Machine Translated by Google

Ajuntament de Barcelona

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

Machine Translated by Google

Ajuntament de Barcelona

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 5
1.4. Gobernanza 10
2. Concepción del servicio 11
2.1. Divulgación de información 11
2.2. Asignación del riesgo 11
2.3. El estudio de impacto algorítmico (EIA) 12
2.4. Espacios de transparencia 13
2.5. Protección de datos 14
3. Contratación 14
3.1. Calidad de los datos 15
3.2. Derechos sobre los datos 16
3.3. Calidad del sistema algorítmico 16
3.4. Transparencia del sistema algorítmico 17
3.5. Protección de datos 18
4. Desarrollo, producción y explotación 19
4.1. Estrategia de gestión de riesgos en el desarrollo del sistema algorítmico 19
4.2. Auditoría y otras tipologías de inspección 20
4.3. Gastos 20
4.4. Espacios de transparencia 20
4.5. Protección de datos 21
5. Desmantelamiento 21
5.1. Destrucción de los datos 21
ANEXO 1. Definiciones 23
ANEXO 2. Modelo de SOLICITUD 26

Machine Translated by Google

1. Introducción

1.1. Antecedentes y motivación

El uso de las tecnologías basadas en sistemas algorítmicos e inteligencia artificial1 ofrece oportunidades al Ayuntamiento de Barcelona para transformar los servicios municipales para que sean más eficientes, proactivos y adaptados a las necesidades de la ciudadanía, así como para agilizar los procesos internos y apoyar la toma de decisiones. Sin embargo, el despliegue de este tipo de sistemas lleva asociados múltiples riesgos2 , por lo que es imprescindible asegurar que su implementación introduce mecanismos de salvaguardia 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 estándares éticos y técnicos adecuados— y que el proceso favorece la participación y deliberación constantes con la ciudadanía y las comunidades impactadas.

Este documento es una de las acciones previstas por la "Medida de Gobierno por la Estrategia Municipal de Algoritmos y Datos para el Impulso Ético de la Inteligencia Artificial" publicada en 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 piloto3 de sistemas algorítmicos confiables. Se entienden como confiables todos aquellos sistemas que cumplan 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 [4] por el Ayuntamiento de Barcelona, independientemente de si éste 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 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 en cuanto a la protección de datos y (4) los estándares éticos definidos por diversas administraciones y organizaciones. El documento hace accionables los principios, recomendaciones y estándares definidos en la "Medida de Gobierno por la Estrategia Municipal de Algoritmos y Datos para el Impulso Ético de la Inteligencia Artificial", el Reglamento General de Protección de Datos (RGPD) y la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales (LOPDGDD), la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial, la Ley 15/2022 , integral por la igualdad de

[1] La inteligencia artificial utiliza conjuntamente la recolección y análisis de datos masivos (big data) con algoritmos y el entrenamiento de sistemas inteligentes (machine learning) para realizar predicciones o recomendaciones más o menos automatizadas. Para más definiciones, consulte el Anexo 1.

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

[3] El Ayuntamiento de Barcelona, a través de BIT Habitat, impulsa al testado de pilotos de soluciones urbanas innovadoras en fase precomercial para validar su utilidad y el interés público en el entorno urbano real antes de que éstas sean comercializadas . En el marco de este programa, el Ayuntamiento de Barcelona testea sistemas de IA aplicados a la mejora de servicios urbanos.

[4] Para "impulsar" entendemos aquellos sistemas algorítmicos de licitados por el Ayuntamiento de Barcelona, sean estas soluciones ya existentes o soluciones desarrolladas expresamente, o aquellos que sean testados en el marco del programa "Barcelona Innova Lab", de BIT Habitat (sobre esto, véase la página 6).

Machine Translated by Google

tratamiento y no discriminación, los de documentos elaborados por otras ciudades5, organizaciones y el marco regulatorio sobre inteligencia artificial que ha propuesto la Comisión Europea (AI Act)6 . Aunque la ley europea sobre inteligencia artificial todavía se encuentra en proceso de tramitación legislativa, este protocolo incorpora varios elementos presentes, como la clasificación de las aplicaciones de IA según el riesgo, para anticiparse y adaptarse de forma proactiva a los estándares que serán de obligado cumplimiento cuando la ley entre en vigor. Cuando las instituciones europeas aprueben el reglamento sobre IA, este protocolo se adaptará a su contenido final.

En cuanto a los sistemas algorítmicos que incluyan procesamiento de datos personales, este protocolo tiene en cuenta los tres criterios personales marcados por el RGPD de cara a la inteligencia artificial: Requerimiento de que el procesamiento sea justo (art. 5), Principio de minimización de datos (art. 5), Evaluaciones de impacto de la protección de datos (art. 35).

La regulación de actuaciones automatizadas municipales el Ayuntamiento de Barcelona tiene como antecedente el decreto de Comisión de Gobierno por el que 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 2022.7

De esta combinación surge un protocolo que define, paso a paso, los mecanismos de garantía y salvaguarda de derechos que deben introducirse en cada momento de la implementación de un sistema de inteligencia artificial por parte de Ayuntamiento de Barcelona, que 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 hecha por la Comisión Europea en el marco de la AI Act8 , y establece distintos 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 mecanismos en función del nivel de riesgo, como explicita este documento. Para los sistemas algorítmicos de alto riesgo -de forma 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.

