



IA urbana ética en la práctica

Mecanismos de política para establecer marcos de gobernanza local

Marta Galceran-Vercher y Alexandra Vidal D'oleo (eds.)

CIDOB

BARCELONA
CENTRE FOR
INTERNATIONAL
AFFAIRS

CIUDADES GLOBALES



Ajuntament de
Barcelona

IA urbana ética en la práctica

Mecanismos de política para establecer marcos de gobernanza local

Marta Galceran-Vercher y Alexandra Vidal D'oleo (eds.)

CIDOB
BARCELONA
CENTRE FOR
INTERNATIONAL
AFFAIRS



© 2024 CIDOB

CIDOB edicions
Elisabets, 12
08001 Barcelona
Tel.: 933 026 495
www.cidob.org
cidob@cidob.org

Impresión: QP Print Global Services
ISBN: 978-84-18977-28-2
Depósito legal: B 1974-2025

Barcelona, diciembre 2024

Foto de la cubierta:
Vehículos en una vía rodeada de edificios altos. 10th Ave.
<https://unsplash.com/photos/vehicles-on-pathway-surrounded-by-high-rise-buildings-3UnFTjhsWWA>

SUMARIO

SOBRE LOS AUTORES	5
RESÚMENES PARTE I	9
INTRODUCCIÓN. IA URBANA ÉTICA EN LA PRÁCTICA	11
Tanya Álvarez, Marta Galceran-Vercher y Alexandra Vidal D'oleo	
PARTE I. LA OPERACIONALIZACIÓN DE PRINCIPIOS ÉTICOS EN ENTORNOS URBANOS	17
Shazade Jameson	
Rendición de cuentas y transparencia en IA urbana	19
Josuan Eguiluz Castañeira y Carlos Fernández Hernández	
Privacidad y gobernanza de datos en IA urbana	29
Leandry Junior Jieutsa	
Equidad y no discriminación en IA urbana	41
María Pérez-Ortiz	
Sostenibilidad en IA urbana	51
PARTE II. CASOS PRÁCTICOS DE MARCOS DE GOBERNANZA DE IA URBANA	63
Alexandra Vidal D'oleo	
Caso práctico 1: Barcelona	66
Caso práctico 2: Ámsterdam	68
Caso práctico 3: Nueva York	71
Caso práctico 4: San José	73
Caso práctico 5: Dubái	75
Caso práctico 6: Singapur	76
CONCLUSIONES. MECANISMOS DE POLÍTICA, DESAFÍOS Y RECOMENDACIONES EN IA URBANA	79
Marta Galceran-Vercher y Alexandra Vidal D'oleo	

SOBRE LOS AUTORES

Tanya Álvarez

Es investigadora del Observatorio de la Mobile World Capital Foundation. Dirige investigaciones sobre la inclusión digital y el uso de sistemas automatizados de toma de decisiones en el sector público. Defiende una perspectiva interdisciplinar del impacto de la tecnología en la sociedad. Es licenciada en Historia del Arte por el Swarthmore College y tiene un máster en Gestión del Patrimonio Cultural por la Universidad de Barcelona.

Josuan Eguiluz Castañeira

Es licenciado en Derecho y TIC por la Universidad de Deusto y posee un doble máster en Acceso a la Abogacía y en Propiedad Intelectual, Nuevas Tecnologías y Protección de Datos por Esade. Inició su carrera profesional como abogado en el Departamento de Propiedad Intelectual, Nuevas Tecnologías y Protección de Datos de Cuatrecasas. En la actualidad, ejerce como asesor jurídico en Adevinta. También es profesor de Derecho Digital en Esade, Deusto y The Legal School, entre otras instituciones. Está realizando un Doctorado Industrial en IA, Áreas de Alto Riesgo y Derechos Fundamentales en la Universidad Pompeu Fabra, Barcelona Supercomputing Center e Infojobs/Adevinta.

Carlos Fernández Hernández

Es licenciado en Derecho y actualmente doctorando en la Universidad Carlos III de Madrid. Tras ejercer como abogado durante varios años, ha dedicado su carrera profesional a la publicación de contenidos jurídicos para profesionales del sector, siempre bajo el sello La Ley. Desde 2013, se ha especializado en el ámbito de las relaciones entre derecho y tecnología, inicialmente en el portal de información jurídica Noticias Jurídicas y, desde 2016, en Diario La Ley Ciberderecho, donde ha publicado más de 1.500 artículos, reportajes y entrevistas. Carlos también ha sido editor de las revistas Derecho Digital e Innovación y La Ley Privacidad. Ha impartido conferencias especializadas sobre estas materias en diversas universidades y centros de postgrado y colaborado en varios libros.

Marta Galceran-Vercher

Es investigadora principal del Programa de Ciudades Globales de CIDOB (Barcelona Centre for International Affairs) y profesora asociada de Relaciones Internacionales en la Universidad Pompeu Fabra (UPF) y en el CEI- Centro de Estudios Internacionales. Es politóloga, doctora en Relaciones Internacionales por la UPF y tiene un máster (con mención honorífica y premio extraordinario) en Relaciones Internacionales por la University of Warwick. Su investigación se centra en la diplomacia

de las ciudades, las alianzas urbanas y la gobernanza global, particularmente en el campo de las transiciones digitales y el humanismo tecnológico. Es autora de varios documentos de política, artículos académicos y otras publicaciones sobre estos temas. También lidera la investigación del Observatorio Global de Inteligencia Artificial Urbana (GOUAI), en el marco de la Cities Coalition for Digital Rights. Antes de unirse a CIDOB, fue la coordinadora del Programa del Smart City Expo World Congress y consultora sénior en Anteverti, donde asesoró a gobiernos locales y organizaciones internacionales (por ejemplo, ONU-Habitat, BID, Comisión Europea) en políticas de internacionalización e innovación urbana.

Shazade Jameson

Es consultora sénior de gobernanza digital para organizaciones de interés público. Es especialista en administraciones públicas urbanas. Como investigadora en ciencias sociales, propone nuevas perspectivas en materia de política digital y tiende puentes para una aplicación eficaz. Fue coautora principal del informe de ONU-Hábitat y Mila Quebec AI Institute «AI and Cities: Risks, Applications and Governance». También es doctoranda en el análisis de la política tecnológica en Singapur, como parte del proyecto Global Data Justice financiado por el Consejo Europeo de Investigación en el Instituto de Derecho, Tecnología y Sociedad de Tilburg. Ha colaborado con CIDOB en la investigación de políticas para el Observatorio Global de Inteligencia Artificial Urbana (GOUAI) y trabaja con la iniciativa United for Smart Sustainable Cities, de la Unión Internacional de Telecomunicaciones, en el marco de la temática Transformación Digital para Ciudades Centradas en las Personas. Desde 2024, trabaja con la Unidad de Ética de la IA de la UNESCO, facilitando el diseño y la implementación de la inteligencia artificial ética en las administraciones públicas.

Leandry Junior Jieutsa

Es consultor en innovación urbana e investigador sobre la gobernanza de la IA en las ciudades. Su carrera ha sido un viaje a través de diversas funciones en la planificación y el desarrollo urbanos, subrayadas por su pasión por la integración de las tecnologías en la planificación urbana. Trabaja como consultor independiente en materia de innovación urbana y gobernanza digital. Leandry es también fundador de Africa Innovation Network, un grupo de reflexión dedicado a cuestiones urbanas. Actualmente está enriqueciendo sus conocimientos realizando un doctorado en la Cátedra UNESCO de Paisaje Urbano de la Universidad de Montreal, centrado en la gobernanza de la inteligencia artificial en las ciudades. Su investigación estudia el impacto de la gobernanza de la IA en el bienestar de las personas en las ciudades. A través de su trabajo, aspira a ayudar a las ciudades a crear urbes inteligentes más centradas en las personas mediante una gobernanza responsable de la IA.

María Pérez Ortiz

Es profesora asociada del AI Centre y del Departamento de Informática del University College de Londres (Reino Unido). También trabaja como investigadora principal en el Ministerio de Asuntos Exteriores, de la Commonwealth y de Desarrollo del Gobierno británico, asesorando sobre inversiones en IA para el desarrollo internacional. En 2022, cofundó el primer programa de máster sobre IA para el desarrollo sostenible, en la intersección de tecnologías emergentes de IA, sostenibilidad y ética. Pérez Ortiz es también adjunta de la Cátedra UNESCO de IA. Su último informe político «Challenging systematic prejudices: an investigation into bias against women and girls in large language models» muestra hasta qué punto los modelos lingüísticos presentan sesgos de género. Su actual línea de trabajo es la IA responsable y cómo estas novedosas tecnologías podrían ayudar a los responsables políticos en escenarios complejos como el cambio climático.

Alexandra Vidal D'oleo

Es investigadora y gestora de proyectos del Programa de Ciudades Globales. Su investigación gira en torno a la digitalización de las ciudades y los procesos democráticos y éticos en el desarrollo de la inteligencia artificial urbana. Actualmente, es parte del equipo de investigación del Observatorio Global de Inteligencia Artificial Urbana (GOUAI) en el marco de Cities Coalition for Digital Rights. Tiene un postgrado en Planificación Territorial con foco en urbanismo sostenible por la Universidad Politécnica de Cataluña y es licenciada en Relaciones Internacionales por Blanquerna – Universidad Ramón Llull, habiendo desarrollado parte de sus estudios en la École de Gouvernance et d'Économie en Rabat. Antes de unirse a CIDOB, fue investigadora y gestora de proyectos en la Fundación Tanja, desarrollando proyectos de cooperación internacional entre Marruecos y España.

Rendición de cuentas y transparencia en IA urbana

Shazade Jameson

Muchas administraciones públicas urbanas ven el potencial que ofrece la implementación de inteligencia artificial (IA), pero no se sienten preparadas para aplicarla de forma responsable. Aunque los enfoques y los marcos basados en una IA responsable son cada vez más frecuentes, muchos de ellos se centran en actores del sector privado o empresarial. Si bien el papel de la IA responsable en el sector público recibe cada vez más atención, existen muchas menos directrices específicas para los gobiernos locales. El objetivo de este capítulo es reducir esta brecha presentando definiciones de rendición de cuentas y transparencia que incluyan tanto perspectivas técnicas acotadas como sociopolíticas más amplias. Para facilitar la rendición de cuentas y la transparencia en el contexto de la aplicación de la IA por parte de las administraciones públicas urbanas, hay dos preguntas orientativas aparentemente sencillas pero fundamentales para la fase de diseño: «¿Debe utilizarse la IA?» y «¿Cómo debe utilizarse la IA?». Tras reflexionar sobre el significado de estas preguntas para los profesionales urbanos, el capítulo presenta un resumen de los mecanismos de política existentes que pueden adaptarse para trabajar en pos de estos objetivos y ofrece algunas lecciones aprendidas de investigaciones anteriores.

Privacidad y gobernanza de datos en IA urbana

Josuan Eguiluz Castañeira y Carlos Fernández Hernández

El desarrollo de la inteligencia artificial (IA) ética en Europa, tal como prevé la Ley de IA de la Unión Europea, debe incluir mecanismos sólidos para la protección de la privacidad y la gestión de los datos. En el contexto de los entornos urbanos públicos, el tratamiento de datos personales a través de sistemas de IA presenta retos específicos que las autoridades públicas deberán abordar cuidadosamente. A la luz de este nuevo marco legislativo, el objetivo de este artículo es, en primer lugar, presentar el marco jurídico y ético que regula el tratamiento de datos personales a través de sistemas de IA en entornos urbanos; en segundo lugar, esbozar mecanismos clave para aplicar el principio de privacidad; y en tercer lugar, examinar los retos asociados a estas prácticas de tratamiento de datos, proporcionando una serie de recomendaciones y buenas prácticas.

Equidad y no discriminación en IA urbana

Leandry Junior Jieutsa

La inteligencia artificial (IA) es una tecnología emergente, disruptiva y ambivalente. En el contexto de su implementación, las ciudades han de crear diversos mecanismos para garantizar que esta tecnología tenga el menor impacto negativo posible en las personas y las comunidades. El objetivo es asegurar que las ciudades sigan siendo espacios equitativos que no dejen a nadie atrás. Este capítulo, subdividido en tres apartados, plantea recomendaciones políticas para integrar aspectos de equidad y no discriminación en la implementación de la IA por parte de las ciudades. En el primer apartado se analizan las nociones de equidad y no discriminación en los entornos urbanos y se presentan los factores determinantes para que la IA sea equitativa y no discriminatoria. En el segundo apartado se exploran las oportunidades y repercusiones de la IA en las ciudades. Por último, en el tercer apartado se formulan recomendaciones políticas para lograr un mayor grado de equidad en la IA aplicada a las ciudades. Estas recomendaciones tienen en cuenta las diferentes funciones que pueden desempeñar las ciudades en la implementación de la IA, como desarrolladoras de soluciones internas, implementadores y reguladoras. Las ciudades deben ser flexibles, basándose en la participación, los enfoques locales, la innovación socio-técnica, la colaboración, etcétera.

Sostenibilidad en IA urbana

María Pérez-Ortiz

El capítulo explora el potencial de la inteligencia artificial (IA) para apoyar el desarrollo de ciudades sostenibles, abordando las dimensiones social, medioambiental y económica de la sostenibilidad. A medida que la urbanización se acelera en todo el mundo, las ciudades se enfrentan a retos cada vez mayores en ámbitos como la movilidad, la vivienda, la contaminación y la gestión de recursos. La IA es prometedora con respecto a la optimización de las infraestructuras urbanas, la reducción de las emisiones de gases de efecto invernadero y la mejora de la eficiencia de los recursos; sin embargo, su implementación también plantea problemas de equidad social, impacto ambiental y perturbación económica. La IA sostenible se propone como marco para que el desarrollo y la aplicación de la IA se ajusten a los objetivos de sostenibilidad, garantizando que esta funcione dentro de los límites ecológicos, promueva la inclusión y apoye un crecimiento económico equitativo y circular. Entre las principales áreas en las que se hace hincapié se encuentran la reducción de la huella de carbono de la IA mediante prácticas energéticamente eficientes, la integración de la equidad en los sistemas urbanos basados en IA y la garantía de una gobernanza transparente. El documento ofrece recomendaciones políticas para guiar la implementación de la IA en entornos urbanos, subrayando la importancia de la colaboración internacional, la gobernanza ética y las políticas económicas para fomentar la resiliencia y la inclusión en las ciudades del mañana.

INTRODUCCIÓN. IA URBANA ÉTICA EN LA PRÁCTICA

Tanya Álvarez

Investigadora, Observatorio de la Mobile World Capital Foundation

Marta Galceran-Vercher

Investigadora sénior, Programa de Ciudades Globales, CIDOB

Alexandra Vidal D'oleo

Investigadora y gestora de proyecto, Programa de Ciudades Globales, CIDOB

La inteligencia artificial (IA) es una de las tecnologías más revolucionarias de la actualidad y promete transformar por completo la sociedad. Esta transformación se produce en distintos niveles y ámbitos, y el espacio urbano no es una excepción. Desde que se concibiera el paradigma de la «ciudad inteligente», urbanistas, empresas tecnológicas y actores municipales recurren cada vez más a los avances tecnológicos para resolver los desafíos urbanos más apremiantes que afrontan nuestras sociedades. En este proceso, el uso de sistemas algorítmicos por parte de los gobiernos locales está muy extendida y da forma al proceso de “hacer ciudad” tal y como lo entendemos.

En este marco, la «IA urbana» es entendida como la relación entre los sistemas de IA y el medio urbano. Tales sistemas, junto con otras tecnologías, se integran en todo tipo de contextos urbanos, como los hogares, los lugares de trabajo, los espacios públicos o las infraestructuras. La digitalización de estas experiencias urbanas crea, además, un entorno híbrido en el que las tecnologías digitales desempeñan un papel de mediación y mejora de la experiencia urbana (Aurigi y De Cindio, 2008). Quienes viven en las ciudades apenas están empezando a percibir el profundo efecto que la IA, como elemento integrado en los entornos urbanos, tiene en la vivencia urbana y en el propio hacer ciudad.

En términos de IA y automatización, las ciudades son un terreno de experimentación idóneo para implementar estas tecnologías. El desarrollo y uso de la IA precisa distintos recursos que se encuentran con facilidad en las urbes: 1) un entorno físico sobre el que actuar, 2) acceso a actividades diversas, 3) abundancia de datos de calidad, y 4) infraestructuras e instalaciones (Cugurullo *et al.*, 2023). En la última década se ha producido un aumento repentino de tecnologías basadas en datos que abordan desafíos urbanos como el mantenimiento de las infraestructuras, la personalización de servicios públicos, la salud, la mejora del transporte, la planificación urbana y el uso eficiente de los recursos, entre otros (Galceran Vercher y Vidal, 2024).

Quienes viven en las ciudades apenas están empezando a percibir el profundo efecto que la IA, como elemento integrado en los entornos urbanos, tiene en la vivencia urbana y en el propio hacer ciudad.

Conforme las tecnologías algorítmicas se generalizan, existe una necesidad imperiosa de que las administraciones locales velen por el uso ético y responsable de estos sistemas.

Cada vez es mayor el número de gobiernos municipales conscientes de las ventajas que la IA aporta a la labor de administración y la prestación de servicios. Los sistemas de IA se adoptan con la esperanza de que reduzcan el trabajo rutinario, automatizando las tareas burocráticas y permitiendo así que los gobiernos locales funcionen con mayor eficiencia. Otro objetivo es que resulten más eficaces al tomar decisiones más acertadas basándose en datos y permitir que los gobiernos locales dispongan de más tiempo para atender mejor a las necesidades de la ciudadanía.

Sin embargo, conforme las tecnologías algorítmicas se generalizan, existe una necesidad imperiosa de que las administraciones locales velen por el uso ético y responsable de estos sistemas. Cabe destacar que la creciente relevancia que los gobiernos locales tienen en el terreno de la política mundial, así como su influencia en millones de vidas, requiere que la gobernanza de la IA local tome en cuenta el impacto en las personas, las comunidades y el medioambiente.

Con la llegada de la regulación de la IA, la gobernanza de estas tecnologías en las ciudades se ha convertido en un motivo especial de preocupación para las personas defensoras de los derechos humanos, las organizaciones de la sociedad civil y las minorías urbanas, que vienen siendo testigos de los posibles peligros que conlleva la implementación de estos sistemas tecnológicos. Por ejemplo, aunque la vigilancia mediante IA ofrece a las ciudades soluciones en relación con la seguridad, la gestión del tráfico o la monitorización de factores ambientales, se ha demostrado que estas medidas resultan invasivas y discriminatorias para determinados sectores de la población. Este ejemplo concreto, entre otros, alerta sobre cómo el aumento de eficacia derivado de la automatización puede tener un gran costo y, por tanto, como los gobiernos municipales deben conocer las implicaciones éticas de los sistemas de IA que pretenden establecer.

El desafío de operacionalizar principios éticos en la IA urbana

Aunque las ciudades puedan estar focalizadas en los beneficios técnicos y operativos que promete la IA, los especialistas sostienen que, al ser sistemas sociotécnicos, las repercusiones de estas tecnologías van más allá de su precisión técnica. En consecuencia, los responsables políticos y administradores locales que sólo se centran en dicha precisión técnica o imparcialidad del sistema no abordan plenamente las implicaciones más amplias que estos sistemas pueden tener a la hora de asegurar una implementación responsable y ética. Aplicar una IA responsable es algo más que desarrollar sistemas cuyos resultados sean correctos o fiables. La IA ética y responsable hace hincapié en la importancia de la moralidad durante todo el ciclo de vida de los sistemas, y garantiza que las herramientas algorítmicas se ajusten a los valores democráticos y protejan los derechos digitales de las personas.

El hecho de que muchas ciudades ya hayan empezado a aplicar y desarrollar mecanismos de política en materia de IA responsable es encomiable. Nueva York, por ejemplo, ha establecido la obligatoriedad de auditar herramientas de contratación de personal que usen IA. En Finlandia, tres ciudades se han reunido para impulsar transparencia mediante los registros de algoritmos, que posibilitan el acceso de la ciudadanía a información

pertinente. En otro caso, la policía de Toronto ha adoptado una política de adquisición pública para sistemas de IA. Además, una investigación reciente del [Observatorio Global de la Inteligencia Artificial Urbana \(GOUAI, por sus siglas en inglés\)](#) muestra que, en todo el mundo, las ciudades están impulsando sistemáticamente mecanismos de política para incentivar sistemas de IA responsable, como la elaboración de principios y directrices específicos, las prohibiciones o moratorias de determinados sistemas algorítmicos de alto riesgo (por ejemplo, los sistemas de reconocimiento facial en tiempo real), los registros públicos de algoritmos, las auditorías y evaluaciones del impacto, el establecimiento de órganos de supervisión externos e independientes o las cláusulas de contratación pública que garantizan el cumplimiento de los derechos humanos. Estas experiencias pueden servir de hoja de ruta para que otros actores del sector público conozcan qué tipo de mecanismos funcionan para desarrollar sistemas de IA responsable.

No obstante, si bien hay ciudades que están dando los primeros pasos para establecer prácticas de IA responsables, existe una creciente necesidad de que los responsables municipales conozcan cómo se está desarrollando la IA urbana y qué buenas prácticas adoptar. El Observatorio Global de la Inteligencia Artificial Urbana antes mencionado es un proyecto colectivo dirigido por CIDOB —con el apoyo las ciudades de Barcelona, Ámsterdam y Londres, la [Coalición de Ciudades por los Derechos Digitales](#) y ONU-Hábitat— que responde a esta necesidad. El resultado es el Atlas del GOUAI, en el que se recopilan casos de IA urbana de todo el mundo que se ajustan a seis principios éticos: transparencia y apertura, protección de la privacidad, equidad y no discriminación, seguridad y ciberseguridad, rendición de cuentas, y sostenibilidad. Una publicación reciente basada en el análisis del Atlas (Galceran Vercher y Vidal, 2024) puso de manifiesto que la creciente expansión de IA urbana va acompañada de un desequilibrio entre ciudades que han adoptado herramientas algorítmicas y las que han establecido políticas o estrategias para garantizar que estas tecnologías se alineen con principios éticos.

En este contexto, la presente monografía CIDOB profundiza en marcos de gobernanza existentes y mecanismos de política específicos para operacionalizar principios éticos concretos y así impulsar una IA urbana responsable en la práctica. El objetivo es generar un documento útil que pueda inspirar la acción de otros actores del sector público y servirles de hoja de ruta.

Estructura de la publicación

La primera parte de esta publicación consta de cuatro capítulos que examinan la aplicación práctica de principios éticos clave —concretamente 1) rendición de cuentas y transparencia, 2) privacidad y gobernanza de datos, 3) equidad y no discriminación, y 4) sostenibilidad— en entornos urbanos mediante mecanismos de política específicos. **Shazade Jameson** sostiene que los gobiernos locales carecen de directrices claras para promover una IA urbana ética en el marco de sus competencias y presenta dos definiciones prácticas de “rendición de cuentas” y “transparencia” que incorporan puntos de vista tanto técnicos como sociopolíticos. Jameson afirma que, para favorecer estos principios, la fase de concepción de la IA debe guiarse por dos preguntas aparentemente sencillas pero indispensables: «¿Debe utilizarse la IA?» y, en caso afirmativo, «¿Cómo debe utilizarse la IA?».

La IA ética y responsable hace hincapié en la importancia de la moralidad durante todo el ciclo de vida de los sistemas, y garantiza que las herramientas algorítmicas se ajusten a los valores democráticos y protejan los derechos digitales de las personas.

Una investigación reciente del Observatorio Global de la Inteligencia Artificial Urbana (GOUAI) muestra que, en todo el mundo, las ciudades están impulsando sistemáticamente mecanismos de política para incentivar sistemas de IA responsable.

En el capítulo siguiente, **Leandry Junior Jieutsa** examina los factores que contribuyen a una IA equitativa y no discriminatoria. En concreto, identifica dos factores principales de discriminación en los sistemas de IA: los sesgos algorítmicos y el uso de estas tecnologías. Jieutsa ofrece recomendaciones políticas dirigidas a que las ciudades que usan IA sean más justas, y hace hincapié en que los gobiernos locales deben adaptarse a los distintos roles que desempeñan como desarrolladores, implementadores y reguladores. El autor sostiene que las ciudades deben servirse de procesos participativos, enfoques contextualizados, innovación sociotécnica y colaboración intersectorial para garantizar que la IA se implanta de forma responsable y equitativa.

A continuación, **Josuan Eguiluz Castañeira** y **Carlos Fernández Hernández** revisan los mecanismos que permiten una gestión robusta de la privacidad y de los datos en el despliegue de la IA. Eguiluz y Fernández analizan los marcos éticos y jurídicos que rigen el procesamiento de datos personales por parte de los sistemas de IA, centrándose en particular en la Ley de IA de la UE. Además, esbozan los mecanismos clave para aplicar implementar el principio de privacidad en entornos urbanos y exploran los desafíos relacionados con estas prácticas de tratamiento de datos, ofreciendo un conjunto de recomendaciones prácticas. Los autores subrayan que la gobernanza de los datos debe constituir la base de las estrategias de IA urbana, primando la calidad, la pertinencia y la protección de los conjuntos de datos utilizados en estos sistemas. Esto incluye realizar evaluaciones del impacto para salvaguardar tanto los datos personales como los derechos fundamentales, garantizando que la privacidad y la seguridad de la ciudadanía no se vean comprometidas.

Por último, en el capítulo de **María Pérez Ortiz** se explora el potencial que tiene la IA para propiciar el desarrollo de ciudades sostenibles, y se abordan las dimensiones social, ambiental y económica de la sostenibilidad. La autora defiende que, si bien es una tecnología muy prometedora, su implantación suscita inquietud en cuanto a equidad social, impactos medioambientales y disrupciones económicas. En este contexto, el marco de la IA sostenible constituye una valiosa herramienta para alinear el desarrollo y aplicación de esta tecnología con los objetivos de sostenibilidad. A su vez, también garantiza que la IA funcione dentro de los límites ecológicos, fomente la inclusión y favorezca un crecimiento económico circular y justo.

En la segunda parte de la publicación se presentan seis casos prácticos que brindan ejemplos de marcos de gobernanza de IA locales que han establecido ciudades de todo el mundo, mediante la adopción de mecanismos de política concretos, para llevar a la práctica una IA urbana ética. En concreto, **Alexandra Vidal D'oleo** analiza la gobernanza de la IA de las ciudades de Barcelona, Ámsterdam, Nueva York, San José, Singapur y Dubái.

Esta monografía CIDOB concluye con un capítulo final en el que **Marta Galceran-Vercher** y **Alexandra Vidal D'oleo** exponen una clasificación de los mecanismos de política extraídos de los capítulos, los casos prácticos y la revisión bibliográfica. En el análisis se examinan los mecanismos de política más utilizados y se estudia su alineamiento con los distintos principios éticos. Las autoras también identifican tendencias y desafíos comunes que presentan las ciudades a la hora de llevar tales principios éticos a la práctica, y recogen un conjunto de recomendaciones generales para abordarlos.

Referencias bibliográficas

Aurigi, A. y De Cindio, F. «Augmented Urban Spaces: Articulating the Physical and Electronic City». Farnham: Ashgate Publishing, 2008

Cugurullo F. *et al.* «Artificial Intelligence and the City. Urbanistic perspectives on AI». Londres: Routledge, 2023

Galceran-Vercher, M. y Vidal, A. «Mapeo de la Inteligencia Artificial urbana: primer informe del Atlas de la Inteligencia Urbana del GOUAI». *CIDOB Briefings*, n.º 56, 2024

PARTE I. LA OPERACIONALIZACIÓN DE PRINCIPIOS ÉTICOS EN ENTORNOS URBANOS

- RENDICIÓN DE CUENTAS Y TRANSPARENCIA
EN IA URBANA

Shazade Jameson

- PRIVACIDAD Y GOBERNANZA DE DATOS
EN IA URBANA

*Josuan Eguiluz Castañeira, Carlos Fernández
Hernández*

- EQUIDAD Y NO DISCRIMINACIÓN EN IA URBANA

Leandry Junior Jieutsa

- SOSTENIBILIDAD EN IA URBANA

María Pérez-Ortiz

Shazade Jameson

Consultora experta en gobernanza digital

1. Introducción

En esta monografía, por «Inteligencia artificial (IA) urbana ética» se entiende aplicar enfoques de IA responsable en las administraciones públicas urbanas. Por consiguiente, todo debate sobre IA responsable ha de ajustarse a las necesidades y situaciones particulares de las administraciones públicas urbanas y su electorado.

