

## **Definición de metodologías de trabajo y protocolos para la implementación de sistemas algorítmicos**

## ÍNDICE

|  |    |
|--|----|
| 1. Introducción  | 3  |
| 1.1. Antecedentes y motivación   | 3  |
| 1.2. El ciclo de vida de un servicio TIC en el Ayuntamiento de Barcelona       | 5  |
| 1.3. El ciclo de vida de un sistema algorítmico                                | 6  |
| 1.4. Gobernanza  | 12 |
| 2. Concepción del servicio   | 13 |
| 2.1. Divulgación de información  | 13 |
| 2.2. Asignación del riesgo   | 13 |
| 2.3. El estudio de impacto algorítmico (EIA)                                   | 14 |
| 2.4. Espacios de transparencia   | 16 |
| 2.5. Protección de datos   | 16 |
| 3. Contratación  | 17 |
| 3.1. Calidad de los datos  | 17 |
| 3.2. Derechos sobre los datos  | 18 |
| 3.3. Calidad del sistema algorítmico   | 19 |
| 3.4. Transparencia del sistema algorítmico                                     | 19 |
| 3.5. Protección de datos   | 21 |
| 4. Desarrollo, producción y explotación  | 21 |
| 4.1. Estrategia de gestión de riesgos en el desarrollo del sistema algorítmico | 22 |
| 4.2. Auditoría y otros tipos de inspección                                     | 22 |
| 4.3. Gastos  | 23 |
| 4.4. Espacios de transparencia   | 23 |
| 4.5. Protección de datos   | 24 |
| 5. Desmantelamiento  | 24 |
| 5.1. Destrucción de los datos  | 24 |
| ANEXO 1. Definiciones  | 26 |
| ANEXO 2. Modelo de solicitud   | 29 |

## 1. Introducció

### 1.1. Antecedents y motivació

El uso de las tecnologías basadas en sistemas algorítmicos e inteligencia artificial<sup>1</sup> ofrece oportunidades al Ayuntamiento de Barcelona para transformar los servicios municipales y hacer que sean más eficientes, proactivos y adaptados a las necesidades de la ciudadanía, así como para agilizar los procesos internos y respaldar la toma de decisiones. Sin embargo, el despliegue de este tipo de sistemas lleva asociados múltiples riesgos,<sup>2</sup> de modo que resulta imprescindible asegurar que su implementación introduce mecanismos de salvaguarda y garantía de los derechos fundamentales; que se lleva a cabo de forma proporcionada, supervisada y fundamentada —de acuerdo con la legislación vigente y los debidos estándares éticos y técnicos—, y que el proceso favorece la participación y deliberación constantes con la ciudadanía y las comunidades afectadas.

Este documento es una de las acciones previstas por la [“Medida de gobierno de la estrategia municipal de algoritmos y datos para el impulso ético de la inteligencia artificial”](#) publicada en el 2021 por el Ayuntamiento de Barcelona, y tiene como finalidad crear un protocolo y una metodología de trabajo común en el Ayuntamiento de Barcelona para la contratación pública o la admisión de pruebas piloto<sup>3</sup> de sistemas algorítmicos confiables. Se entienden como confiables todos aquellos sistemas que cumplan con los principios rectores definidos en la medida de gobierno: acción y supervisión humana; robustez técnica y seguridad; privacidad y gobernanza de los datos; transparencia; diversidad, inclusión y equidad; compromiso social y ambiental, y responsabilidad, rendición de cuentas y control democrático. Este protocolo es aplicable a cualquier sistema algorítmico impulsado<sup>4</sup> por el Ayuntamiento de Barcelona, independientemente de si ya ha sido desarrollado por el proveedor o si el proveedor desarrolla una iniciativa municipal desde el principio.

La metodología recogida en este documento combina: (1) el ciclo de vida de un servicio de tecnología de la información y la comunicación (TIC) en el Ayuntamiento de Barcelona, (2) el ciclo de vida de un sistema algorítmico —y sus particularidades—, (3) los mecanismos existentes en el Ayuntamiento con respecto a la protección de datos, y (4) los estándares éticos definidos por varias administraciones y organizaciones. El documento hace viables los principios, recomendaciones y estándares definidos en la “Medida de gobierno de la estrategia municipal de algoritmos y datos para el impulso ético de la inteligencia artificial”, el Reglamento general de protección de datos (RGPD), la Ley orgánica de protección de datos personales y garantía de los derechos digitales (LOPDGDD), la Propuesta de

<sup>1</sup> La inteligencia artificial (IA) utiliza la recolección y el análisis de datos masivos (*big data*) junto con algoritmos y el entrenamiento de sistemas inteligentes (*machine learning*) para realizar predicciones o recomendaciones más o menos automatizadas. Para más definiciones, véase el anexo 1.

<sup>2</sup> Los riesgos asociados a los sistemas algorítmicos, sobre todo en aquellos casos en los que estos se aplican en servicios sensibles que afectan a derechos fundamentales, tienen que ver con la falta de transparencia y la explicabilidad de las decisiones tomadas; el hecho de que a menudo se prioriza la optimización sin tener en cuenta a quién afecta; la dificultad para atribuir responsabilidades y garantizar la rendición de cuentas; la vulneración de la privacidad de los individuos, y el hecho de que tanto las bases de datos como el código de los algoritmos pueden incorporar sesgos sociales que afecten de manera discriminatoria a colectivos minorizados.

<sup>3</sup> El Ayuntamiento de Barcelona, a través de BIT Habitat, impulsa el testado de pilotos de soluciones urbanas innovadoras en fase precomercial con el fin de validar la utilidad y el interés público en el entorno urbano real antes de que estas sean comercializadas. En el marco de este programa, el Ayuntamiento de Barcelona testea sistemas de IA aplicados a la mejora de servicios urbanos.

<sup>4</sup> Por *impulsados* entendemos aquellos sistemas algorítmicos licitados por el Ayuntamiento de Barcelona, sean soluciones ya existentes o soluciones desarrolladas expresamente, o aquellos que sean testados en el marco del servicio Barcelona Innovat Lab, de BIT Habitat (véase la página 6 sobre esta cuestión).

Reglamento del Parlamento Europeo y del Consejo por el cual se establecen normas armonizadas en materia de inteligencia artificial, y la Ley 15/2022 integral para la igualdad de trato y no discriminación, así como los de documentos elaborados por otras ciudades<sup>5</sup> u organizaciones y el marco regulatorio sobre inteligencia artificial que ha propuesto la Comisión Europea (*AI Act* o Ley IA).<sup>6</sup> Aunque la ley europea sobre inteligencia artificial todavía se encuentra en proceso de tramitación legislativa, este protocolo incorpora varios elementos que estarán presentes en ella, como la clasificación de las aplicaciones de IA según su riesgo, para anticiparse y adaptarse de manera proactiva a los estándares que serán de cumplimiento obligatorio cuando la ley entre en vigor. Cuando las instituciones europeas aprueben el reglamento sobre IA, este protocolo se adaptará a su contenido final.

Con respecto a los sistemas algorítmicos que incluyan tratamiento de datos personales, este protocolo tiene en cuenta los tres criterios personales marcados por el RGPD en cuanto a la inteligencia artificial: requerimiento de que el tratamiento sea justo (art. 5), principio de minimización de datos (art. 5), evaluación de impacto de la protección de datos (art. 35).

La regulación de actuaciones automatizadas municipales del Ayuntamiento de Barcelona tiene como antecedente el Decreto de Comisión de Gobierno por el cual se regulan las actuaciones administrativas automatizadas municipales, aprobado en sesión de 22 de junio de 2022 y publicado en el *BOPB* de 23 de agosto de 2022.<sup>7</sup>

De esta combinación surge un protocolo que define, paso a paso, los mecanismos de garantía y salvaguarda de derechos que tienen que introducirse en cada momento de la implementación de un sistema de inteligencia artificial por parte del Ayuntamiento de Barcelona, los cuales incluyen, entre otros, la elaboración de estudios de impacto algorítmico, la introducción de cláusulas tipo para la contratación de sistemas de IA fiables, la obligación de llevar a cabo auditorías algorítmicas o la creación de espacios de participación y de mecanismos de comunicación ciudadana y transparencia, como los registros de algoritmos.

Este protocolo se basa en la clasificación de riesgos realizada por la Comisión Europea en el marco de la *IA Act*,<sup>8</sup> y establece varios tipos de mecanismos de garantía en función del nivel de riesgo del sistema que se licita. Los sistemas algorítmicos de riesgo inaceptable serán rechazados automáticamente, mientras que para el resto de sistemas se establecerán unos mecanismos en función del nivel de riesgo, tal como explicita este documento. Para los sistemas algorítmicos de alto riesgo —de manera sintética, aquellos que tienen un impacto directo en los derechos fundamentales de las personas—, este protocolo prevé una serie de mecanismos de garantía más estrictos que para los considerados de riesgo limitado o mínimo.

---

<sup>5</sup> Véase, por ejemplo, las cláusulas tipo para la contratación de sistemas algorítmicos de Ámsterdam.

<sup>6</sup> <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>.

<sup>7</sup> Este decreto se circunscribe a las actuaciones administrativas automatizadas dentro del procedimiento administrativo y tiene como finalidad identificar el uso de la tramitación electrónica automatizada en el procedimiento administrativo y establecer el protocolo para su autorización, todo ello en cumplimiento de la normativa administrativa de aplicación y como desarrollo de lo establecido en la Ordenanza reguladora de la administración electrónica.

