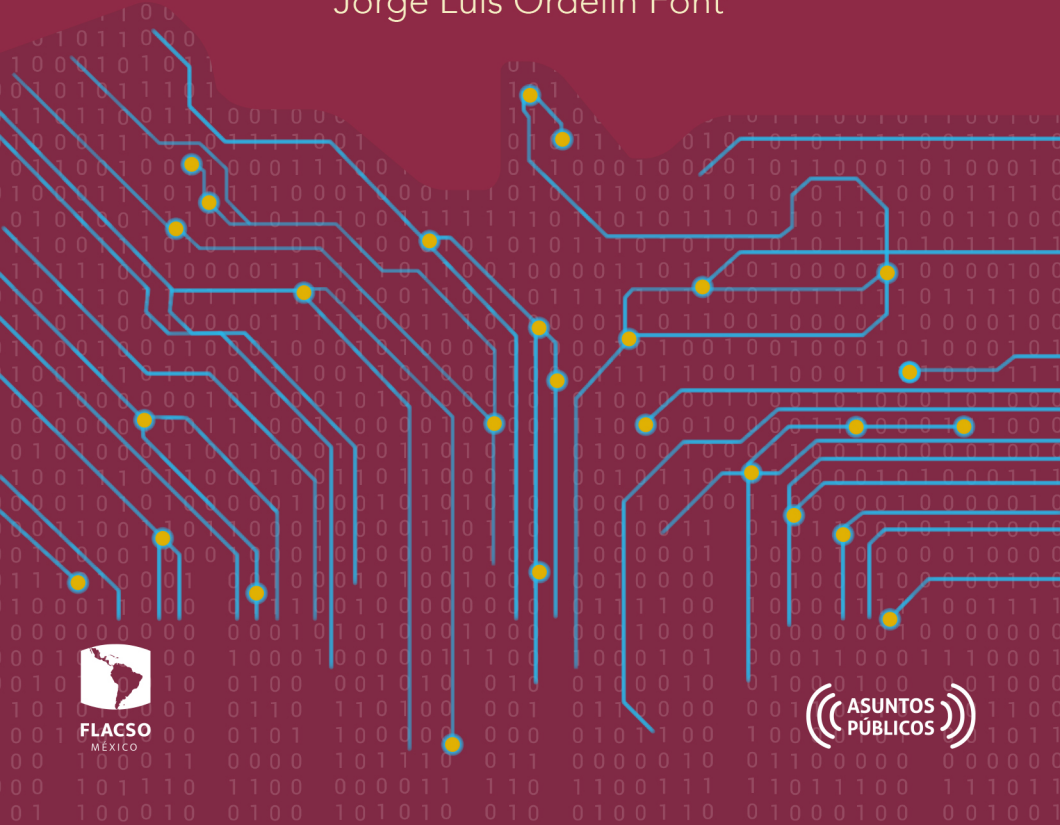
Inteligencia artificial en el ámbito público mexicano

Dónde estamos y hacia dónde vamos

César Rentería
Jorge Luis Ordelin Font



FLACSO MÉXICO

ASUNTOS PÚBLICOS

Colección Asuntos Públicos

Coordinadores
Gloria Del Castillo Alemán
Mauricio I. Dussauge Laguna

Inteligencia artificial en el ámbito público mexicano

Dónde estamos y hacia dónde vamos

César Rentería
Jorge Luis Ordelin Font

FLACSO
MÉXICO

006.30972
R422s Rentería Marín, César
Inteligencia artificial en el ámbito público mexicano : dónde estamos y hacia dónde vamos / César Rentería Marín y Jorge Luis Ordelin Font. México : FLACSO México, 2026.
220 páginas : gráficas ; 21 cm

ISBN 978-607-8817-81-8

1. Inteligencia Artificial -- México. 2. Sector Público – Innovaciones Tecnológicas -- México. 3. Internet en la Administración Pública -- México. 4. Transformación Digital -- México. 5. Algoritmos – Sector Público -- México. I. Ordelin Font, Jorge Luis, autor. II.t.

Primera edición: febrero de 2026
D.R. © 2026, Facultad Latinoamericana de Ciencias Sociales, Sede México
Carretera al Ajusco 377, Héroes de Padierna, Tlalpan, 14200 Ciudad de México
www.flacso.edu.mx | public@flacso.edu.mx | www.flacso.edu.mx/publicaciones/libros

ISBN 978-607-8817-81-8

Este libro fue sometido a un proceso de dictaminación por parte de académicos externos nacionales e internacionales de acuerdo con el Consejo Editorial de la Flacso México. Agradecemos a quienes evaluaron por su contribución a la calidad de esta obra.

Este libro es una obra literaria protegida por la Ley Federal del Derecho de Autor y los tratados internacionales firmados por México. Se prohíbe expresamente su reproducción total o parcial por cualquier medio incluyendo el digital, con o sin propósitos de lucro, si no se cuenta con el consentimiento previo y por escrito de la Facultad Latinoamericana de Ciencias Sociales, Sede México.

Impreso y hecho en México. *Printed and made in Mexico.*

Índice

Colección Asuntos Públicos
Gloria Del Castillo Alemán y Mauricio I. Dussauge Laguna 7
Prólogo
Por Mauricio I. Dussauge Laguna y Danay Quintana Nedelcu 11

Agradecimientos 15
Introducción 17

I. IA y ámbito público 31
Sistemas de IA 31
Sistemas de IA y ámbito público 35

II. Los repositorios como herramientas de transparencia algorítmica 41
Transparencia algorítmica 41
Posicionamientos institucionales sobre transparencia algorítmica 42
Transparencia algorítmica en México 46
Transparencia algorítmica y repositorios públicos 50
Experiencias previas 54

III. Estrategia de mapeo de sistemas de IA 59

IV. Panorama de los sistemas de IA en el ámbito público mexicano 75

V. Sistemas de IA en la práctica gubernamental 89
Análisis descriptivo del uso de los sistemas de IA 93
Casos ilustrativos. 108

VI. La gestión de los sistemas de IA en el servicio público: diagnóstico y recomendaciones 131
La IA como instrumento del gobierno para mejorar su relación con la ciudadanía 131
Implicaciones de la brecha en alfabetización en IA en la adopción de sistemas de IA. 137
Debilidades en capacidades organizacionales dificultan la adopción de sistemas de IA. 150
Desarrollo y compartición de sistemas de IA propios en el sector público 157
Complejidad en los procesos de adquisición y despliegue de sistemas de IA 168
Los retos de la rendición de cuentas en el creciente uso informal de IA generativa 171

Conclusiones 181

Referencias 187
Apéndices. 207

Colección Asuntos Públicos

La Facultad Latinoamericana de Ciencias Sociales, Sede México (FLACSO México), nació en 1975 como un organismo internacional latinoamericano y caribeño cuya misión y compromiso público se centran en la docencia, la investigación, la cooperación técnica y la difusión y divulgación del conocimiento en el ámbito de las ciencias sociales, con los rasgos distintivos de la multidisciplinariedad y pluralidad teórica y metodológica.

De acuerdo con nuestra vocación regional e internacional, privilegiamos la producción de conocimiento orientado a la observación, descripción y comprensión de los fenómenos sociales, económicos y políticos, sin perder de vista el enfoque estratégico que permite identificar y contribuir a la solución de los problemas públicos que afectan a las sociedades en América Latina y el Caribe.

Con esta impronta institucional, desde el segundo semestre de 2018 nos propusimos construir proyectos de naturaleza intelectual que colocaran en el centro de sus inquietudes dos temas cruciales. Primero, conocer qué debatimos a nivel nacional y regional en las ciencias sociales, para identificar, a manera de corte de caja, los avances teóricos, analíticos y metodológicos alcanzados. Segundo, entender, desde una perspectiva multidisciplinaria de las ciencias sociales, qué conexiones y aportes se registran en la comprensión y resolución de problemas públicos globales, regionales, nacionales y locales. El denominador común de esta agenda institucional es reconocer, asumir y promover el conocimiento como motor de cambio.

Bajo esta premisa, hemos impulsado y consolidado dos grandes proyectos intelectuales: el Debate Renovado e Innovador de las Ciencias Sociales, iniciado en septiembre de 2018, y la Cátedra FLACSO México, que nació en noviembre de 2020. Ambos han dado lugar

a diversos eventos y publicaciones bajo el sello editorial de nuestra institución.

En esta ocasión, presentamos la **Colección Asuntos Públicos**, una iniciativa que amplía nuestra agenda con el propósito de generar conocimiento orientado a comprender y resolver problemas públicos.

Se trata de un nuevo espacio para analizar temas sociales, económicos, políticos, demográficos, ambientales y administrativos de actualidad, con especial énfasis en sus implicaciones de política pública. Su objetivo es ofrecer una visión integral de cada asunto, pero basada en diagnósticos, literatura reciente, perspectivas comparadas y ejemplos empíricos concretos, que contribuyan tanto a la discusión y valoración de los asuntos públicos contemporáneos (de ahí el nombre de la colección), como a la formulación de políticas públicas y a la toma de decisiones gubernamentales.

Son volúmenes de formato breve y accesible para una fácil lectura por parte de diversos públicos: académicos, servidores públicos, actores políticos, integrantes de organizaciones sociales, estudiantes de educación superior y ciudadanos interesados. Cada libro incluye tres elementos esenciales:

— Un panorama general del asunto público tratado, tanto en términos de su relevancia académica, como mediante la presentación de información y datos recientes.
— Un análisis de políticas o programas públicos en curso, en México o en otros países.
— Una discusión de hallazgos empíricos, lecciones aprendidas y/o recomendaciones de política sobre el tema presentado.

La **Colección Asuntos Públicos** responde a las preocupaciones e intereses de la agenda de investigación de la FLACSO México, buscando convertirse en un punto de encuentro para académicas y académicos de casa, pero también se encuentra abierta a las contribuciones de especialistas de otras instituciones. En última instancia, se aspira a que la colección se enriquezca con las aportaciones informa-

das de quienes, por sus conocimientos y experiencia, puedan brindar ideas, análisis, comparaciones y valoraciones renovadas a la conversación académica y, sobre todo, al debate en torno a los asuntos públicos que preocupan a nuestras sociedades.

Gloria Del Castillo Alemán y Mauricio I. Dussauge Laguna
Coordinadores de la Colección Asuntos Públicos

Prólogo

Por Mauricio I. Dussauge Laguna y Danay Quintana Nedelcu

Por años vista como tema de ciencia ficción, la inteligencia artificial (IA) es hoy una realidad de la vida cotidiana. De los algoritmos que organizan nuestras preferencias televisivas a los chatbots que atienden nuestras dudas sobre los productos que deseamos comprar, pasando por los autos manejados sin conductores o los nuevos sistemas generadores de texto e imágenes, la IA está presente en múltiples ámbitos de nuestras dinámicas sociales y económicas contemporáneas. Aunque con un ritmo más lento y una menor extensión, la IA también se ha expandido ya a los entornos políticos y gubernamentales: en el diseño de campañas electorales, la atención de quejas ciudadanas o la detección de fraudes al fisco, por citar solo algunos de los ejemplos más conocidos.

El surgimiento de los sistemas de IA (posible gracias a los grandes desarrollos computacionales, la expansión digital y la acumulación de inmensas fuentes de datos de todo tipo) ha venido acompañado de un imaginario social acerca de sus inconmensurables beneficios. El desarrollo y la expansión de la IA supone modernidad, progreso tecnológico, productividad económica, eficiencia gubernamental. Un mundo mejor. En efecto, la experiencia empírica internacional ha demostrado que el uso de sistemas de IA puede mejorar el funcionamiento de las empresas, apoyar los desarrollos médicos, fortalecer la capacidad de prevención de desastres y un largo etcétera de ejemplos interesantes, valiosos y prometedores. Es indudable que está cambiando a nuestras

sociedades y que sus beneficios potenciales (económicos, científicos, de política pública) son inmensos y deben ser aprovechados.

Al mismo tiempo, son también ya numerosos los ejemplos de riesgos, fallas, sesgos y demás consecuencias negativas que los sistemas de IA pueden traer consigo. Campañas políticas manipuladas por algoritmos; reconfiguración del mundo laboral y amenaza de pérdida de numerosos empleos; violencia de género apoyada en el uso de *deepfakes*; estereotipos raciales en los sistemas de reconocimiento facial; discriminación en la asignación de prestaciones gubernamentales; o, simplemente, menor empatía social al sustituir a las personas por robots en las interacciones sociales. Por todo ello, resulta indispensable tomar cierta distancia ante los impresionantes desarrollos tecnológicos, ponderar sus limitaciones éticas y jurídicas, y reflexionar sobre las medidas sociales y normativas que deben diseñarse en paralelo a los avances tecnológicos.

Frente a este complejo escenario tecnológico/económico/político/social que caracteriza a nuestra época, en la Facultad Latinoamericana de Ciencias Sociales, Sede México (Flacso México), decidimos "entrarle al toro por los cuernos" para entender mejor de qué trata esto de la IA. Así, un grupo de docentes empezamos a organizar charlas, diplomados, cursos optativos, encuestas y seminarios para analizar, discutir y aprender sobre las múltiples aristas de la IA: el concepto, el fenómeno social, la herramienta de investigación, el mecanismo de transformación social, el instrumento de política pública, etc. Uno de los espacios que hemos creado es el "Ciclo de seminarios sobre Inteligencia Artificial, digitalización, análisis computacional e innovación en sociedades complejas" (coordinado por Mónica Casalet, Óscar Fontanelli y quienes esto escriben), en donde expertas y expertos nacionales e internacionales han discutido los orígenes de la IA, sus aspectos técnicos, sus dimensiones ético-jurídicas, sus implicaciones sociales y sus aplicaciones gubernamentales. Fue justo en el marco de dicho ciclo de seminarios que se presentó una primera versión de *Inteligencia artificial en el ámbito público mexicano. Dónde estamos y hacia dónde vamos*, libro que afortunadamente ahora se encuentra frente a usted, lectora o lector.

En esta valiosa monografía, César Rentería y Jorge Luis Ordelin Font exploran cómo los sistemas de IA se han ido integrando a las actividades y los procesos de las administraciones públicas mexicanas. Aunque el número de experiencias en nuestro país sigue siendo limitado, el cuidadoso análisis de los autores permite lograr varias cosas a la vez. En términos empíricos, se presenta un completo mapa de los tipos de sistemas de IA que se han desarrollado en los distintos ámbitos del sector público, tanto en el nivel federal como en los niveles subnacionales. Además del recuento general, Rentería y Ordelin profundizan en algunos casos especialmente interesantes, que reflejan no solo la capacidad de innovación de las instituciones públicas mexicanas, sino las posibles contribuciones al bienestar social que la IA puede aportar en países latinoamericanos. Se trata de un esfuerzo pionero por sistematizar el uso de IA en las instituciones público-gubernamentales de México. Por lo tanto, el libro será una útil línea de base para ir documentando la evolución del tema a lo largo de los años. Además, su valor aumenta en el contexto nacional, pero también latinoamericano, en donde diversos actores públicos y sociales están impulsando iniciativas de IA para el aparato y la gestión estatales. Un libro, pues, útil por su diagnóstico y su prospectiva.

En términos más amplios, el libro de Rentería y Ordelin constituye también una aportación significativa a las discusiones generales sobre el uso de sistemas de IA en entornos gubernamentales. Primero, por la manera en que plantean la necesidad de que el desarrollo y aplicación de los sistemas de IA cumplan con ciertas normas de transparencia y trazabilidad. Tomando en cuenta tanto la complejidad del diseño y funcionamiento de los algoritmos, como los posibles sesgos discriminatorios que su uso puede traer consigo al brindar servicios públicos, resulta indispensable que gobiernos y funcionarios entiendan que la IA supone repensar la responsabilidad y la rendición de cuentas públicas. Segundo, el análisis de ambos autores ayuda a comprender el potencial de los sistemas de IA en los ámbitos público-gubernamentales, pero también las complejidades que caracterizan su implementación en este sector. Frente la visión eficientista que subyace a buena parte de los debates sobre la IA en los entornos

productivos del sector privado, estas páginas dejan muy en claro que en el ámbito público es indispensable privilegiar las preocupaciones sobre la privacidad, los derechos humanos, los recursos limitados y la empatía social.

El libro de Rentería y Ordelin adquiere un calibre mayor al dimensionarlo como pieza de un cuerpo de conocimiento científico en crecimiento, contextualizado en experiencias regionales. La obra abona con originalidad a la discusión académica de máxima trascendencia sobre el lugar actual y potencial de México, como parte de América Latina, en el desarrollo tecnológico mundial. La carrera desigual por la innovación no puede ser entendida fuera de las dinámicas geopolíticas y locales, por lo que una investigación como esta plantea resultados y lecciones que nutren importantes debates para la academia, los gobiernos y en general las sociedades latinoamericanas.

Cabe destacar, por último, el conjunto de atinadas recomendaciones de política pública que Rentería y Ordelin presentan como resultado de su investigación. Gracias a la amplísima difusión de las nuevas tecnologías digitales, la IA está hoy al alcance de cualquier persona, particularmente en su variante "generativa". Y, sin embargo, como bien muestran los autores, buena parte de las funcionarias y los funcionarios del país aún desconocen la lógica, los alcances, las posibles aportaciones y los probables riesgos de los sistemas de IA. El nuevo y prometedor mundo que vivimos requiere impulsar una nueva "alfabetización", para que la ciudadanía haga un uso productivo y responsable de estas nuevas tecnologías, y para que los gobiernos tomen decisiones informadas sobre cómo aprovechar los beneficios y cómo regular/controlar los riesgos de la IA en la hechura de políticas públicas.

Como académicos de la Flacso México, agradecemos a los profesores Rentería y Ordelin que hayan aceptado publicar su libro en nuestra casa editorial y confiamos en que su obra servirá para que estudiantes, investigadoras e investigadores, personas de la función pública e interesados en general sigamos aprendiendo sobre inteligencia artificial y sobre lo que los sistemas y tecnologías asociados a ella pueden (o no) aportar al desarrollo, bienestar y transformación de las sociedades mexicana, latinoamericanas y caribeñas.

Agradecimientos

La obra que aquí se presenta no habría sido posible sin el generoso apoyo de numerosas personas e instituciones. Agradecemos a Rodrigo Velázquez y Jacqueline Hernández, quienes facilitaron los recursos organizacionales a su alcance para hacer posible la investigación que constituye la base de esta obra. También reconocemos con gratitud la colaboración de Alejandro Villegas, Tlaltécatl Galván, Dante Rodríguez, Gabriela Gualito y, de manera muy especial, a Lucero Bejarano, cuya dedicación y compromiso fueron fundamentales para concluir la investigación. También, agradecemos a Gisela González y Mauricio Dussauge por su apoyo y acompañamiento en la materialización de este proyecto editorial. Finalmente, extendemos un agradecimiento a Jazmín Aquino, Sarah Muñoz Cadena, Lidya Ugalde y Diego Micher, quienes nos brindaron su apoyo durante el proceso de investigación.

Introducción

Cada día es más frecuente que la administración pública utilice sistemas de inteligencia artificial (IA) con cierto grado de autonomía, tanto en la gestión interna como en la prestación de servicios públicos. Estos sistemas buscan eficientizar procesos, optimizar la toma de decisiones, mejorar los servicios públicos y reducir la burocracia, entre otros usos. Los sistemas de IA utilizados en el ámbito público pueden tomar decisiones, emitir juicios o llegar a conclusiones sin la intervención humana, o bien apoyar la toma de decisiones, proporcionando información para justificar la intervención de un funcionario público.

Sin embargo, el uso de los sistemas de IA en la administración pública genera múltiples cuestionamientos, especialmente sobre si puede cumplir o al menos coadyuvar al cumplimiento de las funciones esenciales de una entidad pública, y al mismo tiempo respetar los valores públicos y los derechos humanos. Por ello, es necesario comprender cómo funcionan y se utilizan estas herramientas tecnológicas en el ámbito público, así como su influencia en la toma de decisiones, el ejercicio de la función pública y, sobre todo, su impacto en la vida de las personas.

Dado el contexto particular de los sistemas de IA aplicados en el ámbito público, se requieren mecanismos e instrumentos que permitan entender su uso y funcionamiento, con el fin de mitigar riesgos asociados a esa tecnología, como sesgos discriminatorios, posibles

errores en la toma de decisiones, despersonalización del servicio público y desigualdad en su acceso, entre otros.

Por otra parte, los repositorios de sistemas de IA, también denominados en algunos casos registros de algoritmos,[1] son herramientas utilizadas por gobiernos, organizaciones de la sociedad civil e instituciones académicas para identificar qué sistemas de IA se emplean en el ámbito público. Estos permiten mapear y reconocer prácticas que estas tecnologías generan en la gestión de asuntos públicos y, además, fomentan la transparencia algorítmica. Su creación y puesta en funcionamiento es la primera etapa en la búsqueda de herramientas que faciliten el seguimiento y la supervisión de los resultados de los sistemas de IA en el sector público. El seguimiento de los sistemas de IA en este sector constituye uno de los principales retos de la administración pública (OCDE, 2024a).

En este sentido, las propuestas de gobernanza de la IA en el sector público reconocen la necesidad de promover un repositorio público nacional de sistemas de IA, por ejemplo, la Carta Iberoamericana de Inteligencia Artificial (CLAD, 2023). Sin embargo, los repositorios son solo uno de los mecanismos de supervisión, seguimiento y transparencia algorítmica que eventualmente se establecerán en cada país. Por lo tanto, es este el primer paso para garantizar el uso efectivo de los sistemas de IA en el sector público y fortalecer la confianza de la ciudadanía en esta tecnología.

En México, se están implementando iniciativas orientadas a promover el uso responsable de la IA en el sector público. Una de ellas es la Agenda Nacional Mexicana de Inteligencia Artificial que recomienda, entre sus líneas de acción, recoger el número de iniciativas o proyectos de IA del gobierno con el propósito de definir un modelo y estrategia para su aplicación (IA2030Mx, 2020). Otro caso es la Política de IA y Gobernanza de Datos del gobierno del estado de Jalisco, la cual reconoce la importancia de analizar y comunicar las limitaciones de los sistemas de IA sin exagerar sus capacidades y

[1] En este trabajo se utiliza el concepto de repositorio de sistemas de IA.

buscando explicar su funcionamiento de manera comprensible, de acuerdo con los principios de transparencia y explicabilidad (Gobierno de Jalisco, 2024). El gobierno federal también ha anunciado la creación del Centro Nacional de Tecnologías Públicas que incluirá un repositorio nacional de productos y sistemas desarrollados con recursos públicos para promover la creación de capacidades tecnológicas de las administraciones públicas federal, estatal y municipal (ATDT, 2025). En particular, la Dirección General de Simplificación Administrativa, dentro de la Agencia de Transformación Digital y Telecomunicaciones, tiene la atribución de integrar un repositorio nacional de código, conforme a lo establecido en el artículo 15, fracción X, del Reglamento Interior de dicha agencia.

En la actualidad, no se han identificado instrumentos que cumplan con los fines y objetivos de un repositorio de sistemas de IA en el sector público mexicano, ni herramientas que desempeñen funciones similares. La información disponible es dispersa, carece de datos claros y estandarizados sobre los sistemas, su nivel tecnológico, funciones, uso territorial, entidades públicas, posibles riesgos, etc. Además, el debate sobre el tema es desarticulado, sin liderazgo identificable ni comunicación efectiva entre las diferentes entidades públicas, niveles de gobierno que utilizan IA, ciudadanos y otros actores.

A pesar de ello, existen algunas investigaciones y proyectos, aunque limitados y dispersos, que buscan identificar casos de usos de IA en el sector público. Como ejemplos destacados se encuentran la convocatoria "Premios sobre casos de uso de la Inteligencia Artificial en el sector público", promovida por el Instituto Federal de Telecomunicaciones (IFT), la Alianza Nacional de Inteligencia Artificial (ANIA), la Asociación Latinoamericana de Internet (ALAI) y u-GOB; así como el informe de investigación *La adopción de la inteligencia artificial en los gobiernos estatales de México* (Ruvalcaba y García, 2023). Asimismo, en el ámbito de la administración de justicia existen algunos estudios que documentan casos específicos de uso de sistemas de IA (México Evalúa, 2025).

Por otra parte, algunas entidades públicas han adoptado prácticas de transparencia proactiva en las que se declara de forma pública

el uso de sistemas de IA para el cumplimiento de sus funciones. Sin embargo, esto todavía no se puede considerar una práctica generalizada dentro de esas entidades. Entre los pocos casos se encuentran las publicaciones de la Agencia Digital de Innovación Pública de Ciudad de México para obtener la constancia digital de no antecedentes penales en Chatbot CDMX; el Poder Judicial del Edomex, con su Asistente Jurídico Virtual, y la Comisión Nacional del Sistema de Ahorro para el Retiro, que promueve el uso de un asistente conversacional.

Cuando los sistemas de IA se limitan a realizar las funciones de una entidad pública, suele haber poca información, ya que esta solo reconoce el uso de modelos como *machine learning*. Dentro de los ejemplos documentados tenemos la Secretaría Ejecutiva del Sistema Nacional Anticorrupción que ha declarado de forma proactiva el uso del "Algoritmo Anticorrupción" para detectar posibles distorsiones y riesgos de parcialidad en programas sociales. También destaca el Plan Maestro 2024 del Servicio de Administración Tributaria que reconoce el uso de *machine learning* para clasificar contribuyentes de riesgo, identificar redes complejas de evasión fiscal y detectar inconsistencias en comprobantes fiscales digitales por internet.

Desde este contexto, la presente investigación busca realizar un diagnóstico y emitir recomendaciones sobre el uso de sistemas de IA en el ámbito público en México, a partir de identificar cuáles son los que actualmente se utilizan en el sector público mexicano, y clasificarlos con base en su función, el servicio público al que se integran y los modelos de gestión y aprovechamiento tecnológico que permiten su implementación efectiva. Para elaborar esta investigación, se tomó como referencia el repositorio de sistemas de IA implementado por la Escuela de Gobierno de la Universidad Adolfo Ibañez (GobLab UAI) de Chile, y la base de datos sobre el sistema de decisión automatizada en el sector público colombiano desarrollado por investigadores de la Universidad del Rosario, en Colombia (Gutiérrez *et al.*, 2023).

Asimismo, se implementó, entre el 1 de marzo y el 1 de julio de 2024, una estrategia de búsqueda de información basada en cuatro

fuentes: solicitudes de transparencia, monitoreo de medios, referencias directas, y otras. Uno de los logros más relevantes fue la delimitación, aplicación y análisis de las solicitudes de transparencia. Se enviaron un total de 2241 (una por sujeto obligado) utilizando la Plataforma Nacional de Transparencia. Estas solicitudes permitieron consultar de forma directa y sistemática a las entidades públicas que pudieran estar desarrollando o utilizando sistemas de IA en el sector público mexicano, así como generar una base de datos con dichas respuestas que se presenta como parte de los resultados de esta investigación.[2]

Adicionalmente, este proyecto de investigación requiere de una definición precisa de *sector público* para delimitar las entidades sujetas a análisis. Para ello, como definición básica se asume que sector público: "es el conjunto de organizaciones, instituciones y empresas que dependen del Estado". Partiendo de esta definición, se han incluido en este estudio las entidades públicas de las tres ramas de gobierno (ejecutivo, legislativo y judicial, y los órganos autónomos) en los tres niveles de gobierno (federal, estatal y municipal) vinculadas con servicios públicos (o servicios a la ciudadanía) o con incidencia en las políticas públicas. Además de entidades no gubernamentales, como organizaciones de la sociedad civil, empresas, agrupaciones ciudadanas y universidades, que si bien están fuera de la definición de sector público, desarrollan sistemas innovadores de IA que atienden problemas públicos.

Definir la IA es complejo, pues no existe un consenso uniforme al respecto en la literatura. Para efectos de este estudio, se decidió utilizar un concepto básico y amplio que la concibe como cualquier herramienta tecnológica o sistema de IA capaz de realizar alguna tarea similar a las que ejecutan los seres humanos (aprender, razonar, tomar decisiones, etc.), siempre que dicha característica represente el componente principal de la tecnología y genere mayor valor en el ámbito público. Esta delimitación permitió establecer criterios de exclusión

[2] Las respuestas obtenidas de estas solicitudes se pueden encontrar en el sitio web https://doi.org/10.17605/OSF.IO/CU6W3

aplicables a aquellas tecnologías que, aunque son clasificadas como IA, no desempeñan funciones en el sector público, pues corresponden a otro tipo de *software*, o se utilizan dentro de la administración pública sin haber sido diseñadas específicamente para ello.

