de la inteligencia artificial en Galicia, que busca establecer un marco legal uniforme para el diseño, adquisición, implementación y uso de sistemas de inteligencia artificial en la Administración General y del Sector Público Autonómico de Galicia y fijar los principios éticos rectores del diseño, adquisición e implantación de los sistemas de inteligencia artificial en Galicia. En definitiva, el objetivo de esta Ley es garantizar que los sistemas de inteligencia artificial que se utilizan en la comunidad son fiables y respetuosos con los derechos fundamentales de las personas.

De este modo, la presente guía constituye un componente esencial dentro de este marco estratégico y normativo con el propósito de proporcionar las claves necesarias para que cualquier administración pública identifique las distintas fases y agentes intervinientes en la gestión de proyectos que integren sistemas de inteligencia artificial.

## Acerca de la Axencia para a Modernización Tecnolóxica de Galicia

La Agencia para la Modernización Tecnológica de Galicia (Amtega) es el instrumento estratégico y ejecutivo de la Xunta de Galicia para impulsar la modernización de la Administración pública, así como el desarrollo económico y social de la región. Constituida a través del Decreto 252/2011, del 15 de diciembre, su actuación responde a los siguientes objetivos:

-  Poner a disposición de la ciudadanía y de las empresas de Galicia contenidos y servicios públicos digitales avanzados
-  Dotar a la ciudadanía de las competencias y recursos para participar activamente en el desarrollo de la sociedad de la información
-  Transformar, mediante un uso acomodado de las TIC, la Administración pública autonómica, con la finalidad de mejorar la calidad de prestación de los servicios públicos
-  Impulsar el hipersector TIC, de forma que se convierta en un soporte para incrementar la competitividad y el empleo
-  Impulsar el uso de los servicios TIC por parte de las empresas gallegas
-  Implantar una red de infraestructuras moderna y sostenible que garantice la integración de Galicia en la sociedad de la información





## ¿Por qué esta guía?

La inteligencia artificial es una tecnología disruptiva que tiene la capacidad de generar un impacto significativo en diversos aspectos de la sociedad y la economía. Entre las razones que hacen de la inteligencia artificial una tecnología revolucionaria, cabe destacar su contribución a una mejor toma de decisiones por parte de las personas, en tanto que es capaz de analizar grandes conjuntos de datos y extraer información valiosa de ellos de forma ágil. La inteligencia artificial también permite la automatización de procesos en diversos ámbitos. En definitiva, en el caso del sector público, estas y otras aplicaciones de la inteligencia artificial permiten que las personas centren sus esfuerzos en tareas de valor añadido y tomen decisiones apoyadas en datos, para dar lugar a mejores servicios o a una mejor atención, entre otros beneficios.

Sin embargo, es importante señalar que el uso de la inteligencia artificial también plantea desafíos éticos y sociales que deben abordarse de manera responsable. La importancia de la inteligencia artificial radica no solo en sus beneficios potenciales, sino también en la necesidad de abordar sus implicaciones éticas y sociales a medida que se avanza en su desarrollo y adopción. Así lo reconoce la Unión Europea, que ha asumido el compromiso de liderar a nivel mundial el establecimiento de normas armonizadas en materia de gestión de sistemas de IA y de los riesgos asociados a dichos sistemas.

Por ello, esta guía, de forma coherente con la normativa europea, pretende servir como soporte en la gestión de proyectos que integren sistemas de inteligencia artificial en la administración pública, aportando los pasos que se han de dar y los controles de riesgos que se han de realizar desde que se concibe una oportunidad hasta que se convierte en un proyecto, de forma que se garantice la óptima adopción de esta tecnología.

## ¿Para quién es esta guía?

Esta guía se dirige al conjunto de las administraciones públicas, organismos y entidades que conforman el sector público. La guía aporta una serie de pautas e indicaciones que pueden ser adaptadas por cada administración pública, de acuerdo con su idiosincrasia.

## Índice de contenidos

### Capítulo 1

#### La inteligencia artificial en las políticas públicas

7

1.1 Qué es la inteligencia artificial

9

1.2 La inteligencia artificial en el ciclo de vida de las políticas públicas

12

### Capítulo 2

#### El ciclo de vida de un sistema de inteligencia artificial

21

2.1 El sistema de inteligencia artificial

23

2.2 El ciclo de vida de un sistema de IA

25

2.3 La gestión del ciclo de vida de un sistema de IA

27

### Capítulo 3

#### Fases del ciclo de vida de un sistema de inteligencia artificial

45

3.1 Fase 1. Identificación y evaluación de oportunidades vinculadas a la adopción de la IA

48

3.2 Fase 2. Diseño, desarrollo e implementación del sistema de IA

62

### Capítulo 4

#### Gestión de riesgos de un sistema de IA

69

4.1 Introducción a la gestión de riesgos

71

4.2 Modelo de gestión de riesgos de un sistema de IA

72

4.3 Ejemplo de gestión de riesgos en el diseño, desarrollo e implantación de un sistema de IA

84

### Conclusiones

96

### Referencias

98



## Capítulo 1

### La inteligencia artificial en las políticas públicas

En un mundo tecnológico en constante evolución, la inteligencia artificial (IA) emerge como una fuerza transformadora en el ciclo de vida de las políticas públicas. Este capítulo explora cómo la IA, al automatizar procesos administrativos y mejorar la toma de decisiones, se convierte en un catalizador clave para la eficiencia y la adaptabilidad en la administración pública. Desde la realización de actos administrativos hasta la anticipación de necesidades mediante análisis predictivos, la IA redefine la interacción entre la administración y la ciudadanía, proporcionando eficiencia, personalización y transparencia. La implementación responsable de la IA no solo aborda ineficiencias, sino que también abre oportunidades para una relación más eficiente y satisfactoria con la sociedad.

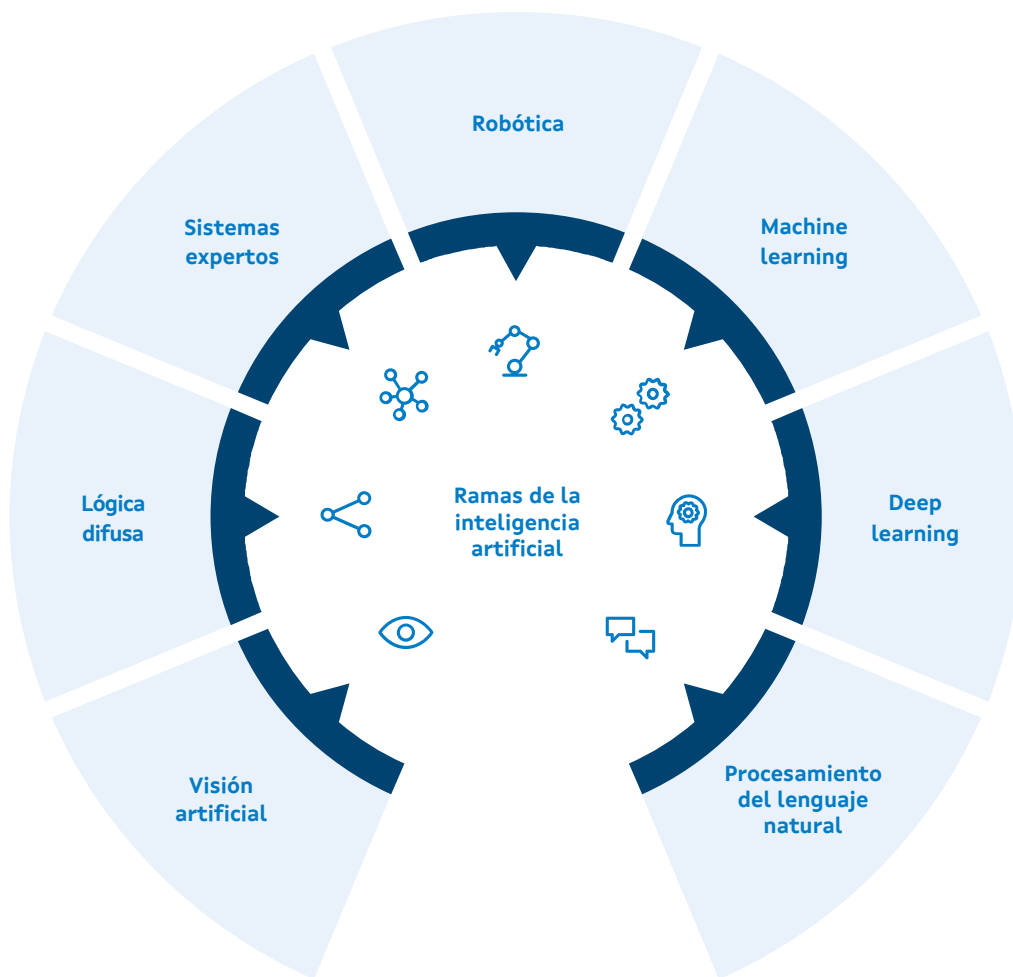




## 1.1. Qué es la inteligencia artificial

En este contexto, la inteligencia artificial (IA) tendrá un impacto relevante en todos los ámbitos de la sociedad. Esta tecnología busca emular la capacidad cognitiva de los seres humanos a través de algoritmos y modelos matemáticos. Se trata de un campo de la informática que se enfoca en desarrollar sistemas y programas capaces de realizar tareas que normalmente requerirían inteligencia humana. Estas tareas incluyen el razonamiento, el aprendizaje, la percepción, la toma de decisiones o la comprensión del lenguaje natural, entre otras.

Si bien su ritmo de evolución conlleva la aparición incesante de nuevas ramas de inteligencia artificial, pueden identificarse las siguientes:





Tiene como objetivo desarrollar técnicas que ayuden a las computadoras a ver y comprender imágenes y videos digitales. Aplicar modelos de aprendizaje automático a imágenes permite a las computadoras identificar objetos, rostros, personas, animales y más. Los modelos algorítmicos ayudan a las computadoras a aprender por sí mismas sobre los contextos de los datos visuales, y con suficientes datos proporcionados a través de un modelo, las computadoras pueden aprender a distinguir una imagen de otra.

Una red neuronal convolucional trabaja junto a un modelo para descomponer las imágenes en píxeles, asignándoles etiquetas. La red neuronal luego utiliza estas etiquetas para realizar convoluciones, que es una operación matemática en dos funciones para producir una tercera función, y hacer predicciones sobre lo que ve.



### Lógica difusa

Es una técnica que ayuda a resolver problemas o afirmaciones que pueden ser verdaderos o falsos. Este método imita las decisiones humanas al considerar todas las posibilidades existentes entre los valores digitales de 'sí' y 'no'. En pocas palabras, mide el grado en que una hipótesis es correcta. Es una forma conveniente y flexible de implementar técnicas de aprendizaje automático e imitar el pensamiento humano de manera lógica.



### Sistemas expertos

Un sistema experto es un programa especializado en una tarea singular. Estos sistemas están diseñados principalmente para resolver problemas intrincados con capacidades de toma de decisiones similares a las humanas. Utilizan un conjunto de reglas, llamadas reglas de inferencia (p.e. nociones lógicas de tipo "si-entonces"), que una base de conocimientos alimentada por datos define para ellos.



### Robótica

La robótica basada en IA se refiere a la integración de sistemas y algoritmos de inteligencia artificial en robots para permitirles realizar tareas de manera más autónoma, adaptativa y eficiente. En este contexto, la IA permite que los robots procesen información, aprendan de experiencias anteriores y tomen decisiones basadas en datos y patrones. En lugar de depender de programación rígida y algoritmos predefinidos, los robots basados en IA pueden aprender y mejorar su rendimiento con el tiempo. Esto implica la capacidad de adaptarse a entornos cambiantes, tomar decisiones en tiempo real y enfrentarse a situaciones nuevas sin intervención humana constante.



### Machine learning (aprendizaje automático)

El aprendizaje automático es la capacidad de las máquinas para aprender automáticamente a partir de datos y algoritmos. El aprendizaje automático mejora el rendimiento utilizando experiencias pasadas y puede tomar decisiones sin estar programado específicamente para hacerlo. El proceso comienza con la recopilación de datos históricos, como instrucciones y experiencia directa, para que se puedan construir modelos lógicos para inferencias futuras. La precisión de la salida depende del tamaño de los datos: una mayor cantidad de datos construirá un mejor modelo, lo que a su vez aumenta su precisión. Los algoritmos de aprendizaje automático se clasifican en tres tipos:

- Aprendizaje supervisado: las máquinas se entrenan con datos etiquetados para predecir el resultado.
- Aprendizaje no supervisado: las máquinas se entrenan con datos no etiquetados, y el modelo extrae información de la entrada para identificar características y patrones, de modo que pueda generar un resultado.
- Aprendizaje por refuerzo: las máquinas aprenden a través de ensayo y error, utilizando retroalimentación para formar acciones.





### Redes neuronales/aprendizaje profundo (deep learning)

Las redes neuronales también son conocidas como redes neuronales artificiales o redes neuronales simuladas. En el corazón de los algoritmos de aprendizaje profundo, las redes neuronales se inspiran en el cerebro humano y copian cómo las neuronas biológicas envían señales entre sí.

Las redes neuronales artificiales tienen capas de nodos, que constan de una capa de entrada, una o más capas ocultas y una capa de salida. Cada nodo, también llamado una neurona artificial, se conecta con otras neuronas y tiene un umbral y un peso asociados. Cuando la salida de un nodo individual supera un valor de umbral especificado, el nodo se activa para enviar datos a la siguiente capa de la red. Las redes neuronales necesitan datos de entrenamiento tanto para aprender como para mejorar la precisión



### Procesamiento del lenguaje natural

El procesamiento del lenguaje natural permite a las computadoras entender tanto el texto como las palabras habladas, de manera similar a los humanos. Combinando aprendizaje automático, lingüística y modelos de aprendizaje profundo, las computadoras pueden procesar el lenguaje humano en datos de voz o texto para comprender el significado completo, la intención y el sentimiento.



## 1.2. La inteligencia artificial en el ciclo de vida de las políticas públicas

El ciclo de vida de la política pública representa la forma en la que se desarrollan las políticas públicas y sirve, además, para planificar y analizar las diferentes fases de este proceso. El ciclo se inicia con la identificación y definición de un problema o asunto a resolver. En este punto, las administraciones han de evaluar las alternativas propuestas por los distintos órganos gestores responsables de esa política. Durante este proceso, la tecnología actúa como un instrumento que puede servir para desarrollar las alternativas de política pública propuestas a partir del contexto y la factibilidad de la solución.

En ese caso, la inteligencia artificial se alza como una herramienta esencial para impulsar la calidad de los servicios públicos. En la administración pública se implementan soluciones de IA para analizar grandes volúmenes de datos, identificar patrones en el comportamiento ciudadano y optimizar la asignación de recursos, entre otros múltiples usos. El uso de la IA contribuye por tanto al impulso de una administración más ágil y adaptativa, permitiendo alcanzar respuestas más efectivas a los retos y adaptadas a los cambios en las necesidades de la ciudadanía.

### 1.2.1. La relación de la ciudadanía con la administración pública

La evolución de la relación entre la administración pública y la ciudadanía ha experimentado un cambio drástico con la aparición de la vía digital. Antes, las interacciones se limitaban a procesos presenciales y trámites poco transparentes, marcados por largas esperas y limitada accesibilidad. La participación ciudadana en el ciclo de vida de las políticas públicas estaba condicionada por barreras burocráticas y una experiencia desigual.

Con la transición hacia la vía digital, este escenario ha sufrido una transformación fundamental. La ciudadanía ahora puede interactuar con la administración pública de manera remota, aprovechando plataformas en línea, portales y aplicaciones móviles. Este cambio ha generado un nuevo paradigma, donde la conveniencia y la agilidad definen la relación entre la ciudadanía y la administración.

La ventaja primordial de la vía digital es la comodidad que ofrece. La posibilidad de realizar una variedad de trámites de forma telemática y acceder a servicios públicos en cualquier momento del día representa un cambio paradigmático en la dinámica de interacción. La agilidad se ha convertido en una característica distintiva de esta vía, facilitada por la presentación electrónica de documentos, la realización de pagos en línea y la capacidad de dar seguimiento en tiempo real al estado de las solicitudes. La digitalización de trámites no solo ha reducido los tiempos de respuesta, sino que también ha mejorado significativamente la experiencia del usuario.

A pesar de estos avances, la vía digital no está exenta de desafíos. La complejidad de algunos procesos, la falta de uniformidad en la atención a la ciudadanía y la brecha digital son aspectos críticos que deben abordarse para garantizar una relación efectiva entre la administración y los ciudadanos. La superación de estos desafíos será crucial para consolidar los beneficios de la transformación digital y garantizar una interacción ciudadanía-administración inclusiva y equitativa. Este nuevo paradigma digital, aunque prometedor, demanda una atención constante a fin de optimizar la relación y fortalecer la confianza en los servicios gubernamentales.

## 1.2.2. La IA como fuerza transformadora de la relación administración-ciudadanía

El poder transformador de IA en la relación entre la administración y la ciudadanía es innegable. La implementación de soluciones basadas en IA redefine la dinámica tradicional al ofrecer eficiencia, personalización y anticipación en la prestación de servicios públicos. En este contexto, se pueden diferenciar dos grandes aplicaciones de la IA en la administración pública:

- **Sistemas para la realización de actos administrativos:** se refiere al ámbito de la IA en el que la tecnología es capaz de “tomar” decisiones y “realizar” las acciones resultantes sin la intervención directa de personas. Esto implica que el sistema asume la ejecución de tareas que anteriormente eran realizadas por personas.

La IA actúa como un catalizador clave para la automatización de procesos administrativos. La capacidad de sistemas avanzados para automatizar tareas y procesos permite que las personas se centren en la resolución de cuestiones complejas y estratégicas. Esta automatización no solo acelera los tiempos de respuesta en la resolución de procedimientos, sino que también contribuye a una toma de decisiones más informada por parte de las personas, garantizando una mayor precisión y consistencia en los servicios públicos.

- **Sistemas para la realización de actuaciones administrativas:** en el contexto de esta tecnología, estos sistemas abarcan aquellas instancias en las que la información producida por los modelos de IA se emplea como elemento clave en el proceso de toma de decisiones humanas.

La IA brinda la capacidad de realizar análisis predictivos, permitiendo a la administración anticipar necesidades y ajustar recursos y servicios de manera proactiva. Esta capacidad de prever posibles obstáculos en los procesos administrativos mejora la eficiencia y la capacidad de respuesta, garantizando que la administración esté preparada para enfrentar desafíos emergentes y cambios en las demandas de la ciudadanía.

La personalización de la interacción, facilitada por la IA, se convierte en un elemento esencial para fortalecer la conexión con la ciudadanía. Los sistemas de asistentes virtuales pueden adaptarse a las necesidades específicas de cada persona, proporcionando información y asistencia de manera contextualizada. Este enfoque personalizado no solo mejora la experiencia del usuario, sino que también contribuye a construir una relación más cercana y satisfactoria entre la administración y la ciudadanía.

Asimismo, la IA facilita una administración más transparente y eficiente. Los algoritmos de IA pueden analizar grandes conjuntos de datos para identificar patrones, tendencias y áreas de mejora. Esta capacidad de análisis profundo proporciona información valiosa para la toma de decisiones informadas, permitiendo que la administración ajuste sus políticas públicas de manera más eficaz.

En conclusión, la implementación responsable de la IA no solo se erige como la solución para las ineficiencias en la administración pública, sino que también representa una fuente de oportunidades a futuro. Esta integración, llevada a cabo de manera consciente, genera una transformación positiva en la prestación de servicios públicos y establece una relación más eficiente y satisfactoria con la ciudadanía. Es imperativo destacar que la correcta adopción de la IA debe ir de la mano de las personas de la administración, quienes son en última instancia partícipes de la aplicación de esta tecnología en el ciclo de definición y ejecución de políticas públicas, asegurando así un enfoque equitativo y adaptativo a las necesidades cambiantes de la sociedad.



## Posibles aplicaciones de la IA en la administración pública

### Resolución de consultas de la ciudadanía

Los chatbots y asistentes virtuales basados en IA pueden proporcionar respuestas rápidas y precisas a las preguntas de la ciudadanía, mejorando la calidad de la atención y reduciendo la carga de trabajo de las personas empleadas.

### Refuerzo de la ciberseguridad

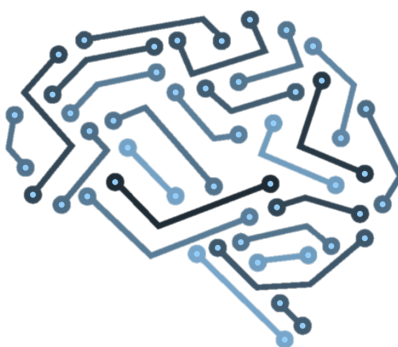
La IA se utiliza en la detección de amenazas cibernéticas y en la protección de infraestructuras críticas, lo que es fundamental en un mundo cada vez más digitalizado.