<sup>8</sup> Puede encontrarse la clasificación de riesgos en la página 7.

Vista la novedad del campo de la IA en el sector público y la rápida evolución de las tecnologías emergentes y de los riesgos que llevan asociados, este protocolo se concibe como un documento vivo que irá evolucionando e incorporando la retroalimentación de los agentes internos y externos implicados, las lecciones aprendidas en el despliegue de proyectos por parte del Ayuntamiento de Barcelona, así como los avances normativos y las recomendaciones realizadas por personas expertas.

## 1.2. El ciclo de vida de un servicio TIC en el Ayuntamiento de Barcelona

Para garantizar que los sistemas algorítmicos contratados por el Ayuntamiento de Barcelona cumplen con los principios rectores definidos en la “Medida de gobierno de la estrategia municipal de algoritmos y datos para el impulso ético de la inteligencia artificial”, hay que alinear el ciclo de vida de un servicio TIC y los procedimientos de contratación pública con las particularidades del ciclo de vida de los sistemas algorítmicos e introducir, en cada una de estas fases, todas las salvaguardas necesarias.

Todo producto o servicio TIC tiene un ciclo de vida: introducción, crecimiento, madurez y declive. En concreto, las fases del ciclo de vida de un producto o servicio en el Ayuntamiento de Barcelona incluyen lo siguiente:

1. **Concepción del servicio:** fase en la que se descubre una nueva necesidad o surge la idea del proyecto y que puede incluir la elaboración de un anteproyecto u otros estudios previos.
2. **Contratación:** redacción de los pliegos para la adquisición del servicio (de desarrollo o de otra naturaleza).
3. **Desarrollo:** creación del código, documentación y otros artefactos, incluyendo el establecimiento de la infraestructura necesaria para su construcción.
4. **Producción:** despliegue del servicio, incluyendo la posible migración de datos y procesos desde uno o más servicios propios.
5. **Explotación:** fase que se extiende durante toda la vida útil del sistema en producción, incluidas las labores de operación y mantenimiento.
6. **Desmantelamiento:** fase que incluye todas las acciones necesarias que hay que realizar una vez que se termina la vida útil del producto o servicio TIC.

Este ciclo de vida, en el Ayuntamiento de Barcelona, tiene dos particularidades:

- La puesta en marcha y el mantenimiento de los servicios TIC corren a cargo del Instituto Municipal de Informática (IMI) del Ayuntamiento de Barcelona, un organismo autónomo cuyo objetivo es suministrar la mayoría de los servicios TIC al Ayuntamiento y a los órganos y empresas públicas que dependen de este.
- El Ayuntamiento de Barcelona, como administración pública del ámbito local, se rige por unos procedimientos administrativos regulados por ley. Estos, especialmente los

procedimientos de contratación pública, necesitan integrarse en el ciclo de vida de los servicios TIC.

### 1.3. El ciclo de vida de un sistema algorítmico

La creación de un sistema algorítmico también tiene unas fases definidas.

1. Formulación del problema: en esta fase, hay que definir claramente qué necesidad se pretende abordar con el sistema algorítmico y establecer qué significa que el modelo funcione bien. También hay que definir los *inputs* o valores de entrada que se introducirán en el sistema y los *outputs* o resultados que se espera obtener.<sup>9</sup>
2. Recogida y procesamiento de los datos.
3. Selección y construcción del modelo: los modelos codifican un conjunto de suposiciones sobre la relación potencial entre el *input* y el *output*, y pueden entenderse como la fórmula matemática que combina los *inputs* con un conjunto de parámetros o coeficientes de ponderación ajustables para obtener el *output* predicho.<sup>10</sup> En caso de que el sistema incorpore el aprendizaje automático, esta fase incluiría el entrenamiento del sistema.
4. Desarrollo y monitorización: dada la naturaleza de los sistemas algorítmicos, que difiere de la de otros sistemas de *software* por el hecho de que el modelo, los datos o los sistemas de aprendizaje pueden cambiar con el tiempo, es importante realizar un seguimiento detallado y riguroso de los sistemas una vez que se han puesto en funcionamiento.

En algunos casos, ciertas fases de este proceso se repiten de manera iterativa: puede ser que, en algún momento del desarrollo del sistema, se recojan nuevos datos y el modelo se reestructure o bien que este incorpore nuevos datos de manera continua (*machine learning*).

Teniendo en cuenta estas fases y las particularidades de los sistemas algorítmicos, en este documento se proponen los siguientes procesos y su relación con cada fase para establecer un protocolo que tendrá que seguir cualquier área del Ayuntamiento de Barcelona que quiera licitar un sistema algorítmico. En las figuras 1 y 2, se sintetiza gráficamente el proceso que se explica detalladamente en las siguientes páginas.

Es necesario enfatizar el hecho de que los sistemas algorítmicos que se liciten en el Ayuntamiento de Barcelona, como servicios TIC, serán licitados potencialmente por el IMI, que centralizará todas las fases que incluye el ciclo de vida de un sistema algorítmico, en estrecha colaboración con el área impulsora.

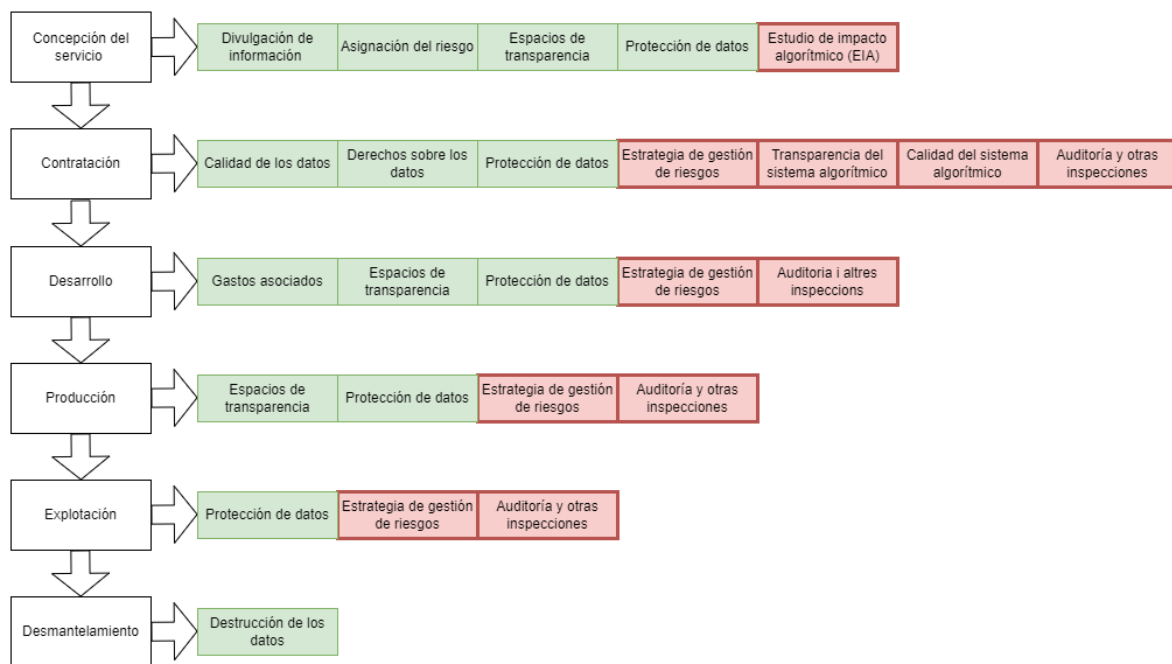
---

<sup>9</sup> Los *outputs* pueden ser resultados binarios (sí, no), una clasificación en categorías o una probabilidad (en una escala del 0 al 100 o en porcentajes). Puede ser que estos *outputs* sirvan simplemente para apoyar una decisión humana o que la decisión esté plenamente automatizada o bien que el sistema esté programado para tener tres posibles resultados: "sí", "no" y "quizás", lo que implicaría la revisión del resultado por parte de un humano.

<sup>10</sup> En la construcción de los modelos de los sistemas que utilizan *machine learning* se utilizan datos y un algoritmo de aprendizaje para establecer los parámetros o coeficientes del modelo, empezando con valores aleatorios que se van ajustando en función de los *outputs*.