Dada 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 el feedback de los actores 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 hechas por expertos.

[5] Véanse, por ejemplo, las Cláusulas tipo para la contratación de sistemas algorítmicos de Ámsterdam
[6] https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai
[7] Este decreto se circunscribe a las actuaciones administrativas automatizadas dentro de los procedimientos administrativos y tiene como finalidad identificar el uso de la tramitación electrónica automatizada en el procedimiento administrativo y establecer el protocolo para su autorización. Todo esto en cumplimiento de la normativa administrativa de aplicación y como desarrollo de lo establecido en la Ordenanza reguladora de la administración electrónica.
[8] Encontrará la clasificación de riesgos en la página 7.

Machine Translated by Google

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 por la Estrategia Municipal de Algoritmos y Datos para el Impulso Ético de la Inteligencia Artificial”, es necesario 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 salvaguardias 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:

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 por 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 la construcción de los mismos.

4. Producción: Despliegue del servicio, incluyendo la posible migración de datos y procesos desde uno o más servicios propios.

5. Explotación: Fase que se entiende durante toda la vida útil del sistema en producción, incluyendo las labores de operación y mantenimiento.

6. Desmantelamiento: Fase que incluye todas las acciones necesarias a realizar una vez terminada 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 servicios TIC corren a cargo del Instituto Municipal de Informática (IMI) del Ayuntamiento de Barcelona, organismo autónomo que tiene por objetivo suministrar la mayoría de los servicios TIC al Ayuntamiento ya los órganos y empresas públicas que dependen del mismo.

- El Ayuntamiento de Barcelona, como administración pública del ámbito local, se rige por un procedimientos administrativos regulados por ley. Éstos, 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:

Machine Translated by Google

1. Formulación del problema: En esta fase, es necesario definir claramente qué necesidad se quiere abordar con el sistema algorítmico y establecer qué significa que el modelo funcione bien. También es necesario definir los inputs o valores de entrada que se introducirán en el sistema y los outputs o resultados que se espera que se obtendrán9 .

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 puede entenderse como la fórmula matemática que combina los inputs con un conjunto de parámetros o coeficientes de ponderación ajustables para obtener el output predicho10. 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 debido a 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 puestos en funcionamiento.

En algunos casos, ciertas fases de este proceso se repiten de manera iterativa: puede que en algún momento del desarrollo del sistema se recojan nuevos datos y el modelo se reestructure, o que éste 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 deberá seguir cualquier área del Ayuntamiento de Barcelona que desee pujar un sistema algorítmico. En las figuras 1 y 2, se sintetiza gráficamente el proceso que se explica de forma detallada en las próximas páginas.

Hay que poner énfasis en que los sistemas algorítmicos que se liciten en el Ayuntamiento de Barcelona, en tanto que servicios TIC, potencialmente serán licitados por el IMI, que centralizará todas las fases que incluye el ciclo de vida de un sistema algorítmico , en colaboración estrecha con el área impulsora.

La mayoría de estos procesos sólo tienen sentido en situaciones donde el sistema algorítmico pueda tener un impacto negativo en la ciudadanía o en comunidades específicas, siguiendo la clasificación propuesta por la Comisión Europea en cuanto al riesgo de los sistemas de inteligencia artificial11. En ningún caso se trata de un protocolo de aplicación general a cualquier sistema algorítmico, sino que se desplegará de forma 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.

[9] 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 que estos outputs sirvan simplemente para apoyar una decisión humana o que la decisión esté plenamente automatizada, o que el sistema esté programado para tener tres posibles resultados: sí, no y “quizás”, el cual implicaría la revisión del resultado por parte de un humano.