A partir de estas definiciones, la estrategia de análisis de datos se enfocó en generar información sobre los sistemas de IA utilizados en el sector público y, en particular, a delimitar sus características tecnológicas, funcionales, contractuales y de transferencia tecnológica. Con este propósito, se diseñó un esquema de codificación de la información recopilada, estructurado en tres secciones principales: 1) información general de los sistemas de IA y de su incorporación al repositorio; 2) características funcionales de los sistemas de IA, y 3) condiciones de adquisición, desarrollo y despliegue.[3]

El levantamiento de información sobre sistemas de IA en el sector público se complementó con una investigación cualitativa que permitió profundizar en el diagnóstico de catalizadores y barreras para su adopción. Esta etapa incluyó la realización de entrevistas semiestructuradas a desarrolladores, funcionarios públicos y tomadores de decisión que intervienen en la contratación, despliegue y uso de sistemas de IA en el sector público mexicano. Además, se levantó una encuesta entre las direcciones generales de Tecnología de la Información y la Comunicación de los organismos públicos federales. Por consiguiente, el resultado final de esta obra, además de ser de carácter descriptivo (identificación y caracterización), también plantea recomendaciones de políticas públicas.

Esta obra es un producto derivado del proyecto "Repositorio de sistemas de IA en el Ámbito Público", promovido por el programa de investigación Algoritmos CIDE.[4] Se divide en seis capítulos más esta introducción y las conclusiones. El primero presenta una revisión de la literatura enfocada a los conceptos de sistemas de IA y sus implicaciones para el sector público. El segundo, se ocupa del

[3] La identificación de cada uno de estos sistemas de IA, así como del levantamiento de información puede consultarse en https://doi.org/10.17605/OSF.IO/CU6W3

[4] Véase el sitio https://algoritmoscide.org/

concepto de transparencia algorítmica, y revisa las características de los repositorios de sistemas de IA como herramienta de transparencia algorítmica y experiencias previas. En el tercero, se describe la metodología utilizada para el levantamiento del repositorio de sistemas de IA, así como los métodos de investigación cualitativa y cuantitativa que complementaron el proceso. El cuarto capítulo ofrece un análisis estadístico descriptivo sobre los sistemas de IA identificados. En el quinto, se analizan seis casos de estudio de sistemas de IA que, según los autores, están generando valor público y tienen potencial de ser replicados en otras entidades públicas. En el sexto, se realiza un diagnóstico y se ofrecen recomendaciones de política pública, basadas en la información recopilada durante la identificación de sistemas de IA, así como en el análisis cualitativo de fuentes de información complementarias. Finalmente, se exponen las conclusiones del estudio.

A manera de resumen ejecutivo, se exponen a continuación los principales resultados del estudio.

A) Mapeo y características de los sistemas de IA en el ámbito público:

— Se identificaron 119 sistemas de IA utilizados en el ámbito público, de los cuales 97 corresponden al sector público y 22 a entidades no gubernamentales del ámbito público.
— La mayoría de los sistemas identificados son relativamente sencillos y presentan bajos niveles de autonomía y complejidad, aunque todavía no se dispone de suficiente información para poder analizar su efectividad e impacto en los contextos en los que se aplica, incluyendo la detección de sesgos, su minimización, rendimientos de calidad y seguridad.
— A nivel nacional, el uso de los sistemas de IA en el ámbito público es heterogéneo, con desigualdades entre las diversas entidades públicas que la implementan, así como entre las entidades no gubernamentales del ámbito público.
— Dentro de las entidades del sector público, el poder ejecutivo es la rama con más sistemas de IA, lo que representa el 54.6% de los

identificados; mientras que su uso en la rama judicial representa el 10.08% de los sistemas, y la legislativa apenas el 1.68%. Además del ejecutivo, son los órganos autónomos (15%)[5] los que poseen mayor porcentaje de sistemas identificados.

— El 45.73% de los sistemas identificados en el sector público pertenecen a los gobiernos estatales, y el 33.61% a entidades del gobierno federal, incluyendo los organismos autónomos federales. Solo se identificaron cuatro sistemas de IA en los gobiernos municipales, lo que representa el 3.36% del total de sistemas identificados.
— Los sistemas de IA más utilizados son los asistentes conversacionales.[6] En el sector público, estos son el 45.36%, en tanto que, en el ámbito público no gubernamental, suman el 54.54%. Otro tipo de sistema sobresaliente es el de predicción con 18 casos en total.
— Si bien los gobiernos estatales concentran la mayor parte de los sistemas desarrollados y declarados, el número por entidad federativa varía. Existe una brecha entre las que poseen un mayor número de sistemas de IA, como Ciudad de México cuyo total es de once, lo que significa casi una quinta parte del total estatal registrado. Algunas entidades solo cuentan con un sistema, y en ocho de ellas no se identificó ninguno.

B) Usos y aplicaciones de los sistemas de IA en el ámbito público:

— La tarea cognitiva de mayor uso tanto en el sector público como en el ámbito público no gubernamental es la de consulta de

---

5 La identificación de órganos autónomos utilizada fue previa a la reforma constitucional que eliminó algunos de ellos y que se publicó en el *Diario Oficial de la Federación* del 20 de diciembre de 2024.

6 Se entiende como asistente conversacional al sistema que se basa en grandes modelos de lenguaje (*large language models*) y que, por lo tanto, cuenta con un amplio repositorio de conocimiento (o datos para consulta) y es capaz aprender y adaptar sus respuestas en función de la interacción con el usuario (Noonia *et al.*, 2024). En la primera sección del capítulo V se detalla la distinción funcional entre "chatbots" y "asistentes conversacionales".

información. En el primer caso, el número de sistemas reportados que se relacionan con ella es de 32.98% de los sistemas identificados en este sector, mientras que, en el segundo, la cifra es el 50% de los sistemas identificados en el ámbito público no gubernamental.

— Los sectores con mayores sistemas de IA para las entidades del sector público, conforme a la clasificación de los Documentos de Marco Sectorial del Banco Internacional de Desarrollo (BID), son los de transparencia[7] e integridad, y seguridad ciudadana y justicia, con un total de 27 y 23 sistemas, respectivamente. Lo que representa el 27.8% y el 23.7%, en el mismo orden, del total de sistemas identificados en este ámbito. Sin embargo, en las entidades no gubernamentales del sector público, el mayor uso de sistemas identificado ocurre en los sectores de género y diversidad, e innovación, ciencia y tecnología, con un total de seis sistemas cada uno, que representan el 27.27%, respectivamente.

— En el sector público, el mayor número de sistemas, un total de 59, responden al objetivo 16, "Paz, justicia e instituciones sólidas", de los Objetivos de Desarrollo Sostenible (ODS). En el ámbito público no gubernamental, la contribución a los ODS es similar en cuanto a cantidades de sistemas: coinciden en una mayor incidencia en el objetivo 16, "Paz, justicia e instituciones sólidas", con 7 sistemas; aunque también sobresalen, educación de calidad (7), igualdad de género (5) y salud y bienestar (3).

7 Esta categorización de transparencia no debe ser entendida como equivalente del concepto de transparencia algorítmica que se utiliza en este texto. Para el BID este criterio está determinado por el incremento del acceso a la información y la facilitación del control y demanda ciudadana. El uso apropiado de tecnologías digitales, entre ellas los sistemas de IA, permite una mayor transparencia e integridad a nivel sectorial y en sentido general en toda la administración pública, ello incluye mejores decisiones por parte de los ciudadanos y actividades de control social (BID, 2020).

C) Condiciones de adquisición, adopción y despliegue:

— La mayoría de los sistemas de IA documentados no declaran o no cuentan con información disponible sobre cómo se adquiere la tecnología, lo cual corresponde a un total de 65 sistemas. En cambio, 18 de estos, una cantidad significativa, han sido desarrollados internamente por las propias entidades públicas, es decir, que los crea el equipo técnico que los utiliza. Son desarrollos que buscan atender preocupaciones propias de la ética del servicio público, como la transparencia y la rendición de cuentas, y aspectos relacionados con la política de transformación digital.
— Existe un alto nivel de desinformación sobre los usos de la IA en las entidades públicas, tanto entre los funcionarios públicos como entre la ciudadanía en general. Esto influye en la percepción del riesgo y lleva a sobreestimar sus aplicaciones, posibilidades y resultados.
— La formación de recursos humanos y el conocimiento sobre los sistemas de IA y sus usos son clave para decidir sobre la adopción de la tecnología en el sector público, así como su contratación o desarrollo interno, y su despliegue responsable.
— Las entidades públicas se perciben con un bajo nivel de capacidades informáticas, incluyendo infraestructura en sistemas de información, manejo de datos y desarrollo tecnológico interno.
— Las capacidades organizacionales determinan la adopción de la IA en el sector público. Destacan: el presupuesto asignado para el desarrollo o la contratación de tecnologías de la información; la experiencia previa en la contratación de sistemas de IA o tecnologías similares; la capacidad institucional para realizar adquisiciones flexibles; el nivel de conocimiento sobre el sector de IA y sus proveedores; y el grado de familiaridad con sistemas de IA que actualmente se utilizan en la administración pública.
— Se observa una desarticulación entre los actores del ecosistema de la IA en la innovación pública, esto es, entidades gubernamentales, el sector privado, la academia y la sociedad civil; así como un desbalance entre la discusión generada en los foros pú-

blicos sobre el tema y los desarrollos tecnológicos que realmente existen en México.
— Existe una narrativa de "búsqueda de soberanía tecnológica" vinculada a prácticas como promoción del desarrollo interno en materia de transformación digital y compartir tecnología.
— Persisten condiciones de incertidumbre en los procesos de contratación y adquisiciones de la tecnología. Los periodos para su implementación, pruebas y supervisión no siempre coinciden con los legalmente previstos en la regulación actual, así como tampoco con las limitaciones que habitualmente suele tener la administración pública federal (APF) en materia de cierre del año fiscal y de contratación.

Con base en los hallazgos descritos arriba, se proponen las siguientes recomendaciones de política pública:

1. Promover entre la ciudadanía la difusión de información sobre la IA y sus aplicaciones. Esto incluye la alfabetización sobre esta materia, la capacidad de identificar cuándo se interactúa con un sistema de IA y el acceso a información significativa sobre estos sistemas.
2. Fomentar la difusión de información y la capacitación de los funcionarios públicos en temas de IA. Esto aplica para quienes participan en procesos de acompañamiento técnico e implementación de esa tecnología en los diferentes niveles de la APF y en el desarrollo de productos tecnológicos públicos, y para los funcionarios expertos en las problemáticas o procesos donde la IA impactará o será utilizada.
3. Incentivar la generación de políticas, herramientas y medidas jurídicas de transparencia algorítmica durante todo el ciclo de vida de un sistema de IA, involucrando a los diversos agentes que participan en dicho proceso.
4. Fortalecer la capacidad de información y de ciencia de datos de los gobiernos, en la generación, procesamiento y uso de la información de calidad.

5. Revalorización de los datos y de la cultura de transparencia en la administración pública, basada en la importancia de considerar el dato como un activo de la administración pública.
6. Consolidar un ecosistema de IA en la innovación pública. En la administración pública, esta tecnología debe ser parte integral de la política nacional de transformación digital.
7. Documentar y difundir casos de éxito de sistemas de IA.
8. Compartir los sistemas creados dentro del sector público como un mecanismo para impulsar su desarrollo y adopción en otras entidades públicas.
9. Promover el debate público sobre nuevos modelos para la adquisición de desarrollos tecnológicos que permitan la contratación ágil y flexible de servicios digitales y, así, garantizar el uso responsable de la IA.

Es importante considerar las limitaciones del presente estudio, en especial aquellas que derivan de las fuentes utilizadas, tales como la estandarización de los datos, la suficiencia en la descripción y la clasificación de los sistemas, así como posibles errores en la aplicación de los criterios de inclusión y exclusión en el listado definitivo. Este primer levantamiento de datos para el repositorio se concibe como un ejercicio inicial cuyos resultados podrán actualizarse de manera periódica, incorporando nuevos sistemas, actualizando la información de los existentes y añadiendo otros criterios de clasificación y comparación, sobre todo en cuanto al nivel de desarrollo tecnológico, el cumplimiento ético y normativo, y la eficacia conforme a los objetivos declarados o propuestos para su uso.

Así pues, esta investigación busca contribuir a la identificación de mecanismos innovadores para fomentar la gobernanza y transparencia de los sistemas de IA en el sector público. La implementación y uso de un repositorio y diagnóstico de sistemas de IA permitirá disponer de información concreta sobre qué tecnología se está utilizando en el cumplimiento de funciones públicas: cómo se la adquiere, qué parámetros éticos se consideran para su contratación, despliegue y adopción, cómo identificar áreas de mejora, entre otros datos.

Se espera que los resultados de esta investigación se difundan entre los estudiantes de política pública, tomadores de decisiones y la sociedad civil, para facilitar la identificación de casos de éxitos en el uso de IA, de entidades líderes en la adopción de estas herramientas, y experiencias que puedan replicarse para resolver problemáticas similares. De este modo, se promoverá el uso de estas tecnologías en la transformación digital a nivel institucional, estatal y federal, optimizando capacidades institucionales, recursos humanos, infraestructuras y tecnologías, además de identificar actores importantes del ecosistema de innovación nacional.

También se busca que este trabajo sea la base para un debate público informado, que permita formular propuestas y tomar decisiones conforme a las necesidades y capacidades del país, con la finalidad de cerrar brechas y promover el uso responsable de la IA. Asimismo, para la ciudadanía y la sociedad civil representa un ejercicio de transparencia que fortalece el derecho a la información y la participación en el debate sobre el uso de esta tecnología en el sector público.

Así pues, es necesario continuar investigando y desarrollando iniciativas en las áreas del sector público donde la IA tiene una mayor capacidad transformadora y puede generar un verdadero valor público. Es importante también avanzar en el diseño e implementación de medidas efectivas que garanticen la supervisión, efectividad y mitigación de los riesgos; para ello es imprescindible la colaboración de todos los actores relevantes del ecosistema de la IA.

I. IA y ámbito público

No existe una definición única de inteligencia artificial. Pero ante la multiplicidad de dimensiones del término y sin ánimo de establecer un debate doctrinal sobre el tema, se considera importante hacer referencia a su sentido y alcance en este libro. Ello permite delimitar el objeto de estudio, y sus implicaciones en el ámbito público. En este sentido, se parte de la idea de que el análisis del uso de los sistemas de IA en el ámbito público debe abordarse desde una dimensión tecnológica, pero también desde la sociotécnica, en la cual se tomen en cuenta las implicaciones y retos económicos de infraestructura, sociales, jurídicos y hasta culturales que puede llegar a representar.

Sistemas de IA

Diferentes documentos, incluyendo regulaciones internacionales y nacionales además de estándares técnicos, han adoptado definiciones más o menos precisas del término *inteligencia artificial*, tratando de abarcar el amplio abanico de tecnologías que actualmente se desarrollan bajo esa denominación. Cualquier definición supone riesgo, pero citamos algunas al respecto.

Cuadro I.1.Definiciones de sistemas de inteligencia artificial

Fuente	*Concepto*	*Tipo*
OCDE. Principios de IA	Sistema basado en máquinas que, con objetivos explícitos o implícitos, deduce, a partir de la información que recibe, cómo generar resultados tales como predicciones, contenido, recomendaciones o decisiones que pueden influir en entornos físicos o virtuales. Los diferentes sistemas de IA varían en sus niveles de autonomía y adaptabilidad tras su implementación.	Principios
Convenio Marco del Consejo de Europa sobre Inteligencia Artificial y Derechos Humanos, Democracia y Estado de Derecho (CETS n.º 225) (art. 2)	"Sistema de inteligencia artificial": Sistema basado en una máquina que, con objetivos explícitos o implícitos, deduce, a partir de la información que recibe, cómo generar resultados tales como predicciones, contenidos, recomendaciones o decisiones que pueden influir en entornos físicos o virtuales. Los diferentes sistemas de inteligencia artificial varían en sus niveles de autonomía y adaptabilidad tras su implementación.	Tratado internacional
Reglamento (UE) 2024/1689 por el que se establecen normas armonizadas sobre inteligencia artificial. (art. 3.1)	"Sistema de IA": Un sistema basado en una máquina que está diseñado para funcionar con distintos niveles de autonomía y que puede mostrar capacidad de adaptación tras el despliegue, y que, para objetivos explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar resultados de salida, como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos físicos o virtuales.	Regulación comunitaria
2021 U.S. Code Title 15 - Commerce and Trade Chapter 119 - National Artificial Intelligence Initiative Sec. 9401 – Definitions 15 U.S.C. 9401(3).	"Inteligencia artificial": Se refiere a un sistema basado en máquinas que, para un conjunto determinado de objetivos definidos por humanos, puede realizar predicciones, recomendaciones o decisiones que influyen en entornos reales o virtuales. Los sistemas de inteligencia artificial utilizan aportaciones de máquinas y humanos para: (A) percibir entornos reales y virtuales; (B) abstraer dichas percepciones en modelos mediante análisis automatizados; y (C) utilizar la inferencia de modelos para formular opciones de información o acción.	Regulación nacional
ISO/IEC 22989	Sistema diseñado para generar resultados tales como contenido, previsiones, recomendaciones o decisiones para un conjunto determinado de objetivos definidos por el ser humano.	Estándar internacional
NIST	Rama de la informática dedicada al desarrollo de "sistemas de procesamiento de datos que realizan funciones normalmente asociadas con la inteligencia humana, como el razonamiento, el aprendizaje y la superación personal"	Estándar técnico

Fuente: Elaboración propia.

Asimismo, el concepto de IA se refiere a la rama de la informática dedicada "al desarrollo de sistemas de procesamiento de datos que realizan funciones normalmente asociadas a la inteligencia humana, como el razonamiento, el aprendizaje y la superación personal" (National Institute of Standards and Technology, 2019), lo que comprende la disciplina de investigación y desarrollo de sistemas de IA. En otros casos, se distingue a la IA como "capacidad de un dispositivo para realizar funciones que normalmente se asocian a la inteligencia humana, como el razonamiento, el aprendizaje y la superación personal" (ISO/IEC, 2022), del concepto de *sistemas de IA*, que identifica al sistema técnico que utiliza la IA para resolver problemas (National Institute of Standards and Technology, 2019).

Para efectos de este estudio, que no tiene como objetivo brindar una definición exacta de IA, utilizaremos el concepto de sistemas de IA, tomando como referencia los elementos que lo identifican: 1) un sistema basado en una máquina; 2) que está diseñado para funcionar con distintos niveles de autonomía; 3) que puede mostrar capacidad de adaptación tras el despliegue; 4) y que, para objetivos explícitos o implícitos; 5) infiere, de la información de entrada que recibe, la manera de generar resultados de salida, 6) como predicciones, contenidos, recomendaciones o decisiones, 7) que pueden influir en entornos físicos o virtuales[1] (Comisión Europea, 2025).

Más que una definición exacta de qué entender por sistemas de IA, que comprende múltiples tecnologías, es necesario considerar que estos sistemas se diseñan con distintos niveles de autonomía y adaptabilidad, que se materializan en su capacidad para aprender y actuar sin intervención humana, y que características como niveles de autonomía, inferencia, capacidades de adaptación e influencia en entornos físicos y virtuales, se encuentran estrechamente relacionadas.

---

[1] Estos elementos son los adoptados de la definición brindada en el Reglamento Europeo de IA y conforme a las Directrices de la Comisión Europea relativas a la definición de sistema de inteligencia artificial establecida en el Reglamento (UE) 2024/1689 (Reglamento de IA).

El concepto incluye un conjunto de técnicas de IA que van, desde el aprendizaje automático (supervisado, no supervisado, autosupervisado y aprendizaje por refuerzo), hasta las basadas en la lógica y el conocimiento (motores de inferencia y deducción, razonamiento simbólico, entre otras). Ello a su vez abre el abanico de posibilidades y áreas de aplicación de estos sistemas, desde la visión por computadora y el procesamiento del lenguaje natural, hasta el reconocimiento de voz y los sistemas inteligentes de apoyo a la toma de decisiones, entre otros. La variedad de posibilidades tecnológicas que el término implica supone retos técnicos, así como de diseño e implementación de estos sistemas, que pueden ser constantemente ajustados conforme a su nivel de autonomía, capacidad de adaptación, rendimiento y resultados.

La capacidad de interactuar e influir en entornos físicos o virtuales de estos sistemas es quizás la que más relación tiene con el objeto de este estudio, dado el impacto que el sistema de IA pudiera suponer para el entorno donde es utilizado. Este impacto, ya sea positivo o negativo, se vincula estrechamente con las condiciones del desarrollo, despliegue y uso del sistema. Lo que se traduce en que el estudiar o analizar cualquier sistema de IA, desde una perspectiva del ámbito público, no se pueda hacer exclusivamente desde un análisis exclusivamente tecnológico, como una herramienta más de innovación, sino como una que tiene un nivel determinado de autonomía, que escapa del control humano en mayor o menor medida, y que incide en el entorno.

Todo lo anterior nos obliga a dimensionar el concepto de IA desde una perspectiva de crítica cultural y social, vinculada a retos y debates sociales más allá de su naturaleza técnica (Pasquale, 2020). Por ello, y para comprender el alcance del concepto de sistemas de IA, en este trabajo su abordaje se da no solo desde una perspectiva exclusivamente tecnológica, sino también desde un enfoque multidisciplinar que parta de una concepción tecnosocial. El enfoque tecnosocial de los sistemas de IA nos obliga a preguntarnos, "¿en qué circunstancias deberían utilizarse, en qué condiciones y si deberían utilizarse en absoluto?" (Jobin y Katzenbach, 2023).

Lo anterior adquiere especial significado cuando se trata de usos y aplicaciones en el ámbito público donde la tecnología debe responder a la solución de problemas públicos, o al menos coadyuvar a su solución. En el ámbito público se reafirma el hecho de que las tecnologías de IA, aparte de ser un tema de índole técnico, se entrelazan a sistemas sociales y a sus problemáticas. Cuestiones estructurales que inciden en su desarrollo e implementación, y en el contexto en el que tienen lugar. No es posible hablar de IA en el ámbito público sin considerar cuestiones de política, economía, cultura y discursos, entre otras.

Sistemas de IA y ámbito público

Se espera que el uso de los sistemas de IA revolucione varios aspectos de la gestión en el ámbito público, como una administración más eficiente y efectiva, una mejor vinculación con la ciudadanía, una mejora en la calidad de los servicios públicos y una mejor focalización de estos, entre muchos más (Wirtz *et al.*, 2021; Criado *et al.*, 2024; Grimmelikhuijsen y Tangi, 2024). A pesar de los retos de su uso en el ámbito público, que incluyen preocupaciones éticas, falta de capacidades institucionales e infraestructura de información o falta de confianza en el uso de este instrumento, la integración de sistemas de IA en las actividades sustantivas de las entidades públicas puede fortalecer significativamente el ejercicio del gobierno, desde la relación de un gobierno con la ciudadanía, pasando por la provisión de servicios públicos, hasta el fortalecimiento de la transparencia y la rendición de cuentas (Medaglia *et al.*, 2024; CEPAL, 2024; G7 Italia, 2024).

En este sentido, el debate sobre el uso de los sistemas de IA en el sector público ha sido abundante y diverso; se ha revisado y discutido el potencial de la IA para mejorar la función pública, casos de éxito, preocupaciones y riesgos percibidos, y la posibilidad de su uso ético y responsable. En esta revisión de la literatura, haremos un breve repaso sobre la información, recomendaciones y posicionamien-

to que han brindado organismos internacionales y otras entidades relevantes respecto al uso de sistemas de IA en el sector público en tres temas sustanciales: i) potenciales beneficios y riesgos de su uso en el sector público, ii) los principales posicionamientos sobre la importancia y la forma de operar la transparencia algorítmica y iii) la función que los repositorios de sistemas de IA cumplen en un contexto de transparencia y mapeo de buenas prácticas.

Tanto los posicionamientos institucionales como la literatura académica sobre el uso de IA en el sector público han presentado argumentos y evidencia relevante sobre sus potenciales riesgos y beneficios para el sector público, así como de las competencias y prácticas de gobernanza necesarias para llevar instrumentaciones de este tipo a buen puerto. Uno de sus potenciales beneficios es su capacidad de mejorar significativamente la eficiencia de las actividades administrativas en el sector público, mediante, por ejemplo, la automatización de tareas rutinarias, la integración de procesos y la toma de decisiones basada en evidencia. De acuerdo con la Comisión Europea (2024), los sistemas de IA han tenido mayor impacto en los dominios de servicios públicos, impuestos o finanzas públicas y salud. Mientras que la Organización para la Cooperación y el Desarrollo Económico (OCDE) ha reseñado el poder transformativo de la IA en la integración de las operaciones de gobierno y la personalización de los servicios (OCDE, 2022a, 2024a). Los sistemas de detección de fraude son particularmente relevantes por su capacidad de identificar anomalías en vastas fuentes de información. Algunos gobiernos están utilizando modelos de aprendizaje de máquinas para monitorear transacciones financieras, pagos de impuestos y desembolso de programas de bienestar para detectar irregularidades y anomalías, lo que permite intervenciones focalizadas y eficientes sobre sospechas de fraude. Por ejemplo, en Brasil se utilizó un sistema de IA para la detección de fraude en programas de bienestar social, lo que ha permitido mayor eficiencia en el gasto de fondos públicos (Nicolás y Sampaio, 2024).

Por otro lado, la IA permite a los gobiernos ofrecer atención personalizada a la ciudadanía, para satisfacer mejor sus distintas necesidades. Al analizar datos sobre la demografía, preferencias y com-

portamientos de los ciudadanos, los sistemas de IA pueden optimizar la prestación de servicios públicos. Por ejemplo, los asistentes conversacionales impulsados por sistemas de IA dan asistencia en tiempo real a los ciudadanos, mientras que los modelos predictivos sugieren medidas proactivas, como alertar a las personas sobre plazos próximos o beneficios a los que podrían calificar. En los servicios de salud, las herramientas personalizadas de IA analizan datos de los pacientes para recomendar planes de atención adaptados, asegurando mejores resultados de salud y una asignación más eficiente de recursos (Fisher y Rosella, 2022).

Finalmente, otro beneficio que se debe enfatizar es su potencial contribución para promover la transparencia y la rendición de cuentas. Cuando se aplica en apego a principios éticos y de responsabilidad, el uso de los sistemas de IA aporta nuevos mecanismos de gestión de la información, los cuales pueden mejorar los procesos de toma de decisiones, reducir errores y generar mayor confianza entre la ciudadanía (OCDE, 2022a; Wirtz, 2020). Sin embargo, es crucial abordar desafíos como el problema de la "caja negra", donde el proceso de toma de decisiones de los sistemas de IA no es fácilmente comprensible, para aprovechar plenamente estos beneficios (Wirtz, 2020; Mergel *et al.*, 2024).

Por otro lado, el uso de los sistemas de IA en el sector público también representa retos que deben enfrentarse proactivamente. Las complejidades de un sistema de IA, combinada con el contexto en el que operan, crean consecuencias no esperadas y, en algunos casos, violan derechos fundamentales de la ciudadanía, como la no discriminación, la justificación y explicación de cualquier acción administrativa o la protección de datos personales. Uno de los riesgos más conocidos y discutidos en el uso de los sistemas de IA en el sector público es el potencial sesgo sistemático que los sistemas de IA pueden estar reproduciendo (Grimmelikhuijsen y Tangi, 2024; Comisión Europea, 2024). Este problema surge cuando los modelos de IA se entrenan con datos sesgados, lo que conduce a resultados injustos o discriminatorios en áreas críticas como la elegibilidad para beneficios sociales y la asignación de recursos públicos. Se han do-

cumentado muchos casos de reproducción de sesgo sistemático en sistemas de IA, por ejemplo, los utilizados en la vigilancia predictiva han sido criticados por dirigirse desproporcionadamente a comunidades minoritarias. Estos sistemas corren el riesgo de perpetuar inequidades sistémicas a menos que sean cuidadosamente monitoreados y auditados para garantizar la equidad. Abordar este problema requiere mecanismos sólidos para evaluar y mitigar el sesgo, como evaluaciones de impacto ético y revisiones de posibles sesgos discriminatorios.