Las administraciones públicas urbanas forman un contexto particular: son guardianas del interés público y actúan mayoritariamente en el plano local. Esto crea un entorno especialmente interesante a la vez que complejo debido a que se encuentran, por un lado, cerca de las complejidades locales y, por el otro, alejadas de las estrategias nacionales. Además, existe una asombrosa diversidad en cuanto a dimensión y capacidades entre las administraciones.

Esto significa que, aunque las administraciones públicas urbanas pueden aprovechar muchas ideas de los enfoques de la «IA ética» y la «IA responsable», la adaptación de estos puede ser limitada por requerir de una perspectiva más amplia de lo que permiten los recursos disponibles. Muchos enfoques de «IA responsable» corresponden al ámbito de la gobernanza empresarial y están orientados a ese contexto: ¿cómo pueden las empresas utilizar la IA para los productos y servicios que ofrecen, y emplearla de forma responsable? Las administraciones públicas urbanas, en cambio, tienen un modelo de gestión diferente, presuntamente dando prioridad al interés público.

Cada vez se presta más atención al papel que los enfoques de IA responsable desempeñan en el sector público (véase, por ejemplo, OCDE, 2024). Sin embargo, las orientaciones para los gobiernos locales en específico son mucho más escasas, sobre todo desde un punto de vista global. El presente capítulo tiene por objeto reducir esta brecha planteando definiciones de rendición de cuentas y transparencia, ubicando estos principios en el marco de la aplicación de IA por parte de administraciones públicas urbanas y, por último, presentando un resumen de los mecanismos de política existentes que pueden adaptarse con el fin de alcanzar estas metas.

2. Principios de rendición de cuentas y transparencia

Cada vez se presta más atención al papel que los enfoques de IA responsable desempeñan en el sector público. Sin embargo, las orientaciones para los gobiernos locales en específico son mucho más escasas, sobre todo desde un punto de vista global.

2.1. Rendición de cuentas

La rendición de cuentas es un concepto que acepta definiciones tanto amplias como acotadas. Es importante que los gobiernos locales, por su posición como organismos públicos, tomen en cuenta ambos tipos de descripciones.

En su forma más elemental, la rendición de cuentas es una forma de relación. La teoría más aceptada en la administración pública (Bovens, 2007) afirma que la rendición de cuentas es una *relación* entre un actor y un foro, y que ese foro tiene la facultad de rechazarla. La rendición de cuentas debe especificar *respecto a qué* y *ante quién* se lleva a cabo. Como forma de relación, la rendición de cuentas es un proceso social que requiere de un compromiso social y una visión social compartida (Wieringa, 2020).

El *respecto a qué* de la rendición de cuentas suele determinarse mediante las normas sustantivas y de procedimiento de la administración pública y la capacidad para evaluar si se han cumplido esas normas. El *ante quién* de la rendición de cuentas es sumamente importante para los gobiernos locales, y puede estar constituido por diversos públicos. ¿Quién tiene la obligación de rendir cuentas, respecto a qué y ante quién? ¿La entidad financiadora? ¿La parte interesada? ¿La ciudadanía afectada? Puesto que las administraciones públicas urbanas deben considerar el interés público, el conjunto de partes interesadas y de actores que deben rendir cuentas es mucho más amplio (Jameson *et al.*, 2021). Algunos casos de uso de IA pueden tocar también cuestiones de rendición de cuentas política, como el caso del escándalo de las prestaciones para el cuidado infantil que llevó a la dimisión del Gobierno neerlandés (Dachwitz, 2022; Amaro, 2021).

Cuando los gobiernos locales diseñan casos de uso de IA de cara al público, es importante que empiecen a trabajar con las comunidades afectadas en la etapa de concepción del proyecto (por ejemplo, ONU-Hábitat y Mila-Quebec AI Institute, 2022). Algunos marcos de IA responsable tienen un alcance limitado y pueden no estar preparados para satisfacer las exigencias de los procesos participativos de mayor amplitud que se requieren en la administración pública. En concreto, el modo en que prejuicios y desigualdades se codifican en los algoritmos como forma de gobernanza sugiere que es necesario incluir nuevas formas de impugnación y respuesta en la reestructuración organizativa en torno a la gobernanza de la IA (Taylor, 2021).

2.2. Transparencia

La transparencia, en lo que respecta a la IA, es un principio que presenta distintas capas. Al igual que la rendición de cuentas, se trata de un mecanismo de gran tradición en la administración pública, así como en la ingeniería de *software* y la informática.

Desde el punto de vista técnico, la transparencia consiste en revelar la información relacionada con un sistema algorítmico a lo largo de su ciclo de vida. La transparencia, en estos niveles técnicos, permite llevar a cabo una labor independiente de investigación y auditoría sobre el uso y la calidad de los modelos. Esto incluye: la finalidad del diseño, las fuentes de datos, los requisitos de *hardware*, las condiciones de funcionamiento, los resultados esperados del sistema y, clave en el caso de los sistemas algorítmicos, la relación entre las variables del modelo y la arquitectura, así como las características de los datos con los que se ha entrenado el modelo. La transparencia exige documentar el proceso de selección de los conjuntos de datos, las variables y los indicadores de calidad necesarios para el desarrollo de los sistemas.

La procedencia de los datos (es decir, su origen) y la calidad de los datos de entrenamiento son factores esenciales a la hora de implementar IA en las administraciones públicas. Constituyen una limitación considerable para la calidad de los modelos algorítmicos y es la principal fuente de sesgo en la implementación de estos sistemas en las administraciones públicas (ONU-Hábitat y Mila-Quebec AI Institute, 2022; Longpre *et al.*, 2023).

La transparencia es un principio general en el ámbito de la IA explicable, que incluye las nociones de explicabilidad e interpretabilidad. Estos conceptos adquirieron popularidad rápidamente como mecanismos de transparencia y rendición de cuentas tanto a nivel técnico como sociopolítico. El propósito general de la IA explicable es abrir la «caja negra» de los algoritmos opacos que no revelan la esencia de su funcionamiento interno (Adadia y Berrada, 2018).

Existen diferentes enfoques para abordar la explicabilidad desde un punto de vista técnico y, a grandes rasgos, se dividen en cuatro categorías (Wierenga, 2020). La primera consiste en explicar el modelo, por ejemplo ofreciendo indicaciones claras sobre los procedimientos que siguen y comprobando si un modelo algorítmico se puede explicar en lenguaje sencillo a una persona que no es experta en el tema. La segunda consiste en explicar el resultado, lo que significa describir en detalle las decisiones concretas que toman los algoritmos y si los mecanismos de esa toma de decisiones pueden comprenderse y evaluarse o no. La tercera consiste en examinar la caja negra, lo que puede incluir técnicas diversas tales como la visualización del funcionamiento interno del algoritmo. Por último, crear una caja transparente es un principio de diseño en el que se utilizan predictores visibles y explícitos. En general, la dificultad que presenta la transparencia desde el punto de vista técnico es que con frecuencia la interpretabilidad puede repercutir en la presión del sistema.

La transparencia incluye también una importante capa sociopolítica que trasciende el plano técnico. Este nivel permite ver la manera en que se usan los sistemas algorítmicos, las decisiones de diseño que toma cada actor y hace explícitos los supuestos de gobernanza. Así, la transparencia se convierte en una condición que hace posible la rendición de cuentas algorítmica al plantear formas de avanzar hacia la impugnación.

La transparencia incluye también una importante capa sociopolítica que trasciende el plano técnico. Este nivel permite ver la manera en que se usan los sistemas algorítmicos, las decisiones de diseño que toma cada actor y hace explícitos los supuestos de gobernanza.

Los principios de transparencia y rendición de cuentas actúan de manera combinada. Las soluciones para la rendición de cuentas se basan con frecuencia en el principio de transparencia, que, a su vez, ha de integrarse en un marco institucional que permita que se establezcan relaciones de rendición de cuentas.

2.3. Interrelación entre principios

Los principios de transparencia y rendición de cuentas actúan de manera combinada. Las soluciones para la rendición de cuentas se basan con frecuencia en el principio de transparencia, que, a su vez, ha de integrarse en un marco institucional que permita que se establezcan relaciones de rendición de cuentas.

Por ejemplo, los registros algorítmicos son instrumentos para la rendición de cuentas cuyo funcionamiento práctico consiste en mostrar con transparencia la información sobre los algoritmos y su uso en un repositorio disponible al público (Jameson y Leal, 2022; Cath y Jansen, 2021). De este modo, la transparencia es un medio que permite evaluar la rendición de cuentas en el diseño de sistemas algorítmicos.

Sin embargo, aunque sea condición necesaria para la rendición de cuentas, la transparencia resulta insuficiente. Por ejemplo, el mero hecho de que un sistema algorítmico esté bien documentado y sea transparente no implica que indique el motivo por el que se decidió que era «suficientemente bueno» para el propósito en cuestión, quién decidió que lo era y quién participó en el proceso. Mientras que la transparencia puede funcionar de forma pasiva, la rendición de cuentas es más activa: abarca no solo el modo en que funciona un sistema, sino el motivo por el que funciona (Wierenga, 2020).

3. Implementación de IA responsable en administraciones públicas urbanas

Al pensar en el uso responsable de la IA, hay dos preguntas básicas que las administraciones públicas urbanas deben plantearse: «¿Debe utilizarse la IA?» y «¿Cómo debe utilizarse la IA?». Brindar respuestas claras a estas preguntas aparentemente sencillas puede abrir una de las sendas más eficaces que conducen a la transparencia y la rendición de cuentas porque en esas contestaciones se hacen visibles los supuestos fundamentales. Este proceso requiere también que se asignen tiempo y recursos.

3.1. ¿Debe utilizarse la IA?

La IA no es neutra. Más bien, la IA incorpora y refuerza los supuestos establecidos en sus datos y su diseño. Si el diseño no se orienta de forma consciente hacia un conjunto de valores que sustenten el interés público, las estructuras y la gobernanza de la IA incorporarán valores de forma inconsciente, lo que entraña riesgos importantes (por ejemplo, ONU-Hábitat y Mila-Quebec AI Institute, 2022). Por tanto, la cuestión de si debe utilizarse la IA no debe tomarse a la ligera.

Para llevar a cabo una verdadera rendición de cuentas, la opción de dejar de utilizar la IA debe estar sobre la mesa. El «no» ha de ser siempre una posibilidad. De lo contrario, la rendición de cuentas se convierte en un mero principio, aludiendo a ella como virtud en lugar de como relación funcional (Wierenga, 2020).

En segundo lugar, la cuestión del «debe» no es solo normativa, sino también operativa. Las administraciones públicas intentan lograr un objetivo determinado, y la IA podría ser la mejor forma de conseguirlo. O puede que no. Hay otras soluciones tecnológicas o basadas en datos que pueden ser más adecuadas. En particular, las aplicaciones de IA y aprendizaje automático requieren una gran cantidad de datos de calidad, por lo que, cuando no se da esa condición, un análisis de datos más sencillo podría bastar.

Los proyectos municipales basados en datos suelen tener que enfrentarse a infraestructuras heredadas, sensores antiguos y bases de datos desconectadas. Esto significa que el éxito de las aplicaciones de aprendizaje automático en contextos urbanos requiere de fases de descubrimiento prolongadas, que a veces llegan a suponer entre el 30 y el 40% del calendario del proyecto. Este periodo comprende una labor de investigación sobre el problema en cuestión, el estado de las infraestructuras y los conjuntos de datos en ese momento, y sobre el tipo de solución que puede ser la más adecuada. Los presupuestos y las expectativas de las partes interesadas deben dar cabida a esta fase exploratoria ampliada.

Dado el entusiasmo y la atención que suscita la aplicación de la IA, existe un riesgo importante de tecnosolucionismo: el antiguo desafío del martillo en busca de un clavo. A veces, aplicar un enfoque social o conductual puede ser una mejor opción para resolver el problema que se plantea. Con frecuencia, distintos tipos de soluciones responden a distintos encuadres, lo que significa que la forma en que acotamos el problema establece los límites del espacio de soluciones. En otras palabras, la forma en que reflexionamos sobre el problema ya define los tipos de soluciones que podemos encontrar. Esto afecta no solo a la IA, sino, más en general, a la interacción entre las personas y la tecnología: un ejemplo sencillo es que, si el problema consiste en que un ascensor es demasiado lento, en lugar de intentar mejorar la velocidad mediante innovaciones de ingeniería mecánica, convendría instalar un espejo para que la gente no se aburriera tanto durante el trayecto. Una fase exploratoria ampliada también permite a las partes interesadas plantearse la pregunta fundamental: ¿cuál es el problema que tratamos de resolver?

La fase exploratoria ampliada incluye también una importante colaboración de las partes interesadas locales. Desarrollar satisfactoriamente IA es casi siempre un asunto colectivo que implica trabajar con universidades, centros de investigación y empresas locales, sobre todo teniendo en cuenta el déficit de capacidad que enfrentan los municipios. En Barcelona, por ejemplo, el algoritmo de aprendizaje automático desarrollado para asistir en la toma de decisiones en el proceso de admisión del centro de acogida de los servicios sociales fue el resultado de una importante colaboración entre varias entidades para poder elaborar un algoritmo bilingüe y pertinente a nivel local (Jameson y Leal, 2022).

3.2. ¿Cómo debe utilizarse la IA?

Aunque existen muchas aplicaciones diferentes de IA en las ciudades, la tendencia de uso entre las administraciones públicas encaja en dos amplias categorías: la automatización de los procesos existentes y las predicciones basadas en datos.

Al pensar en el uso responsable de la IA, hay dos preguntas básicas que las administraciones públicas urbanas deben plantearse: «¿Debe utilizarse la IA?» y «¿Cómo debe utilizarse la IA?». Brindar respuestas claras a estas preguntas aparentemente sencillas puede abrir una de las sendas más eficaces que conducen a la transparencia y la rendición de cuentas.

Dado el entusiasmo y la atención que suscita la aplicación de la IA, existe un riesgo importante de tecnosolucionismo. A veces, aplicar un enfoque social o conductual puede ser una mejor opción para resolver el problema que se plantea.

La automatización consiste en mecanizar parte de los procesos burocráticos o servicios urbanos existentes. En esta categoría hay una lógica o proceso que ya existe, y parte de esa cadena de eventos se va a hacer de forma más rápida o eficaz con la ayuda de la IA. Al considerar cómo aplicar la IA, el punto de partida es el sistema existente.

Las predicciones basadas en datos constituyen un planteamiento diferente, porque el punto de partida es otro: disponer de una gran cantidad de datos. A partir de esa información, el personal de análisis de datos extrae conclusiones que la administración utiliza como fundamento en la creación de nuevos procesos burocráticos para los servicios urbanos. Los modelos predictivos conforman una nueva lógica, basada en datos, de la administración (Kitchin, 2016).

Pese a que estas dos categorías pueden utilizar el mismo tipo de IA desde el punto de vista técnico (por ejemplo, ambas pueden utilizar técnicas de aprendizaje profundo o de reconocimiento de imágenes), la *forma* en que la IA se integra en los procesos de la ciudad es distinta; y esta diferencia modifica los tipos de repercusiones que la IA puede tener y, por tanto, cambia nuestra concepción sobre la rendición de cuentas y la transparencia.

Por ejemplo, cuando la IA se utiliza para automatizar procesos burocráticos existentes, los procesos de revisión vigentes pueden ampliarse con mecanismos adicionales de rendición de cuentas. Por ejemplo, una revisión trimestral puede ampliarse con una evaluación del impacto. Otras innovaciones de procesos pueden complementar prácticas organizativas existentes con el fin de tener en cuenta enseñanzas extraídas de la integración de la IA, como los comentarios de la ciudadanía y del funcionariado público que participa en el proceso.

Por otro lado, el uso de las predicciones basadas en datos requiere un enfoque algo más complejo de la transparencia y la rendición de cuentas, ya que se trata de una nueva forma de generación de conocimiento que las administraciones públicas tradicionales no están preparadas para procesar. Más concretamente, la modelización predictiva cambia el papel de la experiencia local y el lugar en el que se aplica (Kitchin, 2016). Se puede ver así: alguien que lleva 20 años de experiencia podría tener un punto de vista diferente de lo que indican los datos. El conocimiento computacional es distinto del conocimiento experiencial (Van Ewijk y Baud, 2009), y la toma de decisiones asistida por algoritmos puede cambiar el equilibrio entre ambos.

Los procesos de rendición de cuentas requieren un diálogo entre los distintos modos de comprensión, como la diferencia entre el conocimiento computacional y el experiencial. ¿Cómo interpretamos el problema urbano que se quiere resolver? Esta «producción de sentido» o «construcción de significado» consiste en decidir la manera en que valoramos distintas opciones de política y resultados sociales; y podría decirse que es algo para lo que la IA depende totalmente de los humanos (Tan, 2024). Examinar detenidamente y rediseñar los procesos de rendición de cuentas y mecanismos de política constituye una oportunidad para evaluar los distintos tipos de construcción de significado que entran en juego con el fin de garantizar que el uso de la IA en las administraciones públicas sea ético.

4. Mecanismos de política

Un enfoque sociotécnico de la IA reconoce que lo que ocurre con un sistema de IA es el resultado de la interacción entre lo técnico y lo social, entre el sistema y la manera en que este se integra en su contexto. Eso significa que, para entender el funcionamiento de sistemas algorítmicos, es importante comprender cómo interactúan con el entorno y qué mecanismos pueden ser más eficaces en cada momento.

Los sistemas algorítmicos pueden describirse mediante el «ciclo de vida de la IA», que es una forma abreviada de describir el proceso de diseño, desarrollo y utilización. Comprender este concepto es útil porque muchos de los marcos de gestión de riesgos disponibles se basan en variantes de este ciclo de vida de la IA.

Estas son distintas opciones de mecanismos de política disponibles en las distintas fases del ciclo de vida de la IA. También hay mecanismos generales de gobernanza institucional presentes, como trasfondo, durante todo el ciclo de vida de la IA.

Un enfoque sociotécnico de la IA reconoce que lo que ocurre con un sistema de IA es el resultado de la interacción entre lo técnico y lo social, entre el sistema y la manera en que este se integra en su contexto.

Encuadre y diseño:

- Las **evaluaciones del impacto** suelen concretarse en un cuestionario para analizar las posibles consecuencias éticas y sociales antes de la utilización de un sistema de IA. Hay muchas variantes de evaluaciones del impacto, como las relativas a la ética, la privacidad, la equidad, etc. Véase, por ejemplo, [la herramienta de evaluación del impacto ético de la Unesco](#).
- Las **cláusulas de contratación pública** son estipulaciones en los contratos que utilizan los gobiernos para adquirir bienes y servicios, en este caso servicios de IA o relacionados con ella. Aunque aparentemente son una formalidad burocrática pueden convertirse en una palanca estratégica para lograr objetivos de interés público, por ejemplo estableciendo normas de auditabilidad. Véase, por ejemplo, la alianza [GovAI Coalition](#), encabezada por la ciudad de San José (Estados Unidos), que ha elaborado unas plantillas de políticas para que las adapten las administraciones públicas y que incluyen, entre otras, una hoja informativa sobre IA y un acuerdo para entidades proveedoras que obliga a estos actores a cumplir requisitos relativos a los resultados, el sesgo algorítmico, la supervisión humana y otros aspectos. La red Eurocities también está elaborando unas [plantillas de cláusulas de contratación](#) conforme a la Ley de IA de la Unión Europea

Desarrollo:

- Las **auditorías algorítmicas externas** son evaluaciones independientes del funcionamiento de los sistemas algorítmicos para garantizar el cumplimiento de las normas jurídicas y éticas. Véase, por ejemplo, la [lista de control para la auditoría de IA del Comité Europeo de Protección de Datos](#).

Utilización:

- Los **registros de algoritmos y los estándares de transparencia** son listas de acceso público en las que se efectúa un seguimiento del uso que las administraciones públicas hacen de los algoritmos o la IA, con el fin de que la ciudadanía y las partes interesadas puedan acceder a esa información. Estos repositorios se basan en un esquema común de metadatos e información sobre el algoritmo. Véase el [Estándar de Transparencia Algorítmica](#), en el que figura el esquema de códigos que utilizan nueve ciudades europeas. Otra iniciativa similar es la [Norma sobre el Registro en materia de Transparencia Algorítmica de Reino Unido](#)

Contexto de política y gobernanza:

- Los **comités interdisciplinarios de supervisión de la gobernanza** reúnen a personas expertas de diversos ámbitos —como el derecho, la ética y las ciencias sociales— y representantes de las comunidades afectadas con el fin de presentar un conjunto diverso de puntos de vista en el proceso de supervisión. Para ser eficaces, estos consejos de supervisión deben ser independientes y tener un auténtico poder de veto.
- Los **procesos participativos**, especialmente con las comunidades afectadas, implican de forma activa y significativa a las personas en todas las fases del ciclo de vida de la IA, desde el momento del encuadre y diseño y no solo *a posteriori*. Mediante un proceso más equitativo, estos grupos pueden ayudar a codiseñar resultados más equitativos.
- El **diseño con intervención humana** supone que las personas siguen tomando las decisiones clave en todos los puntos de un sistema para reducir errores y permitir el control manual. Pese a que los sistemas algorítmicos nunca se desvinculan totalmente de las personas porque todos los sistemas incorporan sus valores de diseño (y muchos son de propiedad empresarial), este enfoque de diseño sigue siendo útil para hacer hincapié en el hecho de que las personas deben seguir siendo quienes tomen las decisiones finales.

5. Lecciones aprendidas

En una investigación previa del CIDOB (Jameson y Leal, 2022) se analizaron varios casos prácticos y experiencias de administraciones municipales que aplicaban mecanismos de rendición de cuentas y transparencia en materia de IA urbana. En concreto, se estudiaron el registro de algoritmos de Ámsterdam, el registro de IA de Helsinki y un caso de aprendizaje automático explicable desarrollado para los servicios sociales de Barcelona. En este capítulo se destacan algunas de las recomendaciones y lecciones aprendidas para que las iniciativas de transparencia y rendición de cuentas tengan éxito.

Diseño:

- Las iniciativas de rendición de cuentas y transparencia de IA funcionaban bien cuando se encuadraban como asuntos de interés público, vinculándolas con cuestiones sociales más amplias, y no como meros problemas técnicos.
- De la identificación de las prioridades municipales derivan las definiciones locales de éxito, lo que significa que las iniciativas pueden variar de una localidad a otra. En varios casos, estas variantes obedecían a los sucesos y la información de la zona.
- Las personas tienen distintas expectativas sobre lo que puede lograr una iniciativa de rendición de cuentas de IA en la administración pública. Para funcionar, los proyectos requerían una gran cantidad de energía y que se designara un «propietario o propietaria» del proyecto como la principal persona de referencia, que dedicaba mucho tiempo a gestionar las expectativas de las partes interesadas.

Proceso:

- Determinar definiciones claras y comprensibles para todo el funcionariado, que no es experto en la materia. Los términos clave para garantizar el alineamiento son: el algoritmo, el objeto de la transparencia, la fecha de publicación, el foro ante el cual se rinden cuentas, y el propietario o propietaria del producto para cada elemento del proyecto.
- Identificar las prácticas organizativas que pueden ampliarse con los procesos de rendición de cuentas. Por ejemplo, se consideró que las reuniones trimestrales sobre los informes económicos eran el momento en que el equipo directivo ya estaba reunido y podía examinar otras innovaciones técnicas.
- Empezar poco a poco e iterar. Los cambios en el funcionamiento de la administración pública llevan tiempo, y funcionan mejor cuando se aplican de forma gradual y no de una sola vez.

Capacidad:

- En todas las iniciativas de rendición de cuentas se requería de inversión en capacitación a fin de actualizar la formación del funcionariado, así como de tiempo para que el personal se familiarizase con los nuevos enfoques.
- Vincularse con redes de intercambio de conocimientos, como la Coalición de Ciudades por los Derechos Digitales, en las que se ponen en común e intercambian experiencias sobre la adaptación de los mecanismos de transparencia y rendición de cuentas.

Referencias bibliográficas

Adadi, A. y Berrada, M. «Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)». *IEEE Access*, vol. 6 (2018), p. 52138–52160.

Amaro, S. «[Dutch government resigns after childcare benefits scandal](#)» *CNBC* (enero 2021) [Fecha de consulta 18.01.2021]

Bovens, M. «Analysing and Assessing Accountability: A Conceptual Framework». *European Law Journal*, vol. 13, n.º 4 (2007), p. 447–468.

Cath, C. y Jansen, F. «Dutch Comfort: The Limits of AI Governance through municipal registers». *ArXiv* (septiembre 2021).

Dachwitz, I. «[Childcare benefits scandal: Dutch government to pay million Euro fine over racist data discrimination](#)». *Netzpolitik.org* (enero 2022)

Van Ewijk, E. y Baud, I. «Partnerships between Dutch municipalities and municipalities in countries of migration to the Netherlands; knowledge exchange and mutuality». *City-to-City Co-Operation*, vol. 33, n.º 2 (2009), p. 218-226.

Jameson, S. y Leal, A. «Transparency and accountability in urban artificial intelligence: Lessons from city initiatives». *Observatorio Global de la Inteligencia Urbana (GOUAI)*, CIDOB (2022).

Jameson, S., Taylor, L. y Noorman, M. «Data Governance Clinics: a New Approach to Public-Interest Technology in Cities». *SSRN Scholarly Paper. Rochester (Nueva York): Social Science Research Network* (septiembre 2021).

Rob, K. «The ethics of smart cities and urban science». *Philosophical Transactions of the Royal Society*, vol. 374, n.º 2083 (diciembre 2016): 20160115

Longpre, S. et al. «The Data Provenance Initiative: A Large Scale Audit of Dataset Licensing & Attribution in AI». *ArXiv* (noviembre 2023).

Organización para la Cooperación y el Desarrollo Económico (OCDE), «[G7 Toolkit for Artificial Intelligence in the Public Sector](#)» (2024)

Tan, V. «[AI's meaning-making problem](#)». Boletín de *Substack*. *The Uncertainty Mindset (Soon to Become Tbd)*, blog (mayo 2024)

Taylor, L. «[Fairness and AI governance – responsibility and reality](#)». *Global Data Justice*, blog (abril 2021)

ONU-Hábitat y Mila-Quebec AI Institute. «[AI & Cities: Risks, Applications and Governance](#)» (2022)

Wieringa, M. «What to account for when accounting for algorithms: a systematic literature review on algorithmic accountability». *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, (2020), p. 1-18. Nueva York (Nueva York, Estados Unidos): Association for Computing Machinery.

Josuan Eguiluz Castañeira

Asesor jurídico, Adevinta

Carlos Fernández Hernández

Miembro del consejo asesor, Global LegalTech

1. Introducción

Desde que en 2018 la Unión Europea (UE) comenzó el diseño de su marco normativo sobre la inteligencia artificial (IA), puso especial énfasis en que esta tecnología debe ser «fiable» (*trustworthy*). Se considera una IA fiable aquella que respeta el marco normativo aplicable y que es ética y robusta, tanto desde el punto de vista técnico como social, puesto que los sistemas de IA, incluso si las intenciones son buenas, pueden provocar daños accidentales (Grupo de expertos de alto nivel sobre inteligencia artificial, 2019).

En consecuencia, el enfoque europeo en esta materia, incentiva el desarrollo y la adopción de una IA ética y fiable en toda la economía de la UE, a partir del principio de que dicha tecnología debe estar al servicio de las personas y ser una fuerza positiva para la sociedad (Libro Blanco sobre la inteligencia artificial, 2020, § 6).

Dado que la disponibilidad de datos es fundamental para entrenar a los sistemas algorítmicos, y que muchos de esos datos son de carácter personal, un componente de la IA ética es que debe incluir mecanismos de gestión de la privacidad y de los datos (Comisión Europea, Revisión de 2021 del plan coordinado sobre la IA). Esta exigencia ha sido plenamente acogida por el Reglamento europeo en materia de inteligencia artificial (RIA) (junio de 2024), que fija como uno de sus objetivos promover la adopción de una IA centrada en el ser humano y fiable (art. 1), respetando, a la vez, el marco normativo vigente en materia de protección de datos, constituido –principal pero no exclusivamente–, por el Reglamento general de protección de datos de 2016 (RGPD).