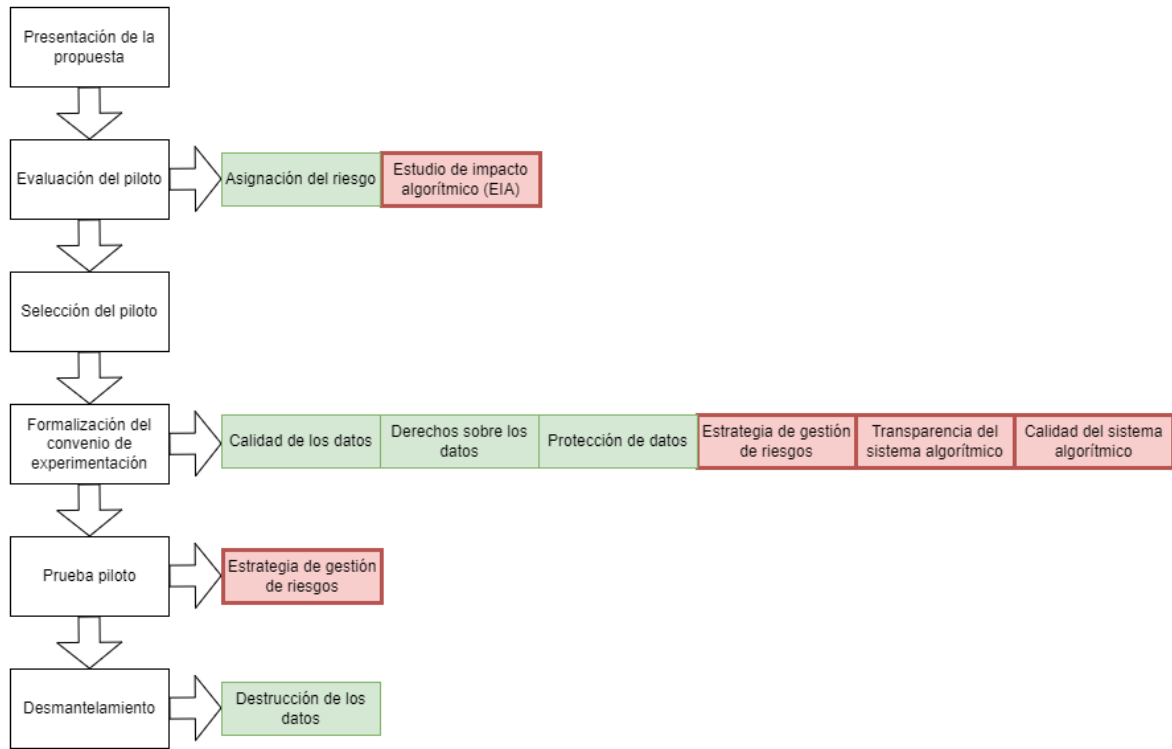
La mayoría de estos procesos solo tienen sentido en situaciones en las que el sistema algorítmico pueda tener un impacto negativo en la ciudadanía o en comunidades específicas, siguiendo la clasificación propuesta por la Comisión Europea con respecto al riesgo de los sistemas de inteligencia artificial.<sup>11</sup> En ningún caso se trata de un protocolo de aplicación general a cualquier sistema algorítmico, sino que se desplegará de manera sistemática en las aplicaciones de los sistemas algorítmicos de alto riesgo, y los órganos de gobernanza de la IA ética del Ayuntamiento podrán decidir si lo aplican, también, en aplicaciones de riesgo limitado y mínimo. Este riesgo será determinado por la Oficina Técnica de Inteligencia Artificial del IMI.

**Figura 1. Relación de las fases de desarrollo de un sistema algorítmico licitado por el Ayuntamiento de Barcelona con los mecanismos definidos en este protocolo**



**Figura 2. Relación de las fases del despliegue de un proyecto piloto de un sistema algorítmico impulsado por el Ayuntamiento de Barcelona con los mecanismos definidos en este protocolo**

<sup>11</sup> <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>.





## Testado de pilotos de inteligencia artificial en el Ayuntamiento de Barcelona

Aparte de los casos en los que el Ayuntamiento de Barcelona licita un sistema algorítmico, otra vía de entrada de este tipo de sistemas en la ciudad es a través del servicio Barcelona Innova Lab, impulsado por BIT Habitat y el IMI. En el marco de este servicio, se impulsa el testado de pilotos de sistemas algorítmicos aplicados a servicios urbanos en fase precomercial con el fin de validar su utilidad e interés público en el entorno urbano real antes de que los sistemas sean comercializados. En este caso, el Ayuntamiento de Barcelona no contrata el sistema algorítmico a través de una licitación, sino que facilita a la empresa experimentadora las herramientas para testar el sistema (como, por ejemplo, estableciendo diálogos con las áreas municipales implicadas para garantizar el acceso al espacio urbano, los activos municipales y la conectividad 5G). En estos casos, los mecanismos de salvaguarda y garantía de derechos se introducen de dos formas:

- a) La admisión de los pilotos se lleva a cabo a través de una convocatoria con criterios estrictos de elegibilidad. Estos criterios, en caso de que el piloto involucre la aplicación de algoritmos de riesgo alto, pueden incluir una valoración del Consejo Asesor en Inteligencia Artificial, Ética y Derechos Digitales.
- b) Se introducen las cláusulas adecuadas en el convenio de experimentación que se firma con las empresas experimentadoras, siguiendo los criterios que se aplican a las cláusulas tipo para la contratación de sistemas de IA fiables (véase el apartado 3 de este documento).

### Clasificación de riesgos propuesta por la Comisión Europea

La clasificación de riesgos de cualquier sistema de IA licitado o impulsado por el Ayuntamiento de Barcelona será la que determine el reglamento europeo de IA, en proceso de tramitación. La Oficina Técnica del IMI será la que adjudicará la categoría de riesgo a las solicitudes de sistemas.

Estos riesgos hacen referencia al contexto de aplicación del sistema y a la finalidad que se le da. Los aspectos relacionados con la protección de datos personales y la ciberseguridad se regulan de acuerdo con lo previsto en las leyes vigentes sobre la materia.

### RIESGO INACEPTABLE

Se prohibirán todos los sistemas de IA que se consideren una clara amenaza para la seguridad y los derechos de las personas, desde el *social scoring* de los gobiernos hasta los juguetes que utilicen asistencia de voz y que fomenten comportamientos peligrosos. Se prohibirán también los sistemas de identificación biométrica remota en espacios públicamente accesibles.

### RIESGO ALTO

Aquí se incluye la tecnología de IA utilizada en los siguientes casos:

- identificación biométrica remota en tiempo real o en diferido de personas físicas;
- infraestructuras críticas (por ejemplo, el transporte) que podrían poner en riesgo la vida y la salud de las personas;
- formación educativa o profesional, que puede determinar el acceso a la educación y el desarrollo profesional de alguien (por ejemplo, la evaluación de los exámenes);
- los componentes de seguridad de los productos (por ejemplo, la aplicación de la IA en la cirugía asistida por robots);
- la ocupación, la gestión de los trabajadores y el acceso al autoempleo (por ejemplo, software de clasificación de currículums para los procedimientos de contratación);
- servicios privados y públicos esenciales (por ejemplo, la cualificación de créditos que nieguen a la ciudadanía la oportunidad de obtener un préstamo);
- una aplicación de la ley que puede interferir en los derechos fundamentales de las personas (por ejemplo, evaluación de la fiabilidad de las pruebas);
- la gestión de la migración, el asilo y el control de fronteras (por ejemplo, verificación de la autenticidad de los documentos de viaje), y
- la administración de justicia y los procesos democráticos (por ejemplo, la aplicación de la ley a un conjunto concreto de hechos).

Los sistemas de IA de alto riesgo estarán sujetos a estrictas obligaciones antes de poder comercializarse o implementarse. Deberán tener:

- sistemas adecuados de evaluación y mitigación de riesgos;
- alta calidad de los conjuntos de datos que alimentan el sistema para minimizar los riesgos y los resultados discriminatorios;
- registro de la actividad para garantizar la trazabilidad y transparencia de los resultados;
- documentación detallada que proporcione toda la información necesaria sobre el sistema y su finalidad para que las autoridades puedan evaluar su conformidad;
- información clara y adecuada a cada persona usuaria;
- medidas adecuadas de supervisión humana para minimizar el riesgo, o
- alto nivel de solidez, seguridad y precisión.

Todos los sistemas de identificación biométrica a distancia se consideran de alto riesgo y están sujetos a requisitos estrictos. El uso de la identificación biométrica a distancia en espacios de acceso público con finalidades policiales está, en principio, prohibido.

Se definen y regulan estrictamente algunas excepciones, como cuando es necesario buscar a un niño desaparecido, prevenir una amenaza terrorista específica e inminente o detectar, localizar, identificar o procesar a una persona autora o sospechosa de un delito grave.

Esta utilización está sujeta a la autorización de un órgano judicial u otro órgano independiente y a unos límites adecuados en cuanto a tiempo, alcance geográfico y bases de datos consultadas.

### RIESGO LIMITADO

El riesgo limitado se refiere a los sistemas de IA con obligaciones específicas de transparencia. Cuando se utilizan sistemas de IA, como los *chatbots* o asistentes de voz, las personas usuarias tienen que ser conscientes de que están interactuando con una máquina para poder tomar una decisión informada sobre si continúan o retroceden.

### RIESGO MÍNIMO O NINGÚN RIESGO

La propuesta permite el uso libre de la IA de riesgo mínimo. Eso incluye aplicaciones como los videojuegos con IA o los filtros de correo basura. La gran mayoría de los sistemas de IA utilizados actualmente en la Unión Europea (UE) entran en

esta categoría.

#### 1.4. Gobernanza

En el marco de la “Medida de gobierno de la estrategia municipal de algoritmos y datos para el impulso ético de la inteligencia artificial”, se han creado órganos de gobernanza internos para garantizar que los mecanismos definidos en este protocolo serán aplicados en la contratación o desarrollo piloto de cualquier sistema algorítmico por parte del Ayuntamiento de Barcelona:

- **Comisión Transversal para el Impulso de la IA Ética en el Ayuntamiento de Barcelona.** En esta comisión están representadas las áreas municipales y entidades dependientes relacionadas con la implementación de proyectos de inteligencia artificial. Entre otras, tiene la función de validar los sistemas basados en inteligencia artificial que se apliquen en el Ayuntamiento, propuestos por las áreas municipales y entidades dependientes, de dirimir los conflictos que puedan surgir en la aplicación/uso de algoritmos de inteligencia artificial en los procesos internos y servicios públicos prestados por el Ayuntamiento y entidades dependientes, y de velar para que la voz de la ciudadanía sea tenida en cuenta en el proceso de diseño e implantación de servicios basados en IA.
- **Oficina Técnica de IA del IMI.** Esta oficina se encarga de definir, desde el punto de vista técnico, los estándares, herramientas, uso e implementación de tecnologías de IA sobre los proyectos desarrollados/implantados en el Ayuntamiento de Barcelona, así como de desarrollar proyectos piloto de IA, prestar apoyo para conceptualizar las soluciones a escala tecnológica, metodológica y ética, y evaluar los riesgos.