Otra preocupación recurrente en el uso de los sistemas de IA es la falta de transparencia o rendición de cuentas del instrumento; varios de ellos son considerados "cajas negras", en cuanto a que las personas no pueden hacer sentido de la forma en que generan sus resultados de predicción o de recomendación. Es decir, algunos de tales sistemas carecen de un mecanismo de interpretación de los datos inteligible para las personas, lo que dificulta la capacidad de explicar o justificar decisiones basadas en recomendaciones de los sistemas de IA. Para mitigar estos retos, el informe de gobernanza del Joint Research Centre (Tangi *et al.*, 2023) enfatiza la importancia de las prácticas de gobernanza procedimental para garantizar la transparencia en esos sistemas. No obstante, sin una explicación clara, los gobiernos pueden enfrentar desafíos al justificar las decisiones tomadas por la IA, lo que podría erosionar la confianza pública. Un ejemplo destacado proviene de los Países Bajos, donde se abandonó un sistema de detección de fraude en beneficios sociales impulsado por IA debido a su falta de transparencia y rendición de cuentas.

De manera similar, los sistemas de IA representan un reto mayúsculo con respecto a la privacidad de datos de los ciudadanos. Evidentemente, el desempeño de los sistemas de IA (no generativa) depende en gran medida de grandes conjuntos de datos y una riqueza de información importante sobre las personas, lo que genera riesgos importantes sobre la privacidad de sus datos. Las administraciones públicas suelen ser responsables de información sensible, y las brechas de seguridad pueden tener graves consecuencias. La Comisión Europea (2024) destaca cómo una gobernanza de datos inadecuada puede comprometer

la privacidad y llevar a violaciones de los derechos de los ciudadanos. Por ejemplo, las tecnologías de reconocimiento facial utilizadas en la vigilancia pública han originado debates sobre el equilibrio entre seguridad y privacidad. Estos sistemas, aunque efectivos para ciertas aplicaciones, conllevan el riesgo de abuso y uso indebido.

Tanto para capitalizar los beneficios como para mitigar los riesgos, los gobiernos requieren fortalecer sus capacidades de información y gestionar prácticas de gobernanza de IA que permitan navegar estos retos de manera efectiva. Para ello son necesarias capacidades técnicas básicas en las entidades públicas. El informe del Joint Research Centre (Tangi *et al.*, 2023) enfatiza que las competencias técnicas son fundamentales para diseñar, implementar y mantener soluciones de IA. Por ejemplo, las iniciativas de IA en Estonia demuestran de qué forma la experiencia técnica puede optimizar los servicios públicos, como las herramientas de salud predictiva y los sistemas de identidad digital. Las competencias gerenciales, que incluyen la gestión de proyectos y la supervisión estratégica, son esenciales para integrar la IA en los sistemas existentes. Pero también son fundamentales las capacidades gerenciales para la implementación de sistemas de IA, como liderazgo, gestión del cambio y asignación de recursos.

II. Los repositorios como herramientas de transparencia algorítmica

A la par de que se estudia y trabaja en la implementación efectiva de los sistemas de IA en el ámbito público, se debate sobre los mecanismos y herramientas para garantizar que sus riesgos y sesgos puedan supervisarse y minimizarse. Cuáles y cómo deberán operar sigue siendo un reto para el sector público, debido, entre otros aspectos, a la complejidad de la tecnología, a sus mecanismos de adquisición, y a su despliegue y uso, pero también por el propio papel que puede desempeñar en el sector público. En este sentido, uno de los temas pendientes es la implementación efectiva de mecanismos de transparencia algorítmica.

Transparencia algorítmica

El concepto de *transparencia algorítmica* ha sido entendido como principio, norma o estándar de la gobernanza de los sistemas de IA. Pero, desde una perspectiva amplia, implica comprender cómo se construyen, gestionan y mantienen los sistemas de IA (IAPP, 2024), mientras que de forma particular está relacionado con la obligación de facilitar información en la medida que sea necesaria, razonable y técnicamente posible, durante el desarrollo, la prestación y el uso de un sistema o servicio de IA. Todo ello en función del contexto social en el que este servicio se utilice, garantizando, al propio tiempo, su verificabilidad (Japan AI Safety Institute, 2024).

El reporte *Recomendación sobre la ética de la IA* de la Unesco reconoce que el objetivo del principio de transparencia es proporcionar información adecuada a los destinatarios para comprenderla y fomentar la confianza. Esto es, que dicho principio permite comprender cómo se implementa cada etapa de un sistema de IA, proporcionando información sobre los factores que influyen en una predicción o decisión específicas, y si existen o no garantías adecuadas. A la vez, la transparencia se relaciona estrechamente con la explicabilidad dado que hace inteligibles los resultados de estos sistemas y facilita información sobre ellos, y su comprensibilidad y trazabilidad (Unesco, 2022).

Sin embargo, no siempre es determinar cómo opera el sistema de IA ni comprender cómo ha llegado a tomar determinadas decisiones o resultados; una limitación que obedece a la propia tecnología y no a decisiones humanas. En estos casos se habla de la IA de caja negra, la cual se relaciona con la opacidad tecnológica, un aspecto que suelen presentar estos sistemas de IA (Burrell, 2016). No se trata de un tema de conocimientos o comprensión por parte de las personas y la IA, sino de que los propios expertos y/o desarrolladores están imposibilitados de comprender el funcionamiento del sistema.

No obstante, dado que la gobernanza de la IA está diseñada a partir de principios, leyes, políticas, procesos, normas, estándares, marcos, mejores prácticas de la industria y otras herramientas que están incorporadas al diseño, desarrollo, despliegue y uso de sistemas de IA (IAPP, 2024), es lógico advertir que la transparencia algorítmica adquiere diversas facetas, según se le reconozca como principio, obligación técnica-jurídica o un estándar. A continuación, se analizan varios de estos posicionamientos conforme a varios de los diseños adoptados.

Posicionamientos institucionales sobre transparencia algorítmica

Desde el punto de vista del *soft law*, son innumerables las propuestas que han reconocido la transparencia algorítmica como un principio,

aunque también se ha admitido que el acceso a la información y los mandatos de transparencia es uno de los enfoques reguladores de la IA (Gutiérrez, 2024) e inclusive que es un estándar.

La *Recomendación sobre la ética en el uso de la IA* de la Unesco reconoce que la transparencia algorítmica es un principio funcional o instrumental de la gobernanza de la IA, que es parte de las condiciones previas fundamentales para garantizar el respeto, la protección y la promoción de los derechos humanos, las libertades fundamentales y los principios éticos, a fin de determinar responsabilidades eficazmente, conocer los motivos por los que se ha tomado una decisión y, en consecuencia, poder interponer un recurso efectivo que posibilita el escrutinio público y, en sentido general, crear sociedades más pacíficas, justas, democráticas e inclusivas (Unesco, 2022).

En similar dirección, la *Recomendación sobre la IA* de la Organización para la Cooperación y el Desarrollo Económicos (OCDE) establece como uno de los principios de la administración responsable de una IA fiable, el de transparencia y explicabilidad. En este sentido, llama a los diferentes actores a comprometerse con la transparencia y divulgación responsable en relación con dichos sistemas y proporcionar información significativa de acuerdo con cada situación en particular y el estado de la tecnología (OCDE, 2024a).

En América Latina, la *Declaración de Santiago* también promueve la transparencia en las operaciones de IA, en particular respecto a la obligación de elaborar y adoptar medidas de política pública y normas jurídicas que contengan este principio (Cumbre Ministerial y de Altas Autoridades de América Latina y el Caribe, 2023).

Existen otros documentos que, aunque carecen de carácter vinculante, son adoptados por instituciones, organizaciones y otros actores que reconocen la transparencia algorítmica como un principio de gobernanza de la IA, y los han adoptado en la forma de declaraciones o directrices. Entre estos se encuentran las *Directrices éticas para una IA* fiable presentadas por el grupo de expertos de alto nivel sobre la IA de la Unión Europea en 2019, en las cuales se admite que la transparencia es uno de los requisitos para una IA fiable, reconociendo así las obligaciones de rastreabilidad, explicabilidad y comunicación en

relación con los elementos pertinentes de un sistema de IA; léase datos, el sistema y los modelos de negocio.

La Association for Computing Machinery (ACM, por sus siglas en inglés) adoptó la *Declaración sobre transparencia algorítmica y rendición de cuentas*, donde se establece un conjunto de principios relacionados con la transparencia y responsabilidad algorítmica, entre los que destaca la concienciación por parte de los propietarios, diseñadores, usuarios y partes interesadas en los sesgos implicados en su diseño, implementación y uso, así como el posible daño de estos sistemas; además de los principios de acceso y reparación, rendición de cuentas, explicación, procedencia de los datos, auditabilidad, y validación y pruebas (ACM, 2017).

La *Carta de Derechos Digitales del Gobierno Español*, por su parte, reconoce dentro de los derechos en los entornos de la IA (derechos digitales en entornos específicos), la transparencia, auditabilidad, explicabilidad, trazabilidad, supervisión humana, y gobernanza, en el desarrollo y ciclo de vida de los sistemas de IA y que la información facilitada sea accesible y comprensible (Ministerio de Asuntos Económicos y Transformación Digital, 2021).

En la *Declaración de Bletchley* de los países que asistieron a la cumbre de seguridad en torno a la IA en 2023 (AI Safety Summit), se demanda mayor transparencia de los actores privados que desarrollan capacidades de IA, en concomitancia con métricas de evaluación, pruebas de seguridad y desarrollo de capacidades e investigación científica relevante en el sector público. En tanto que la *Declaración Conjunta sobre los Derechos y Principios Digitales para la Década Digital* reconoce, dentro de la libertad de elección, las interacciones con sistemas de IA y concretamente la obligación de "velar por un nivel adecuado de transparencia en el uso de los algoritmos y la IA para que las personas estén informadas y capacitadas cuando interactúen con ellos" (Parlamento Europeo, Consejo y Comisión, 2023).

Un ulterior documento, la *Carta Iberoamericana de Inteligencia Artificial en la Administración Pública* del Centro Latinoamericano de Administración para el Desarrollo (CLAD, 2023), también reconoce el Principio de Transparencia, Trazabilidad y Explicabilidad, el cual se

fundamenta en la accesibilidad y disponibilidad de la información en todo momento del ciclo de vida de los sistemas, datos significativos, contextualizados y consistentes, así como en la interacción consciente e informada con los sistemas de IA, y conocimiento de los resultados de su uso y potenciales efectos.

Si bien se ha señalado que pocos países han adoptado obligaciones de transparencia algorítmica para los organismos públicos a través de la regulación (Unesco, 2024a), lo cierto es que comienza a aparecer el reconocimiento de este principio en instrumentos jurídicos vinculantes, como la Convención Marco del Consejo de Europa sobre Inteligencia Artificial y Derechos Humanos, Democracia y Estado de Derecho, que es el primer instrumento internacional jurídicamente vinculante. Este instrumento reconoce que la "Transparencia y Supervisión" es uno de los principios fundamentales que deben cumplir los sistemas de IA, incluyendo su utilización por parte de las autoridades y agentes privados que actúan en nombre de estas últimas (Consejo de Europa, 2024).

En el Reglamento de IA de la Unión Europea se han establecido normas armonizadas de transparencia aplicables a determinados sistemas de IA, entre las que se encuentran varias obligaciones: a) de transparencia de los proveedores y responsables del despliegue de determinados sistemas de IA; b) garantizar un nivel de transparencia suficiente para que los responsables del despliegue interpreten y usen correctamente los resultados de salidas, y c) la referente a la documentación técnica que los proveedores de modelos de IA de propósito general deberán proporcionar para los proveedores posteriores que integren el modelo en su sistema de IA (Parlamento Europeo y Consejo, 2024).

A pesar de sus diferentes perspectivas y alcances, en este estudio se considera que es posible alcanzar la transparencia algorítmica a partir de las cuatro obligaciones básicas reconocidas por la OCDE: 1) fomentar una comprensión general de los sistemas de IA, incluidas sus capacidades y limitaciones; 2) poner en conocimiento de las partes interesadas sus interacciones con los sistemas de IA, incluido el lugar de trabajo; 3) suministrar cuando resulte viable y útil la informa-

ción transparente y comprensible sobre las fuentes de datos/entradas, los factores, los procesos o el razonamiento que subyace a las predicciones, contenidos, recomendaciones y decisiones, para permitir a los afectados por un sistema de IA comprender los resultados, y 4) proporcionar información a estos últimos para que puedan cuestionarlo (OCDE, 2024a).

Al mismo tiempo se reconoce que, para que la transparencia algorítmica deje de ser una deuda en el ámbito público y en el privado, es necesario identificar mecanismos que la garanticen durante todo su ciclo de vida en los sistemas de IA. Además del reconocimiento jurídico, es imprescindible delimitar y concebir los mecanismos que la hagan viable y efectiva en cada país.

Transparencia algorítmica en México

En México, se han realizado distintos tipos de propuestas que abordan directa o indirectamente el tema de la transparencia algorítmica y que provienen de sectores, como el gobierno, las iniciativas legislativas y organizaciones de la sociedad civil. Pero es en el sector civil donde se han elaborado más número de propuestas reconociéndose en la mayoría de los casos que la transparencia algorítmica es un principio de la gobernanza de la IA, mientras que, en otros, se proponen medidas concretas para su implementación.

Entre las propuestas que declaran la transparencia algorítmica como principio de la gobernanza de la IA está la de la Agenda Nacional de la Inteligencia Artificial para México 2024-2030, elaborada por la Alianza Nacional de Inteligencia Artificial, allí se presenta un conjunto de recomendaciones de política pública, regulación y gobernanza de la IA. Esta iniciativa promueve el reconocimiento de la transparencia y rendición de cuentas en todas las etapas del desarrollo de la tecnología, y recomendaciones de regulación y gobernanza basada en la transparencia y demás principios éticos (Lagunes *et al.*, 2024).

Por su parte, el cuarto de los principios para el desarrollo confiable, responsable y seguro de la IA en México del Centro Méxi-

co Digital propone la Transparencia, Explicabilidad y Auditabilidad (Centro México Digital, 2024). Y el grupo de personas expertas de Alianza Global para la Inteligencia Artificial que trabaja en temas sobre la investigación, desarrollo, incidencia y la política pública en IA en el país, al elaborar sus recomendaciones en materia de innovación y comercialización para una política soberana de IA en México, propone adoptar políticas que promuevan la transparencia y trazabilidad en el uso de sistemas de IA, así como en la gobernanza y soberanía de datos, en los procesos de toma de decisiones, y en la definición de estrategias para garantizar la transparencia desde un punto de vista legislativo (Ricaurte *et al.*, 2024).

Algunas propuestas contienen medidas concretas para operacionalizar la transparencia algorítmica en distintos escenarios y ámbitos. En este sentido, la Agenda Nacional Mexicana de IA elaborada por la Coalición IA2030Mx recomienda el establecimiento de estándares de transparencia para las empresas que dan servicios de IA al gobierno local o federal. Mientras que el Proyecto Open Loop México que propone un Prototipo de Políticas Públicas sobre Transparencia y Explicabilidad de Sistemas de IA, recomienda medidas de transparencia respecto a la divulgación e información, incluyendo la obligación de los implementadores de tomar las medidas técnicas y organizativas necesarias para garantizar que el uso del sistema sea transparente y los resultados explicables, además de deberes de notificación e información, y divulgación de elementos relevantes, que los distintos destinatarios podrían tener en cuenta (Del Pozo *et al.*, 2023). En la misma ruta, la Alianza Nacional de Inteligencia Artificial, entre sus acciones concretas, promueve la adopción de estándares de transparencia algorítmica en espacios controlados de pruebas, y el desarrollo de directrices y estándares éticos y legales que regulen el uso y aplicación de la tecnología en el sistema judicial, como parte del modelo de gobernanza para el acceso a la información pública, privacidad y protección de datos personales y datos abiertos (Lagunes *et al.*, 2024).

De manera similar, en su evaluación del estadio de preparación de la IA, la Unesco México realiza recomendaciones en torno a la transparencia algorítmica que son específicas para el diseño institu-

cional y de gobernanza de los sistemas de IA. Algunas se dirigen a la Procuraduría Federal del Consumidor (Profeco) en lo relativo a los derechos del consumidor y la IA al momento de aceptar términos y condiciones, y al Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos Personales (INAI) en materia de privacidad. Asimismo se hacen recomendaciones en ciberseguridad y el marco jurídico de la IA, aunque en este último caso solo es en cuanto a la transparencia de datos (Unesco, 2024a).

La transparencia algorítmica ha comenzado a ser parte de las agendas regulatorias para el diseño, desarrollo y uso responsable de la IA, incluso desde iniciativas de instituciones públicas, las cuales se encuentran en diversos niveles de desarrollo. Las conclusiones de las mesas temáticas sobre IA en las Telecomunicaciones y Radiodifusión llevadas a cabo por la Subsecretaría de Comunicaciones en 2019 recomendaban un marco regulatorio adecuado que garantizara la transparencia del desarrollo e implementación de la IA, y proponía, entre los mecanismos, asegurar la transparencia a partir de sistemas de código abierto (Secretaría de Comunicaciones y Transporte, 2019).

La Ley General de Mecanismos Alternativos de Solución de Controversias reconoce el principio de transparencia algorítmica aplicable a los procedimientos de solución de controversias en línea (Congreso de la Unión, 2024a). Mientras que el Acuerdo del Gobernador Constitucional del Estado de Jalisco, mediante el cual se emiten las Políticas y Disposiciones Generales en materia de IA y gobernanza de datos de la administración pública del Gobierno del Estado de Jalisco, reconoce el principio de transparencia y explicabilidad, además de la necesidad de analizar y comunicar las limitaciones de los sistemas de IA sin exagerar sus capacidades y explicar su funcionamiento de manera comprensible para los usuarios y partes interesadas (Gobierno de Jalisco, 2024).

Actualmente, se están diseñando mecanismos que permiten operacionalizar la obligación de transparencia algorítmica en temas concretos, por ejemplo, en el decreto por el que se adicionan disposiciones respecto a plataformas digitales en la Ley Federal del Trabajo, donde se especifican obligaciones claras acerca de la transparencia algorítmica,

incluida la de elaborar un documento de política de gestión de sistemas de IA y de reglas para la asignación de tareas, servicios, obras o trabajos (Congreso de la Unión, 2024b).

Por último, la propuesta de ley para el uso de la IA y el tratamiento de datos personales por sujetos obligados en Ciudad de México concibe la transparencia como un principio que permite trazabilidad y explicabilidad adecuada. Esta propuesta aborda asimismo mecanismos y herramientas para llevar a cabo la transparencia algorítmica, tales como la transparencia de datos de entrenamiento, carpetas de evidencia, evaluaciones de transparencia algorítmica, publicación y accesibilidad a resultados de las auditorías de sistemas de IA (Info CDMX, 2024).

Aunque no son disposiciones específicas para la transparencia algorítmica, existen dentro del ordenamiento jurídico mexicano regulaciones jurídicas que podrían ser aplicadas, como las establecidas en los artículos 6 y 16 de la Constitución Política de los Estados Unidos Mexicanos, o en la Ley Federal de Protección de Datos Personales en Posesión de los Particulares (por ejemplo, artículo 26.2 sobre el tratamiento automatizado de datos) y el Reglamento de esta misma ley (art. 112, entre otros), en la Ley General de Protección de Datos Personales en Posesión de Sujetos Obligados (por ejemplo, el art. 41.2 sobre el tratamiento automatizado de datos), y en la Ley General de Transparencia y Acceso a la Información Pública. Sin embargo, no puede pensarse que estos sean en sí mismos mecanismos de transparencia algorítmica, y aunque pueden coadyuvarlos, no fueron creados con dicho fin. La transparencia en tanto principio rector de la administración pública no se limita a la transparencia algorítmica, esta última no debe ser únicamente incorporada al sector público ni quedar limitada al acceso a la información pública. Subsumirla dentro de la transparencia como principio de la administración pública nos podría llevar a realizar un análisis limitado de ambas figuras y los alcances, fines y mecanismos que deben ser implementados para cumplir cada una de las obligaciones que conllevan.

Aunque el sistema de transparencia pública en México existe, tiene protección constitucional y es operativo, el sistema de transparen-

cia algorítmica aún requiere de un diseño orgánico que contemple diversos mecanismos para acceder a información relevante sobre los sistemas de IA a lo largo de todo su ciclo de vida. Este también debe incluir obligaciones diferenciadas para cada uno de los actores que intervienen en este ciclo. En el caso particular de los sistemas de IA que se utilizan en el ámbito público, deberán pensarse a la par de las obligaciones vigentes sobre transparencia pública, y de las herramientas de supervisión de sistemas de IA.

Transparencia algorítmica y repositorios públicos

La transparencia algorítmica es parte de la gobernanza técnica de la IA, por ende, se materializa a partir de herramientas y mecanismos que se diseñan para su aplicación y cumplimiento (Reuel y Bucknall, 2024). Uno de los aspectos más relevantes es la suficiencia de la información, así como los mecanismos claves que se utilizan para divulgar y poner en conocimiento de los distintos agentes dicha información, sin afectar otros derechos. El grado de transparencia debe ser adecuado al contexto y al efecto, considerando el ámbito de su aplicación, la utilización prevista, los destinatarios y la viabilidad de cada sistema de IA en particular (Unesco, 2022). La divulgación de la información debe realizarse conforme a las necesidades de los tipos de usuarios y agentes que intervienen en el ciclo de vida de la IA, tomando en cuenta los conocimientos del destinatario, y su papel en la interacción con el sistema (IAPP, 2024).

En el ámbito de la transparencia en el uso y desarrollo de sistemas de IA, se han establecido diversas medidas técnico-jurídicas para garantizar un uso ético, responsable y comprensible de estas tecnologías. Entre tales disposiciones se encuentra la obligación de informar de manera clara y precisa cuando un usuario interactúa con un sistema de IA, asegurando que tenga pleno conocimiento de sus capacidades y limitaciones. También es fundamental informar si un producto o servicio se proporciona directa o indirectamente con la asistencia de IA.

Asimismo, surge la necesidad de documentar exhaustivamente toda la información relevante a lo largo del ciclo de vida de un sistema de IA. Esta documentación debe abarcar aspectos técnicos y no técnicos, incluyendo decisiones de diseño, datos empleados en el entrenamiento, estructura del modelo, criterios de despliegue y uso, y especificaciones sobre arquitectura, rendimiento, limitaciones, pruebas realizadas y usos recomendados. Derivado de lo anterior, se impone la obligación de conservar registros detallados que puedan ser utilizados por proveedores en etapas posteriores, facilitando su introducción en el mercado y permitiendo el acceso a las autoridades nacionales competentes de ser requerido.

Otro aspecto crucial en materia de transparencia es la identificación de los contenidos generados o modificados mediante IA. En este sentido, se considera esencial la elaboración de resúmenes de los conjuntos de datos utilizados en el entrenamiento de modelos de IA de propósito general,[1] para proporcionar una visión clara del origen y características de la información con la que han sido desarrollados. Además, se establecen obligaciones específicas en cuanto a la detección y etiquetado de contenidos generados o manipulados de manera artificial.

De igual modo, se impone la obligación de proporcionar acceso al código fuente o a los conjuntos de datos empleados en su desarrollo.Y se exige que los proveedores, contratistas y beneficiarios de subvenciones públicas adopten prácticas sólidas de gobernanza y gestión de sistemas de IA, particularmente en lo que respecta a los modelos sujetos a contratación o financiamiento público. Este requisito busca instalar mecanismos de control adecuados que aseguren el cumplimiento de estándares éticos y normativos en el uso de los sistemas de IA en el sector público y privado (NTIA, 2024).

[1] Se entiende por IA de propósito general al sistema de IA que puede realizar, o ser adaptado para realizar, un rango diverso de tareas, incluidas algunas para las que no fue diseñado (Gutiérrez *et al.*, 2023).Véase la sección "Casos excluidos" en el capítulo III de este libro.

En este orden de ideas, para garantizar un acceso equitativo y estandarizado a la información, se plantea la necesidad de desarrollar protocolos claros y ejecutables en todas las fases del ciclo de vida de los sistemas de IA. Esto implica la adopción de formatos accesibles y redactados en un lenguaje sencillo, tales como tarjetas descriptivas de modelos o sistemas y hojas de datos de IA. Finalmente, se promueve la creación de repositorios públicos de sistemas de IA empleados en el sector gubernamental, con el propósito de incrementar la transparencia y la confianza en la implementación de estas tecnologías. Es por ello que se fomenta la adopción de prácticas de cocreación mediante la participación de la ciudadanía en los procesos de diseño e implementación de sistemas de IA, lo que contribuirá a una mayor legitimidad y aceptación pública de estas innovaciones (European Union, 2024).

Los repositorios de IA se han conceptualizado como una ventana a los sistemas de IA utilizados por una organización gubernamental (Haataja *et al.*, 2020). Funcionan como directorios consolidados de información e incluyen información estandarizada sobre dichos sistemas (Valderrama *et al.*, 2022). En este sentido, son un mecanismo de transparencia sobre los sistemas de IA, pues permiten el acceso público a la información sobre los sistemas y procesos algorítmicos (Ada Lovelace Institute, AI Now Institute y Open Government Partnership, 2021).

Dichos repositorios devienen en un importante mecanismo de transparencia algorítmica, especialmente relacionado con el acceso a la información; sin embargo, no puede pensarse que su adopción es suficiente para cumplimentar los requerimientos de aquella. En otras palabras, como instrumento de transparencia algorítmica son necesarios, pero no suficientes, y pueden ser independientes o estar integrados a otros mecanismos de transparencia algorítmica.

Los repositorios de IA constituyen herramientas de transparencia relevantes tanto para las propias entidades públicas como para la sociedad en sentido general. Son parte del modelo de gobernanza para la adopción, uso ético, responsable, transparente y explicable de sistemas de IA en el sector público, por ello, no es suficiente saber cuáles son

las herramientas tecnológicas que existen, también es importante que la información contenida sea completa, posible, abierta, comprensible, accesible y gratuita (CDDO y RTA, 2024).

La implementación de un repositorio no solo permite materializar la transparencia algorítmica y la rendición de cuentas significativa, también debería ayudar a adoptar decisiones informadas y basadas en la evidencia sobre políticas públicas en relación con el tema de la IA, sus usos y retos en la sociedad, incluyendo cuestiones de índole regulatoria y gobernanza.

Tampoco eso debe ser visto únicamente como la aportación de información concreta sobre sistemas de IA en el contexto público, sino que se debe ir más allá. Para que cumpla sus funciones y objetivos previstos es importante que su construcción y utilización sean de forma compartida, no solo por la administración pública; es necesario que intervengan también la academia, investigadores y sociedad en general, quienes deben utilizar la información de forma responsable y convertirse en actores de su configuración, además de participantes directos en la generación de confianza que necesita el ecosistema de uso de la IA en la toma de decisiones públicas.

La implementación de los repositorios de sistemas de IA presenta retos relacionados con su disponibilidad y funcionamiento, y con las funciones que debería cumplir dentro de la transparencia algorítmica. Estos retos abarcan diversas dimensiones, desde la suficiencia y calidad de la información, hasta la sostenibilidad financiera, promover la innovación y la adecuada comunicación de sus resultados. O el establecimiento de estándares uniformes e interoperables que permitan cumplir con los estándares de transparencia necesarios, pero que permitan contrastar la información entre diferentes repositorios. Los estándares son esenciales para la configuración del repositorio y deben ser adoptados con base en un criterio mínimo. Un principio fundamental deberá ser el de mayor apertura posible sin poner en riesgo información sensible, derechos humanos, privacidad y seguridad de los datos.

Uno de los aspectos más controvertidos de esta herramienta es el carácter obligatorio o proactivo de la información que se suministra. La declaración de la herramienta tecnológica es vital para el

repositorio, tanto desde el punto de vista de su cantidad como de su calidad. ¿Quién debe suministrar esta información? ¿Cómo y cuándo hacer ese suministro? ¿Cuánta información es necesaria? Son algunas de las preguntas que deben responderse. Además, la información se relaciona estrechamente con el ciclo de vida de la tecnología, por lo tanto, no será siempre la misma, de modo que puede ser revisada, contrastada, modificada e incluso eliminada cuando no cumpla los fines para los cuáles se creó.

La sostenibilidad y el financiamiento representan otro desafío en la consolidación de repositorios de sistemas de IA, ya que su creación y mantenimiento requieren de recursos económicos, humanos y tecnológicos. Para que estos repositorios cumplan su función de manera efectiva, es indispensable garantizar su actualización permanente, e implantar mecanismos que permitan su adecuado mantenimiento y la difusión de sus resultados.