[10] 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. https://

[11] digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

Machine Translated by Google

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

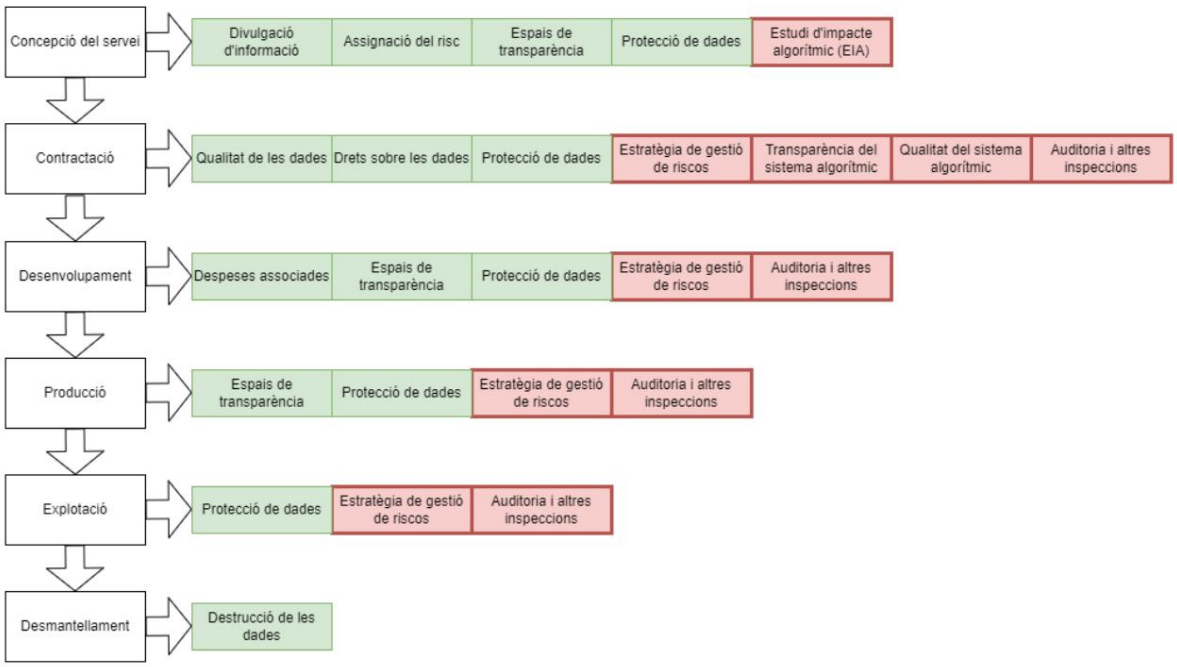
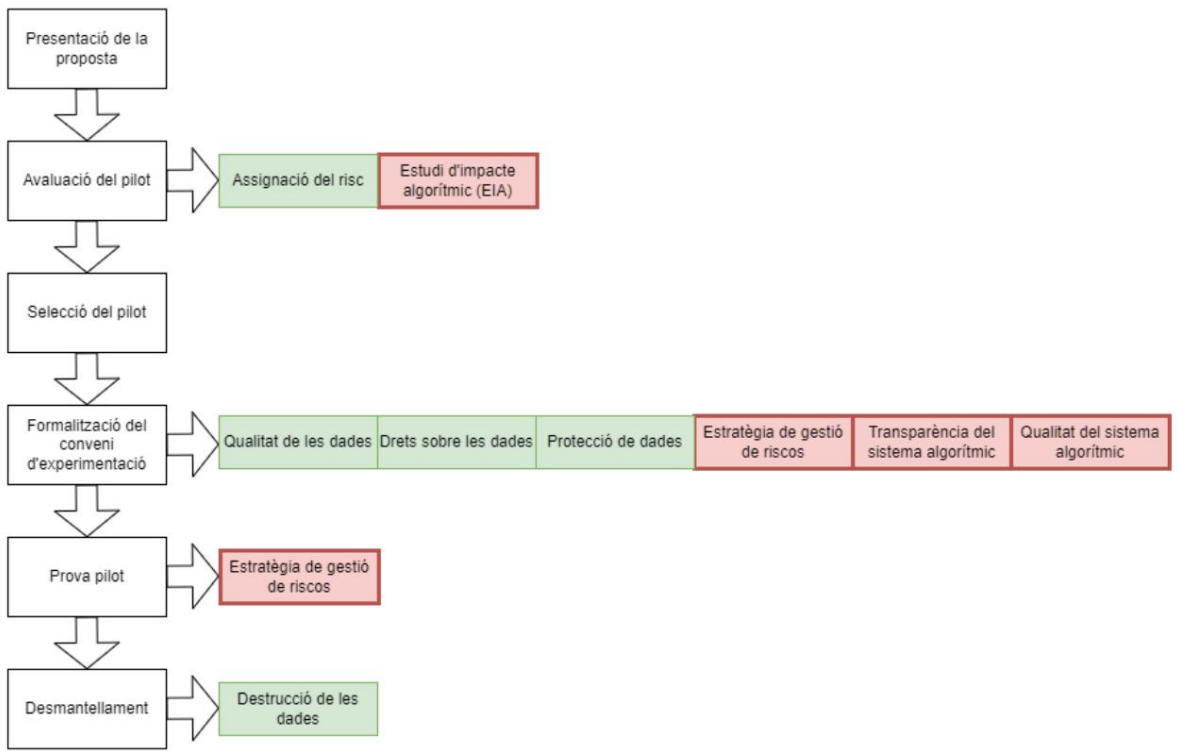


Fig. 2. Relación de las fases del despliegue de un PROYECTO PILOTO de un sistema algorítmico impulsado por el Ayuntamiento de Barcelona con los mecanismos definidos en este protocolo



Machine Translated by Google

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 programa del Barcelona Innova Lab impulsado por BIT Habitat y el IMI. En el marco de este programa, se impulsa el testado de pilotos de sistemas algorítmicos aplicados a servicios urbanos en fase precomercial para validar su utilidad e interés público en el entorno urbano real antes de que éstos 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 acceso al espacio urbano, los activos municipales y la conectividad 5G). En estos casos, los mecanismos de salvaguardia y garantía de derechos se introducen mediante dos mecanismos:

a) La admisión de los pilotos se lleva a cabo a través de una convocatoria con estrictos criterios 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 de IA ética.

b) La introducción de las cláusulas adecuadas al 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 apartado 3 de este documento) .

Machine Translated by Google

Ajuntament de Barcelona

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. Será la Oficina Técnica del IMI quien adjudicará la categoría de riesgo a las solicitudes de sistemas.