Al mismo tiempo, también en el marco del despliegue de su “Medida de gobierno de la estrategia municipal de algoritmos y datos para el impulso ético de la inteligencia artificial”, el Ayuntamiento de Barcelona ha creado dos órganos de participación para que la ciudadanía, el mundo académico, las instituciones y el mundo empresarial puedan participar en el proceso de deliberación y ejercer una supervisión externa sobre los sistemas algorítmicos que se desarrollan. (Véase el anexo 4 para la lista completa de miembros de cada órgano):

- **Consejo Asesor en Inteligencia Artificial, Ética y Derechos Digitales.** Este órgano estará integrado por miembros del mundo académico, organizaciones de la sociedad civil y empresas tecnológicas, y tiene, entre otras, la función de evaluar externamente el impacto de los sistemas algorítmicos. Como se define más adelante, el Consejo Asesor emitirá estudios de impacto algorítmico (EIA) preceptivos y no vinculantes sobre los algoritmos considerados de alto riesgo.
- **Pacto de ciudad para el impulso del humanismo tecnológico.** Este pacto, cuyos agentes firmantes se reunirán una vez al año, funcionará como órgano participativo amplio de entidades, empresas e instituciones comprometidas con el desarrollo ético de la inteligencia artificial y la protección de los derechos digitales.

Dentro del pacto se crearán grupos de trabajo de temas específicos, según las necesidades, intereses y problemas que vayan emergiendo con respecto a la gobernanza de las tecnologías digitales en la ciudad. En el marco de este pacto se producirán diálogos de carácter más general (y no tan técnico, como en el Consejo), entre otras cuestiones, sobre la idoneidad e impacto social de los sistemas de IA que se impulsan desde el Ayuntamiento de Barcelona.

## 2. Concepción del servicio

El momento que transcurre entre que se detecta la necesidad de desarrollar un sistema algorítmico en el Ayuntamiento de Barcelona y el desarrollo de dicho sistema (ya sea a través de un proyecto piloto o de los procedimientos de contratación estándares) es clave para garantizar que su implementación se realiza de forma transparente, segura, legítima y proporcionada y que se incluye a las personas afectadas y las personas externas expertas en el proceso de deliberación, en aquellas aplicaciones que tienen un riesgo más elevado.

### 2.1. Divulgación de información

Obligatorio: **RIESGO ALTO** **RIESGO MEDIO** **RIESGO MÍNIMO O NINGÚN RIESGO**

Como parte del estudio previo a la adquisición de un sistema algorítmico, el área municipal impulsora de este estudio rellenará una ficha de solicitud de servicio (en adelante, la Solicitud; véase el anexo 2 para la ficha completa), en la cual detallará:

- Su definición interna de sistema de toma de decisiones automatizadas.
- La información sobre la necesidad que tiene, por qué se considera que la necesidad podría ser resuelta con un sistema algorítmico, como podría responder este a esa necesidad, y si se han considerado otras alternativas.
- Los detalles de cada sistema automatizado de toma de decisiones que se plantea licitar, incluyendo detalles sobre su finalidad, alcance, datos que se utilizarán (si los hay), política de utilización y calendario de aplicación.
- La definición preliminar de posibles riesgos de inexactitud, sesgos o daños a comunidades específicas, que definirá medidas de mitigación de riesgos.
- La definición de los mecanismos para permitir el acceso de personas expertas externas que revisarán el sistema una vez que haya sido desplegado.

### 2.2. Asignación del riesgo

Obligatorio: **RIESGO ALTO** **RIESGO MEDIO** **RIESGO MÍNIMO O NINGÚN RIESGO**

La Oficina Técnica de IA del IMI recibirá y evaluará la Solicitud y, en un plazo de dos semanas, emitirá un informe que incluirá una extensión de la información proporcionada por el área municipal y una valoración del riesgo. Concretamente, se detallará:

- La información sobre el sistema de toma de decisiones automatizadas, incluyendo detalles sobre su finalidad, alcance, política de utilización y calendario de aplicación.
- El estudio de posibles riesgos de inexactitud, sesgos o daños a comunidades afectadas, y definición de medidas de mitigación dirigidas a minimizar dichos

riesgos, incluyendo diálogos proactivos que fomenten el compromiso con miembros de esas comunidades.

- La explicación de los mecanismos adecuados para permitir el acceso a personas expertas externas que se ofrezcan a revisar el sistema una vez que haya sido desplegado.
- Determinación del riesgo (inaceptable/alto/limitado/mínimo o ningún riesgo).

En función del riesgo determinado por la Oficina Técnica, se decidirán los siguientes pasos:

- Riesgo inaceptable: la Solicitud del área municipal será denegada.
- Riesgo alto: la Oficina Técnica informará de ello a la Comisión Transversal y realizará una petición de estudio de impacto algorítmico preceptivo no vinculante remunerado (a partir de ahora, EIA) al Consejo Asesor en IA, el cual, en un máximo de tres semanas,<sup>12</sup> emitirá dicho estudio.
- Riesgo limitado y mínimo o ningún riesgo: No será necesario ningún EIA y el área municipal implicada podrá seguir adelante con la redacción de los pliegos.

### 2.3. El estudio de impacto algorítmico (EIA)

Obligatorio: **RIESGO ALTO**

Para los sistemas de alto riesgo, el Consejo Asesor emitirá un informe de impacto algorítmico (EIA). Este tipo de estudios describen lo que hace el sistema algorítmico y quién es parte responsable en caso de que surja algún problema.<sup>13</sup> Se lleva a cabo en el momento prelicitación, antes de que el Ayuntamiento de Barcelona haya comprometido la utilización de un sistema algorítmico, con el fin de conocer los posibles resultados indeseables e inseguros que afectan a los derechos fundamentales de las personas y negociar e introducir las medidas necesarias para corregirlos.

Los EIA son cruciales para ofrecer a todas las partes interesadas la oportunidad de discutir y evaluar la adopción del sistema, asegurar la transparencia y garantizar el acceso de la ciudadanía a la información que la afecta, así como para evaluar su idoneidad teniendo en cuenta el caso de uso al que se quiere aplicar, tanto antes de la puesta en producción como durante la misma. En los casos de alto riesgo, los EIA correspondientes y las medidas de mitigación que se introduzcan servirán como buenas prácticas tanto para el Ayuntamiento de Barcelona como para otras administraciones locales. El presente protocolo irá actualizándose y mejorando con estas experiencias.

Partiendo del informe “Algorithmic Impact Assessments: A Practical Framework For Public Agency Accountability”,<sup>14</sup> un EIA está diseñado para alcanzar cuatro objetivos:

---

<sup>12</sup>Se establece un tiempo de entrega del informe de tres semanas con el fin de evitar la ralentización del proceso de contratación.

<sup>13</sup> Las evaluaciones de impacto no son nada nuevo. Se han aplicado en ámbitos científicos y políticos tan amplios como la protección del medio ambiente, los derechos humanos, la protección de datos y la privacidad. El estudio de impacto algorítmico se basa en estos marcos combinándolos con la investigación de personas expertas científicas y políticas que se ha ido desarrollando en torno a la responsabilidad algorítmica.

<sup>14</sup> <https://ainowinstitute.org/aiareport2018.pdf>

1. Respetar el derecho de la ciudadanía a saber qué sistemas afectan a sus vidas, enumerando y describiendo públicamente como son y de qué manera estos sistemas afectan significativamente a personas y comunidades.
2. Aumentar la experiencia y la capacidad interna de los organismos públicos para evaluar los sistemas que construyen o adquieren, y así poder anticiparse a los aspectos que pueden suscitar preocupación, como los impactos negativos o las violaciones de garantías procesales.
3. Garantizar la restitución de cuentas de los sistemas algorítmicos y facilitar que personas expertas externas revisen, auditen y evalúen estos sistemas utilizando métodos que les permitan identificar y detectar fallos o problemas.
4. Garantizar a la ciudadanía la oportunidad de responder y, si es necesario, impugnar el uso de un determinado sistema o el enfoque realizado sobre la restitución de cuentas de un sistema algorítmico determinado.

De manera orientativa y no restrictiva, los EIA preceptivos y no vinculantes contendrán la siguiente información relacionada con el sistema algorítmico que se quiere licitar:

- Descripción, propósito, alcance, política y calendario de utilización
- Descripción del contexto de aplicación
- Necesidad y proporcionalidad del sistema
- Identificación de partes implicadas
- Revisión ética: valores y conflictos (*trade-offs*)
- Impacto en los derechos fundamentales de las personas y comunidades afectadas
- Supervisión humana
- Definición de posibles riesgos y de medidas de mitigación
- Recomendaciones

En el caso del Ayuntamiento de Barcelona, y en línea con los debates que se han producido en los últimos años en el marco de la creación de una regulación de la inteligencia artificial en la Comisión Europea, los estudios de impacto serán elaborados por personas expertas externas al Ayuntamiento, reunidas en el Consejo Asesor en Inteligencia Artificial, Ética y Derechos Digitales, con el fin de garantizar la máxima imparcialidad sobre los sistemas algorítmicos. Estos EIA serán consultables por la ciudadanía y por cualquier agente interesado en el registro público de algoritmos —junto con el resto de información pública sobre el sistema, tal como se definirá más adelante—.