Por otro lado, la innovación debe ser promovida como un eje transversal en la implementación de estos repositorios, fomentando el aprendizaje a partir de experiencias previas y la aplicación de tecnología avanzada para optimizar su funcionamiento. La adopción de buenas prácticas derivadas de casos de éxito puede contribuir a la mejora continua de los repositorios, y a la reducción de costos dentro del ámbito público.

Finalmente, la comunicación y difusión de los repositorios de sistemas de IA es un aspecto fundamental para garantizar que la sociedad pueda acceder a información clara, comprensible y actualizada sobre el impacto de la IA en su vida cotidiana. Es imperativo que los ciudadanos tengan acceso a datos concretos sobre las aplicaciones de estas tecnologías, sobre sus beneficios y riesgos, de manera que puedan tomar decisiones informadas y ejercer sus derechos de manera efectiva.

Experiencias previas

Aunque el estado del arte sobre las iniciativas de repositorios de sistemas de IA se encuentra en el documento *Algorithmic Transparency*

in the Public Sector: A State-of-the-Art Report of Algorithmic Transparency Instruments elaborado por el GPAI (2024a), es importante abordar algunas de las más importantes, tanto para ser utilizadas como referentes, como para comprender los distintos actores y experiencias que existen, provenientes desde los propios gobiernos y sus instituciones, y de la academia y organizaciones de la sociedad civil (NYC Automated Decision Systems Task Force, 2019).

La idea inicial de los repositorios de sistemas de IA fue desarrollada desde el contexto de los niveles de gobierno locales, con la finalidad de saber cómo la IA impacta la vida de los ciudadanos en estos contextos. Algunos ejemplos son el repositorio de sistemas de IA de Ámsterdam (City of Amsterdam, s. f.), Helsinki (City of Helsinki, s. f.), Nueva York (NYC Office of Technology and Innovation, 2023), y Nantes (Nantes Métropole, s. f.), entre otros.

A nivel regional tenemos el caso del Laboratorio del Foro Digital de Eurocities, que funciona en colaboración con Ámsterdam, Barcelona, Bruselas, Eindhoven, Mannheim, Rotterdam y Sofía, el cual ayuda a ciudades europeas a proporcionar información sobre los sistemas de IA que apoyan sus decisiones; con tal propósito ha desarrollado un estándar de transparencia algorítmica, disponible públicamente (The Algorithm Register, s. f.). Un segundo caso es el *datasets* de casos seleccionados de instituciones del sector público en Europa que fue utilizado en el Informe 129301 del Centro Común de Investigación (CCI) "AI Watch. Panorama europeo del uso de la Inteligencia Artificial por el sector público", y que se actualiza en el Observatorio de Vigilancia Tecnológica del Sector Público (European Commission, s. f.).

A nivel de país, por ejemplo, en Reino Unido el Central Digital and Data Office (CDDO) y la Responsible Technology Adoption Unit (RTA) han implementado un centro de registro de normas de transparencia algorítmica que ayuda a las organizaciones del sector público a proporcionar información clara sobre los sistemas de IA que utilizan para apoyar las decisiones. En este sentido han adoptado el *algorithmic transparency recording standard*, para estandarizar la información en las instituciones públicas (CDDO y RTA, 2024). Con similar fin el gobierno

federal de los Estados Unidos exige a los organismos federales que informen sobre el uso que hacen de la IA y publica un inventario de los casos de uso (U. S. Government, s. f.).

En América Latina, uno de los proyectos pioneros fue el Repositorio de Algoritmos Públicos de la Escuela de Gobierno de la Universidad Adolfo Ibáñez (GobLab UAI) que constituye una plataforma digital donde se "recopila y muestra la información sobre sistemas de soporte o toma de decisiones automatizadas que han sido desarrollados o implementados por instituciones públicas en Chile" (Algoritmos Públicos Chile, s. f.). También es posible encontrar la base de datos sobre el sistema de decisión automatizada en el sector público colombiano (Gutiérrez *et al.*, 2023) desarrollado por investigadores de la Universidad del Rosario; el Informe sobre Sistemas de Decisiones Automatizadas [SDA] en el Estado Nacional de Argentina, elaborado por Algorithmic Avengers, como organización independiente, que lleva a cabo el proyecto Observatorio Algorítmico, en el cual se monitorean y divulgan los sistemas de IA con impacto social que se utilizan en el sector público argentino (Algoritmos Estatales, s. f.). En Uruguay existe el Observatorio de Uso de Inteligencia Artificial en el Estado, el cual, en 2023, elaboró un procedimiento para relevar los casos de uso de IA en el Estado (Agencia de Gobierno Electrónico y Sociedad de la Información y el Conocimiento, AGESIC, s. f.).

Además de las prácticas y casos descritos, hay un esfuerzo por institucionalizar estas iniciativas, ya sea porque se ordena la elaboración de un repositorio o porque las entidades públicas están obligadas a notificar el uso de esta tecnología en sus procesos internos o en la prestación de servicios públicos. Algunos de estos esfuerzos han quedado como propuestas regulatorias; otras ya se han adoptado.

El gobierno de Canadá, por ejemplo, ha implementado la *Directiva sobre la toma de decisiones automatizada*. Y aunque la finalidad de esta no es la creación de un repositorio, sí contempla, entre sus principales resultados, recabar los datos e información sobre el uso de sistemas automatizados de toma de decisiones en las instituciones federales para poner a disposición del público (Treasury Board

of Canada Secretariat, s. f.). En Estados Unidos, la Orden Ejecutiva 13960 del 3 de diciembre de 2020 del presidente Joe Biden, *Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government*, establecía el inventario de casos de usos y determinaba la obligación de cada agencia de preparar un inventario de sus casos de uso no clasificados y no sensibles de IA (Office of the Federal Register, 2020). En Chile, el Consejo para la Transparencia, en sus recomendaciones sobre transparencia algorítmica, pide a los sujetos obligados que mantengan a disposición pública permanente, a través de sus sitios electrónicos, la información relacionada con los sistemas de decisiones automatizadas y semiautomatizadas (Consejo para la Transparencia de Chile, 2024). En Australia, se puso en marcha y en 2024 el *Standard for AI Transparency Statements*, como parte de una política para uso responsable de los sistemas de IA en el gobierno. Se contempla la obligación de las agencias de realizar una declaración pública sobre su uso y adopción. Y al igual que en Chile, se determina la información que se necesita publicar.

A nivel de propuestas, la *Public Authority Algorithmic and Automated Decision-Making Systems Act* 2024 (HL Bill 27), de Reino Unido, además de las obligaciones de las autoridades de llevar a cabo evaluaciones de impacto de los sistemas de IA, busca implementar un registro de transparencia algorítmica que debe ser publicado en formato accesible, en un plazo de 30 días a partir de la finalización del registro (UK Parliament, s. f.). En el mismo sentido, la propuesta de regulación SB 5116 - 2021-22, "Establishing guidelines for government procurement and use of automated decision systems in order to protect consumers, improve transparency, and create more market predictability", del estado de Washington, propone la obligación de poner a disposición del público un inventario completo de todos los sistemas de decisión automatizados que han sido propuestos o están siendo utilizados, desarrollados o adquiridos por agencias estatales, incluidos todos los informes de responsabilidad sobre el uso de sistemas de IA (Washington State Legislature, 2021).

Es imposible resumir todas las experiencias en cuanto a los repositorios de sistemas de IA públicos a nivel global. Representan

múltiples retos, pero muchos países trabajan en su implementación o perfeccionamiento. Esto se debe, entre otras razones, a que son una valiosa herramienta para promover la adopción transparente de sistemas de IA en el sector público.

III. Estrategia de mapeo de sistemas de IA

Teniendo en cuenta el objeto de la investigación y las fuentes de información que podrían utilizarse para el presente estudio, se adoptó una estrategia de métodos mixtos para el mapeo de los sistemas de IA. Esto se definió de acuerdo a cuatro etapas: a) delimitación de las entidades que forman parte del objeto de estudio, b) delimitación de fuentes de información, c) determinación de criterios de selección y exclusión de sistemas de IA, y d) definición de una técnica para el análisis de datos. En las siguientes subsecciones se explica cada caso.

Delimitación del objeto de estudio

El marco de mapeo del estudio se compone, principalmente, del *sector público* en México, mismo que se define como el conjunto de instituciones, agencias, empresas y otras entidades mediante las cuales el Estado cumple o hace cumplir la política o voluntad expresada en las leyes del país. No obstante, es importante considerar que la literatura en administración pública ha señalado un proceso paulatino de transformación del ámbito público hacia formas de nueva gobernanza, donde la atención a problemas —u oportunidades— de interés público, ya no recae únicamente en las entidades controladas por el Estado, sino también en las redes de actores independientes (estatales, privados y de la sociedad civil) que participan en la formulación

e implementación de políticas públicas (Rhodes, 1996). Desde este enfoque, el *sector público* se plantea como parte del aparato estatal tradicional (como fue definido arriba), cuya actividad se relativiza en un *ámbito público* más amplio, que incorpora a una variedad de entidades del *ámbito público no gubernamental*.

Desde tal perspectiva, se considera relevante mapear la incorporación de sistemas de IA en entidades del sector público, y en las del ámbito público no gubernamental que atienden problemas públicos. Más aún, la literatura en innovación pública ha demostrado que este último tipo de entidades juega un papel fundamental en el desarrollo de prácticas innovadoras que las entidades del primero incorporan a su quehacer eventualmente (Criado *et al.*, 2025); al respecto, un ejemplo reciente es la evolución de los espacios de deliberación pública ocurrida en varios países, donde algunos formatos de deliberación no convencionales, pero creados por ciudadanos, van encontrando poco a poco su lugar en las prácticas de gobierno e instituciones públicas (Pitti, 2018). Se entiende por entidades del *ámbito público no estatal* a organizaciones de carácter social, académico o privado que, sin formar parte del aparato gubernamental, realizan actividades vinculadas, directa o indirectamente, con la atención de los problemas públicos (Rhodes, 1996).

En consecuencia, el objeto de estudio del levantamiento corresponde al ámbito público mexicano, definición que incorpora a entidades del sector público y a las del ámbito público no estatal. Esta delimitación incluye a las entidades del aparato estatal tradicional, y también a universidades, centros de investigación y organizaciones de la sociedad civil que participan en el desarrollo de soluciones tecnológicas para la atención de problemas públicos. Entidades en esta esfera están impulsando prácticas innovadoras que ofrecen lecciones importantes de los desarrollos generados en el ámbito público no estatal, los cuales podrían incorporarse al sector público. Por ejemplo, OlimpIA, un sistema de IA desarrollado por Defensoras Digitales, es un asistente virtual que brinda apoyo a mujeres víctimas de violencia digital en más de treinta idiomas. Por su parte, el Consejo Ciudadano para la Seguridad y Justicia de la Ciudad

de México creó el asistente virtual Adelita Sin Delito, en el que la ciudadanía puede denunciar reportes de robo, extorsión, violencia familiar, maltrato animal o trata de personas, entre otros. El sector académico también ha hecho su aportación incorporando sistemas de IA innovadores, como el caso de la Universidad Autónoma de Tamaulipas (UAT), que generó un sistema que traduce información del español a lenguas indígenas.

Acotado el objeto de estudio del levantamiento, enseguida se detallan las decisiones operativas que delimitaron el marco de mapeo, procediéndose a realizar lo mismo con las entidades que serán mapeadas. En primer lugar, respecto al sector público, resalta que el gobierno mexicano se conforma de una amplia gama de tipos de entidad pública, y no todas son objeto de interés para el presente estudio. Las entidades públicas mexicanas se pueden organizar en cuatro ramos: los tres poderes de los gobiernos federal y estatales (ejecutivo, legislativo y judicial), y los órganos autónomos (véase el apéndice B de este libro). A su vez, cada ramo de gobierno se compone de diferentes tipos de entidades públicas. Se utilizó la clasificación de *sujetos obligados* de la Ley General de Transparencia y Acceso a la Información Pública como fundamento para identificarlos, esto es: organismos descentralizados, descentralizados no sectorizados, desconcentrados, centros de investigación, administración pública centralizada, instituciones de banca descentralizada, dependencias del poder ejecutivo, entidades paraestatales, organismos técnicos especializados, órganos reguladores, organismos jurisdiccionales, organismos públicos autónomos, organismos del poder judicial, organismos del poder ejecutivo, órgano imponedor de justicia, órganos públicos constitucionales y órganos constitutivos autónomos.

Se incluyeron en el marco de mapeo solo las entidades públicas vinculadas, directa o indirectamente, con la provisión de servicios públicos, operación de políticas públicas o servicios a la ciudadanía. Esto, con el argumento de que los sistemas de IA de interés para el mapeo son los de incidencia en el interés colectivo o en la atención de problemas públicos. Con este criterio, se excluyó del mapeo a los siguientes sujetos obligados: sindicatos, fondos, consejos, fideicomisos,

asociaciones civiles, mandatos, patronatos, partidos políticos y empresas productivas del Estado.

Los gobiernos municipales representaron un reto importante debido a su cantidad y al número de entidades públicas que los componen. Por restricciones de tiempo y recursos, se optó por focalizar el mapeo de estas entidades, circunscribiendo el marco de mapeo a los gobiernos municipales considerados como *metrópolis*, según la Secretaría de Desarrollo Agrario, Territorial y Urbano (Sedatu, 2024, p. 43): "Una metrópoli es el conjunto de municipios cuya extensión urbana puede rebasar sus límites político-administrativos, concentra 50 mil o más habitantes y tienen un alto grado de integración física o funcional; también se incluyen aquellos que son de gran relevancia en términos económicos o alojan a la sede del gobierno estatal". Para 2024, la Sedatu identificó 48 metrópolis en el país, compuestas por 345 municipios, los cuales quedaron considerados como parte del marco de mapeo de este estudio.

Un segundo reto del mapeo de gobiernos municipales corresponde a la cantidad de entidades públicas (o sujetos obligados) que los componen: ayuntamiento, sistemas e institutos municipales, organismos operadores. Una composición que varía entre municipios, es decir, que no hay uniformidad en la cantidad o tipos de entidades públicas que forman parte de un gobierno municipal. Por esta razón, el marco de mapeo utilizado aquí considera solo a los ayuntamientos como el sujeto obligado seleccionado para cada uno de los 345 gobiernos municipales mapeados.

Por último, el marco de mapeo contempla el análisis de entidades del *ámbito público no gubernamental*, es decir, organizaciones de la sociedad civil, universidades[1] y asociaciones profesionales. A diferencia de las del sector público, donde se puede acceder a un marco bien delimitado de las entidades que cumplen con dicho criterio, las del ám-

[1] En México, algunas universidades públicas son formalmente sujetos obligados. Debido a esto, las que tienen esta característica fueron parte del mapeo mediante solicitudes de transparencia, la misma técnica de levantamiento utilizada para las entidades del sector público.

bito público no gubernamental difícilmente son identificables; estos casos emergieron en nuestro mapeo de manera casuística, en particular mediante estrategias de monitoreo. En el cuadro III.1, se sintetizan los tipos de entidades del sector público y del ámbito público no gubernamental, que conforman nuestro objeto de estudio.

Cuadro III.1. Tipos de entidades que conforman el marco de mapeo

Tipo	*Descripción*
Sector público	Conjunto de ministerios y agencias que componen la administración pública gubernamental (federal, estatales y municipales) y que tienen competencias legales y administrativas vinculadas a la provisión de bienes y servicios públicos.
Ámbito público no gubernamental	Organizaciones de carácter social, académico o privado que, sin formar parte del aparato gubernamental, realizan actividades de interés público.

Fuente: Elaboración propia.

Recolección de datos

Para el mapeo se recolectó la información mediante cuatro fuentes: 1) solicitudes de transparencia, 2) monitoreo de medios, 3) referencias directas y 4) otros. Las solicitudes de transparencia constituyeron la principal fuente de información del proceso de recolección de datos, no solo por la magnitud del levantamiento, sino también porque permitió mapear la totalidad de las entidades del sector público que cumplían con los criterios de inclusión para el estudio. Se enviaron 2241 solicitudes a través de la Plataforma Nacional de Transparencia, es decir, una por sujeto obligado (cuadro III.2). El periodo de levantamiento de información con esta estrategia inició el 1 de marzo de 2024 y concluyó el 1 julio de 2024. La tasa de respuesta a las solicitudes por parte de los sujetos obligados osciló entre el 67 y el 96%.

Asimismo, la tasa de respuestas positivas alcanzó del 0.7% en entidades de gobiernos municipales al 12% en las del federal. Las respuestas positivas son aquellas en las que las entidades públicas afirmaron contar con sistemas de IA en su organización. Algunas entidades

públicas reportaron más de un sistema de IA, por lo que, de las 190 respuestas positivas, se obtuvo un listado de 356 posibles sistemas de IA. Después de un análisis pormenorizado de dicho total reportado, se identificaron únicamente 49 sistemas de IA. En la siguiente subsección se explican los criterios que guiaron la exclusión de los 307 casos restantes.

Cuadro III.2. Relación de solicitudes enviadas y respuestas obtenidas

Nivel de gobierno	*Solicitudes enviadas*	*Tasa de respuestas*	*Tasa de respuestas positivas*
Federal	225	217 (96%)	27 (12%)
Estatal	1671	1401 (83%)	138 (8%)
Municipal	345	230 (67%)	25 (0.7%)
Total	2241	1848 (82%)	190 (8.5%)

Fuente: Elaboración propia.

Todas las solicitudes de transparencia siguieron el mismo esquema: requisitar información sobre el nombre del sistema de IA, el puesto laboral de la persona responsable de su uso, el objetivo de uso del *software* y, en caso de que el sistema se haya adquirido por contratación o convenio, la provisión de estos. Con la información obtenida generamos dos bases de datos: la del repositorio de sistemas de IA, y la que integró todas las respuestas a las solicitudes de transparencia y de los contratos o convenios asociados.[2]

La segunda fuente de información para identificar sistemas de IA fue un monitoreo de medios de comunicación y de redes sociales, mismo que se realizó entre el 20 de junio de 2024 y el 12 de noviembre de 2024. Para ello se utilizaron como medios de consulta el buscador de Google y las redes sociales LinkedIn, X (antes Twitter), Instagram y Facebook. Adicionalmente, se dio seguimiento a páginas de internet de agencias de noticias, conferencias, así como organismos

[2] Dichas bases de datos pueden consultarse en https://doi.org/10.17605/OSF.IO/CU6W3

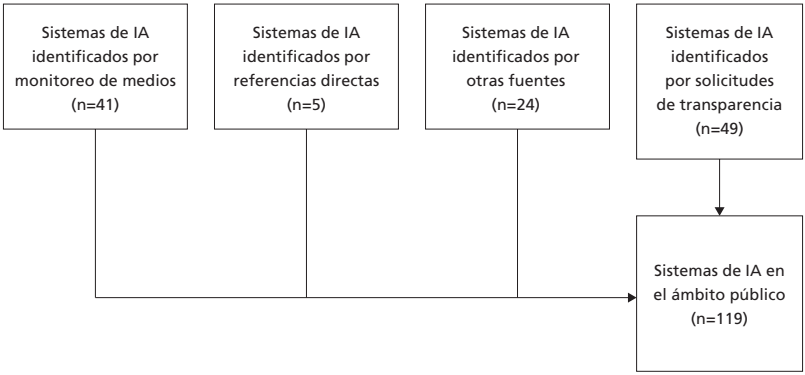
nacionales e internacionales, entre otros. Con esta fuente de información se identificaron 41 sistemas de IA.

La tercera fuente de información fueron las referencias directas que el equipo de investigación obtuvo a lo largo del proceso de investigación. Este concepto se refiere a formas de consulta directa con actores relevantes del sector de IA en México (p. ej., empresas de desarrollo tecnológico, directores de tecnologías de la información, investigadores especializados y desarrolladores), a través de entrevistas semiestructuradas, interacción con expertos en foros, conferencias o seminarios, y comunicaciones escritas. Esta fuente de información arrojó cinco sistemas de IA que no se identificaron con las otras fuentes de información.

Finalmente, el estudio se enriqueció de otras dos fuentes de información coyunturales acontecidas durante el proceso de levantamiento. Primero, por medio del apoyo de la Coordinación de la Estrategia Digital Nacional (CEDN), que brindó la oportunidad al equipo de investigación de enviar una consulta a las direcciones generales de Tecnología de la Información y la Comunicación de los organismos públicos federales vinculadas a las actividades de la CEDN. Se envió una encuesta a dichas entidades por correo electrónico, de las que se recibieron 71 respuestas. Otra fuente de información coyuntural fue el listado de las entidades participantes en el concurso "Premio IA en Gobierno 2024", organizado por el Instituto Federal de Telecomunicaciones;[3] y cuya información fue facilitada al equipo organizador por los responsables del concurso con la finalidad de robustecer el repositorio. Luego de la consulta, se identificaron 21 sistemas. Otro concurso relevante fue "Las más innovadoras 2024", organizado por la empresa Netmedia; allí se presentaron sistemas utilizados en el sector público. En total, con este conjunto de fuentes se identificaron 24 sistemas de IA. En el gráfico III.1, se presenta un resumen de los sistemas identificados, según la fuente de información.

[3] https://www.ift.org.mx/transformacion-digital/blog/premio-ia-en-gobierno.

Gráfico III.1. Resumen de sistemas de IA identificados, por fuente de información



Fuente: Elaboración propia.

Casos excluidos

La recolección de casos requirió de un análisis preliminar para determinar si se les incluía o no en el repositorio en función de sus cualidades como sistemas de IA. Se definieron tres criterios de exclusión de casos: 1) *software* convencional, 2) IA de propósito general, y 3) sistemas de IA sin vínculo con la función pública.

La categoría *software convencional* se refiere, en términos generales, a los sistemas informáticos basados en reglas o programación determinista, en contraste con los basados en aprendizaje automático. Si bien el objetivo de este sistema es identificar falsos positivos de IA, es cada vez más difícil hacer una clara distinción entre los sistemas informáticos que tienen (o no) un sistema de IA embebido en sistemas convencionales, ya que es posible que las empresas incorporen funcionalidades basadas en IA en un *software convencional*. Por ejemplo, la paquetería de Microsoft Office utiliza aplicaciones de IA embebidas en sus aplicaciones más conocidas como Word, Excel o PowerPoint, aunque el servicio que brindan estos sistemas no depende directamente de la funcionalidad de los sistemas de IA (aprendizaje, predicción o conversación), como es el caso de sistemas tipo ChatGPT o

Midjourney. Por tanto, se excluyeron del repositorio los sistemas informáticos donde la posible funcionalidad de la IA no es clara, como Microsoft Office (Excel, Word, PowerPoint), Contpaq, Saacg, Aspel, Nomipaq o sistemas de antivirus.

La categoría *IA de propósito general* se refiere, en un sentido práctico, a sistemas de IA "listos para usarse" y desarrollados para usos amplios, que son adaptables a muchos ámbitos de trabajo, incluyendo el sector público. En un sentido más amplio, se entiende por IA de propósito general aquel sistema de IA que puede realizar, o ser adaptado para realizar, un rango diverso de tareas, incluidas algunas para las que el sistema no fue diseñado (Gutiérrez *et al.*, 2023). Se trata de sistemas con una generalidad funcional (afín a la noción tradicional de tecnologías de propósito general), que le brinda la capacidad de generar valor en un rango amplio de tareas, problemas o contextos. Este tipo de sistemas son, por ejemplo, ChatGPT, DALL-E, Copilot, Midjourney, DeepL o Amazon Transcribe. Un caso ilustrativo es el de la Comisión de Búsqueda del Estado de Coahuila que utiliza DJI Terra, un *software* bajo licenciamiento que facilita el estudio de terrenos potencialmente relacionados con restos inhumados de personas desaparecidas. Este *software* permite la creación de modelos 3D y el mapeo de áreas mediante el uso de IA.

Si bien son sistemas cuyo uso está generando valor en la función pública, ellos se caracterizan por su uso individual e informal en el espacio de trabajo y porque se adquieren mediante servicios de suscripción, sin necesidad de contratos, licitaciones públicas o convenios de colaboración para hacer usar el sistema en la organización. Se trata de un tipo de sistema de IA que abre líneas de investigación con respecto a la incorporación de sistemas de IA en el servicio público, por su carácter polifuncional y de uso informal —hasta el momento— en las entidades públicas. En este sentido, se observó que una función destacada de este tipo de sistemas es la generación de contenidos multimedia, que, a través de las solicitudes de transparencia, varias entidades públicas reportaron que utilizan para apoyo en las labores de comunicación institucional.

La tercera categoría de exclusión, *sistemas de IA sin vínculo con la función pública*, se refiere a la identificación de casos en varias fuentes de información que no fueron desarrollados para su uso en la función pública o no inciden en la esfera pública. Varios provienen de universidades públicas mapeadas (en su carácter de sujetos obligados), de monitoreo de medios o de premios. Por ejemplo, el Centro de Investigación en Matemáticas desarrolló un modelo de predicción de fallos en maquinaria y conteo de inventario a través de análisis de imágenes. O el programa de actualización docente sobre la IA generativa en contextos educativos reportada por el Tecnológico de Estudios Superiores de Coacalco.

Dichos criterios de exclusión fueron aplicados por los investigadores en todas las consultas. En cuanto al monitoreo de medios, referencias directas y otras fuentes, ello se hizo directamente, por lo que no se guarda registro de los casos excluidos. Respecto a las solicitudes de información, enseguida se expone un análisis de estadística descriptiva de los casos excluidos que, además, ofrece un panorama del nivel de familiarización de los funcionarios públicos con los sistemas de IA: en la solicitud de información, se presentó a los sujetos obligados una definición de sistemas de IA[4] para minimizar la confusión con otro tipo de sistemas informáticos, sin embargo, un amplio porcentaje de los 356 casos obtenidos mediante esta consulta fueron excluidos. Se excluyeron 220 sistemas por ser considerados *software* convencional; se hizo lo mismo con otros 49 bajo el concepto de IA de propósito general y 38 a partir del criterio de sistemas de IA sin vínculo con la función pública (cuadro III.3). Así, de los 356 casos reportados por los sujetos obligados, solo se incorporaron 49 casos al repositorio.

[4] La presentación completa de la solicitud fue la siguiente: "Toda la información sobre la contratación y/o desarrollo interno de cualquier instrumento de *software* que haga uso de algoritmos de inteligencia artificial, herramienta tecnológica o equivalente desde el primero de enero de 2020 a la fecha de la solicitud. En caso de que exista un *software* que haga uso de algoritmos de inteligencia artificial, herramienta tecnológica o equivalente, solicito la siguiente información: […]."

Cuadro III.3. Casos excluidos para las respuestas a las solicitudes de transparencia

Criterio de exclusión	*Casos excluidos*
Software convencional	220 (62%)
IA de propósito general	49 (14%)
Sistemas de IA sin vínculo con la función pública	38 (11%)
Total	307 de 356 (87%)

Nota: Los casos incluidos fueron 49, cifra que se considera al contabilizar un total de 356 sistemas de IA.
Fuente: Elaboración propia.

El proceso de recolección de datos comenzó el 1 de marzo de 2024 y concluyó el 12 de noviembre de 2024. Como resultado final, se identificaron 119 sistemas de IA para la función pública, de los cuales 97 corresponden al sector público y 22 a entidades no gubernamentales del ámbito público (cuadro III.4).

Cuadro III.4 Sistemas de IA para la función pública identificados en el ámbito público mexicano

Tipo de entidad	*Casos*
Entidades del sector público	97 (82%)
Entidades no gubernamentales del ámbito público	22 (18%)
Total	119 (100%)

Fuente: Elaboración propia.

En suma, la selección de sistemas de IA incorporó los que se usan para la función pública y que se caracterizan por estar diseñados o adaptados para tareas administrativas, regulatorias o de prestación de servicios públicos específicos de la función pública o ajustados al contexto del ámbito público. Naturalmente, es difícil una distinción clara entre sistemas informáticos con una funcionalidad de IA o entre sistemas adaptados para la función pública y sistemas de propósito general —también adaptados, a su manera, para la función pública—. Por tanto, se asume como una limitación de la investigación el incurrir en la inclusión o exclusión errónea de casos en el repositorio.

Análisis de datos

El análisis de los datos recolectados consistió en investigar las características de los sistemas de IA o del contexto en el que funcionan. Este proceso está determinado por los objetivos de investigación del repositorio: generar información sobre características tecnológicas, funcionales, contractuales y de transferencia tecnológica de los sistemas de IA que se utilizan en el ámbito público. Con base en este propósito, se elaboró un esquema de codificación de los datos dividido en tres secciones. La primera comprende variables sobre información general de los sistemas de IA y de su incorporación al repositorio, como el nombre del sistema de IA, la entidad pública que lo utiliza, fuente de información mediante la que se documentó y la fecha de incorporación al repositorio, entre otras (cuadro III.5).