Estos riesgos se refieren al contexto de aplicación del sistema ya 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 en vigor sobre el tema.

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 utilizan asistencia de voz y que fomentan comportamientos peligrosos. Se prohibirán también los sistemas de identificación biométrica remota en espacios públicamente accesibles.

RIESGO ALTO

Estos incluyen la tecnología de IA utilizada en:

- 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
- la 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);
- el empleo, la gestión de los trabajadores y el acceso al autoempleo (por ejemplo, software de clasificación de CV para los procedimientos de contratación);
- servicios privados y públicos esenciales (por ejemplo, la calificación de créditos que niega a los ciudadanos la oportunidad de obtener un préstamo);
- aplicación de la ley que puede interferir con los derechos fundamentales de las personas (por ejemplo, evaluación de la fiabilidad de las pruebas);
- gestión de la migración, el asilo y el control de fronteras (por ejemplo, verificación de la autenticidad de los documentos de viaje);
- 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:

- 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 al usuario;
- medidas adecuadas de supervisión humana para minimizar el riesgo;
- 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 fines 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 un autor o sospechoso de un delito grave.

  Esta utilización está sujeta a la autorización de un órgano judicial u otro órgano independiente ya 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, los usuarios deben ser conscientes de que están interactuando con una máquina para poder tomar una decisión informada de continuar o retroceder.

RIESGO MÍNIMO O NINGÚN RIESGO

La propuesta permite el uso libre de la IA de riesgo mínimo. Esto incluye aplicaciones como los videojuegos con IA o los filtros de spam.
La gran mayoría de los sistemas de IA utilizados actualmente en la UE entran en esta categoría.

Machine Translated by Google

1.4. Gobernanza

En el marco de la “Medida de Gobierno por 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 entes dependientes relacionados con la implementación de proyectos de inteligencia artificial. Entre otros, tiene la función de validar los sistemas basados en inteligencia artificial de aplicación al Ayuntamiento propuesto por las áreas municipales y entes dependientes, 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 entes dependientes y velar por 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, a nivel técnico, los estándares, herramientas, uso e implementación de tecnologías de IA sobre los proyectos desarrollados/implantados en el Ayuntamiento de Barcelona; desarrollar proyectos piloto de IA; apoyar para conceptualizar las soluciones a nivel tecnológico, metodológico y ético y evaluar sus riesgos.

Asimismo, también en el marco del despliegue de su “Medida de Gobierno por 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 ejercitar una supervisión externa sobre los sistemas algorítmicos que se desarrollan. (Ver 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, de organizaciones de la sociedad civil y empresas tecnológicas y tiene, entre otras cosas, la función de evaluar externamente el impacto de los sistemas algorítmicos. Como se define más adelante, el Consejo Asesor emitirá preceptivos y no vinculantes Estudios de Impacto Algorítmico (EIA) para los algoritmos considerados de alto riesgo.

- Pacto de ciudad para el impulso del humanismo tecnológico. Este Pacto, cuyos actores firmantes se reúnen 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 problemáticas que vayan emergiendo en lo que se refiere a la gobernanza de las tecnologías digitales en la ciudad. En el marco de este Pacto se

Machine Translated by Google

producirán diálogos de carácter más general (y no tan técnico, como en el Consell), entre otras temáticas, 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 existe entre que se detecta la necesidad de desarrollar un sistema algorítmico en el Ayuntamiento de Barcelona y el desarrollo del mismo (ya sea a través de un proyecto piloto o de los procedimientos de contratación estándares), es clave para garantizar que su implementación se hace de forma transparente, segura, legítima y proporcionada y que se incluye a las personas afectadas ya 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 éste llenará una ficha de solicitud de servicio (en adelante, LA SOLICITUD, véase Anexo 2 para la ficha completa), en la que se detallará:

- Su definición interna de sistema de toma de decisiones automatizadas.
- La información sobre la necesidad que tiene, porque se considera que la necesidad podría ser resuelta con un sistema algorítmico, como éste podría responder a esa necesidad y si se han considerado alternativas a éste.
- 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 hubiere), política de utilización y calendario de aplicación.
- La definición preliminar de posibles riesgos de inexactitud, sesgos, o daños a comunidades específicas, definirá medidas de mitigación de riesgos.
- La definición de los mecanismos para permitir el acceso de los expertos externos que revisarán sistema una vez 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 2 semanas emitirá un informe que incluirá una extensión de la información proporcionada por el área municipal y una valoración del riesgo. De forma concreta, 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 estos riesgos, incluyendo diálogos proactivos que fomenten el compromiso con los miembros de estas comunidades.

Machine Translated by Google

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

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

- Riesgo inaceptable: LA SOLICITUD del área municipal será denegada
- Riesgo alto: La Oficina Técnica informará 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, que en un máximo de 3 semanas12 emitirá este estudio.
- Riesgo limitado y mínimo o ningún riesgo: No será necesaria ninguna EIA y el área municipal implicada podrá llevar adelante 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 el sistema algorítmico hace y quién es parte responsable en caso de que surja algún problema13. Se lleva a cabo en el momento prelicitación, antes de que el Ayuntamiento de Barcelona haya comprometido la utilización de un sistema algorítmico, para conocer los posibles resultados indeseables e inseguros que afectan a los derechos fundamentales de las personas y negociar e introducir las medidas necesarias para su corrección.