### Detección y respuesta a desastres naturales

La IA puede ayudar en la detección temprana de desastres naturales, como terremotos, inundaciones y incendios forestales, y en la coordinación de la respuesta de emergencia, lo que puede salvar vidas y reducir el impacto.

### Predicción de epidemias

La IA puede analizar grandes volúmenes de datos para identificar tendencias y patrones que podrían llevar a la detección temprana de problemas o la anticipación de necesidades. Por ejemplo, en el ámbito de la salud pública, la IA puede predecir brotes de enfermedades antes de que se propaguen.



### Traducción y servicios multilingües

La IA se utiliza para la traducción automática de documentos y servicios multilingües, lo que es fundamental en una Unión Europea con múltiples idiomas oficiales. Esto facilita la comunicación y la colaboración en un entorno diverso.

### Detección de fraudes y corrupción

La IA puede ayudar en la identificación de actividades fraudulentas y corruptas mediante el análisis de datos financieros y transaccionales. Esto es esencial para garantizar la integridad y la transparencia en el sector público.

### Optimización del tráfico y transporte público

En áreas urbanas, la IA se utiliza para optimizar el tráfico y el transporte público. Los algoritmos de IA pueden mejorar la gestión de semáforos, predecir congestiones y proporcionar rutas más eficientes para el transporte público, lo que reduce la congestión y mejora la movilidad.

### Planificación urbana y gestión de residuos

La IA se puede utilizar en la planificación urbana para optimizar el uso de recursos, prever necesidades de infraestructura y gestionar eficientemente los residuos y la energía en las ciudades.

### 1.2.3. Aplicaciones de la IA en la administración pública a nivel mundial

El sector público global no es ajeno al impacto de las aplicaciones de IA y ha comenzado ya a aplicar esta tecnología en sus procesos. Estos ejemplos ilustran cómo la IA se ha integrado en diversas áreas de las administraciones públicas en todo el mundo, mejorando la eficiencia, la calidad de los servicios y la toma de decisiones, entre otros aspectos.









**Transformando la atención médica: el NHS del Reino Unido adopta chatbots e ia para mejorar el acceso y la investigación médica**

En el Reino Unido, la implementación de inteligencia artificial en el Servicio Nacional de Salud (NHS) ha revolucionado la atención médica. La introducción de chatbots y asistentes virtuales ha mejorado significativamente la accesibilidad a la información de salud para los pacientes, ofreciendo respuestas rápidas a consultas comunes y facilitando la gestión de citas médicas. Además, se han explorado aplicaciones más avanzadas de IA en la investigación médica, mejorando la precisión en el diagnóstico y avanzando en el desarrollo de tratamientos personalizados.



**Transformando la atención médica: el NHS del Reino Unido adopta chatbots e ia para mejorar el acceso y la investigación médica**

En Estados Unidos, la Administración de Alimentos y Medicamentos (FDA) ha adoptado la inteligencia artificial para agilizar la revisión y aprobación de medicamentos y dispositivos médicos. Los algoritmos de IA analizan datos clínicos y científicos de manera eficiente, acelerando el proceso de aprobación y mejorando la capacidad de tomar decisiones informadas sobre la seguridad y eficacia de nuevos productos médicos.



**Smart Nation: Singapur utiliza la IA para una gestión urbana eficiente y sostenible**

Singapur, a través de su iniciativa "Smart Nation", ha impulsado el uso de la inteligencia artificial en la gestión del tráfico, la seguridad pública y la mejora de servicios públicos. La ciudad utiliza tecnologías avanzadas, como sensores y análisis de datos, para tomar decisiones informadas en tiempo real, creando un entorno urbano más eficiente y sostenible.



**Infraestructuras inteligentes: Países Bajos moderniza la gestión de carreteras y puentes con IA**

La implementación de la inteligencia artificial en los Países Bajos se ha centrado en la gestión de infraestructuras críticas. La Agencia Holandesa de Carreteras ha adoptado tecnologías avanzadas para el mantenimiento y la inspección de carreteras y puentes, mejorando la eficiencia operativa y reduciendo los costos asociados con la infraestructura vital.



**Avanzando en la salud y la movilidad: Francia implementa la IA en diagnósticos médicos y transporte urbano**

Francia ha aplicado la inteligencia artificial en la salud y la movilidad urbana. En el sector de la salud, la IA se ha utilizado para el análisis de imágenes médicas y la gestión electrónica de registros médicos, mejorando la precisión del diagnóstico y la eficiencia en la atención médica. En el ámbito de la movilidad, proyectos de IA han contribuido a una gestión más eficiente del transporte público y la reducción de congestiones en las ciudades francesas.



**Educación personalizada y robótica: Japón lidera con sistemas de tutoría basados en IA y asistencia a personas mayores**

Japón destaca por la aplicación de inteligencia artificial en la educación y la robótica. Los sistemas de tutoría personalizados basados en IA han mejorado la experiencia educativa de los estudiantes japoneses, adaptándose a sus necesidades individuales. Además, la implementación de la IA en proyectos de robótica y asistencia a personas mayores ha contribuido a la mejora de la calidad de vida y la autonomía de la población mayor.



**Energía del futuro: Australia impulsa la sostenibilidad con sistemas de IA en la gestión de redes eléctricas**

Australia ha implementado la inteligencia artificial con un enfoque especial en la gestión de la energía. Los sistemas de IA se utilizan en la distribución de recursos en redes eléctricas, mejorando la eficiencia y la resiliencia del sistema energético. Este enfoque ha contribuido a la transición hacia fuentes de energía más sostenibles y ha fortalecido la infraestructura energética del país.



**Fronteras eficientes: la Agencia Canadiense de Servicios Fronterizos de Canadá implementa la IA para facilitar el comercio transfronterizo y reforzar la seguridad**

La Agencia Canadiense de Servicios Fronterizos (CBSA) ha adoptado la inteligencia artificial para agilizar los procesos en los cruces fronterizos. Los sistemas de IA ayudan en la detección de riesgos y facilitan el comercio transfronterizo al mejorar la eficiencia en el procesamiento de viajeros y mercancías. Esta implementación ha fortalecido la seguridad fronteriza y ha impulsado la fluidez en el intercambio comercial.

### 1.2.2. Aplicaciones de la IA en la administración pública a nivel nacional, autonómico y local

La administración pública en España también es conocedora del potencial de la IA y está demostrando un creciente interés y compromiso en su adopción. Tanto a nivel nacional, como autonómico y local, la administración pública cuenta con numerosos ejemplos de aplicación práctica de la IA en su actividad.

La Secretaría de Estado de Digitalización e Inteligencia Artificial (SEDIA) del Ministerio de Asuntos Económicos y Transformación Digital ha impulsado la elaboración de un sello nacional que certifique la seguridad, calidad y responsabilidad de los sistemas de inteligencia artificial desarrollados en España.





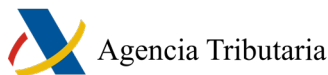
### Detección temprana de enfermedades

El Ministerio de Sanidad ha utilizado la IA en la detección temprana de brotes de enfermedades, como la gripe, a través del análisis de datos y la monitorización de síntomas. Asimismo, durante la pandemia de COVID-19, se han desarrollado modelos de IA para predecir la propagación del virus y gestionar recursos de atención médica.



### Chatbots de atención a la ciudadanía

El SEPE ha implementado chatbots de IA en su sitio web para ayudar a la ciudadanía a obtener información sobre el desempleo, la búsqueda de trabajo y otros servicios relacionados con el empleo.



### Identificación de fraude

La Agencia Tributaria ha utilizado la IA para identificar evasión fiscal y fraude a través del análisis de grandes conjuntos de datos. La IA ayuda a detectar patrones de evasión y a tomar medidas para prevenirlo.



### Gestión del tráfico

El Gobierno Vasco ha utilizado la IA en la optimización del tráfico y el transporte público. Se han desarrollado sistemas de IA para mejorar la gestión de semáforos y la movilidad urbana.



### Asistente virtual de atención turística

El Gobierno de Navarra ha creado un asistente virtual impulsado por IA generativa con el objetivo de mejorar la atención turística a visitantes y ciudadanía y promover la adopción de nuevas tecnologías en el sector turístico.



### Gestión de la movilidad urbana

Barcelona ha utilizado la IA para mejorar la gestión de la movilidad urbana. Han implementado sensores y algoritmos de IA para controlar el tráfico, predecir atascos y mejorar la movilidad en la ciudad.



### Gestión de la calidad del aire

El Ayuntamiento de Valencia ha desarrollado sistemas de IA para la monitorización de la calidad del aire y la gestión de la contaminación. Estos sistemas proporcionan información en tiempo real sobre la calidad del aire en la ciudad y proporcionan recomendaciones de actuación para la mejora de los indicadores.



## Capítulo 2

### El ciclo de vida de un sistema de inteligencia artificial

En este capítulo, exploramos el rápido avance de la inteligencia artificial (IA) y la definición precisa de un sistema de IA según la normativa europea. Se detalla el ciclo de vida de estos sistemas, dividido en dos fases: identificación y evaluación de oportunidades, y diseño, desarrollo e implantación. La gestión integral, destacada en la futura Ley de Inteligencia Artificial, se enfoca en la transparencia, objetividad y confiabilidad. Se resalta la importancia de la gestión de riesgos en todas las etapas, con la clasificación de los sistemas según su nivel de riesgo y la imposición de obligaciones proporcionales.



## 2.1. El sistema de inteligencia artificial

La inteligencia artificial evoluciona a un ritmo imparable, generando innovaciones constantes en tiempos cada vez más cortos que impulsan nuevos debates sobre su uso ético y sobre la gestión de los riesgos asociados a esta tecnología. Este vertiginoso progreso plantea desafíos y oportunidades, requiriendo una adaptabilidad continua en la comprensión de esta tecnología en constante cambio.

La Unión Europea, pionera en la regulación de la IA a nivel global, es consciente de que es necesario que la definición de un sistema de IA sea tan concreta que permita acotar su naturaleza, pero que al mismo tiempo sea lo suficientemente flexible para abarcar cualquier innovación emergente en torno a esta tecnología. Por ello, la normativa comunitaria aporta una definición que combina claridad con adaptabilidad, lo que garantiza una comprensión completa de su alcance y permite a su vez entender mejor los elementos clave para su correcta gestión.

**Un sistema de IA se define como “el software que se desarrolla empleando una o varias de las técnicas y estrategias de IA y que puede, para un conjunto determinado de objetivos definidos por seres humanos, generar información de salida como contenidos, predicciones, recomendaciones o decisiones que influyan en los entornos con los que interactúa”.**

Se constata que esta definición refleja una comprensión integral y coherente de la IA en el ámbito de la administración pública, reconociendo su capacidad de autonomía y su función asistida en la toma de decisiones humanas. En el primer caso, la capacidad de “generar información de salida” y “tomar decisiones” sin intervención directa de personas concuerda con la noción de sistemas para la realización de actos administrativos. En el segundo caso, la referencia a “objetivos definidos por seres humanos” y la generación de información que influye en los entornos se alinea con la idea de sistemas para la realización de actuaciones administrativas, donde la información producida por los modelos de IA se emplea en el proceso de toma de decisiones humanas.



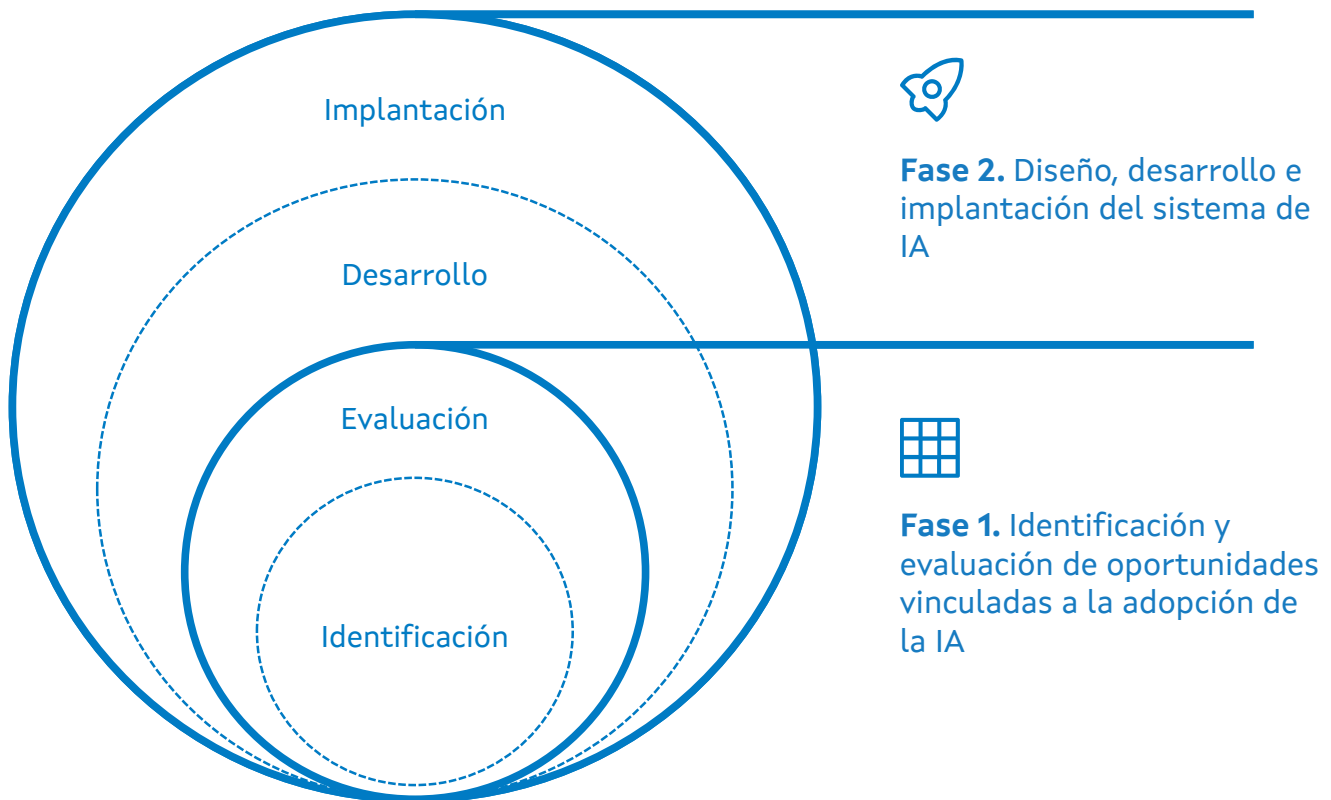




## 2.2. Gestión del ciclo de vida de un sistema de IA

### 2.2.1. El ciclo de vida de un sistema de IA

El ciclo de vida de un sistema de IA es el proceso por el cual una administración pública identifica una barrera o área de mejora que puede ser cubierta mediante la aplicación de IA, analiza la idoneidad de adoptar un sistema de IA y conceptualiza y ejecuta un proyecto para su implantación. Este proceso implica el análisis continuo del riesgo de la implantación de la IA, así como la evaluación de las capacidades de las que dispone la administración para la adopción del sistema. En términos globales, pueden diferenciarse dos grandes fases en el ciclo de vida de un sistema de IA:





## Fase 1. Identificación y evaluación de oportunidades vinculadas a la adopción de la IA

El ciclo de vida de un sistema de IA en la administración pública comienza con la identificación de la barrera o área de mejora y el análisis de la idoneidad de aplicar IA para su cobertura. En este punto, se vuelve esencial que aquellas personas encargadas de identificar oportunidades puedan definir de manera nítida el desafío en términos de políticas públicas o gestión interna al cual buscan dar respuesta mediante la incorporación de un sistema de IA. Es necesario que la administración pública tenga en su ámbito de actuación la competencia para hacer frente al problema o circunstancia de la que deriva la oportunidad.

Una vez identificada la barrera o área de mejora, resulta imprescindible que la administración pública cuente con elementos objetivos y transparentes para su evaluación y potencial transformación en un proyecto para la implantación de un sistema de IA. Estos elementos permiten, por un lado, asegurar la factibilidad del futuro proyecto; y, por otro lado, asegurar su alineación con los requerimientos normativos, éticos o de otra índole que deban observarse en la adopción de la IA.



## Fase 2. Diseño, desarrollo e implantación del sistema de IA

A partir de la evaluación de las oportunidades y de la verificación de su factibilidad, la administración pública ha de concretar los requisitos de arquitectura y funcionalidad que deberá incorporar el sistema de IA que dará materializará dicha oportunidad. Una vez desarrollado el sistema de IA con base en dichos requisitos, adquiere especial relevancia el entrenamiento del sistema, en tanto que permite que el sistema aprenda y mejore su capacidad para hacer análisis más precisos.

El ciclo de vida del sistema de IA concluye con su validación, despliegue y mantenimiento. Estas etapas son procesos continuos en los que se requiere una colaboración constante para asegurar que el sistema de IA opera de forma confiable, conforme a los objetivos que se habían definido en un inicio y manteniendo la observancia de los requerimientos normativos asociados a la adopción del sistema.

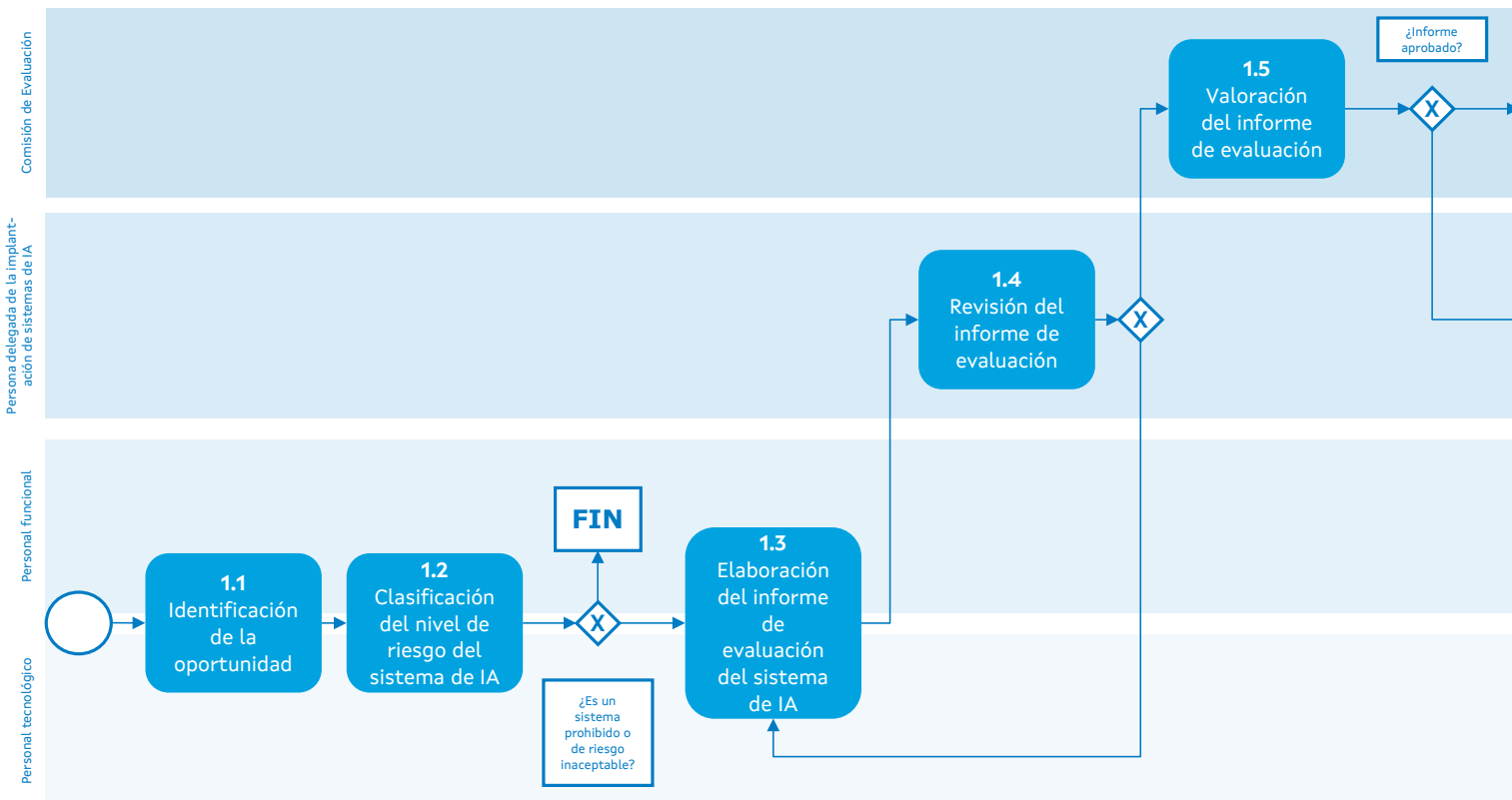
## 2.3. Flujo de gestión del ciclo de vida de un sistema de IA

La gestión de un sistema de IA abarca no solo las actividades necesarias para la gestión de su ciclo de vida, sino especialmente las actividades vinculadas con la gestión de los riesgos inherentes al diseño, desarrollo e implantación de las soluciones de IA. La utilización de

herramientas estándares de soporte a estas actividades aporta al proceso global la transparencia, objetividad y confiabilidad necesarias para la óptima gestión del sistema de inteligencia artificial.