Como han señalado algunos autores (Almonacid Lamelas, 2024), el RIA representa un desafío no menor para los gobiernos locales, en tanto que deberán adaptar sus procesos, políticas y estrategias para cumplir con las nuevas exigencias. Sin embargo, también supone una oportunidad para mejorar su funcionamiento, así como la calidad y confiabilidad de los servicios basados en IA ofrecidos a los ciudadanos (*ibid.*). Esto explica la proliferación de sistemas de «inteligencia artificial urbana» (*urban AI*), un concepto que hace referencia a «la recopilación, interpretación y análisis

Los sistemas de IA deben garantizar la protección de los datos a lo largo de todo el ciclo de vida de dichos sistemas. Esto incluye la información inicialmente facilitada por el usuario, así como la que se genera sobre él en el contexto de su interacción con el sistema.

de datos urbanos con el fin de apoyar la toma de decisiones relacionadas con las políticas, así como el desarrollo de soluciones que se utilizan, o podrían utilizarse, en un contexto urbano» (Galcerán-Vercher, 2023).

Con todo, el tratamiento de datos personales en el ámbito público-urbanístico puede plantear problemas específicos, desde la legitimidad de dicho tratamiento para una finalidad para la que originalmente no fue consentida, hasta la necesidad de realizar evaluaciones de impacto en los derechos fundamentales de las personas. Estos, inequívocamente, deberán ser tenidos en cuenta por los organismos públicos.

A la luz del nuevo marco legislativo, el objeto de este artículo es (i) presentar el marco jurídico y ético que regula el tratamiento de datos personales en el ámbito urbanístico por medio de sistemas de IA, en especial a nivel europeo (RIA); (ii) identificar los principales mecanismos para implementar el principio de privacidad, y (iii) analizar los desafíos que plantea este tipo de tratamiento de datos y ofrecer un conjunto de recomendaciones y buenas prácticas para minimizarlos o eliminarlos.

2. IA ética y privacidad

Una IA fiable debe ser ética, y para serlo debe, entre otros requisitos, respetar la privacidad de las personas. El RIA establece como objetivo específico «promover la adopción de una IA centrada en el ser humano y fiable». Con este fin, las normas comunes que establece para los sistemas de IA de alto riesgo deben ser coherentes con la Carta de los Derechos Fundamentales de la Unión Europea (2000), y tener en cuenta tanto la Declaración Europea sobre los Derechos y Principios Digitales para la Década Digital (2022), como las Directrices éticas para una IA fiable del Grupo independiente de expertos de alto nivel sobre inteligencia artificial (2019). Según estas directrices, en un contexto de rápido cambio tecnológico,

«La fiabilidad es un requisito previo para que las personas y sociedades desarrollen, desplieguen y utilicen sistemas de IA. Si estos sistemas –y las personas que se encuentran detrás de ellos– no demuestran ser merecedores de confianza, pueden producirse consecuencias no deseadas que obstaculicen su adopción, impidiendo el logro de los enormes beneficios económicos y sociales que pueden acarrear los sistemas de IA.» (Introducción, punto 13)

La fiabilidad de la IA se apoya en tres componentes, que deben satisfacerse a lo largo de todo el ciclo de vida del sistema de IA:

1. Debe ser lícita, de modo que se garantice el respeto de todas las leyes y normativa aplicables;
2. Ha de ser ética, es decir, debe asegurar el cumplimiento de los principios y valores éticos, y, finalmente,
3. Debe ser robusta, tanto desde el punto de vista técnico como social, puesto que los sistemas de IA, incluso si las intenciones son buenas, pueden provocar daños accidentales.

Por tanto, se debe establecer la ética como pilar fundamental para garantizar y expandir una IA fiable. Esto implica que es preciso garan-

tizar que se cumplan unas normas éticas básicas, así como las medidas que establece el RIA para la protección de los derechos fundamentales.

En este sentido, la protección de datos es un derecho fundamental que se ve especialmente afectado por los sistemas de IA, y que guarda una estrecha relación con el principio de prevención del daño. Dicho principio de prevención comienza por una adecuada gestión de esos datos, que abarque la calidad y la integridad de aquellos que sean utilizados, su pertinencia en contraste con el ámbito en el que se desplegarán los sistemas de IA, sus protocolos de acceso y la capacidad para procesar datos sin vulnerar la privacidad.

Entre esas medidas se incluye el hecho de que los sistemas de IA dispongan de un mecanismo de gestión de la privacidad y de los datos que incluya tanto el respeto de la privacidad, como la calidad y la integridad de dichos datos y el acceso a estos.

Además, los sistemas de IA deben garantizar la protección de los datos a lo largo de todo el ciclo de vida de dichos sistemas. Esto incluye la información inicialmente facilitada por el usuario, así como la que se genera sobre él en el contexto de su interacción con el sistema (por ejemplo, los productos que genere el sistema de IA para determinados usuarios o la respuesta de estos a ciertas recomendaciones). Los registros digitales del comportamiento humano pueden posibilitar que los sistemas de IA no solo infieran las preferencias de las personas, sino también su orientación sexual, edad, género u opiniones políticas y religiosas. Para permitir que los individuos confíen en el proceso de recopilación de datos, es preciso garantizar que la información recabada sobre ellos no se utilizará para discriminarlos de forma injusta o ilegal.

Del cumplimiento de estos requisitos tienen que encargarse los operadores, en particular, los desarrolladores de los sistemas de IA y los responsables del despliegue (que deben asegurarse de que los sistemas que utilizan y los productos y servicios que ofrecen cumplen los requisitos establecidos). Por otro lado, las personas que se vean afectadas por el funcionamiento de un sistema de IA tendrán derecho a estar informadas de dicha afectación y, en su caso, a presentar una reclamación por infracción del RIA (arts. 85 y 86).

2.1. La privacidad en el Reglamento europeo de IA

El art. 2.7 del RIA recoge el principio general de que el mismo respeta en su integridad el marco regulador de la Unión en materia de protección de datos establecido por el RGPD.

En primer lugar, las normas armonizadas que se establecen en el RIA deben aplicarse en todos los sectores y deben entenderse sin perjuicio del derecho vigente de la Unión. Es importante destacar, pues, que el RIA no pretende afectar a la aplicación del derecho de la Unión que regula el tratamiento de datos personales, incluidas las funciones y competencias de las autoridades de supervisión independientes que vigilan el cumplimiento de dichos instrumentos. Del mismo modo, tampoco afecta a las obligaciones previas de los proveedores y los encargados del despliegue de sistemas de IA en su papel de responsables del tratamiento

[E] RIA recoge el principio general de que el mismo respeta en su integridad el marco regulador de la Unión en materia de protección de datos establecido por el RGPD.

Para las ciudades, garantizar que sus sistemas de IA cumplen con regulaciones como el RGPD o el RIA a lo largo de todo el ciclo de vida de la IA, resulta fundamental para salvaguardar los derechos de los ciudadanos y mantener la confianza pública.

de datos personales. En particular, el RIA no debe afectar a las prácticas actualmente prohibidas por el derecho de la Unión, incluidos los derechos en materia de protección de datos.

En paralelo, el hecho de que un sistema de IA sea clasificado como de alto riesgo no debe interpretarse como indicador de que su uso sea lícito con arreglo a otros actos del derecho de la Unión o del derecho nacional, por ejemplo, en materia de protección de los datos personales. Todo uso de ese tipo de sistemas de IA debe seguir realizándose exclusivamente en consonancia con los requisitos oportunos derivados de la Carta, del derecho derivado de la Unión y del derecho nacional.

Además, el RIA no constituye un fundamento jurídico para el tratamiento de datos personales, incluidas las categorías especiales de dichos datos, salvo que se disponga específicamente otra cosa. Por ello, los interesados siguen disfrutando, tras la entrada en vigor del RIA, de todos los derechos y las garantías que les confiere el derecho de la Unión, incluidos los relacionados con las decisiones individuales totalmente automatizadas, como la elaboración de perfiles. Las normas armonizadas que establece el RIA, pues, deben permitir el ejercicio de los derechos y otras vías de recurso de los interesados garantizados por el derecho de la Unión en materia de protección de datos personales y otros derechos fundamentales.

Finalmente, a fin de facilitar el cumplimiento del derecho de la Unión en materia de protección de datos, en determinadas condiciones, el RIA proporciona la base jurídica para que, en un espacio controlado de pruebas, los proveedores (también los potenciales) utilicen datos personales recabados para otros fines para desarrollar determinados sistemas de IA en favor del interés público.

3. Mecanismos políticos para implementar el principio de privacidad en el ámbito urbanístico

La protección de la privacidad y los datos en la implementación de IA urbana requiere la adopción de mecanismos políticos específicos. Estos mecanismos permiten a las ciudades cumplir con las normativas vigentes y asegurar que la IA se despliegue de manera ética y responsable, respetando los derechos de los ciudadanos. A continuación, se identifican y explican los principales mecanismos políticos para implementar este principio ético.

a) Garantizar la conformidad legal

El cumplimiento de la regulación es un requisito ético esencial en la protección de la privacidad y los datos en la implementación de sistemas de IA en entornos urbanos por parte de las autoridades públicas. Para las ciudades, garantizar que sus sistemas de IA cumplen con regulaciones como el RGPD o el RIA a lo largo de todo el ciclo de vida de la IA, resulta fundamental para salvaguardar los derechos de los ciudadanos y mantener la confianza pública. Esto incluye la adhesión a requerimientos clave como la calidad e integridad de los datos utilizados, su pertinencia en contraste con el ámbito en el que se desplegarán los sistemas de IA, sus

protocolos de acceso y la capacidad para procesar datos sin vulnerar la privacidad (Grupo de expertos de alto nivel sobre IA, 2018).

Precisamente, estos requerimientos se materializan en obligaciones concretas en el propio RIA, específicamente diseñadas para casos de alto riesgo, tales como los sistemas de IA de identificación biométrica remota –p. ej., el programa ABIS (Pascual, 2024)– o aquellos utilizados para evaluar la admisibilidad de las personas físicas para beneficiarse de servicios y prestaciones esenciales de asistencia pública –p. ej., el caso Syri (Digital Future Society, 2022)–.

b) Sistemas de gestión de riesgo y gobernanza de datos

El RIA incluye obligaciones específicas (arts. 9 y 10) estrechamente vinculadas al principio de privacidad y protección de datos. Por un lado, el artículo 9 se centra en la creación de un sistema de gestión de riesgos que sea capaz de identificar, documentar y mitigar aquellos que vayan asociados al uso de IA en las ciudades. Estos sistemas de gestión de riesgos deberán establecer procesos iterativos continuos, planificados y ejecutados a lo largo de todo el ciclo de vida de las tecnologías de IA, que, por supuesto, requerirán revisiones y actualizaciones sistemáticas periódicas. De hecho, no se trata solo de evaluar los posibles riesgos antes de la introducción en el mercado o puesta en servicio de estos sistemas de IA, sino también de establecer y/o supervisar el funcionamiento de un sistema de vigilancia poscomercialización para gestionar riesgos emergentes –arts. 17.1 h), 26.5 y 72 RIA–.

Por otro lado, la gobernanza de datos regulada en el artículo 10 exige que los conjuntos de datos de entrenamiento, validación y prueba utilizados en sistemas de IA de alto riesgo se sometan a prácticas de gobernanza y gestión de datos adecuadas para su finalidad prevista. Las prácticas a implementar por las ciudades para asegurar una gobernanza de datos efectiva y legal se centrarán en cuestiones como los procesos de recogida y origen de los datos, la finalidad del tratamiento, la evaluación de la disponibilidad, la cantidad y la adecuación de los conjuntos de datos necesarios, el examen de posibles sesgos que puedan afectar a la salud, la seguridad o los derechos fundamentales de las personas, etc.

c) Evaluaciones de impacto

El artículo 35 del RGPD impone a los responsables del tratamiento (p. ej., ayuntamientos) la obligación de realizar una evaluación de impacto relativa a la protección de datos (EIPD). Esta evaluación se realizará cuando sea probable que un tipo de tratamiento, por su naturaleza, alcance, contexto o fines (en particular si utiliza nuevas tecnologías), entrañe un alto riesgo para los derechos y las libertades de las personas físicas (AEPD, 2018; Grupo de Trabajo del art. 29, 2017; Friedewald *et al.*, 2022). Este enfoque preventivo es vital en los entornos urbanos para anticipar posibles vulnerabilidades en la protección de datos y tomar las medidas necesarias para corregirlas a tiempo.

Asimismo, para los sistemas de IA de alto riesgo, el artículo 27 RIA introduce la obligación de realizar una evaluación de impacto relativa a los

derechos fundamentales (en sus siglas en inglés, FRAI) (Gobierno de Holanda, 2022; Instituto Danés de Derechos Humanos, 2020), que complementa la EIPD. Esta evaluación tiene como objetivo determinar los riesgos específicos para los derechos de las personas que probablemente se vean afectadas y definir las medidas que deben adoptarse en caso de que se materialicen dichos riesgos (Cdo. 96 RIA). Cabe destacar que las evaluaciones de impacto (Manzoni, M. *et al.*, 2022) deben centrarse no solo en el retorno de la inversión, sino también en la sostenibilidad y el impacto ético de la tecnología, abordando aspectos financieros, humanos y medioambientales (OECD, 2024).

d) Realización de auditorías

Sentado lo anterior, será necesario poder demostrar ante las autoridades, las partes interesadas y los ciudadanos que se cumple con la legislación y todos sus requisitos de implementación específicos. En este sentido, se realizarán auditorías internas y externas, y se obtendrán certificaciones que verifiquen que los sistemas operan dentro de los marcos legales establecidos. Para ello, las ciudades europeas, por ejemplo, deberán realizar evaluaciones de conformidad (art. 43 RIA) con el fin de garantizar y demostrar que han cumplido con los requisitos asociados a sistemas de alto riesgo, alineándose con las normas armonizadas publicadas en el *Diario Oficial de la Unión Europea* (art. 41 RIA). También deberán seguir las especificaciones comunes establecidas por la Comisión Europea, asegurando así una implementación estandarizada y segura de los sistemas de IA (p. ej., certificaciones ISO).

Las auditorías de IA se consideran un mecanismo de gobernanza fundamental para asegurar que la implementación y operación de los sistemas de IA cumplen con las normativas legales y los estándares éticos y técnicos establecidos (Fernández, C. y Eguiluz, J. A., 2024). Con carácter general, estas auditorías deberán ser realizadas por entidades independientes y competentes. El proceso de auditoría incluye metodologías que incorporan evaluaciones de impacto ético (UNESCO, 2024; CEN-CENELEC, 2017), asegurando que los sistemas de IA se comportan de manera responsable y que sus efectos en la sociedad y los individuos son debidamente monitoreados y mitigados. No obstante, es recomendable plantear las auditorías de IA desde un punto de vista multidisciplinar –legal, técnico y ético– (Mökander, J., 2023). En esta dirección, surgen propuestas como la de «algo-scores» para clasificar y evaluar de manera accesible el nivel de conformidad de los sistemas algorítmicos en materias como el cumplimiento ético, la gobernanza de la IA, la equidad del modelo y su vigilancia posterior, siguiendo un enfoque similar al de las etiquetas de eficiencia energética (Galdon Clavell, 2024).

e) Repositorios de algoritmos y registros de sistemas de IA

En último lugar, conviene recordar la importancia de los repositorios de algoritmos públicos y los registros de sistemas de IA (art. 49 RIA), que promueven la transparencia en la toma de decisiones automatizadas en el sector público y desempeñan un papel crucial en la protección de la privacidad y los datos personales. Al hacer accesibles los detalles sobre cómo se diseñan, implementan y operan estos sistemas, los repositorios

y registros permiten a los ciudadanos y a las organizaciones entender cómo y con qué finalidades se utilizan sus datos personales en estos procesos. Dichos repositorios incluyen también información sobre las fuentes de datos utilizadas y los mecanismos de supervisión, lo cual es esencial para evaluar el impacto de los sistemas en la privacidad de los individuos y garantizar que las medidas de protección de datos sean efectivas (Gutiérrez & Muñoz-Cadena, 2024).

4. Desafíos y recomendaciones

La gestión inadecuada de los datos es una de las principales limitaciones para la implementación de IA en el sector público. Igualmente, lo es también la falta de acceso a volúmenes suficientes de datos de alta calidad. Este problema se ve exacerbado por el intercambio insatisfactorio de datos entre organizaciones debido a la falta de estándares unificados y una gobernanza de datos subdesarrollada. Además, la desconfianza en los sistemas de IA agrava estos desafíos. Las leyes dispersas y la falta de conocimiento sobre los impactos de la IA también generan barreras significativas (Manzoni *et al.*, 2023). Asimismo, el aumento de ciberataques ha llevado a que la [Directiva NIS 2](#) (2022) refuerce la seguridad y la responsabilidad legal para los administradores. En 2023, la administración pública fue uno de los sectores más afectados, con el 19% de los incidentes reportados, destacando entre ellos el crecimiento de ataques como el *ransomware* y DDoS (ENISA, 2023).

El complejo panorama regulatorio también representa un desafío significativo. La interacción entre las normativas urbanísticas a nivel europeo, nacional y local crea un entramado de reglas que complica la implementación efectiva de IA en las ciudades. La legislación urbanística y las regulaciones específicas de cada municipio deberán alinearse con normativas europeas como el Reglamento sobre la Europa Interoperable (Interoperable Europe Act) (2022), que busca mejorar la interoperabilidad de los servicios públicos digitales (Tangi *et al.*, 2023).

Otra limitación importante es la falta de experiencia y conocimiento técnico dentro de las administraciones locales, lo que dificulta la correcta implementación de la IA. La escasez de profesionales en esta materia a nivel global, junto con la creciente competencia por el talento, representan una barrera significativa para las ciudades que intentan desarrollar y desplegar estos sistemas de manera efectiva (OECD, 2024).

Por otra parte, la recolección masiva de datos personales, necesaria para entrenar estos sistemas, puede infringir el derecho de los ciudadanos a controlar sus datos, ya que estos pueden ser sensibles o ser gestionados de manera inapropiada. Mientras que las aplicaciones de IA, como las utilizadas en el control policial, pueden intensificar la vigilancia masiva y comprometer aún más la privacidad de los individuos (Véliz, 2020; Agarwal, 2018; Dwivedi *et al.*, 2019).

Para superar estas barreras, es esencial promover mecanismos de innovación, como los *sandboxes* regulatorios (Madiaga, 2022), que permiten a las ciudades experimentar con IA en un entorno controlado mientras se garantiza el cumplimiento regulatorio (Tangi *et al.*, 2023). Asimismo, la coordinación entre las autoridades nacionales (en el caso español, la

Es esencial promover mecanismos de innovación, como los *sandboxes* regulatorios, que permiten a las ciudades experimentar con IA en un entorno controlado mientras se garantiza el cumplimiento regulatorio.

Agencia Española de Supervisión de la Inteligencia Artificial - AESIA) y europeas (Oficina Europea de IA) es crucial para garantizar que los sistemas de IA cumplan con las normativas vigentes y se implementen de manera segura y responsable.

La interoperabilidad y la colaboración son igualmente claves. Iniciativas como el sistema SALER - Sistema de Alertas Rápidas, utilizado en la Comunidad Valenciana para prevenir la corrupción en la administración, demuestran cómo la IA puede utilizarse de manera efectiva para mejorar los procesos de gobernanza (Digital Future Society, 2023). Igualmente, resulta esencial que la financiación pública esté condicionada a la disponibilidad de ciertos resultados por parte de las distintas administraciones (p. ej., generar conjuntos de datos públicos) (Comisión Europea, 2022). En este sentido, la Comisión Europea ha publicado, mediante el [Reglamento de ejecución \(UE\) núm. 2023/138](#), una lista de conjuntos de datos específicos de alto valor que deberán estar disponibles para su reutilización gratuita, destacando el potencial de los datos públicos en beneficio de la sociedad, el medio ambiente y la economía (Comisión Europea, 2022). Además, el acceso a datos multilingües para entrenar modelos locales de IA que reflejen las características específicas de cada región (OECD, 2024) y la recopilación de casos de uso de IA en el sector público a nivel europeo (Comisión Europea, 2021) mejorarán la efectividad y equidad de los sistemas de IA, a la vez que proporcionarán una valiosa fuente de información sobre cómo se están implementando estas tecnologías en diversos contextos.

5. Conclusiones

El marco general de protección de datos en la UE está ya asentado en unos principios conocidos y sólidamente interpretados por los organismos administrativos y jurisdiccionales de la Unión. Sin embargo, la IA plantea unos problemas específicos, de naturaleza tecnológica y jurídica, que se encuentran en un momento incipiente de conocimiento y tratamiento.

Por ello, serán necesarios todavía numerosos estudios, experiencias y precisiones para dotarles de un marco jurídico que garantice la proclamada finalidad de que la IA debe estar centrada en el ser humano, ser una herramienta para las personas y tener por objetivo último aumentar su bienestar.

La implementación de mecanismos políticos específicos es esencial para garantizar que las ciudades utilicen sistemas de IA de manera ética y respetuosa con los derechos de la ciudadanía. El cumplimiento de regulaciones como el RGPD y el RIA resulta indispensable para salvaguardar la privacidad y los datos personales en entornos urbanos. Del mismo modo, es fundamental que las ciudades establezcan sistemas de gestión de riesgos que aborden de manera iterativa las contingencias asociadas a todo el ciclo de vida de la IA, incluyendo revisiones periódicas y auditorías externas que aseguren el cumplimiento normativo.

Por otra parte, la gobernanza de datos debe estar en el centro de las estrategias urbanas de IA. Las ciudades deben implementar prácticas de gobernanza y gestión de datos sólidas, centradas en la calidad, la pertinencia y la protección de los conjuntos de datos utilizados en los

sistemas de IA. Esto incluye la realización de evaluaciones de impacto tanto para la protección de datos personales como para los derechos fundamentales, asegurando que las tecnologías implementadas no vulneren la privacidad ni la seguridad de la ciudadanía.

En definitiva, lograr una IA centrada en el ser humano exigirá un esfuerzo conjunto entre los responsables de desarrollar políticas públicas, las instituciones académicas y los sectores privados, que deben colaborar para asegurar que los sistemas de IA que sean implementados por las ciudades se alineen con los valores y principios éticos fundamentales.

Como se ha señalado, el futuro de las ciudades inteligentes se caracterizará por la síntesis de múltiples tecnologías orientadas a satisfacer el intrincado mosaico de necesidades humanas. Esta convergencia requerirá de una precisa optimización de las tecnologías aplicadas que asegure que la digitalización de los espacios urbanos se alinee con prácticas sostenibles y equitativas, así como de la atención a las dimensiones éticas que estas innovaciones conllevan. Por ello, es imperativo que la integración de la IA en el corazón de las ciudades inteligentes se rija por principios que defiendan la privacidad, la seguridad y la inclusión. En este sentido, y como apuntan Zhenjun *et al.* (2023): «Garantizar que los beneficios del desarrollo de las ciudades inteligentes se compartan equitativamente será esencial para evitar fracturas sociales y fomentar un entorno en el que la tecnología sirva de puente hacia una vida urbana más ilustrada y armoniosa».

Referencias bibliográficas

Agencia Española de Protección de Datos (AEPD). «[Gestión del riesgo y evaluación de impacto en tratamientos de datos personales](#)». Madrid, 2018 [Fecha de consulta: 02.09.2024]

Agencia Española de Protección de Datos. «[Guía sobre Protección de Datos y Administración Local](#)». Actualizada en 2023 [Fecha de consulta: 02.09.2024]

AI Ethics Impact Group. «[From Principles to Practice: An interdisciplinary framework to operationalise AI ethics](#)». 1 de abril de 2020 [Fecha de consulta: 02.09.2024]

Alan Turing Institute. «[Urban analytics](#)» [Fecha de consulta: 02.09.2024]

Alan Turing Institute. «[Why the public sector needs to know about AI ethics \(and how we're helping\)](#)». 2 de noviembre de 2023 [Fecha de consulta: 02.09.2024]

Almonacid Lamelas, V. «[Reglamento \(europeo\) de Inteligencia Artificial: impactos y obligaciones que genera en los Ayuntamientos](#)». El Consultor de los Ayuntamientos, LA LEY, 15 de julio de 2024 [Fecha de consulta: 02.09.2024]

Burbano, L. (1). «[Privacy protection in smart cities: How are they taking care of citizens' most precious information?](#)» *Tomorrow.city*, 23 de enero de 2024 [Fecha de consulta: 02.09.2024]

Burbano, L. (2). «AI urbanism: risks and benefits of a seemingly unstoppable movement». *Tomorrow.city*, 22 de febrero de 2024 [Fecha de consulta: 02.09.2024]

Canda, J. «AI in Urban Planning and Smart City Development». *Medium*, 7 de abril de 2024 [Fecha de consulta: 02.09.2024]

CEN-CENELEC. «Ethics assessment for research and innovation - Part 2: Ethical impact assessment framework». CWA 17145-2, junio 2017 [Fecha de consulta: 02.09.2024]

Centro Latinoamericano de Administración para el Desarrollo (CLAD). «Carta Iberoamericana de Inteligencia Artificial en la Administración Pública». 20 de noviembre de 2023 [Fecha de consulta: 02.09.2024]

Comisión Europea (1). «Selected AI cases in the public sector (JRC129301)». Joint Research Centre (JRC), 2021 [Fecha de consulta: 02.09.2024]

Comisión Europea (2). «Revisión de 2021 del plan coordinado sobre la inteligencia artificial», Anexos de la Comunicación de la Comisión al Parlamento Europeo, al Consejo Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones. Fomentar un planteamiento europeo en materia de inteligencia artificial, COM(2021) 205 final, 2021 [Fecha de consulta: 02.09.2024]

Comisión Europea (3). «Opportunities and challenges of artificial intelligence technologies for the cultural and creative sectors». Oficina de Publicaciones de la UE, Luxemburgo, 2022 [Fecha de consulta: 02.09.2024]

Comisión Europea (4). «Second Report on the application of the General Data Protection Regulation», COM(2024) 357 final, Bruselas, 25 de julio de 2024 [Fecha de consulta: 02.09.2024]

Digital Future Society (1). «Algorithmic discrimination in Spain: limits and potential of the legal framework», agosto de 2022 [Fecha de consulta: 02.09.2024]

Digital Future Society (2). «El acceso digital en las ciudades, entendido como algo más que un derecho fundamental», junio de 2023 [Fecha de consulta: 02.09.2024]

Digital Future Society (3). «El uso de algoritmos en el sector público en España: cuatro estudios de caso sobre ADMS», febrero de 2023 [Fecha de consulta: 02.09.2024]

European Data Protection Board (EDPB). «Statement 3/2024 on data protection authorities' role in the Artificial Intelligence Act framework», 16 de julio de 2024 [Fecha de consulta: 02.09.2024]

European Union Agency for Cybersecurity (ENISA). «Threat Landscape 2023», octubre de 2023 [Fecha de consulta: 02.09.2024]

Fernández, C. y Eguiluz, J. A. «Diez puntos críticos del Reglamento europeo de Inteligencia Artificial». *Diario LA LEY*, núm. 85, Sección Ciberderecho, 28 de junio de 2024 [Fecha de consulta: 02.09.2024].

Friedewald, M. et al. «Data Protection Impact Assessments in Practice: Experiences from Case Studies». *Computer Security, ESORICS 2021 International Workshops*, febrero de 2022, p. 424-443 [Fecha de consulta: 02.09.2024]

Galceran-Vercher, M. «Trustworthy Cities: Ethical Urban Artificial Intelligence». The GovLab, Course «AI Ethics, Global perspectives», diciembre de 2023 [Fecha de consulta: 02.09.2024]

Galceran-Vercher, M., y Vidal, A. «Mapeo de la inteligencia artificial urbana: primer informe del Atlas de la Inteligencia Artificial Urbana del GOUAI». Global Observatory of Urban AI (GOUAI), 2024 [Fecha de consulta: 02.09.2024]

Galdon, G. «AI Auditing. Proposal for Algo-scores». EDPB, 27 de junio de 2024 [Fecha de consulta: 02.09.2024]

Ghisleni, C. «Artificial Intelligence and Urban Planning: Technology as a Tool for City Design». *ArchDaily*, 8 de febrero de 2024 [Último acceso: 02.09.2024]

Gobierno de Holanda. «Impact Assessment Fundamental Rights and Algorithms», 31 de marzo de 2022 [Fecha de consulta: 02.09.2024].