Cuadro III.5. Información general del sistema de IA[a]

Atributo	*Descripción*
ID	Número de identificación único en el repositorio de sistemas de IA
Nombre	Nombre del sistema de IA
Entidad pública	Entidad pública donde se utiliza el sistema de IA
Acrónimo	Acrónimo de la entidad pública
Rama	Rama del poder público de la entidad pública
Nivel de gobierno	Nivel de gobierno de la entidad pública
Entidad federativa	Nombre de la entidad federativa en que se encuentra la entidad pública
Código INEGI	Clave de registro único de la entidad política, de acuerdo con la clasificación del INEGI[b]
Incorporación	Fecha de incorporación del sistema de IA al repositorio
Fuente	Fuente de información mediante la que se documentó el registro del sistema de IA
Enlace	Enlace de internet que provee información adicional sobre el sistema de IA
Estatus	Situación de utilización en la que se encuentra el sistema de IA
Verificación	Fecha de última consulta en las fuentes de información para verificar el estatus del sistema de IA

Notas: [a]Solo se presenta el nombre del atributo y su descripción; la de las categorías analíticas correspondientes a cada atributo se puede consultar en el apéndice A de este libro. [b]Se utiliza el código correspondiente al Catálogo Único de Claves de Áreas Geoestadísticas Estatales, Municipales y Localidades del INEGI (https://www.inegi.org.mx/app/ageeml/).

Fuente: Elaboración propia.

La segunda sección del esquema de códigos describe las características tecnológicas y funcionales del sistema de IA (cuadro III.6).[5] Se compone de una descripción del sistema presentada por los sujetos obligados (que proviene de las respuestas de las solicitudes de transparencia o que se extrajo de la información pública disponible) para los casos identificados mediante otras fuentes de consulta. Esta categorización es particularmente difícil porque existen diversas formas de clasificación de los sistemas de IA (OCDE, 2022b, 2024b) y porque la información disponible es limitada en la mayoría de los casos. Por lo tanto, se optó por una categorización corta que reflejara, lo mejor posible, la forma en que se presenta el sistema de IA en la información disponible. También se identificó la tarea que el sistema de IA realiza para apoyar la actividad rutinaria de la organización, como el análisis de datos, la estimación de riesgos, el diagnóstico de salud, el procesamiento de información o la gestión de trámites, reportes o denuncias, entre otros. El atributo de función se refiere al modo en que el sistema de IA contribuye a los objetivos de la entidad pública. En este rubro, clasificamos los sistemas en tres grandes categorías: a) los que contribuyen a mejorar la entrega de servicios públicos (p. ej., sistemas de consulta para informar a la ciudadanía, reducción de cargas administrativas a la ciudadanía, facilitar trámites o denuncias a la ciudadanía); b) los utilizados para operaciones internas de la entidad pública (p. ej., operación de programas públicos, reducción de carga laboral de los servidores públicos o la planeación y toma de decisiones), y c) los que apoyan las gestión del desempeño de los empleados o programas públicos (p. ej., evaluación de programas o formación de capital humano).

Otros rubros relevantes en esta sección son el sector de desarrollo conforme a la clasificación del Banco Internacional de

[5] Salvo el rubro de descripción (que es información obtenida por fuentes externas), los atributos de esta sección fueron codificados por el equipo de investigación, de acuerdo con categorías analíticas predefinidas (véase el apéndice C de este libro).

Desarrollo (BID) y el Objetivo de Desarrollo Sostenible en el que se enmarca el sistema de IA. La clasificación del BID utilizada se basa en los Documentos del Marco Sectorial de Innovación, Ciencia y Tecnología del BID, los cuales ofrecen información precisa sobre los sectores de desarrollo de cada país, a donde tributan específicamente cada uno de los sistemas de IA utilizados. Por su parte, la clasificación basada en los Objetivos del Desarrollo Sostenible (ODS) permite tener una idea sobre el impacto de los sistemas de IA en el cumplimiento de los retos comunes y globales en dicha materia.

Cuadro III.6. Descripción de las características tecnológicas y funcionales del sistema de IA

Atributo	*Descripción*
Descripción	Descripción general del sistema de IA
Tipo de sistema	Tipo de sistema de IA implementado
Tarea	Tarea cognitiva en actividades rutinarias de la entidad pública en la que se utiliza el sistema de IA
Función	Función que cumple el sistema en la entidad pública
Sector	Sector de desarrollo en el que se enmarca el sistema de IA[a]
ODS	Objetivo de Desarrollo Sostenible en el que se enmarca el sistema de IA

Nota: [a]Esta clasificación se fundamenta en el sistema de clasificación de sectores del Banco Interamericano de Desarrollo. Véase https://idbinvest.org/es/sectores

Fuente: Elaboración propia.

La tercera sección del esquema de códigos corresponde a las condiciones de adquisición, desarrollo y despliegue de los sistemas de IA. Estas condiciones se describen tomando en cuenta las formas en que la entidad pública obtuvo la tecnología, ya sea porque la generó internamente o porque la adquirió con terceros (cuadro III.7). Como son diversas las fuentes de información utilizadas y porque la información es inexistente o incompleta, no fue posible precisar para todos los casos la forma de adquisición. En cambio, los esquemas de adquisición entre desarrollos internos y externos se clasificaron con mayor claridad.

Cuadro III.7. Condiciones de adquisición, desarrollo y despliegue

Atributo	*Descripción*
Adquisición	Forma de adquisición del sistema de IA
Desarrollador	Entidad que desarrolló el sistema de IA
Ejecutor	Entidad pública que despliega y/o utiliza el sistema de IA
Financiador	Entidad que financia el desarrollo o implementación del sistema de IA
Costo	Costo del desarrollo (en USD).

Fuente: Elaboración propia.

Entre las formas de adquisición con terceros se encuentran desde las contempladas en la "Ley de adquisiciones, arrendamientos y servicios del sector público para adquirir bienes y servicios", licitación pública, invitación a cuando menos tres personas, y adjudicación directa, hasta otras formas menos claras, como donaciones y acuerdos de colaboración, entre otros.[6] Debe señalarse que, en relación con el grado de madurez de la tecnología, no siempre puede ser acreditado, ni es posible saber si el sistema continúa siendo o no utilizado. Acerca de esto es posible decir que se han incluido sistemas diseñados, desarrollados, así como verificados y validados y, al menos, que se encuentran en fase de prueba en entornos controlados. Esto incluye tanto a los proyectos piloto donde se prueba el sistema de IA en condiciones cercanas a la realidad en un espacio controlado, como cuando el sistema ya está implementado o desplegado en una entidad concreta para resolver problemas específicos.

Para determinar si el sistema se mantiene en uso se debería aplicar una revisión particular, por ejemplo, el sistema Ánimo Tuitero, que

6 Estas eran las formas de contratación aplicables según la "Ley de adquisiciones, arrendamientos y servicios del sector público" vigente al momento de realizar esta investigación. Una nueva "Ley de adquisiciones, arrendamientos y servicios del sector público" fue publicada en el *Diario Oficial de la Federación* el 16 de abril de 2025. A diferencia de su predecesora, esta incluye, además de los referidos, otros procedimientos de contratación como son el diálogo competitivo, la adjudicación directa con estrategia de negociación, la asignación de un contrato específico derivado de la suscripción de un acuerdo marco y la asignación de órdenes de suministro derivados de la Tienda Digital del Gobierno Federal o de servicios conforme a los catálogos electrónicos.

implementa el INEGI, y que fue desarrollado a través de un convenio de colaboración entre el propio INEGI, el Centro de Investigación e Innovación en TIC (Infotec) y el Centro de Investigación en Ciencias de Información Geoespacial (Centro Geo). Ánimo Tuitero tiene por objetivo analizar el estado de ánimo de las personas en X (antes Twitter) en México; concretamente explota estadísticas experimentales para producción de información mediante ciencia de datos. Sin embargo, a causa de las modificaciones en los términos y condiciones para acceder a la fuente de información, el 23 de junio de 2023 se dejó de actualizar este programa.

En suma, el diseño metodológico del presente estudio permitió mapear los sistemas de IA en el sector público mexicano, identificando sus características y condiciones de adopción. La combinación de solicitudes de transparencia, monitoreo de medios y referencias directas aseguró una recolección de datos representativa. No obstante, persisten desafíos como la falta de criterios unificados para clasificar estos sistemas y la opacidad en su contratación. Estos hallazgos resaltan la necesidad de fortalecer la transparencia algorítmica y mejorar la disponibilidad de información para una adopción responsable de los sistemas de IA en el sector público.

IV. Panorama de los sistemas de IA en el ámbito público mexicano

En este capítulo se presentan los principales resultados de nuestro análisis estadístico. La información utilizada está disponible en la base de datos *Repositorio de sistemas de IA en el ámbito público mexicano (Versión 1.2)*.[1] Además, se han elaborado gráficas con estadísticas descriptivas acompañadas de su breve análisis. El análisis estadístico-descriptivo provee información relevante sobre el estado de los sistemas de IA en el ámbito público en México, lo que contribuye a su diagnóstico y caracterización en áreas como número de sistemas de IA por rama del poder público, por nivel de gobierno, entidades federativas, tipos de sistemas utilizados, funciones públicas a las que coadyuvan, y en cuáles ODS inciden, entre otros.

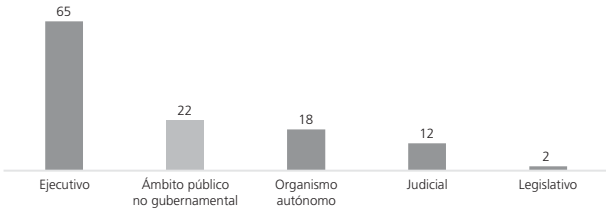
El gráfico IV.1 muestra el número de sistemas de IA por rama del poder público. Cuántos sistemas son utilizados en cada rama, y los que pertenecen al ámbito público no gubernamental, apreciándose que el mayor número de sistemas identificados (97) son utilizados en entidades del sector público, lo que representa el 81.5%. Mientras que solo 22 (18.5%) se utilizan en entidades no gubernamentales del ámbito público.

Entre las entidades del sector público se observa que el poder ejecutivo es la rama que más sistemas de IA utiliza con un total de

[1] Véase Rentería y Ordelin. (2025). *Repositorio de sistemas de IA en el ámbito público mexicano. (Versión 1.2)*. [Data set]. Disponible en https://doi.org/10.17605/OSF.IO/CU6W3

65, es decir, el 54.6% de los sistemas identificados en esas entidades. Mientras que el uso de sistemas en la rama judicial representa el 10.08% de los sistemas, y la legislativa apenas 1.7%. Es notable que, además del ejecutivo, son los órganos autónomos (15%)[2] los que poseen mayor porcentaje de sistemas identificados, por lo que ocupan el segundo lugar.

Gráfico IV.1. Número de sistemas de IA por rama del poder público



Nota: n = 119.

Fuente: Elaboración propia.

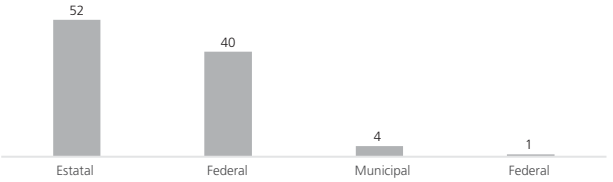
En la rama judicial, el Chatbot Sor Juana ha tenido un alto impacto mediático, tras devenir en el primer sistema de IA conversacional sobre proyectos de sentencia de la Suprema Corte de Justicia de la Nación (SCJN). Hasta enero de 2025 había procesado 21 225 preguntas sobre un total de 75 sentencias, y contabilizaba 7428 usuarios únicos no rastreables. Fue desarrollado por la ponencia de la ministra de la SCJN, Ana Margarita Ríos Farjat, y representa un ejemplo del impacto de la tecnología en la administración de justicia. Una segunda versión del proyecto prevé la inclusión de más asuntos, así como el desarrollo de un nuevo apartado de búsqueda y recomendaciones de asuntos.[3]

2 La identificación de órganos autónomos utilizada fue la que existía antes de la reforma constitucional que eliminó algunos de ellos, publicada en el *Diario Oficial de la Federación* del 20 de diciembre de 2024.

3 Desglose por mes del Proyecto Sor Juana, facilitado por el magistrado Juan Jaime González Varasel, magistrado de circuito, coordinador de la ponencia de la ministra Ana Margarita Ríos Farjat en la Primera Sala de la SCJN.

El gráfico IV.2 muestra el número de sistemas de IA por nivel de gobierno al que pertenece la entidad pública que los utiliza; se excluyen aquellos que pertenecen al ámbito público no gubernamental. El 45.4% de los identificados se utilizan en los gobiernos estatales, lo que representa más de la mitad de los que existen en entidades gubernamentales del país. El 33.6% se utiliza en entidades del gobierno federal incluyendo sus organismos autónomos. Se registraron solo cuatro sistemas de IA en los gobiernos municipales, es decir, el 3.4% del total de los identificados.

Gráfico IV.2. Número de sistemas de IA por nivel de gobierno



Nota: n = 97.

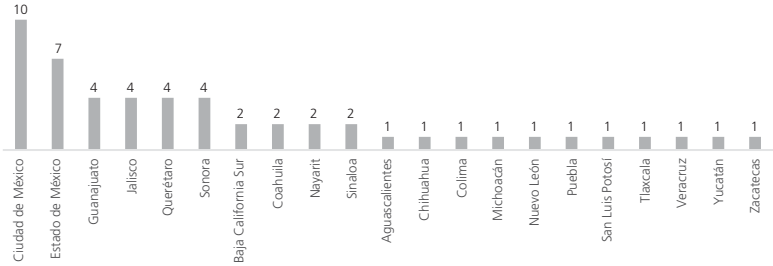
Fuente: Elaboración propia.

Podrían mencionarse múltiples explicaciones para justificar los datos arrojados en el gráfico, las cuáles van desde la disponibilidad de recursos hasta la infraestructura a nivel estatal, entre otros factores. Sin embargo, como se podrá apreciar posteriormente, el comportamiento a nivel de las entidades estatales no es homogéneo; existen diferencias sustanciales en el número de sistemas por entidades federativas. No obstante, los datos arrojados muestran una importante área de oportunidad para desplegar sistemas de IA a nivel local, más cercanas a las personas y a la gestión de asuntos públicos.

El gráfico IV.3 recupera el número de sistemas por entidad federativa, aplicados en cualquier rama del gobierno: ejecutivo, judicial, legislativo, o en órganos autónomos. Si bien los gobiernos estatales son los que más sistemas han desarrollado y declaran poseer, lo

cierto es que la cantidad de sistemas por entidad federativa es dispar, apreciándose una brecha entre las que poseen un mayor número de sistemas como Ciudad de México (11), lo que representa el 17% y prácticamente la quinta parte de los identificados, y las entidades que cuentan con un sistema, o las ocho en las que no se identificó su uso en el ámbito público.

Gráfico IV.3. Número de sistemas por entidad federativa



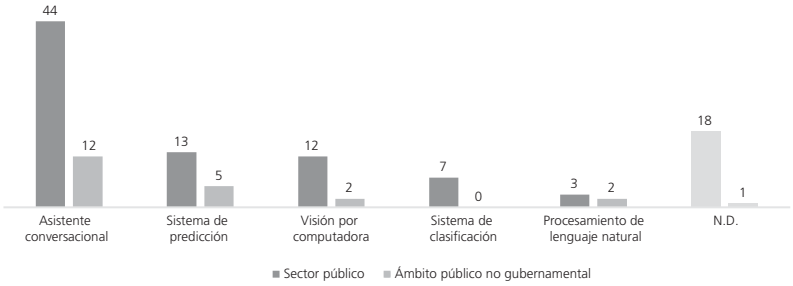
Nota: Solo se contabilizan los sistemas de IA de entidades federativas. (n = 52).

Fuente: Elaboración propia.

La desproporción que se aprecia entre entidades puede obedecer a múltiples factores, pero llama la atención que las que más sistemas reportan son aquellas con mayor aportación al PIB nacional en valores nominales (INEGI, 2024), esto es, Ciudad de México y Estado de México, con 11 y 8 sistemas, respectivamente.

El gráfico IV.4 muestra la cantidad de sistemas de IA según el tipo utilizado. Los tipos se relacionan con las técnicas de aprendizaje y procesamiento de datos que usa cada sistema. En esta variable es importante reconocer las limitaciones del presente estudio, pues al no tener suficiente información, se optó por clasificar conforme a la información disponible en las fuentes originales. Aun así, esta estrategia nos da un panorama de los tipos de sistemas disponibles en el sector público.

Gráfico IV.4. Número de sistemas de IA, según su tipo



Nota: n = 119.

Fuente: Elaboración propia.

Los sistemas de IA más utilizados son los asistentes conversacionales,[4] tanto en el sector público como en el ámbito público no gubernamental con un total de 56 (44 y 12, respectivamente), esto es, el 47% de los identificados. En el sector público los asistentes conversacionales son el 45.4%, mientras que, en el ámbito público no gubernamental, suman el 54.5%. Otro tipo que sobresale es el sistema de predicción con 18 en total. Los sistemas menos identificados son los de clasificación y procesamiento de lenguaje natural. En el ámbito público no gubernamental no se identificó ningún sistema de clasificación, mientras que en el procesamiento de lenguaje natural solo cinco: tres en el sector público y dos en el ámbito público no gubernamental.

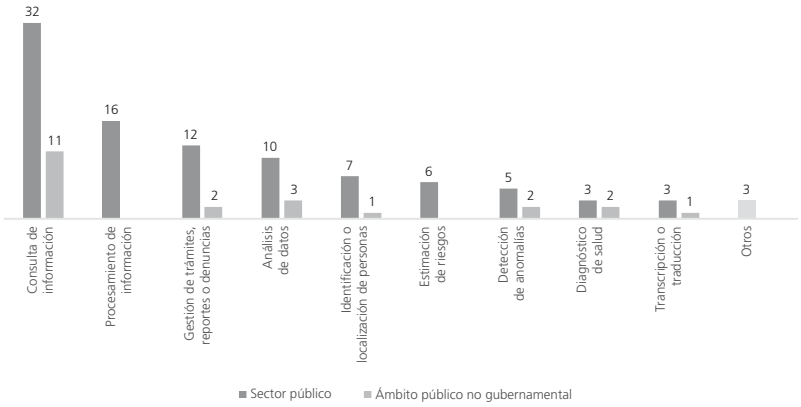
Otro dato por considerar en este caso es la imposibilidad de clasificar los sistemas cuando no hay información contrastable y/o que los describa. Sobresale que, en el sector público, al menos la quinta parte

[4] Se entiende como asistente conversacional a un sistema que se basa en grandes modelos de lenguaje (*large language models*) y que, por lo tanto, cuenta con un amplio repositorio de conocimiento (o datos para consulta) y es capaz de aprender y adaptar sus respuestas en función de la interacción con el usuario (Noonia *et al.*, 2024). En la sección "Análisis descriptivo del uso de los sistemas de IA" del capítulo V, se detalla la distinción funcional entre "chatbots" y "asistentes conversacionales".

de los sistemas identificados (18.5%) no pudo clasificarse, y que en el ámbito público no gubernamental esta cifra se redujo al 4.5%.

El gráfico IV.5 indica la cantidad de sistemas de IA según la tarea que realicen, es decir, su tarea cognitiva o función. Básicamente, la tarea está determinada por la habilidad humana que los sistemas de IA imitan o automatizan, y pueden realizar más de una; sin embargo, para aplicar este criterio de clasificación tuvimos en cuenta la tarea cognitiva o función principal que se declara en correspondencia con los objetivos perseguidos con su aplicación en el ámbito público.

Gráfico IV.5. Número de sistemas de IA por tarea



Nota: n = 119.

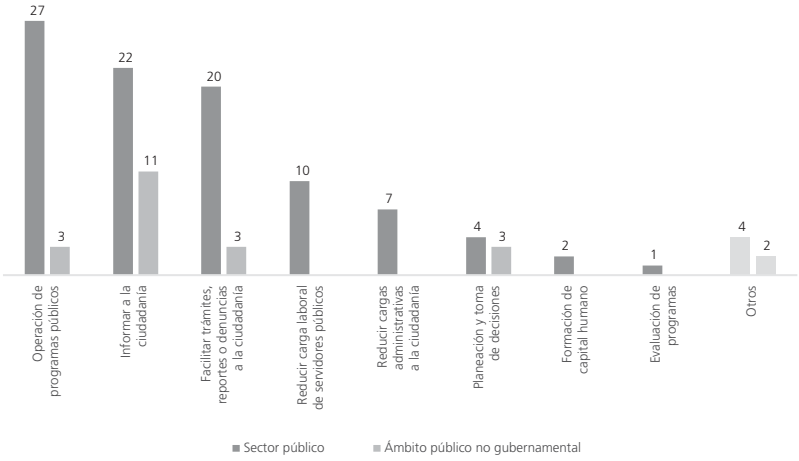
Fuente: Elaboración propia.

La tarea cognitiva de mayor uso en el sector público es la consulta de información, lo que se repite en el ámbito público no gubernamental. En el primer caso, el 33% de los sistemas identificados realizan esa tarea; en el segundo, la cifra representa el 50%. Los tipos de tareas difieren considerablemente entre un sector y otro. En el ámbito público las principales son, después de la mencionada consulta, el procesamiento de información (16.5%), gestión de trámites

(12.4%) y análisis de datos (10.3%). En el ámbito público no gubernamental son análisis de datos (13.6%), gestión de trámites (9.1%), detección de anomalías (9.1%) y diagnóstico de salud (9.1%).

Como se evidencia, en el ámbito público existe una notable diferencia del doble entre la tarea cognitiva identificada en el mayor número de sistemas y la segunda que más sistemas posee. Esta diferencia se acrecienta en el ámbito público no gubernamental donde es cinco veces superior. En gran medida, ello se relaciona con los tipos de sistemas que se utilizan (gráfico IV.4) y la función en este ámbito (gráfico IV.6).

Gráfico IV.6. Número de sistemas de IA por función en el ámbito público



Nota: n = 119.

Fuente: Elaboración propia.

Tanto en el ámbito público no gubernamental como en el sector público la tarea de menos aplicación es la de transcripción o traducción con cuatro sistemas, tres en el sector público y uno en el ámbito público no gubernamental. Es decir, 3.1% y 4.5% de las tareas identificadas que utilizan los sistemas de IA en cada sector, respectivamente.

El número de sistemas de IA por el tipo de función que desempeñan en el ámbito público se vincula con su actividad dentro de la administración pública (gráfico IV.6). En este sentido, se observa que la principal aplicación de estos sistemas es en la operación de programas públicos en entidades del sector público con el 27.8% del total de sistemas identificados en estas entidades, mientras que en las no gubernamentales la función más frecuente de los sistemas es informar a la ciudadanía (50%). Sobresale el hecho de que las entidades no gubernamentales carecen de sistemas que coadyuven a funciones típicas del sector público, como reducir carga laboral de servidores públicos o cargas administrativas de la ciudadanía.

La función que desempeñan los sistemas de IA en el ámbito público es, generalmente, su parte más conocida o externa, y en gran medida ella resume su valor dentro de la administración pública por el problema que coadyuva a resolver. Por tal causa este dato no debe analizarse exclusivamente en término estadísticos sin tener en cuenta el valor de la tecnología en la resolución de problemas públicos, este pudiera ser el caso, por ejemplo, de los sistemas de IA que, aunque no representan un número significativo, cumplen funciones de especial relevancia, como la localización e identificación de personas y la atención a poblaciones indígenas. En este sentido, destaca el proyecto "Regresa" de la Universidad Nacional Autónoma de México (UNAM) que conjunta las herramientas de la inteligencia artificial con los conocimientos y técnicas de la antropología física y social a fin de apoyar en la búsqueda, localización e identificación de menores desaparecidos en el país (toma de decisiones); y el sistema de IA "Mexin" de la UAT, que permite la traducción de texto en español a varias lenguas indígenas, a través del emparejamiento semántico y modelos de aprendizaje automático (informar a la ciudadanía).

El BID clasifica los sectores de desarrollo en diferentes categorías con la finalidad de enfocar los proyectos que impulsan el crecimiento económico y social de los países, y para identificar los principales desafíos que existen en un contexto y sector específico. Es importante considerar que esta tarea o función no coincide necesariamente con el sector al que pertenece la entidad que utiliza la IA. Cuando la fun-

ción del sistema de IA no corresponde con una función específica de la entidad pública, entonces se considera el sector al que esta última pertenece. Clasificar el sector con base en la función que realiza el sistema y no el sector de la entidad que la utiliza permite comprender con más precisión cuáles son los principales sectores estratégicos a los cuáles realmente aporta el sistema de IA según su función en el ámbito público.

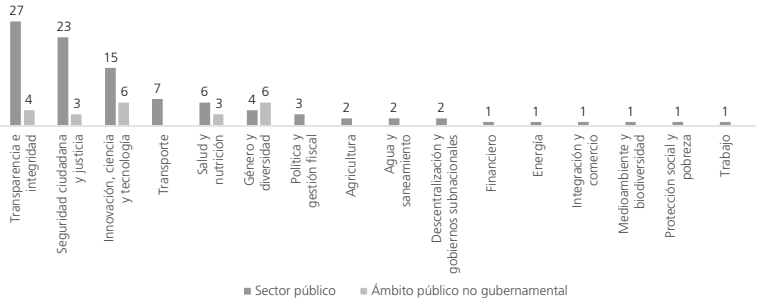
Como se observa en el gráfico IV.7, en las entidades del sector público, los sectores con más sistemas de IA son los de transparencia[5] e integridad y los de seguridad ciudadana y justicia, con 27 y 23 sistemas, respectivamente, que representan 27.8% y 23.7% del total identificado en ese ámbito. Esto no coincide con las entidades no gubernamentales del sector público, donde el uso de sistemas de IA se identifica mayormente en el género y diversidad e innovación ciencia y tecnología, con seis sistemas cada uno, o 27.3%, respectivamente.

El sector de innovación, ciencia y tecnología sobresale en las entidades del sector público con 23 sistemas que son el 23.7% del total identificado. Es importante precisar que, para efectos de este trabajo, dicha clasificación incluye a los sistemas de IA que se aplican en el ámbito de la educación y no exclusivamente los que inciden en temas relacionados con actividades de innovación. En las entidades públicas, merece mención especial el uso cada vez más frecuente de sistemas de IA en el sector de género y diversidad, aun cuando el ámbito no gubernamental supera en estos usos al sector público. Un ejemplo es el sistema Vera, del municipio de Veracruz, que es uno de los canales de comunicación que el Instituto Municipal de las Mujeres del

[5] Esta categorización de transparencia no debe entenderse como equivalente del concepto de transparencia algorítmica utilizado en este texto. Para el BID este criterio está determinado por el incremento del acceso a la información y la facilitación del control y la demanda ciudadana. El uso apropiado de tecnologías digitales, incluidos los sistemas de IA, permite mayor transparencia e integridad, no solo a nivel sectorial sino en sentido general en toda la administración pública, ello abarca, además, mejores decisiones por parte de los ciudadanos y actividades de control social (BID, 2020).

Ayuntamiento de Veracruz pone a disposición de la población femenina para coordinar esfuerzos en prevenir, atender, sancionar y erradicar la violencia de género.

Gráfico IV.7. Número de sistemas por sector de acuerdo con la clasificación del BID



Nota: n = 119.

Fuente: Elaboración propia.

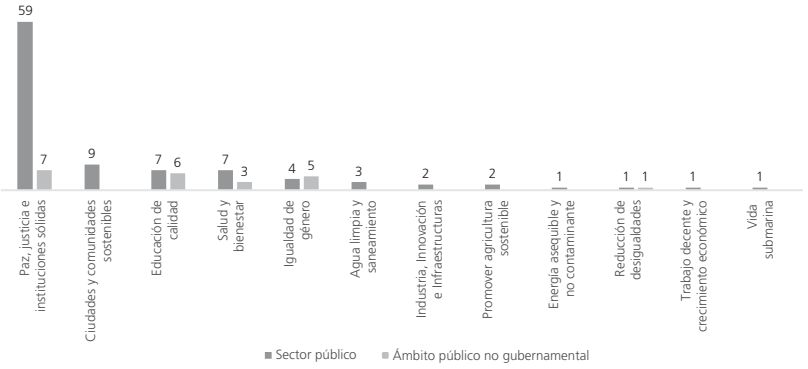
Entre los sectores con menor número de sistemas están los de trabajo, energía, descentralización, medioambiente y biodiversidad, integración y comercio, así como protección social y pobreza, con uno en cada caso. En estos sectores surge un área de oportunidad para la implementación de sistemas de IA. Se debe destacar el "Algoritmo en programa sociales", que actualmente utiliza la secretaria ejecutiva del Sistema Anticorrupción, dado que permite conocer e identificar de forma aproximada los criterios con los que se valora el desempeño real de un programa social.