Los EIA son cruciales para dar a todas las partes interesadas la oportunidad de discutir y evaluar la adopción del sistema, asegurar su transparencia y garantizar el acceso de la ciudadanía a la información que le afecta, así como para evaluar la su idoneidad teniendo en cuenta el caso de uso en el 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 de buenas prácticas tanto para el Ayuntamiento de Barcelona como para otras administraciones locales. Este Protocolo se irá actualizando y mejorando con estas experiencias.

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

1. Respetar el derecho de la ciudadanía a saber qué sistemas afectan a sus vidas, enumerando y describiendo públicamente cómo son y de qué manera estos sistemas afectan significativamente a las 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.

12Se establece un tiempo de entrega del informe de tres semanas para evitar la ralentización del proceso de contratación.

13 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 los expertos científicos y políticos que se ha ido desarrollando en torno a la responsabilidad algorítmica.

14 https://ainowinstitute.org/aiareport2018.pdf

Machine Translated by Google

3. Garantizar la rendición de cuentas de los sistemas algorítmicos y facilitar que expertos externos revisen, auditen y evalúen estos sistemas utilizando métodos que les permitan identificar y detectar fallas 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 rendición de cuentas de un sistema algorítmico determinado.

De forma 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 durante 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 expertos externos al Ayuntamiento, reunidos en el Consejo Asesor en inteligencia artificial, Ética y Derechos Digitales, a fin de garantizar la máxima imparcialidad sobre los sistemas algorítmicos. Estos EIA serán consultables por la ciudadanía y cualquier actor interesado en el Registro Público de Algoritmos -junto con el resto de información pública sobre el sistema, como se definirá más adelante.

Cabe destacar que los EIA no son la solución para los problemas que plantean algunos sistemas algorítmicos, y estos van acompañados de los instrumentos necesarios para garantizar una implementación legítima, segura y transparente, como definen los siguientes capítulos de este documento. Sin embargo, son herramientas prácticas para introducir el debate sobre la utilización de estos sistemas, proporcionar información a la ciudadanía y para determinar su idoneidad.
Una vez licitado, el impacto del sistema algorítmico se determinará y revisará a través de las auditorías algorítmicas llevadas a cabo por una entidad externa en el 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: pueden incluir los sesgos existentes en las bases de datos (por falta de representatividad o por introducción de variables proxy, por ejemplo15); el proceso por el que llegan a una decisión puede ser opaco o poco transparente y

[15] Las variables proxy son variables aparentemente no relevantes, pero que actúan como una variable inobservable o inmedible. Como señala el AI Primero de Nueva York, ejemplo de ello es el código postal, que en EE.UU. a menudo se relaciona de forma directa con la racialización o el nivel de ingresos y por tanto, utilizarlo como variable puede conducir a decisiones sesgadas. Estas variables deben tratarse de forma muy cuidadosa para evitar comportamientos inacurados, inadecuados e incluso perjudiciales.

Machine Translated by Google

explicable, sobre todo en los sistemas que incorporan el aprendizaje automático o machine-learning -porque no se pueden explicar los pesos a los coeficientes que ponderan el resultado final a partir de los inputs dados-; a menudo se prioriza la optimización sin tener en cuenta a quien afecta; entre otros. Todo esto hace que el Ayuntamiento de Barcelona se encuentre ante la necesidad de velar por la legitimidad democrática, respeto de los derechos fundamentales, robustez técnica y seguridad, transparencia y explicabilidad de cualquier sistema algorítmico que se desarrolle a nivel municipal.

De forma concreta, 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 se quiere licitar (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 pre-licitación.
- Como prevé la instrucción en protección de datos, habrá una 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 su naturaleza y adecuación normativa, 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 personas16 .

Si el sistema incluye datos personales o puede inferirlos y pasa a la fase de contratación y explotación, deberá registrarse como nuevo tratamiento en el Registro de Actividades de Tratamiento (RAT) municipal, de acuerdo al 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 de las recomendaciones emitidas al EIA en los pliegos.
- Riesgo limitado y mínimo o ningún riesgo: No será necesario un EIA y el área municipal podrá tirar adelante la redacción de los pliegos.

[16] A modo de ejemplo, en la medida en que tratemos datos de carácter personal o indirectamente se pueda llegar a identificar personas, deberá identificarse una base legal de acuerdo al artículo 6 del RGPD, teniendo en cuenta que el interés legítimo no será de aplicación 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.

Machine Translated by Google

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 Ajuntament de Barcelona. Por eso 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 Fiables17 , en redacción por 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 sólo a los sistemas algorítmicos calificados como "riesgo alto". Al inicio de cada apartado, se indica en qué casos se aplicará la condición: "independiente del riesgo" significa que la condición se aplicará en cualquier sistema algorítmico que se desarrolle, y "alto riesgo" significa que la condición sólo 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, éstos pueden incorporar imprecisiones y errores construidos socialmente18. Las medidas que debe adoptar el contratista deben estar focalizadas al evitarlo. Por este motivo, es necesario
establecer unos requisitos de calidad de los datos.