Grupo de expertos de alto nivel sobre inteligencia artificial. «Directrices éticas para una IA fiable», abril de 2019 [Fecha de consulta: 02.09.2024]

Grupo de Trabajo del Art. 29. «Guidelines on Data Protection Impact Assessments (DPIA) and determining whether processing is “likely to result in a high risk” for the purposes of Regulation 2016/679», Bruselas, 2017 [Fecha de consulta: 02.09.2024]

Gutiérrez, J. D. y Muñoz-Cadena, S. «Algorithmic transparency in the public sector: A state-of-the-art report of algorithmic transparency instruments». *Global Partnership on Artificial Intelligence*, mayo de 2024 [Fecha de consulta: 02.09.2024]

Hadec, J., Di Leo, M. y Kotsev, A. «AI generated synthetic data in policy applications». *Science for Policy Brief*, European Commission, Joint Research Center, 2024 [Fecha de consulta: 02.09.2024]

Imdat As, P.; Basu, P. y Talwar, P. *Artificial Intelligence in Urban Planning and Design. Technologies, Implementation, and Impacts*. Amsterdam: Elsevier, 2022.

Instituto Danés de Derechos Humanos. «Guidance on Human Rights Impact Assessment of Digital Activities». Copenhagen, 2020 [Fecha de consulta: 02.09.2024]

Madiega, T. y Van De Pol, A. L. «Artificial intelligence act and regulatory sandboxes». European Parliamentary Research Service (EPRS), PE 733.544, junio de 2022 [Fecha de consulta: 02.09.2024]

Manzoni, M. et al. *AI Watch. Road to the adoption of Artificial Intelligence by the public sector*. Oficina de Publicaciones de la UE. Luxemburgo, 2022. JRC129100 [Fecha de consulta: 02.09.2024]

Mökander, J. «Auditing of AI: Legal, Ethical and Technical Approaches». *Digital Society*, vol. 2, artículo núm. 49, 2023 [Fecha de consulta: 02.09.2024]

Organización de Naciones Unidas (ONU). «Recommendation on the Ethics of Artificial Intelligence», 23 de noviembre de 2021 [Fecha de consulta: 02.09.2024]

Organización para la Cooperación y el Desarrollo Económico (OCDE) (1). «Advancing Accountability in AI: governing and managing risks throughout the lifecycle for trustworthy AI», *OECD Artificial Intelligence Papers*, núm. 349, 23 de febrero de 2023 [Fecha de consulta: 02.09.2024]

Organización para la Cooperación y el Desarrollo Económico (OCDE) (2). «Governing with artificial intelligence: Are governments ready?», *OECD Digital Economy Papers*, núm. 20, junio de 2024 [Fecha de consulta: 02.09.2024]

Pascual, M. G. «La Policía española ya usa en sus investigaciones un sistema automático de reconocimiento facial». *El País*, 28 de mayo de 2024 [Fecha de consulta: 02.09.2024]

Pellegrin, J., Colnot, L. y Delponte, L. «Artificial Intelligence and Urban Development». Research for REGI Committee, European Parliament, Policy Department for Structural and Cohesion Policies, Bruselas, 2021 [Fecha de consulta: 02.09.2024]

Tangi, L. et al. «Artificial Intelligence for Interoperability in the European Public Sector: an exploratory study», Oficina de Publicaciones de la UE, Luxemburgo, 2023 [Fecha de consulta: 02.09.2024]

Tangi, L. et al. «AI Watch. European Landscape on the Use of Artificial Intelligence by the Public Sector», Oficina de Publicaciones de la UE, Luxemburgo, 2022 [Fecha de consulta: 02.09.2024]

Timan, T., Van Veenstra, A. F. y Bodea, G. «Artificial Intelligence and public services». Policy Department for Economic, Scientific and Quality of Life Policies, PE 662.936, julio de 2021 [Fecha de consulta: 02.09.2024]

Véliz, C. *Privacidad es poder: Datos, vigilancia y libertad en la era digital*. Madrid: Debate, 2021.

Verhulst, S. G. «Are we entering a Data Winter? On the urgent need to preserve data access for the public interest». *Frontiers Policy Labs*, 2024 [Fecha de consulta: 02.09.2024]

Yan, Z. et al. «Intelligent urbanism with artificial intelligence in shaping tomorrow's smart cities: current developments, trends, and future directions». *Journal of Cloud Computing*, 18 de diciembre de 2023 [Fecha de consulta: 02.09.2024]

Leandry Junior Jieutsa

*Investigador sobre la gobernanza de la IA en las ciudades,
Cátedra UNESCO de Paisaje Urbano*

1. Introducción

La equidad y la no discriminación son valores fundamentales de la inteligencia artificial (IA) urbana en ciudades inteligentes centradas en las personas. El creciente debate entre investigadores y responsables políticos atestigua la importancia cada vez mayor de abordar la parcialidad y la discriminación en los sistemas de IA. La equidad se deriva del juicio moral, es decir, el proceso por el cual los individuos determinan lo que es moralmente correcto o incorrecto (Weinkauf, 2023). Aunque la IA ofrece muchas ventajas a las ciudades, su implementación pone a prueba la búsqueda de una urbe equitativa al crear o reforzar la discriminación y las desigualdades. Así pues, la integración de los principios de equidad y no discriminación en el ciclo de vida de la IA urbana es crucial para garantizar el bienestar de las personas y las comunidades en las ciudades inteligentes. No obstante, la puesta en práctica de este principio sigue siendo compleja y ambigua. Para lograrlo, las ciudades necesitan articular sus diversas funciones en la gobernanza de la IA, ya sea en calidad de desarrolladoras de soluciones internas, responsables de la implementación de sistemas externos o reguladoras. Esto requiere la adopción de diversos mecanismos, como la innovación sociotécnica, el establecimiento de normas locales para la equidad en la IA y estándares de contratación. Además, debe introducirse una legislación urbana que proteja a las personas más vulnerables y garantice a la ciudadanía el ejercicio de sus derechos digitales. Sin embargo, estas medidas requieren recursos, que las ciudades pueden movilizar fomentando la cooperación y el trabajo en red.

2. Un sistema de IA equitativo debe ser imparcial y utilizarse de forma responsable

La equidad y la no discriminación son conceptos complejos y cruciales en la sociedad contemporánea (Barocas *et al.*, 2023). Según el Diccionario de Cambridge, «equidad» se refiere a la cualidad de tratar a los individuos de forma igualitaria y de una manera que sea justa o razonable. El concepto de equidad se basa en el respeto a las personas tanto en su calidad de individuos como de miembros de la sociedad. Este concepto, que se refiere a individuos o grupos, está constituido por tres elementos principales, articu-

La puesta en práctica de este principio sigue siendo compleja y ambigua. Para lograrlo, las ciudades necesitan articular sus diversas funciones en la gobernanza de la IA, ya sea en calidad de desarrolladoras de soluciones internas, responsables de la implementación de sistemas externos o reguladoras.

lados en dimensiones distributiva y socio-relacional (Barocas *et al.*, 2023b): la igualdad equitativa de oportunidades, el derecho a la justificación y la igualdad en las relaciones (Giovanola y Tiribelli, 2022). Una sociedad equitativa requiere considerar a cada individuo o grupo de individuos en función de sus características y circunstancias específicas para garantizar un trato y unos resultados equitativos (Giovanola y Tiribelli, 2022; Lyu *et al.*, 2023). Así, este concepto incorpora la noción de no discriminación, que implica que ninguna persona debe ser excluida. Los individuos o grupos vulnerables son los más susceptibles de sufrir discriminación.

La aparición de tecnologías disruptivas como la inteligencia artificial (IA) desafía las dimensiones de la equidad. En el contexto de la discriminación en relación con la IA intervienen dos factores principales: los sesgos algorítmicos y el uso de sistemas basados en IA (Ferrara, 2023; O'Neil, 2016; Wachter *et al.*, 2021).

El primer factor, el sesgo algorítmico, distorsiona los datos de entrenamiento originales o el algoritmo de IA, dando lugar a resultados parciales y potencialmente perjudiciales (Holdsworth, 2023). Estos sesgos reducen la precisión y el potencial de la IA con diversos grados de impacto en función de la aplicación. En la IA existen dos categorías principales de sesgos: el sesgo de automatización y el sesgo de correlación (Barocas *et al.*, 2023a; González-Sendino *et al.*, 2023). El sesgo de automatización es la propagación a gran escala, a través de los procesos de los sistemas de IA, de sesgos sociales y culturales profundamente arraigados en los datos históricos de entrenamiento utilizados para alimentar el sistema de IA. Esta categoría incluye el sesgo humano, el sesgo de la base de datos, el sesgo de aprendizaje y el sesgo de implementación. El sesgo de correlación se produce cuando la correlación involuntaria de variables protegidas (por ejemplo, el género, la raza) permite inferir sesgos, a pesar de los esfuerzos por excluirlas de los datos de entrenamiento.

El segundo factor es el empleo de sistemas basados en la IA. De hecho, cuando se utilizan para la elaboración de perfiles o el control social, los sistemas vulneran los derechos digitales (Calzada, 2021; Cugurullo *et al.*, 2022). Al recopilar y utilizar información personal, las tecnologías de reconocimiento facial, por ejemplo, violan la privacidad y los datos personales de la ciudadanía (ONU-Hábitat, 2023). Los derechos digitales se interpretan como los derechos humanos existentes que deben protegerse en el contexto de las tecnologías digitales, ya que los espacios físicos y digitales están cada vez más entrelazados (ONU-Hábitat, 2020).

La equidad algorítmica se fundamenta en variables interrelacionadas (Weinkauf, 2023). Un sistema de decisión automatizado se considera equitativo cuando no se basa en datos sensibles como el género o las creencias religiosas, no perjudica a las minorías y se utiliza de forma responsable.

3. El dilema de la IA: equilibrio entre las oportunidades y las repercusiones de los sistemas de IA en las ciudades

Históricamente, la planificación urbana ha contribuido a crear y reforzar diferentes formas de desigualdad y discriminación urbanas (Fainstein, 2009; Hall, 2014). Las poblaciones más afectadas son sobre

todo las minorías y las personas más vulnerables, que varían según el contexto. En consecuencia, han surgido numerosos conceptos, como el «derecho a la ciudad» de Henri Lefebvre o la «ciudad justa» (Fainstein, 2009; Fincher e Iveson, 2012; Harvey y Potter, 2009; Lefebvre, 1968). Estos conceptos tienen por objeto hacer que las ciudades sean más equitativas, en particular mediante el acceso a los servicios y las oportunidades urbanas, con el fin de mejorar la calidad de vida.

La aparición de la IA supone un reto para la ciudad justa, ya que brinda oportunidades para crear ciudades más inclusivas, al tiempo que genera y refuerza distintas formas de desigualdad y discriminación. De hecho, los sistemas de IA poseen la capacidad de filtrar y procesar volúmenes considerables de datos conectados a redes extensas y al entorno urbano. En consecuencia, pueden permitir tomar decisiones complejas de forma autónoma o semiautónoma (Marvin *et al.*, 2022; Sherman, 2023; Yigitcanlar *et al.*, 2021). Las metodologías de IA explicable (XAI, por sus siglas en inglés) pueden ayudar a los municipios a comprender el cálculo de la equidad y su mejora (Lyu *et al.*, 2023). La aplicación de la IA facilita una mayor participación de los ciudadanos en la administración municipal y optimiza la prestación de servicios, especialmente para las poblaciones más vulnerables.

Por ejemplo, las herramientas de aprendizaje profundo mejoran la gestión de datos espaciales para optimizar la prestación de servicios en barrios desfavorecidos de [Durban \(Sudáfrica\)](#). La IA generativa facilita los procesos de planificación participativa a través de la generación de escenarios urbanos en tiempo real, lo que permite una planificación urbana más inclusiva que incorpora diversas perspectivas. Además, los chatbots municipales, como los que se han introducido en [Helsinki \(Finlandia\)](#) o [Saint-Lin-Laurentides \(Canadá\)](#), automatizan la interacción con los ciudadanos. Esto mejora la gestión de los servicios públicos, especialmente para las personas que no están familiarizadas con los procedimientos administrativos, a menudo complejos, o que tienen dificultades para acceder a los servicios en persona.

Sin embargo, como se ha señalado anteriormente, los sistemas de IA y el énfasis en la competitividad económica de las ciudades desafían la ciudad justa al producir resultados no equitativos y discriminatorios. Por otra parte, la discriminación automatizada por algoritmos –a diferencia de la que se produce en las formas tradicionales– es más abstracta u opaca, así como poco intuitiva, sutil, intangible, difícil de detectar y a gran escala (Kleinberg *et al.*, 2018; O’Neil, 2016; Sanchez *et al.*, 2024; Wachter *et al.*, 2021)

Por ejemplo, algoritmos fiscales que singularizaban «nombres que sonaban extranjeros» y personas con «doble nacionalidad» llevaron a [miles de familias racializadas a ser acusadas falsamente de fraude](#) en los Países Bajos. A escala mundial, los sistemas de vigilancia predictiva, como Clearview AI, plantean problemas de privacidad al tiempo que refuerzan los sesgos (Dauvergne, 2022; O’Neil, 2016). En 2021, *Forbes* informó de un [sesgo algorítmico en las solicitudes de hipotecas](#), con un 80% de solicitudes denegadas a personas solicitantes negras. Del mismo modo, *The Markup* (2021) descubrió que las personas solicitantes de color tenían [entre un 40 % y un 80 % más de probabilidades de que se les denegara un préstamo](#), lo que subraya el impacto discriminatorio de la IA.

La discriminación automatizada por algoritmos –a diferencia de la que se produce en las formas tradicionales– es más abstracta u opaca, así como poco intuitiva, sutil, intangible, difícil de detectar y a gran escala.

Debido a la opacidad y a la escala del impacto, resulta difícil para las personas afectadas hacer valer sus derechos [...] ya que los sistemas algorítmicos de toma de decisiones perturban los recursos y procedimientos jurídicos tradicionales destinados a detectar, investigar, prevenir y corregir la discriminación.

Además, la concentración de la riqueza en las grandes ciudades, como consecuencia de la IA urbana, conduce a la gentrificación urbana (Sánchez *et al.*, 2024). El acceso a los servicios, en particular a la vivienda, para los habitantes de escasos recursos es cada vez más difícil, cuando no imposible. Las políticas de implementación de IA urbana contribuyen así a reforzar las asimetrías entre territorios y las desigualdades urbanas.

Por tanto, los sistemas de IA tienen importantes repercusiones en las ciudades y las sociedades. Esta ambivalencia plantea la necesidad de una gobernanza eficaz. Además, debido a la opacidad y a la escala del impacto, resulta difícil para las personas afectadas defenderse o hacer valer sus derechos. Esta situación pone en entredicho el derecho a la no discriminación de que goza la ciudadanía, ya que los sistemas algorítmicos de toma de decisiones perturban los recursos y procedimientos jurídicos tradicionales destinados a detectar, investigar, prevenir y corregir la discriminación (Wachter *et al.*, 2021).

4. Recomendaciones políticas

De acuerdo con las recomendaciones de la UNESCO, los actores implicados en IA deben adoptar un enfoque inclusivo dirigido a hacer que los beneficios de las tecnologías de IA estén disponibles y sean accesibles para todas las personas, teniendo en cuenta las necesidades específicas de los diferentes grupos (UNESCO, 2023). A nivel municipal, la aplicación de la equidad y la no discriminación en los sistemas urbanos de IA requiere la articulación de las diversas funciones que asumen los municipios como desarrolladores de soluciones internas (aunque con relativa poca frecuencia debido a limitaciones financieras y técnicas), implementadores y reguladores. Además, la mejora de la equidad de los sistemas de IA implica tener en cuenta todo su ciclo de vida, abordando diversos aspectos a lo largo de los procesos de diseño, desarrollo y aplicación. Asimismo, la aportación de soluciones eficaces a las disparidades en los resultados de los sistemas de IA comienza con la identificación de sus causas subyacentes.

4.1. Recomendaciones generales:

- **Definir una estrategia:** Las ciudades deben aplicar estrategias de IA estructuradas en torno a los principios de equidad y no discriminación. Estos documentos estratégicos permiten a las ciudades dotarse de una base sólida y realizar una evaluación precisa de sus objetivos relacionados con la IA. Este enfoque es esencial para planificar la integración de la IA con el fin de maximizar sus beneficios y al mismo tiempo mitigar los posibles riesgos. Estas estrategias deben elaborarse mediante un proceso participativo e ir acompañadas de planes de acción que definan medidas concretas para garantizar una integración equitativa de la IA que no deje a ninguna persona atrás.
- **Establecer niveles de riesgo en función de las aplicaciones:** Las ciudades deben identificar las aplicaciones de IA de alto riesgo dentro de sus jurisdicciones, teniendo en cuenta las disparidades y desigualdades existentes en el territorio. A la identificación de estas aplicaciones de alto riesgo debe seguir la aplicación de mecanismos de

protección. Las aplicaciones relacionadas con servicios sociales esenciales deberían clasificarse como de alto riesgo y se debería prohibir que funcionen de forma completamente autónoma. Por ejemplo, la [ciudad de San José](#) ha puesto en marcha un registro de IA articulado en torno a una evaluación rigurosa de los sistemas de IA. Este proceso implica un análisis de riesgos, seguido de una evaluación del impacto más exhaustiva, en función del nivel de riesgo, todo ello documentado a través de una «Hoja de impacto» y una «Hoja informativa sobre IA».

4.2. Recomendaciones específicas para las ciudades como desarrolladoras de soluciones internas

- **Hacer hincapié en la innovación sociotécnica inclusiva.** Incorporar a diversas partes interesadas no técnicas a lo largo del ciclo de vida de la IA. Según ONU-Hábitat, dicho ciclo de vida comprende cinco fases: encuadre, diseño, aplicación, implementación y mantenimiento. Cuando las decisiones en estas distintas fases son tomadas principalmente por actores del ámbito técnico o grupos homogéneos, existe un riesgo importante de que sus prejuicios se integren en el sistema. Este riesgo es especialmente pronunciado si la herramienta se aplica posteriormente o se generaliza a segmentos de población más amplios. Los gobiernos locales deben hacer mayor hincapié en la interdisciplinariedad y la multidisciplinariedad para integrar a los distintos grupos sociales en el ciclo de vida de la IA urbana.
- **Aplicar técnicas equitativas,** como el preprocesamiento de datos (que consiste en identificar y abordar sesgos en los datos antes del entrenamiento del modelo), la selección de modelos (que se centra en la utilización de métodos de selección que prioricen la equidad) y las decisiones posteriores al procesamiento (que consiste en ajustar el resultado de los modelos de IA para mitigar los sesgos y garantizar la equidad) (Ferrara, 2023).
- **Aumentar la diversidad en la construcción de bases de datos abarcando tres dimensiones: equipos, datos y modelos.** Para minimizar los sesgos, es fundamental crear equipos diversos e interdisciplinarios y aplicar una formación continua en materia de equidad y ética. Respecto a los datos, una mejor recopilación de atributos sensibles (por ejemplo, sexo, raza, etnia) junto con la documentación de las decisiones relacionadas con los datos promueven la transparencia y permiten abordar con mayor facilidad las desigualdades en el mundo real. En lo que concierne a los modelos, ofrecer un acceso abierto a la comunidad para la realización de pruebas, garantizar la transparencia de la documentación y utilizar IA explicable puede ayudar a identificar y mitigar los sesgos, garantizando así resultados equitativos (González-Sendino *et al.*, 2023).
- **Integrar la correlación compensatoria en los sistemas de IA.** Como indican Giovanola y Tiribelli (2022), garantizar una justa igualdad de oportunidades en los sistemas de IA no puede limitarse a eliminar los sesgos discriminatorios en los datos de entrenamiento. Los sistemas de IA urbana deben diseñarse teniendo en cuenta las desigualdades existentes en sus respectivos contextos e incorporar mecanismos para compensarlas. Por ejemplo, en una ciudad donde

existen disparidades entre comunidades o grupos sociales, las IA urbanas deben tener en cuenta estas disparidades y aplicar medidas compensatorias. Esto puede manifestarse, por ejemplo, en forma de contenidos personalizados.

- **Integrar técnicas de mitigación en el ciclo de vida de la IA.** Neutralizar efectos discriminatorios presentes en los datos durante la fase de preentrenamiento mediante métodos como el remuestreo (alteración del tamaño del conjunto de datos que afecta a la distribución sin transformar los datos), la representación equitativa (conseguida eliminando la información que puede asociar a un individuo con un grupo protegido) y la reponderación (utilizada para transformar los datos modificando su peso en el conjunto de datos). Durante la fase de formación, emplear la regularización y la formación adversarial, que son los métodos más comunes para este fin. Otros enfoques emergentes son el aprendizaje descentralizado, la regresión lineal equitativa, el método DeepFair, los modelos multimodales y la agrupación de fair-lets. Durante la fase posterior al entrenamiento, aplicar probabilidades igualadas, probabilidades igualadas calibradas y rechazar la clasificación de opciones.

4.3. Recomendaciones específicas para las ciudades como implementadores y reguladores

- **Establecer normas locales para una IA equitativa.** La discriminación y las desigualdades pueden manifestarse de forma diferente según el contexto, afectando a individuos o grupos sociales de diversas maneras y a diferentes escalas. Por tanto, las ciudades deben aplicar normas de equidad para la IA urbana que tengan en cuenta estas especificidades locales. Estas normas deben incorporar principios generales integrando al mismo tiempo consideraciones locales. El objetivo es garantizar que la IA urbana no refuerce la discriminación existente ni cree nuevas formas de prejuicio que afecten desproporcionadamente a las personas más vulnerables. Estas normas deben elaborarse consultando a las comunidades locales y abarcar todo el ciclo de vida de la IA.
- **Establecer normas de contratación para una IA equitativa.** Las ciudades deben asegurarse de que las entidades que les prestan servicios se ajustan a los principios de equidad. Para ello es necesario establecer mecanismos de contratación que obliguen a los proveedores de servicios a cumplir las normas de IA equitativa de la ciudad. Los proveedores deben cumplir los requisitos de conformidad relativos a sus algoritmos como condición para que la ciudad los utilice. Por ejemplo, [la iniciativa GOV AI](#), una coalición impulsada por la ciudad estadounidense de San José, ha adoptado e introducido la mencionada «Hoja informativa sobre IA» para sistemas procedentes de terceros. Se trata de una plantilla armonizada para que los proveedores faciliten información detallada sobre sus productos de IA, que abarca aspectos como la finalidad del sistema, los datos de entrenamiento, los detalles del modelo, las métricas de rendimiento, la gestión de sesgos, la solidez y la interacción entre los seres humanos y el ordenador.
- **Aplicar una legislación urbana que garantice el derecho a la justificación.** Este derecho permite a las personas afectadas por un sistema

de IA entender el razonamiento que subyace a una decisión algorítmica, lo que permite a la ciudadanía comprender y controlar cómo es tratada por estos sistemas. Cuando este derecho no se respeta adecuadamente, las personas deben tener la capacidad de cuestionar y modificar los parámetros subyacentes a la decisión. Por tanto, las ciudades deben considerar, a lo largo de todo el proceso, si implementan o retiran un sistema de IA, en particular si no es posible satisfacer la petición de explicación de una persona. Esta medida permite a los individuos que sufren discriminación reivindicar sus derechos digitales.

- **Crear órganos consultivos para investigar, prevenir y mitigar los posibles usos malintencionados de la IA.** Los gobiernos locales deben crear órganos consultivos multidisciplinares que incluyan organizaciones comunitarias, el mundo académico, las empresas y otras partes interesadas. Estos órganos desempeñarán un papel de auditoría para limitar la discriminación algorítmica a través de la evaluación de los modelos de IA de la ciudad basándose en parámetros de equidad. Los objetivos de su evaluación serán los siguientes: en primer lugar, identificar posibles sesgos que puedan afectar a la equidad, en segundo lugar seleccionar parámetros para medir la equidad de los sistemas de IA y, en tercer lugar, mitigar el impacto causado por dichos sesgos. Además, dichos órganos consultivos actuarán para orientar a las ciudades en sus acciones y políticas relativas a una IA equitativa.

5. Limitaciones

Lograr la equidad en la IA es complejo. Las intervenciones orientadas a alcanzar la equidad en la IA urbana pueden crear tensiones con los propios objetivos de los mismos algoritmos. Esto implica que las ciudades deben adoptar un enfoque basado en el compromiso para equilibrar las ganancias y los beneficios, priorizando al mismo tiempo el bienestar de las personas y las comunidades. Sin embargo, este principio puede parecer abstracto y dar lugar a interpretaciones divergentes, lo que complica la operacionalización de medidas de éxito e impacto (Sadek *et al.*, 2024). Por lo tanto, las ciudades deben aplicar un enfoque local para hacer operativas la equidad y la no discriminación en la IA urbana. Este enfoque holístico tiene en cuenta la configuración socioeconómica y cultural de la ciudad a lo largo de todo el ciclo de vida de la IA.

Desde una perspectiva técnica, una IA urbana equitativa requiere recursos humanos diversos e infraestructuras adaptadas (Du *et al.*, 2023; Marvin *et al.*, 2022; Yigitcanlar *et al.*, 2020, 2023). Esto requiere, a su vez, importantes inversiones financieras (Bettoni *et al.*, 2021). Asimismo, las ciudades deben hacer frente a costes adicionales para la formación y educación continuas del personal y las comunidades (Sadek *et al.*, 2024; Varanasi, 2023) y prever costes legales y de cumplimiento, incluidas auditorías y ajustes del sistema para satisfacer normas reglamentarias. Estas inversiones pueden representar gastos considerables, especialmente para las ciudades pequeñas y medianas.

Para superar estas limitaciones, las ciudades pueden recurrir a *networking*. Estas redes ofrecen oportunidades para el intercambio de conocimientos, la innovación política y las respuestas coordinadas a problemas globales. Algunos ejemplos al respecto son:

Cities Coalition for Digital Human Rights: una plataforma para promover un desarrollo inclusivo y democrático de las nuevas tecnologías en las ciudades.

City AI Connect: una comunidad de aprendizaje global y una plataforma digital para que las ciudades ensayen y avancen en el uso de la inteligencia artificial generativa para mejorar los servicios públicos.

GovAI: una coalición compuesta por más de 1.000 miembros y más de 350 entidades locales, estatales y federales unidas en la misión de promover una IA responsable y orientada en el sector público.

AI4Cities: un proyecto que permitió a Helsinki, Ámsterdam, Copenhague, el Gran París, Stavanger y Tallin desafiar al mercado a proponer soluciones basadas en IA para reducir emisiones de CO2 en los ámbitos de la energía y la movilidad.

Referencias bibliográficas

Barocas, S., Hardt, M., & Narayanan, A. «Classification». En *Fairness and Machine Learning : Limitations and Opportunities*. MIT Press, 2023a

Barocas, S., Hardt, M., & Narayanan, A. «Relative notions of fairness». En *Fairness and Machine Learning: Limitations and Opportunities*. MIT Press, 2023b

Bettoni, A. et al. «An AI adoption model for SMEs: A conceptual framework». *IFAC-PapersOnLine*, 2021, 54(1), pp. 702708.

Calzada, I. «The right to have digital rights in smart cities». *Sustainability (Switzerland)*, 2021, 13(20).

Cugurullo, F. et al. «Urban AI in China: Social control or hyper-capitalist development in the post-smart city? ». *Frontiers in Sustainable Cities*, 2022

Dauvergne, P. «Facial recognition technology for policing and surveillance in the Global South: A call for bans». En *Third World Quarterly*, 2022, (Vol. 43, Número 9, p. 23252335). Routledge.

Du, J. et al. «Artificial intelligence enabled participatory planning: A review». *International Journal of Urban Sciences*, 2023

Fainstein, S. «Planning and the Just City». En P. Marcuse (Ed.), *Searching for the just city: Debates in urban theory and practice*. Routledge, 2009

Ferrara, E. «Fairness and bias in artificial intelligence: A brief survey of sources, impacts, and mitigation strategies». *Sci*, 2023, 6(1), 3.