Uno de los mayores retos de la implementación de sistemas de IA es su utilización efectiva para contribuir al bien común y a la prosperidad de las personas. Por ello y con ese fin se ha establecido un parámetro, determinado por la contribución de dichos sistemas al cumplimiento de los ODS de la Organización de las Naciones Unidas (ONU), objetivos globales trazados por consenso de la comunidad internacional para erradicar la pobreza, proteger el planeta y asegurar

la prosperidad para todas las personas. Aunque es cierto que no todos los sistemas de IA en el ámbito público contribuyen directamente al cumplimiento de tales objetivos y/o metas por la función que desempeñan dentro de la entidad pública o la no gubernamental.

Con base en la anterior precisión, en el gráfico IV.8 se constata que el mayor número de sistemas (59 o 60.8% del total) utilizados en el sector público responden al objetivo 16 de los ODS, esto es, "Paz, justicia e instituciones sólidas", lo que incluye las metas de instituciones eficaces y transparentes que rindan cuentas, y la adopción de decisiones inclusivas, participativas y representativas, entre otras. Del ámbito público llama la atención la diferencia entre el mayor número de sistemas y su contribución a los ODS y el segundo mayor número de sistemas, así como el resto que contribuye a los ODS, lo que tal vez evidencia un impacto limitado de las contribuciones de la tecnología a la resolución de problemas sociales sustantivos. Estas diferencias no aparecen en los sistemas del ámbito público no gubernamentales, donde la contribución a los ODS es similar en cuanto a cantidades de sistemas y su contribución a la paz, justicia e instituciones sólidas (7), educación de calidad (7), igualdad de género (5) y salud y bienestar (3).

Gráfico IV.8. Número de sistemas según su contribución a los ODS



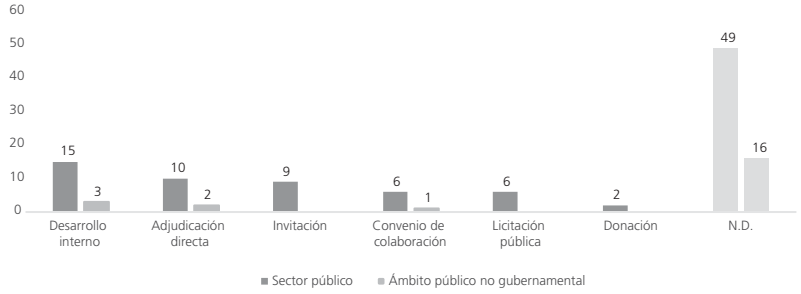
Nota: n = 119.

Fuente: Elaboración propia.

La contribución de los sistemas de IA al cumplimiento de los ODS se relaciona estrechamente con los tipos de funciones que realiza el sistema de IA dentro de la entidad pública, y con clasificación de los sistemas de IA conforme al sector del BID al cual se aplican.

El gráfico IV.9 presenta el número de sistemas según la forma de adquisición de la tecnología. Conforme a la anterior "Ley de adquisiciones, arrendamientos y servicios del ámbito público", bajo la cual se realizaron los procedimientos de contratación para adquirir bienes y servicios en la administración pública, solo eran tres los mecanismos de contratación: licitación pública; invitación a cuando menos tres personas, o adjudicación directa (Congreso de la Unión, 2000). Sin embargo, no se puede afirmar que dicha adquisición siguiera exclusivamente a dichas figuras. Durante el estudio se pudo acreditar que existen diversas vías por las cuales una entidad pública puede adquirir la tecnología: desde los procedimientos referidos hasta desarrollos internos, donaciones y convenios de colaboración interinstitucional, entre otros.

Gráfico IV.9. Número de sistemas por esquema de adquisición



Nota: N = 119.

Fuente: Elaboración propia.

El acuerdo por el que se emiten las políticas y disposiciones para impulsar el uso y aprovechamiento de la informática, el gobierno di-

gital, las tecnologías de la información y comunicación, y la seguridad de la información en administración pública federal no señala tampoco disposiciones particulares sobre cómo adquirir tecnología de IA, aunque se pudieran aplicar algunas de sus disposiciones relacionadas con los aplicativos de cómputo (Coordinación de la Estrategia Digital Nacional, 2021).

La mayoría de los sistemas de IA documentados no declaran o no cuentan con información sobre cómo se adquiere la tecnología (65 casos o 55%); una distribución desigual entre los sistemas del sector público y los del ámbito público no gubernamental. Acerca de esto, no se tiene información sobre la forma de adquisición en el 50.5% de los sistemas del primero, y en el 72.7% de los del segundo.

De los sistemas de IA con información sobre la forma de adquisición, una cantidad significativa se ha desarrollado internamente: 18 sistemas (33.3%), que se desglosan en 15 del sector público (31.2%) y tres del no gubernamental (50%). Asimismo, las formas de adquisición varían entre el sector público y en el no gubernamental, mientras en el primero hay diversidad de mecanismos de adquisición, en el no gubernamental solo se identificaron tres: desarrollo interno (3), adjudicación directa (2) y convenio de colaboración (1).

Las formas de adquisición que corresponden a invitación, licitación pública y donación, solo se identificaron en el sector público. De hecho, apenas seis proyectos declaran haberse adquirido por licitación pública; son ejemplo el Sistema de CCTV del Poder Judicial del Estado de México, que utiliza reconocimiento facial para la detección de personas en sus edificios, y el sistema de administración biométrica de comparación de huellas en la Fiscalía General de Ciudad de México que permite llevar a cabo investigaciones criminales efectivas y precisas, para facilitar la colaboración y el intercambio de información con otras entidades.

Los hallazgos de este capítulo muestran un uso predominante de sistemas de IA en el poder ejecutivo y en los gobiernos estatales, especialmente las áreas de transparencia y seguridad. Sin embargo, persisten desafíos como la baja adopción en gobiernos municipales, la disparidad entre entidades federativas y la falta de información sobre

su adquisición. Estos resultados destacan la necesidad de fortalecer la gobernanza de la IA para garantizar una adopción más equitativa y transparente en el ámbito público.

V. Sistemas de IA en la práctica gubernamental

El propósito de este capítulo es presentar con detalle las actividades gubernamentales en las que se usa un sistema de IA para sustentar actividades rutinarias de la entidad pública. Para ello, primero se discute el estado del arte con respecto a la caracterización de las tareas que dicho sistema puede cumplir en el sector público, y en segundo presentamos casos ilustrativos para los distintos tipos de prácticas gubernamentales identificados.

La literatura especializada en sistemas de IA en el sector público sigue explorando los espacios de la función pública en donde podrían generar valor. Son diversas las perspectivas que analizan las actividades, funciones y tareas en las que esos sistemas apoyan el servicio público. Los estudios clásicos, por ejemplo, señalan la analítica de datos para la acción pública, que sugiere el uso de sistemas de IA para identificar patrones en grandes volúmenes de datos o la detección de eventos y anomalías en tiempo real o en cuasi tiempo real, en aras de lograr intervenciones de política pública precisas y oportunas (Wirtz *et al.*, 2021; Medaglia *et al.*, 2024; Criado *et al.*, 2024). Wirtz *et al.* (2021) destacan que estas funciones son el núcleo potencial de dichos sistemas, al tiempo que advierten de los ya muy conocidos riesgos en su uso: falta de transparencia y de explicabilidad, y posibles sesgos discriminatorios. La literatura también argumenta que los modelos pueden servir como insumos para "encuadrar" problemas antes de que escalen, lo que habilitaría acciones tempranas, p. ej., seguros paramétricos ante

sequías mediante satélites. Estas prácticas concretan la idea de políticas públicas anticipatorias: un modo de decisión que "escanea el horizonte" e integra sistemas de previsión, aprendizaje organizacional y retroalimentación para actuar oportunamente a eventos no deseados (Kimbell y Vesnić-Alujević, 2020; Maffei *et al.*, 2020).

Otros autores han señalado la función de los sistemas de IA como herramientas de supervisión, inspección y control de la función pública o de los comportamientos individuales que pudiesen incurrir en un problema social (Meijer *et al.,* 2021; Criado *et al.*, 2024; Mergel *et al.*, 2024). Criado *et al.* (2024) sostienen que estos usos habilitan nuevas capacidades de vigilancia y monitoreo en las entidades públicas, aunque las exponen a dilemas de legitimidad y valores públicos. Esto es, que los sistemas de IA transforman los mecanismos de fiscalización y control, y reconfiguran los dilemas normativos de la administración pública, al insertar prácticas de supervisión en un marco de tensiones entre eficiencia y legitimidad democrática (Nam y Bell, 2024).

Una tercera categoría de uso se relaciona con la habilitación de nuevos espacios de interacción con la ciudadanía y nuevas formas de prestación de servicios públicos para ella. Esta es quizá el área que registra más crecimiento, tanto en desarrollo como en uso de sistemas de IA en el sector público, pues está asociada con la explosión de la IA generativa en el sector público. Esta literatura ha evolucionado en dos corrientes de investigación diferenciadas. Por un lado, los estudios sobre asistentes conversacionales gubernamentales, que destacan cómo estas herramientas se han consolidado como la forma más extendida de mediación digital en los encuentros ciudadano–Estado. Investigaciones recientes muestran que los asistentes conversacionales permiten consultas de información, trámites básicos y, en versiones más avanzadas, la negociación de servicios personalizados. Su adopción se asocia a beneficios como disponibilidad 24/7, reducción de cargas administrativas y mejoras en la accesibilidad de la información, aunque también muestran limitaciones técnicas, riesgos de pérdida de empatía en el servicio y dilemas sobre la rendición de cuentas (Alishani *et al.*, 2025).

Por otro lado, existe un campo emergente que analiza el uso informal de sistemas de propósito general, como ChatGPT, por parte

de funcionarios públicos. Aquí la literatura subraya que estas herramientas no se emplean para interacción ciudadana directa, sino para el procesamiento interno de información y como apoyo en el trabajo administrativo. Estudios sobre burocracia de calle sugieren que la incorporación de GenAI transforma la identidad profesional y las dinámicas laborales, dando origen a tensiones entre la automatización y la ampliación de las capacidades humanas (Jung y Camarena, 2024). Evidencia empírica muestra que la dependencia de ChatGPT puede mejorar la autoeficacia y el esfuerzo laboral de servidores públicos, aunque a la vez abre debates sobre riesgos de desplazamiento o dependencia excesiva (Zhang *et al.,* 2025).

Con esta revisión de la literatura, se llevó a cabo un análisis inductivo de los sistemas mapeados y un análisis constante comparativo entre los casos identificados y las categorías conceptuales para identificar los tipos de tareas que los sistemas de IA están cumpliendo en las entidades públicas. Este concepto permite vincularlos con los procesos administrativos concretos en los que se insertan, indicando el nivel de sofisticación técnica de la herramienta, y la naturaleza de las actividades públicas que son objeto de automatización, apoyo o transformación. Como lo señala el gráfico IV.5, se identificó el uso de sistemas de IA para ocho tareas relevantes: consulta de información, procesamiento de información, gestión de trámites, reportes o denuncias, análisis de datos, identificación o localización de personas, estimación de riesgos, detección de anomalías, diagnóstico de salud, y transcripción o traducción.

La subsección "Análisis descriptivo del uso de los sistemas de IA" de este capítulo está dedicada a describir con detalle estas categorías analíticas y a presentar casos observados en el mapeo. Son seis los casos ilustrativos que ejemplifican con mayor profundidad algunas de estas tareas. Para documentarlos, se realizaron un análisis documental y entrevistas con los actores involucrados en el desarrollo del sistema de IA. Son casos que nos permiten conocer las características principales del sistema y su funcionamiento, el problema público o la generación de valor público que se busca atender con el sistema, y las lecciones aprendidas durante su desarrollo, despliegue y uso.

El primer caso (*consulta de información*) es "Vera" que, desarrollado por el Instituto Municipal de las Mujeres del gobierno municipal de Veracruz, brinda información confiable —salvaguardando el anonimato— a víctimas de violencia. Para exponer el *procesamiento de información*, se aborda el caso del "Modelo de Prospectiva Económica", generado por el Centro de Investigación y Docencia Económicas (CIDE) que tiene como usuarios principales a la Secretaría de Economía y las embajadas de México. Este es un sistema dedicado a fortalecer las actividades de investigación de mercados e inteligencia de negocios para la sustitución de importaciones. Los sistemas de *estimación de riesgos* se ilustran con el "Sistema de inteligencia de datos para la inspección laboral" (SIDIL), que es utilizado por inspectores de la Secretaría de Trabajo y Previsión Social para clasificar y priorizar los centros de trabajo sujetos a inspección de cumplimiento de normatividad laboral. El cuarto caso (*análisis de datos*) es Siempre Seguras, un sistema desarrollado por investigadoras de la Universidad Autónoma de Zacatecas (UAZ) para monitorear incidentes de acoso sexual callejero. Las tareas de *identificación o localización de personas* se ejemplifican con el sistema "Regresa", creado por la UNAM y cuyo objetivo es generar imágenes de rostros de menores desaparecidos para apoyar las labores de búsqueda. La categoría de diagnóstico de salud se explica con la "IA para bebés", proyecto desarrollado por investigadores de la UNAM, la Universidad Autónoma Metropolitana (UAM) y el Instituto Nacional de Perinatología (INP). Esta tecnología se utiliza para identificar oportunamente los riesgos de salud en fetos. En el cuadro V.1 se reseñan los tipos de tareas identificados y su correspondiente caso ilustrativo.

Con base en esta tipología de usos, este capítulo muestra cómo se despliega en el sector público esta diversidad de tareas —que va de la consulta de información, a la transcripción o traducción—. Por tanto, la clasificación del cuadro V.1 ofrece un marco para entender cómo la práctica gubernamental traduce el potencial tecnológico de los sistemas de IA en funciones específicas del ámbito público. En la siguiente subsección se muestra un panorama de las tareas más frecuentes que cumplen los sistemas de IA y sus implicaciones. Mientras que en la subsección "Casos ilustrativos" se exami-

nan seis ejemplos que permiten observar con mayor detalle cómo los sistemas de IA apoyan en dichas tareas.

Cuadro V.1. Tipología de usos de los sistemas de IA

Tarea	*Descripción*	*Caso ilustrativo*
Consulta de información	Responder preguntas frecuentes y facilitar a la ciudadanía el acceso rápido y simplificado a información de gobierno.	Vera
Procesamiento de información	Ordenar, clasificar y transformar grandes volúmenes de datos administrativos en insumos útiles para la gestión y toma de decisiones.	Modelo de Prospectiva Económica
Gestión de trámites, reportes o denuncias	Agilizar la recepción, registro y seguimiento de solicitudes ciudadanas.	Sin caso ilustrativo
Análisis de datos	Identificación de patrones, tendencias o correlaciones con el fin de proveer información estadística descriptiva.	Siempre Seguras*
Identificación o localización de personas	Verificar identidades o ubicar individuos en contextos de seguridad, servicios públicos o gestión administrativa.	Regresa*
Estimación de riesgos	Anticipar la posibilidad de ocurrencia de eventos adversos, con el fin de apoyar la prevención y la gestión proactiva de problemas públicos.	SIDIL
Detección de anomalías	Identificar irregularidades o desviaciones inusuales en datos; orientada a fortalecer el control, la supervisión y vigilancia.	Sin caso ilustrativo
Diagnóstico de salud	Análisis de información clínica o epidemiológica para apoyar la detección temprana de enfermedades u optimizar la atención médica.	IA para bebés*
Transcripción o traducción	Conversión de audios en texto o traducción automática entre idiomas, con el fin de eficientizar la gestión administrativa o mejorar la accesibilidad e inclusión de personas a servicios públicos.	Sin caso ilustrativo

Notas: a) Los sistemas marcados con asterisco pertenecen al ámbito público no gubernamental. b) Las categorías están ordenadas de mayor a menor según la frecuencia reportada en el gráfico IV.5.

Fuente: Elaboración propia.

Análisis descriptivo del uso de los sistemas de IA

Para caracterizar los distintos usos de los sistemas de IA en el sector público, en este apartado se explican las categorías analíticas del cuadro V.1 y los usos identificados en el mapeo.

Consulta de información. Los sistemas de IA facilitan el acceso a datos, documentos y respuestas relevantes a los usuarios, tanto ciudadanos como funcionarios públicos. Estos sistemas cumplen funciones de intermediarios informativos entre entidades públicas y usuarios que, en

muchas ocasiones, pese a ser información pública es de difícil acceso. Sus objetivos principales son reducir las asimetrías de información entre usuarios y entidades públicas, la agilización de interacciones, la instrumentación de la transparencia proactiva, y la reducción de barreras para la consulta o interacción entre usuarios y entidades públicas.

Se observa en este caso que el uso de chatbots o asistentes conversacionales está fuertemente vinculado con dicha tarea. A lo largo del mapeo, quedó claro que el término *chatbot* se usa indistintamente para referirse a sistemas con o sin IA que simulan una conversación humana. No obstante, algunos de respuesta basados en reglas y no diseñados para aprender, razonar o adaptarse sobre al flujo de conversación, mismos que la literatura denomina *chatbots*[1] (Noonia *et al.*, 2024), no forman parte de los sistemas elegibles para ser reportados aquí. Por otro lado, se entiende como asistente conversacional al sistema que se basa en grandes modelos de lenguaje (*large language models*) y que, por lo tanto, cuenta con un amplio repositorio de conocimiento o datos para consulta y es capaz de aprender y adaptar sus respuestas en función de la interacción con el usuario (Noonia *et al.*, 2024). En consecuencia, el objetivo del estudio es identificar y seleccionar los sistemas de IA que cumplen o dicen cumplir con la definición de asistentes conversacionales. Aunque todos los casos identificados en el

[1] El uso del término *chatbot* ha sido incorporado en los ámbitos académico y social para referirse genéricamente a cualquier sistema informático diseñado para efectuar interacciones de conversación similares a las humanas (Connolly, 2025). En las últimas décadas, los sistemas conversacionales evolucionaron considerablemente, pasando de sistemas basados en reglas y sin aprendizaje a sistemas basados en grandes modelos de lenguaje con un amplio repositorio de conocimiento (o datos para consulta) y capaces aprender y adaptar sus respuestas en función de la interacción con el usuario. Esto ha derivado en la necesidad de distinguir los tipos de sistema conversacional en al menos tres tipos: chatbots, asistentes conversacionales y agentes virtuales (Noonia *et al.*, 2024). La literatura tiende a reservar el término *chatbot* para los sistemas basados en reglas y que no aprenden; a utilizar *asistentes conversacionales* para referirse a los sistemas basados en grandes modelos de lenguaje; en tanto que se está consolidando el uso de *agentes virtuales* para referirse a sistemas conversacionales más completos y complejos, con capacidades de autonomía y automatización avanzadas y que pueden integrar una secuencia de múltiples tareas (incluso múltiples agentes) (Schneider, 2025).

repositorio *fueron declarados* (por los sujetos obligados o por el equipo de investigación) como asistentes conversacionales, debe señalarse que incluso bajo esta etiqueta conviven sistemas de IA de distintos grados de sofisticación tecnológica.

Los sistemas de IA que comprenden esta categoría son, en su mayoría, instrumentos para apoyar la transparencia, el derecho al acceso a la información pública, así como para la atención de dudas de la ciudadanía sobre trámites y servicios públicos. Por ejemplo, el chatbot ATIC de Info CDMX agiliza la respuesta a solicitudes de información pública; en tanto que el municipio de Purísima del Rincón habilitó un asistente conversacional para proveer información a la ciudadanía sobre trámites y servicios del gobierno municipal; y el chatbot Constanza ofrece atención a los derechohabientes del Instituto del Fondo Nacional de la Vivienda para los Trabajadores para consultas sobre trámites y servicios relativos a la vivienda (cuadro V.2). En la subsección "Consulta de información: brindar información confiable a víctimas de violencia" se presenta el sistema Vera, un caso ilustrativo de esta categoría desarrollado por el Instituto Municipal de las Mujeres de Veracruz.

Cuadro V.2. Casos de consulta de información

Sistema de IA	*Descripción*	*Institución*
Chatbot ATIC	Permite a la ciudadanía enviar solicitudes de información pública con respuestas automatizadas y rápidas.	Info CDMX
Asistente virtual	Es un canal informativo 24/7 en el sitio web del gobierno municipal, que proporciona a la ciudadanía respuestas estandarizadas sobre trámites y servicios comunes (p. ej., cartilla militar o pago predial).	Municipio Purísima del Rincón
Chatbot del SAR	Ofrece un canal de interacción digital que responde a dudas frecuentes de los usuarios sobre trámites y normativas vinculadas a los fondos que oferta el SAR.	Comisión Nacional del Sistema de Ahorro para el Retiro
Chatbot Cavinai	Atiende las consultas más frecuentes dirigidas al Instituto, ofreciendo respuestas inmediatas a la ciudadanía.	Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos Personales
Chatbot Constanza	Ofrece atención continua 24/7 para responder consultas de derechohabientes, empleadores y público en general sobre trámites y servicios relacionados con la vivienda.	Instituto del Fondo Nacional de la Vivienda para los Trabajadores
MarcIA	Permite a los usuarios consultar gratuitamente expedientes de signos distintivos dentro del Sistema de Propiedad Industrial. Busca ampliar el acceso público a datos especializados y promover la transparencia en materia de propiedad intelectual.	Instituto Mexicano de la Propiedad Industrial

Cuadro V.2 (*Continuación*)

Sistema de IA	*Descripción*	*Institución*
Victoria	Facilitar el acceso directo y actualizado a la oferta gubernamental, mediante un canal de interacción que responde consultas ciudadanas sobre programas, trámites y acciones de la administración en curso.	Secretaría de Gobierno de la CDMX
Archi	Brinda a los usuarios información sobre conceptos, instructivos, trámites y herramientas de gestión sobre documentos del archivo del Tribunal.	Tribunal Electoral del Poder Judicial de la Federación

Fuente: Elaboración propia.

Procesamiento de información. Sistemas de IA cuya tarea principal es identificar, estructurar, sintetizar o transformar datos (principalmente de texto), lo que amplifica las capacidades humanas en el trabajo analítico y la toma de decisiones. A diferencia de la categoría anterior (que plantea sistemas de intermediación y acceso a información), el procesamiento de información plantea el apoyo en operaciones cognitivas necesarias para cumplir una tarea, por ejemplo, clasificar, resumir, extraer o redactar información para la gestión documental o procedimientos administrativos. Los propósitos del procesamiento de información son variados. Por ejemplo, Locatel utiliza un sistema de IA predictiva para clasificar y despachar reportes o denuncias ciudadanas presentados en la plataforma *311; el Poder Judicial del Estado de México utiliza un sistema para obtener, clasificar y sintetizar información de sentencias públicas, acuerdos o jurisprudencias y utiliza otro más para la redacción de demandas laborales y sentencias; la Comisión Nacional Bancaria y de Valores (CNBV) puso a disposición el sistema chatbot Aria como herramienta de analítica de datos para apoyar las actividades de supervisión y fiscalización de sus inspectores; y el Instituto Mexicano del Seguro Social dispuso el Asistente Médico Virtual para que los médicos puedan conocer los datos de sus pacientes (su historial clínico) de manera ágil (cuadro V.3). Más adelante se presenta el Modelo de Prospectiva Económica elaborado por investigadores del CIDE, que fue desarrollado para transformar grandes volúmenes de datos económicos, culturales y sectoriales en insumos estratégicos comprensibles para la toma de decisiones, y facilitar la inserción de pymes mexicanas en el comercio internacional.

Cuadro V.3. Casos de procesamiento de información

Sistema de IA	*Descripción*	*Institución*
Servicio *311	Clasifica y canaliza en tiempo real los reportes ciudadanos ante las autoridades competentes, lo que fortalece el punto de acceso del gobierno donde la ciudadanía registra incidencias.	Agencia Digital de Innovación Pública
Sistema de búsqueda de sentencias y jurisprudencia	Organiza y pone a disposición pública resoluciones, acuerdos y criterios judiciales mediante procesos de clasificación y tratamiento automatizado de documentos. Busca facilitar la gestión de información jurídica especializada.	Poder Judicial del Estado de México
Chatbot Aria	Brinda apoyo a los supervisores de la CNBV en la consulta de información vinculada con reportes regulatorios y actividades realizadas durante visitas de inspección a entidades financieras.	Comisión Nacional Bancaria y de Valores
Asistente Médico Virtual	Genera de forma ágil resúmenes del historial clínico de cada paciente, facilitando que los médicos dispongan de información precisa durante la consulta. Busca optimizar el acceso a datos clínicos relevantes y contribuir tanto a la eficiencia diagnóstica como a la calidad de la atención.	Instituto Mexicano del Seguro Social
Modelo de Prospectiva Económica*	Este sistema permite mapear y vincular actores clave en mercados internacionales con pequeñas y medianas empresas mexicanas. Su finalidad es proveer información macroeconómica y cultural, e identificar empresas y socios potenciales en países como Japón, China y Corea.	Centro de Investigación y Docencia Económicas

Nota: Los sistemas marcados con asterisco pertenecen al ámbito público no gubernamental.

Fuente: Elaboración propia.

Procesamiento de información. Se observó en esta categoría un predominio de herramientas orientadas a actividades jurídico-administrativas, particularmente concentradas en la gestión de documentos internos de la entidad pública. En su mayor parte, estas herramientas suponen usuarios especializados en el campo en cuestión (p. ej., salud o justicia). Asimismo, se advierte variación en las funciones cognitivas de los sistemas de IA, partiendo de actividades de *procesamiento descriptivo*, como identificación de textos o clasificación, hasta actividades de *procesamiento normativo* (o de interpretación de la información), como la elaboración de síntesis jurídicas o médicas. Para esta categoría se presenta como caso ilustrativo el Modelo de Prospectiva Económica desarrollado por el CIDE y utilizado por embajadas de México y la Secretaría de Economía. Esta herramienta se utiliza para la investigación de mercados (o de inteligencia de negocios) que, a partir de un diseño de ingeniería de *prompt* y una base de datos especializados en sustitución de importaciones, ofrece procesamientos

descriptivos (p. ej., identificación de empresas) y normativos (p. ej., elaboración de una estrategia de comercialización, o análisis legal para la importación de bienes y servicios). En este caso, el procesamiento de la información consiste en la gestión de información sectorial valiosa para vincular empresas mexicanas y extranjeras para el comercio internacional.

Gestión de trámites, reportes o denuncias. Se refiere a los sistemas de IA que habilitan a la ciudadanía para iniciar un trámite administrativo o procedimiento de denuncia, dotando a la interacción digital de las condiciones de un acto administrativo con potenciales consecuencias de gestión de servicios públicos e incluso jurídicas. A diferencia de la consulta de información —que también se basa en instrumentos de interacción con la ciudadanía—, estos sistemas abren canales para que el ciudadano active trámites administrativos que deben ser procesados y gestionados por las autoridades correspondientes.

Cuadro V.4. Casos de gestión de trámites, reportes o denuncias

Sistema de IA	*Descripción*	*Institución*
Chatbot TEO	Permite a la ciudadanía reportar casos de corrupción o abuso de autoridad cometidos por servidores públicos o elementos policiales.	Secretaría de Seguridad Pública, Puebla
Chatbot Movimex	Opera a través de WhatsApp y permite a los ciudadanos presentar denuncias, emitir quejas, tramitar licencias de conducir y reportar fallas de infraestructura como baches.	Gobierno del Estado de México
Inés	Herramienta donde la ciudadanía pueda reportar al Instituto posibles noticias falsas o inexactas relacionadas con el proceso electoral de 2024 y pueda recibir información verificada al respecto.	Instituto Nacional Electoral
Chatbot del Portal SFP	Es un canal que habilita interacciones digitales que activan procedimientos administrativos formales ante la administración pública. Con ello, busca facilitar la realización de trámites administrativos como solicitudes de baja en el ISSSTE, gestiones en el sistema Declaranet o consultas de semanas cotizadas.	Secretaría de la Función Pública
Chatbot AMIMgo	Permite a la ciudadanía levantar reportes relacionados con señalamientos viales, infraestructura ciclista, el programa MiBici y el funcionamiento de semáforos.	Agencia Metropolitana de Infraestructura para la Movilidad, Jalisco.
Güicho	Permite a la ciudadanía reportar fallas como baches o luminarias dañadas, canalizar esos reportes directamente a las dependencias responsables y recibir seguimiento del cumplimiento.	Gobierno del Estado de San Luis Potosí

Fuente: Elaboración propia.