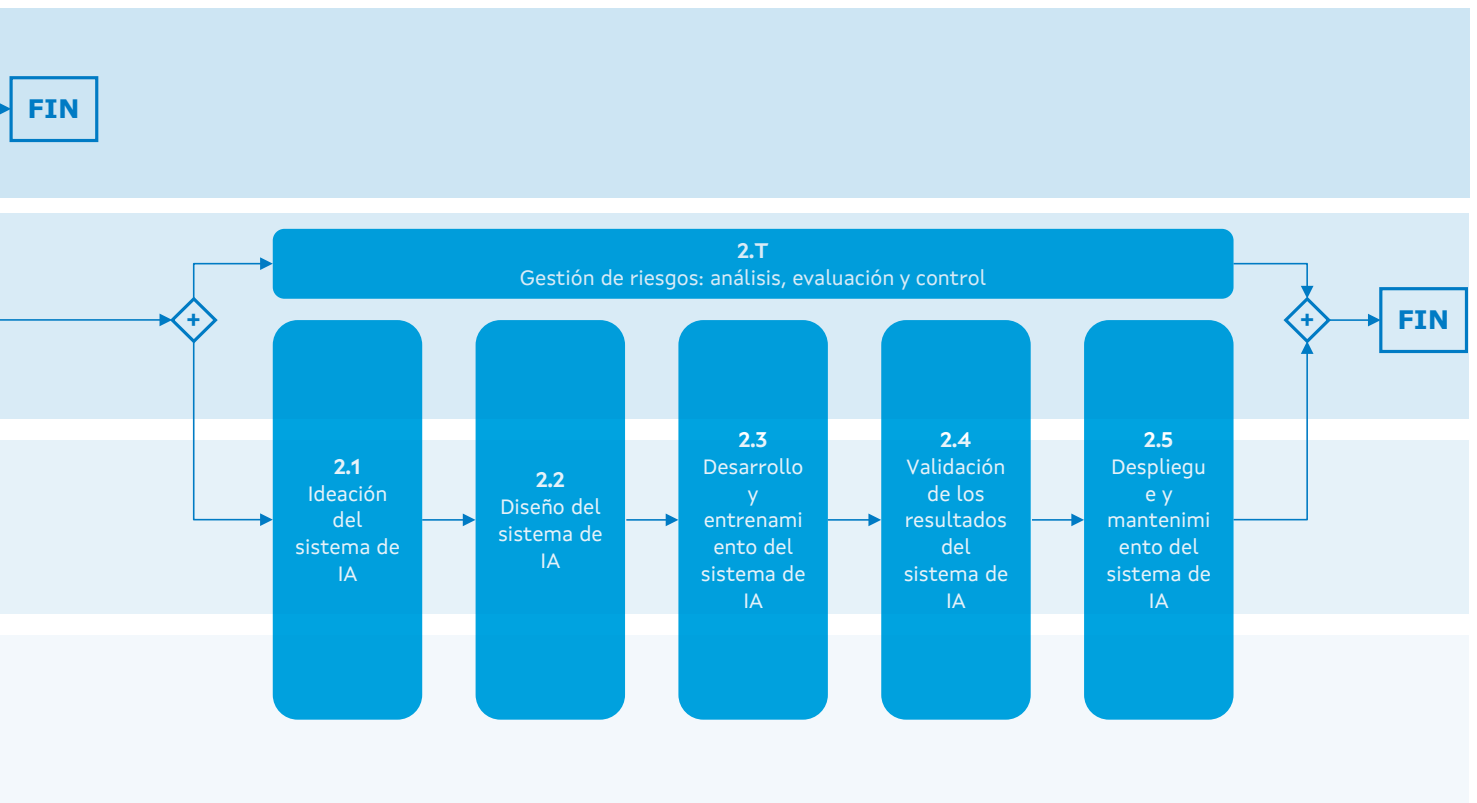


### Fase 1. Identificación y evaluación de oportunidades vinculadas a la adopción de la IA





## Fase 2. Diseño, desarrollo e implantación del sistema de IA





### 2.3.1. Agentes intervinientes en la gestión de un sistema de inteligencia artificial

La relación de los agentes que intervienen en la gestión de un sistema de IA puede variar de una administración pública a otra, tanto en su nomenclatura como en su ámbito de actuación específico. No obstante, las actividades que se desarrollan en la gestión del ciclo de vida de un sistema de IA han de ser cubiertas en cualquier administración pública y deben ser ejecutadas por agentes cuyo rol y funciones deben asemejarse a las que aquí se plantean.



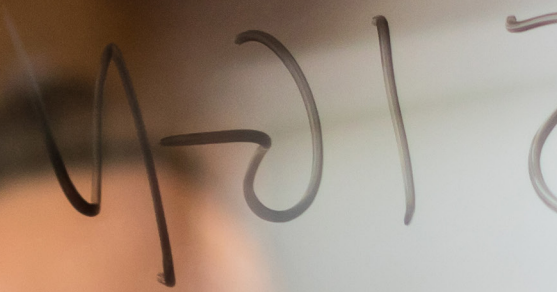
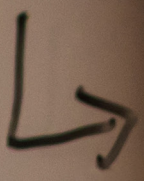


REVIEW

ASSESS

[Redacted]

[Redacted]





## Persona delegada de la implantación de sistemas de IA

La persona delegada de la implantación de sistemas de IA es una persona de la administración pública cuyas funciones se vinculan a supervisar las oportunidades relacionadas con la IA que se identifiquen en la organización, ofrecer asesoramiento y evaluar las mismas para su elevación y debate en la Comisión de Evaluación; y apoyar la gestión de riesgos de las mismas a lo largo de su diseño, desarrollo e implantación.

### ¿Cuál es su rol?

Supervisar y validar el análisis de las oportunidades de IA para su elevación a la Comisión de Evaluación de IA.

### ¿Cuáles son sus funciones?

- Garantizar el cumplimiento de la normativa de IA
- Monitorizar el estado del arte de la IA y apoyar la identificación de oportunidades de IA
- Supervisar la evaluación de oportunidades de IA
- Apoyar la gestión de riesgos de proyectos y sistemas de IA



## Comisión de evaluación

La Comisión de Evaluación es un órgano formado por personal funcional, personal TI y personal experto de la administración, encargado de evaluar las oportunidades vinculadas a la IA que le sean elevadas así como de supervisar la gestión de riesgos a lo largo del diseño, desarrollo e implantación de la solución.

La Comisión de Evaluación se reúne de forma periódica para revisar las oportunidades de IA que le hayan sido elevadas, así como para revisar la gestión de riesgos de las soluciones de IA que se encuentren en fase de diseño, desarrollo e implantación. El método de decisión de este órgano ha de fomentar la unanimidad de sus integrantes y ha de garantizar que las personas responsables de las oportunidades de IA que han sido elevadas sean escuchadas de cara a la valoración de las mismas.

### ¿Cuál es su rol?

Evaluar las oportunidades vinculadas a la IA que le sean elevadas; y autorizar y supervisar la implantación de soluciones de IA.

### ¿Cuáles son sus funciones?

- Evaluar las oportunidades de IA que le sean elevadas
- Revisar la gestión de riesgos de las soluciones de IA





## Personal funcional y personal TI

El personal funcional y el personal TI se encuentran involucrados en el conjunto de las actividades de carácter ejecutivo que componen el proceso de gestión del ciclo de vida de un sistema de IA en la administración pública, desde la identificación y evaluación preliminar de la oportunidad de IA hasta su implementación final. La responsabilidad de estas actividades es compartida entre el personal funcional y el personal TI.

### ¿Cuál es su rol?

Liderar las actividades de carácter ejecutivo que componen el proceso de gestión del ciclo de vida de un sistema de IA en la administración pública.

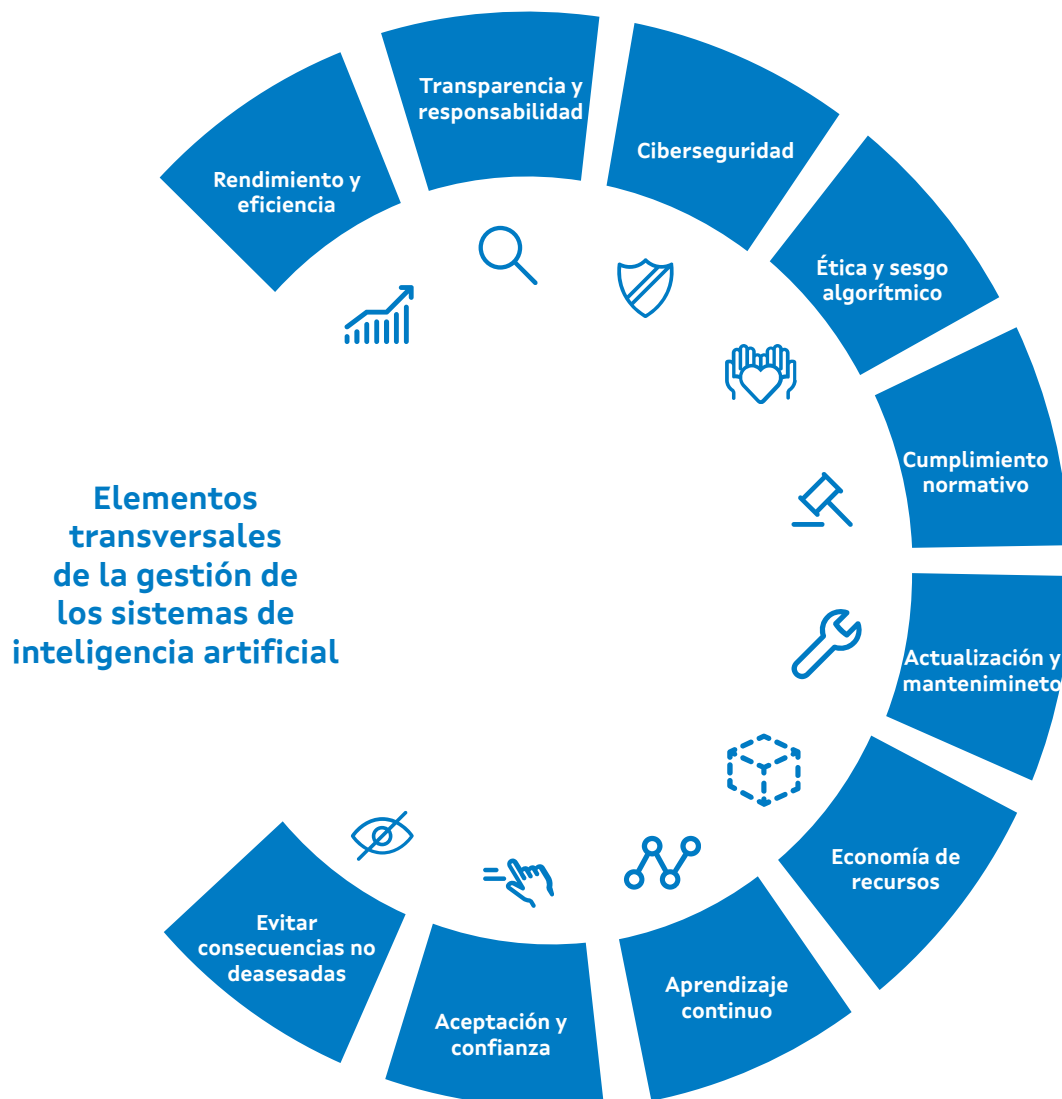
### ¿Cuáles son sus funciones?

- Identificar oportunidades de IA
- Clasificar el nivel de riesgo del sistema de IA
- Evaluar el sistema de IA de forma global de cara a su valoración por la Comisión de Evaluación
- Liderar el diseño, desarrollo e implantación del sistema de IA



### 2.3.2. Elementos transversales de gestión

La gestión del ciclo de vida de un sistema de IA es un proceso integral que permite asegurar que el sistema observa en todo momento las condiciones que garantizan su idoneidad, en términos de adecuación a la necesidad que cubre y de rendimiento; en términos de uso confiable y sostenible en el tiempo; en términos de cumplimiento normativo y de los requerimientos éticos asociados a la tecnología; y en términos de mejora continua.





#### **Rendimiento y eficiencia**

Una gestión adecuada garantiza que el sistema de IA funcione de manera eficiente y cumpla con sus objetivos de rendimiento. Esto implica optimizar los algoritmos, los recursos informáticos y los conjuntos de datos para lograr un funcionamiento eficiente y maximizar el rendimiento.



#### **Transparencia y responsabilidad**

La gestión adecuada de un sistema de IA incluye la documentación y la trazabilidad de las decisiones que toma. Esto es esencial para garantizar la transparencia y la rendición de cuentas en el caso de problemas o decisiones incorrectas.



#### **Ciberseguridad**

La ciberseguridad es una preocupación importante en la gestión de sistemas de IA, especialmente cuando se trata de aplicaciones críticas o sensibles. Una gestión adecuada debe abordar la seguridad de los datos y proteger el sistema contra amenazas y ataques maliciosos.



#### **Ética y sesgo algorítmico**

La gestión de la IA implica abordar cuestiones éticas, como el sesgo en los algoritmos. Es importante garantizar que el sistema de IA no discrimine a ciertos grupos o individuos y que se ajuste a principios éticos sólidos.



#### **Cumplimiento normativo**

La regulación en torno a la IA está en constante evolución. La gestión adecuada debe garantizar el cumplimiento de las leyes y regulaciones aplicables, lo que incluye el cumplimiento de regulaciones de privacidad de datos y derechos de los ciudadanos.



#### **Economía de recursos**

La gestión eficiente de los recursos, como el poder de procesamiento y el almacenamiento de datos, es esencial para evitar el desperdicio y garantizar un uso rentable de la infraestructura de inteligencia artificial.



#### **Aprendizaje continuo**

La IA se beneficia del aprendizaje continuo. La gestión adecuada incluye la capacidad de adaptar y mejorar el sistema a medida que se recopilan y analizan más datos y se obtienen más conocimientos.



#### **Aceptación y confianza**

La gestión de la IA también se relaciona con la percepción y la confianza de los usuarios y las partes interesadas. Un sistema de IA gestionado adecuadamente es más propenso a ganar la aceptación y la confianza de los usuarios y, en última instancia, a ser más efectivo.



#### **Evitar consecuencias no deseadas**

La IA puede tener impactos significativos en la sociedad y en la toma de decisiones. La gestión adecuada ayuda a evitar consecuencias no deseadas, como la discriminación, la falta de privacidad y la pérdida de empleo, y a mitigar posibles riesgos.

### 2.3.3. La importancia de la gestión de riesgos

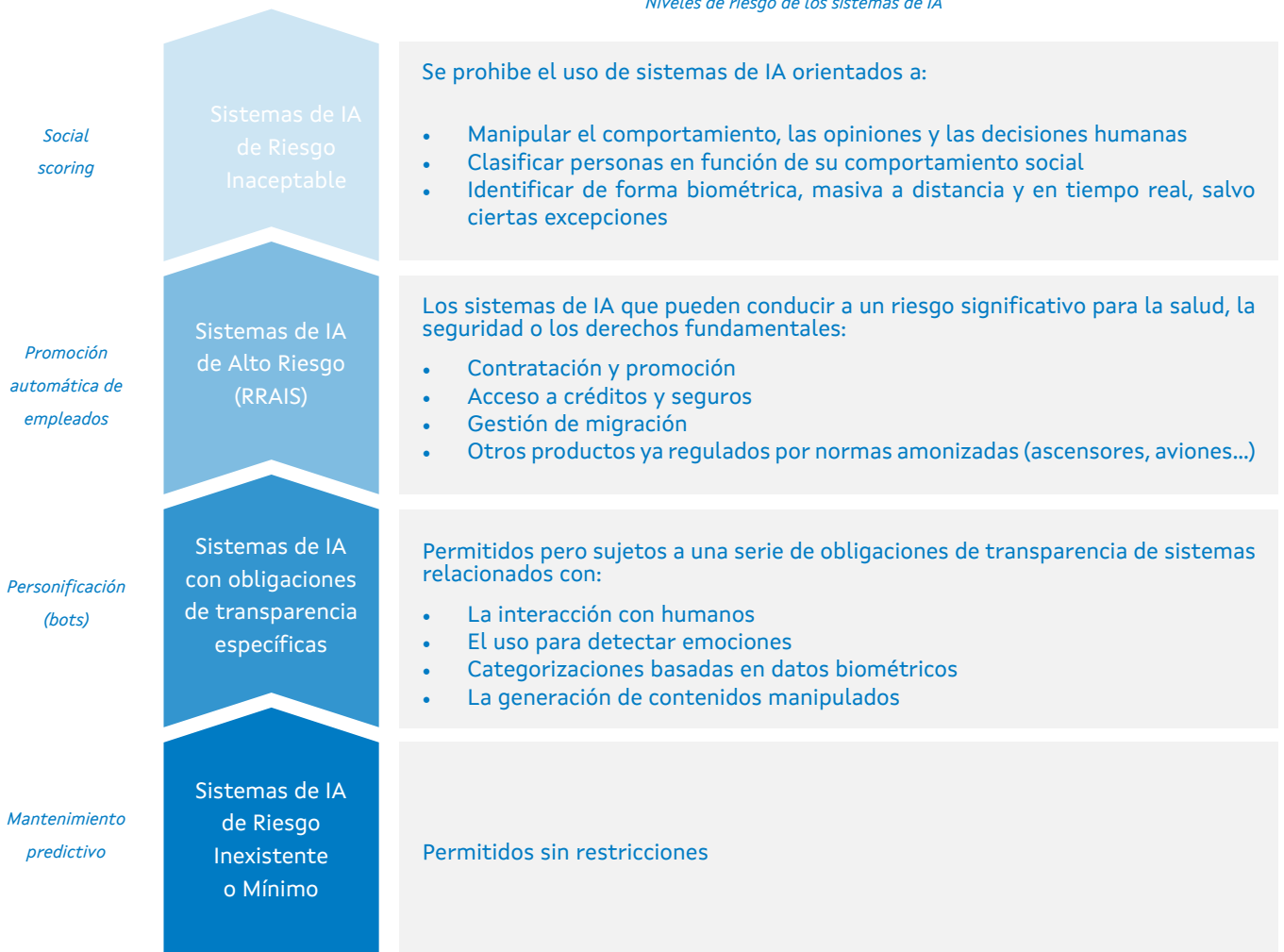
La gestión de riesgos desempeña un papel fundamental en el éxito de cualquier proyecto, especialmente en el ámbito tecnológico. Los proyectos tecnológicos a menudo involucran avances rápidos y cambios en las condiciones del entorno, lo que puede generar riesgos imprevistos. La identificación temprana de estos riesgos permite desarrollar estrategias de mitigación y contingencia, minimizando así el impacto negativo en el desarrollo de los proyectos.

En el caso específico de proyectos de IA, la gestión de riesgos cobra una relevancia aún mayor. La complejidad inherente a la creación de sistemas de IA, como la disponibilidad de datos adaptados, la interpretación de los modelos de IA o el uso ético de esta tecnología ponen de manifiesto esta importancia. Por ello, implementar

medidas de prevención o mitigación de riesgos es esencial para garantizar la equidad y la seguridad en el despliegue de soluciones de IA.

Consciente de esta realidad, la normativa europea sitúa el riesgo como elemento central de la gestión de los sistemas de IA. En concreto, la normativa permite trasladar los elementos transversales de gestión previamente señalados a una escala de riesgos a través de los cuales se materializa la gestión del ciclo de vida del sistema. Así, los sistemas de IA deben clasificarse en función de su nivel de riesgo (inaceptable, alto, limitado y mínimo) y las personas responsables de su gestión deben, de forma consecuente, aplicar determinadas obligaciones y restricciones en función de cada nivel de riesgo.