Cabe destacar que los EIA no son la solución a los problemas que plantean algunos sistemas algorítmicos y que tienen que ir acompañados de los instrumentos necesarios para garantizar una implementación legítima, segura y transparente, como se define en los siguientes capítulos de este documento. Aun así, son herramientas prácticas para introducir el debate sobre la utilización de estos sistemas, proporcionar información a la ciudadanía y determinar su idoneidad. Una vez licitado el sistema algorítmico, su impacto se determinará y revisará a través de las auditorías algorítmicas llevadas a cabo por una entidad externa al Ayuntamiento de Barcelona. El proceso de contratación de las auditorías lo liderará el IMI.

## 2.4. Espacios de transparencia

Obligatorio: **RIESGO ALTO** **RIESGO MEDIO** **RIESGO MÍNIMO O NINGÚN RIESGO**

Los sistemas algorítmicos plantean muchos riesgos y retos: (1) pueden incluir los sesgos existentes en las bases de datos (por falta de representatividad o por introducción de variables *proxy*,<sup>15</sup> por ejemplo); (2) el proceso por el cual llegan a una decisión puede ser opaco o poco transparente y explicable, sobre todo en aquellos sistemas que incorporan el aprendizaje automático o *machine learning* —porque no se pueden explicar los pesos en los coeficientes que ponderan el resultado final a partir de los *inputs* aportados—; (3) a menudo se prioriza la optimización sin tener en cuenta a quien afecta, etcétera. Todo esto hace que el Ayuntamiento de Barcelona se encuentre ante la necesidad de velar por la legitimidad democrática respecto de los derechos fundamentales, la robustez técnica y la seguridad, y la transparencia y la explicabilidad de cualquier sistema algorítmico que se desarrolle a escala municipal.

Concretamente, la ciudadanía en general y las comunidades afectadas dispondrán de mecanismos de participación antes, durante y después de la compra y despliegue del sistema algorítmico. En el momento prelicitación, se establecerán los siguientes mecanismos de participación y transparencia:

- **Publicación de la información pública** del servicio que quiere licitarse (en cualquier tipo de sistema algorítmico, independientemente de su riesgo) en el registro de algoritmos.
- **Revisión del sistema de IA por parte del Consejo Asesor y elaboración del estudio de impacto algorítmico prelicitación.**
- Tal como prevé la instrucción de protección de datos, fase de información al público en el caso de supuestos presentes en la lista de alto riesgo publicada por las autoridades de control de protección de datos.

## 2.5. Protección de datos

Obligatorio: **RIESGO ALTO** **RIESGO MEDIO** **RIESGO MÍNIMO O NINGÚN RIESGO**

En materia de protección de datos, durante la fase de concepción se revisará la inclusión de datos personales y la potencialidad de obtener decisiones automatizadas, así como la naturaleza y la adecuación normativa de dichas decisiones, según lo previsto en el RGPD y la normativa española sobre las restricciones de las decisiones automatizadas que puedan tener efectos jurídicos o equivalentes para las personas.<sup>16</sup>

Si el sistema incluye datos personales o los puede inferir y pasa a la fase de contratación y explotación, tendrá que registrarse como nuevo tratamiento en el Registro de actividades de

<sup>15</sup> Las variables *proxy* son variables aparentemente no relevantes, pero que actúan como una variable inobservable o inmensurable. Tal como señala el manual [Artificial Intelligence Primer de Nueva York](#), un ejemplo podría ser el código postal, que en EE.UU. suele relacionarse directamente con la racialización o el nivel de ingresos; por lo tanto, utilizarlo como variable puede conducir a decisiones sesgadas. Estas variables tienen que tratarse con sumo cuidado para evitar comportamientos inexactos, inadecuados e incluso perjudiciales.

<sup>16</sup> A modo de ejemplo, en la medida en que tratemos datos de carácter personal o cuando indirectamente se pueda llegar a identificar a personas, se tendrá que determinar una base legal de acuerdo con el artículo 6 del RGPD, teniendo en cuenta que el interés legítimo no es aplicable a las administraciones públicas. En el caso de decisiones automatizadas que puedan afectar a las personas, se cumplirá con la obligación normativa de transparencia definida en el RGPD.



tratamiento (RAT) municipal, de acuerdo con el procedimiento previsto en la instrucción de protección de datos.

### 3. Contratación

La clasificación de riesgos y los EIA tendrán el siguiente efecto en la redacción de los pliegos:

- *Riesgo inaceptable*: No se redactarán los pliegos.
- *Riesgo alto*: El área solicitante valorará la idoneidad de continuar con el proceso de licitación y, en caso afirmativo, y valorará la incorporación en los pliegos de las recomendaciones emitidas en el EIA.
- *Riesgo limitado y mínimo o ningún riesgo*: No será necesario un EIA y el área municipal podrá seguir adelante con la redacción de los pliegos.

En esta etapa es fundamental introducir los mecanismos necesarios para garantizar que el sistema algorítmico contratado cumple con los estándares éticos de inteligencia artificial definidos por el Ayuntamiento de Barcelona. Por ese motivo se incluirán, en los pliegos de licitación de cualquier sistema algorítmico contratado por el IMI, las cláusulas tipo para la contratación de sistemas de IA fiables,<sup>17</sup> que están en redacción por parte de los servicios de contratación del Ayuntamiento de Barcelona.

Estas cláusulas obligan al contratista a desarrollar el sistema algorítmico en cuestión bajo las condiciones que se explican a continuación. Cabe destacar que algunas de las cláusulas se aplican solo a los sistemas algorítmicos calificados como de “riesgo alto”. Al inicio de cada apartado, se indica en qué casos se aplicará la condición: “independiente del riesgo” quiere decir que la condición se aplicará en cualquier sistema algorítmico que se desarrolle, y “alto riesgo” quiere decir que la condición solo se aplicará a los sistemas de alto riesgo.

#### 3.1. Calidad de los datos

Obligatorio: **RIESGO ALTO** **RIESGO MEDIO** **RIESGO MÍNIMO O NINGÚN RIESGO**

La calidad de los datos utilizados en el desarrollo de un sistema algorítmico puede afectar de forma notoria a la calidad del sistema algorítmico en sí. Cuando se recopilan datos, estos pueden incorporar imprecisiones y errores contruidos socialmente.<sup>18</sup> Las medidas que tiene que adoptar el contratista deberán centrarse en evitarlo. Por este motivo, hay que establecer unos requisitos de calidad de los datos.

Por otra parte, hay que diferenciar entre los casos en los que el contratista desarrolla un sistema algorítmico con los datos proporcionados por el Ayuntamiento de Barcelona y los casos en los que es el contratista quien desarrolla el sistema algorítmico a partir de los

<sup>17</sup> El documento que contiene las cláusulas se encuentra en construcción y está basado en la propuesta de las [cláusulas estándares de contratación de sistemas de IA fiables](#) elaboradas por el Ayuntamiento de Amsterdam.

<sup>18</sup> Cabe recordar que la recolección y clasificación de los datos que entrenan y alimentan a los sistemas algorítmicos son realizadas por personas humanas que pueden tener prejuicios —racistas o machistas, por ejemplo— o visiones sesgadas que pueden llevarlas a crear etiquetas o categorías para la ordenación de los datos que incluyan sesgos. Aparte de la introducción de la subjetividad humana, las bases de datos también pueden llevar a imprecisiones debido a la falta de información sobre colectivos minorizados que no queden representados de forma significativa. Un sistema algorítmico entrenado solo con datos de un colectivo mayoritario muy probablemente ofrecerá resultados sesgados cuando se aplique a otros colectivos de características diferentes.

datos obtenidos a través de otras fuentes. El motivo es que, si el contratista recibe los datos del Ayuntamiento de Barcelona, se le puede considerar responsable del análisis, la ordenación y la edición de los datos, pero no de su recogida. Al fin y al cabo, en este caso, es el Ayuntamiento de Barcelona el que los ha recogido y, por lo tanto, es lógico que la responsabilidad recaiga sobre el Ayuntamiento.

Para asegurar la calidad de los datos utilizados en el desarrollo de un sistema algorítmico, el Ayuntamiento de Barcelona exigirá la utilización de un método fundamentado en la recogida (cuando sea el caso), el análisis, la organización y la edición de los datos. Este método tiene que basarse, siempre que sea posible, en pautas científicas y en otras normas existentes en la práctica real.

Otro objetivo de la exigencia de que la recogida (cuando sea el caso), el análisis, la organización y la edición de los datos se lleven a cabo siguiendo un método fundamentado es evitar, hasta donde sea posible, sesgos en los datos. Lo que se considere sesgo tendrá que evaluarse en cada caso en concreto. En cualquier caso, se cree que existe sesgo cuando la utilización de los datos conduce o puede conducir a una forma de discriminación protegida por las leyes y normativas aplicables, incluyendo el artículo 14 del Convenio europeo de derechos humanos.<sup>19</sup>

Los conjuntos de datos utilizados (tanto para entrenar los algoritmos como para hacerlos funcionar) pueden presentar sesgos sociales inadvertidos, ya sea por falta de representatividad de colectivos, por introducción de variables *proxy*, por lagunas o por modelos de gestión incorrectos. Todo sesgo demostrable y discriminatorio tiene que ser identificado y, siempre que sea posible, el contratista tiene que eliminarlo tan pronto como sea posible. Si no es posible eliminarlo, tendrá que advertir al Ayuntamiento de Barcelona tan pronto como sea posible y de forma clara. Con el fin de asegurar que el contratista cumple con su responsabilidad de mantener la calidad de los datos, se llevará a cabo una auditoría algorítmica anual.