Fincher, R., & Iveson, K. «Justice and injustice in the city». *Geographical research*, 2012, 50(3), 231241.

Giovanola, B., & Tiribelli, S. «Weapons of moral construction? On the value of fairness in algorithmic decision-making». *Ethics and Information Technology*, 2022, 24(1), 3.

González-Sendino, R. et al. «A review of Bias and Fairness in Artificial Intelligence». *International Journal of Interactive Multimedia and Artificial Intelligence*, In press, 2023

Hall, P. «Cities of tomorrow: An intellectual history of urban planning and design since 1880(Fourth edition)». *Wiley-Blackwell*, 2014

Harvey, D., & Potter, C. «The right to the Just City». En P. Marcuse (Ed.), *Searching for the just city: Debates in urban theory and practice*. *Routledge*, 2009

Holdsworth, J. «[What is AI bias?](#) », 2023 [Fecha de consulta: 20 enero 2024]

Kleinberg, J. et al. «Discrimination in the Age of Algorithms». *Journal of Legal Analysis*, 2018, 10(2005), pp. 113174.

Lefebvre, H. «[The Right to the City](#)», 1968 [Fecha de consulta: 20 enero 2024]

Lyu, Y. et al. «IF-City: Intelligible fair city planning to measure, explain and mitigate inequality». *IEEE Transactions on Visualization and Computer Graphics*, 2023

Martinez, E., & Kirchner, L. «[The secret bias hidden in Mortgage-Approval algorithms](#)», 2021 [Fecha de consulta: 20 enero 2024]

Marvin, S. et al. «Urban AI in China: Social control or hyper-capitalist development in the post-smart city? ». *Frontiers in Sustainable Cities*, 2022, 4, 1030318.

O'Neil, C. «Weapons of math destruction: How big data increases inequality and threatens democracy (First edition) ». *Crown*, 2016

Rezende, I. N. «Facial recognition in police hands: Assessing the 'Clearview case' from a European perspective». *New Journal of European Criminal Law*, 2020, 11(3), pp. 375389.

Sadek, M. et al. «Challenges of responsible AI in practice: Scoping review and recommended actions». *AI & SOCIETY*, 2024

Sanchez, T. W., Brenman, M., & Ye, X. «The Ethical Concerns of Artificial Intelligence in Urban Planning». *Journal of the American Planning Association*, 2024, 0(0).

Sherman, S. «The Polyopticon: A diagram for urban artificial intelligences». *AI and Society*, 2023, 38(3), pp. 12091222.

UNESCO. «Readiness assessment methodology: A tool of the recommendation on the ethics of artificial intelligence». *UNESCO*, 2023

UN-Habitat. «[Mainstreaming human rights in the digital transformation of cities-A guide for local governments](#)». *United Nations Human Settlements Programme*, 2020. [Fecha de acceso 20 enero 2024]

UN-Habitat. «[Human rights in the digital era](#)». United Nations Human Settlements Programme, 2023, pp. 156. [Fecha de consulta: 20 enero 2024]

Varanasi, R. A. «“It is currently hodgepodge”: Examining AI/ML Practitioners’ Challenges during Co-production of Responsible AI Values», 2023

Wachter, S., Mittelstadt, B., & Russell, C. «Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI». *Computer Law & Security Review*, 2021, 41, 105567.

Weinkauf, D. «[Privacy Tech-Know Blog: When Worlds Collide – The Possibilities and Limits of Algorithmic Fairness \(Part 1\)](#)». *Office of the Privacy Commissioner of Canada*, 2023. [Fecha de consulta: 20 enero 2024]

Yigitcanlar, T. *et al.* «Artificial intelligence in local governments: Perceptions of city managers on prospects, constraints and choices». *AI and Society*, 2023, 38(3), pp. 11351150.

Yigitcanlar, T. *et al.* «Responsible urban innovation with local government artificial intelligence (Ai): A conceptual framework and research agenda». *Journal of Open Innovation: Technology, Market, and Complexity*, 2021, 7(1), pp. 116.

Yigitcanlar, T. *et al.* «Contributions and Risks of Artificial Intelligence (AI) in Building Smarter Cities: Insights from a Systematic Review of the Literature», 2020

María Pérez-Ortiz

*Profesora titular, Centro de Inteligencia Artificial
de University College London (Reino Unido)*

1. Introducción

En 2050, las Naciones Unidas prevén que casi el 70% de la población mundial vivirá en zonas urbanas (frente al 56% actual [ONU-Hábitat, 2022]). La urbanización del planeta avanza a un ritmo sin precedentes, por lo que los retos que enfrentan las ciudades –como la movilidad, la provisión de servicios y vivienda, la contaminación y la salud urbana, y el uso de los recursos– aumentan de manera cada vez más acuciante. Esta gran velocidad del crecimiento urbano conlleva la necesidad urgente de desarrollar soluciones innovadoras para garantizar que las ciudades sean habitables, favorables al desarrollo humano, eficientes y respetuosas con el medioambiente.

Uno de los avances más debatidos a la hora de abordar estos desafíos es la incorporación de la tecnología, en concreto la inteligencia artificial (IA), en los entornos urbanos. No obstante, a medida que se generaliza su adopción, surgen motivos de preocupación por los efectos que tiene para la sostenibilidad, tanto ambiental como social. En el presente texto se examina el concepto de IA sostenible, prestando especial atención al papel que debería desempeñar en la implementación de este tipo de tecnologías en los entornos urbanos. Además, se analizan los factores económicos, sociales y ambientales de la utilización de la IA en las ciudades, poniendo de relieve las ventajas, desafíos y tendencias de esta tecnología en la búsqueda de unos futuros urbanos equitativos y sostenibles.

2. La búsqueda de la sostenibilidad

La sostenibilidad, en particular la relativa al desarrollo, tal y como la definen actualmente las Naciones Unidas (Keeble, 1988), se refiere a la capacidad de lograr que el desarrollo satisfaga las necesidades humanas presentes y futuras (como la salud y bienestar, la educación de calidad, el trabajo digno y la igualdad social) y de conseguirlo dentro de los límites socioecológicos actuales. La sostenibilidad suele separarse en tres pilares:

La IA sostenible abarca el uso de estas tecnologías en favor de la sostenibilidad económica, ambiental y social, pero también de la sostenibilidad de la propia IA, y aborda diversos motivos de preocupación, como el consumo de energía, el uso de los recursos y la equidad social.

- **Sostenibilidad ambiental:** vivir con los recursos naturales de los que disponemos, y proteger y mantener nuestros ecosistemas. En los entornos urbanos, un desafío clave es reducir las emisiones de gases de efecto invernadero y mejorar la calidad del aire. Por ejemplo, las ciudades pueden implementar zonas de bajas emisiones, ampliar el transporte público y promover edificios eficientes desde el punto de vista energético para reducir la contaminación urbana y atenuar el efecto de isla térmica.
- **Sostenibilidad social:** lograr de forma sistemática un nivel adecuado de bienestar social. En las ciudades, esto suele significar fomentar la inclusión, la accesibilidad, la equidad y un modo de vida más sostenible en los procesos de urbanización. Un desafío actual es generar viviendas asequibles y garantizar el acceso equitativo a los espacios verdes, el transporte y los servicios básicos. Por ejemplo, las ciudades pueden proyectar barrios de usos múltiples en los que se pueda caminar, dar prioridad a la integración social y ajardinar zonas con especies vegetales autóctonas para incorporar entornos más ecológicos.
- **Sostenibilidad económica:** utilizar los recursos con eficiencia y responsabilidad. La dimensión económica de la sostenibilidad se centra en las economías circulares, ecológicas y justas, y no simplemente en el crecimiento económico constante. En el contexto de la sostenibilidad urbana, la atención se centra en el desarrollo de economías resilientes y con bajas emisiones de carbono. Un ejemplo de esto es el fomento de infraestructuras urbanas sostenibles que favorecen el empleo verde y los principios de la economía circular, como la creación de sistemas de movilidad urbana sostenible (p. ej., el uso compartido de bicicletas o la infraestructura de recarga de vehículos eléctricos) y el apoyo a las empresas ecológicas locales.

Pese a que el concepto de sostenibilidad suele reducirse a su pilar ambiental, los tres ejemplos anteriores de sostenibilidad económica, ambiental y social ponen de manifiesto la interconexión de todos estos elementos y la necesidad de examinarlos juntos. Como afirmó el naturalista John Muir: “Cuando intentamos seleccionar algo por sí mismo, nos damos cuenta de que está atado a todo lo demás en el Universo”. (Muir, 1911). Del mismo modo, las ciudades son sistemas complejos que requieren de análisis del sistema en su conjunto y enfoques que aborden objetivos múltiples y tomen en cuenta los pilares económico, ambiental y social.

3. La tecnología se encuentra con la sostenibilidad

¿Podrían los avances tecnológicos influir en nuestra capacidad para garantizar un futuro sostenible? (GACGC, 2019). Resulta fundamental comprender los efectos que la tecnología tiene en la sostenibilidad, sobre todo porque las tecnologías radicalmente innovadoras, como la IA, aceleran el cambio global y dirigido por el ser humano a un ritmo sin precedentes y sin una agenda unificada clara.

En los últimos años ha aparecido un conjunto nuevo de principios deseados para la sostenibilidad en relación con los sistemas de IA (Vinuesa, 2020; Van Wynsberghe, 2021). En concreto, la IA sostenible

hace referencia a un marco en rápida evolución que pretende configurar el desarrollo, la implementación y el uso de esta tecnología de manera responsable desde el punto de vista ético, social y ambiental, y aspira a encontrar un equilibrio entre los riesgos y las oportunidades que se presentan.

En esencia, la IA sostenible también considera las tres dimensiones principales de la sostenibilidad: la social, la ambiental y la económica. No obstante, es más que la suma de sus partes: la IA sostenible abarca el uso de estas tecnologías en favor de la sostenibilidad económica, ambiental y social, pero también –y es importante señalarlo– de la sostenibilidad de la propia IA (Van Wynsberghe, 2021), y aborda diversos motivos de preocupación, como el consumo de energía, el uso de los recursos y la equidad social, para garantizar que las soluciones de IA supongan con el tiempo una contribución positiva a la sociedad y el medioambiente.

Uno de los desafíos principales para desarrollar IA sostenible es afrontar su propio impacto ambiental. Los sistemas de IA, sobre todo los modelos de gran escala como los utilizados en el aprendizaje profundo, precisan una gran capacidad de cálculo, lo que se traduce en un nivel considerable de consumo de energía y emisiones de carbono. Por ejemplo, en 2019 se calculaba que el entrenamiento de un solo modelo grande de lenguaje emitía tanto carbono como cinco automóviles a lo largo de su vida útil (Strubell *et al.*, 2019), estimación que probablemente haya aumentado notablemente desde entonces. Este cálculo incluye únicamente el entrenamiento, sin tomar en cuenta los efectos de su uso (p. ej., se calcula que hacer una imagen con IA generativa consume tanta energía como cargar el teléfono [Heikkilä, 2023]). Además de las emisiones operativas, es importante tener en cuenta las incorporadas, es decir, la huella de carbono asociada con la fabricación, el transporte y el desecho del *hardware* utilizado para la IA, como los servidores y las unidades de procesamiento de gráficos (GPU, por sus siglas en inglés). La IA sostenible trata de reducir al mínimo las repercusiones ambientales, tanto operativas como incorporadas, promoviendo algoritmos más eficientes desde el punto de vista energético tales como la destilación o la cuantificación de modelos.

Otra preocupación acuciante, además de las consecuencias ambientales, es la sostenibilidad social de la IA. Si no se conciben con cuidado, los sistemas de IA pueden reforzar las desigualdades sociales existentes mediante algoritmos sesgados o un acceso desigual a este tipo de tecnologías. La IA sostenible defiende la creación de sistemas que promuevan la inclusión y la justicia sociales y garanticen que los grupos marginados no se vean perjudicados por decisiones basadas en esta tecnología en ámbitos tales como la contratación de personal, la vivienda o la justicia penal. Esto implica integrar aspectos éticos y los derechos humanos en el diseño y aplicación de la IA, con unos marcos sólidos en materia de transparencia, rendición de cuentas y gobernanza que ofrezcan protección frente a los daños involuntarios. La imparcialidad de los sistemas de IA se analiza más a fondo en el capítulo de esta monografía que trata sobre la equidad.

En lo que respecta a la sostenibilidad económica, la IA tiene potencial tanto para impulsar como para obstaculizar el desarrollo sostenible. Esta

La IA sostenible no consiste únicamente en disponer de un subconjunto de tecnologías concebidas expresamente para favorecer la sostenibilidad, sino en reestructurar todo el campo de la IA para garantizar que contribuya sistemáticamente al bienestar económico, ambiental y social a largo plazo.

Las tecnologías de IA de las ciudades tienen efectos significativos en el medioambiente debido a su dependencia de los centros de datos y los recursos informáticos. Un enfoque sostenible de la IA urbana implicaría la optimización de la eficiencia energética de estos sistemas.

tecnología puede optimizar las infraestructuras, las redes de energía y los sistemas de transporte urbano, y de este modo fomentar un uso más eficiente de los recursos y reducir el despilfarro en las ciudades. También puede facilitar la transición a una economía circular mejorando procesos tales como la gestión de las cadenas de suministro, la reducción de los residuos y la optimización del ciclo de vida de los productos. Sin embargo, si no se gestiona con cuidado, la IA puede intensificar la desigualdad económica automatizando puestos de trabajo sin crear nuevas oportunidades de empleo o aumentando la concentración de riqueza. La IA sostenible, por tanto, defiende modelos económicos basados en la IA que dan prioridad a los beneficios sociales a largo plazo frente al corto plazo y garantizan que las ganancias económicas procedentes de esta tecnología se distribuyen de forma equitativa.

Los marcos de IA sostenible proponen que se incorporen en el diseño de los sistemas de IA preguntas como (Vinuesa, 2020; Van Wynsberghe, 2021):

- ¿Qué equilibrio hay entre las repercusiones directas e indirectas que la tecnología de IA tiene en la sociedad, el medioambiente y la economía? ¿Cómo podemos diseñar sistemas de IA más sostenibles desde el principio? ¿Qué marcos de evaluación de riesgos pueden ayudarnos a prever las consecuencias no deseadas antes de que se produzcan?
- ¿Cómo podemos abordar el sistema sociotécnico más general que rodea a la IA, especialmente las repercusiones sociales para las personas que la utilizan o se ven afectadas por ella? ¿Qué medidas podemos tomar para desarrollar una IA que sea compatible con la preservación de los recursos ambientales para las generaciones actuales y futuras, favorezca modelos económicos sostenibles y respete los valores sociales fundamentales de las distintas comunidades?
- ¿Cómo podemos promover cambios a lo largo de todo el ciclo de vida de la IA –desde la generación de ideas, el entrenamiento y el perfeccionamiento hasta la evaluación, la aplicación y la gobernanza– orientados a una mayor sostenibilidad ecológica y equidad social? ¿Qué medidas son necesarias para garantizar el funcionamiento de los sistemas de IA dentro de los límites ecológicos del planeta en relación con el consumo de energía, el uso de agua dulce y la dependencia de minerales escasos, entre otros factores?

En última instancia, la IA sostenible defiende un enfoque en el que esta tecnología sea un factor multiplicador de fuerzas para lograr los objetivos de sostenibilidad y potencie los esfuerzos para mitigar el cambio climático, reducir la desigualdad y fomentar unas economías resilientes e inclusivas, al tiempo que vela por que el desarrollo y el uso de las propias tecnologías de IA respondan a los principios de sostenibilidad. Es importante reconocer que las tecnologías, incluida la IA, pueden representar riesgos extrínsecos e intrínsecos para la sostenibilidad. En el caso de los riesgos intrínsecos, e incluso cuando no se aplican directamente a los desafíos de la sostenibilidad, las tecnologías defectuosas, sin solidez o inequitativas pueden menoscabar involuntariamente los objetivos de sostenibilidad al intensificar la desigualdad, los daños ambientales o la inestabilidad económica por vías indirectas. Por tanto, la IA sostenible no consiste únicamente en disponer de un

subconjunto de tecnologías concebidas expresamente para favorecer la sostenibilidad, sino en reestructurar todo el campo de la IA para garantizar que contribuya sistemáticamente al bienestar económico, ambiental y social a largo plazo.

4. IA sostenible en entornos urbanos

Aunque están empezando a aparecer nuevos marcos para la sostenibilidad de la IA (Vinuesa, 2020; Van Wynsberghe, 2021; Wu, 2022; Wilson, 2022; Nishant, 2020), actualmente hay muy pocos trabajos centrados en futuros urbanos (Yigitcanlar, 2020; Bibri, 2021; Pastor Escuredo, 2022). En este sentido, el presente texto tiene por objeto exponer un punto de vista que ilustra los fundamentos que serían necesarios para garantizar que los sistemas de IA implantados en las ciudades sean no solo tecnológicamente avanzados, sino también responsables, equitativos y beneficiosos tanto para el medioambiente como para las poblaciones urbanas.

Repercusiones ambientales de los sistemas de IA urbana. Las consecuencias ambientales, incluido –entre otros elementos– el elevado nivel de consumo de energía y uso de agua dulce, constituyen uno de los grandes desafíos de la IA (Luccioni, 2024). En entornos urbanos, donde la IA se utiliza cada vez más en aplicaciones tales como los sistemas de tráfico inteligentes, las redes de energía y la administración de edificios, el total acumulativo de las demandas de estos sistemas puede llegar a ser considerable. Las tecnologías de IA de las ciudades tienen efectos significativos en el medioambiente debido a su dependencia de los centros de datos y los recursos informáticos. Un enfoque sostenible de la IA urbana implicaría la optimización de la eficiencia energética de estos sistemas con métodos tales como el aprendizaje automático en dispositivos pequeños, las prácticas de computación ecológica en materia de ingeniería de *software*, la destilación del conocimiento, y la poda o la cuantificación de modelos. Esto también supondría aprender de conjuntos de datos más pequeños y de calidad (es decir, hacer más con menos), utilizar energías renovables, favorecer modelos de consumo y producción sostenibles y reducir al mínimo la huella de carbono del total de soluciones de IA implantadas en las ciudades. Por ejemplo, las redes inteligentes que funcionan con IA podrían ajustar de forma dinámica el uso de energía basándose en datos obtenidos en tiempo real, lo que disminuye el despilfarro y favorece la integración de fuentes de energía renovables como la solar y la eólica. El aprendizaje federado abre también una vía prometedora para la sostenibilidad de la IA urbana. En lugar de depender de centros centralizados de datos para el entrenamiento de modelos en gran escala, el aprendizaje federado permite entrenar modelos más pequeños directamente en dispositivos descentralizados, como los sensores de la Internet de las Cosas integrados en la infraestructura urbana. Este cambio reduce la necesidad de transmitir cantidades enormes de datos a servidores centralizados, con lo que se ahorra en procesos de tratamiento y almacenamiento de los datos, que consumen mucha energía. Al aprovechar los recursos informáticos locales existentes, el aprendizaje federado también reduce la demanda agregada de *hardware* nuevo y mitiga el impacto ambiental de la actividad de IA masiva. Además, este enfoque aumenta la privacidad y la seguridad de los datos al mantener la información confidencial en los dispositivos locales, y reduce la necesidad de intercambiar datos al tiempo que favorece las prácticas de IA sostenible.

La verdadera sostenibilidad social de estos sistemas también debe tener en cuenta la eticidad del proceso de su desarrollo y garantizar prácticas equitativas a lo largo de todo el ciclo de vida de la IA. De este modo se contribuye a impulsar la sostenibilidad social, fomentando ciudades más justas e inclusivas, al tiempo que se da respuesta a las desigualdades globales presentes en la producción de IA.

Para que sea sostenible en los entornos urbanos, la IA también debe ser viable a largo plazo desde el punto de vista económico, lo que implica desarrollar sistemas que se integren con facilidad en la infraestructura urbana existente, se adapten a las necesidades futuras y se conciban para un uso prolongado.

Sostenibilidad social: equidad e imparcialidad. Puesto que las ciudades adoptan cada vez más la IA para el funcionamiento de servicios tales como la seguridad, la atención sanitaria y la asignación de recursos públicos, es fundamental que estos sistemas contribuyan a la sostenibilidad social promoviendo la imparcialidad, la equidad y la inclusión. En materia de ordenamiento urbano, por ejemplo, la IA sostenible podría instrumentalizarse para identificar y abordar las desigualdades, por ejemplo garantizando que los barrios desfavorecidos reciban acceso equitativo al transporte, la atención médica y la educación. No obstante, el componente social de la IA sostenible no abarca solo la finalidad de su uso, el diseño de algoritmos que minimicen el sesgo y la garantía de que las comunidades marginadas no se vean perjudicadas por las decisiones basadas en la IA, sino también la manera de abordar las implicaciones éticas del desarrollo e implementación de estos sistemas. Muchos de los programas son entrenados y mantenidos por personal escasamente remunerado y sobrecargado de trabajo de las comunidades del Sur global (Rowe, 2023), a menudo subcontratado. Esta fuerza laboral –decisiva para el entrenamiento de muchos sistemas de IA– pone de manifiesto las enormes desigualdades presentes en la cadena mundial de suministro de la IA, ya que suele soportar malas condiciones de trabajo al tiempo que asume la mayor parte de las tareas repetitivas y poco apreciadas. La verdadera sostenibilidad social de estos sistemas también debe tener en cuenta la eticidad del proceso de su desarrollo y garantizar prácticas equitativas a lo largo de todo el ciclo de vida de la IA. De este modo se contribuye a impulsar la sostenibilidad social, fomentando ciudades más justas e inclusivas, al tiempo que se da respuesta a las desigualdades globales presentes en la producción de IA.

Gobernanza ética y rendición de cuentas. Los sistemas de IA urbana deben regirse por unos marcos éticos sólidos que primen la transparencia y la rendición de cuentas. Los gobiernos municipales y las partes interesadas deben garantizar que los sistemas de IA sean explicables y que los procesos de toma de decisiones resulten claros para la ciudadanía. Este tipo de medidas generarían confianza y garantizarían que los errores o consecuencias no deseadas pudieran identificarse y abordarse con prontitud. Por ejemplo, los sistemas de IA utilizados para la vigilancia o la aplicación de la ley en las ciudades deben diseñarse con estructuras claras de rendición de cuentas que protejan la privacidad y los derechos civiles de la población. En el capítulo de esta monografía dedicado a la transparencia y la rendición de cuentas se analiza más a fondo la operativización de estos principios en los sistemas de IA.

Sostenibilidad económica en las ciudades. Para que sea sostenible en los entornos urbanos, la IA también debe ser viable a largo plazo desde el punto de vista económico, lo que implica desarrollar sistemas que se integren con facilidad en la infraestructura urbana existente, se adapten a las necesidades futuras y se conciban para un uso prolongado. Las ciudades pueden favorecer los modelos de economía circular fomentando la reutilización y reciclado de las tecnologías, los datos y el *hardware* de la IA, con lo que se reducen los residuos y se disminuyen los costes.

Alineamiento con los objetivos de sostenibilidad urbana. Los sistemas de IA implantados en las ciudades no deben limitarse a favorecer el logro de los objetivos de sostenibilidad urbana –como la disminución de la contaminación, la mejora de la salud pública y el aumento de la calidad de vida–, sino también garantizar que el uso de estas mismas tecnologías

contribuya a la sostenibilidad. Una forma de conseguirlo es utilizar con nuevos fines la energía y los recursos empleados por la infraestructura de IA. Por ejemplo, utilizar el excedente de calor de los centros de datos – importante subproducto de las demandas computacionales de la IA– puede contribuir a la sostenibilidad urbana al reducir el consumo total de energía. El proyecto Parques de Datos de Estocolmo¹ ha puesto de manifiesto que el calor residual de los centros de datos puede redirigirse para calentar los edificios residenciales y comerciales, y demostrado que la infraestructura de IA puede incorporarse en un modelo de economía circular, lo que se ajusta a los objetivos climáticos al tiempo que reduce las necesidades públicas de energía. Más allá de la eficiencia de los recursos, relacionada con la idea de la sostenibilidad de la propia IA, la IA puede servir para optimizar los sistemas urbanos de sostenibilidad. Al aprovechar estas tecnologías para mejorar la gestión de los recursos, reducir el consumo de energía y favorecer las iniciativas de resiliencia climática, las ciudades pueden hacer frente a desafíos acuciantes tales como el cambio climático y la urbanización. Por ejemplo, las zonas urbanas son especialmente vulnerables a los efectos del cambio climático, como el aumento de las temperaturas y los fenómenos meteorológicos extremos. La IA puede desempeñar un papel fundamental en el aumento de la resiliencia climática urbana facilitando análisis predictivos avanzados y sistemas de alerta temprana de riesgos relacionados con el clima; estos últimos pueden alertar a las autoridades de los posibles peligros ambientales y facilitar una respuesta rápida, contribuyendo a mitigar los efectos de la contaminación urbana sobre el medioambiente y la salud.

IA para la planificación y el desarrollo urbanos. La IA transforma la manera de planificar y desarrollar las ciudades (Jha, 2021) y fomenta una mayor sostenibilidad económica, ambiental y social. Mediante el análisis de grandes conjuntos de datos sobre el crecimiento demográfico, el uso del suelo, las modalidades de transporte y los factores ambientales, entre muchas otras aplicaciones, la IA puede ayudar a urbanistas y a las autoridades responsables a concebir ciudades más eficientes y sostenibles por medio de los denominados gemelos digitales. Por ejemplo, los modelos de IA pueden prever el efecto que tendrán los cambios en las infraestructuras –como la construcción de nuevas carreteras o los sistemas de transporte público– sobre el comportamiento del tráfico, los niveles de contaminación y el consumo de energía. Así, las instancias de planificación pueden anticiparse al futuro y realizar previsiones responsables, lo que permite tomar decisiones más fundamentadas que promuevan la sostenibilidad a largo plazo. La IA también puede servir para optimizar el uso del suelo y las políticas de zonificación, y garantizar el equilibrio entre el desarrollo urbano y la conservación de los espacios verdes y recursos naturales. Esta cuestión es especialmente importante en ciudades de rápido crecimiento, en las que la demanda de viviendas e infraestructuras suele dar lugar a la expansión urbana y la pérdida de ecosistemas valiosos. Sin embargo, más allá de su aplicación a los sistemas urbanos, es fundamental que el desarrollo e implementación de la propia IA sea compatible con las prácticas sostenibles. Aunque los gemelos digitales son herramientas potentes para simular situaciones posibles en materia de ordenamiento urbano, la sostenibilidad de estos escenarios depende de la eficiencia de los modelos de IA subyacentes y de la infraestructura de apoyo. No obstante, cabe observar que, tal y como señalan numerosos estudios (Andersson, 2021), los gemelos digitales creados con IA pueden constituir un método más eficiente desde el punto de vista de los recursos que sus simulaciones homólogas basadas en las leyes de la física, ya que se ejecutan en cuestión de segundos en las unida-

1. <https://stockholmdataparks.com/>

des centrales de procesamiento (CPU) de los ordenadores portátiles frente a la necesidad de utilizar superordenadores durante varios días.