Niveles de riesgo de los sistemas de IA





La normativa comunitaria orienta una serie de reflexiones para que la administración pública que está valorando un potencial sistema de IA conozca su nivel de riesgo. Como se indicaba anteriormente, en función del resultado de dichas reflexiones, las personas responsables deberán aplicar unas u otras medidas de gestión de riesgos, conforme al modelo de gestión de riesgos propio de la administración pública.

**Reflexiones para la clasificación del nivel del riesgo de un sistema de IA:**





¿El sistema de IA se sirve de técnicas subliminales que pueden trascender la conciencia de una persona con el objetivo de alterar de manera sustancial su comportamiento de un modo que provoque o sea razonablemente probable que provoque perjuicios físicos o psicológicos a esa persona o a otra, o que tenga ese efecto?

¿El sistema de IA puede aprovechar alguna de las vulnerabilidades de un grupo específico de personas derivadas de su edad o discapacidad, o de una situación social o económica específica, con el objetivo de alterar de manera sustancial el comportamiento de una persona que pertenezca a dicho grupo de un modo que provoque o sea razonablemente probable que provoque perjuicios físicos o psicológicos a esa persona o a otra, o que tenga ese efecto?

¿El sistema de IA ha sido diseñado con el fin de evaluar o clasificar a las personas físicas durante un período determinado de tiempo atendiendo a su comportamiento social o a características personales o de su personalidad conocidas o predichas, de forma que la puntuación ciudadana resultante provoque una o varias de las situaciones siguientes mencionadas a continuación?

i. un trato perjudicial o desfavorable hacia determinadas personas físicas o grupos de personas físicas en contextos sociales que no guarden relación con los contextos donde se generaron o recabaron los datos originalmente.

ii. un trato perjudicial o desfavorable hacia determinadas personas físicas o grupos de personas físicas que es injustificado o desproporcionado con respecto a su comportamiento social o la gravedad de este.



### Según la finalidad del sistema

¿La finalidad de su sistema IA es la identificación biométrica remota(\*)?

*(\*)Designa un sistema de IA destinado a identificar a personas físicas generalmente a distancia, sin su participación activa, comparando sus datos biométricos con los que figuran en un repositorio de datos de referencia.*

¿La finalidad de su sistema IA es ser utilizado como componentes de seguridad en la gestión y funcionamiento de infraestructuras digitales críticas, el tráfico rodado y del suministro de agua, gas, calefacción y electricidad?

¿La finalidad de su sistema IA es determinar el acceso o la admisión de personas físicas a programas o centros educativos y de formación profesional a todos los niveles o para asignar a personas físicas a dichos programas o centros?

¿La finalidad de su sistema IA es evaluar los resultados del aprendizaje, también cuando dichos resultados se utilicen para orientar el proceso de aprendizaje de las personas físicas en programas o centros educativos y de formación profesional a todos los niveles?

¿La finalidad de su sistema IA es la contratación o la selección de personas físicas, en particular para publicar anuncios de empleo específicos, analizar y filtrar las solicitudes de empleo y evaluar a los candidatos?

¿La finalidad de su sistema IA es tomar decisiones relativas a la promoción y a la rescisión de relaciones contractuales de índole laboral, para la asignación de tareas a partir de comportamientos individuales o rasgos o características personales y para realizar un seguimiento y una evaluación del rendimiento y el comportamiento de las personas en el marco de dichas relaciones?

¿La finalidad de su sistema IA es ser utilizado por las autoridades públicas o en su nombre, para evaluar la admisibilidad de las personas físicas para acceder a servicios y ayudas esenciales de asistencia pública, así como para conceder, reducir, retirar o recuperar dichos servicios y ayudas?



¿La finalidad de su sistema IA es evaluar la solvencia de personas físicas o establecer su calificación crediticia, salvo los sistemas de IA puestos en servicio por proveedores que son microempresas y pequeñas empresas, tal como se definen en el anexo de la Recomendación 2003/361/CE de la Comisión, para su uso propio?

¿La finalidad de su sistema IA es el envío o el establecimiento de prioridades en el envío de servicios de primera intervención en situaciones de emergencia, por ejemplo bomberos y servicios de asistencia médica?

¿La finalidad de su sistema IA es la evaluación de riesgos y la fijación de precios en relación con las personas físicas en el caso de los seguros de vida y de salud, con excepción de los sistemas de IA puestos en servicio por proveedores que sean microempresas y pequeñas empresas, tal como se definen en el anexo de la Recomendación 2003/361/CE de la Comisión, para su propio uso.

¿La finalidad de su sistema IA es ser utilizado por las autoridades encargadas de la aplicación de la ley o en su nombre para evaluar el riesgo de que una persona física cometa una infracción o reincida o el riesgo de que una persona física se convierta en posible víctima de infracciones penales?

¿La finalidad de su sistema IA es ser utilizado por las autoridades encargadas de la aplicación de la ley o en su nombre como polígrafos y herramientas similares, o para detectar el estado emocional de una persona física?

¿La finalidad de su sistema IA es ser utilizado por las autoridades encargadas de la aplicación de la ley o en su nombre para evaluar la fiabilidad de las pruebas durante la investigación o el enjuiciamiento de infracciones penales?

¿La finalidad de su sistema IA es ser utilizado por las autoridades encargadas de la aplicación de la ley o en su nombre para predecir la comisión o reiteración de una infracción penal real o potencial a partir de la elaboración de perfiles de personas físicas mencionada en el artículo 3, punto 4, de la Directiva (UE) 2016/680, o para evaluar rasgos y



características de la personalidad o comportamientos delictivos pasados de personas físicas o grupos?

¿La finalidad de su sistema IA es ser utilizado por las autoridades encargadas de la aplicación de la ley o en su nombre para elaborar perfiles de personas físicas, como se menciona en el artículo 3, punto 4, de la Directiva (UE) 2016/680, durante la detección, la investigación o el enjuiciamiento de infracciones penales.

¿La finalidad de su sistema IA es ser utilizado por las autoridades públicas competentes o en su nombre como polígrafos y herramientas similares, o para detectar el estado emocional de una persona física?

¿La finalidad de su sistema IA es ser utilizado por las autoridades públicas competentes o en su nombre para evaluar un riesgo, como un riesgo para la seguridad, la salud o relativo a la migración irregular, que plantee una persona física que tenga la intención de entrar o haya entrado en el territorio de un Estado miembro;

¿La finalidad de su sistema IA es ser utilizado por las autoridades públicas competentes o en su nombre para examinar las solicitudes de asilo, visado y permiso de residencia, y las reclamaciones asociadas con respecto a la admisibilidad de las personas físicas solicitantes.

¿La finalidad de su sistema IA es ser utilizado por una autoridad judicial o en su nombre para interpretar hechos o la ley, así como para aplicar la ley a un conjunto concreto de hechos?





## Según el sector de aplicación del sistema

¿Su sistema se constituye en sí mismo como un producto o un componente de seguridad de éste dentro de los sectores o ámbitos listados a continuación?

- a. Máquinas
- b. Seguridad de los juguetes
- c. Embarcaciones de recreo y motos acuáticas
- d. Ascensores y componentes de seguridad para ascensores
- e. Aparatos y sistemas de protección para uso en atmósferas potencialmente explosivas
- f. Equipos radioeléctricos
- g. Equipos a presión
- h. Instalaciones de transporte por cable
- i. Equipos de protección individual
- j. Aparatos que queman combustibles gaseosos
- k. Productos sanitarios
- l. Productos sanitarios para diagnóstico in vitro
- m. Seguridad de la aviación civil
- n. Homologación de vehículos de dos o tres ruedas y cuatriciclos
- o. Homologación de vehículos agrícolas o forestales, y a la vigilancia del mercado de dichos vehículos
- p. Equipos marinos
- q. Interoperabilidad del sistema ferroviario dentro de la Unión Europea
- r. Homologación y vigilancia del mercado de los vehículos de motor y sus remolques y de los sistemas, los componentes y las unidades técnicas independientes destinados a dichos vehículos
- s. Normas comunes en el ámbito de la aviación civil y por el que se crea una Agencia de la Unión Europea para la Seguridad Aérea en la medida en la que afecte al diseño, la producción y la comercialización de aeronaves contemplados en el artículo 2 (IA Act), apartado 1, letras a) y b), cuando se refiera a aeronaves no tripuladas y sus motores, hélices, componentes y equipos para controlarlas a distancia.
- t. Requisitos de homologación de tipo de los vehículos de motor y de sus remolques, así como los sistemas, componentes y unidades técnicas independientes destinados a esos vehículos, en lo que respecta a su seguridad general y a la protección de los ocupantes de los vehículos y de los usuarios vulnerables de la vía pública.

## SISTEMAS DE IA CON OBLIGACIONES DE TRANSPARENCIA ESPECÍFICAS



Si el sistema de IA está destinado a interactuar con personas físicas, ¿está diseñado y desarrollado de forma que dichas personas estén informadas de que están interactuando con un sistema de IA, excepto en las situaciones en las que esto resulte evidente desde el punto de vista de una persona jurídica que esté razonablemente informada, observadora y circunspecta, dadas las circunstancias y el contexto de utilización?

*(Conteste no a esta pregunta si el sistema de IA será autorizado por la ley para fines de detección, prevención, investigación o enjuiciamiento de infracciones penales, con sujeción a las correspondientes salvaguardas de los derechos y libertades de terceros, salvo que este sistema esté a disposición del público para denunciar una infracción penal).*

Si el sistema de IA es o utiliza un sistema de categorización biométrica, ¿se informa del funcionamiento del sistema a las personas físicas expuestas a él?

*(Conteste no a esta pregunta si el sistema de IA será utilizados para la categorización biométrica, autorizado por ley para detectar, impedir e investigar infracciones penales, con sujeción a las correspondientes salvaguardas de los derechos y libertades de terceros.)*

Si el sistema IA es o utiliza un sistema de IA que genere o manipule contenido de imagen, sonido o vídeo que se asemeje notablemente a personas, objetos, lugares u otras entidades o sucesos existentes, y que pueda inducir erróneamente a una persona a pensar que son auténticos o verídicos (ultrafalsificación), ¿hace público que el contenido ha sido generado de forma artificial o manipulado?

*(Conteste no a esta pregunta si el uso del sistema de IA será autorizado por ley para detectar, impedir, investigar y enjuiciar infracciones penales o cuando el contenido forme parte de una obra o programa manifiestamente creativo, satírico, artístico o ficticio, con sujeción a las correspondientes salvaguardas de los derechos y libertades de terceros.)*

Si el sistema IA es o utiliza un sistema de IA que genere o manipule contenido de imagen, sonido o vídeo que se asemeje notablemente a personas, objetos, lugares u otras entidades o sucesos existentes, y que pueda inducir erróneamente a una persona a pensar que son auténticos o verídicos (ultrafalsificación), ¿hace público que el contenido ha sido generado de forma artificial o manipulado?

*(Conteste no a esta pregunta si el uso del sistema de IA será autorizado por ley para detectar, impedir, investigar y enjuiciar infracciones penales o cuando el contenido forme parte de una obra o programa manifiestamente creativo, satírico, artístico o ficticio, con sujeción a las correspondientes salvaguardias de los derechos y libertades de terceros.)*



## MODELO FUNDACIONAL O SISTEMA DE IA GENERATIVA



¿El sistema de IA constituye sistema de IA Generativa o se basa en el desarrollo de un modelo fundacional?

Modelo fundacional: un modelo de IA que se entrena con gran cantidad de datos a escala, diseñado para proporcionar resultados a nivel general y puede adaptarse a una amplia gama de tareas distintivas.

Sistema de IA generativa: modelos fundacionales utilizados en sistemas de IA destinados específicamente a generar, con distintos niveles de autonomía, contenidos como texto complejo, imágenes, audio o vídeo.



## Capítulo 3

### Fases del ciclo de vida de un sistema de inteligencia artificial

En este capítulo, exploraremos en detalle las distintas fases que conforman el ciclo de vida de un sistema de inteligencia artificial (IA), desde su concepción hasta su plena implementación. Proporcionaremos una visión integral del proceso, destacando las actividades esenciales que componen el flujo de gestión de este ciclo. En cada etapa, se abordarán las cuestiones clave, se identificarán los roles responsables y se analizarán errores comunes que pueden surgir durante la ejecución.



A continuación, se presenta la visión global y se desarrollan las actividades que conforman el flujograma de gestión del ciclo de vida de un sistema de IA, reflejando en cada caso las cuestiones clave, los roles responsables y errores comunes en su ejecución.



## **Fase 1.** Identificación y evaluación de oportunidades vinculadas a la adopción de la IA

**1.1** Identificación del área de mejora y análisis de la idoneidad de aplicar IA

**1.2** Clasificación del nivel de riesgo del sistema de IA

**1.3** Evaluación estándar del sistema de IA

**1.4** Revisión preliminar de la evaluación

**1.5** Valoración final de la evaluación



## **Fase 2.** Diseño, desarrollo e implantación del sistema de IA

**2.1** Ideación del sistema de IA

**2.2** Diseño del sistema de IA

**2.3** Desarrollo y entrenamiento del sistema de IA

**2.4** Validación de los resultados del sistema de IA

**2.5** Despliegue y mantenimiento del sistema de IA

## 3.1

### Fase 1. Identificación y evaluación de oportunidades vinculadas a la adopción de la IA

Las actividades de identificación y evaluación de oportunidades de IA permiten que todas las partes interesadas en la implantación de sistemas de IA puedan conocer y evaluar de forma objetiva y compartida posibles sistemas de IA a implantar en la administración pública. En todo caso, estas actividades pueden modularse a las capacidades o modelo de funcionamiento de cada administración pública que las implementa.

#### 3.1.1 Identificación del área de mejora y análisis de la idoneidad de aplicar IA



El proceso global de incorporación de la IA en la administración pública empieza por definir de manera precisa la barrera, problema o área de mejora que se busca abordar a través de la implantación de la IA. Es esencial acotar el área de mejora, identificar sus implicaciones tanto para usuarios externos como internos, y justificar la implementación de la IA como una solución valiosa, desde la perspectiva de un acto administrativo (toma de decisiones autónoma) o una actuación administrativa (apoyo a la toma de decisiones o mejora de servicios).

Para ello, es necesario determinar si la administración pública tiene la competencia legal y normativa para abordar el problema, si se requiere asociación con otros agentes, y si cuenta con los recursos necesarios. Además, tras definir el problema, es crucial evaluar el estado actual de respuesta de la administración, identificar limitaciones y oportunidades (prefactibilidad), y entender cómo la implementación de la IA podría mejorar la situación.

##### Cuestiones clave

La oportunidad la puede identificar tanto el personal funcional como el personal TI de la administración pública.

La persona que identifique la oportunidad deberá comunicarla a la persona delegada de la implantación de sistemas de IA, dando inicio al proceso de gestión de un sistema de IA.

##### Algunas reflexiones

¿Cuál es el problema que se busca solucionar o la necesidad que se busca cubrir?

¿Cuáles son los agentes (internos o externos) que se ven afectados por el problema o la necesidad? ¿Es posible cuantificar dichos agentes?

¿Por qué es prioritario resolver este problema o cubrir esta necesidad?

¿Existe algún caso parecido de uso de la IA que haya sido implementado anteriormente? ¿Cuál?

¿Cuáles son los objetivos de la potencial implementación de la IA (o los beneficios que se espera obtener) y cómo impactan en los agentes afectados por el problema o la necesidad?

##### Errores comunes

- **Falta de comprensión:** no comprender completamente las capacidades y limitaciones de la IA puede llevar a identificar oportunidades poco realistas o inapropiadas.
- **Falta de enfoque en los problemas correctos:** identificar áreas que no se beneficiarían significativamente de la IA o que tienen necesidades más urgentes puede llevar a la asignación ineficiente de recursos.
- **Falta de colaboración:** no involucrar a agentes clave o no buscar colaboraciones externas puede limitar la eficacia de la implementación.
- **Sobrestimar los beneficios:** tener expectativas poco realistas sobre los beneficios de la IA puede llevar a decepciones y desilusión.
- **No involucrar a los colectivos impactados:** no consultar a los colectivos y partes interesadas puede llevar a soluciones que no se ajusten a sus necesidades o valores.



### 3.1.2 Clasificación del nivel de riesgo del sistema de IA



La clasificación del nivel de riesgo de un sistema de IA es una actividad esencial en el proceso de gestión de un sistema de IA, en tanto que permite determinar si la oportunidad puede dar lugar a un sistema aceptable en los términos de riesgo que establece la normativa europea. La propia normativa concreta una serie de requisitos para que la administración pública que está valorando un potencial sistema de IA conozca su nivel de riesgo.

#### Cuestiones clave

Como primer elemento de filtrado de las oportunidades de IA, se rellenará un cuestionario para clasificar el potencial sistema de IA según los niveles de riesgo previstos en el AI Act.

En el caso de que se trate de un sistema prohibido o un nivel de riesgo inaceptable, el proceso terminará en este punto.

En función del nivel de riesgo resultante (salvo en el caso previamente indicado), se incorporarán preguntas adicionales en el "Informe de evaluación".



La clasificación del nivel de riesgo de un sistema de IA debe realizarse a través de un "**Cuestionario de clasificación del nivel de riesgo del sistema de IA**" que recoge las todas preguntas necesarias para que la administración pública conozca en qué nivel de riesgo se clasifica el sistema de IA que está valorando. En función del riesgo, se podrá optar por continuar con la evaluación detallada del sistema de IA.

#### Errores comunes

- **Subestimar riesgos éticos y de privacidad:** centrarse excesivamente en los riesgos técnicos y operativos, sin considerar los aspectos éticos y de privacidad, puede resultar en falta de cumplimiento de la normativa.
- **Falta de evaluación del impacto social:** no analizar cómo el sistema de IA puede afectar a la sociedad en su conjunto, incluyendo el empleo, la economía y la equidad, puede llevar a consecuencias no deseadas.
- **Falta de transparencia:** no proporcionar información adecuada a las partes interesadas sobre cómo funciona el sistema de IA y cómo se toman las decisiones puede generar desconfianza.

### 3.1.3 Evaluación estándar del sistema de IA



Después de establecer el área de mejora y confirmar que la IA es la herramienta adecuada para respaldar la solución, y antes de determinar el interés de acometer un proyecto para su implantación, es crucial evaluar una serie de aspectos fundamentales para garantizar la viabilidad del proyecto. Esta evaluación debe contener elementos estándares que permitan que todas las partes interesadas puedan analizar el sistema de forma objetiva, lo que potencia la transparencia y la cohesión de las decisiones que se adopten respecto a la implantación de la IA.

#### Cuestiones clave

Una vez clasificado el nivel de riesgo del sistema de IA, se procederá a la elaboración de un informe de evaluación del sistema.

La colaboración del personal funcional y el personal TI es indispensable para el éxito de esta evaluación, así como la comunicación continua con la persona delegada de la implantación de sistemas de IA.



La evaluación de un sistema de IA se articula a través de un **"Informe de evaluación"** que valora la disponibilidad de datos adaptados a un modelo de IA, la disponibilidad de capacidades y competencias internas para la implantación del sistema de IA, los requisitos de integración del sistema de IA con los sistemas informáticos o la necesidad de recursos económicos y humanos para la implantación del sistema

#### Errores comunes

- **No evaluar el sesgo y la equidad:** no considerar la posibilidad de sesgo en los datos de entrenamiento y en los algoritmos puede llevar a decisiones discriminatorias y perjudiciales.
- **No considerar el tamaño del conjunto de datos:** contar con un conjunto de datos pequeño o no representativo puede dar como resultado modelos de IA deficientes y poco confiables.
- **Ignorar la ciberseguridad:** no evaluar adecuadamente las vulnerabilidades de seguridad en el sistema de IA puede dar lugar a riesgos de ataques cibernéticos y brechas de datos.
- **No considerar la robustez y la resiliencia:** no evaluar cómo el sistema de IA puede comportarse en situaciones inesperadas o bajo presión puede ser un error significativo.
- **Ignorar la infraestructura necesaria:** no evaluar adecuadamente si la infraestructura tecnológica existente es compatible con el sistema de IA puede llevar a problemas de integración y retrasos.
- **Subestimar la carga computacional:** no estimar adecuadamente los recursos necesarios para entrenar y ejecutar el sistema puede dar como resultado problemas de rendimiento y costes imprevistos.



### **Estructura y contenido del informe de evaluación de un sistema de IA en la administración pública**

A partir de estas premisas, a continuación, se presenta la estructura y contenido del informe de evaluación de un sistema de IA en la administración pública:



## a) Descripción de la oportunidad de IA

Descripción del sistema de IA objeto de evaluación, indicando la necesidad a la que da respuesta y los principales beneficios esperados mediante su implantación (ver apartado "3.1.1. Identificación del área de mejora y análisis de la idoneidad de aplicar IA").

