

BANCO DE INNOVACIÓN EN LAS ADMINISTRACIONES PÚBLICAS

TÍTULO

El Ayuntamiento de Barcelona aprueba un protocolo que garantiza un uso ético de la inteligencia artificial en sus servicios públicos digitales

INFORMACIÓN INICIAL:

PROBLEMA:

El rápido desarrollo e implementación de la inteligencia artificial (IA) y de técnicas algorítmicas está empujando a las organizaciones a establecer estándares para un empleo seguro, eficaz y solvente. Ante la celeridad e imprevisibilidad de estos cambios y sus consecuencias, algunas Administraciones y organismos internacionales han destacado la necesidad de establecer reglas y armonizar estándares. Especialmente, sobre aquellas herramientas basadas en algoritmos e inteligencia artificial que pueden afectar a derechos de los ciudadanos, a su esfera personal o al funcionamiento de las propias Administraciones públicas.

Reveladores son, a este respecto, los casos de empleo de la IA por parte del gobierno de Reino Unido y de Países Bajos, que han vulnerado derechos de la ciudadanía. Así, en plena pandemia, y ante la imposibilidad de hacer las pruebas de acceso a la universidad de forma presencial, el gobierno de Reino Unido, decidió las notas a través de un sistema de IA. En este caso, se comprobó que los estudiantes con buen expediente veían rebajada su nota por estar estudiando en centros donde la media del centro era baja. Por otra parte, en Países Bajos, un tribunal prohibió el uso de un sistema de IA predictiva que identificaba a personas beneficiarias de prestaciones sociales que eran supuestamente más proclives a cometer fraude (lo que generaba discriminaciones y estigmatización de perfiles sociales concretos).

Estos episodios evidencian los riesgos inherentes al empleo de la IA en los servicios públicos. Una de las principales preocupaciones deriva de los sesgos de los propios sistemas de toma de decisiones basados en la IA. En este sentido, se pueden encontrar sesgos involuntarios o propios de las limitaciones técnicas de estas tecnologías, así como de los derivados del diseño, entrenamiento y desarrollo de los sistemas algorítmicos; siendo crítico incluso el mismo lenguaje con el que se entrena y diseña los algoritmos.

“El aprendizaje y entrenamiento automático es tan propenso al sesgo como la cognición humana”, según manifestaba Julio Antonio Gonzalo Arroyo, catedrático de la ETSI de Informática por la UNED, del Departamento de Procesamiento del Lenguaje Natural y Recuperación de Información, en un [seminario](#) sobre inteligencia artificial organizado por la UNED. Así mismo, este experto considera que los distintos aspectos de la evaluación de sistemas que pueden ser fuentes de sesgo como las métricas de evaluación, los datos de prueba, el diseño experimental y las herramientas de análisis de resultados. En su charla revisó los distintos aspectos de la evaluación de sistemas que pueden ser fuentes de sesgo: las métricas de evaluación, los datos de prueba, el diseño experimental y las herramientas de análisis de resultados.

De ahí que las instituciones y otras entidades hayan advertido de la imprescindible necesidad de que los sistemas algoritmos que se emplean en el proceso de toma de decisiones sean justos, de tal forma que no se vean condicionadas por sesgos de raza o género, entre otros, y que normalmente vienen de los propios datos (véase, por ejemplo, la afirmación de Antoine Bordes, el responsable de Inteligencia Artificial de la red social Facebook que en declaraciones al diario [ABC](#) señalaba que para la inteligencia artificial es muy difícil entender el contexto).

SOLUCIÓN GLOBAL:

Con el fin de anticiparse, generar un espacio seguro y garantizar un correcto despliegue de estas novedosas herramientas, el Ayuntamiento de Barcelona ha lanzado el protocolo [Definición de metodologías de trabajo y protocolos para la implementación de sistemas algorítmicos](#) para la implementación de sistemas algorítmicos en el ámbito municipal.

Se trata de un documento pionero en cuanto a la regulación y gobernanza de los sistemas algorítmicos en el ámbito de la Administración local, de tal forma que la aplicación de herramientas basadas en inteligencia artificial para la prestación de servicios públicos se adecuó a los estándares legales, éticos y técnicos definidos, a fin de que su acción sea proporcionada, supervisada y fundamentada.

COSTE APROXIMADO

El coste de esta política es indeterminado, puesto que habrá de valorarse *a posteriori*, en tanto en cuanto debe evaluarse la corrección de los estándares, los daños o riesgos evitados y el nivel de desarrollo e implementación

de los sistemas algorítmicos en los servicios municipales del Ayuntamiento de Barcelona, así como su impacto en la prestación de los mismos.