Por otro lado, cabe diferenciar entre los casos en los que el contratista desarrolla un sistema algorítmico con los datos proporcionados por el Ayuntamiento de Barcelona del caso en que es el contratista quien desarrolla el sistema algorítmico a partir de los datos obtenidos a través de otros fuentes. El motivo es que, si el contratista recibe los datos del Ayuntamiento de Barcelona, se le puede considerar responsable del análisis, ordenación y edición de los datos, pero no de su recogida. Al fin y al cabo, en este caso, es el Ayuntamiento de Barcelona quien lo ha hecho y, por tanto, es lógico que la responsabilidad recaiga en el nombre del Ayuntamiento de Barcelona.

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 el edición de los datos. Este método debe estar basado, cuando sea posible, en pautas científicas y 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 las datos. Qué se considera sesgo deberá evaluarse en cada caso en concreto. En cualquier caso, se considera que existe sesgo cuando la utilización de los datos

[17] El documento que contiene las Cláusulas se encuentra en construcción, y está basado en la propuesta de Cláusulas estándares de contratación de sistemas de IA fiables elaboradas por el Ayuntamiento de Ámsterdam.

[18] Cabe recordar que la recolección y clasificación de los datos que entrenan y alimentan los sistemas algorítmicos se lleva a cabo por humanos, que pueden tener prejuicios -racistas o machistas, por ejemplo- o visiones sesgadas que pueden conducir 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 conducir a imprecisiones debido a la falta de información sobre colectivos minorizados que no quedan representados de forma significativa. Un sistema algorítmico entrenado sólo con datos de un colectivo mayoritario probablemente ofrecerá resultados sesgados cuando se aplique a otros colectivos con características diferentes.

Machine Translated by Google

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 Humanos19 .

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 modelos de gestión incorrectas. Todo sesgo demostrable y discriminatorio debe ser identificado y, siempre que sea posible, el contratista debe eliminarlo lo antes posible. Si no es posible eliminarlo, deberá advertir al Ayuntamiento de Barcelona lo antes 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, no el responsable de los datos, puesto que quien definirá la finalidad del tratamiento y los medios a utilizar es el Ayuntamiento. Esto será de aplicación tanto en 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, es necesario 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 por la materia, en que el adjudicatario resultante cumpla con la normativa de protección de datos20 . El Ayuntamiento de Barcelona establecerá acuerdos contractuales específicos con cada proveedor sobre los derechos u obligaciones relativos a los datos que se intercambian.

Como definen las Cláusulas Tipo para la Contratación de Sistemas de IA Fiables creadas por Amsterdam, de manera general se acordará que el contratista no tiene derecho a utilizar ni los datos proporcionados por el Ayuntamiento de Barcelona ni los datos generados durante el servicio para 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 los entregará al Ayuntamiento de Barcelona, como definen 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

[19] https://www.echr.coe.int/Documents/Convention_CAT.pdf

[20] Cabe recordar que no es posible poseer los datos en el sentido del derecho de propiedad, puesto 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 tanto, los datos no suelen estar 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. Sin embargo, el derecho sobre las bases de datos sólo ofrece protección frente a la reutilización de toda la base de datos o una parte sustancial de la misma, 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

Machine Translated by Google

Cuando en la fase prelicitación se haya establecido que un sistema algorítmico es de alto riesgo habrá que añadir todos los mecanismos necesarios a las Cláusulas Tipo para la Contratación de Sistemas de IA Fiables para asegurar que éste se desarrolla de forma segura, transparente , explicable, legítima, respetando los derechos fundamentales y que lo hace de acuerdo con el marco legal aplicable. Por tanto, será necesario evaluar qué regulación es aplicable a cada caso en concreto (pe: uso de tecnologías de anonimización y GDPR) y establecer la obligación de que el sistema algorítmico se desarrolle siguiendo un método fundamentado, basado, cuando sea posible, el 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 debe proporcionar el resultado por el que fue concebido. Esto significa, entre otras cosas, que el sistema algorítmico debe ser apto para su utilización prevista, y debe tener la capacidad de realizar decisiones correctas (eg, clasificar correctamente la información en las categorías adecuadas) o respecto a su capacidad realizar adecuadamente predicciones, formular recomendaciones o tomar decisiones correctas basadas en datos o modelos. Esto también implica que el sistema algorítmico debe ser seguro y fiable. Es necesario que el sistema algorítmico sea robusto desde el punto de vista técnico y el contratista debe 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

Como señala un informe de Algorithm Watch21, la transparencia no asegura, de por sí, conformidad plena con los requerimientos éticos, pero es una condición necesaria para alcanzar esta 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 le afectan.

El Ayuntamiento de Barcelona diferencia entre transparencia del procedimiento, transparencia técnica y explicabilidad, según los distintos 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 aplicación del mismo y de los datos utilizados en este contexto. En cualquier caso, la información suministrada debe ser suficiente para garantizar que, si surgen daños, se puede rastrear su causa, así como para garantizar la comprensión de las opciones y supuestos asumidos y las categorías de datos utilizadas en el desarrollo del sistema algorítmico. Debe incluir también la forma en que se da 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

[21] Algorithm Watch (Junio 2021) “Automated Decision-Making Systems in the Public Sector: Public Authorities”, Michele Loi, Anna Mätzener, Angela Müller, Matthias Spielkamp. Disponible online.