## b) Tareas a realizar y ramas de la IA empleadas por el sistema

Descripción de en qué medida es idóneo que la necesidad sea cubierta mediante una de las siguientes tecnologías de IA (o por una tecnología de IA alternativa); y en qué medida dicha necesidad no puede ser cubierta por una tecnología que no aplique IA.

Tareas del sistema	Descripción	Rama de la IA a emplear	Justificación
Reconocimiento	Categorización de imágenes, textos y videos, por medio de la identificación de sus características claves	<i>*texto libre*</i>	<i>*texto libre*</i>
Detección de eventos	Detección de patrones y anomalías en un conjunto de datos	<i>*texto libre*</i>	<i>*texto libre*</i>
Predicción	Predicción de un estado futuro basado en datos históricos	<i>*texto libre*</i>	<i>*texto libre*</i>
Personalización	Desarrollo de perfiles de usuario que, con base en datos generados por sus acciones, aumentan el nivel de personalización con el tiempo	<i>*texto libre*</i>	<i>*texto libre*</i>



Tareas del sistema	Descripción	Rama de la IA a emplear	Justificación
Apoyo a la interacción	Desarrollo de herramientas que potencian la interacción entre personas y máquinas	<i>*texto libre*</i>	<i>*texto libre*</i>
Optimización enfocada en objetivos	Optimización de los procesos dentro de una organización a través de la simulación de escenarios	<i>*texto libre*</i>	<i>*texto libre*</i>
Razonamiento con estructuras de conocimiento	Establecimiento de relaciones causales entre los datos disponibles y un evento futuro no existente, focalizándose en la causalidad entre variables	<i>*texto libre*</i>	<i>*texto libre*</i>
+	<i>*texto libre*</i>	<i>*texto libre*</i>	<i>*texto libre*</i>

### c) Nivel de riesgo del sistema de IA

Identificación del nivel de riesgo del sistema de IA y concreción de las medidas para la gestión de los riesgos del sistema (ver apartado "2.3.3. La importancia de la gestión de riesgos" y apartado "3.1.2. Clasificación del nivel de riesgo del sistema de IA").





## d) Disponibilidad de datos adaptados a un modelo de IA

Descripción del sistema de IA objeto de evaluación, indicando la necesidad a la que da respuesta y los principales beneficios esperados mediante su implantación (ver apartado "3.1.1. Identificación del área de mejora y análisis de la idoneidad de aplicar IA").

Preguntas	Sí	No	Comentarios
¿Tiene suficientes datos para que el modelo aprenda?	<input type="checkbox"/>	<input type="checkbox"/>	<i>*texto libre*</i>
¿Entiende los efectos posteriores del uso de datos de esta manera?	<input type="checkbox"/>	<input type="checkbox"/>	<i>*texto libre*</i>
¿Los datos son precisos y completos y con qué frecuencia se actualizan?	<input type="checkbox"/>	<input type="checkbox"/>	<i>*texto libre*</i>
¿Los datos han sido recopilados utilizando fuentes de medición adecuadas, confiables e imparciales?	<input type="checkbox"/>	<input type="checkbox"/>	<i>*texto libre*</i>
¿Los datos están seguros y tiene permiso para usarlos?	<input type="checkbox"/>	<input type="checkbox"/>	<i>*texto libre*</i>
¿Dónde se encuentran los datos?	<input type="checkbox"/>	<input type="checkbox"/>	<i>*texto libre*</i>
¿Tiene acceso a los datos y con qué rapidez puede acceder a ellos?	<input type="checkbox"/>	<input type="checkbox"/>	<i>*texto libre*</i>
¿En qué formato están los datos? ¿Requieren una preparación significativa para estar listos para el modelado?	<input type="checkbox"/>	<input type="checkbox"/>	<i>*texto libre*</i>
¿Los datos están estructurados? p.e. ¿se pueden almacenar en una tabla o no están estructurados, como e-mails o páginas web?	<input type="checkbox"/>	<input type="checkbox"/>	<i>*texto libre*</i>
¿Existen restricciones en los datos? p.e. ¿contienen información confidencial, como domicilios particulares?	<input type="checkbox"/>	<input type="checkbox"/>	<i>*texto libre*</i>



Preguntas	Sí	No	Comentarios
¿Puedes vincular variables clave dentro y entre conjuntos de datos?	<input type="checkbox"/>	<input type="checkbox"/>	<i>*texto libre*</i>
¿El conjunto de datos contiene patrones claros para que el modelo explore al hacer una predicción?	<input type="checkbox"/>	<input type="checkbox"/>	<i>*texto libre*</i>
¿El conjunto de datos contiene patrones claros pero accidentales, lo que resulta en sesgos de aprendizaje del modelo?	<input type="checkbox"/>	<input type="checkbox"/>	<i>*texto libre*</i>
¿Tiene suficientes datos para que el modelo aprenda?	<input type="checkbox"/>	<input type="checkbox"/>	<i>*texto libre*</i>

## e) Capacidades y competencias internas

Descripción de en qué medida se dispone actualmente en el equipo interno de los perfiles necesarios para la implantación de la IA, concretando su rol en dicha implantación; y, en caso de no disponer de ellos, identificación de las acciones de mitigación que se plantea para dar cobertura a las habilidades que se asocian a dicho rol.

Perfiles	Carácter	Disponibilidad	Rol	Justificación	Mitigación
<b>Arquitecto de datos</b> Establecer la visión para el uso de datos de la organización, a través del diseño de datos, para satisfacer las necesidades comerciales	Recomendable	Sí/No	<i>*texto libre*</i>	<i>*texto libre*</i>	<i>*texto libre*</i>
<b>Científico de datos</b> Identificar problemas comerciales complejos mientras se aprovecha el valor de los datos	Recomendable	Sí/No	<i>*texto libre*</i>	<i>*texto libre*</i>	<i>*texto libre*</i>
<b>Ingeniero de datos</b> Desarrollar la entrega de productos y servicios de datos en sistemas y procesos comerciales	Obligatorio	Sí/No	<i>*texto libre*</i>	<i>*texto libre*</i>	<i>*texto libre*</i>



Perfiles	Carácter	Disponibilidad	Rol	Justificación	Mitigación
<b>Responsable ético</b> Proporcionar juicios éticos y evaluaciones sobre las entradas (datos) del modelo de IA	Obligatorio	Sí/No	<i>*texto libre*</i>	<i>*texto libre*</i>	<i>*texto libre*</i>
<b>Técnico funcional</b> Identificar problemas comerciales complejos mientras se aprovecha el valor de los datos	Obligatorio	Sí/No	<i>*texto libre*</i>	<i>*texto libre*</i>	<i>*texto libre*</i>

## f) Integración con los sistemas de la administración pública

Identificación de las principales necesidades de integración de la solución de IA con los sistemas actuales de la administración pública, según las diferentes capas en torno a las que resulta necesario realizar dicha integración. Asimismo, identificación de las acciones necesarias para llevar a cabo dicha integración.

Capa	Necesidades de integración	Acciones de integración
<b>Capa de adquisición/ interconexión</b> Integra la información desde las fuentes de datos y suministra información a la capa de conocimiento	<i>*texto libre*</i>	<i>*texto libre*</i>
<b>Capa de conocimiento</b> Modela semánticamente los datos de la ciudad y procesa analíticamente los datos de las capas de adquisición e interoperabilidad	<i>*texto libre*</i>	<i>*texto libre*</i>
<b>Capa de interoperabilidad</b> Facilita la prestación de servicios. Para la interoperabilidad semántica, la información debe ser compatible con un modelo de datos estándar	<i>*texto libre*</i>	<i>*texto libre*</i>
<b>Capa de servicios</b> Constituida por servicios conectados a través de la capa de interoperabilidad	<i>*texto libre*</i>	<i>*texto libre*</i>



Capa	Necesidades de integración	Acciones de integración
<b>Capa de soporte</b> Capa transversal encargada de ofrecer servicios comunes como auditoría, monitorización o seguridad	<i>*texto libre*</i>	<i>*texto libre*</i>

### g) Interés de adquirir IAaaS, desarrollar o comprar soluciones

Descripción de en qué medida es idónea la selección de una alternativa (adquirir IAaaS, desarrollar o comprar soluciones) para la implantación de la solución de IA.

Alternativa	Descripción	Justificación
<b>IaaS (IA as a Service)</b> Es una alternativa de interés para aquellas situaciones en las que se desee implementar de forma ágil y flexible servicios basados en IA, sin la necesidad de integrar soluciones en los sistemas de la administración pública	<ul style="list-style-type: none"><li>• <b>Necesidad exclusiva o genérica:</b> aporta flexibilidad en ambos escenarios.</li><li>• <b>Madurez de los productos disponibles:</b> más adecuada para la implantación de productos maduros en el mercado, si bien aporta flexibilidad en ambos escenarios.</li><li>• <b>Necesidad de integración con los sistemas de administración pública:</b> más adecuada para soluciones que no requieran una integración con los sistemas de administración pública.</li></ul>	<ul style="list-style-type: none"><li>• Necesidad exclusiva o genérica: <i>*texto libre*</i></li><li>• Madurez de los productos disponibles: <i>*texto libre*</i></li><li>• Necesidad de integración con los sistemas de administración pública: <i>*texto libre*</i></li></ul>



Alternativa	Descripción	Justificación
<p><b>Compra de una solución</b></p> <p>Esta opción es más adecuada si se está buscando una aplicación común de IA (p.e. reconocimiento óptico de caracteres), si bien presenta mayores barreras en términos de integración con los sistemas de la administración pública</p>	<ul style="list-style-type: none"><li>• <b>Necesidad exclusiva o genérica:</b> resulta adecuada para responder a necesidades genéricas.</li><li>• <b>Madurez de los productos disponibles:</b> más adecuada para la implantación de productos muy maduros en el mercado.</li><li>• <b>Necesidad de integración con los sistemas de administración pública:</b> presenta mayores barreras en términos de integración con los sistemas de la administración pública.</li></ul>	<ul style="list-style-type: none"><li>• Necesidad exclusiva o genérica: <i>*texto libre*</i></li><li>• Madurez de los productos disponibles: <i>*texto libre*</i></li><li>• Necesidad de integración con los sistemas de administración pública: <i>*texto libre*</i></li></ul>
<p><b>Desarrollo de una solución</b></p> <p>Se puede optar por crear o adaptar modelos de IA listos para usar o algoritmos de código abierto, siempre y cuando el equipo interno cuente con las habilidades para construir un proyecto de IA internamente y el equipo de operaciones puede ejecutar y mantener una solución de IA interna.</p>	<ul style="list-style-type: none"><li>• <b>Necesidad exclusiva o genérica:</b> resulta adecuada para responder a necesidades específicas (no genéricas en el mercado).</li><li>• <b>Madurez de los productos disponibles:</b> no requiere la existencia de productos maduros en el mercado.</li><li>• <b>Necesidad de integración con los sistemas de administración pública:</b> la integración con los sistemas de la administración pública no supone una alta barrera, en el sentido de que el desarrollo se realizará desde un primer momento considerando dicha integración.</li></ul>	<ul style="list-style-type: none"><li>• Necesidad exclusiva o genérica: <i>*texto libre*</i></li><li>• Madurez de los productos disponibles: <i>*texto libre*</i></li><li>• Necesidad de integración con los sistemas de administración pública: <i>*texto libre*</i></li></ul>





## h) Estimación de coste total de propiedad (TCO) vinculados al sistema

Indicación de la estimación de recursos económicos necesarios para la implantación y mantenimiento del sistema de inteligencia artificial.

Fase	Tipología	Tipología	Justificación
Implantación de la solución	Costes directos de adquisición	<i>*texto libre*</i>	<i>*texto libre*</i>
	Costes indirectos de adquisición	<i>*texto libre*</i>	<i>*texto libre*</i>
	Otros eventuales costes	<i>*texto libre*</i>	<i>*texto libre*</i>
	Gastos generales de estructura	<i>*texto libre*</i>	<i>*texto libre*</i>
Mantenimiento de la solución	Costes de propiedad	<i>*texto libre*</i>	<i>*texto libre*</i>
	Costes de mantenimiento	<i>*texto libre*</i>	<i>*texto libre*</i>
	Costes de uso	<i>*texto libre*</i>	<i>*texto libre*</i>
	Costes de eliminación	<i>*texto libre*</i>	<i>*texto libre*</i>

## i) Estimación de recursos humanos internos vinculados al sistema

Indicación la estimación de recursos humanos internos (FTE) necesarios para la implantación y mantenimiento del sistema, así como el área del que deberán proceder (de la Consellería correspondiente y/o la administración pública).

Fase	Tipología	Tipología	Justificación
Implantación de la solución	<i>*texto libre*</i>	<i>*texto libre*</i>	<i>*texto libre*</i>
Mantenimiento de la solución	<i>*texto libre*</i>	<i>*texto libre*</i>	<i>*texto libre*</i>

### 3.1.4 Revisión preliminar de la evaluación



Es esencial que la evaluación preliminar de un sistema de IA cuente con la supervisión de una figura experta en IA, de forma que se garantice que, de cara a la decisión final sobre la implantación de un sistema de IA, se valoran únicamente aquellos sistemas con garantías de viabilidad.

#### Cuestiones clave

La persona delegada de la implantación de sistemas de IA deberá revisar el informe de evaluación para que este pueda ser elevado a la Comisión de Evaluación.

En el caso de identificar aspectos a modificar, la persona delegada de la implantación de sistemas de IA informará al personal funcional y el personal TI, que procederá a su incorporación en el informe hasta la validación definitiva del informe.

### 3.1.5 Valoración final de la evaluación



La Comisión de Evaluación ha de tomar la decisión final respecto a la adopción de un sistema de IA, con base en criterios conocidos vinculados a los elementos de la evaluación preliminar. Esta garantía de transparencia y objetividad es clave para asegurar la confianza en las decisiones que tome la administración pública en relación con la IA y para fomentar el consenso entre todas las partes interesadas.

#### Cuestiones clave

La Comisión de Evaluación se reunirá periódicamente para evaluar el conjunto de las oportunidades que le sean elevadas.

Resulta de interés que la Comisión decida por unanimidad o mediante mayorías reforzadas, de cara a impulsar la confianza y el consenso en torno a las decisiones que se tomen.

En el caso de una decisión desfavorable, se considera acreditado que el proyecto de potencial implantación del sistema de IA debe ser descartado de forma definitiva.

Las decisiones que se adopten por la Comisión de Evaluación deben reflejar la priorización e incluso la calendarización de la implantación de los sistemas de IA que hayan recibido la aprobación, de forma coherente con los recursos de los que dispone la administración pública para realizar dichas implantaciones.

## 3.2

### Fase 2. Diseño, desarrollo e implantación del sistema de IA adopción de la IA

El proceso de diseño, desarrollo e implantación de una solución de IA se centra en abordar las oportunidades identificadas, utilizando datos y algoritmos para construir sistemas inteligentes que mejoren la toma de decisiones y la eficiencia en diversas aplicaciones. La colaboración de equipos multidisciplinares y una mentalidad de mejora continua son clave para el éxito en la adopción de estos sistemas.

#### 3.2.1 Ideación del sistema de IA



La ideación de un sistema de Inteligencia Artificial (IA) se refiere al proceso de generación de ideas y conceptos para diseñar un sistema basado en IA que aborde un problema o una necesidad específica. Es una etapa fundamental en el desarrollo de soluciones de IA y suele preceder a la fase de diseño y desarrollo del sistema de IA. La ideación es un proceso creativo y flexible en el que se fomenta la exploración de múltiples enfoques y soluciones. Puede requerir iteraciones y ajustes a medida que se obtiene más información y retroalimentación.

##### Cuestiones clave

La ideación implica la definición clara y objetiva de la necesidad o problema que se desea abordar con el sistema de IA.

La ideación también implica la concreción los objetivos específicos que se desean alcanzar y cómo la IA ayudará a resolver dichas necesidades o problemas.

Una vez ideado el sistema de IA, esta actividad puede incluir el prototipado o experimentación inicial del sistema, es decir, crear un prototipo o una versión de prueba del sistema de IA para evaluar su funcionamiento en condiciones controladas.

##### Errores comunes

- **Sesgo en la generación de ideas:** la falta de diversidad en el equipo de ideación o la influencia de sesgos personales pueden limitar la variedad de soluciones propuestas.
- **Enfoque excesivo en la tecnología:** centrarse demasiado en la tecnología en lugar de considerar las necesidades del usuario o las aplicaciones prácticas puede llevar a soluciones de IA que no se ajustan a las situaciones del mundo real.
- **Falta de consideración ética y de privacidad:** no tener en cuenta las implicaciones éticas y de privacidad desde el principio puede resultar en problemas legales y de confianza en etapas posteriores de la implantación.
- **No explorar suficientes enfoques de IA:** limitarse a un único enfoque de IA o no considerar una variedad de técnicas puede resultar en soluciones subóptimas.
- **Falta de alineación con los objetivos de la organización:** no asegurarse de que las ideas de IA estén alineadas con los objetivos estratégicos de la empresa o la organización puede llevar a soluciones que no agreguen valor.
- **Ignorar la retroalimentación de las partes interesadas:** no recopilar retroalimentación temprana de personas relevantes puede llevar a la falta de ajustes y mejoras en las ideas iniciales

### 3.2.2 Diseño del sistema de IA



El diseño de un sistema de Inteligencia Artificial (IA) implica la definición de la arquitectura y la funcionalidad del sistema. Comprende la selección de algoritmos de IA, la estructura de datos, la interfaz de usuario y la infraestructura tecnológica. Se abordan cuestiones como la escalabilidad, la robustez y la eficiencia. Además, se consideran aspectos éticos y de privacidad. El diseño se basa en los resultados de la etapa de ideación, donde se generan ideas para resolver un problema o aprovechar una oportunidad. Es un paso crucial que sienta las bases para el desarrollo e implementación efectivos del sistema de IA.

#### Cuestiones clave

Esta actividad consiste en la definición de las características y especificaciones técnicas del sistemas de IA, siendo las principales:

- **Tipología de datos:** se refiere a la naturaleza y el tipo de datos que el sistema de IA utilizará. Esto incluye datos estructurados, no estructurados y semiestructurados.
- **Algoritmos y modelos:** los algoritmos y modelos determinan cómo se procesarán los datos para hacer predicciones, tomar decisiones o realizar tareas específicas. La elección de algoritmos y modelos depende del tipo de problema que se esté abordando.
- **Diseño de la arquitectura:** el diseño de la arquitectura incluye la organización de componentes, flujos de datos, procesamiento en lotes o en tiempo real, y la selección de hardware y software para el funcionamiento del sistema.
- **Interfaz de usuario:** la interfaz de usuario es la parte que permite a los usuarios interactuar con el sistema de IA. Esto puede incluir interfaces gráficas de usuario (GUI), interfaces de línea de comandos (CLI) o integración en aplicaciones y servicios existentes.
- **Ética y responsabilidad:** esto implica garantizar que el sistema no perpetúe sesgos, que cumpla con las regulaciones y estándares éticos, y que se respete la privacidad de los datos y la transparencia en las decisiones



de la IA. La evaluación y mejora continua son procesos esenciales en el diseño de un sistema de IA. Implica la realización de pruebas, la recopilación de retroalimentación de usuarios y la supervisión del rendimiento del sistema.

### Errores comunes

- **Sesgo en los datos:** utilizar conjuntos de datos sesgados o no representativos puede dar lugar a sistemas de IA que perpetúan prejuicios y discriminación.
- **Selección inapropiada de algoritmos:** elegir algoritmos de IA inadecuados para el problema en cuestión puede resultar en un bajo rendimiento o ineficiencia.
- **No realizar evaluaciones éticas y pruebas de sesgo:** no evaluar y mitigar sesgos en el sistema puede resultar en decisiones discriminatorias y problemas de equidad.
- **Diseño de arquitectura poco escalable:** no planificar la arquitectura para adaptarse a volúmenes crecientes de datos o demanda puede limitar la capacidad del sistema.
- **Interfaz de usuario deficiente:** no diseñar una interfaz de usuario efectiva puede dificultar la adopción y el uso del sistema.
- **Falta de transparencia y explicabilidad:** no proporcionar información sobre cómo el sistema toma decisiones puede dar lugar a desconfianza y falta de comprensión.