### 3.2. Derechos sobre los datos

Obligatorio: **RIESGO ALTO** **RIESGO MEDIO** **RIESGO MÍNIMO O NINGÚN RIESGO**

El contratista será el encargado del tratamiento, pero no el responsable de los datos, ya que el que definirá la finalidad del tratamiento y los medios que se utilicen será el Ayuntamiento. Eso será aplicable tanto a las situaciones en las que el Ayuntamiento de Barcelona proporciona datos al contratista como en aquellas en las que los datos se recogen en el contexto del contrato. En ambas situaciones, hay que partir del principio base de que los derechos sobre los datos corresponden exclusivamente al Ayuntamiento de Barcelona, que es el responsable, y no al contratista.

Los contratos de sistemas algorítmicos incluirán acuerdos claros respecto de los derechos y obligaciones sobre los datos, poniendo especial énfasis en la sensibilidad hacia la materia y en que el adjudicatario resultante cumpla con la normativa de protección de datos.<sup>20</sup> El

<sup>19</sup> [https://www.echr.coe.int/Documents/Convention\\_CAT.pdf](https://www.echr.coe.int/Documents/Convention_CAT.pdf)

<sup>20</sup> Hay que recordar que no es posible poseer los datos en el sentido del derecho de propiedad, ya que el derecho de propiedad se limita a los elementos materiales que pueden ser objeto del control humano. Los datos tampoco suelen estar sujetos a derechos de propiedad intelectual, ya que el desarrollo de los datos no suele implicar decisiones creativas y, por lo

Ayuntamiento de Barcelona establecerá acuerdos contractuales específicos con cada proveedor sobre los derechos o las obligaciones relativos a los datos que se intercambian.

Tal como indican las cláusulas tipo para la contratación de sistemas de IA fiables creadas por Ámsterdam, de manera general se acordará que el contratista no tiene derecho a utilizar los datos proporcionados por el Ayuntamiento de Barcelona ni los datos generados durante el servicio para ningún otro fin que no sea la ejecución del contrato. A petición del Ayuntamiento de Barcelona y al final del contrato, el contratista destruirá los datos o se los entregará al Ayuntamiento de Barcelona, tal como determinan las obligaciones previstas para encargados de tratamiento de las cláusulas diseñadas para utilizar en el Ayuntamiento de Barcelona.

### 3.3. Calidad del sistema algorítmico

Obligatorio: **RIESGO ALTO**

En la medida de lo posible: **RIESGO MEDIO** **RIESGO MÍNIMO O NINGÚN RIESGO**

Cuando en la fase de prelicitación se haya establecido que un sistema algorítmico es de alto riesgo, tendrán que añadirse todos los mecanismos necesarios a las cláusulas tipo para la contratación de sistemas de IA fiables para asegurar que dicho sistema se desarrolla de forma segura, transparente, explicable, legítima, respetando los derechos fundamentales y de acuerdo con el marco legal aplicable. Por lo tanto, será necesario evaluar qué regulación es aplicable en cada caso concreto (por ejemplo, uso de tecnologías de anonimización y RGPD) y establecer la obligación de que el sistema algorítmico se desarrolle siguiendo un método fundamentado y basado, siempre que sea posible, en pautas científicas y otras normas existentes en la práctica real.

El Ayuntamiento de Barcelona establece un requisito de calidad específico para el sistema algorítmico: el sistema algorítmico no puede presentar ningún defecto y tiene que proporcionar el resultado para el cual fue concebido. Eso significa, entre otras cosas, que el sistema algorítmico tiene que ser apto para su utilización prevista y que debe tener la capacidad de tomar decisiones correctas (por ejemplo, clasificar correctamente la información en las categorías adecuadas) o, con respecto a su capacidad, de realizar adecuadamente predicciones, formular recomendaciones o tomar decisiones correctas basadas en datos o modelos. Eso implica también que el sistema algorítmico tiene que ser seguro y fiable. Es necesario que el sistema algorítmico sea robusto desde el punto de vista técnico y el contratista tiene que garantizar que no se pueda destinar a usos malintencionados.

### 3.4. Transparencia del sistema algorítmico

Obligatorio: **RIESGO ALTO**

En la medida de lo posible: **RIESGO MEDIO** **RIESGO MÍNIMO O NINGÚN RIESGO**

---

tanto, los datos no estarían sujetos a derechos de autor. En determinadas circunstancias, una base de datos que contenga datos puede estar sujeta a derechos sobre las bases de datos. No obstante, el derecho sobre las bases de datos solo ofrece protección ante la reutilización de toda la base de datos o de una parte sustancial de esta, pero no sobre la reutilización de datos específicos. Para más información sobre la adecuación del RGPD al tratamiento de datos que incorpora inteligencia artificial, véase: <https://www.aepd.es/sites/default/files/2020-02/adecuacion-rgpd-ia.pdf>.

Como señala un informe de Algorithm Watch,<sup>21</sup> la transparencia no asegura, por sí misma, una conformidad plena con los requerimientos éticos, pero es una condición necesaria para alcanzar dicha conformidad. El Ayuntamiento de Barcelona entiende la transparencia como una herramienta necesaria para garantizar que el sistema algorítmico se desarrolla de acuerdo con los estándares éticos y de seguridad técnica, y que la ciudadanía tiene acceso a la información sobre los sistemas que la afectan.

El Ayuntamiento de Barcelona diferencia entre transparencia del procedimiento, transparencia técnica y explicabilidad, según los tipos de información que puede haber sobre un sistema algorítmico.

- **Transparencia del procedimiento:** se trata del suministro de información pública relativa al sistema algorítmico o al proceso seguido en el desarrollo y la aplicación de dicho sistema y de los datos utilizados en este contexto. En cualquier caso, la información suministrada ha de ser suficiente para garantizar que, si surgen daños, puede rastrearse la causa, así como para garantizar la comprensión de las opciones y de los supuestos asumidos y las categorías de datos utilizadas en el desarrollo del sistema algorítmico. Tiene que incluir también la forma en la que tendrá lugar la intervención humana en el sistema algorítmico, la metodología utilizada para identificar los riesgos, los riesgos identificados y las medidas adoptadas para mitigarlos, así como las partes que han participado en el desarrollo del sistema algorítmico y sus funciones.
- **Transparencia técnica:** se trata del suministro de información por parte del proveedor al Ayuntamiento de Barcelona —respetando la propiedad intelectual— que permita comprender el funcionamiento técnico del sistema algorítmico. Esta información tiene que incluir la divulgación del código fuente del sistema, las especificaciones técnicas y los datos utilizados en el desarrollo del sistema, la información técnica sobre cómo se han obtenido los datos, el proceso de desarrollo utilizado, la justificación de la elección del modelo concreto y sus parámetros, así como la información sobre el rendimiento del sistema.
- **Explicabilidad:** información sobre por qué un sistema algorítmico lleva a una decisión o resultado concreto. Esto incluirá una indicación clara de los factores clave que han llevado a un sistema algorítmico a un resultado concreto y los cambios que tienen que hacerse en los datos de entrada con el fin de obtener un resultado diferente. Hacer que un sistema algorítmico sea explicable incluye proporcionar toda la información técnica y de otro tipo necesaria para explicar, en los espacios de participación o en procedimientos como estudios de impacto, auditorías algorítmicas o procedimientos de impugnación, de recurso u otros procedimientos judiciales, cómo se ha llegado a una decisión y ofrecer a la otra parte o a cualquier parte interesada el acceso a la información y una protección jurídica satisfactoria.

La transparencia en el procedimiento y la transparencia técnica se centran en la obtención de la información que el Ayuntamiento de Barcelona considere necesaria para evaluar el sistema algorítmico. Además, la transparencia en el procedimiento se centra también en proporcionar a la ciudadanía y a otras partes interesadas información general sobre la utilización del sistema algorítmico.

---

<sup>21</sup> Algorithm Watch, "Automated Decision-Making Systems in the Public Sector: An Impact Assessment Tool for Public Authorities", de Michele Loi, Anna Mätzener, Angela Müller, Matthias Spielkamp (junio de 2021). Disponible en línea.

Con respecto a la explicabilidad, aunque de manera general se obligará al contratista a proporcionar la información detallada sobre los factores clave que han llevado a un sistema algorítmico a un resultado particular o sobre los cambios que deberán realizarse en los datos de entrada para obtener resultados diferentes, podría darse el caso de que, en una situación específica, eso no fuera técnicamente factible o necesario. En estos casos, se opta deliberadamente por ofrecer a las partes la posibilidad de establecer una excepción en el contrato que se formalice.