5. Recomendaciones de política y observaciones finales

La innovación y la tecnología desempeñan un papel cada vez más destacado en la planificación de futuros urbanos sostenibles (ONU-Hábitat, 2022). Como se expone en el listado de recomendaciones políticas que figura a continuación, el diseño e implementación de la tecnología deben adaptarse a la gran diversidad del contexto urbano:

a. Recomendaciones sobre sostenibilidad ambiental

- La urgencia por descarbonizar las economías urbanas debería impulsar la **confluencia de las tecnologías inteligentes y ecológicas**. Las políticas deberían hacer hincapié en la eficiencia energética, la preservación del medioambiente y la resiliencia. Esto incluye el establecimiento de normas ecológicas para la IA que primen los algoritmos y el *hardware* eficientes desde el punto de vista energético, así como la creación de economías circulares en torno a los centros de datos, por ejemplo reciclando el excedente de calor. Además, las políticas de gestión del ciclo de vida deberían promover el abastecimiento responsable, la reutilización y el reciclaje del *hardware* de IA a fin de reducir al mínimo los residuos electrónicos.
- Las evaluaciones del impacto deberían sopesar cuidadosamente si la implementación de la **IA para los proyectos sobre sostenibilidad justifica el coste ambiental de la tecnología**, tal como señalan trabajos anteriores (Dixon, 2022). Los nuevos marcos son fundamentales para medir y comparar los costes de todo el ciclo de vida de la IA y garantizar la evaluación exhaustiva de su sostenibilidad.
- Están comenzando a aparecer **marcos de clasificación energética por estrellas de la IA** (Luccioni, 2024) que deberían incorporarse a los dispositivos urbanos de la Internet de las Cosas, lo que ofrecería a las personas usuarias una información valiosa para comprender mejor las repercusiones ambientales de las herramientas que utilizan y para adoptarlas de forma más responsable.
- La colaboración y las asociaciones público-privadas pueden impulsar el desarrollo de tecnologías de IA sostenible en las zonas urbanas. La creación de **consorcios de IA para lograr ciudades sostenibles** puede fomentar alianzas entre gobiernos, empresas tecnológicas e instituciones de investigación con el fin de abordar desafíos urbanos tales como la gestión de la energía, el transporte y la reducción de los residuos. Las ciudades deberían incentivar el desarrollo sostenible de la IA ofreciendo bonificaciones fiscales o subvenciones a las empresas que desarrollan soluciones de IA respetuosas con el medioambiente.
- La resiliencia urbana y las infraestructuras inteligentes deberían favorecerse mediante **políticas que alienten la IA en pro de la resiliencia en materia climática y de biodiversidad**. Esto incluye el uso de sistemas de alerta temprana basados en la IA para hacer

frente a las catástrofes naturales, los puntos críticos en materia de pérdida de biodiversidad y los fenómenos climáticos extremos.

b. Recomendaciones sobre sostenibilidad social

- Dado que todas las dimensiones de la IA sostenible son interdependientes, la implementación responsable y ética de la IA es también un aspecto fundamental en los entornos urbanos. Los sistemas de IA deben someterse a auditorías de equidad para evitar la discriminación y la desigualdad social. La creación de **comités locales de ética en materia de IA** debería garantizar que los proyectos de IA urbana cumplan las normas sobre privacidad, equidad y responsabilidad.
- La protección y la seguridad de la información son también áreas prioritarias. Deberían promulgarse **leyes firmes sobre la privacidad de los datos urbanos** a fin de proteger los datos personales recogidos por los sensores, las cámaras y las aplicaciones móviles, con medidas tales como la anonimización y el uso del consentimiento explícito. Además, se precisa disponer de marcos transparentes y seguros para el intercambio de información entre gobiernos, empresas privadas y entidades desarrolladoras de IA para garantizar un uso responsable de los datos de la ciudadanía que no ponga en riesgo la privacidad.
- Para fomentar el apoyo y la comprensión de la población, las políticas deberían promover **la participación social y la alfabetización digital**. Las ciudades deberían fomentar **modelos de gobernanza participativa** que involucren a la población en los procesos de toma de decisiones sobre la IA, y poner en marcha al mismo tiempo campañas de alfabetización digital para informar a la población sobre este tipo de tecnologías y los efectos que producen, y el modo de proteger sus derechos.
- Garantizar la **igualdad de acceso a los servicios públicos basados en la IA**, sobre todo para las comunidades marginadas e insuficientemente atendidas, es esencial para promover la inclusión.
- Promover la **ciencia abierta** es fundamental y propicia la realización de auditorías públicas de estos sistemas, a la vez que garantiza la aplicación de medidas firmes en materia de ciberseguridad para proteger los datos confidenciales y los servicios públicos. La transparencia, mediante normativas que exijan que los sistemas de IA utilizados en los servicios públicos sean explicables y permitan que tanto las partes interesadas como la población en general comprendan la manera en que se toman las decisiones, también es clave.

c. Recomendaciones generales

- Para amortiguar las repercusiones económicas de la IA, las políticas deberían **apoyar la transición laboral y el desarrollo profesional de la fuerza de trabajo**. Deberían asignarse fondos públicos a programas de reciclaje profesional que ayuden a las personas a cambiar de empleo, sobre todo en sectores emergentes en los que la automa-

tización puede provocar la pérdida de puestos de trabajo. Promover el crecimiento del empleo verde basado en la IA, como el relacionado con la gestión de las energías renovables y las infraestructuras urbanas sostenibles, puede impulsar aún más el crecimiento económico sostenible.

- Los marcos de supervisión y rendición de cuentas son esenciales para garantizar que los sistemas de IA sean compatibles con los objetivos de sostenibilidad a lo largo del tiempo. Debería establecerse la **obligatoriedad de realizar evaluaciones del impacto de la IA**, similares a las evaluaciones del impacto ambiental, para analizar los efectos ambientales, económicos y sociales de la implementación de estas tecnologías en las ciudades. La supervisión y auditoría continuas de los sistemas de IA urbana pueden ayudar a garantizar que estos sigan siendo adaptables a los desafíos y las consideraciones éticas que surjan.
- **Incorporar en las normas ambientales, sociales y de gobernanza aspectos** que tengan en cuenta la sostenibilidad de los datos, algoritmos y recursos informáticos utilizados por las empresas, así como el apoyo proporcionado a las fuentes de energía renovables y las economías computacionales circulares.
- Las normas reguladoras de la infraestructura urbana inteligente deben garantizar **que las tecnologías de IA sean flexibles, compatibles entre sí y adaptables a las necesidades urbanas futuras**, especialmente en ámbitos tales como la gestión del tráfico, la reducción de residuos y la eficiencia energética.
- Por último, en el plano mundial, deberían alentarse la **colaboración y estandarización** internacionales. Las ciudades deberían aunar esfuerzos para elaborar unas normas globales de sostenibilidad para la IA e intercambiar buenas prácticas (Strubell, 2019), y garantizar la compatibilidad con los objetivos internacionales. Las plataformas de intercambio de conocimientos entre ciudades pueden contribuir a acelerar la adopción de prácticas de IA sostenible en todo el mundo.

Referencias bibliográficas

Andersson, T. R. *et al.* «Seasonal Arctic sea ice forecasting with probabilistic deep learning». *Nature Communications*, vol. 12, n.º 1 (2021): 5124.

Bibri, S. E. «Data-driven smart sustainable cities of the future: Urban computing and intelligence for strategic, short-term, and joined-up planning». *Computational Urban Science*, vol.1, n.º 1 (2021): 8.

Dixon, B., Pérez Ortiz, M. y Bieker, J. «Comparing the carbon costs and benefits of low-resource solar nowcasting». Taller de NeurIPS sobre la lucha contra el cambio climático mediante el aprendizaje automático (2022).

Consejo Asesor Alemán sobre el Cambio Global (GACGC), *Towards Our Common Digital Future*, informe emblemático, 2019.

Heikkilä, M. «Making an image with generative AI uses as much energy as charging your phone», *MIT Technology Review* (2023).

Jha, A. K. *et al.* «A review of AI for urban planning: Towards building sustainable smart cities». Sexto Congreso Internacional sobre Tecnologías Informáticas Creativas (ICICT), 2021. IEEE, 2021.

Keeble, B. R. «The Brundtland Report: “Our Common Future”». *Medicine and War*, vol. 4, n.º 1 (1988), pp. 17-25.

Luccioni, S., Trevelin, B. y Mitchell, M. «The Environmental Impacts of AI-Primer», 2024.

Muir, J. «My first summer in the Sierra», en: *British Politics and the Environment in the Long Nineteenth Century*. Routledge, 1911 (reeditado en 2023), pp. 291-296. [Traducción propia]

Nishant, R., Kennedy, M. y Corbett, J. «Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda». *International Journal of Information Management*, vol. 53 (2020): 102104.

Pastor Escuredo, D., Treleaven, P y Vinuesa, R. «An Ethical framework for artificial intelligence and sustainable cities». *Ai*, vol. 3, n.º 4 (2022), pp. 961-974.

Rowe, N. «Millions of Workers Are Training AI Models for Pennies», WIRED (2023).

Strubell, E., Ganesh, A. y McCallum, A. «Energy and policy considerations for modern deep learning research». *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 34. N.º 9 (2020).

ONU-Hábitat, *Envisaging the Future of Cities (Visualizando el futuro de las ciudades)*, Informe Mundial de las Ciudades 2022.

Van Wynsberghe, A. «Sustainable AI: AI for sustainability and the sustainability of AI». *AI and Ethics*, vol.1, n.º 3 (2021), pp. 213-218.

Vinuesa, R. *et al.* «The role of artificial intelligence in achieving the Sustainable Development Goals». *Nature Communications*, vol. 11, n.º 1 (2020), pp. 1-10.

Wilson, C. y Van Der Velden, M. «Sustainable AI: An integrated model to guide public sector decision-making». *Technology in Society*, vol. 68 (2022): 101926.

Wu, C-J. *et al.* «Sustainable AI: Environmental implications, challenges and opportunities». *Proceedings of Machine Learning and Systems*, vol. 4 (2022), pp. 795-813.

Yigitcanlar, T. y Cugurullo, F. «The sustainability of artificial intelligence: An urbanistic viewpoint from the lens of smart and sustainable cities». *Sustainability*, vol. 12, n.º 20 (2020): 8548.

PARTE II. CASOS PRÁCTICOS DE MARCOS DE GOBERNANZA DE IA URBANA

- CASO PRÁCTICO 1: BARCELONA
- CASO PRÁCTICO 2: ÁMSTERDAM
- CASO PRÁCTICO 3: NUEVA YORK
- CASO PRÁCTICO 4: SAN JOSÉ
- CASO PRÁCTICO 5: DUBÁI
- CASO PRÁCTICO 6: SINGAPUR

PARTE II. CASOS PRÁCTICOS DE MARCOS DE GOBERNANZA DE IA URBANA

Alexandra Vidal D'oleo

*Investigadora y gestora de proyecto,
Programa de Ciudades Globales, CIDOB*

El primer informe del Atlas de la Inteligencia Artificial Urbana del GOUAI (Galceran-Vercher y Vidal, 2024) revela que, aunque muchas ciudades están experimentando actualmente con inteligencia artificial (IA), solo un pequeño porcentaje ha puesto en marcha políticas o estrategias generales para regular su uso y garantizar el cumplimiento de principios éticos fundamentales. Hasta ahora, se ha hecho hincapié en abordar los retos urbanos inmediatos con un pragmatismo orientado a la búsqueda de soluciones. Como resultado, existe una brecha significativa entre la adopción de IA y el establecimiento de marcos de gobernanza eficaces. Sin embargo, impulsados por el discurso público y una corriente importante que aboga por una regulación global de los algoritmos y la IA, algunos gobiernos locales han tomado la iniciativa de crear sus propios marcos de gobernanza. Se prevé que esta tendencia crezca exponencialmente en los próximos años.

En las páginas siguientes se ofrece una recopilación de casos prácticos de ciudades de todo el mundo que se han esforzado por establecer marcos de gobernanza local de IA mediante la adopción de diferentes mecanismos de política para gobernar la IA de forma integral. Todos los mecanismos de política pertenecen a una de las siguientes categorías: (1) Principios, estrategias y directrices; (2) Normativa y legislación locales; (3) Mecanismos de transparencia y explicabilidad; (4) Evaluación del impacto algorítmico; (5) Auditorías e inspección reglamentaria; (6) Supervisión humana, rendición de cuentas, procedimientos de audiencia y recurso; (7) Condiciones de contratación pública; (8) Órganos consultivos y de supervisión externos/independientes; (9) Alianzas, comunidades de práctica y grupos de aprendizaje; (10) Programas de desarrollo de capacidades; (11) Fomento de la innovación, el conocimiento y la experimentación locales; (12) Participación de la comunidad; (13) Gobernanza de datos; y (14) Otras políticas y medidas. Cabe señalar que, aunque se cree que algunas de las políticas presentadas están en fase de desarrollo, es posible que ya se hayan aplicado sin preaviso.

Las ciudades elegidas varían en cuanto a situación geográfica, tamaño y renta per cápita. En concreto, se presentan los siguientes casos prácticos:

1. Barcelona (España)
2. Ámsterdam (Países Bajos)
3. Nueva York (Estados Unidos de América)
4. San José (Estados Unidos de América)
5. Dubái (Emiratos Árabes Unidos)
6. Singapur (República de Singapur)

Caso práctico 1: Barcelona

Población: 1,655,956 (2023)	Renta per cápita: €31,531 (2022)	Región: Europa
------------------------------------	---	-----------------------

MARCO DE GOBERNANZA DE IA		
[1] Principios, estrategias y directrices	Estrategia de IA	<p>Estrategía municipal de algoritmos y datos para el impulso ético de la inteligencia artificial: define un conjunto de principios guías y 20 acciones para una implementación ética de la IA, que incluyen:</p> <ul style="list-style-type: none"> • el uso de la IA para sistemas automatizados de recomendación, más que para sistemas de toma de decisiones. • la transparencia y la auditabilidad: los modelos algorítmicos y las bases de datos deben ser accesibles, comprensibles y auditables por el público en general. • el establecimiento de regímenes de responsabilidad por cualquier daño o pérdida que pueda producirse.
	Principios locales para la IA	Incluidos en la estrategia municipal de IA: (1) Acción y supervisión humana; (2) Solidez técnica y seguridad; (3) Privacidad y gobernanza de datos; (4) Transparencia; (5) Diversidad, inclusión y equidad; (6) Compromiso social y ambiental; (7) Responsabilidad, rendición de cuentas y control democrático
	Protocolos internos para el uso de IA	Definición de metodologías de trabajo y protocolos para la implementación de sistemas algorítmicos: define los mecanismos para cada etapa de la licitación e implementación de sistemas de IA por parte del Ayuntamiento y establece los órganos de gobernanza y supervisión que velarán por que el impacto de la IA se alinee con los principios éticos.
[3] Mecanismos de transparencia y explicabilidad	Registro público de algoritmos (*)	Creación de un registro municipal de algoritmos actuales y futuros que afecten a trámites y servicios municipales. El registro será público y servirá también para clasificar los algoritmos según el riesgo que entrañen, con explicaciones claras para la ciudadanía y otras partes interesadas. Para cada algoritmo registrado, se incluirá un punto de contacto público al que la ciudadanía pueda dirigirse.
	Estándar de transparencia algorítmica	Desarrollo de un modelo algorítmico común que garantice el uso adecuado de los datos. Proyecto desarrollado junto con otras ocho ciudades europeas en el marco de Eurocities.
	Página web municipal que divulgue toda información pertinente sobre IA	Barcelona Ciudad Digital : sitio web municipal que divulga todos los recursos disponibles en la ciudad para impulsar la digitalización, incluida toda la información relevante sobre proyectos e iniciativas de IA.
[4] Evaluación del impacto algorítmico	Evaluación y gestión de riesgos	Análisis de riesgos (incluido en el protocolo): La Oficina Técnica de IA evalúa los algoritmos utilizados por la Administración municipal y emite un informe que incluye la evaluación del riesgo. En función del riesgo que determine la oficina, se establecen los siguientes pasos: se deniegan todos los algoritmos considerados de «riesgo inaceptable», mientras que los algoritmos de «alto riesgo» están obligados a someterse a una evaluación del impacto algorítmico obligatoria por parte de la Comisión transversal municipal (véase más abajo).
	Evaluación del impacto sobre los derechos humanos	Evaluaciones obligatorias del impacto algorítmico en los sistemas de alto riesgo (incluidas en la estrategia y el protocolo).
[5] Auditorías e inspección reglamentaria	Auditorías	Auditorías obligatorias de los sistemas de alto riesgo (incluidas en la estrategia y el protocolo). Las conclusiones se harán públicas a través del registro de algoritmos.

[7] Condiciones de contratación pública	Cláusulas de contratación pública*	Inclusión de cláusulas relacionadas con los derechos digitales en la licitación de soluciones basadas en inteligencia artificial.
[8] Órganos consultivos y de supervisión externos/independientes	Consejo asesor	Consejo Asesor en Inteligencia Artificial, Ética y Derechos Digitales : formado por quince expertos independientes y multidisciplinares. Su misión es asesorar a la Administración municipal en el uso de IA, realizar estudios del impacto algorítmico sobre sistemas algorítmicos de alto riesgo y evaluar el desarrollo de la Estrategia municipal de IA, entre otras.
[9] Alianzas, comunidades de práctica y grupos de aprendizaje	Miembro de una comunidad de práctica	Cities Coalition for Digital Rights (CC4DR) (miembro fundador): una red mundial de más de 60 ciudades que trabajan en el campo de la formulación de políticas basadas en derechos digitales.
[10] Programas de desarrollo de capacidades	Desarrollo de capacidades municipal	Formación del personal municipal (incluida en la estrategia).
	Organismo municipal	Comisión transversal para el impulso de la IA ética : constituida por 25 miembros, su misión es orientar y alinear las políticas municipales que desarrollan herramientas que utilizan IA y promover la colaboración interdepartamental. Ha elaborado la Estrategia de IA.
[11] Fomento de la innovación, el conocimiento y la experimentación locales	Observatorios locales de IA	Observatorio Global de IA Urbana (GOUAI) : una iniciativa de Barcelona, liderada por CIDOB, establecida en colaboración con Ámsterdam y Londres en el marco de la CC4DR y con el apoyo de ONU-Hábitat. Investiga sobre la IA urbana desde un punto de vista ético.
[12] Participación de la comunidad	Participación popular	Canales de comunicación con la ciudadanía (incluidos en la estrategia). Promoción de espacios de reflexión y debate sobre el impacto de la IA en los servicios públicos (incluida en la estrategia).
[13] Gobernanza de datos	Transparencia de datos	Open data BCN : el portal de datos abiertos de la ciudad. Su objetivo es aprovechar al máximo los recursos públicos disponibles, haciendo que los datos generados o custodiados por organismos públicos sean accesibles, gratuitos y utilizables por todas las personas.
	Derechos sobre los datos	Proyecto Decode : iniciativa de colaboración de la UE para reforzar los derechos de la ciudadanía sobre sus datos y permitirle controlarlos y compartirlos por el bien común.

* Mecanismo de política previsto, no aplicado plenamente en diciembre de 2024.

Caso práctico 2: Ámsterdam

Población: 921.402 (2022)	Renta per cápita: 54.700€ (2022)	Región: Europa
----------------------------------	---	-----------------------

MARCO DE GOBERNANZA DE IA		
[1] Principios, estrategias y directrices	Agenda de IA	<p>Agenda de Inteligencia de Ámsterdam (2020-2024): destacó los objetivos de la ciudad en el ámbito de los algoritmos, en particular la IA. Los objetivos se centran en adoptar medidas que mejoren la calidad de vida de los habitantes de Ámsterdam, reduzcan los efectos perjudiciales de la digitalización e impulsen sus resultados beneficiosos.</p> <p>Visión de Ámsterdam sobre la IA (2024): en su nueva política, la ciudad esboza cómo debe integrarse la IA en la vida urbana y cómo debe influir en la ciudad. La visión se elaboró mediante debates con los habitantes de Ámsterdam, expertos y personal municipal. Para ponerla en práctica, la Administración municipal está elaborando una nueva Agenda de IA, cuya publicación está prevista para 2025. En ella se establecerán directrices para el uso responsable de IA en el municipio, garantizando que la tecnología se aplique de forma ética, integradora y sostenible.</p>
	Principios locales para la IA	<p>Adhesión a los principios Tada (del Manifiesto Tada), que guían la Agenda Digital de Ámsterdam y también se aplican al ámbito de la IA para que esta sea: (1) Inclusiva; (2) Controlada; (3) Adaptada a las personas; (4) Legítima y supervisada; (5) Abierta y transparente; (6) De todas y para todas las personas.</p> <p>Principios guía locales (incluidos en «Visión de Ámsterdam sobre la IA», 2024): (1) Centrada en el ser humano; (2) Fiable; (3) Preparada para el futuro.</p>
	Directrices, guías y manuales	<p>Guía de algoritmos: documento orientativo que establece el enfoque integral y las herramientas de política de la ciudad para un uso responsable de los algoritmos. El Enfoque del Ciclo de Vida de los Algoritmos consta de siete herramientas para gestionar, evaluar los riesgos e investigar los algoritmos a lo largo de su ciclo de vida: (1) Registro de algoritmos; (2) Condiciones contractuales; (3) Procedimiento de objeción; (4) Definición de gobernanza y modelo de ciclo de vida; (5) Auditoría; (6) Modelo de análisis de sesgos; y (7) Modelo de análisis del impacto sobre los derechos humanos.</p> <p>Manual de equidad: normativa para el análisis de sesgos, un plan paso a paso para evaluar un modelo, detectar sesgos y mitigar sus efectos.</p>
	Protocolos internos para el uso de IA	<p>Establecimiento de la gobernanza y modelo de ciclo de vida: especifica tareas y responsabilidades, medidas que deben tomarse para prevenir riesgos al aplicar algoritmos, la información que debe documentarse y quién es responsable si un algoritmo no cumple el objetivo previsto. El Enfoque del Ciclo de Vida de los Algoritmos describe el proceso de un algoritmo de principio a fin (incluido en la Guía de algoritmos).</p>
[3] Mecanismos de transparencia y explicabilidad	Registro público de algoritmos	Registro de algoritmos: visión general de los sistemas y algoritmos de IA utilizados por la Administración municipal.
	Estándar de transparencia algorítmica	Desarrollo de un modelo algorítmico común que garantice el uso adecuado de los datos. Proyecto desarrollado junto con otras ocho ciudades europeas en el marco de Eurocities.
	Página web municipal que divulgue toda información pertinente sobre IA	Portal municipal accesible que divulga toda la información y los recursos pertinentes de la IA, como proyectos emblemáticos y documentos estratégicos.

	Evaluación del impacto sobre los derechos humanos	Modelo de evaluación del impacto sobre los derechos humanos : elaborado a partir de las lecciones prácticas aprendidas y las herramientas existentes (incluido en la Guía de algoritmos).
[4] Evaluación del impacto algorítmico	Análisis de sesgos	Plan detallado para evaluar un modelo para detectar sesgos, que incluye, entre otros, los siguientes componentes (a) definición de los grupos («sensibles») que deben estudiarse; (b) redacción de hipótesis sobre las características que pueden dar lugar a sesgos indirectos; (c) selección de métricas que se ajusten al proyecto; (d) análisis de sesgos directos e indirectos; (e) análisis de sesgos en variables no medibles; (f) ponderación y revisión de los sesgos encontrados con una gestión responsable; (g) mitigación de los sesgos cuando sea necesario; y (h) redacción de conclusiones. El mencionado Manual de equidad proporciona las directrices para llevar a cabo el análisis de sesgos.
[5] Auditorías e inspección reglamentaria	Auditorías	Auditorías anuales : encargadas por la oficina del CIO y realizadas por el Servicio de Auditoría ACAM. Se ha elaborado un marco de normas para estas auditorías (incluido en la Guía de algoritmos).
[6] Supervisión humana, rendición de cuentas, procedimientos de audiencia y recurso	Procedimientos de respuesta y objeción accesibles a la ciudadanía	Procedimientos de objeción y directrices para los responsables de esta (incluidos en la Guía de algoritmos).
[7] Condiciones de contratación pública	Cláusulas de contratación pública	Cláusulas estándar para la contratación de sistemas algorítmicos fiables : el estándar pionero que estipula las condiciones contractuales y los requisitos de información exigidos a los proveedores de sistemas de IA contratados.
[8] Órganos consultivos y de supervisión externos/independientes	Órgano consultivo y de supervisión externo	Comité de Datos Personales de Ámsterdam : asesora a la administración municipal sobre algoritmos, ética de los datos, derechos humanos digitales y exposición de datos personales. Incluye la evaluación ética en el uso de algoritmos. El comité defiende la transparencia organizando reuniones públicas y emitiendo dictámenes.
	Miembro de una comunidad de práctica	Cities Coalition for Digital Rights (CC4DR) (miembro fundador)
[9] Alianzas, comunidades de práctica y grupos de aprendizaje	Colaboraciones multiactor	Amsterdam AI : colaboración entre el Ayuntamiento, las instituciones del conocimiento, centros de investigación, centros médicos y el Consejo Económico de Ámsterdam. Esta colaboración se focaliza en IA responsable con un enfoque centrado en el ser humano. NL AI Coalition : la ciudad forma parte del grupo de trabajo nacional sobre IA en el sector público, junto con la coalición «Amsterdam AI». Smart Health Amsterdam network : la red regional para la innovación impulsada por los datos y la IA en el sector de las ciencias de la vida y la salud en Ámsterdam (también forma parte de la coalición «Amsterdam AI»).
[10] Programas de desarrollo de capacidades	Desarrollo de capacidades municipal	Formación del personal municipal: todos los funcionarios deben realizar el curso nacional de IA. Los servicios jurídicos, de contratación pública y de auditoría deben realizar cursos periódicos de actualización. Creación de equipos municipales de IA.
[11] Fomento de la innovación, el conocimiento y la experimentación locales	Centros, hubs y laboratorios de innovación de IA	Civic AI Lab : creado por el Ayuntamiento, la universidad pública y un ministerio nacional, el objetivo es investigar cómo la IA puede contrarrestar la desigualdad social o evitar que la refuerce. DataLab : crea soluciones tecnológicas abiertas y responsables.

[13] Gobernanza de datos	Transparencia de datos	Data Amsterdam (versión beta): el portal de datos abiertos de la ciudad. Su objetivo es aprovechar al máximo los recursos públicos disponibles, haciendo que los datos generados o custodiados por organismos públicos sean accesibles, gratuitos y utilizables por todas las personas.
	Derechos sobre los datos	Proyecto Decode
	Mecanismos de intercambio de datos	Amsterdam Data Exchange (AMdEX) : para el intercambio de datos entre organizaciones mediante la creación de un notario digital.

* Mecanismo de política previsto, no aplicado plenamente en diciembre de 2024.

Caso práctico 3: Nueva York

Población: 8.258.000 (2023)	Renta per cápita: 44.537€ (2022)	Región: Norteamérica
------------------------------------	---	-----------------------------

MARCO DE GOBERNANZA DE IA		
[1] Principios, estrategias y directrices	Estrategia de IA	Estrategia de IA (2021-2023): esfuerzo fundacional para fomentar un ecosistema local de IA sano e intersectorial. El documento estableció una base de información sobre IA para ayudar a garantizar que los responsables de la toma de decisiones trabajasen a partir de una comprensión compartida de la tecnología y los problemas que plantea. Incluía un marco de ética, gobernanza y política.
	Plan de acción de IA	Plan de Acción de IA (2023-2025): incluye 37 acciones en siete iniciativas con el objetivo de crear gobernanza para el uso municipal de la IA. Obliga a elaborar un informe anual sobre los avances del plan.
	Principios locales para la IA	Principios de IA (incluidos en su Plan de Acción de IA): (1) Validez y fiabilidad; (2) Responsabilidad social; (3) Privacidad de la información; (4) Ciberseguridad; (5) Confianza y transparencia. Principios incluidos en su anterior Estrategia de IA: (1) Rendición de cuentas; (2) Justicia; (3) Privacidad y seguridad; (4) Implicación y participación de la comunidad.
	Directrices, guías y manuales	Directrices sobre principios y definiciones de IA : especifica los principios de IA de la ciudad, ofrece definiciones concretas de términos relacionados con la IA y detalla las leyes, las políticas, los requisitos y los procesos relacionados que se aplican. Orientaciones preliminares de uso: Inteligencia Artificial Generativa: incluye términos y definiciones, funciones y responsabilidades, orientación para el uso de la inteligencia artificial generativa por parte de la Administración municipal, y especifica las leyes, políticas, requisitos y procesos aplicables.
	Protocolos internos para el uso de IA	Protocolos internos especificados en el Plan de Acción, que incluyen un informe anual obligatorio por parte de los organismos municipales para ampliar la información pública sobre IA.
[2] Normativa y legislación locales	Regulación de aplicaciones controvertidas de IA	Ley de protección de datos biométricos para empresas. Tecnología de contratación obligada a auditar su parcialidad.
[3] Mecanismos de transparencia y explicabilidad	Página web municipal que divulgue toda información pertinente sobre IA	Portal municipal accesible que divulga toda la información y los recursos pertinentes de la IA, como proyectos emblemáticos y documentos estratégicos.
	Directorio municipal de herramientas de IA adquiridas (para uso interno)	Establecimiento de un directorio interno de herramientas de IA adquiridas y orientación sobre su uso adecuado, compartido con todos los departamentos para apoyar la visibilidad y el acceso. Incluido en el plan de acción.
[4] Evaluación del impacto algorítmico	Evaluación y gestión de riesgos (*)	Evaluación de riesgos de la IA y proceso de revisión del proyecto (incluido en el plan de acción).
[7] Condiciones de contratación pública	Cláusulas de contratación pública (*)	Normas de contratación específicas para la IA.