### 3.2.3 Desarrollo y entrenamiento del sistema de IA



El desarrollo y entrenamiento de un sistema de IA es una etapa crítica en la implantación de soluciones basadas en IA, etapa que comienza con la implementación del modelo de IA diseñado en la fase de diseño. Este modelo sirve como el "cerebro" del sistema, y su tarea es realizar predicciones, tomar decisiones o ejecutar tareas específicas.

Para entrenar el modelo, se requieren datos adecuados. Estos datos pueden ser etiquetados, lo que significa que se conocen las respuestas correctas, o no etiquetados, en cuyo caso el modelo aprende patrones sin la guía explícita de respuestas. Durante el entrenamiento, el modelo se expone repetidamente a ejemplos de datos, ajustando sus parámetros internos para optimizar su capacidad para realizar la tarea específica. Este proceso es esencialmente un proceso de aprendizaje, donde el modelo mejora su capacidad para hacer predicciones precisas.

#### Cuestiones clave

El desarrollo y entrenamiento del sistema de IA consiste en alimentar el modelo del sistema de IA con un conjunto de datos de entrenamiento que permitan al sistema aprender de forma iterativa, de manera que se ajusten los parámetros del modelo para minimizar los errores y maximizar el rendimiento del sistema tras cada iteración.

Es necesario asegurar que el conjunto de datos de entrenamiento sea representativo y abarque los diferentes escenarios que se desea que el sistema pueda reconocer, además de dividir los datos en conjuntos de entrenamiento, validación y prueba para evaluar el rendimiento del sistema de IA.

El desarrollo y entrenamiento de un sistema de IA es una combinación de ciencia y tecnología que requiere experiencia técnica y una comprensión profunda de los algoritmos, los datos y el dominio específico del problema a abordar. La precisión y la eficacia del modelo dependen en gran medida de la calidad y cantidad de los datos de entrenamiento, así como de la capacidad para ajustar y validar adecuadamente el modelo.

#### Errores comunes

- **Selección inadecuada de datos:** utilizar conjuntos de datos de baja calidad o no representativos puede llevar a un modelo de IA poco preciso y sesgado.
- **Sobreajuste del modelo:** ajustar el modelo excesivamente a los datos de entrenamiento puede dar como resultado un modelo que funciona bien en esos datos, pero no generaliza bien a nuevas situaciones.
- **Subestimación de la cantidad de datos:** no contar con suficientes datos de entrenamiento puede limitar la capacidad del modelo para aprender patrones con precisión.
- **No considerar sesgos:** no abordar los sesgos en los datos de entrenamiento puede resultar en modelos de IA que perpetúan prejuicios y discriminación.
- **Selección inadecuada de algoritmos:** elegir el algoritmo de IA incorrecto para el problema en cuestión puede dar lugar a un rendimiento subóptimo.
- **Falta de optimización de hiperparámetros:** no ajustar adecuadamente los hiperparámetros del modelo puede impedir que alcance su máximo potencial.
- **No documentar el proceso:** la falta de documentación adecuada del proceso de desarrollo y entrenamiento puede dificultar la comprensión y la colaboración entre equipos.

### 3.2.4 Validación de los resultados del sistema de IA



Esta etapa tiene como objetivo determinar si el modelo de IA entrenado funciona de manera precisa y generaliza bien a nuevas situaciones. Una vez que el modelo se entrena, se somete a una fase de validación para evaluar su rendimiento. Se utilizan conjuntos de datos adicionales que no se usaron durante el entrenamiento para verificar si el modelo es capaz de adaptar su funcionamiento a nuevas situaciones. Si el rendimiento es insatisfactorio, se ajustan los parámetros y se vuelve a validar. Esto puede implicar ajustar configuraciones clave que afectan el comportamiento del modelo, como la velocidad de aprendizaje o la complejidad del modelo.

#### Cuestiones clave

El modelo de IA se evalúa utilizando el conjunto de prueba o validación. Se comparan las predicciones del modelo con las respuestas reales y se calculan métricas de rendimiento, como la precisión, la recuperación, la precisión, el F1-score, entre otros, según el tipo de problema que aborde la IA.

La validación se realiza en múltiples aspectos, incluyendo:

**Precisión:** ¿con qué precisión el modelo realiza predicciones correctas?

- **Recuperación:** ¿cuántos de los casos positivos reales son identificados por el modelo?
- **Robustez:** ¿cómo se comporta el modelo frente a datos ruidosos, atípicos o cambiantes?
- **Sesgo:** ¿existe sesgo en las predicciones hacia ciertos grupos o clases?
- **Interpretabilidad:** ¿se pueden entender y explicar las decisiones del modelo?

La validación puede requerir ajustes en el modelo, como la optimización de hiperparámetros o la mitigación de sesgos. Es un proceso iterativo que busca mejorar continuamente el rendimiento y la capacidad del sistema de IA para tomar decisiones precisas y éticas. Una validación exhaustiva es esencial para garantizar que el sistema de IA funcione en el mundo real y se ajuste a las necesidades de los usuarios conforme criterios de ética y responsabilidad.

#### Errores comunes

- **No dividir adecuadamente los datos:** no separar de manera adecuada los conjuntos de entrenamiento, validación y prueba puede resultar en una evaluación inadecuada del modelo y en una sobreestimación de su rendimiento.
- **No utilizar métricas adecuadas:** seleccionar métricas de rendimiento incorrectas o inapropiadas para el problema en cuestión puede llevar a una evaluación inexacta del sistema de IA.
- **No tener un conjunto de validación independiente:** la falta de un conjunto de validación independiente para ajustar hiperparámetros y evaluar el rendimiento del modelo puede dar como resultado modelos mal calibrados.
- **No abordar el desequilibrio de clases:** en problemas de clasificación desequilibrados, no tratar el desequilibrio de clases puede llevar a un modelo que se sesga hacia la clase dominante.
- **No considerar la robustez:** no evaluar la robustez del modelo frente a datos ruidosos, atípicos o adversariales puede resultar en un sistema de IA vulnerable a ataques o situaciones inesperadas.
- **Ignorar sesgos:** no evaluar y mitigar el sesgo en los datos y las predicciones puede resultar en decisiones discriminatorias y problemas de equidad.
- **No documentar el proceso:** la falta de documentación adecuada del proceso de validación puede dificultar la comprensión y la revisión del rendimiento del modelo.
- **Falta de retroalimentación del usuario:** no buscar la retroalimentación de los usuarios finales puede dar lugar a la implementación de un modelo que no se ajusta a sus necesidades.
- **No realizar validación continua:** no llevar a cabo evaluaciones periódicas del modelo después de su implementación en producción puede llevar a un sistema de IA desactualizado.

### 3.2.5 Despliegue y mantenimiento del sistema de IA



Una vez que el modelo ha sido entrenado y evaluado satisfactoriamente, se procede al despliegue del sistema de IA en un entorno de producción, planificando la implementación y pruebas antes de su lanzamiento. Igualmente, se puede proceder a la implementación de los mecanismos para evaluar el rendimiento del sistema de IA y los procesos de mejora continua de los resultados alcanzados.

El despliegue y mantenimiento son procesos continuos que requieren una colaboración interdisciplinaria y una gestión efectiva para asegurar que el sistema de IA opere de manera confiable y se mantenga alineado con sus objetivos y necesidades cambiantes a lo largo del tiempo.

La implementación implica integrar el modelo en la infraestructura existente, lo que puede requerir el desarrollo de interfaces de usuario y servicios. Posteriormente, se realizan pruebas exhaustivas en un entorno de producción para verificar su funcionamiento correcto y garantizar su eficiencia. La monitorización constante de este proceso es esencial. Se establecen sistemas de monitorización para rastrear el rendimiento en tiempo real, lo que incluye el seguimiento de métricas clave y la detección de posibles problemas de rendimiento o decisiones inesperadas del modelo. La actualización y mejora continua son también parte del proceso. Los modelos de IA necesitan actualizarse para mantener su precisión y relevancia. Esto implica reentrenar el modelo con nuevos datos y ajustar parámetros para optimizar su rendimiento.

La seguridad y la privacidad son consideraciones críticas del despliegue y mantenimiento del sistema de IA, por lo que es necesario implementar medidas para proteger los datos y el modelo de ciberseguridad y garantizar el cumplimiento de las regulaciones de privacidad. La evaluación ética del sistema también resulta fundamental, lo que conlleva realizar una evaluación continua para detectar y mitigar sesgos y decisiones discriminatorias del modelo, asegurando la equidad y la imparcialidad.

#### Cuestiones clave

En esta actividad, el modelo entrenado se implementa en un entorno de producción y se garantiza su funcionamiento continuo y eficiente. Comprende varias etapas clave:

- **Implementación:** el modelo de IA se integra en la infraestructura tecnológica existente, lo que puede requerir el desarrollo de aplicaciones, interfaces de usuario o servicios que interactúen con el modelo.
- **Pruebas en entorno de producción:** se realizan pruebas exhaustivas en un entorno de producción real para garantizar que el modelo funcione correctamente y que no haya problemas de rendimiento ni errores inesperados.

- **Monitorización y supervisión:** se establecen sistemas para rastrear el rendimiento del modelo en tiempo real. Esto incluye el seguimiento de métricas de rendimiento, la detección de anomalías y la identificación de problemas.
- **Actualización y mejora continua:** el modelo se actualiza periódicamente para adaptarse a cambios en los datos, las necesidades del usuario o la evolución de la tecnología. Las actualizaciones pueden incluir el reentrenamiento del modelo con nuevos datos.

La implementación de medidas de seguridad y privacidad, la evaluación ética del sistema y el soporte a la resolución de problemas y consultas a personas usuarias son también elementos clave del despliegue y mantenimiento del sistema de IA.

### Errores comunes

- **Falta de pruebas exhaustivas:** no realizar pruebas en profundidad en un entorno de producción puede llevar a problemas no detectados y errores críticos en el sistema.
- **No establecer una monitorización adecuada:** la falta de sistemas de monitoreo en tiempo real puede retrasar la detección de problemas y anomalías, lo que afecta la confiabilidad del sistema.
- **Actualización infrecuente o inexistente:** no actualizar el modelo de IA o los algoritmos puede llevar a la obsolescencia y la pérdida de precisión con el tiempo.
- **No abordar la seguridad:** ignorar la seguridad de datos y modelos puede exponer el sistema a amenazas cibernéticas, robo de datos o manipulación.
- **Falta de consideración ética y de privacidad:** no abordar cuestiones éticas o de privacidad en la gestión de datos y decisiones del modelo puede resultar en problemas legales y de confianza.
- **No abordar problemas de sesgo y equidad:** la falta de evaluación y mitigación de sesgos puede resultar en decisiones discriminatorias y problemas de equidad.
- **Ausencia de actualizaciones de políticas:** no adaptar las políticas y regulaciones de acuerdo con los cambios en la tecnología de IA puede llevar a incumplimientos normativos.
- **Falta de soporte y capacitación para usuarios:** no proporcionar soporte adecuado para los usuarios y no capacitarlos puede llevar a una mala adopción del sistema.
- **No documentar cambios y decisiones:** la falta de documentación adecuada de cambios y decisiones en el sistema dificulta la transparencia y la auditoría.

## Capítulo 4

### Gestión de riesgos de un sistema de IA

Este capítulo se adentra en la gestión de riesgos específicos para sistemas de inteligencia artificial (IA). Se abordan procesos clave, desde la identificación y evaluación de riesgos hasta la implementación de estrategias de tratamiento. Se destaca la importancia de considerar riesgos operativos y aquellos que impactan la salud, seguridad y derechos fundamentales. Se explora la determinación del "apetito al riesgo", se detalla el contexto del sistema y se subraya la evaluación de riesgos basada en la probabilidad e impacto.





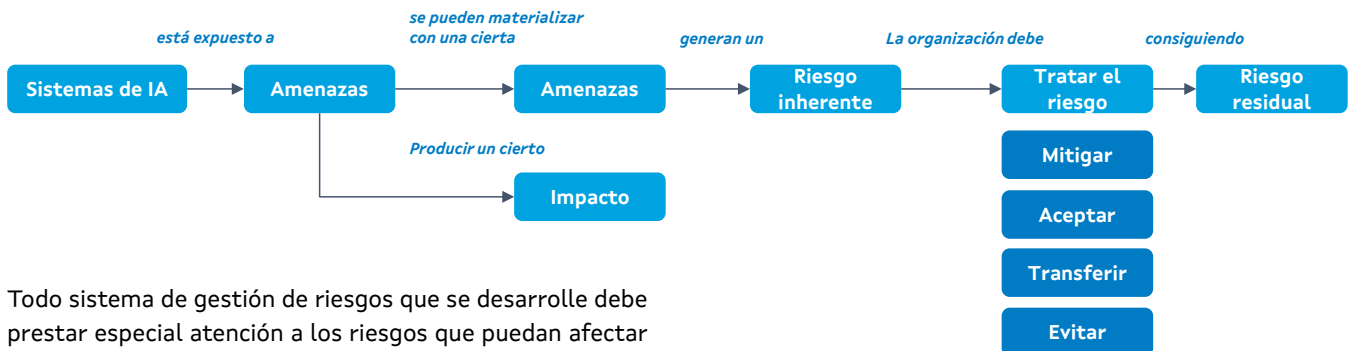
## 4. Gestión de riesgos de un sistema de IA

Un sistema de gestión de riesgos es un sistema de gestión cuyo objetivo es el análisis y el tratamiento de los riesgos a los que una organización puede estar expuesta. El análisis de riesgos permite a la organización evaluar sus activos y anticipar posibles escenarios, mientras que el tratamiento de riesgos se enfoca en desarrollar estrategias para reducir esos riesgos de manera efectiva.

### 4.1 Introducción a la gestión de riesgos

Un sistema de IA puede estar expuesto a diferentes amenazas que podrían terminar suponiendo un riesgo y causar daño al sistema. Estas amenazas pueden materializarse con una cierta probabilidad y producir un cierto impacto, dando origen a un riesgo inherente. La organización debe abordar los riesgos identificados y trabajar para reducirlos a un nivel residual aceptable que permanecerá en el sistema. Para lograrlo, existen estrategias como mitigar, aceptar, evitar o transferir dichos riesgos.

#### Ecosistema de la gestión de riesgos



Todo sistema de gestión de riesgos que se desarrolle debe prestar especial atención a los riesgos que puedan afectar a la salud, la seguridad y los derechos fundamentales de las personas, de forma que se deberá poner el foco en la identificación y análisis de cualquier riesgo que pueda tener especial afectación en dichos elementos, así como en la implementación de las medidas que permitan dar cobertura a los mismos.

## 4.2 Modelo de gestión de riesgos de un sistema de IA

Un sistema de gestión de riesgos se fundamenta en la ejecución de una serie de procedimientos destinados a identificar y evaluar diversos riesgos, además de definir e implementar medidas adecuadas para gestionar el impacto de dichos riesgos. A continuación, se presentan los procesos que conforman un modelo de gestión de riesgos, que ha de entenderse como un proceso transversal que debe ser actualizado a medida que se identifican nuevos riesgos en cada una de las etapas del ciclo de vida del sistema de IA, desde el diseño y desarrollo hasta su implantación y producción.

Modelo de gestión de riesgos de un sistema de inteligencia artificial



### 4.2.1 Definición del apetito al riesgo

La definición del apetito al riesgo es la definición del nivel de riesgo que una organización está dispuesta a asumir. Se define de manera cuantitativa a partir de la configuración de una escala de riesgo y el objetivo es establecer un umbral en el nivel de riesgo que la organización considere oportuno. En esta decisión, se debe tener en consideración los riesgos que puedan afectar el funcionamiento del propio sistema y al tratarse de un sistema de IA y estar regulado por la normativa comunitaria, también aquellos riesgos que el sistema pueda suponer para la salud, seguridad y derechos fundamentales de las personas. No se trata de cuantificar esta escala con precisión numérica, sino valorarlo en una escala conceptual general para toda la organización, y ser coherente en su aplicación a todos los casos. En general, esta escala se configura definiendo un valor mínimo (generalmente igual a 1) y un valor máximo.

El valor o nivel de riesgo se calcula como el producto del impacto por la probabilidad de que suceda. Por ello, es importante que las escalas que cuantifican el impacto y la probabilidad de que materialice una amenaza estén definidas de manera coherente con la primera.



#### 4.2.2. Contexto del sistema de IA

El contexto de un sistema se refiere al conjunto de elementos que influyen en su funcionamiento, tanto desde el entorno externo como el interno. La finalidad de este proceso es recopilar un inventario de todos los elementos del contexto que pueden impactar en el análisis de riesgos. Este proceso suele ser mayoritariamente cualitativo y subjetivo, y la precisión de la recopilación de estos elementos del entorno es esencial para su completa comprensión. A continuación, se detalla el ejemplo de algunos de los elementos más relevantes que completan este entorno:



Los factores sociales, culturales, políticos, legales, regulatorios, financieros, tecnológicos, económicos y ambientales.



Las principales tendencias tecnológicas, avances en áreas de la IA y las implicaciones sociales y políticas del despliegue de estas tecnologías que pueden afectar a nuestro sistema.



Las principales partes interesadas, sus percepciones, valores, necesidades y expectativas.



La complejidad de las redes y sus dependencias, que pueden aumentar con el uso de tecnologías de IA.



Los factores internos organizacionales que giran en torno a la visión, la misión, los valores, la cultura, la estrategia, el modelo de gobierno, las políticas, normas y procedimientos adoptadas y las relaciones contractuales y compromisos.



La cultura de la organización, así como guías, modelos y normas adoptadas.



Las capacidades de la organización, recursos y conocimientos (capital, tiempo, personas, procesos, tecnologías, etc.).



El uso de datos y flujos de información. Los sistemas de IA pueden utilizarse para automatizar, optimizar y mejorar el tratamiento de los datos.



### 4.2.3. Identificación del riesgo

La identificación del riesgo implica el análisis e identificación de todos los posibles riesgos que pueden afectar al sistema de IA. En este proceso se debe tener en consideración aquellos riesgos que puedan poner en peligro el funcionamiento del sistema, la seguridad de las personas que lo utilizan, así como los derechos fundamentales de las personas.

En esta fase se evalúan y documentan los riesgos con el fin de tomar medidas para prevenir, mitigar o gestionarlos de una manera efectiva y garantizar un funcionamiento seguro y confiable del sistema. Para ello, en primer lugar, se identifican los componentes del sistema y, en segundo lugar, a partir de las fuentes de riesgo más comunes, se analizan e identifican los riesgos asociados a los componentes identificados.