Machine Translated by Google

permita comprender el funcionamiento técnico del sistema algorítmico. Esta información debe 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 del 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 deben realizarse en los datos de entrada de cara a obtener un resultado diferente. Hacer que un sistema algorítmico sea explicable incluye proporcionar toda la información técnica y otra tipología necesaria para explicar en los espacios de participación o en procedimientos como los estudios de impacto, las auditorías algorítmicas o en procedimientos de impugnación, de recurso o en otros procedimientos judiciales, como se ha llegado a una decisión y ofrecer a la otra o 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 considera necesaria para la evaluación del sistema algorítmico. Además, la transparencia en el procedimiento se centra también en proporcionar a la ciudadanía y otras partes interesadas información general sobre la utilización del sistema algorítmico.

En cuanto a la explicabilidad, aunque de forma general se obligará al contratista a proveer la información detallada sobre los factores clave que han llevado a un sistema algorítmico a un resultado particular o de los cambios a realizar en los datos de entrada para obtener resultados distintos, podría darse el caso de que en una situación específica esto 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 por objetivo que el Ayuntamiento de Barcelona, en las auditorías anuales correspondientes (encargadas a una entidad tercera), pueda verificar si el contratista debe cumplido con sus obligaciones. Ésta puede incluir información comercial confidencial y es por ello 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 en el Ayuntamiento de Barcelona, se estipulará en el contrato que el contratista deberá 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, se tendrán que aplicar los criterios de privacidad desde el diseño y por defecto que marca el RGPD (art. 25), incluidos

Machine Translated by Google

aquellas funcionalidades que sirvan para salvaguardar los derechos ARSOLP 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 salga 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 la RAT de la Sede Electrónica, se deberá 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 poner en funcionamiento el sistema algorítmico. En estas fases del ciclo de vida (desarrollo, producción y explotación) será necesario garantizar que el sistema funciona según su objetivo inicial y que los riesgos que emergen son mitigados. En caso de que el sistema incorpore el aprendizaje automático, habrá que velar para que las operaciones llevadas a cabo por éste sigan siendo explicables y el sistema no incorpore sesgos.

Las buenas prácticas y aprendizajes que emerjan durante esta fase alimentarán este protocolo y quedarán reflejadas en este documento vivo para mejorar el proceso de cara a los 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 pujas de los sistemas algorítmicos del Ayuntamiento de Barcelona incluirán el diseño y la ejecución de una estrategia de gestión de riesgos sistemática, que deberá ir siendo actualizada. Se propone seguir la lista de comprobación propuesta en el documento “Directrices Éticas por una IA confiable” del grupo de expertos de alto nivel sobre IA facilitado por la Comisión
Europea22 .

La estrategia de gestión de riesgos que se siga debe 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 deberá ponerse a disposición del Ajuntament de Barcelona mediante la solicitud de transparencia del procedimiento. Además, el Ayuntamiento de Barcelona evaluará, mediante auditorías anuales, si se ha llevado a cabo una estrategia de gestión de riesgos correcta.

[22] Independiente High-Level Expert Group on Artificial Intelligence (2019). Ethics Guidelines For Trustworthy Ai, p. 26-31. Disponible online: https://www.aepd.es/sites/default/files/2019-12/ai-ethics-guidelines.pdf

Machine Translated by Google

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

4.2. Auditoría y otras tipologías de inspección

Obligatorio: RIESGO ALTO

El Ayuntamiento de Barcelona debe poder comprobar si el contratista cumple con los requerimientos especificados en el contrato y, a tal fin, obligará al contratista a colaborar en estas comprobaciones ya proporcionar todos los datos necesarios para que una entidad externa lleve a una auditoría algorítmica. Cada sistema algorítmico que se licite deberá 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 subsanar los incumplimientos identificados (y descritos en el informe) en un plazo razonable (también fijado por el Ayuntamiento de Barcelona). Si el contratista no subsane 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.

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 estos efectos. De forma preliminar y no restrictiva[23], se prevé que se publiquen:

- Información básica sobre el sistema algorítmico: •
  - Finalidad del sistema algorítmico
  - Área municipal que lo impulsa
  - Actores implicados en las decisiones del sistema
  - Alcance del sistema algorítmico
  - Datos que se utilizarán

[23] 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 nivel europeo, el Ayuntamiento de Barcelona trabaja de forma constante con otras administraciones, organizaciones internacionales, la Comisión Europea y expertos independientes para mantener al día y consolidar unos estándares comunes e interoperables, entre otros temas, sobre el tipo de información que debe ponerse al alcance de la ciudadanía para garantizar una IA ética, segura, robusta y confiable.

Machine Translated by Google

- Incorporación de aprendizaje automático
- Política de utilización
- Calendario de aplicación
- Riesgos que podrían aparecer y estrategia de gestión de éstos
- 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 (sólo 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 al funcionamiento de los sistemas algorítmicos

Aparte del Registro de Algoritmos, aquellos sistemas que incorporen tratamiento de datos personales deberán incorporarse al RAT de la Sede Electrónica.