[8] Órganos consultivos y de supervisión externos/independientes	Órgano consultivo y de supervisión externo	Red de asesoramiento sobre IA : reúne a expertos independientes del sector empresarial privado, el mundo académico, el mundo laboral y las organizaciones cívicas para apoyar los esfuerzos de la ciudad en materia de IA con carácter consultivo.
[9] Alianzas, comunidades de práctica y grupos de aprendizaje	Miembro de una comunidad de práctica	Cities Coalition for Digital Rights (CC4DR) (miembro fundador)
[10] Programas de desarrollo de capacidades	Organismo municipal	NYC Automated Decision Systems Task Force : creado y encargado de emitir recomendaciones sobre cómo la ciudad debía gestionar el uso de algoritmos. Fue el primero en el país y culminó con la publicación de un informe accesible. Comité Directivo de AI : reúne a las partes interesadas de todo el gobierno de la ciudad para proporcionar información y supervisar las actividades de la IA.
[12] Participación de la comunidad	Participación popular	Sesiones de escucha públicas
[13] Gobernanza de datos	Transparencia de datos	NYC Open data : el portal de datos abiertos de la ciudad. Su objetivo es aprovechar al máximo los recursos públicos disponibles, haciendo que los datos generados o custodiados por organismos públicos sean accesibles, gratuitos y utilizables por todas las personas.

* Mecanismo de política previsto, no aplicado plenamente en diciembre de 2024.

Caso práctico 4: San José

Población: 969.655 (2023)	Renta per cápita: 139.761€ (2023)	Región: Norteamérica
----------------------------------	--	-----------------------------

MARCO DE GOBERNANZA DE IA		
[1] Principios, estrategias y directrices	Manual de IA	Manual de IA: ofrece una orientación exhaustiva sobre cómo cumplir la política de IA de la ciudad. Incluye: (1) Política de IA, (2) Revisión de IA (necesaria para todas las contrataciones públicas y las iniciativas de datos); (3) Gobernanza de IA (el marco para gestionar y supervisar el ciclo de vida de la IA); (4) Directrices sobre IA generativa
	Directrices, guías y manuales	Directrices sobre IA generativa: fue el primer paso de un proceso de colaboración para desarrollar la política general de IA de la ciudad. Se invitó a los usuarios registrados a unirse al Departamento de Informática en un grupo de trabajo para compartir su experiencia y codesarrollar las políticas de IA de la ciudad.
	Principios locales para la IA	Principios de IA de San José: (1) Eficacia (fiabilidad); (2) Transparencia; (3) Equidad; (4) Rendición de cuentas; (5) Diseño centrado en el ser humano; (6) Privacidad; (7) Seguridad y protección; (8) Empoderamiento de los trabajadores.
[3] Mecanismos de transparencia y explicabilidad	Registro público de algoritmos	Inventario de IA (punto 5 del proceso de revisión): descripción general de los sistemas y algoritmos de IA utilizados por la administración municipal.
	Página web municipal que divulgue toda información pertinente sobre IA	Portal municipal accesible que divulga toda la información y los recursos pertinentes de la IA, como proyectos emblemáticos y documentos estratégicos.
[4] Evaluación del impacto algorítmico	Evaluación y gestión de riesgos	Modelo de análisis del umbral de riesgo de la IA: realizado por la Oficina de Privacidad Digital (punto 2 del proceso de revisión).
	Evaluación del impacto sobre los derechos humanos	Evaluación del impacto algorítmico: realizada por el Ayuntamiento cuando los sistemas de IA adquiridos se consideren de riesgo medio-alto según el análisis de riesgos (punto 3 del proceso de revisión).
[6] Supervisión humana, rendición de cuentas, procedimientos de audiencia y recurso	Procedimientos de retroalimentación y oposición accesibles a la ciudadanía	Formulario de comentarios públicos: a través de este formulario la ciudadanía puede presentar comentarios sobre proyectos que impliquen un nuevo uso de información personal. La información sobre nuevos proyectos puede consultarse en línea.
	Seguimiento e informes internos	Informe anual de uso: el departamento titular de la empresa del sistema algorítmico debe presentar un informe anual de uso en el que se detallen: 1. Resumen del proyecto 2. Métricas de rendimiento requeridas 3. Planes futuros para la iniciativa tecnológica El público puede hacer comentarios en línea sobre el uso de los datos y las actualizaciones anuales (punto 6 del proceso de revisión).

[7] Condiciones de contratación pública	Protocolos internos para la adquisición de IA	Marco de revisión de IA: evalúa los beneficios y riesgos de la contratación municipal. Proceso de revisión: 1. Solicitud de contratación 2. Análisis de riesgos 3. Evaluación del impacto algorítmico (para los sistemas de riesgo medio-alto): incluye la evaluación municipal del impacto algorítmico y la hoja informativa sobre IA del proveedor 4. Revisión final 5. Preparación previa al lanzamiento: protocolo de uso de datos, formación de usuarios y publicación del inventario de IA 6. Seguimiento continuo
	Cláusulas de contratación pública	Hoja informativa sobre IA del proveedor: incluye una hoja informativa y un cuestionario de evaluación del impacto algorítmico para el proveedor (punto 3 del proceso de revisión).
[8] Órganos consultivos y de supervisión externos/ independientes	Órgano consultivo y de supervisión multiactor	Grupo de asesoramiento sobre IA: dirigido por el ayuntamiento, las partes interesadas externas asesoran a los departamentos municipales y al CIO sobre las políticas y actividades relacionadas con la gobernanza de IA. Está formado por expertos en IA del sector empresarial, el mundo académico, los derechos civiles y miembros del público. El Grupo de asesoramiento se reúne trimestralmente y el poder de decisión sigue estando en manos del ayuntamiento.
[9] Alianzas, comunidades de práctica y grupos de aprendizaje	Alianza multiactor	GovAI Coalition: coalición liderada por San José que reúne a organismos públicos, sociedad civil, instituciones académicas y empresas para promover una IA responsable en el sector público. Está compuesta por más de 1.500 miembros y más de 500 organismos locales, estatales y federales.
	Miembro de una comunidad de práctica	Cities Coalition for Digital Rights (CC4DR)
[10] Programas de desarrollo de capacidades	Grupo de trabajo municipal	Grupo de Trabajo sobre AI (AIWG): el personal de varios departamentos municipales debate sobre cuestiones y proyectos relacionados con la IA en la ciudad. Compuesto por responsables de IA de los departamentos y potencialmente por representantes de otros departamentos.
[12] Participación de la comunidad	Participación popular	Si un sistema de IA contratado se considera de interés público, el ayuntamiento lleva a cabo actividades de divulgación en línea y en persona (dirigidas a comunidades con un acceso en línea limitado). Los comentarios de la comunidad se incorporan al protocolo de uso de datos.
[13] Gobernanza de datos	Transparencia de datos	San Jose CA Open Data Portal: el portal de datos abiertos de la ciudad. Su objetivo es aprovechar al máximo los recursos públicos disponibles, haciendo que los datos generados o custodiados por organismos públicos sean accesibles, gratuitos y utilizables por todas las personas.
	Protocolo de datos	Protocolo de uso de datos: protocolo para sistemas de riesgo medio-alto que regula la recopilación, el acceso, el tratamiento y el uso compartido de datos y garantiza el cumplimiento de la Política de Privacidad Digital de la ciudad (punto 5 del proceso de revisión).

Caso práctico 5: Dubái

Población: 3.051.000 (2023)	Renta per cápita: 47.995 € (2023)	Región: Oriente Medio y el Norte de África
------------------------------------	--	---

MARCO DE GOBERNANZA DE IA		
[1] Principios, estrategias y directrices	Plan y hoja de ruta de IA	La Hoja de Ruta de IA (2024) forma parte del Plan Universal de Inteligencia Artificial de Dubái y apoya los objetivos de la Agenda Económica de Dubái D33.
	Principios locales para la IA	<p>Principios y directrices de ética de la IA explicados a continuación:</p> <ul style="list-style-type: none"> • Principios éticos de IA: (1) Ética (justa, responsable, transparente y explicable); (2) Seguridad (segura y protegida); (3) Humanidad; (4) Inclusividad. • Directrices éticas de IA, para que los sistemas de IA sean: (1) Justos; (2) Responsables; (3) Transparentes; (4) Explicables.
[3] Mecanismos de transparencia y explicabilidad	Página web municipal que divulgue toda información pertinente sobre IA	Portal municipal accesible que divulga toda la información y los recursos pertinentes de la IA, como proyectos emblemáticos y documentos estratégicos.
[5] Auditorías e inspección reglamentaria	Herramienta de autoevaluación	Herramienta de autoevaluación ética de IA : creada para permitir a las organizaciones desarrolladoras u operadoras de IA evaluar el nivel ético de un sistema de IA, utilizando las Directrices éticas de IA de Dubái.
[11] Fomento de la innovación, el conocimiento y la experimentación locales	Centros, <i>hubs</i> y laboratorios de innovación de IA	AI Lab : creado en colaboración con IBM, trabaja con una creciente red de socios del sector público y privado. Dirige la Hoja de Ruta de IA de Dubái.
[13] Gobernanza de datos	Transparencia de datos	Dubai Pulse : el portal de datos abiertos de la ciudad. Su objetivo es aprovechar al máximo los recursos públicos disponibles, haciendo que los datos generados o custodiados por organismos públicos sean accesibles, gratuitos y utilizables por todas las personas.
	Mecanismos de intercambio de datos	Herramienta de intercambio de datos : ofrece orientación y recursos a particulares y organizaciones privadas y públicas para preparar y diseñar una iniciativa de puesta en común de datos.
	Protección de datos	Marco de datos sintéticos : diseñado para ayudar a las organizaciones a adoptar tecnología de IA evitando cualquier violación de la privacidad.

Caso práctico 6: Singapur

Población: 5.918.000 (2023)	Renta per cápita: 79.996€ (2023)	Región: Asia Oriental y el Pacífico
------------------------------------	---	--

MARCO DE GOBERNANZA DE IA		
[1] Principios, estrategias y directrices	Estrategia de la IA	Estrategia Nacional de IA 2.0 (incluye un manual).
	Principios locales para la IA	Principios guía del Marco de gobernanza de IA modelo : (1) Explicabilidad, transparencia y justicia; (2) Soluciones centradas en el ser humano. Principios de gobernanza de AI Verify : (1) Transparencia; (2) Explicabilidad; (3) Repetibilidad/Reproducibilidad; (4) Seguridad; (5) Protección; (6) Solidez; (7) Justicia; (8) Gobernanza de datos; (9) Rendición de cuentas; (10) Supervisión y agencia humanas; (11) Crecimiento inclusivo y bienestar social y medioambiental.
	Directrices, guías y manuales	Manual de IA para el sector público : un recurso de la Estrategia de IA para la Administración. El manual explica la IA, muestra las aplicaciones más comunes en el sector público, explica cómo iniciar un proyecto de IA y cómo desarrollar las capacidades municipales en este campo. Marco de gobernanza de IA modelo (2ª edición) : ofrece orientaciones detalladas y de fácil aplicación sobre cómo traducir los principios éticos en recomendaciones prácticas que las organizaciones pueden adoptar para implementar IA de forma responsable. Marco de gobernanza modelo para IA generativa Guías sobre el uso de datos personales en sistemas de recomendación y toma de decisiones de IA : elaborado mediante consultas públicas . <i>Otras directrices sectoriales:</i> Directrices para IA en sanidad : ofrece recomendaciones para fomentar el desarrollo y la implementación seguros de dispositivos médicos basados en la IA y otras aplicaciones de la IA en la asistencia sanitaria. Guía para rediseñar el empleo en la era de la IA
[3] Mecanismos de transparencia y explicabilidad	Página web municipal que divulgue toda información pertinente sobre IA	Portal municipal accesible que divulga toda la información y los recursos pertinentes de IA, como proyectos emblemáticos y documentos estratégicos.
[8] Órganos consultivos y de supervisión externos/independientes	Órgano consultivo y de supervisión externo	Consejo asesor sobre el uso ético de la IA y los datos : compuesto por once miembros multidisciplinares, asiste a la Administración en la elaboración de normas éticas y marcos de gobernanza, y publica directrices consultivas, guías prácticas y códigos de prácticas para su adopción voluntaria por parte del sector empresarial.
[9] Alianzas, comunidades de práctica y grupos de aprendizaje	Coaliciones multiactor	AI Singapore : reúne a centros de investigación y al ecosistema empresarial para investigar sobre IA fiable, gobernanza ética, crear herramientas de código abierto y fomentar talento para los esfuerzos nacionales en materia de IA. Consortio Veritas : compuesto por socios de la industria y la Autoridad Monetaria de Singapur, su objetivo es permitir que las instituciones financieras evalúen sus soluciones basadas en IA y datos en relación con los principios de equidad, ética, rendición de cuentas y transparencia.
	Miembro de una comunidad de práctica	A.I. Verify Foundation : una comunidad mundial de código abierto que reúne a propietarios de aplicaciones de IA, proveedores de soluciones, usuarios y responsables políticos para crear una IA fiable.

[10] Programas de desarrollo de capacidades	Desarrollo de capacidades municipal	Formación del personal municipal: personalizada para los distintos tipos de usuarios municipales, Singapur cuenta con un directorio de cursos para lograr diversas competencias (incluido en el Manual de IA para el sector público).
[11] Fomento de la innovación, el conocimiento y la experimentación locales	Centros, <i>hubs</i> y laboratorios de innovación de IA	Centro para la IA y la Gobernanza de Datos (CAIDG, por sus siglas en inglés) : centro de investigación interdisciplinar con asociaciones multiactor, desde agencias gubernamentales a organizaciones intergubernamentales, empresas, universidades, grupos de reflexión, ONG y OSC.
	Entornos controlados (<i>regulatory sandboxes</i>)	Entornos controlados (<i>sandboxes</i>): GenAI Sandbox ; Sandboxes tecnológicos de protección de la privacidad
[13] Gobernanza de datos	Transparencia de datos	Portal de datos de Singapur : el portal de datos abiertos de la ciudad. Su objetivo es aprovechar al máximo los recursos públicos disponibles, haciendo que los datos generados o custodiados por organismos públicos sean accesibles, gratuitos y utilizables por todas las personas.
[14] Otras políticas y medidas	Marcos e instrumentos de ensayo	A.I. Verify : un marco de ensayo de gobernanza de IA y un conjunto de herramientas de <i>software</i> para empresas que sirven para validar el rendimiento de sistemas de IA con respecto a un conjunto de principios éticos reconocidos internacionalmente mediante ensayos estándares. Guía de implementación y autoevaluación para organizaciones (ISAGO por sus siglas inglés) : ayuda a organizaciones a evaluar la adecuación de sus prácticas de gobernanza de IA respecto al Marco de gobernanza modelo. Kit de herramientas de código abierto Veritas : permite el uso responsable de IA en el sector financiero.
	Certificaciones verdes	Hoja de ruta para la certificación verde de centros de datos : traza una vía sostenible para el crecimiento continuo de centros de datos en Singapur. Certificado verde de centros de datos : para los operadores que han implantado con éxito las mejores prácticas ecológicas en centros de datos, demostrando una sostenibilidad y unos resultados medioambientales superiores.

CONCLUSIONES. MECANISMOS DE POLÍTICA, DESAFÍOS Y RECOMENDACIONES EN IA URBANA

Marta Galceran-Vercher

Investigadora sénior, Programa de Ciudades Globales, CIDOB

Alexandra Vidal D'oleo

*Investigadora y gestora de proyectos,
Programa de Ciudades Globales, CIDOB*

Las administraciones locales de todo el mundo recurren cada vez con más frecuencia a sistemas algorítmicos para mejorar la prestación de servicios públicos. Sin embargo, hay pruebas crecientes de que estos sistemas pueden ocasionar daños no deseados y carecen de transparencia en su aplicación. En consecuencia, la adopción de sistemas algorítmicos suele ir acompañada de la elaboración de principios rectores para el uso responsable de las tecnologías de inteligencia artificial (IA), principalmente a nivel nacional, supranacional o global. Algunos ejemplos notables son los Principios de la OCDE sobre IA (2019), los Principios del G20 sobre IA (2019), el grupo de redacción del Convenio del Consejo de Europa sobre IA (2022-2024), la Declaración Ministerial de la Asociación Mundial sobre Inteligencia Artificial (2022), la Declaración de los Ministros del G7 (2023), la Declaración de Bletchley (2023), la Declaración Ministerial de Seúl (2024), la Ley de la Unión Europea sobre IA (2024) o el Informe de la ONU «Gobernar la IA para la Humanidad» (2024). Sin embargo, estos marcos suelen limitarse a proporcionar directrices generales sobre lo que constituye un uso responsable de la IA, ofreciendo escasas indicaciones prácticas sobre cómo deben aplicarse estos principios en contextos reales.

En respuesta a estos desafíos, muchos gobiernos están recurriendo a marcos reguladores e instrumentos de política para hacer operativos estos principios. Aunque estos esfuerzos están surgiendo con rapidez, su alcance y enfoque varían considerablemente. Además, gran parte de los análisis existentes sobre los instrumentos normativos del ámbito público tienden a centrarse en perspectivas de alcance nacional (por ejemplo, OCDE, 2024), pasando por alto a menudo el contexto y los desafíos específicos a los que se enfrentan los gobiernos locales.

La presente monografía CIDOB ha tratado de llenar este vacío mediante la identificación de los principales mecanismos de política y marcos impulsados por gobiernos locales para garantizar que su adopción de sistemas algorítmicos se alinea con principios éticos fundamentales tales como la transparencia y la rendición de cuentas, la equidad y la no discriminación, la privacidad y gobernanza de datos, y la sostenibilidad. Este

La presente monografía CIDOB ha identificado los principales mecanismos de política y marcos impulsados por gobiernos locales para garantizar que su adopción de sistemas algorítmicos se alinea con principios éticos fundamentales.

El criterio utilizado para establecer esta categorización se basan en la función y objetivos primarios de los mecanismos de política.

análisis se complementa con una serie de casos prácticos que ilustran la aplicación práctica de estos mecanismos de política por parte de las principales ciudades, lo que da lugar a marcos locales de IA integrales.

En este capítulo final, ofrecemos una categorización de los mecanismos de política identificados a lo largo de esta publicación, así como una visión de cuáles son los mecanismos más empleados por los gobiernos locales y cómo se alinean con los principios éticos antes mencionados. Asimismo, se analizan los desafíos a los que se enfrentan los gobiernos locales y se ofrecen recomendaciones para avanzar hacia un uso más responsable de los sistemas de IA en los entornos urbanos.

1. Categorización de los mecanismos de política

Tras una revisión exhaustiva de los mecanismos de política presentados a lo largo de los capítulos de esta monografía CIDOB, complementada con una revisión bibliográfica de publicaciones relevantes sobre mecanismos de política para gobiernos locales y/o administraciones públicas (entre ellas, los informes elaborados por Ada Lovelace Institute, AI Now Institute y Open Government Partnership, 2021; Ben Dhaou *et al.*, 2024; Jordan *et al.*, 2024; United 4 Smart Sustainable Cities, 2024), hemos identificado catorce categorías distintas de políticas que están siendo implementadas actualmente por gobiernos locales de todo el mundo (véase el cuadro 1)¹.

El criterio utilizado para establecer esta categorización se basa en la función y objetivos primarios de los mecanismos de política. Esto incluye proporcionar orientación normativa para el desarrollo y uso de sistemas de IA, evaluar los riesgos potenciales de los algoritmos, garantizar el acceso público a la información sobre los sistemas algorítmicos y exigir rendición de cuentas.

Otros posibles criterios podrían haberse centrado en si los mecanismos están orientados a procesos administrativos internos (por ejemplo, directrices para el personal municipal o la creación de comisiones municipales de IA) o a acciones dirigidas a audiencias externas, como la publicación de registros públicos de algoritmos o la imposición de prohibiciones a aplicaciones de IA que susciten controversia. Además, nuestra categorización no distingue entre mecanismos específicos de la IA (por ejemplo, una estrategia de IA) y mecanismos indirectos que contribuyen a la gobernanza ética de la IA (por ejemplo, las estrategias de gobernanza de datos, que, aunque tienen un alcance más amplio, son fundamentales para la gobernanza de la IA debido a la importancia de los datos en los sistemas de IA).

1. Es importante señalar que la escasa literatura sobre este tema emplea una serie de nombres y términos para los diversos mecanismos de política, y que no existe un vocabulario común para sus componentes básicos.

Cuadro 1. Categorías de mecanismos de política para una gobernanza responsable de la IA en las ciudades

CATEGORÍA	DESCRIPCIÓN Y FINALIDAD	MECANISMOS DE POLÍTICA
[1] Principios, estrategias y directrices	Documentos de política que ofrecen orientación normativa no vinculante sobre principios y valores éticos para las administraciones locales, esbozando directrices generales para desarrollar y utilizar la IA y, al mismo tiempo, gestionar los riesgos asociados. Aunque el formato varía, estos documentos suelen identificar objetivos políticos de alto nivel y su relevancia para el uso de sistemas algorítmicos por parte de los organismos públicos. En algunos casos, también ofrecen directrices prácticas para aplicar dichos principios en el diseño y la implementación de tales sistemas. En última instancia, estas directrices establecen estándares normativos que permiten a los organismos y al público evaluar el uso ético de los sistemas algorítmicos.	<ul style="list-style-type: none"> • Estrategias de IA éticas, planes de acción, agendas, hojas de ruta, estatutos, manuales, etc. • Definición de principios éticos locales para la IA: declaración y/o adhesión • Directrices, guías y manuales sobre cómo implementar una IA ética • Protocolos internos para el uso de IA • Marcos estratégicos no relacionados específicamente con la IA que repercuten en la gobernanza de la IA (por ejemplo, marcos de digitalización o gobernanza de datos)
[2] Normativa y legislación locales	Herramientas destinadas a establecer normas, leyes y reglamentos que garanticen el cumplimiento y aborden las repercusiones sociales.	<ul style="list-style-type: none"> • Normativa y leyes locales (por ejemplo, normativa para garantizar el derecho a la justificación, etc.) • Mecanismos de cumplimiento legal: para garantizar el cumplimiento de los marcos normativos regionales, nacionales o supranacionales • Normas reglamentarias (por ejemplo, normas ecológicas de IA, normas de transparencia) • Cumplimiento de las normas reglamentarias internacionales • Regulación de aplicaciones controvertidas de IA: prohibiciones, moratorias, etc.
[3] Mecanismos de transparencia y explicabilidad	Mecanismos para establecer el acceso público a la información sobre sistemas y procesos algorítmicos. Tienen por objeto proporcionar información sobre los sistemas algorítmicos al público en general (por ejemplo, personas afectadas, medios de comunicación o sociedad civil) para que pueda conocer el uso de estos sistemas y exigir explicaciones y justificaciones relacionadas con dicho uso. Estos mecanismos pueden funcionar de forma independiente o como parte de marcos más amplios de rendición de cuentas en materia de algoritmos. Es importante distinguir los mecanismos de transparencia pública de los derechos de audiencia y explicación, que otorgan a las personas el derecho a una explicación de las decisiones algorítmicas específicas que se hayan tomado sobre ellas.	<ul style="list-style-type: none"> • Registros públicos de algoritmos • Página web municipal que divulgue toda información pertinente sobre IA • Estándar de transparencia algorítmica • Directorio municipal de herramientas de IA adquiridas (de uso interno) • Requisitos de transparencia del código fuente • Explicaciones de lógica algorítmica
[4] Evaluación del impacto algorítmico	Instrumento normativo utilizado por los organismos públicos para evaluar los riesgos y daños potenciales de los sistemas algorítmicos. Estas evaluaciones tienen como finalidad comprender, clasificar y abordar los posibles efectos negativos de los algoritmos antes o durante su implementación. Las evaluaciones del impacto algorítmico (EIA) se basan en marcos establecidos en otros ámbitos, como las evaluaciones del impacto ambiental, los derechos humanos y las evaluaciones del impacto de la protección de datos (EIPD).	<ul style="list-style-type: none"> • Procedimientos de evaluación y gestión de riesgos (incluido el análisis de sesgos) • Evaluaciones del impacto sobre los derechos humanos • Mecanismos de evaluación del impacto ambiental
[5] Auditorías e inspección reglamentaria	Las auditorías abarcan una serie de prácticas dirigidas a examinar el funcionamiento de un sistema algorítmico específico. Su objetivo principal es comprender las operaciones del sistema y evaluar su prestación con respecto a normas predefinidas. Aunque las auditorías presentan similitudes con las evaluaciones del impacto algorítmico (EIA), tienen un origen distinto y suelen realizarse durante o después de la implementación del sistema. En cambio, las EIA suelen realizarse antes o durante la implementación. Las auditorías pueden ser realizadas por entidades internas, externas o terceras, en función del alcance y la naturaleza de la evaluación. En una auditoría de terceros, una organización externa evalúa el sistema basándose únicamente en sus resultados. Una auditoría de segunda parte es realizada por un evaluador externo que tiene acceso tanto a la administración del sistema del sistema como a sus resultados. Las auditorías de primera parte están a cargo de miembros internos de la organización.	<ul style="list-style-type: none"> • Auditorías de los sistemas algorítmicos • Evaluación de procesos • Herramientas de autoevaluación

[6] Supervisión humana, rendición de cuentas, procedimientos de audiencia y recurso	Mecanismos para supervisar y exigir responsabilidades a los sistemas de IA. En concreto, se refieren a mecanismos que obligan a que las decisiones tomadas con la ayuda de sistemas algorítmicos sigan procedimientos específicos diseñados para salvaguardar la equidad y proporcionar vías para que las personas puedan obtener reparación en caso de resultados sesgados o erróneos. Estas garantías de procedimiento ofrecen a las personas o grupos afectados la oportunidad de impugnar o recurrir las decisiones que les afectan.	<ul style="list-style-type: none"> • Seguimiento e informes internos • Requisitos para la participación humana • Procedimientos de respuesta y objeción accesibles a la ciudadanía • Obligación de notificar la decisión y la oportunidad de explicarse a las partes afectadas • Obligación de motivar y explicar las decisiones • Mecanismos para garantizar el derecho de las partes afectadas a presentar pruebas, recurrir e impugnar las decisiones automatizadas
[7] Condiciones de contratación pública	Las normas que rigen la adquisición de sistemas algorítmicos por parte de gobiernos y organismos públicos son cruciales para garantizar su uso responsable. Muchos de los sistemas algorítmicos utilizados por las administraciones públicas se subcontratan a proveedores privados, ya sea mediante la compra de productos o mediante contratos de servicios. Por consiguiente, los proveedores desempeñan un papel importante en el diseño y la implementación de estos sistemas. Las condiciones de los contratos de adquisición son esenciales para configurar el desarrollo y la aplicación de estos sistemas. Para garantizar que los sistemas adquiridos cumplan las normas éticas y se utilicen de forma responsable, pueden aplicarse condiciones de contratación específicas, tales como requisitos de transparencia y no discriminación.	<ul style="list-style-type: none"> • Cláusulas de contratación pública • Directrices, marcos y protocolos internos para la adquisición de IA
[8] Órganos consultivos y de supervisión	Órganos de supervisión independientes, que tienen por objeto supervisar y dirigir el uso de los sistemas algorítmicos por parte de los organismos públicos. Estos mecanismos de supervisión independientes tienen por objeto garantizar la rendición de cuentas mediante la supervisión de las acciones de los organismos públicos y la formulación de recomendaciones, sanciones o decisiones sobre el uso que estos organismos hacen de los sistemas algorítmicos.	<ul style="list-style-type: none"> • Órganos consultivos: consejos, comités, juntas, redes, grupos, etc.
[9] Alianzas, comunidades de práctica y grupos de aprendizaje	Mecanismos destinados a fomentar la cooperación y las asociaciones a escala local, nacional e internacional.	<ul style="list-style-type: none"> • Comunidades de práctica sobre aprendizaje a nivel local/nacional/internacional: redes de ciudades, grupos de trabajo, etc. • Colaboraciones multiactor locales/nacionales/internacionales en el ámbito de la IA: redes, plataformas, coaliciones, etc. • Colaboración y asociación entre los sectores público y privado
[10] Programas de desarrollo de capacidades	Mecanismos para ampliar los conocimientos y crear competencias en torno a la inteligencia artificial ética. Estas iniciativas pueden dirigirse al personal municipal implicado –directa o indirectamente– en el diseño, la implementación o el uso de sistemas algorítmicos, así como al público en general, para promover una ciudadanía informada y fomentar una comprensión más amplia de la ética de la IA.	<ul style="list-style-type: none"> • Formación del personal municipal (enfoque socio-técnico) • Creación de equipos municipales de IA • Organismo municipal de IA: organismo o comisión transversal que coordina/supervisa el uso municipal de la IA • Enfoque multidisciplinar: creación de equipos diversos
[11] Fomento de la innovación, el conocimiento y la experimentación locales	Mecanismos que ofrecen un espacio para la experimentación, la innovación y la realización de pruebas en entornos reales.	<ul style="list-style-type: none"> • Promoción y colaboración con centros, <i>hubs</i> y laboratorios locales de innovación de IA • Observatorios locales de IA • Entornos controlados (regulatory sandboxes) locales para la IA • Iniciativas para promover y apoyar los ecosistemas locales de IA
[12] Participación de la comunidad	Herramientas para involucrar a los ciudadanos, las comunidades y las partes interesadas en los procesos de toma de decisiones relacionados con la IA; fomentar discusiones y debates, y garantizar que las políticas de IA reflejen las preocupaciones y aportaciones de la población.	<ul style="list-style-type: none"> • Participación popular: procesos participativos, modelos de gobernanza participativa, sesiones de escucha a la ciudadanía, fomento de espacios de reflexión y debate, canales de comunicación con la ciudadanía, etc. • Educación de la ciudadanía (alfabetización digital) • Comités locales de ética en materia de IA
[13] Gobernanza de datos		<ul style="list-style-type: none"> • Medidas de transparencia de datos, como portales de datos abiertos • Mecanismos de intercambio de datos • Derechos sobre los datos • Protocolos de utilización de los datos • Protección de datos: leyes de protección de datos, marcos de datos sintéticos, etc. • Sistemas de gobernanza de datos
[14] Otras políticas y medidas		<ul style="list-style-type: none"> • Marcos e instrumentos de ensayo • Incentivos fiscales como bonificaciones fiscales, subvenciones, etc. • Programas de reciclaje profesional • Marcos de clasificación (por ejemplo, marcos de clasificación por estrellas de la IA, certificaciones verdes, etc.)