Al analizar los casos, se observa el uso de sistemas de IA para ampliar el ámbito de acción de la ciudadanía, con el establecimiento de espacios para, por ejemplo, desde denunciar corrupción, presentar quejas, levantar reportes de fallas en el sistema de transporte público o en la infraestructura pública urbana, hasta la denuncia de presuntas noticias falsas. Se identificaron varios sistemas relacionados con la denuncia de casos de corrupción o abuso de autoridad, como el chatbot TEO de la Secretaría de Seguridad Pública de Puebla; otras funciones administrativas identificadas fueron la realización de trámites administrativos para los ciudadanos (chatbot Movimex), trámites administrativos para funcionarios públicos (chatbot del Portal SFP), reporte de noticias falsas (Inés) o de malfuncionamiento de infraestructura pública urbana (Güicho) (cuadro V.4). Asimismo, es notoria la preferencia por enmarcar estos sistemas de IA en canales de comunicación de mayor adopción y uso entre la ciudadanía, como WhatsApp, Facebook Messenger o formatos tradicionales de chat.

Estos casos ilustran la plasticidad de los sistemas de IA para adaptarlos a funciones administrativas heterogéneas, pero siempre convergentes en el propósito de mejorar los instrumentos de acción del ciudadano para ejercer sus derechos. Si bien este propósito representa un principio valioso de reforma administrativa o modernización de la función pública, no es clara la capacidad de estos casos de detonar, efectivamente, acciones administrativas y soluciones a las quejas y denuncias. Para ello, es fundamental conocer si la implementación de sistemas de IA (una modificación en el *front-end* de la interacción con la ciudadanía) fueron acompañados con cambios en los procedimientos administrativos o en la normatividad organizacional (procesos relacionados con el *back-end* de la interacción), lo cual es fundamental para dotar de validez legal a los reportes, quejas o denuncias dispuestos por estos innovadores canales de comunicación. La estrategia de mapeo del presente estudio no contempla tal profundidad de análisis, por lo que es un área de oportunidad para futuras investigaciones.

En otro orden de ideas, en la literatura se ha abordado ampliamente la función de análisis de riesgos, que se ha proyectado bajo

distintos términos (p. ej., sistemas de alerta temprana, detección de eventos o de anomalías y predicción de riesgos).[2] De manera general, esta categoría analítica se refiere a funciones preventivas o prospectivas de problemas públicos, así como de amenazas o fallas en los procesos administrativos o técnicos en la gestión pública. A diferencia de las categorías anteriores, los usuarios primarios de esta categoría no son la ciudadanía, el análisis de riesgos se circunscribe a los funcionarios públicos como usuarios previstos de los sistemas de IA, quienes los utilizan como instrumentos de apoyo para decisiones de supervisión, acción pública o gestión.

El análisis de riesgos puede tener distintas aristas, como la prevención o anticipación de eventos, la priorización de casos o la detección de irregularidades. Si bien todas comparten la noción de riesgos que los gobiernos buscan identificar y atender oportunamente, se considera que el análisis de riesgos debería distinguirse en dos funciones sustantivas: la estimación de riesgos y la detección de anomalías. La categoría *estimación de riesgos* describe sistemas que estiman una probabilidad de ocurrencia de un evento o asignan un puntaje de riesgo a las entidades analizadas, en función del análisis de los datos disponibles. Mientras que la categoría *detección de ano-*

[2] Esta categoría se ha discutido ampliamente desde diversos términos que refieren en general a la misma función. Así, hay autores que denominan *sistemas de alerta temprana* a modelos que procesan datos para anticipar eventos —a veces en tiempo real— y generar alertas oportunas (Wirtz *et al.*, 2019). El término *detección de eventos* —predominante en informática— describe sistemas con la capacidad de identificar, en grandes volúmenes de datos, la ocurrencia de hechos relevantes (Roehl, 2023). Otro término —también común en informática— es *detección de anomalías*, estrategia que enfatiza la desviación estadística y está orientada a estimar patrones atípicos, desviaciones respecto de un patrón o la probabilidad de ocurrencia de un evento en el marco de una distribución de probabilidad conocida o estimada (Tomaževič *et al.*, 2024). El término más utilizado en la literatura de administración pública es *predicción de riesgos* que no tiene una caracterización técnica, pero que se ha vinculado con la identificación de sujetos o casos "de riesgo" o la estimación de probabilidades de resultados adversos, lo cual permite priorizar intervenciones de política pública (Meijer *et al.*, 2021; Selten y Klievink, 2024; Alon-Barkat y Busuioc, 2022; Whitford y Whitford, 2024).

malías constituye un análisis de puntos de datos que se desvían de los patrones estadísticamente esperados (p. ej., valores predichos) o que representan datos atípicos —de acuerdo con ciertas distribuciones de probabilidad o modelos estadísticos—. En los siguientes párrafos se detallan ambas categorías.

Estimación de riesgos. Constituye el núcleo de las políticas públicas anticipativas (*anticipatory policymaking*), las cuales proponen la habilitación de sistemas de analítica predictiva para formular intervenciones preventivas, activar alertas oportunas y priorizar esfuerzos del gobierno (Kimbell y Vesnić-Alujević, 2020; Maffei *et al.*, 2020). Este tipo de sistemas se alimenta de fuentes de información heterogéneas y no convencionales (p. ej., sensores, satélites o registros administrativos) para una amplia gama de dominios de política pública. En general, se trata de sistemas capaces de interpretar patrones en los datos para proyectar escenarios o clasificar, de acuerdo a un puntaje estimado, casos de riesgo. En el mapeo, se observaron estimaciones de riesgo relacionadas con una expectativa de que el gobierno pueda "anticiparse" a futuras situaciones no deseables, lo que plantea un análisis o proyección en el plano temporal. Por ejemplo, entre los sistemas de IA orientados a la anticipación, se detectó uno que adelanta la reaparición de vehículos robados y otro más de alertas tempranas sobre el riesgo de deserción escolar (cuadro V.5). Por otro lado, se identificaron sistemas de clasificación de casos para determinar la prioridad de atención por parte del gobierno y donde el factor tiempo —o el interés por anticiparse a un problema— juega un papel menos relevante. Estos sistemas parecen estar más orientados a optimizar los recursos del gobierno al atender un problema público. Es el caso de la Secretaría del Trabajo y Previsión Social que utiliza un sistema que estima el riesgo de incumplimiento de la normatividad laboral en centros de trabajo, y que sirve de insumo para seleccionar aquellos que han de ser inspeccionados. (Más adelante se expone este caso en mayor profundidad.) Y también la Secretaría de Administración Tributaria utiliza un sistema que le permite identificar casos potenciales de redes de evasión fiscal (cuadro V.5).

Cuadro V.5. Casos de estimación de riesgos

Sistema de IA	*Descripción*	*Institución*
	Estimación de riesgos orientada a la anticipación	
Sistema de Detección de Vehículos Robados	Se emplean modelos predictivos basados en datos históricos y lecturas de puntos de monitoreo para anticipar el posible desplazamiento de unidades reportadas como robadas, emitiendo alertas inmediatas a través de canales digitales. Busca prever escenarios de reaparición y activar respuestas oportunas para la recuperación vehicular	Coordinación de Seguridad, Jalisco
Sistema de Actuación Temprana para la Permanencia Escolar	Utiliza técnicas de análisis predictivo sobre información académica, familiar y contextual para identificar estudiantes en riesgo de deserción. Genera alertas tempranas que permiten intervenciones preventivas y favorecen la continuidad educativa.	Secretaría de Educación de Guanajuato
	Estimación de riesgos orientada a la selección de casos	
Sistema de Inteligencia de Datos para la Inspección Laboral	Estima un puntaje de riesgo de incumplimiento en distintas submaterias de la normatividad laboral (p. ej., salarios mínimos, subcontratación, trabajo infantil), que permite a los inspectores ordenar los centros de trabajo en función de su probabilidad de incumplir. Busca focalizar la labor de inspección sobre los centros de trabajo con mayor puntaje de riesgo de incumplimiento.	Secretaría de Trabajo y Previsión Social
Sistema de IA para la fiscalización	Mediante analítica de datos, clasifica contribuyentes de riesgo, identifica redes de evasión fiscal y detecta inconsistencias en comprobantes digitales vinculadas con contrabando o empresas fachada. Busca optimizar la selección de casos a fiscalizar mediante la detección predictiva de conductas irregulares.	Servicio de Administración Tributaria

Fuente: Elaboración propia.

Mención aparte merecen los sistemas de IA dedicados al *diagnóstico en salud* que, si bien caen en la categoría de sistemas de alerta temprana, su contribución específica al sector salud amerita una descripción específica. Su función es generar inferencias clínicas aplicables a las personas, en general mediante clasificación sintomática o interpretación de imágenes médicas, con el propósito de orientar decisiones de salud. Pueden asistir en funciones de tamizaje, triaje o confirmación diagnóstica y su valor público radica en su capacidad de brindar atención oportuna a pacientes, desaturar los servicios de atención en salud y expandir la cobertura mediante la telemedicina o la atención multilingüe. Por ejemplo, el sistema Retina IA Jalisco realiza la detección temprana de retinopatía diabética con un tamizaje sistemático asistido con herramientas tecnológicas que habilitan la consulta por telemedicina. Asimismo, el sistema Ixtlilton, además de realizar diagnósticos clínicos, es capaz de comunicar los diagnósticos a los usuarios en el idioma tutunakú. Otros sistemas posibilitan diagnósticos basados en

otras fuentes de información, como imágenes (para tumores cerebrales), micrografías de gotas secas (para enfermedades cardiovasculares) o ultrasonido (para afecciones en fetos).

Cuadro V.6. Casos de diagnósticos de salud

Sistema de IA	*Descripción*	*Institución*
Diagnóstico de enfermedades*	Mediante el análisis de micrografías de gotas secas, el sistema identifica patrones asociados a la calidad de medicamentos, enfermedades cardiovasculares y agua contaminada con equipos de bajo costo. Busca detectar tempranamente condiciones de salud o ambientales que permitan activar intervenciones preventivas.	Centro de Investigación y Estudios Avanzados del Politécnico Nacional
Retina IA Jalisco	Mediante el análisis de imágenes de retina, detecta casos de retinopatía diabética en etapas tempranas con un proceso de validación por especialistas. Busca agilizar el tamizaje y la referencia de pacientes en riesgo, reduciendo los tiempos de diagnóstico y ampliando la cobertura de atención en un contexto de escasez de oftalmólogos en el sistema de salud.	Gobierno de Jalisco
Ixtlilton*	Diagnostica hasta 21 enfermedades a partir de la sintomatología reportada por el paciente, disponible en español y en tutunakú, lengua originaria de Papantla. Busca derivar recomendaciones sobre la necesidad de atención médica, reducir la saturación en centros hospitalarios y mejorar el acceso a cuidado en comunidades remotas.	Colegio Nacional de Educación Profesional Técnica
Diagnóstico de tumores cerebrales*	Por medio de la clasificación de imágenes médicas, identifica la presencia y tipo de tumor con alta precisión. Busca producir inferencias sobre la condición tumoral del paciente, lo que permite priorizar tratamiento o referencia, especialmente en situaciones donde la capacidad diagnóstica es escasa.	Universidad Tecnológica del Valle de Toluca
IA para bebés*	El sistema analiza imágenes de ultrasonido en el primer y segundo trimestre de gestación, lo que permite identificar señales tempranas de riesgo en el desarrollo fetal (p. ej., mediciones anómalas del fémur o la circunferencia craneal o abdominal). Mediante este análisis, el sistema genera un diagnóstico preliminar que facilita decisiones médicas oportunas y amplía la cobertura diagnóstica en zonas con acceso limitado a especialistas.	Universidad Nacional Autónoma de México

Nota: Los sistemas marcados con asterisco pertenecen al ámbito público no gubernamental.

Fuente: Elaboración propia.

En los sistemas dedicados a diagnósticos en salud, se observó que, salvo Retina IA Jalisco, todos los sistemas identificados forman parte del ámbito público no gubernamental (cuadro V.6). Incluso Retina IA Jalisco es una iniciativa conjunta entre el Banco Interamericano de Desarrollo, el Tec de Monterrey, campus Guadalajara, el Gobierno de Jalisco y la agencia de innovación C-Minds. Por lo tanto, se considera que el uso de sistemas de IA en diagnósticos en salud es un área de

oportunidad relevante para su adopción en las entidades públicas del sector salud. Para ejemplificar este caso, más adelante se presenta el sistema desarrollado por la UNAM, en conjunto con la UAM y el INP, que utiliza redes neuronales convolucionales para analizar imágenes de ultrasonido y detectar señales de alarma en fetos, buscando mejorar la atención materno infantil.

Detección de anomalías. Se refiere a los sistemas de IA orientados a identificar patrones en bases de datos, señalando casos que se desvían de la normalidad estadística o el comportamiento esperado. Estos sistemas podrían concebirse como una capacidad ampliada del análisis de datos, que posibilita la identificación de problemas emergentes o difícilmente observables mediante una inspección tradicional de los datos. En el mapeo se observaron sistemas utilizados para la detección de enfermedades en cultivos, fugas de agua, irregularidades financieras vinculadas al lavado de dinero u objetos inusuales en sistemas de movilidad, entre otros (cuadro V.7). El caso del sistema del servicio de detección de facturas falsas del Sistema de Administración Tributaria (SAT) ilustra este tipo de sistemas. En dicho caso, el sistema de IA combina las técnicas de análisis de redes en los sistemas de registro de facturación del SAT y modelos de aprendizaje de máquina para identificar contribuyentes con patrones de interacción que se desvían del comportamiento típico del universo fiscal. Esta aproximación permite señalar transacciones y actores atípicos con alta precisión y priorizar auditorías (Zumaya *et al.*, 2021).

Cuadro V.7. Casos de detección de anomalías

Sistema de IA	*Descripción*	*Institución*
Sensores de IA para detectar fugas de agua	Reconoce patrones sonoros anómalos en las tuberías para identificar filtraciones.	Comisión Estatal de Agua, Querétaro
Sistema de detección de irregularidades	Identificar patrones atípicos en las operaciones de entidades financieras, incluyendo conductas vinculadas con lavado de dinero.	Comisión Nacional de Banca y Valores
Sufragio seguro*	Detecta estados emocionales atípicos (p. ej., estrés o coacción) durante el proceso de emisión del voto, con el fin de identificar posibles intentos de coacción.	Tecnológico Nacional de México

Nota: Los sistemas marcados con asterisco pertenecen al ámbito público no gubernamental.

Fuente: Elaboración propia.

Análisis de datos. Comprende sistemas de IA que fortalecen la capacidad de análisis estadístico de sus usuarios, sin incidir directamente en la predicción de riesgos o la orientación (o recomendación) en la toma de decisiones. Estos sistemas fortalecen las actividades de recopilación, procesamiento y clasificación de datos para analizar patrones estadísticos de manera más ágil, accesible y clara para los usuarios. A diferencia de la estimación de riesgos o la detección de anomalías, no buscan anticipar eventos ni priorizar casos, sino producir evidencia descriptiva que visibilice fenómenos sociales, identifique patrones generales o caracterice condiciones de grupos poblacionales. Algunas de estas iniciativas sugieren la idea del uso de sistemas de IA como instrumentos para los *observatorios temáticos*, es decir, entidades dedicadas a monitorear el dinamismo de problemas públicos, sin detonar acciones públicas sobre individuos o casos concretos. En el mapeo, se ubicaron casos de monitoreo y clasificación descriptiva de temas de interés (p. ej., clasificación de precios, análisis dinámico del aforo vehicular en vialidades, estado de ánimo de las personas tuiteras) y análisis exploratorio de datos (p. ej., estadística descriptiva sobre el desempeño de programas sociales, análisis de patrones de violencia contra la mujer en espacios públicos o análisis de ausentismo o deserción escolar) (cuadro V.8). Más adelante se presenta con detalle el caso Siempre Seguras, desarrollado por la UAZ que analiza datos sobre incidentes de acoso sexual callejero en México para visibilizar el problema e informar para políticas públicas sobre violencia de género.

Cuadro V.8. Casos de análisis de datos

Sistema de IA	*Descripción*	*Institución*
Clasificador de precios	Permite determinar la categorización de precios en una canasta dada, con el fin de comparar niveles de precio entre productos genéricos y servicios esenciales. Permite organizar y clasificar precios para visibilizar diferencias y patrones.	Banco de México
Ánimo Tuitero	Recolecta tuits públicos y georreferenciados de Twitter en México para clasificar automáticamente su carga emotiva (positiva, negativa o neutra), para generar índices diarios de estado de ánimo y visualizar variaciones territoriales.	Instituto Nacional de Estadística y Geografía.
Algoritmo en programas sociales	Permite estimar los criterios que efectivamente influyen en el desempeño de programas sociales. Busca generar evidencia descriptiva sobre cómo los programas operan en la práctica para mejorar transparencia y monitoreo institucional.	Secretaría Ejecutiva del Sistema Anticorrupción

Cuadro V.8 (*Continuación*)

Siempre Seguras*	Mediante análisis de datos de redes sociales, el sistema clasifica y analiza reportes de acoso sexual callejero en México. Busca generar evidencia descriptiva sobre patrones de violencia de género en el espacio público.	Universidad Autónoma de Zacatecas
Mujer segura*	Mediante técnicas de aprendizaje de máquina, recopila y analiza datos sobre violencia contra la mujer, identificando patrones descriptivos de ocurrencia. Su finalidad es visibilizar y caracterizar el fenómeno mediante evidencia estadística.	Tecnológico Nacional de México

Nota: Los sistemas marcados con asterisco pertenecen al ámbito público no gubernamental.

Fuente: Elaboración propia.

Identificación o localización de personas. Se refiere a sistemas de IA dedicados a tareas de verificación, identificación y localización de personas en bases de datos. En estos modelos predominan los datos biométricos y las imágenes, por lo que el tipo de sistema de IA más recurrente en esta categoría es la visión por computadora. Se observaron cuatro funcionalidades distintas. Primero, la confirmación de que una persona es quien dice ser, lo que se utiliza como medio de acreditación para acceder a un servicio o validar una transacción. Por ejemplo, el Instituto Mexicano del Seguro Social usa un sistema de verificación facial para evitar fraudes por suplantación de identidad. Segundo, la búsqueda de una persona en bases de datos generadas en tiempo real. Por ejemplo, el Poder Judicial del Estado de México implementó sistemas de CCTV con reconocimiento facial en edificios públicos para identificar posibles coincidencias con la identidad de personas en listas (negras o blancas). Tercero, como herramienta de localización o apoyo forense, donde se analizan bases de datos de imágenes para buscar o reconstruir trayectorias de personas desaparecidas. Por ejemplo, la UNAM combina un sistema de IA con técnicas de antropología física y social para generar imágenes de progresión de edad de menores desaparecidos, apoyando la labor de búsqueda y contribuyendo a la investigación en antropología forense. Más adelante se expone este caso con mayor profundidad. Cuarto, los sistemas se utilizan como herramientas de analítica de presencia y agrupación de multitudes (*crowd analytics*), es decir, se detectan personas o aglomeraciones como eventos de riesgo (p. ej., picos de aforo o concentraciones inusuales).

Por ejemplo, el Sistema Estatal de Seguridad Pública de Nayarit implementa un sistema de IA para detectar la concentración inusual (y de riesgo) de personas en algunos espacios públicos (cuadro V.9).

Cuadro V.9. Casos de identificación o localización de personas

Sistema de IA	*Descripción*	*Institución*
Verificador de identidad	Integra tecnología de reconocimiento facial para autenticar a los usuarios a partir de sus rasgos biométricos únicos. Su finalidad es la verificación de identidad con el propósito de prevenir fraudes y suplantaciones en el acceso a servicios de seguridad social.	Instituto Mexicano del Seguro Social
Sistema de CCTV	Incorpora algoritmos de reconocimiento facial para identificar personas en tiempo real y cotejarlas con listas negras o blancas en los accesos a edificios públicos.	Poder Judicial del Estado de México
Regresa UNAM*	El sistema es capaz de realizar progresiones faciales automatizadas, generando imágenes predictivas de cómo evolucionaría el rostro de un menor a lo largo del tiempo. Se utiliza para para apoyar en la búsqueda e identificación de menores desaparecidos.	Universidad Nacional Autónoma de México
Detección de concentraciones inusuales	El sistema utiliza visión por computadora para identificar en tiempo real la formación de aglomeraciones de personas frente a cámaras. Su finalidad es proporcionar información oportuna sobre patrones de presencia colectiva que se desvían de lo esperado y que pueden representar riesgos de seguridad.	Sistema Estatal de Seguridad Pública, Nayarit

Nota: Los sistemas marcados con asterisco pertenecen al ámbito público no gubernamental.

Fuente: Elaboración propia.

Transcripción o traducción. Finalmente, se reseñan dos tareas relacionadas con el procesamiento de audio y texto. Primero, algunas entidades públicas utilizan sistemas de IA para la conversión de audio grabado (por ejemplo, de audiencias) a texto, lo que reduce la carga administrativa de los funcionarios públicos y, potencialmente, reduce los errores tipográficos en la generación de actas o registros oficiales. Otra función sustantiva puede ser la traducción automatizada, por ejemplo, la desarrollada por la UAT, que habilita la traducción de textos en español a múltiples lenguas indígenas de México. Por otro lado, el sistema OlivIA for Her es un sistema de IA para transcripción de audio que recopila y sistematiza información sensible para apoyar a mujeres víctimas de violencia mediante la conversión automática de audio a texto y la sistematización de sus relatos en expedientes digitales. Busca reducir la revictimización y agilizar el acceso a datos confiables para la atención institucional, al

mismo tiempo que fortalece la trazabilidad y el análisis de la información para la toma de decisiones. Por tanto, los sistemas de IA usados para transcripción o traducción pueden ser herramientas administrativas (p. ej., en la producción de registros administrativos) o instrumentos de inclusión y democratización del acceso a la información.

En síntesis, la caracterización realizada mostró que los sistemas de IA en el sector público cumplen funciones heterogéneas que van desde tareas rutinarias de consulta de información y procesamiento de datos hasta aplicaciones más complejas como la estimación de riesgos, el diagnóstico en salud o la identificación de personas. Estas categorías reflejan la amplitud de ámbitos en los que la IA está siendo utilizada y la diversidad de finalidades que persiguen las instituciones públicas al adoptarla. Al mismo tiempo, la clasificación pone de relieve que, si bien algunas aplicaciones buscan mejorar la eficiencia administrativa mediante la automatización de procesos, otras se orientan a fortalecer la toma de decisiones estratégicas o a ampliar la accesibilidad de los servicios. Este panorama general constituye una base para analizar en la siguiente subsección los casos y prácticas específicas que ilustran la forma en que estas tareas se despliegan en la gestión gubernamental.

Casos ilustrativos

En este apartado, se presentan seis casos ilustrativos que muestran cómo las categorías analíticas previamente descritas se materializan en experiencias concretas de adopción de IA en el sector público. Los ejemplos abarcan desde asistentes virtuales para la consulta de información hasta sistemas especializados en diagnóstico de salud, pasando por aplicaciones orientadas al procesamiento de información, la estimación de riesgos, el análisis de datos y la identificación de personas. En conjunto, nos dejan en claro la diversidad de propósitos y diseños tecnológicos que caracterizan la incorporación de la IA en la gestión gubernamental, así como los problemas públicos específicos que cada sistema busca atender.

Consulta de información: brindar información confiable a víctimas de violencia

El asistente virtual Vera es un sistema de IA desarrollado por el Instituto Municipal de las Mujeres de Veracruz (Immuver).[3] Este asistente conversacional, desplegado en la aplicación Telegram, está diseñado para ofrecer información confiable y accesible a mujeres víctimas de violencia, actuando como un primer canal de contacto para orientarlas sobre los servicios disponibles y resguardar su anonimato. De acuerdo con la tipología del capítulo, Vera cumple la tarea de consulta de información, pues facilita el acceso inmediato y seguro a datos relevantes en contextos de alta vulnerabilidad.

El sistema consiste en un asistente virtual diseñado para dar apoyo a las usuarias, de manera que ellas puedan acceder a los servicios que ofrece el instituto desde cualquier lugar y en cualquier horario, incluso el de fuera de oficina. Este asistente virtual se presenta a las usuarias en forma de un chat de la aplicación Telegram. De esta manera las personas interesadas podrán hacer el primer contacto con Vera mediante el escaneo de un código, lo cual permite que las usuarias pueden empezar a interactuar mediante el chat con Vera.

Vera actúa como un primer contacto para que las usuarias hagan uso de los servicios del Immuver. A través de sus capacidades conversacionales, este sistema puede identificar las necesidades de las usuarias y ofrecer una gran variedad de servicios informativos, como los requisitos para la atención integral de primer contacto, los horarios de atención del instituto o información de emergencia. También brinda información sobre infografías, cursos y talleres del instituto. Como resultado de este primer contacto, Vera recopila datos de las usuarias, quienes posteriormente pueden acudir presencialmente al Immuver para dar seguimiento a sus necesidades. Sumado a ello, el sistema está diseñado para salvaguardar cualquier dato personal que pueda poner en peligro a las usuarias.

---

[3] La información de este caso se obtuvo en entrevista con Blanca L. Aquino Santiago y Luis Gómez Wong, desarrolladores de este sistema de IA.

Este sistema de IA está construido en el lenguaje de programación JavaScript, y se basa en los servicios de los agentes conversacionales de Telegram. Así pues, Vera es el resultado de la combinación de un algoritmo de captación de datos, con las capacidades conversacionales de dicho agente. Anteriormente ya se había hecho el despliegue del sistema de IA, sin embargo, por cuestiones institucionales y diversas actualizaciones, se tuvo que reconstruir el código, desplegar un nuevo agente conversacional y, con ello, construir un nuevo código QR de acceso.

Motivación

La principal motivación para el desarrollo de Vera es aprovechar la potencia de las redes sociales y las aplicaciones móviles para hacer más accesibles los servicios y la información del Immuver. Al respecto una de las personas del equipo desarrollador mencionó: "Hay muchas mujeres que no utilizan radio, no les gusta o no tienen tiempo, pero sí usan redes. Notamos que hay muchas mujeres que sufren algún tipo de violencia, pero sí lo suben a redes". Además, para el equipo desarrollador es fundamental salvaguardar el anonimato de las usuarias, de modo que se han tomado las medidas para reservarlo.

Resultados hasta ahora

Los principales resultados son que las usuarias ya han utilizado Vera.

Lecciones aprendidas

"Todo ha sido fácil". La comunicación, la claridad en la idea del proyecto y la coordinación de las partes involucradas evitaron las trabas burocráticas para el desarrollo de Vera. Incluso pese a que esto se dio en dos administraciones distintas.

¿Qué se puede esperar en el futuro?

Actualmente, Vera se encuentra en fase de despliegue. Posteriormente se buscará implementar una campaña publicitaria que lo dé a conocer en diferentes espacios, como escuelas o festividades.

Procesamiento de información: inteligencia de negocios para la sustitución de importaciones

El Modelo de Prospectiva Económica es un sistema de IA desarrollado por el CIDE y utilizado por la Secretaría de Economía y las embajadas de México.[4] Esta herramienta de inteligencia de negocios emplea técnicas de IA generativa para organizar y sintetizar información de mercados internacionales, con el fin de apoyar a las pequeñas y medianas empresas en procesos de sustitución de importaciones y vinculación comercial con socios extranjeros. De acuerdo con la tipología del capítulo, este sistema cumple la tarea de procesamiento de información, al transformar grandes volúmenes de datos sectoriales en reportes estratégicos que orientan la toma de decisiones en política económica y comercio exterior.

El principal objetivo del Modelo es impulsar a las pequeñas y medianas empresas (pymes) mexicanas y así contribuir a crear una industria nacional fuerte. Para ello, el sistema se enfoca en facilitar la conexión entre las pymes y potenciales socios comerciales del extranjero, principalmente de Japón, China y Corea. En palabras del entrevistado, este sistema de IA pretende "Proveer de información estratégica y específica para que puedan [los dueños de una pyme] orientar y conectar con otras empresas en el mundo". Esto se pretende lograr a partir de la creación personalizada y automatizada de un reporte que permita a los interesados saber cómo y a quién acercarse para hacer crecer su negocio.