Aparte de la información pública disponible en el Registro de Algoritmos, el Ayuntamiento de Barcelona facilitará la participación de la sociedad civil, empresas y entidades a través del Pacto de Ciudad por el Humanismo Tecnológico, dando la oportunidad a todas las partes implicadas debatir la aplicación de los sistemas algorítmicos y conocer sus resultados. Además, el Ayuntamiento de Barcelona contará con los expertos del Consejo Asesor en inteligencia artificial, Ética y Derechos Digitales para realizar un seguimiento detallado de los sistemas de IA que se encuentren en funcionamiento. De manera concreta, se añadirá un punto al orden del día de las reuniones de ambos órganos en los que el área municipal impulsora del sistema algorítmico y el IMI tendrán que informar sobre el desarrollo del sistema algorítmico y los miembros de los órganos podrán hacer preguntas sobre éste y debatirlos.

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, se tendrán que aplicar 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 ARSOLP 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, debería tenerse muy claro el principio de accountability 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

Machine Translated by Google

Cuando finalice el contrato con el proveedor, éste 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. Por el resto de datos (no de carácter personal), se fomentará la reutilización de los mismos, estableciendo acuerdos ad hoc con cada proveedor.

En cuanto al código fuente, en caso de que el Ayuntamiento de Barcelona no haya adquirido la propiedad intelectual, el proveedor podrá venderlo a otros clientes y seguirlo utilizando, siempre que éste 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 mismo, 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 en nivel técnico 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.

Machine Translated by Google

ANEXO 1. Definiciones

Algoritmo, en un sentido muy amplio: desde modelos de regresión y árboles de decisión que pueden realizar 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 que no haya intervenido de forma 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 sobre otra de forma diferente a la función prevista del algoritmo.

Datos personales: toda información sobre una persona física identificada o identificable («el interesado»); se considerará persona física identificable 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 online o uno o diversos 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 que se comuniquen datos personales, se trate o no de un tercero. Sin embargo, no se considerarán destinatarios las autoridades públicas que puedan recibir datos personales en el marco de una búsqueda concreta de conformidad con el derecho de la Unión o de los Estados miembros.
El tratamiento de tales datos por estas autoridades públicas será 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 al rendimiento profesional, situación económica, salud, preferencias personales, intereses, fiabilidad, comportamiento, ubicación o movimientos de esa 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/explicable: Poder explicar a nivel individual por qué un sistema algorítmico conduce a una decisión o resultado en particular. Salvo que las Partes estén expresamente

Machine Translated by Google

de acuerdo con lo contrario, esto incluirá una indicación clara de los factores clave que han llevado a un sistema algorítmico a un resultado particular y los cambios que deben realizarse para que el sistema conduzca a un resultado diferente. Hacer un sistema algorítmico explicable incluye la provisión de toda la información técnica para explicar, en los procedimientos de objeción, apelación u otros procedimientos judiciales, cómo se ha producido una decisión y ofrecer a las partes afectadas oa cualquier otra parte interesada la oportunidad de evaluar la forma en que se ha adoptado una decisión, con la finalidad única de ofrecer una protección jurídica realista.

Inteligencia artificial: disciplina científica que incluye diversos 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, la programación, la representación y el razonamiento de conocimientos, la búsqueda y la optimización), y la robótica (que incluye el control, la percepción, los sensores y accionadores), así como la integración de todas las demás técnicas en sistemas “ciberfísicos”. Según la definición del Consejo de Europa: “La IA se entiende 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 marcado 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, sólo 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, el responsable del tratamiento o los criterios específicos para su nombramiento podrá establecer su 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 dado a partir de unos datos y tome una decisión en función de ese análisis.

Tercero: persona física o jurídica, autoridad pública, servicio u organismo distinto del interesado, del responsable del tratamiento, del encargado del tratamiento y de las personas autorizadas para tratar los datos personales bajo la autoridad directa del responsable o de el encargado.

Tratamiento: cualquier operación o conjunto de operaciones realizadas sobre datos personales o conjuntos de datos personales, ya 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, cotejo 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 del mismo y los datos utilizados. Se incluirá 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 los riesgos, así como las partes implicadas en el desarrollo del sistema algorítmico y sus funciones.

Machine Translated by Google

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

Machine Translated by Google

Ajuntament de Barcelona

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 forma de responder a esta necesidad?

¿Se han considerado alternativas a éste? [SÍ/NO]

En caso de que se haya respondido SÍ, indique qué 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 inexactitud, sesgos o daños a comunidades específicas podrían aparecer: ¿Qué medidas de mitigación de riesgos se utilizarán?:

Machine Translated by Google

ANEXO 3. Aprendizajes

- Cabe pensar en modelos en los que participen los propios usuarios, como propietarios de la futura herramienta, y en guías para el propio desarrollo del modelo con auditorías posteriores.

- Evitar que los mecanismos descritos en este protocolo entorpezcan y hagan más pesados los procesos.
- Habrá que afinar este protocolo en función de si el sistema que se licita ya ha sido desarrollado o si éste se desarrolla específicamente para el Ayuntamiento de Barcelona.

- Habrá que definir cómo se trabajará con soluciones y métodos propiedad de grandes empresas (Google, OpenAI, etc).
- Valorar la introducción de nuevos criterios en las licitaciones que, tal y como ocurre con la solvencia técnica y económica, introduzcan el concepto de “solvencia ética” de los proveedores.