Fuente: Autoras

2. Alineamiento de los mecanismos de política con principios éticos

El análisis de los mecanismos de política identificados a lo largo de esta publicación² revela que pueden establecerse varias tipologías en relación con su alineamiento con principios éticos específicos (véase el cuadro 2), que se discuten a continuación.

- a) **Mecanismos de política destinados a defender determinados principios éticos.** Por ejemplo, una evaluación del impacto ambiental persigue principalmente fines de sostenibilidad, mientras que una medida de intervención humana puede defender tanto el principio de rendición de cuentas como el de equidad al garantizar que una persona supervisa el correcto funcionamiento de un sistema de IA, remitiendo en última instancia la decisión final a un ser humano.
- b) **Mecanismos de política que defienden todos los principios éticos.** Debido a su naturaleza transversal o a través de la personalización, algunos mecanismos de política pueden promover todos los principios éticos colectivamente. Por ejemplo, mecanismos de política como las estrategias de IA pueden personalizarse para incluir todos los principios éticos. Del mismo modo, un comité de supervisión puede encargarse de supervisar la protección de la privacidad, la rendición de cuentas o todo el conjunto de principios éticos.

A pesar de su potencial, no todos los mecanismos de política son aplicados con la misma frecuencia por las ciudades que se esfuerzan por establecer marcos locales éticos de IA. Algunos ejemplos destacables entre estos mecanismos son:

- **Principios, estrategias y directrices:** uno de los mecanismos más frecuentemente aplicados por las ciudades de todo el mundo. Las ciudades aplican sistemáticamente estos mecanismos de política desde un punto de vista ético para proporcionarse unos fundamentos y orientación. Estos mecanismos son especialmente utilizados por los municipios para demostrar su voluntad de comprometerse con una implementación de IA responsable.
- **Cláusulas de contratación pública:** dado que los ayuntamientos carecen a menudo de recursos para desarrollar sus propios sistemas internos de IA, otro mecanismo de política comúnmente aplicado son las cláusulas de contratación pública, que han llegado a ser un elemento esencial y una solución práctica. Las cláusulas permiten a los ayuntamientos aprovechar su poder adquisitivo para dotarse de sistemas de IA, al tiempo que promueven un desarrollo ético de la IA por parte de los proveedores del sector privado.
- **Mecanismos orientados al exterior:** cada vez son más las ciudades que recurren a **órganos de asesoramiento y supervisión** constituidos por expertos externos e independientes que asesoran al municipio sobre los dilemas éticos y supervisan su uso de sistemas algorítmicos. Del mismo modo, muchos municipios están participando en **alianzas, comunidades de práctica y grupos de aprendizaje** para abordar conjuntamente desafíos e identificar formas de utilizar la IA de forma segura.

El análisis de los mecanismos de política identificados a lo largo de esta publicación revela que pueden establecerse varias tipologías en relación con su alineamiento con principios éticos específicos.

2. Extraídos únicamente de los capítulos (primera parte) y casos prácticos (segunda parte) de esta monografía.

- **Gobernanza de datos:** aunque la gobernanza de datos puede no percibirse inicialmente como un mecanismo de política directamente relacionado con la IA, sienta las bases para una correcta gestión y salvaguardia de los datos de la ciudadanía, y es crucial para establecer sistemas no discriminatorios, lo que la convierte en un elemento esencial para una implementación ética de IA. La gobernanza de datos es un componente básico que permite la transparencia de los datos, la protección de los derechos sobre los datos, la privacidad de los datos y el uso sostenible de los datos en los sistemas de IA. Algunos ejemplos son los protocolos para la anonimización de datos personales o el uso de datos sintéticos para entrenar sistemas de IA, con el fin de resolver problemas de privacidad y equidad.
- **Auditorías:** a pesar de ser universalmente reconocidas por expertos y funcionarios como uno de los mecanismos de política más importantes para salvaguardar la ética, las auditorías siguen infrautilizándose. Su uso poco frecuente se debe en gran medida a las limitaciones impuestas por los proveedores privados de IA y a la falta de capacidad técnica interna.

c) **Mecanismos de política no asociados a ningún principio ético específico.** Aunque no están vinculados a ningún principio ético en particular, estos mecanismos de política se consideran clave para una implementación ética de la IA. Establecen procesos estructurados a seguir, coordinan su implementación o proporcionan los conocimientos y la competencia necesarios para una toma de decisiones informada. Algunos ejemplos notables de estos mecanismos son:

- **Protocolos internos para el uso de IA:** la mayoría de las ciudades elaboran protocolos internos para orientar su uso de IA, proporcionando estructuras detalladas que facilitan su aplicación por parte de la administración municipal. Algunas ciudades, aunque una minoría, los complementan además con **protocolos exhaustivos relativos a la contratación de servicios de IA.** Estos últimos pueden incluir evaluaciones del impacto obligatorias, análisis de sesgos y otros mecanismos de política destinados a garantizar el respeto de principios éticos específicos.
- **Centros, hubs y laboratorios de innovación de IA:** un número significativo de ciudades colaboran, promueven o han creado centros, hubs y laboratorios de innovación de IA para generar conocimientos prácticos y desarrollar soluciones de IA. La investigación que llevan a cabo estas instituciones está orientada desde un punto de vista ético.
- **Programas de desarrollo de capacidades:** uno de los mecanismos menos aplicados es la creación de **equipos municipales dedicados a la IA** con experiencia para auditar o desarrollar sistemas internos de IA. Esto se debe principalmente a limitaciones técnicas y financieras por parte de los municipios. En cambio, muchos han creado **organismos municipales responsables de coordinar el uso de la IA** en todos los departamentos. Estos organismos desempeñan un papel fundamental a la hora de facilitar el seguimiento transversal de la implementación de la IA en el municipio, garantizando un enfoque más organizado y responsable de la gobernanza de la IA.

Cuadro 2. Alineamiento de los mecanismos de política con principios éticos

MECANISMOS DE POLÍTICA		PRINCIPIOS ÉTICOS			
		Rendición de cuentas y transparencia	Privacidad y gobernanza de datos	Equidad y no discriminación	Sostenibilidad
[1]	Estrategias de IA	X	X	X	X
	Principios éticos locales para la IA	X	X	X	X
	Directrices, guías y manuales	X	X	X	X
	Protocolos internos para el uso de IA				
[2]	Normativas y leyes locales	X	X	X	X
	Mecanismos de cumplimiento legal	X	X	X	X
	Normas reglamentarias	X	X	X	X
	Normas reglamentarias internacionales	X	X	X	X
	Regulación de aplicaciones controvertidas de IA		X	X	
[3]	Registro público de algoritmos	X	X	X	
	Web municipal que divulgue toda información pertinente sobre IA	X			
	Directorio municipal de herramientas de IA adquiridas (de uso interno)	X			
[4]	Evaluación y gestión de riesgos	X	X	X	X
	Evaluaciones del impacto sobre los derechos humanos	X	X	X	X
	Evaluación del impacto ambiental				X
[5]	Auditorías	X	X	X	X
	Herramientas de autoevaluación	X	X	X	X
[6]	Seguimiento e informes internos				
	Participación humana	X		X	
	Procedimientos de respuesta y objeción accesibles a la ciudadanía	X		X	
[7]	Cláusulas de contratación pública	X	X	X	X
	Protocolos internos para la adquisición de IA				
[8]	Órganos consultivos y de supervisión	X	X	X	X
[9]	Alianzas, comunidades de práctica y grupos de aprendizaje	X	X	X	X
[10]	Formación del personal municipal	X	X	X	X
	Equipo municipal de IA				
	Organismo municipal de IA				
	Enfoque multidisciplinar			X	
[11]	Centros, hubs y laboratorios de innovación de IA				
	Observatorios locales de IA				
	Entornos controlados (regulatory sandboxes)		X		
[12]	Participación popular	X		X	X
	Educación de la ciudadanía (alfabetización digital)		X	X	X
	Comités locales de ética en materia de IA	X	X	X	X
[13]	Gobernanza de datos	X	X	X	X
[14]	Marcos y herramientas de evaluación	X	X	X	X
	Incentivos fiscales (bonificaciones fiscales, subvenciones, etc.)				X
	Programas de reciclaje profesional				X
	Marcos de clasificación	X	X	X	X

Fuente: Autoras

Leyenda de la tabla: Amarillo (mecanismos de política alineados con un principio ético específico o varios simultáneamente); Azul (mecanismos de política transversales o personalizables que sirven a todos los principios éticos); Verde (mecanismos de política no asociados a un principio ético específico pero relevantes para una implementación responsable de los sistemas algorítmicos en general). Gris claro (mecanismos de política no mencionados en los artículos); Gris oscuro (mecanismos de política mencionados explícitamente en los artículos, véase el anexo I).

Independientemente de su nivel de especificidad, alcance o eficacia, los mecanismos de política no se aplican de manera uniforme por las ciudades, sino que varían en frecuencia debido a las restricciones de recursos, las limitaciones técnicas o las diferentes prioridades locales. El cuadro 3 ofrece una comparación de las ciudades incluidas en los casos prácticos, destacando la frecuencia con la que se aplican determinados mecanismos de política, cuáles son los más empleados y a cuáles se recurre raramente. Aunque se ha considerado pertinente ofrecer una comparativa de las ciudades analizadas en la segunda parte de esta publicación, se reconoce a su vez que las conclusiones extraídas de los casos prácticos seleccionados están limitadas por el número reducido de ciudades estudiadas.

Cuadro 3. Comparación de los casos prácticos (mecanismos de política más utilizados)

MECANISMOS DE POLÍTICA		CIUDADES					
		Barcelona	Ámsterdam	Nueva York	San José	Dubái	Singapur
[1]	Estrategias de IA, agendas, planes de acción, manuales, hojas de ruta, etc.	x	x	x	x	x	x
[1]	Principios éticos locales para la IA	x	x	x	x	x	x
[3]	Página web municipal que divulgue toda información pertinente sobre IA	x	x	x	x	x	x
[13]	Portal de datos abiertos (transparencia de datos)	x	x	x	x	x	x
[1]	Directrices, guías y manuales		x	x	x	x	x
[8]	Órganos consultivos y de supervisión	x	x	x	x		x
[9]	Miembro de una comunidad de práctica	x	x	x	x		x
[1]	Protocolos internos para el uso de IA	x	x	x	x		
[4]	Evaluación y gestión de riesgos	x	x	x	x		
[7]	Cláusulas de contratación pública	x	x	x	x		
[12]	Participación popular	x	x	x	x		
[3]	Registro público de algoritmos	x	x		x		
[4]	Evaluaciones del impacto sobre los derechos humanos	x	x		x		
[9]	Colaboraciones multiactor en el ámbito de la IA		x		x		x
[10]	Formación del personal municipal (desarrollo de capacidades municipal)	x	x				x
[10]	Organismo municipal de IA	x		x	x		
[11]	Centros, <i>hubs</i> y laboratorios de innovación de IA		x			x	x
[13]	Otras políticas de gobernanza de datos (derechos sobre los datos, mecanismos de intercambio de datos, protocolos de datos, etc.)	x			x	x	
[2]	Estándar de transparencia algorítmica	x	x				
[6]	Seguimiento e informes internos			x	x		
[6]	Procedimientos de respuesta y objeción accesibles a la ciudadanía		x		x		
[5]	Auditorías	x	x				
[2]	Regulación de aplicaciones controvertidas de IA			x			
[3]	Directorio municipal de herramientas de IA adquiridas (de uso interno)			x			
[10]	Creación de equipos municipales de IA (desarrollo de capacidades municipal)		x				
[11]	Observatorio local de IA	x					
[11]	Entornos controlados (regulatory sandboxes)						x
[14]	Marcos e instrumentos de ensayo						x
[14]	Marcos de clasificación						x

Fuente: Autoras

Nota: La lista de mecanismos de política se ha ordenado, en primer lugar, de más a menos recuente; en segundo lugar, por categorización.

3. Desafíos y recomendaciones

- a. **Existen pocas referencias sobre cómo poner en práctica eficazmente principios éticos para la IA a nivel local.** La mayoría de las directrices, estudios y programas de capacitación existentes no tienen en cuenta los retos específicos a los que se enfrentan las administraciones urbanas, que con frecuencia están desconectadas de las estrategias y políticas nacionales. Este vacío se ve agravado por la naturaleza heterogénea de las administraciones locales, que varía mucho en términos de tamaño, recursos y capacidades.

Recomendaciones:

- Localizar (es decir, adaptar al contexto local) los principios éticos regionales, nacionales o globales y los mecanismos de política de gobernanza de la IA mediante la creación de definiciones locales de éxito y la identificación de prioridades locales.
- b. **Las administraciones locales se enfrentan a una importante falta de talento y conocimientos técnicos**, un reto que se agudiza por la escasez mundial de profesionales especializados en IA, lo que dificulta atraer a personas cualificadas a nivel local. En consecuencia, las administraciones municipales suelen comprender de forma limitada las posibles repercusiones e implicaciones de los sistemas algorítmicos.

Recomendaciones:

- Dar prioridad a programas de capacitación, como parte de las estrategias y marcos de política municipales de gobierno y regulación de sistemas algorítmicos. Para ello, habría que asignar recursos específicos a los programas de formación municipales, invertir en campañas de concienciación de la ciudadanía y promover iniciativas para crear conocimientos y habilidades fundamentales en torno a la IA ética dentro de la administración pública.
 - Para superar los desafíos que supone atraer talento local, las administraciones locales deben invertir en estrategias que faciliten el intercambio y la adaptación de conocimiento de actores locales. Además, el establecimiento de alianzas y conexiones sólidas con redes de intercambio puede ayudar a reducir lagunas de conocimientos.
 - Adoptar un enfoque holístico para el desarrollo de capacidades, fomentando debates públicos e iniciativas de sensibilización en las comunidades locales. Estos esfuerzos deben centrarse en educar a la ciudadanía sobre las oportunidades y los riesgos asociados al uso de sistemas algorítmicos.
- c. Garantizar la **transparencia y la rendición de cuentas de sistemas algorítmicos** utilizados por administraciones locales plantea varios retos, incluidos: la gestión de la percepción de la ciudadanía y las posibles reacciones negativas ante sistemas de IA orientados al público, la adaptación de la cultura organizativa y dinámicas de trabajo para sistemas de IA internos, y el fomento de la responsabilidad compartida en el conjunto de la administración pública (en esencia, la IA no debería considerarse responsabilidad únicamente del departamento de informática).

Independientemente de su nivel de especificidad, alcance o eficacia, los mecanismos de política no se aplican de manera uniforme por las ciudades, variando en frecuencia debido a las restricciones de recursos, las limitaciones técnicas o las diferentes prioridades locales.

Recomendaciones:

- Incorporar la transparencia como objetivo central, más allá de políticas específicas de IA. Esto incluye fomentar una cultura de transparencia y rendición de cuentas a lo largo de todo el ciclo de vida de la IA.
 - Fomentar el uso de código abierto, que puede aumentar la confianza y permitir un mayor grado de control y colaboración.
 - Aclarar las responsabilidades mediante la creación de una persona de contacto específica o «responsable del proyecto» para todas las iniciativas de IA, garantizando la rendición de cuentas y una comunicación eficiente.
 - Crear múltiples canales de retroalimentación e integrar evaluaciones en las fases clave del proyecto para garantizar la mejora continua y la capacidad de respuesta.
 - Asignar presupuesto para fases explicativas exhaustivas, garantizando que las partes interesadas, tanto internas como externas, comprendan plenamente los sistemas de IA y sus implicaciones.
- d.** Garantizar la **protección de la privacidad** presenta retos específicos, entre ellos un complejo panorama normativo en el que las administraciones locales a menudo tienen dificultades para desenvolverse. Las cuestiones relacionadas con datos están estrechamente vinculadas a mecanismos de política diseñados para salvaguardar la protección de datos. En particular, la disponibilidad de datos de alta calidad en entornos urbanos es limitada, lo que puede atribuirse a varios factores: prácticas inadecuadas de gestión de datos, riesgos e inquietudes éticas derivados de la recopilación de datos a gran escala, y escaso intercambio de datos entre administraciones debido a la ausencia de estándares unificados y a marcos de gobernanza subdesarrollados.

Recomendaciones:

- Generar conjuntos de datos públicos de gran valor mejorando las prácticas de recopilación y gestión de datos para aumentar su calidad y utilidad.
 - Promover la interoperabilidad y la colaboración entre organismos y sectores para facilitar el intercambio y la puesta en común de datos.
 - Crear marcos seguros y transparentes para el intercambio de datos que garanticen la protección de la privacidad al tiempo que permiten la innovación.
 - Fomentar la innovación y la experimentación en contextos regulados, como los entornos controlados (*regulatory sandboxes*), para probar nuevas soluciones basadas en datos de forma segura y responsable.
- e.** Los sistemas basados en algoritmos pueden reforzar las desigualdades urbanas existentes y crear nuevas formas de discriminación, de ahí la importancia de prestar atención a la **noción de equidad y no discriminación** cuando las administraciones locales implementan

sistemas de IA. Un reto específico a este respecto incluye el hecho de que la discriminación automatizada mediante IA es más abstracta, opaca, difícil de detectar (cajas negras) y a gran escala y, por tanto, altera los recursos y procedimientos jurídicos tradicionales que suelen emplear las administraciones locales para detectarla, prevenirla y corregirla.

Recomendaciones:

- Considerar los múltiples papeles de las administraciones públicas – como desarrolladoras, implementadoras y reguladoras– a la hora de diseñar iniciativas para mejorar la equidad y la no discriminación de los sistemas algorítmicos.
 - Localizar marcos de política existentes para abordar la discriminación no intencionada en sistemas algorítmicos, con el fin de garantizar que se adaptan a los retos únicos que plantean los entornos urbanos.
 - Introducir un enfoque holístico para la gobernanza de la IA en las administraciones locales, teniendo en cuenta las repercusiones socioeconómicas a lo largo de todo el ciclo de vida de la IA, desde el diseño hasta la implementación.
 - Garantizar la diversidad entre los equipos que participan en el diseño y la implementación de sistemas algorítmicos para reducir el riesgo de resultados sesgados y promover soluciones inclusivas.
- f. El principal reto asociado al principio de **sostenibilidad medioambiental de la IA** tiene que ver con el hecho de que el uso de la IA para la sostenibilidad a menudo entra en conflicto con la sostenibilidad de la propia IA. Al mismo tiempo, existen pocos marcos para la sostenibilidad de la IA con un enfoque urbano.

Recomendaciones:

- Evaluar el impacto de todo el ciclo de vida de los sistemas de IA para determinar si sus beneficios compensan sus costes medioambientales. Minimizar la huella de carbono del despliegue de IA urbana dando prioridad a sistemas energéticamente eficientes, adoptando prácticas informáticas ecológicas, utilizando *Tiny Machine Learning* y alimentando centros de datos con energías renovables.
- Fomentar una economía circular en torno a los centros de datos reduciendo los residuos electrónicos. Promover el abastecimiento responsable, la reutilización y el reciclaje del *hardware* utilizado para la IA. Fomentar la reutilización y el reciclaje de tecnologías, datos e infraestructuras relacionadas con la IA.
- Reutilizar la energía y los recursos empleados por la infraestructura de IA e implementar sistemas de IA que se integren sin dificultades en la infraestructura urbana existente, optimizando tanto el uso de la energía como la eficiencia del sistema.

Por último, es importante reconocer las limitaciones de la investigación presentada en esta monografía CIDOB. El estudio se vio limitado tanto por su marco temporal relativamente corto como por los retos inherentes que

supone recopilar información en el contexto del Observatorio Global de la Inteligencia Artificial Urbana-GOUAI (véase: Galceran Vercher y Vidal, 2024). Como resultado, se reconoce que algunos ejemplos clave de mecanismos de política algorítmica y marcos de gobernanza pueden no estar incluidos. Además, la mayoría de los datos proceden de políticas promovidas por el Norte Global, principalmente a través de intervenciones dirigidas por gobiernos locales de Estados Unidos y Europa. Este enfoque geográfico es otra limitación que las futuras investigaciones en el contexto del GOUAI tratarán de abordar. Se es consciente de que un análisis de las políticas y prácticas de gobernanza del Sur Global más sistemático podría aportar nuevas perspectivas, al revelar diferentes enfoques políticos, prioridades y desafíos de aplicación.

Referencias bibliográficas

Ada Lovelace Institute, AI Now Institute and Open Government Partnership. «Algorithmic Accountability for the Public Sector», 2021

Ben Dhaou, S., et al. «Global Assessment of Responsible Artificial Intelligence in Cities: Research and recommendations to leverage AI for people-centred smart cities». Nairobi: Programa de las Naciones Unidas para los Asentamientos Humanos (ONU-Habitat), 2024

Jordan, C.; Glickman, J. y Panettieri, A. «AI in cities: Report and Toolkit». Washington: Liga Nacional de Ciudades y Google, 2024

OECD «OECD AI Policy Observatory: Catalogue of Tools & Metrics for Trustworthy AI», 2024

United 4 Smart Sustainable Cities. «Guiding principles for artificial intelligence in cities». Ginebra: Unión Internacional de Telecomunicaciones, 2024

ANEXO 1. Lista de mecanismos de política mencionados en los capítulos de los principios éticos (Parte I)

Principio ético	Mecanismos de política
Rendición de cuentas y transparencia	<ul style="list-style-type: none"> • Evaluaciones del impacto [4] • Cláusulas de contratación pública [7] • Auditorías algorítmicas externas [5] • Registros de algoritmos [3] • Estándares de transparencia [3] • Comités interdisciplinarios de supervisión de la gobernanza [8] • Procesos participativos (a lo largo de todo el ciclo de vida de la IA) [12] • Diseño con intervención humana [6] • Formación del funcionariado [10] • Vinculación con redes de intercambio de conocimientos [9] • Colaboración de las partes interesadas locales [9]
Privacidad y gobernanza de datos	<ul style="list-style-type: none"> • Conformidad legal [2] • Sistemas de gestión de riesgos [4] • Sistemas de gobernanza de datos [13] • Evaluaciones del impacto [4] • Auditorías [5] • Repositorios de algoritmos y registros de sistemas de IA [3] • Sandboxes regulatorios [11] • Estrategias urbanas de IA [1] • Colaboraciones multiactor [9]
Equidad y no discriminación	<ul style="list-style-type: none"> • Estrategias de IA [1] • Análisis de riesgos y mecanismos de protección [4] • Evaluaciones del impacto [4] • Normas locales para una IA equitativa [2] • Normas de contratación para una IA equitativa [7] • Legislación urbana para garantizar el derecho a la justificación [2] • Órganos consultivos multidisciplinares [8] • Equipos diversos e interdisciplinarios • Auditorías [5] • Técnicas de mitigación en el ciclo de vida de la IA [14] • Redes de intercambio de conocimientos [9] • Formación y educación del personal [10] • Formación y educación de las comunidades [12] • Principios [1] • Análisis de sesgos [4]
Sostenibilidad	<ul style="list-style-type: none"> • Normas ambientales, sociales y de gobernanza (por ejemplo, normas ecológicas para la IA) [2] • Evaluaciones del impacto [4] • Supervisión y auditoría [5] • Consorcios de IA para lograr ciudades sostenibles [9] • Incentivos fiscales: bonificaciones fiscales o subvenciones [14] • Comités locales de ética en materia de IA [12] • Leyes firmes sobre la privacidad de los datos urbanos [13] • Participación social [12] • Modelos de gobernanza participativa [12] • Campañas de alfabetización digital [12] • Programas de reciclaje profesional [14] • Marcos de clasificación energética por estrellas de la IA [14] • Colaboración internacional [9] • Estandarización internacional [2]

Fuente: Autoras

Nota: Los mecanismos de política aquí descritos conservan la terminología original de los artículos sobre los principios éticos (primera parte) y se han clasificado según la categorización de las autoras que figura en el capítulo final de la monografía. Es posible que los mecanismos no aparezcan en la «sección de mecanismos de política» de los artículos, sino que se encuentren a lo largo del propio artículo.

La inteligencia artificial (IA) promete revolucionar los espacios urbanos y ofrecer soluciones a los retos urbanos más importantes a los que se enfrenta la humanidad hoy en día. Las ciudades son un campo de pruebas ideal para la IA, ya que los gobiernos locales están adoptando tecnologías basadas en datos para automatizar tareas rutinarias, mejorar la eficiencia y tomar decisiones rentables. Sin embargo, la rápida implementación de sistemas algorítmicos plantea problemas éticos, sobre todo en relación con los derechos de la ciudadanía y las repercusiones medioambientales y sociales. Disponer de una gobernanza responsable de IA es crucial para evitar consecuencias negativas no deseadas, al tiempo que se aprovechan los beneficios que esta tecnología puede aportar. Las ciudades que ya están aplicando políticas de IA ética pueden proporcionar a otras valiosos ejemplos, destacando la necesidad de que los responsables políticos locales adopten enfoques que prioricen una implementación ética, transparente e inclusiva de la IA.

En este contexto, CIDOB presenta mecanismos políticos específicos y marcos de gobernanza existentes que promueven sobre el terreno una IA urbana responsable. Por un lado, se analiza la aplicación práctica de principios éticos clave - concretamente 1) rendición de cuentas y transparencia, 2) privacidad y gobernanza de datos, 3) equidad y no discriminación, y 4) sostenibilidad- en los entornos urbanos mediante políticas concretas. Por otro lado, se presentan casos prácticos de ciudades de todo el mundo que han establecido marcos integrales de gobernanza de IA. En última instancia, esta monografía CIDOB pretende servir como documento práctico que inspire a la acción y sirva de guía a otros actores del sector público local en el largo camino que queda por recorrer para garantizar una implementación ética de IA urbana.