#### a) Identificación de componentes de un sistema de IA

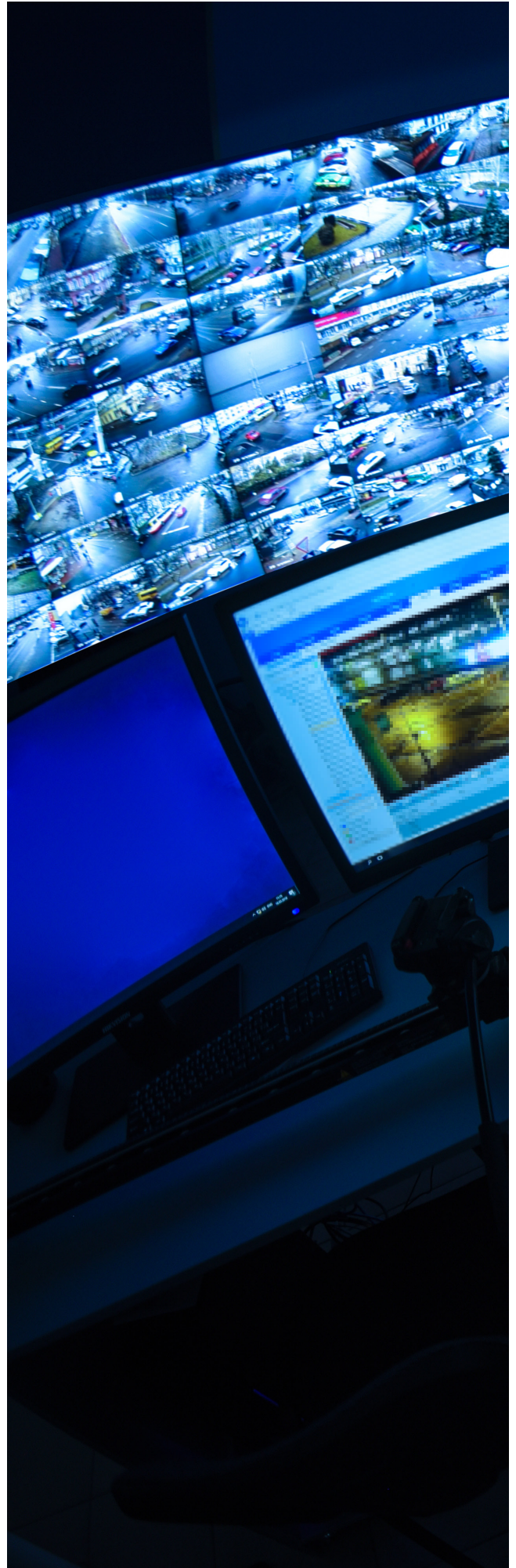
En esta fase de identificación de componentes, se deben incluir todos los elementos que componen un sistema de IA en todas las etapas de su ciclo de vida tales como actores principales o stakeholders, datos, modelos, procesos, entorno y herramientas. A continuación, se detallan ejemplos de algunos de los componentes de un sistema de IA:



<b>Propietario de datos</b>	Persona responsable en la organización del dato, encargada de su definición, clasificación, protección, uso y calidad
<b>Científico de datos</b>	Profesionales que aplican la estadística, el aprendizaje automático y los enfoques analíticos para analizar diferentes conjuntos de datos de distintos tamaños y formas y resolver problemas complejos y críticos
<b>Datos de entrenamiento</b>	Datos iniciales que se utilizan para desarrollar un sistema de IA, a partir de los cuales el sistema adapta sus parámetros internos para refinar sus reglas
<b>Datos de prueba</b>	Conjunto de datos utilizado para proporcionar una evaluación no sesgada de un sistema de IA ajustado al conjunto de datos de entrenamiento. Los datos de prueba se utilizan para probar el sistema
<b>Parámetros del sistema de IA</b>	Variables de configuración internas del sistema de IA y cuyo valor puede estimarse a partir de los datos dados
<b>Preprocesamiento de datos</b>	Comprensión, preparación y limpieza de los datos
<b>Ingesta de datos</b>	Proceso relacionado con el transporte de datos desde múltiples fuentes para componer puntos de datos multidimensionales; los datos pueden colocarse en un medio de almacenamiento al que se pueda acceder, utilizarlos y analizarlos, o bien, el flujo de datos puede utilizarse directamente por el sistema de IA
<b>Etiquetado de datos</b>	Proceso de detección y etiquetado de las muestras de datos; el proceso puede ser manual y lento y estar asistido por software

## b) Fuentes comunes en el proceso de identificación de riesgos de la IA

El proceso de identificación de riesgos se basa en el análisis e identificación de riesgos relacionados con un componente específico del sistema en un elemento particular del contexto. Un mismo componente, cuando se analiza en un entorno de contexto diferente, puede llevar a la identificación de riesgos distintos. A continuación, se listan algunas de las fuentes más comunes, que pueden servir como punto de partida para este análisis e identificación:



### Falta de transparencia

La transparencia es la capacidad de entender y explicar cómo un sistema de IA toma sus decisiones o realiza sus acciones. Cuando un sistema de IA no es transparente, resulta difícil rastrear y comprender por qué tomó una determinada decisión. Esto puede generar preocupaciones éticas, legales y prácticas, ya que sus decisiones pueden ser difíciles de justificar o pueden llevar a resultados injustos. Para mitigar este riesgo, es importante trabajar en hacer que los sistemas de IA sean más transparentes, lo que significa que sus procesos de toma de decisiones deben ser comprensibles y explicables para las personas que los utilizan y confían en ellos.

---

### Falta de explicabilidad

La explicabilidad es la capacidad de comprender y explicar por qué un sistema de IA elige ciertas acciones o decisiones en lugar de otras. Cuando la explicabilidad es insuficiente, resulta difícil entender los factores que influenciaron las decisiones de la IA o su razonamiento subyacente. Si no se pueden explicar estos factores, la validación del sistema de IA y la confianza en el sistema se ven afectados negativamente, ya que no está claro por qué el sistema ha tomado una decisión y si tomará la decisión correcta en todos los casos.

---

### Vigilancia humana y automatización

El nivel de automatización es el nivel de responsabilidad y control se le concede a un sistema de IA para tomar decisiones y realizar tareas sin intervención humana. Si el nivel de automatización es alto y el sistema de IA tiene un papel significativo en la toma de decisiones o en la ejecución de tareas críticas, existe un riesgo potencial. Esto se debe a que la IA puede cometer errores o tomar decisiones subóptimas en ciertas situaciones, especialmente si no se entiende completamente su funcionamiento. Es importante equilibrar el nivel de automatización para garantizar que haya supervisión y control humano adecuados.

---

### Fuentes de riesgo relacionadas con el aprendizaje automático

El comportamiento de los sistemas de IA depende, no solo de los algoritmos en uso, sino también de los datos con los que se entrenan los sistemas. Existen diversos riesgos derivados del uso de los datos. Por ejemplo:

- La calidad inadecuada de los datos podría afectar a varios objetivos como la equidad, la seguridad y la solidez.
- Los datos pueden dejar de ser representativos del dominio de aplicación, lo que conlleva riesgos para los objetivos del negocio.
- El recabado y almacenamiento de datos puede incurrir en riesgos éticos y legales significativos.

### Problemas de hardware del sistema

Las fuentes de riesgo relacionadas con problemas de hardware incluyen, por ejemplo, errores basados en componentes defectuosos. El desarrollo de sistemas de IA podría verse limitado debido a las diferentes capacidades de hardware de los sistemas en términos de potencia de procesamiento, memoria y la disponibilidad de aceleradores de hardware de IA dedicados.

---

### Problemas del ciclo de vida del sistema

Los métodos, procesos y también el uso inadecuado o insuficiente de un sistema de IA a lo largo de su ciclo de vida pueden generar riesgos. Por ejemplo, un proceso de diseño defectuoso puede no anticipar los contextos en los que se utilizará el sistema de IA, lo que hace que falle inesperadamente cuando se usa en estos contextos.

---

### Preparación tecnológica

La preparación tecnológica indica cuán madura es una tecnología dada en un contexto de aplicación determinado. Las tecnologías menos maduras utilizadas en el desarrollo y la aplicación de sistemas de IA pueden añadir riesgos desconocidos para la organización o difíciles de evaluar. Para las tecnologías maduras, se puede disponer de una mayor variedad de datos de experiencia, lo que facilita la identificación y la evaluación de los riesgos.

---

### Complejidad del entorno

La complejidad del entorno de un sistema de IA determina la gama de situaciones que un sistema de IA puede soportar en su contexto operativo. Uno de los riesgos más relevantes, por ejemplo, es el relacionado con el grado de comprensión del entorno del sistema de IA. Una comprensión parcial del entorno dará lugar a un nivel de incertidumbre que es una fuente de riesgo especialmente relevante en la fase del diseño de los sistemas de IA.

---

### Otras potenciales fuentes de riesgos

- Dificultad en la identificación de responsabilidades y rendición de cuentas.
- Uso inadecuado, incorrecto o fraudulento del sistema de IA.
- Amenazas de seguridad y ciberseguridad.
- Posibles amenazas a la privacidad de las personas si no se desarrolla un adecuado gobierno de los datos.
- Posibles perturbaciones o manipulaciones en los datos no deseadas.
- Los posibles sesgos en los datos también pueden suponer una amenaza para el adecuado uso del sistema de IA.

#### 4.2.4. Análisis y evaluación de riesgos

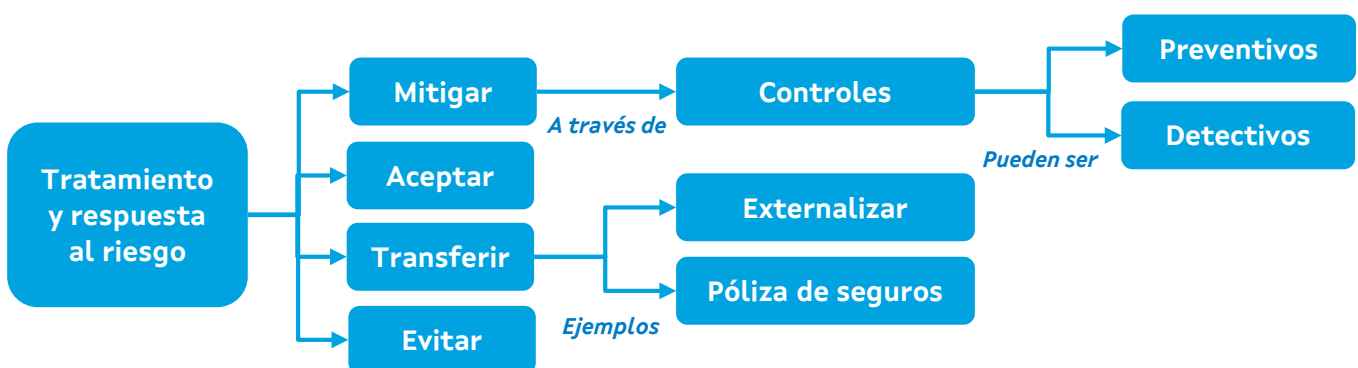
El proceso de análisis y evaluación de riesgos se centra en calcular, para cada uno de los riesgos identificados en el proceso anterior, su nivel de riesgo. Este valor de riesgo se obtiene a partir del producto de la probabilidad de ocurrencia del riesgo y la magnitud de su impacto. Ambos factores deben ser evaluados utilizando escalas predefinidas que midan tanto el impacto como la probabilidad.

Las escalas utilizadas para evaluar estos factores deben ser coherentes con la escala previamente definida para el apetito al riesgo, ya que el siguiente paso implica la comparación de estos valores. Una vez se han establecido los niveles de impacto y probabilidad y se ha calculado el valor del riesgo, se procede a compararlo con el umbral de apetito al riesgo preestablecido. Si el nivel de riesgo obtenido supera el umbral, es necesario determinar qué enfoque de respuesta es el más adecuado para la situación: evitar, transferir, mitigar o aceptar el riesgo.

#### 4.2.5. Respuesta al riesgo

Las respuestas al riesgo son estrategias que una organización puede adoptar para abordar y gestionar los riesgos en un contexto específico. La organización debe determinar qué tipo de medidas de tratamiento de riesgos se seleccionarán para abordar cada uno de los riesgos identificados y analizados. Las opciones principales son las que se indican a continuación:

*Visión global de las alternativas para la respuesta al riesgo*



### Mitigar

La mitigación del riesgo implica tomar medidas proactivas de carácter preventivas o de detección para reducir la probabilidad de que ocurra un riesgo o minimizar su impacto si ocurre. Esto puede incluir la implementación de controles, la adopción de mejores prácticas, la capacitación del personal o la optimización de procesos para reducir la exposición al riesgo hasta alcanzar el umbral de apetito definido.

---

### Aceptar

Aceptar un riesgo significa que la organización reconoce la existencia del riesgo, pero decide no tomar medidas activas al respecto. Esta respuesta puede ser adecuada cuando el riesgo es considerado tolerable o cuando el coste de mitigar el riesgo supera los posibles daños que podría causar.

---

### Trasferir

La transferencia de riesgo implica delegar la responsabilidad de gestionar el riesgo a otra entidad, generalmente a través de acuerdos contractuales o de seguros. Por ejemplo, una organización puede transferir el riesgo de pérdida por daños a la propiedad a una compañía de seguros mediante la compra de una póliza de seguro.

---

### Evitar

Evitar un riesgo implica tomar medidas para eliminar completamente la exposición al riesgo. Esto puede lograrse evitando actividades, procesos o decisiones que puedan dar lugar al riesgo. Es una estrategia radical que se utiliza cuando el riesgo es inaceptable o no puede ser gestionado de manera efectiva de ninguna otra manera.

Una vez definidas las medidas de tratamiento del riesgo se deben definir planes de acción, planificar su implementación y designar a un responsable para llevar a cabo su ejecución. Estratégicamente, la implementación de aquellos tratamientos que mitiguen los riesgos más relevantes debe ser priorizada.



#### **4.2.6. Documentación**

Se debe mantener un registro de los riesgos identificados y analizados en la documentación asociada al sistema, así como establecer periodos de revisión y seguimiento del modelo de gestión de riesgos (se deben incorporar todos aquellos nuevos riesgos que se identifiquen, así como evaluar y definir las medidas de tratamiento oportunas).

#### **4.2.7. Comunicación y consulta**

La comunicación y consulta tienen como objetivo generar consciencia y comprensión de los riesgos. Es fundamental mantener actualizadas a todas las partes interesadas apropiadas, tanto internas como externas, en todas las fases del proceso de gestión de riesgos.

Todos los participantes en el diseño, desarrollo e implementación del sistema de IA deben tener conocimiento del sistema de gestión de riesgos y deben participar activamente en su implementación y en las actualizaciones correspondientes, especialmente las áreas de la organización que estén directamente involucradas en el proceso.

#### **4.2.8. Seguimiento y mejora continua**

La finalidad del seguimiento y mejora continua es garantizar la seguridad y calidad del sistema de IA, manteniéndolo en constante evolución y logrando que el sistema se adapte a los cambios en el entorno. Esta fase implica la actualización de los procesos previamente descritos acorde a la periodicidad establecida o ante cambios significativos en la organización o en el sistema: desde la actualización del contexto del sistema a la revisión y actualización de los planes de acción para el tratamiento de los riesgos.

Este seguimiento y actualización debe llevarse a cabo en todas las etapas del ciclo de vida de un sistema de IA. El contexto, los componentes del sistema y riesgos que se pueden identificar varían según la etapa de desarrollo en el que se encuentre el sistema, por lo que los procesos descritos en este capítulo serán repetidos a medida que el sistema se diseñe, desarrolle e implante.

#### 4.2.9. Liderazgo y compromiso

En la gestión de riesgos, el liderazgo y el compromiso son componentes transversales que deben estar presentes de manera constante y coherente en todas las actividades relacionadas con la gestión de riesgos en la organización. Los líderes deben desempeñar un papel activo en todas las etapas, desde la identificación y evaluación de riesgos hasta la implementación de medidas de control y la revisión periódica del sistema.

Este enfoque transversal asegura que la gestión de riesgos no sea simplemente un proceso aislado, sino una parte integral de la cultura y la operación diaria de la empresa, lo que contribuye a una toma de decisiones más efectiva y a la protección sostenible de los activos y el cumplimiento de los objetivos organizativos.



### **4.3. Ejemplo de gestión de riesgos en el diseño, desarrollo e implantación de un sistema de IA**

A lo largo de esta sección, se aterrizará el proceso completo de gestión de riesgos de un sistema de IA mediante el siguiente ejemplo:

Se decide diseñar, desarrollar e implantar un sistema de distribución de subvenciones a personas físicas y empresas basado en IA para ayudar a las administraciones públicas en la toma de decisiones, con el objetivo de mejorar su eficiencia, ofrecer una mayor transparencia en el proceso de asignación y proporcionar una respuesta más ágil a situaciones de emergencia económica.

### 4.3.1. Ideación del sistema de IA

#### Paso 0. Determinación del apetito al riesgo

El apetito al riesgo debe definirse de forma cuantitativa. Para este ejemplo, se decide establecer una escala del apetito al riesgo del 1 al 15 (siendo el 15 el valor máximo de riesgo que la organización está dispuesta a aceptar).

Esta escala debe alinearse con otras dos escalas: una que va del 1 al 5 para clasificar los riesgos en cinco niveles de impacto (desde muy bajo o inexistente hasta muy alto o crítico, en orden ascendente) y otra que abarca del 1 al 3 para evaluar la probabilidad de que estos riesgos se materialicen (desde improbable hasta muy probable, en orden ascendente). Para el ejemplo que se presenta, se determina el apetito al riesgo en un nivel intermedio, un valor 4 sobre 15.

#### Paso 1. Análisis de riesgos en la ideación del sistema de IA

Contexto						
Cumplimiento con el derecho fundamental de la protección de datos de carácter personal						
Identificación de componentes	Identificación de riesgos					
<b>Proveedor de datos</b>	La falta de evaluación de riesgos relacionados con la privacidad y la ética puede llevar a la recopilación y el uso no ético de datos, lo que puede resultar en problemas legales y reputacionales.					
	<table border="1"><thead><tr><th>Riesgo inherente (Impacto x Probabilidad)</th><th>Conclusión de la evaluación</th></tr></thead><tbody><tr><td>5 (Impacto crítico) x 1 (Improbable) = 5</td><td rowspan="2">El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.</td></tr><tr><td>Riesgo inherente (5) &gt; Apetito (4)</td></tr></tbody></table>	Riesgo inherente (Impacto x Probabilidad)	Conclusión de la evaluación	5 (Impacto crítico) x 1 (Improbable) = 5	El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.	Riesgo inherente (5) > Apetito (4)
Riesgo inherente (Impacto x Probabilidad)	Conclusión de la evaluación					
5 (Impacto crítico) x 1 (Improbable) = 5	El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.					
Riesgo inherente (5) > Apetito (4)						
	<table border="1"><thead><tr><th>Control</th><th>Mitigación y riesgo social</th></tr></thead><tbody><tr><td>Realizar una evaluación exhaustiva de los riesgos de privacidad y ética asociados con la recopilación y el uso de datos. Establecer políticas y prácticas claras para garantizar el cumplimiento de las regulaciones de</td><td>Mitigación: 70% Riesgo residual: 1,5</td></tr></tbody></table>	Control	Mitigación y riesgo social	Realizar una evaluación exhaustiva de los riesgos de privacidad y ética asociados con la recopilación y el uso de datos. Establecer políticas y prácticas claras para garantizar el cumplimiento de las regulaciones de	Mitigación: 70% Riesgo residual: 1,5	
Control	Mitigación y riesgo social					
Realizar una evaluación exhaustiva de los riesgos de privacidad y ética asociados con la recopilación y el uso de datos. Establecer políticas y prácticas claras para garantizar el cumplimiento de las regulaciones de	Mitigación: 70% Riesgo residual: 1,5					

protección de datos y la ética en la IA. Esto puede incluir la anonimización de datos sensibles y la obtención de consentimiento informado cuando sea necesario.

### Identificación de componentes

#### Propietario de datos

### Identificación de riesgos

Falta de claridad en la propiedad, responsabilidad y autoridad sobre los datos que se utilizarán en el sistema de IA. Puede resultar en disputas legales o problemas éticos relacionados con el acceso y el uso de los datos.

#### Riesgo inherente (Impacto x Probabilidad)

5 (Impacto crítico) x 1 (Improbable) = 5

Riesgo inherente (5) > Apetito (4)

#### Conclusión de la evaluación

El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.

#### Control

Realizar un inventario completo de los datos relevantes y establecer acuerdos claros sobre la propiedad y el uso de los datos tales como acuerdos de licencia de datos, acuerdos de confidencialidad y contratos que establezcan los términos de uso.

#### Mitigación y riesgo social

Mitigación: 70%  
Riesgo residual: 1,5

## Paso 2. Documentación, comunicación y consulta durante la fase de ideación del sistema de IA

El apetito al riesgo debe definirse de forma cuantitativa. Para este ejemplo, se decide establecer una escala del apetito al riesgo del 1 al 15 (siendo el 15 el valor máximo de riesgo que la organización está dispuesta a aceptar).

Esta escala debe alinearse con otras dos escalas: una que va del 1 al 5 para clasificar los riesgos en cinco niveles de impacto (desde muy bajo o inexistente hasta muy alto o crítico, en orden ascendente) y otra que abarca del 1 al 3 para evaluar la probabilidad de que estos riesgos se materialicen (desde improbable hasta muy probable, en orden ascendente). Para el ejemplo que se presenta, se determina el apetito al riesgo en un nivel intermedio, un valor 4 sobre 15.