TERRITORIO:

Municipio de Barcelona (Barcelona, Cataluña).

PÚBLICO DESTINATARIO:

Los vecinos y usuarios de la localidad de Barcelona, así como también los que en un futuro adquieran la condición de vecinos o usuarios de los servicios públicos de la localidad.

ENTIDAD QUE LA HA LLEVADO A CABO:

Ayuntamiento de Barcelona.

DESCRIPCIÓN DE LA POLÍTICA O PROGRAMA:

La Revolución Industrial 4.0 está propiciando la aparición de nuevas herramientas y aplicaciones para el empleo por parte de las personas y las organizaciones. Si bien, a diferencia de revoluciones anteriores, esta está creciendo a un ritmo exponencial. Un pilar fundamental de la misma es la inteligencia artificial (IA), que combina el *big data* (recolección y el análisis de datos masivos) y la utilización de algoritmos y el entrenamiento de sistemas inteligentes (*machine learning*), para la toma de decisiones en el seno de un procedimiento u organización.

El uso de las tecnologías basadas en sistemas algorítmicos e inteligencia artificial está ofreciendo enormes oportunidades para transformar el funcionamiento de organizaciones, coadyuvar en procedimientos y cadenas de valor y, en suma, mejorar procesos y productos, de tal forma que sean más eficientes, ágiles y comprensivos.

Son varias las organizaciones públicas que están apostando por la incorporación de herramientas de inteligencia artificial en su organización y en la prestación de servicios públicos. Podemos citar, por ejemplo, el *chatbot Clara*, puesto en marcha por el Ayuntamiento de Madrid, orientado a la participación ciudadana y la tramitación de consultas reales en cualquier momento y de forma inmediata. En línea con las innovaciones digitales, el consistorio de la capital también ha puesto en marcha el proyecto piloto “AirTech” de medición de

calidad del aire en la ciudad. Por su parte, el Ayuntamiento de Fuenlabrada ha impulsado el proyecto piloto Konekta2, que emplea inteligencia artificial en los programas de atención domiciliaria de mayores; y, en la misma línea en Benidorm se ha lanzado durante 2022 un proyecto de teleasistencia.

El Ayuntamiento de Barcelona es otra de las administraciones españolas a la vanguardia de este proceso innovador. De hecho, Barcelona se encuentra entre las entidades públicas europeas más innovadoras en la implementación de la inteligencia artificial para la prestación de servicios. A este respecto, podemos destacar la puesta en marcha de diversos proyectos piloto por parte del ayuntamiento que emplean sistemas inteligentes para hacer los servicios municipales más eficientes y proactivos; destacando el sistema de monitorización de la afluencia en playas, el sistema IRIS de apoyo para la clasificación de quejas y sugerencias de la ciudadanía y, en servicios sociales, la herramienta piloto de IA entrenada que se emplea en el proceso de toma de decisiones, coadyuvando al técnico en relación con los recursos que deben destinarse a cada persona solicitante de ayuda.

Así, siguiendo la hoja de ruta establecida en la *Estrategia municipal de algoritmos y datos para el impulso ético de la inteligencia artificial*, aprobada en abril de 2021, el Ayuntamiento ha presentado a principios de año el protocolo titulado [Definición de metodologías de trabajo y protocolos para la implementación de sistemas algorítmicos](#), con el fin de suplir la inexistencia de un instrumento interno que permitiera ordenar estos esfuerzos y sentar las bases para poder escalar el uso de la IA en la administración municipal y la provisión de servicios públicos por parte del Ayuntamiento de Barcelona, más si cabe cuando la propia administración debe ser consciente de que los sistemas de IA no están exentos de riesgos (como ya hemos visto, al mencionar los casos de Reino Unido y Países Bajos).