El Ayuntamiento de Barcelona será libre de divulgar la información facilitada por el proveedor relacionada con la transparencia del procedimiento. En cambio, la información relacionada con la transparencia técnica que facilitarán los proveedores no será de carácter público y tendrá como objetivo que el Ayuntamiento de Barcelona, en las auditorías anuales correspondientes (encargadas a una entidad tercera), pueda verificar si el contratista ha cumplido con sus obligaciones. Esta puede incluir información comercial confidencial, por lo que el Ayuntamiento de Barcelona asumirá la obligación de mantenerla en secreto y destruirla una vez que la auditoría haya finalizado. En caso de que el contratista no quiera proporcionar la información relacionada con la transparencia técnica al Ayuntamiento de Barcelona, se estipulará en el contrato que el contratista tendrá que proporcionarla directamente a la entidad tercera que lleve a cabo la auditoría.

### 3.5. Protección de datos

Obligatorio: **RIESGO ALTO** **RIESGO MEDIO** **RIESGO MÍNIMO O NINGÚN RIESGO**

En el diseño de los proyectos de IA, así como de los algoritmos implicados, tendrán que aplicarse los criterios de privacidad desde el diseño y por defecto que marca el RGPD (art. 25), incluyendo aquellas funcionalidades que sirvan para salvaguardar los derechos de acceso, rectificación, supresión, oposición, portabilidad, olvido y el derecho de limitación al tratamiento (ARSOPOL) de la ciudadanía, permitiendo la recuperación, modificación o cancelación de cualquier dato personal, así como la oposición o limitación de algunos de sus procesos.

Cuando un área municipal saque adelante la contratación de un sistema algorítmico, en caso de que se traten directa o indirectamente datos de carácter personal, para proceder al registro del tratamiento en el RAT de la Sede Electrónica, tendrá que rellenar también el documento DANTE (documento de evaluación de nuevos tratamientos).

## 4. Desarrollo, producción y explotación

Una vez licitado el servicio, el proveedor pasará a desarrollar y a poner en funcionamiento el sistema algorítmico. En estas fases del ciclo de vida (desarrollo, producción y explotación) habrá que garantizar que el sistema funciona según su objetivo inicial y que los riesgos que emerjan se mitigan. En caso de que el sistema incorpore el aprendizaje automático, habrá que velar por que las operaciones llevadas a cabo por este último continúen siendo explicables y el sistema no incorpore sesgos.

Las buenas prácticas y aprendizajes que emerjan durante esta fase alimentarán el presente protocolo y quedarán reflejadas en este documento vivo con el fin de mejorar el proceso de cara a proyectos posteriores.

#### 4.1. Estrategia de gestión de riesgos en el desarrollo del sistema algorítmico

Obligatorio: **RIESGO ALTO**

Es importante que cualquier riesgo que pueda darse durante la utilización de un sistema algorítmico se identifique durante el desarrollo del mismo y que se proponga una estrategia de mitigación de riesgos adecuada.

Las licitaciones de los sistemas algorítmicos del Ayuntamiento de Barcelona incluirán el diseño y la ejecución de una estrategia de gestión de riesgos sistemática, que tendrá que irse actualizando. Se propone seguir la lista de comprobación propuesta en el documento “Directrices éticas para una IA fiable” del Grupo Independiente de Expertos de Alto Nivel sobre Inteligencia Artificial, facilitado por la Comisión Europea.<sup>22</sup>

La estrategia de gestión de riesgos que se siga tiene que estar en consonancia con la naturaleza del sistema algorítmico. Los sistemas algorítmicos de riesgo alto serán sometidos a una estrategia de gestión de riesgos más intensa y estricta que la que se aplicará a los sistemas algorítmicos desplegados en casos de riesgo limitado o bajo.

El resultado de la ejecución de una estrategia de gestión de riesgos tendrá que ponerse a disposición del Ayuntamiento de Barcelona mediante la solicitud de transparencia del procedimiento. Además el Ayuntamiento de Barcelona evaluará, mediante auditorías anuales, si se ha llevado a término una estrategia de gestión de riesgos correcta.

En el caso de que se utilice un sistema algorítmico ya desarrollado por el contratista, este tendrá que demostrar que ya se ha llevado a cabo una estrategia de gestión de riesgos.

#### 4.2. Auditoría y otros tipos de inspección

Obligatorio: **RIESGO ALTO**

El Ayuntamiento de Barcelona tiene que poder comprobar si el contratista cumple con los requisitos especificados en el contrato y, con esta finalidad, obligará al contratista a colaborar en estas comprobaciones y a proporcionar todos los datos necesarios para que una entidad externa lleve a cabo una auditoría algorítmica. Cada sistema algorítmico que se licite tendrá que pasar una auditoría anual durante todos los años que esté en funcionamiento.

Los resultados de la auditoría se recogerán siempre en un informe. Si el Ayuntamiento de Barcelona establece que el contratista no cumple con las obligaciones contractuales, el contratista quedará obligado a enmendar los incumplimientos identificados (y descritos en el informe) en un plazo razonable (también fijado por el Ayuntamiento de Barcelona). Si el contratista no enmienda los incumplimientos descritos por el Ayuntamiento de Barcelona en el plazo de enmiendas fijado, el contratista podrá ser penalizado según lo establecido en el contrato.

El Ayuntamiento de Barcelona tendrá derecho a difundir públicamente las conclusiones del informe en el registro de algoritmos.

<sup>22</sup> Independent High-Level Expert Group on Artificial Intelligence (2019). Ethics Guidelines For Trustworthy AI, p. 26-31. Disponible en línea: <https://www.aepd.es/sites/default/files/2019-12/ai-ethics-guidelines.pdf>



### 4.3. Gastos

Obligatorio: **RIESGO ALTO** **RIESGO MEDIO** **RIESGO MÍNIMO O NINGÚN RIESGO**

El Ayuntamiento de Barcelona no tendrá que abonar un importe adicional al contratista por los trabajos que se deriven de las acciones descritas anteriormente.

### 4.4. Espacios de transparencia

Obligatorio: **RIESGO ALTO** **RIESGO MEDIO** **RIESGO MÍNIMO O NINGÚN RIESGO**

El Ayuntamiento de Barcelona pondrá a disposición de la ciudadanía y de todas las partes interesadas la información pública relacionada con los sistemas algorítmicos que licite en el registro público de algoritmos que se creará a este efecto. De forma preliminar y no restrictiva<sup>23</sup>, se prevé que se publique lo siguiente:

- Información básica sobre el sistema algorítmico:
  - Finalidad del sistema algorítmico
  - Área municipal que lo impulsa
  - Agentes implicados en las decisiones del sistema
  - Alcance del sistema algorítmico
  - Datos que se utilizarán
  - Incorporación del aprendizaje automático
  - Política de utilización
  - Calendario de aplicación
  - Riesgos que podrían aparecer y su estrategia de gestión
  - Legitimación de la finalidad en el caso de tratamiento de datos personales, previamente alineado en el RAT
- Los estudios de impacto algorítmico llevados a cabo por el Consejo Asesor (solo de los sistemas que acaben siendo licitados)
- El modelo de cláusulas tipo para la contratación de sistemas de IA fiables
- Los informes de auditoría algorítmica
- Informe de balance final
- Información clara y detallada de los mecanismos de comunicación ciudadana disponibles para tramitar peticiones de información o consultas, incidencias, quejas y sugerencias en relación con el funcionamiento de los sistemas algorítmicos

Aparte del registro de algoritmos, aquellos sistemas que incorporen tratamiento de datos personales tendrán que incorporarse al RAT de la Sede Electrónica.

Además de la información pública disponible en el registro de algoritmos, el Ayuntamiento de Barcelona facilitará la participación de la sociedad civil, de empresas y de entidades a través del Pacto de ciudad por el humanismo tecnológico, para dar la oportunidad a todas las partes implicadas de debatir la aplicación de los sistemas algorítmicos y de conocer sus resultados. Además, el Ayuntamiento de Barcelona contará con las personas expertas del

---

<sup>23</sup> Dada la novedad del campo de la IA en el sector público y el hecho de que todavía no se ha aprobado una regulación armonizada y común a escala europea, el Ayuntamiento de Barcelona trabaja constantemente con otras administraciones, organizaciones internacionales, la Comisión Europea y personas expertas independientes para mantener al día y consolidar unos estándares comunes e interoperables, entre otras cuestiones, sobre el tipo de información que tiene que ponerse al alcance de la ciudadanía para garantizar una IA ética, segura, robusta y confiable.

Consejo Asesor en Inteligencia Artificial, Ética y Derechos Digitales para hacer un seguimiento detallado de los sistemas de IA que se encuentren en funcionamiento. En concreto, se añadirá un punto en el orden del día de las reuniones de ambos órganos en el que el área municipal impulsora del sistema algorítmico y el IMI tendrán que informar sobre el desarrollo del sistema algorítmico y las personas miembros de los órganos podrán realizar preguntas al respecto y debatirlo.

#### 4.5. Protección de datos

Obligatorio: **RIESGO ALTO** **RIESGO MEDIO** **RIESGO MÍNIMO O NINGÚN RIESGO**

En el diseño de los proyectos de IA, así como de los algoritmos implicados, tendrán que aplicarse los criterios de privacidad desde el diseño y por defecto que marca el RGPD (art. 25), incluyendo aquellas funcionalidades que sirvan para salvaguardar los derechos de acceso, rectificación, supresión, oposición, portabilidad, olvido y el derecho de limitación al tratamiento (ARSOPOL) de la ciudadanía, permitiendo la recuperación, modificación o cancelación de cualquier dato personal, así como la oposición o limitación de algunos de sus procesos.