### 4.3.2. Diseño del sistema de IA

#### Paso 3. Análisis de riesgos en el diseño del sistema de IA

Contexto						
Las capacidades de la organización, recursos y conocimientos destinados al diseño, desarrollo e implantación del sistema.						
Identificación de componentes	Identificación de riesgos					
<b>Científicos de datos e Ingenieros de datos</b>	Posibilidad de que el diseño de la arquitectura del sistema de IA no sea el más adecuado y conlleve la necesidad de cambios significativos en las etapas posteriores del desarrollo.					
	<table border="1"> <thead> <tr> <th>Riesgo inherente (Impacto x Probabilidad)</th> <th>Conclusión de la evaluación</th> </tr> </thead> <tbody> <tr> <td>3 (Impacto medio) x 2 (Probable) = 6</td> <td rowspan="2">El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.</td> </tr> <tr> <td>Riesgo inherente (6) &gt; Apetito (4)</td> </tr> </tbody> </table>	Riesgo inherente (Impacto x Probabilidad)	Conclusión de la evaluación	3 (Impacto medio) x 2 (Probable) = 6	El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.	Riesgo inherente (6) > Apetito (4)
Riesgo inherente (Impacto x Probabilidad)	Conclusión de la evaluación					
3 (Impacto medio) x 2 (Probable) = 6	El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.					
Riesgo inherente (6) > Apetito (4)						
	<table border="1"> <thead> <tr> <th>Control</th> <th>Mitigación y riesgo social</th> </tr> </thead> <tbody> <tr> <td>Realizar una revisión exhaustiva de las opciones arquitectónicas disponibles, teniendo en cuenta los requisitos del proyecto y realizar pruebas de concepto y prototipado para evaluar diferentes enfoques antes de tomar una decisión final.</td> <td>Mitigación: 80% Riesgo residual: 1,2</td> </tr> </tbody> </table>	Control	Mitigación y riesgo social	Realizar una revisión exhaustiva de las opciones arquitectónicas disponibles, teniendo en cuenta los requisitos del proyecto y realizar pruebas de concepto y prototipado para evaluar diferentes enfoques antes de tomar una decisión final.	Mitigación: 80% Riesgo residual: 1,2	
Control	Mitigación y riesgo social					
Realizar una revisión exhaustiva de las opciones arquitectónicas disponibles, teniendo en cuenta los requisitos del proyecto y realizar pruebas de concepto y prototipado para evaluar diferentes enfoques antes de tomar una decisión final.	Mitigación: 80% Riesgo residual: 1,2					
Identificación de componentes	Identificación de riesgos					
<b>Plataforma de ingesta de datos</b>	El sistema de IA no puede escalar eficazmente para manejar un aumento en la cantidad de datos o en la carga de trabajo, resultando en un sistema ineficiente o en la necesidad de realizar mejoras costosas.					
	<table border="1"> <thead> <tr> <th>Riesgo inherente (Impacto x Probabilidad)</th> <th>Conclusión de la evaluación</th> </tr> </thead> <tbody> <tr> <td>4 (Impacto alto) x 2 (Probable) = 8</td> <td rowspan="2">El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.</td> </tr> <tr> <td>Riesgo inherente (8) &gt; Apetito (4)</td> </tr> </tbody> </table>	Riesgo inherente (Impacto x Probabilidad)	Conclusión de la evaluación	4 (Impacto alto) x 2 (Probable) = 8	El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.	Riesgo inherente (8) > Apetito (4)
Riesgo inherente (Impacto x Probabilidad)	Conclusión de la evaluación					
4 (Impacto alto) x 2 (Probable) = 8	El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.					
Riesgo inherente (8) > Apetito (4)						
	<table border="1"> <thead> <tr> <th>Control</th> <th>Mitigación y riesgo social</th> </tr> </thead> <tbody> <tr> <td></td> <td></td> </tr> </tbody> </table>	Control	Mitigación y riesgo social			
Control	Mitigación y riesgo social					



Realizar una revisión exhaustiva de las opciones arquitectónicas disponibles, teniendo en cuenta los requisitos del proyecto y realizar pruebas de concepto y prototipado para evaluar diferentes enfoques antes de tomar una decisión final.

Mitigación: 80%  
Riesgo residual: 1,2

#### **Paso 4. Documentación, comunicación y consulta durante la fase de diseño del sistema de IA**

A continuación, se deben ejecutar la documentación, comunicación y consulta asociadas a los riesgos.

### 4.3.3. Desarrollo y entrenamiento del sistema de IA

#### Paso 5. Análisis de riesgos en el desarrollo y entrenamiento del sistema de IA

Contexto						
Las capacidades de la organización, recursos y conocimientos: el origen, la cantidad y la propiedad de los datos y modelos usados en el desarrollo, entrenamiento y validación del sistema.						
Identificación de componentes	Identificación de riesgos					
<b>Datos de entrenamiento</b>	Escasez de datos que puede comprometer la viabilidad de la IA y/o comprometer y limitar sus resultados. Esto puede ser explotado deliberada o involuntariamente durante la ingestión de datos.					
	<table border="1"> <thead> <tr> <th>Riesgo inherente (Impacto x Probabilidad)</th> <th>Conclusión de la evaluación</th> </tr> </thead> <tbody> <tr> <td>3 (Impacto medio) x 2 (Probable) = 6</td> <td rowspan="2">El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.</td> </tr> <tr> <td>Riesgo inherente (6) &gt; Apetito (4)</td> </tr> </tbody> </table>	Riesgo inherente (Impacto x Probabilidad)	Conclusión de la evaluación	3 (Impacto medio) x 2 (Probable) = 6	El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.	Riesgo inherente (6) > Apetito (4)
Riesgo inherente (Impacto x Probabilidad)	Conclusión de la evaluación					
3 (Impacto medio) x 2 (Probable) = 6	El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.					
Riesgo inherente (6) > Apetito (4)						
	<table border="1"> <thead> <tr> <th>Control</th> <th>Mitigación y riesgo social</th> </tr> </thead> <tbody> <tr> <td>Definir y supervisar los indicadores del buen funcionamiento del modelo y garantizar que los sistemas de IA cumplen los requisitos de seguridad y calidad de datos a lo largo de todo el ciclo de vida.</td> <td>Mitigación: 70% Riesgo residual: 1,8</td> </tr> </tbody> </table>	Control	Mitigación y riesgo social	Definir y supervisar los indicadores del buen funcionamiento del modelo y garantizar que los sistemas de IA cumplen los requisitos de seguridad y calidad de datos a lo largo de todo el ciclo de vida.	Mitigación: 70% Riesgo residual: 1,8	
Control	Mitigación y riesgo social					
Definir y supervisar los indicadores del buen funcionamiento del modelo y garantizar que los sistemas de IA cumplen los requisitos de seguridad y calidad de datos a lo largo de todo el ciclo de vida.	Mitigación: 70% Riesgo residual: 1,8					
Identificación de componentes	Identificación de riesgos					
<b>Propietario de datos</b>	Los propietarios de los datos pueden tratar de ocultar la información que se va a suministrar al sistema de IA como parte de sus intereses comerciales y, en consecuencia, introducir sesgos.					
	<table border="1"> <thead> <tr> <th>Riesgo inherente (Impacto x Probabilidad)</th> <th>Conclusión de la evaluación</th> </tr> </thead> <tbody> <tr> <td>3 (Impacto medio) x 2 (Probable) = 6</td> <td rowspan="2">El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.</td> </tr> <tr> <td>Riesgo inherente (6) &gt; Apetito (4)</td> </tr> </tbody> </table>	Riesgo inherente (Impacto x Probabilidad)	Conclusión de la evaluación	3 (Impacto medio) x 2 (Probable) = 6	El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.	Riesgo inherente (6) > Apetito (4)
Riesgo inherente (Impacto x Probabilidad)	Conclusión de la evaluación					
3 (Impacto medio) x 2 (Probable) = 6	El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.					
Riesgo inherente (6) > Apetito (4)						
	<table border="1"> <thead> <tr> <th>Control</th> <th>Mitigación y riesgo social</th> </tr> </thead> <tbody> <tr> <td></td> <td></td> </tr> </tbody> </table>	Control	Mitigación y riesgo social			
Control	Mitigación y riesgo social					

Definición de KPI para medir los sesgos en los grupos protegidos, umbrales de tolerancia y planes de remediación para eliminar las posibles imparcialidades para garantizar que los modelos sean imparciales.

Mitigación: 80%  
Riesgo residual: 1,2

## **Paso 6.** Documentación, comunicación y consulta durante la fase de desarrollo y entrenamiento del sistema de IA

A continuación, se deben ejecutar la documentación, comunicación y consulta asociadas a los riesgos.

#### 4.3.4. Validación de los resultados del sistema de IA

### Paso 7. Análisis de riesgos en la validación de los resultados del sistema de IA

Contexto						
Las capacidades de la organización, recursos y conocimientos: el origen, la cantidad y la propiedad de los datos y modelos usados en el desarrollo, entrenamiento y validación del sistema.						
Identificación de componentes	Identificación de riesgos					
<b>Datos de validación</b>	Introducción de datos adversos en el conjunto de datos de validación que comprometan la adecuada validación del modelo entrenado.					
	<table border="1"> <thead> <tr> <th>Riesgo inherente (Impacto x Probabilidad)</th> <th>Conclusión de la evaluación</th> </tr> </thead> <tbody> <tr> <td>5 (Impacto crítico) x 1 (Improbable) = 5</td> <td rowspan="2">El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.</td> </tr> <tr> <td>Riesgo inherente (5) &gt; Apetito (4)</td> </tr> </tbody> </table>	Riesgo inherente (Impacto x Probabilidad)	Conclusión de la evaluación	5 (Impacto crítico) x 1 (Improbable) = 5	El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.	Riesgo inherente (5) > Apetito (4)
Riesgo inherente (Impacto x Probabilidad)	Conclusión de la evaluación					
5 (Impacto crítico) x 1 (Improbable) = 5	El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.					
Riesgo inherente (5) > Apetito (4)						
	<table border="1"> <thead> <tr> <th>Control</th> <th>Mitigación y riesgo social</th> </tr> </thead> <tbody> <tr> <td>Establecer políticas y procedimientos claros que prohíban el acceso o la modificación de los datos de validación durante el proceso de entrenamiento y realizar verificaciones y auditorías periódicas de los conjuntos de datos de validación para identificar anomalías.</td> <td>Mitigación: 80% Riesgo residual: 1</td> </tr> </tbody> </table>	Control	Mitigación y riesgo social	Establecer políticas y procedimientos claros que prohíban el acceso o la modificación de los datos de validación durante el proceso de entrenamiento y realizar verificaciones y auditorías periódicas de los conjuntos de datos de validación para identificar anomalías.	Mitigación: 80% Riesgo residual: 1	
Control	Mitigación y riesgo social					
Establecer políticas y procedimientos claros que prohíban el acceso o la modificación de los datos de validación durante el proceso de entrenamiento y realizar verificaciones y auditorías periódicas de los conjuntos de datos de validación para identificar anomalías.	Mitigación: 80% Riesgo residual: 1					
Identificación de componentes	Identificación de riesgos					
<b>Modelos</b>	Un archivo de modelo legítimo es sustituido por completo por un archivo de modelo envenenado. Esta amenaza puede materializarse explotando las posibles debilidades de los proveedores de la nube.					
	<table border="1"> <thead> <tr> <th>Riesgo inherente (Impacto x Probabilidad)</th> <th>Conclusión de la evaluación</th> </tr> </thead> <tbody> <tr> <td>4 (Impacto alto) x 2 (Probable) = 8</td> <td rowspan="2">El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.</td> </tr> <tr> <td>Riesgo inherente (8) &gt; Apetito (4)</td> </tr> </tbody> </table>	Riesgo inherente (Impacto x Probabilidad)	Conclusión de la evaluación	4 (Impacto alto) x 2 (Probable) = 8	El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.	Riesgo inherente (8) > Apetito (4)
Riesgo inherente (Impacto x Probabilidad)	Conclusión de la evaluación					
4 (Impacto alto) x 2 (Probable) = 8	El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.					
Riesgo inherente (8) > Apetito (4)						
	<table border="1"> <thead> <tr> <th>Control</th> <th>Mitigación y riesgo social</th> </tr> </thead> <tbody> <tr> <td></td> <td></td> </tr> </tbody> </table>	Control	Mitigación y riesgo social			
Control	Mitigación y riesgo social					

Considerar factores de seguridad, como las medidas de protección de datos implementadas, la reputación del proveedor en cuanto a seguridad y las garantías contractuales en caso de incidentes de seguridad.

Mitigación: 80%  
Riesgo residual: 1,6

### **Paso 8. Documentación, comunicación y consulta durante la fase de validación de los resultados del sistema de IA**

A continuación, se deben ejecutar la documentación, comunicación y consulta asociadas a los riesgos.

### 4.3.5. Despliegue y mantenimiento del sistema de IA

#### Paso 9. Análisis de riesgos en el despliegue y mantenimiento del sistema de IA

Contexto						
Los recursos destinados al mantenimiento del sistema, así como cambios en las principales tendencias tecnológicas y otros factores legales y regulatorios.						
Identificación de componentes	Identificación de riesgos					
<b>Parámetros de entrenamiento</b>	Desafío y gasto asociado a la necesidad de reentrenar periódicamente el modelo de IA para mantener su rendimiento y precisión a lo largo del tiempo.					
	<table border="1"> <thead> <tr> <th>Riesgo inherente (Impacto x Probabilidad)</th> <th>Conclusión de la evaluación</th> </tr> </thead> <tbody> <tr> <td>4 (Impacto alto) x 2 (Probable) = 8</td> <td rowspan="2">El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.</td> </tr> <tr> <td>Riesgo inherente (8) &gt; Apetito (4)</td> </tr> </tbody> </table>	Riesgo inherente (Impacto x Probabilidad)	Conclusión de la evaluación	4 (Impacto alto) x 2 (Probable) = 8	El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.	Riesgo inherente (8) > Apetito (4)
	Riesgo inherente (Impacto x Probabilidad)	Conclusión de la evaluación				
	4 (Impacto alto) x 2 (Probable) = 8	El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.				
Riesgo inherente (8) > Apetito (4)						
<table border="1"> <thead> <tr> <th>Control</th> <th>Mitigación y riesgo social</th> </tr> </thead> <tbody> <tr> <td>Utilizar técnicas de transferencia de aprendizaje y fine-tuning para aprovechar modelos preentrenados para complementar al sistema e implementar herramientas y procesos automatizados que optimicen el reentrenamiento del modelo.</td> <td>Mitigación: 80% Riesgo residual: 1,6</td> </tr> </tbody> </table>	Control	Mitigación y riesgo social	Utilizar técnicas de transferencia de aprendizaje y fine-tuning para aprovechar modelos preentrenados para complementar al sistema e implementar herramientas y procesos automatizados que optimicen el reentrenamiento del modelo.	Mitigación: 80% Riesgo residual: 1,6		
Control	Mitigación y riesgo social					
Utilizar técnicas de transferencia de aprendizaje y fine-tuning para aprovechar modelos preentrenados para complementar al sistema e implementar herramientas y procesos automatizados que optimicen el reentrenamiento del modelo.	Mitigación: 80% Riesgo residual: 1,6					
Identificación de componentes	Identificación de riesgos					
<b>Herramientas de retención y borrado de datos</b>	Cambio de regulación: cambios regulatorios (por ejemplo, protección de datos y privacidad) que pudieran afectar la disponibilidad y el uso de los datos necesarios para el sistema de IA.					
	<table border="1"> <thead> <tr> <th>Riesgo inherente (Impacto x Probabilidad)</th> <th>Conclusión de la evaluación</th> </tr> </thead> <tbody> <tr> <td>3 (Impacto medio) x 2 (Probable) = 6</td> <td rowspan="2">El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.</td> </tr> <tr> <td>Riesgo inherente (6) &gt; Apetito (4)</td> </tr> </tbody> </table>	Riesgo inherente (Impacto x Probabilidad)	Conclusión de la evaluación	3 (Impacto medio) x 2 (Probable) = 6	El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.	Riesgo inherente (6) > Apetito (4)
	Riesgo inherente (Impacto x Probabilidad)	Conclusión de la evaluación				
	3 (Impacto medio) x 2 (Probable) = 6	El riesgo es mayor al apetito definido, se decide implantar medidas de control que mitiguen el riesgo.				
Riesgo inherente (6) > Apetito (4)						
<table border="1"> <thead> <tr> <th>Control</th> <th>Mitigación y riesgo social</th> </tr> </thead> <tbody> <tr> <td></td> <td></td> </tr> </tbody> </table>	Control	Mitigación y riesgo social				
Control	Mitigación y riesgo social					



Realizar un seguimiento constante de las regulaciones aplicables y ajustar los procesos en consecuencia.

Mitigación: 80%  
Riesgo residual: 1,2

## **Paso 10. Documentación, comunicación y consulta durante la fase de despliegue y mantenimiento del sistema de IA**

A continuación, se deben ejecutar la documentación, comunicación y consulta previamente descritas.





## 5. Conclusiones

La inteligencia artificial ha emergido como un pilar fundamental de la transformación digital y es la herramienta con mayor capacidad para determinar la forma en la que las sociedades abordarán los desafíos y oportunidades en el futuro. Su relevancia radica en su capacidad para automatizar tareas, procesar datos a gran escala y tomar decisiones basadas en información. En la administración pública, la IA se ha convertido en un facilitador clave para aumentar la eficiencia, la productividad y la calidad de los servicios públicos y, en definitiva, para evolucionar la forma en la que la administración se relaciona con la ciudadanía.

**No obstante, la relevancia de la IA también trae consigo desafíos, como la ética en el uso de datos y la preocupación por la automatización del empleo. La IA es relevante tanto en términos de oportunidades como de desafíos, lo que subraya la importancia de su gestión efectiva.**

Contar con un proceso claro, transparente y objetivo para gestionar la IA en la administración pública es una herramienta indispensable para alcanzar estos beneficios. La colaboración de todas las partes interesadas, la coordinación de los distintos perfiles con conocimiento en torno a la IA y el establecimiento de medidas de validación o control pueden ayudar de forma significativa a desarrollar una gestión responsable, ética y segura de la IA en la administración pública.

La gestión de riesgos constituye un elemento esencial de este proceso. Tal y como fija la normativa europea, la gestión de los sistemas de IA debe considerar como parte central la clasificación del nivel de riesgo de los sistemas de IA y la adopción de medidas coherentes con cada nivel de riesgo. La identificación clara y continua de riesgos, su evaluación periódica y la implantación de medidas de respuesta son aspectos clave del éxito de la adopción de esta tecnología por parte del sector público.



## 6. Referencias

Agencia de la Unión Europea para la Ciberseguridad (ENISA). (s. f.). *AI Cybersecurity Challenges - Threat Landscape*. <https://europa.eu>

Agencia Gallega de Conocimiento en Salud. (s. f.). *Plataforma de innovación sanitaria*. <https://sergas.es>

Banco Interamericano de Desarrollo (BID). (s. f.). *Uso responsable de IA para política pública: manual de formulación de proyectos*. <https://iadb.org>

Britainthinks. (s. f.). *AI Governance*. <https://publishing.service.gov.uk>

CAF (Banco de Desarrollo de América Latina). (s. f.). *Conceptos fundamentales y uso responsable de la Inteligencia Artificial en el Sector Público*. <https://caf.com>

Centre for Data Ethics and Innovation. (s. f.). *The Roadmap to an effective AI assurance ecosystem*. <https://publishing.service.gov.uk>

Comisión Europea. (s. f.). *Inteligencia artificial: La Comisión continúa su trabajo sobre directrices éticas*. <https://europa.eu>

Emiratos Árabes Unidos. (s. f.). *AI Adoption Guideline in Govern*.

GAIN. (s. f.). *Guía de buenas prácticas para favorecer la CPI en Galicia*. <https://galiciainnovacion.es>

Gartner. (s. f.-a). *Definition of Total Cost of Ownership*. <https://gartner.com>

Gartner. (s. f.-b). *How to Staff Your AI Team*. <https://gartner.com>

Gobierno de España. (s. f.-a). *Diez años cooperando en evaluación de tecnologías sanitarias*. <https://sanidad.gob.es>

Gobierno de España. (s. f.-b). *Guía sobre Compra Pública Innovadora*. <https://mineco.gob.es>

Gobierno de Reino Unido. (s. f.-a). *A guide to using artificial Intelligence in the public sector*. <https://publishing.service.gov.uk>

Gobierno de Reino Unido. (s. f.-b). *Guidelines for AI procurement*. <https://publishing.service.gov.uk>

Gobierno de Reino Unido. (s. f.-c). *Managing your artificial intelligence project*. <https://www.gov.uk>

Gobierno de Reino Unido. (s. f.-d). *Survey on AI Assurance Methods*. <https://www.gov.uk>

IDRC. (s. f.). *Government Artificial Intelligence Readiness Index*.

Montaña Merchán Arribas. (s. f.). *Guía de uso de la Inteligencia Artificial en el Sector Público*. <https://wordpress.com>

Neptune. (s. f.). *How to Build Machine Learning Teams That Deliver*. <https://neptune.ai>

Organización Internacional de Normalización (ISO). (s. f.-a). *ISO 31000:2018 – Gestión de riesgos – Directrices*.

Organización Internacional de Normalización (ISO). (s. f.-b). *ISO/IEC 23894 - Tecnología de la información - IA - Orientaciones sobre la gestión de riesgos*.

SERGAS. (s. f.). *Guía para la elaboración y adaptación de informes rápidos de evaluación de tecnologías sanitarias*. <https://sergas.gal>

World Economic Forum. (s. f.). *Lineamientos para los gobiernos sobre adquisiciones de sistemas de inteligencia artificial*. <https://weforum.org>





XUNTA  
DE GALICIA



AXENCIA PARA A  
MODERNIZACIÓN  
TÉCNOLÓXICA DE GALICIA