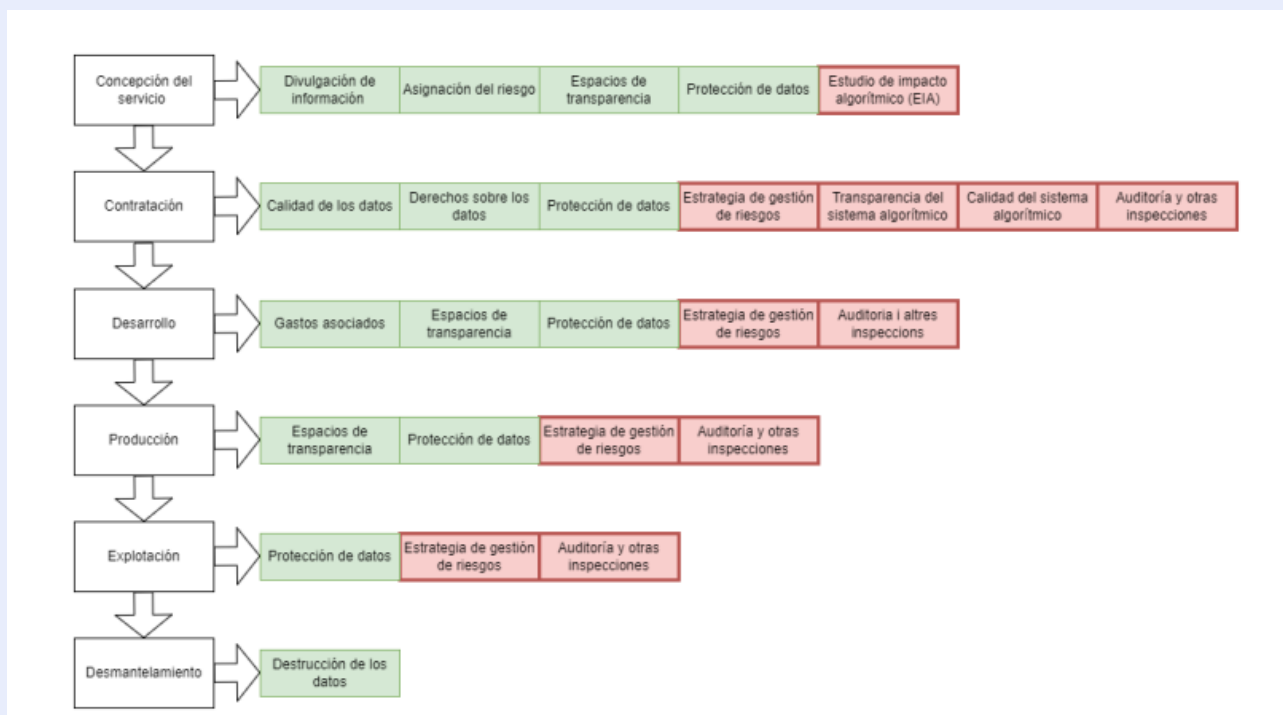
Este protocolo se posiciona como «un documento pionero con respecto a la regulación y gobernanza de los sistemas algorítmicos en el ámbito local, ya que es el primero en combinar el proceso de contratación pública, el ciclo de vida de un sistema algorítmico, los mecanismos que tiene el Ayuntamiento con respecto a la protección de datos y los estándares éticos definidos por varias administraciones y organizaciones».

Resulta aplicable a cualquier sistema algorítmico impulsado por el Ayuntamiento de Barcelona y tiene como objetivo garantizar que estos sistemas se utilizan de manera proporcionada, supervisada y fundamentada, de acuerdo con estándares legales, éticos y técnicos.

Con respecto al ciclo de vida de un algoritmo, el protocolo identifica cuatro fases:

- i) Formulación del problema y necesidad a abordar;
- ii) Recogida y procesamiento de los datos;
- iii) Selección y construcción del modelo;
- iv) Desarrollo y monitorización del sistema algoritmo.

Infografía 1: Fases de desarrollo un sistema algorítmico licitado por el Ayuntamiento de Barcelona con los mecanismos definidos en este protocolo



Fuente: Ayuntamiento de Barcelona.

A este respecto, el protocolo contempla los mecanismos de salvaguarda de derechos que tiene que haber en cada etapa de la licitación e implementación de un sistema algorítmico por parte del Ayuntamiento de Barcelona, y establece los órganos de gobernanza y supervisión que velarán por que el impacto de la IA se alinee con los principios éticos.

Del mismo, el protocolo del Ayuntamiento de Barcelona se hace eco de las propuestas de la Comisión Europea relativas a la medición de riesgos de los sistemas algorítmicos. De esta forma, la clasificación de riesgos de cualquier sistema de IA licitado o impulsado por el Ayuntamiento de Barcelona se efectúa en los siguientes niveles:

- a) Riesgo inaceptable. Se prohibirán todos los sistemas de IA que se consideren una clara amenaza para la seguridad y los derechos de las personas. Se impedirán sistemas de identificación biométrica remota en espacios públicamente accesibles. Uno de los ejemplos más discutidos al respecto, es el empleo del reconocimiento facial en espacios públicos, al que se le unen las desviaciones por sesgos de discriminación u otros.
- b) Riesgo alto. Aquí se incluye la tecnología de IA utilizada en casos como: 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); 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; información clara y adecuada a cada persona usuaria; etc.