En la fase de explotación, tendría que tenerse muy claro el principio de *accountability* o responsabilidad proactiva exigido por el RGPD, tanto con las autoridades de control como con las personas afectadas.

### 5. Desmantelamiento

#### 5.1. Destrucción de los datos

Obligatorio: **RIESGO ALTO** **RIESGO MEDIO** **RIESGO MÍNIMO O NINGÚN RIESGO**

Cuando finalice el contrato con el proveedor, este estará obligado a destruir, de acuerdo con los requerimientos del RGPD, todos los datos de carácter personal relacionados con el desarrollo y explotación del sistema algorítmico. En cuanto al resto de datos (no de carácter personal), se fomentará su reutilización, estableciendo acuerdos *ad hoc* con cada proveedor.

Con respecto al código fuente, en caso de que el Ayuntamiento de Barcelona no haya adquirido su propiedad intelectual, el proveedor podrá venderlo a otros clientes y seguir utilizándolo, siempre que este no incluya ningún tipo de dato recogido en el marco del contrato con el Ayuntamiento de Barcelona.

Una vez desmantelado el sistema algorítmico, el área impulsora del sistema, con el apoyo del IMI, elaborará un informe de balance final que incluirá una actualización del estudio de impacto algorítmico, indicadores sobre el funcionamiento del sistema a escala técnica y social, cualquier otra información sobre la transparencia técnica, así como una explicación de cómo se reutilizará el conocimiento adquirido en el servicio y cómo se integrarán las buenas prácticas en futuros proyectos.





## ANEXO 1. Definiciones

**Algoritmo, en un sentido muy amplio:** desde modelos de regresión y árboles de decisión que pueden hacer predicciones y agilizan procesos, hasta sistemas más complejos, como las redes neuronales y los modelos bayesianos, que funcionan con aprendizaje automático a medida que van haciendo cálculos y predicciones avanzadas.

**Análisis de datos:** debe entenderse en un sentido amplio. El análisis de datos puede incluir combinación, limpieza, ordenación, clasificación y derivación de datos.

**Actuación administrativa automatizada:** de acuerdo con la normativa administrativa aplicable, se entiende por *actuación administrativa automatizada*, la actuación realizada íntegramente a través de medios electrónicos, en el marco de un procedimiento administrativo, en la cual no haya intervenido de manera directa el personal al servicio del Ayuntamiento y que utiliza alguno de los sistemas de firma establecidos en la normativa administrativa de aplicación, sello electrónico o código seguro de verificación.

**Sesgo algorítmico:** errores sistemáticos y repetibles en un sistema informático que crean resultados “injustos”, como “privilegiar” una categoría por encima de otra de manera diferente a la función prevista del algoritmo.

**Datos personales:** toda información sobre una persona física identificada o identificable (“la persona interesada”); se considerará como persona física identificable a toda persona cuya identidad pueda determinarse, directa o indirectamente, en particular mediante un identificador, como, por ejemplo, un nombre, un número de identificación, datos de localización, un identificador en línea o uno o varios elementos propios de la identidad física, fisiológica, genética, psíquica, económica, cultural o social de esa persona.

**Destinatario:** la persona física o jurídica, autoridad pública, servicio u otro organismo al cual se le comuniquen datos personales, se trate o no de una persona tercera. No obstante, no se considerarán destinatarios las autoridades públicas que puedan recibir datos personales en el marco de una investigación concreta, de conformidad con el derecho de la Unión o de los Estados miembros. El tratamiento de tales datos por parte de estas autoridades públicas se realizará conforme a las normas en materia de protección de datos aplicables a los fines del tratamiento.

**Elaboración de perfiles:** toda forma de tratamiento automatizado de datos personales consistente en utilizar datos personales para evaluar determinados aspectos personales de una persona física, en particular para analizar o predecir aspectos relativos a rendimiento profesional, situación económica, salud, preferencias personales, intereses, fiabilidad, comportamiento, ubicación o movimientos de esta persona física.

**Encargado del tratamiento, o encargado:** la persona física o jurídica, autoridad pública, servicio u otro organismo que trate datos personales por cuenta del responsable del tratamiento.

**Explicable/explicabilidad:** poder explicar individualmente por qué un sistema algorítmico lleva a una decisión o resultado en particular. A no ser que las partes estén expresamente de acuerdo con lo contrario, eso incluirá una indicación clara de los factores clave que han llevado a un sistema algorítmico a un resultado particular y los cambios que tienen que hacerse para que el sistema lleve a un resultado diferente. Hacer explicable un sistema algorítmico incluye proporcionar toda la información técnica para explicar, en los procedimientos de objeción, de apelación u otros procedimientos judiciales, cómo se ha producido una decisión y ofrecer a las partes afectadas o a cualquier otra parte interesada la oportunidad de evaluar cómo se ha adoptado una decisión, con la finalidad única de ofrecer una protección jurídica realista.

**Inteligencia artificial:** disciplina científica que incluye varios enfoques y técnicas, como el aprendizaje automático (del que derivan el aprendizaje profundo y el aprendizaje por refuerzo), el razonamiento automático (que incluye la planificación, programación, representación y razonamiento de los conocimientos, la búsqueda y la optimización) y la robótica (que incluye el control, la percepción, los sensores y los accionadores), así como la integración de todo el resto de técnicas en sistemas “ciberfísicos”. Según la definición del Consejo de Europa: “La IA es entendida como un conjunto de ciencias, teorías y técnicas cuyo propósito es reproducir mediante una máquina las capacidades cognitivas de un ser humano”.

**Limitación del tratamiento:** el marcaje de los datos de carácter personal conservados, con el fin de limitar su tratamiento en el futuro.

**Responsable del tratamiento, o responsable:** la persona física o jurídica, autoridad pública, servicio u otro organismo que, solo o junto con otros, determine los fines y medios del tratamiento; si el derecho de la Unión o de los estados miembros determina los fines y medios del tratamiento, entonces el responsable del tratamiento o los criterios específicos para su nombramiento podrá establecerlos el derecho de la Unión o de los estados miembros.

**Sistemas inteligentes:** algoritmos de decisión automatizada, capaces de tomar decisiones sin intervención humana. La única exigencia es que el algoritmo analice el escenario proporcionado a partir de unos datos y tome una decisión en función de este análisis.

**Tercero:** persona física o jurídica, autoridad pública, servicio u organismo diferente de la persona interesada, del responsable del tratamiento, del encargado del tratamiento y de las personas autorizadas a tratar los datos personales bajo la autoridad directa del responsable o del encargado.

**Tratamiento:** cualquier operación o conjunto de operaciones realizadas sobre datos personales o conjuntos de datos personales, sea por procedimientos automatizados o no, como la recogida, registro, organización, estructuración, conservación, adaptación o modificación, extracción, consulta, utilización, comunicación por transmisión, difusión o cualquier otra forma de habilitación de acceso, confrontación o interconexión, limitación, supresión o destrucción.

**Transparencia del proceso:** información sobre el propósito del sistema algorítmico y el proceso seguido en el desarrollo y aplicación de dicho sistema, así como los datos utilizados. Tiene que incluir información sobre las elecciones y suposiciones asumidas, los tipos de datos utilizados en el desarrollo del sistema, el método utilizado para identificar los riesgos, los riesgos identificados y las medidas adoptadas para mitigar dichos riesgos, así como las partes implicadas en el desarrollo del sistema algorítmico y sus funciones.

**Transparencia técnica:** información que permite que las partes afectadas entiendan cómo funciona el sistema algorítmico técnicamente, lo que incluye la divulgación del código fuente, las especificaciones técnicas utilizadas, los datos utilizados, la información técnica sobre la obtención y el procesamiento de los datos, la información sobre el método de desarrollo y el proceso de desarrollo utilizados, la elección de un modelo particular y sus parámetros, y la información sobre el rendimiento del sistema algorítmico.



## ANEXO 2. Modelo de solicitud

### **INFORMACIÓN GENERAL**

Área:

Definición del área \_\_\_\_\_ de “sistema de toma de decisiones automatizadas”:

Descripción de la necesidad y del sistema que se plantea implementar para solucionarla:

¿Por qué el sistema algorítmico es la manera de responder a esa necesidad?

¿Se han considerado otras alternativas? [SÍ/NO]

En caso de que se haya respondido SÍ, indíquese cuáles son esas alternativas:

### **DETALLES DEL SISTEMA QUE SE QUIERE LICITAR**

Finalidad del sistema algorítmico:

Alcance del sistema algorítmico:

Datos que se utilizarán:

Política de utilización:

Calendario de aplicación:

¿Qué riesgos de inexactitudes, sesgos o daños a comunidades específicas podrían aparecer?

¿Qué medidas de mitigación de riesgos se utilizarán?

### ANEXO 3. Aprendizajes

- Es necesario pensar en modelos en los que participen las propias personas usuarias, como propietarias de la futura herramienta, y en guías para el desarrollo del modelo con auditorías posteriores.
- Hay que evitar que los mecanismos descritos en este protocolo entorpezcan y hagan más pesados los procesos.
- Tendrá que perfilarse el presente protocolo en función de si el sistema que se licita ya ha sido desarrollado o si se desarrolla específicamente para el Ayuntamiento de Barcelona.
- Tendrá que definirse cómo se va a trabajar con soluciones y métodos propiedad de grandes empresas (Google, OpenAI, etc.).
- Se valorará la introducción de nuevos criterios en las licitaciones que, como sucede con la solvencia técnica y económica, introduzcan el concepto de *solvencia ética* de los proveedores.