- c) 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.
- d) Riesgo mínimo. 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.

Por último, debemos hacer referencia a la gobernanza de esta política de innovación. El propio protocolo del Ayuntamiento de Barcelona establece los órganos encargados de la gobernanza y de velar por el correcto funcionamiento de los sistemas algorítmicos en último término.

Por un lado, encontramos a la Comisión Transversal para el Impulso de la IA Ética en el Ayuntamiento de Barcelona que, entre otras, tiene la función de validar los sistemas basados en inteligencia artificial que se apliquen en el Ayuntamiento y, a su vez, 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.

Por otro lado, encontramos la Oficina Técnica de IA del Institut Municipal d'Informàtica (IMI), encargada 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.

A estos dos, hay que añadir la existencia de un tercero: el Consejo Asesor en Inteligencia Artificial, Ética y Derechos Digitales, que se instituye como un órgano de consulta y colaboración, que emitirá los informes de impacto algorítmico y tiene la función de evaluar externamente el impacto de los sistemas algorítmicos.

OBSTÁCULOS SUPERADOS

La implementación de nuevas herramientas tecnológicas viene acompañada de una serie de obstáculos que, en este caso, se ven reforzados. El principal es el grado de digitalización de las Administraciones públicas, en general, no del todo consolidado en nuestro país. Este factor, ligado enormemente a la falta de competencias digitales profesionales por parte de los empleados públicos y, en menor medida, también la resistencia al cambio dentro de la cultura organizativa, son determinantes en el despliegue de sistemas algorítmicos.

En este sentido, también es importante hacer referencia a las dificultades asociadas a la alineación o compatibilidad de algunos de los servicios prestados con su monitorización e implementación de sistemas algorítmicos; más si cabe, si situamos en el marco de las exigencias singulares que la prestación pública de servicios exige (entre otros, equidad, universalidad, sostenibilidad, sometimiento al derecho administrativo), así como las relativas a la información personal a tratar y los deberes impuestos a la Administración tratante.

IMPACTO

Dada la recentísima publicación del protocolo (enero de 2023), aún es prematuro evaluar el impacto en el despliegue e implementación de sistemas algorítmicos en el seno de los servicios públicos municipales del Ayuntamiento de Barcelona.

No obstante, sí podemos referirnos a las cuestiones clave de una futura y eventual revisión o evaluación del impacto. En tal sentido, deberían tenerse en cuenta factores tales como: el acceso a los datos, así como su tratamiento; el grado de observancia de los principios éticos y legales establecidos; el número y gravedad de las violaciones de derechos y, en su caso, la gravedad de las mismas; la efectividad de los mecanismos de garantía en cada una de las fases del ciclo de vida y aplicación del sistema algorítmico; y, la mejora en la prestación de los servicios públicos en los que se ha implementado una herramienta de IA o algoritmo (eficacia, eficiencia, calidad del servicio, satisfacción del usuario o ciudadano, etc.).

CALENDARIO DE IMPLANTACIÓN Y REFERENCIA TEMPORAL

El diseño y elaboración del protocolo *Definición de metodologías de trabajo y protocolos para la implementación de sistemas algorítmicos* se ha efectuado dentro del marco temporal previsto en la Estrategia municipal de algoritmos y datos para el impulso ético de la inteligencia artificial (2021-2023). A partir de ahora, su aplicación comenzará tanto con los sistemas algorítmicos que licite el Ayuntamiento como los que se diseñen y produzcan en el seno de la entidad pública.

DOCUMENTACIÓN DE CONSULTA Y APOYO:

- [Protocolo “Definición de metodologías de trabajo y protocolos para la implementación de sistemas algorítmicos”](#) (2023). Ayuntamiento de Barcelona.
- [Estrategia municipal de algoritmos y datos para el impulso ético de la inteligencia artificial](#) (2021). Ayuntamiento de Barcelona.
- [Estrategia de inteligencia artificial de Cataluña](#). (2020). Departamento de Transformación Digital, Generalitat de Cataluña.
- [Directrices éticas para una IA fiable](#) (2019). Grupo de Expertos de Alto Nivel sobre Inteligencia Artificial, Comisión Europea.

- Barcelona Declaration for the Proper Development and Usage of Artificial Intelligence in Europe [AIdeclaration | \(csic.es\)](#)
- Universidad Nacional de Educación a Distancia. [Julio Gonzalo, catedrático de la UNED, señala que en IA se habla poco del sesgo producido por mecanismos de evaluación inadecuada \(2020\).](#)