Para conseguir ese objetivo, el proyecto se basa en IA generativa. La cual se eligió como la herramienta adecuada por su capacidad para generar respuestas tomando en cuenta la información que se le proporciona como contexto. Así pues, uno de los primeros pasos del desarrollo fue la conformación del universo de información que serviría como contexto y de entrenamiento de los modelos utilizados en el proyecto. Para la integración de este contexto

[4] La información de este caso se obtuvo en una entrevista con Daniel Reyes Morales, investigador del CIDE y *senior data scientist* del sistema de IA.

se utilizaron técnicas de *retrieval augmented generation* (RAG), en conjunto con información concreta sobre los mercados, los principales actores, y las costumbres y maneras de comerciar de los países mencionados.

El conocimiento de las costumbres se considera esencial pues, como el entrevistado señalaba "comerciar en Asia no es lo mismo que comerciar en México", en referencia a que las diferencias culturales juegan un papel muy importante al concretar alianzas comerciales. Conformado el contexto inicial, se diseñó una estructura de tres agentes. Cada uno de estos se basa en el uso de una derivación de un LLM particular, y cumple una tarea en específico. Así pues, el primer agente se basa en el modelo Llama 70B, y tiene el objetivo de enriquecer el universo de información que se tomará como contexto.

Esta "investigación" se realiza por medio del análisis de diferentes páginas web y/o bases de datos certificadas por las empresas. Este primer agente está activo de manera continua, pues es necesario que la información se mantenga actualizada, y está diseñado para contrastar diferentes fuentes de información buscando evitar las noticias falsas (*fake news*).

El segundo agente se basa en un modelo derivado de Qwen y DeepSeek, y se encarga de sintetizar y evaluar toda la información del contexto recuperando aquella que pueda ser relevante para lo que el usuario solicite. Este segundo agente enriquece la información del contexto mediante procesos de razonamiento profundo. El tercer agente se basa en un modelo de DeepSeek, su tarea es sintetizar toda la información recuperada por el segundo agente y transformarla en un informe que contenga el plan de acción que se le presentará al usuario.

Se pretende que el acceso de los usuarios sea mediante un mensaje de Telegram, de manera que pueda chatear con el sistema de IA. Así, por ejemplo, el usuario puede decirle al asistente conversacional sobre qué es su negocio y preguntarle cómo podría orientarlo. El sistema le responderá con un informe detallado que contiene la lista de personas y empresas con las que se puede poner en contacto, además de un plan de acción, puntos que debe de investigar, información ma-

croeconómica relevante y cuestiones culturales que debe tomar en cuenta para establecer el contacto y las negociaciones.

En principio el proyecto se realizó en el lenguaje de programación Python, sin embargo, debido a la complejidad, se decidió utilizar JavaScript. Así, la estructura del sistema de IA tiene la suficiente flexibilidad de adaptarse para ofrecer otros tipos de servicios o incorporar nuevas bases de datos o nueva información. Finalmente, para mantener la seguridad de la información utilizada, el sistema se ejecuta de manera local.

Motivación

Una primera motivación para el desarrollo de este proyecto es ayudar a las pymes mexicanas a conectarse con empresas extranjeras para establecer una relación comercial sinérgica, de una forma accesible. En palabras del entrevistado: "Esto [el proyecto] tendría impacto nacional. Ayudaría a las pymes mexicanas a conectarse con el mundo".

La segunda motivación es, según el entrevistado, "Tener otra trinchera desde donde ayudar al país". Es decir, a menudo el producto de la investigación académica son las publicaciones, las cuales son muy importantes, pero también lo es desarrollar herramientas útiles para el Estado, es decir, herramientas prácticas que aterricen la investigación en algo concreto que apoye a las personas y así mover el país hacia algo mejor.

Resultados hasta ahora

— El proyecto se encuentra en fase de prueba, y se espera que pronto sea lanzado al público. En estas pruebas la herramienta ha generado reportes y estrategias con datos reales.
— Se contactado con diversos actores clave y tomadores de decisiones, los cuales mostraron interés en utilizar este tipo de herramientas para entender mejor los mercados.

Lecciones aprendidas

Los principales aprendizajes organizacionales de este proyecto son:

— "Para solicitar cosa tienes que pedírselo a uno que se lo pide al otro que se lo pide al otro…". A menudo, este tipo de desarrollos requieren el uso de infraestructura especializada. Sin embargo, los trámites burocráticos necesarios para su obtención suelen ser engorrosos y tardados, aun cuando exista una buena relación con los actores administrativos.
— "Si crees que lo puedes sacar, hazlo". La institución en general, y la división en particular, son propicias para el desarrollo de estos proyectos. Lo que ha permitido sortear algunas trabas burocráticas, confiando en que el proyecto llegará a buen término.
— "Nos ha favorecido el día a día". Las relaciones de los investigadores del proyecto, las de la división de estudios para el desarrollo o las del propio CIDE, han facilitado contactar con posibles actores de interés, lo que facilita la generación del proyecto.
— "Bajar [el conocimiento] del artículo de investigación a la *praxis* es lo que mueve al planeta". La manera en que se suele trabajar desde la academia conlleva a que existan pocas oportunidades para proyectos como este, a menos que se use el propio tiempo libre. Esto hace que los investigadores deban ver cómo encajar este tipo de proyectos con sus responsabilidades de investigadores, pero no siempre es posible. Por ello, si bien muchas veces estos proyectos podrían tener más impacto que la publicación de un artículo de investigación, los investigadores a menudo tienen las manos atadas, y se ven orillados a preferir las publicaciones.

¿Qué se puede esperar en el futuro?
Se pretende seguir con el desarrollo del sistema de IA, y los siguientes logros que se pretende alcanzar son:

a) Crecer en potencia e infraestructura con la adquisición de un clúster de múltiples computadoras y un servidor dedicado, para que el servicio opere todo el tiempo.
b) Seguir alimentando el sistema con otras fuentes de información, como nuevas bases de datos.

c) Establecer contacto con actores clave que puedan poner el desarrollo en un punto ventajoso y accesible. Se piensa principalmente en la Secretaría de Economía y la Secretaría de Relaciones Exteriores.

Estimación de riesgos: inspección de incumplimiento de normatividad laboral

El Sistema de Inteligencia de Datos para la Inspección Laboral (SIDIL) es una herramienta de IA desarrollada y utilizada por la Secretaría de Trabajo y Previsión Social (STPS).[5] Este sistema emplea técnicas de aprendizaje automático para procesar información de diversas fuentes y generar matrices de riesgo que estiman la probabilidad de incumplimiento de la normatividad laboral en los centros de trabajo. De acuerdo con la tipología del capítulo, el SIDIL corresponde a la tarea de estimación de riesgos, ya que permite priorizar inspecciones con base en indicadores de riesgo y focalizar los recursos de supervisión en los casos con mayor probabilidad de irregularidad.

Su objetivo principal es proporcionar información confiable y oportuna que ayude a optimizar las inspecciones que realiza la STPS y así promover el apego a la normatividad de las prácticas laborales. Esto se logra prediciendo el riesgo de incumplimiento de la normatividad por parte de los centros de trabajo, con lo cual se generan indicadores de riesgo que sirven de evidencia para decidir cuáles centros de trabajo deben ser inspeccionados.

Para conseguir esto, el SIDIL utiliza esencialmente la información relacionada con indicadores de condiciones laborales y registros de inspecciones laborales de años pasados. Es decir, que sus fuentes de in-

[5] La información de este caso se obtuvo de fuentes de acceso público, una entrevista con Martín Levenson, responsable del desarrollo, así como de evidencia documental proporcionada por él. Materiales consultados: *a*) American Institute for Research (AIR) (2025) sistema de Inteligencia de Datos para la Inspección Laboral (SIDIL); *b*) Hernández, Gerardo (2024), "STPS usará IA para definir inspecciones de trabajo" en *El Economista*.

sumos provienen del IMSS (patrones, empleados, accidentes), del INEGI y de la propia STPS (AIR, 2025). Puesto que los datos provienen de diversas fuentes, es necesario procesarlos para que tengan las características adecuadas y puedan ser utilizados.

A partir de tales datos y de la selección de atributos e indicadores laborales específicos, se utiliza la técnica *random forest*. Como resultado, el SIDIL forma una matriz de riesgos, en la cual el sistema asigna una predicción sobre la probabilidad de que un determinado centro de trabajo no cumpla con la normatividad en un ámbito específico. Es decir, para cada centro de trabajo se calcula un porcentaje de riesgo en función de cada submateria de las inspecciones, por ejemplo, en las submaterias de "salarios mínimos", "subcontratación" o "trabajo infantil" (AIR, 2025).

Esta matriz de riesgo es utilizada en conjunto con otras características que interesen a los inspectores para generar un "universo inteligente de inspección". Así pues, para elegir el centro de trabajo que se quiere inspeccionar, el usuario final no solamente aplica filtros anteriores, como, por ejemplo, el tamaño del centro laboral, sino que también la matriz de riesgo le permite filtrar por la probabilidad de que dicho centro de trabajo esté incumpliendo la normativa en el ámbito que sea de interés. Esto genera una lista ordenada en función del riesgo (universo inteligente de inspección), la cual sirve de base para que el tomador de decisiones puede enfocar los recursos de las inspecciones de manera óptima, pues ya no se depende del azar para seleccionar qué centros de trabajo serán supervisados.

Asimismo, se creó una interfaz que facilita el uso del SIDIL. De esta forma, para generar el "universo inteligente de inspección" los usuarios solo deben seleccionar los atributos, características de la empresa y la submateria que les interesa en ese momento.

Los beneficios del SIDIL se dan en dos ámbitos. Por un lado, aumenta la eficacia de las inspecciones, lo cual ayuda a optimizar los recursos con los que cuenta la institución. Por otro, el SIDIL incentiva a que los centros de trabajo acaten la normativa, pues se envía el mensaje de que los centros que no están en regla serán más fácilmente descubiertos y sancionados.

Motivación

La inspección del trabajo es la principal herramienta con la que cuentan las autoridades para monitorear el cumplimiento de la normatividad en materia laboral y confirmar que las condiciones en que los trabajadores desempeñan sus funciones sean las adecuadas. Sin embargo, en México existen millones de centros de trabajo, por lo cual, pretender inspeccionarlos todos cada año es inviable. Por ello, lo más razonable es que se elijan los que tengan una mayor probabilidad de incumplir con la normativa. Por tal motivo, se desarrolló una herramienta que, a partir de la ciencia de datos, permitiera dar insumos que optimicen las inspecciones (AIR, 2025).

Resultados hasta ahora

Hasta el momento se han identificado dos resultados. El primero es la mejora en los niveles de eficacia de los procesos de inspección (Hernández, 2024). Las pruebas realizadas muestran que las inspecciones derivadas del uso del SIDIL pueden identificar hasta tres veces más el número de incumplimientos. Este sistema, en conjunto con otros cambios en los modelos de inspecciones de la STPS, ha permitido que la efectividad de estas últimas pase de 55 a 94% en los últimos seis años (Hernández, 2024).

El segundo resultado se da en el ámbito administrativo, pues, con la llegada del SIDIL, se impartieron capacitaciones dentro de la organización para usar adecuadamente esa herramienta. Esto significó un aumento en el conocimiento y las capacidades institucionales de procesamiento de información. Sumado a ello, la llegada del SIDIL fomentó la coordinación entre distintas áreas de la STPS y otras instituciones externas, lo cual significa un cambio positivo en la cultura organizacional de la institución.

En resumen, el SIDIL ha permitido dar a conocer a los tomadores de decisiones de la STPS que "es posible desarrollar estrategias y políticas basadas en evidencia y datos para mejorar la eficiencia de la función pública". Algo relevante, pues el SIDIL es un ejemplo de los beneficios de la introducción de la IA en las labores de la administración gubernamental.

Lecciones aprendidas

Los principales aprendizajes organizacionales del desarrollo e implementación de SIDIL son:

— "El desarrollo e implementación de un sistema de IA es parecido a una carrera de obstáculos" debido a que cada paso es un reto, desde la diversidad de los datos disponibles y la calidad de estos, hasta los obstáculos organizacionales para compartir información o el desarrollo de las capacidades para implementar adecuadamente la herramienta.
— "Ningún estado de la república podría absorber la carga [de generar un modelo como el SIDIL] pero sí se le puede compartir la información, la voluntad existe". Para que este tipo de desarrollos sea posible se debe contar con infraestructura y capacidades institucionales especializadas, y un presupuesto considerable. Por ello, sería complicado que instituciones con menos recursos desarrollaran algo similar. Sin embargo, lo que sí es posible es compartir los resultados del modelo. Actualmente, la STPS colabora con ocho estados para generar "universos inteligentes de inspección" que les puedan ser útiles.
— "Es necesaria la voluntad de incorporar la herramienta ya en la operación". Implementar estas herramientas solo es posible si existe voluntad de las partes involucradas. A menudo, en las organizaciones hay resistencia para su adopción, lo que obstaculiza su desarrollo. Para superar este obstáculo es necesario que las autoridades de la organización la adopten y promuevan.
— "Las palabras clave son predicción de riesgo". La cuarta lección es identificar la potencial utilidad de los sistemas predictivos para mejorar la actividad de otras entidades públicas. La predicción de riesgo puede ser una herramienta muy útil para ámbitos como los riesgos ambientales o la emisión de permisos de construcción. La predicción de riesgos no solo puede usarse para la toma de decisiones, también es útil para monitorear o controlar actividades. Por ejemplo, tomando en cuenta un trámite, se pueden caracterizar a los interesados en función del riesgo de incumplimiento a la

normatividad, y con ello crear más cargas administrativas orientadas a los actores más riesgosos.

¿Qué se puede esperar en el futuro?
Los planes para el futuro del SIDIL se basan en tres pilares: 1) La incorporación de más fuentes de datos que permitan aumentar el poder de predicción del sistema. 2) La implementación y evaluación de operativos de inspección basados cien por ciento en el SIDIL; la expectativa es que sean al menos mil inspecciones. 3) Se espera que el uso del SIDIL se institucionalice y sea una práctica común en la planeación de las inspecciones.

Análisis de datos: observatorio de incidentes de acoso sexual callejero

Siempre Seguras es un sistema de IA desarrollado por investigadoras de la Universidad Autónoma de Zacatecas.[6] Este proyecto utiliza algoritmos de procesamiento de lenguaje natural para recolectar y analizar datos de redes sociales sobre incidentes de acoso sexual callejero, para visibilizar esta problemática y ofrecer evidencia útil para la formulación de políticas públicas sobre violencia de género. De acuerdo con la tipología del capítulo, este sistema de IA se ubica en la tarea de análisis de datos, al organizar y transformar información dispersa en patrones descriptivos que permiten comprender la frecuencia, localización y emociones asociadas a este fenómeno.

Siempre Seguras puede ser considerado como el primer observatorio que mapea los incidentes de acoso sexual callejero en México (Ortiz-Espinoza y Trujillo, 2024). En este sentido, su objetivo

[6] La información de este caso se obtuvo mediante fuentes de acceso público, y entrevistas con Ángeles Ortiz Espinoza y Angelina Espejel Trujillo, responsables del proyecto. Y además, Ortiz-Espinoza, Á., y Trujillo, A. E. (2024). "Siempre seguras": observatorio de acoso sexual callejero basado en inteligencia artificial. *Regiones y Desarrollo Sustentable*, *24*(45).

principal es "contribuir al diagnóstico del acoso sexual callejero" (Ortiz-Espinoza y Trujillo, 2024).

Este proyecto se llevó a cabo en dos fases. La primera consistió en la creación de un sistema de IA cuya función fuera recolectar datos de la plataforma X (antes Twitter) referentes al acoso sexual callejero, a partir de la app de la plataforma. Se recolectaron datos de entre junio y agosto de 2019 (Ortiz-Espinoza y Trujillo, 2024). Una vez descargados, se procesaban y filtraban a partir de palabras y correlaciones de palabras (n-gramas), tales como "acoso sexual callejero" (Ortiz-Espinoza y Trujillo, 2024). Después se usaron herramientas de procesamiento de lenguaje natural para identificar el sentimiento asociado a este tipo de incidentes, como el miedo o la tristeza.

Para facilitar el análisis de los datos obtenidos, se les visualizaba en forma de gráficas, nubes de palabras y mapas de incidencia (Ortiz-Espinoza y Trujillo, 2024). Esta primera fase se desarrolló con el lenguaje R, desde el cual se usaron paqueterías especializadas en técnicas de procesamiento de lenguaje natural y análisis de sentimientos. Una de las limitaciones de esta etapa es que arrojaba datos geográficamente imprecisos, pues a pesar de que se pensó para usarse en México, recogía reportes de otros países fronterizos.

Debido a las limitaciones, y para aprovechar la accesibilidad que permiten los dispositivos móviles y tener más cercanía con los usuarios, se decidió pasar al desarrollo de un sistema de IA para cualquier dispositivo con sistema operativo Android e iOS.

Desde este sistema, las personas usuarias tienen la posibilidad de reportar los incidentes de acoso callejero, como víctima o testigo, dando información básica sobre el tipo y el lugar, o más detallada como si se conoce al agresor. El sistema está diseñado para procurar el bienestar de la víctima, por lo cual no guarda sus datos personales. Además, los datos de la víctima se encriptan para que sea imposible identificar a quien hizo los reportes.

A diferencia de la primera fase, el sistema de IA identificaba de manera más exacta la ubicación geográfica de estos reportes, pues está vinculada al servicio de Google Maps. Asimismo, este sistema permite ver si han existido reportes en un área específica, lo cual ayuda a que

otras usuarias identifiquen su riesgo potencial. En la fase de desarrollo del Siempre Seguras existió el acercamiento del Instituto Municipal de las Mujeres de Querétaro, el cual otorgó un apoyo económico y solicitó un diagnóstico de la problemática. Sin embargo, por cuestiones administrativas, eso solo quedó en la entrega de un reporte a dicho instituto. Después ya no fue posible contar con su apoyo para escalar y difundir el sistema de IA.

Motivación

El desarrollo de Siempre Seguras se trató de una iniciativa ciudadana con fines de activismo. Fue impulsado por las motivaciones personales surgidas a partir de experiencias propias y de compañeras de trabajo, pues se cayó en cuenta de que este tipo de acoso podía ocurrir en cualquier lugar. Una de las encargadas del proyecto señalaba: "Hablando con compañeras [de trabajo] resulta que todas […] en algún momento tuvimos incidentes de acoso sexual cerca de la oficina."

Lo anterior dio pie a que las autoras emprendieran una búsqueda de datos que sirvieran para caracterizar al fenómeno, pero no encontraron los suficientes. Así, la segunda motivación fue la generación de datos científicos útiles para caracterizar el problema, visibilizarlo y obtener insumos para que los tomadores de decisiones diseñaran políticas públicas más eficientes (Ortíz-Espinoza y Trujillo, 2024). Al respecto una de las desarrolladoras refería: "Cuando nosotras iniciamos el proceso no había datos formales, no había estudios científicos al respecto aquí en México".

Así pues, el proyecto pretendía dar información sobre cómo había sido el incidente, dónde había ocurrido, sus causas, y ¿por qué estás ubicaciones y no otras? La idea era relacionar estos incidentes con la habitabilidad y el tipo de infraestructura pública ligada a la frecuencia de incidentes. Finalmente, el tercer motivo, es que, en ese momento, si bien no había este tipo de desarrollos en México, en otros países sí existían. Al respecto, una de las desarrolladoras comentó: "Encontramos información de otros observatorios que habían abordado el problema […] ¿Cómo es que en México no había algo así?".

Resultados hasta ahora

El proyecto está en la fase de consolidar un marco de colaboración con instituciones públicas. De su desarrollo se han identificado los siguientes resultados:

— Sentimientos relacionados al acoso sexual callejero. Los sentimientos predominantes que arrojó el análisis son de enojo, miedo, tristeza y disgusto (Ortíz-Espinoza y Trujillo, 2024)
— Obtención de datos que permitieran caracterizar el problema. La información recopilada dio detalles sobre los espacios y horarios en los que es más común la ocurrencia del acoso sexual callejero. Por ejemplo, uno de los lugares más susceptibles es el transporte público. El horario en el que se registran más incidentes es temprano en la mañana y las actividades más susceptibles ocurren mientras se hace ejercicio. La mayoría de los incidentes se registraron en la zona metropolitana de Querétaro (43%).

Lecciones aprendidas

Los principales aprendizajes organizacionales que deja el proyecto Siempre Seguras son:

— "Un candidato andaba proponiendo un proyecto curiosamente parecido". Posterior al desarrollo, otro presidente municipal propuso el lanzamiento de un sistema de IA muy parecido a Siempre Seguras, lo cual se interpretó como un intento de apropiación. Las autoras solicitaron aclaraciones, pero nunca hubo respuesta; ese otro sistema nunca vio la luz.
— "El problema público no existe hasta que alguien dice que existe". La visibilización de un problema es clave para las políticas públicas. Se necesitan actores relevantes para que el problema se reconozca como público y prioritario. Desarrollos como Siempre Seguras pueden contribuir a esto.
— "Tener una ONG es demasiado demandante. Se necesita tiempo, dinero, oportunidades". Se pretendía evolucionar el proyecto a una organización de la sociedad civil para tener acceso a otro tipo

de recursos, sin embargo, los que tenían las desarrolladoras fueron insuficientes para tal propósito.

— "Se preocupan más porque el reporte esté bien estructurado que por el impacto de la aplicación". La vinculación con el gobierno es complicada, pues el cumplimiento de la normativa y los procesos burocráticos provocan que el proyecto se pierda de vista.

¿Qué se puede esperar en el futuro?

Pese a los esfuerzos de las desarrolladoras para publicitar y visibilizar al proyecto, ante la falta de apoyo institucional y cuestiones personales fue imposible continuarlo. No obstante, se publicó un artículo con los resultados principales (Ortiz-Espinoza y Trujillo, 2024).

Identificación o localización de personas: imágenes predictivas de rostros de menores desaparecidos

Regresa UNAM es un sistema de IA desarrollado por investigadores del Instituto de Investigaciones Antropológicas de la UNAM.[7] Este proyecto aplica técnicas de aprendizaje profundo para generar imágenes predictivas que muestran cómo podrían evolucionar los rostros de menores desaparecidos, con el objetivo de apoyar a las familias buscadoras y a las autoridades en procesos de identificación. De acuerdo con la tipología del capítulo, este proyecto corresponde a la tarea de identificación o localización de personas, al utilizar algoritmos de visión por computadora para fines de reconstrucción forense y localización en contextos de desaparición.

Para lograrlo, al sistema se le entrena con información científica para que "aprenda" cómo se desarrolla cada una de las partes del rostro a lo largo del tiempo. En particular se trabaja el contorno, la textura, el envejecimiento, la pigmentación, la pérdida de elasticidad, etc. Es decir, el sistema debe aprender las dinámicas y los ritmos

[7] La información de este caso se obtuvo en fuentes de acceso público y en una entrevista con Ana Itzel Juárez Martín, encargada del proyecto.

diferenciados de las distintas partes del rostro. Dicho entrenamiento se hace con fotos de un mismo individuo en diferentes etapas de su vida hasta los 21 años, que es la edad en la que cesa el crecimiento.

Como insumo para dicho entrenamiento, en las primeras fases del desarrollo se utilizaron bancos de imágenes públicos que pueden usarse para fines científicos. Además, se usaron fotos de personajes públicos famosos de los cuales existieran imágenes disponibles de todas las etapas de sus vidas. Pero uno de los principales objetivos de Regresa UNAM es que sea aplicado a la población mexicana, y que dé cuenta del crecimiento y la diversidad de la población de México.

Para tal objetivo, era necesario contar con un banco de imágenes propio que reflejara las características particulares de los rostros de la población mexicana. Mediante diversas convocatorias abiertas al público y empleando un protocolo de toma de imágenes forenses y científicas, se logró su creación con alrededor de 1500 fotografías de 71 voluntarios. Este banco refleja las características de la población del centro de México.

Una vez obtenidas las fotografías, el segundo paso fue realizar una calificación somatoscópica de los distintos rasgos del rostro. Para ello, en una ficha y de manera manual, se capturan y codifican los distintos tipos. Estas fichas se utilizan como insumo para el entrenamiento del sistema de IA, y permiten dar cuenta de la evolución diferenciada de cada característica del rostro. Este no se desarrolla de manera uniforme.

El tercer paso fue procesar todos los insumos a partir de dos redes neuronales. Una preentrenada del tipo LSTM, la cual extrae las características del rostro (*landmarks)*, y da como resultado una malla conformada por 468 puntos. La segunda, utiliza procesos de *deep learning* para procesar la malla, aprender de ella, y generar un rostro sintético. Posteriormente, se lleva a cabo la progresión del rostro a lo largo de los años por medio de una superposición del rostro sintético en las fotos de la base de datos. Es así que se logra obtener imágenes predictivas del rostro evolucionado en el tiempo. El desarrollo utiliza principalmente el lenguaje de programación Python.

Para este proyecto se ha contado con el apoyo de distintas instituciones, en primer lugar, la UNAM y el Centro de Investigaciones Antropológicas. En 2022, a partir de una convocatoria, se contó con el apoyo de la empresa Huawei, y el de la Secretaría de Relaciones Exteriores. Asimismo, se ha tenido el acercamiento de distintas instituciones, en particular de distintas fiscalías. Pero se ha optado por mantener el proyecto en manos de la UNAM, esto con la finalidad de mantener el proyecto lo más abierto y libre posible; además de no perder "en términos éticos el acuerdo de garantía que tenemos con las y los voluntarios". Una vez que esté finalizado el sistema se pretende que su uso sea libre, vía solicitud, de tal manera que las familias buscadoras que lo soliciten solo tengan que proporcionar las imágenes de la persona desaparecida. Incluso se pretende tener un repositorio GitHub.

Motivación

La idea del proyecto surgió hace más de diez años, y la principal motivación se ha mantenido a lo largo del tiempo: atender la problemática de desaparición de menores y que las familias puedan estar juntas, pues, de acuerdo con la encargada del proyecto, "es un problema en el cual se tiene la urgencia de actuar". Posteriormente, se tomó conciencia de las posibilidades que traían las nuevas herramientas de IA para aportar elementos certeros de búsqueda.

Uno de los principales retos para atender la problemática es que los rostros de los menores cambian mucho y muy rápido, por lo cual fotos de hace dos o más años ya no son tan efectivas. El reto se agrava si se considera que los menores son una población vulnerable. Así pues, el sistema permitirá profundizar en el conocimiento de las condiciones sociodemográficas de las personas que sufren estos delitos y así ayudar a que las políticas públicas se basen en información científica.

Adicionalmente, se pretende que el proyecto sirva para apoyar a otro tipo de población, como la búsqueda de migrantes desaparecidos, ayudar a adultos que sospechen de haber sufrido sustracción de sus hogares cuando eran menores, etc. Finalmente, este sistema de IA también

puede ser utilizado con fines científicos, pues permitirá dar cuenta de la diversidad del país y de sus distintos momentos ontogenéticos.

Resultados hasta ahora

Por ahora el desarrollo sigue en la fase de entrenamiento del sistema, pero se han logrado los siguientes resultados:

— Conformación de un equipo de expertos para desarrollar el proyecto.
— Construcción y aprobación de los instrumentos a utilizar. Destacan los consentimientos informados para obtener el banco de imágenes, el cual se basa en valores bioéticos.
— La creación de un banco de imágenes propio que refleje las características de la población mexicana y que cumpla con estándares éticos y científicos.
— Creación de un piloto que en cierta medida puede dar cuenta de la diversidad del país en sus distintos momentos ontogenéticos.

Lecciones aprendidas

Los principales aprendizajes organizacionales que ha dejado el desarrollo de Regresa UNAM son:

— "Resolver con poco". En distintos momentos del proyecto no se contaron con los recursos óptimos para su desarrollo. Entonces se recurrió a la creatividad del equipo para seguir adelante con lo que se tenía a la mano.
— "Ha ocurrido la cosecha del esfuerzo voluntario que se ha hecho". El proyecto persigue una causa noble y legítima, lo cual ha atraído el esfuerzo voluntario de un gran número de actores, desde expertos que cooperan de manera extraordinaria con materiales e ideas, hasta voluntarios que aportan sus fotos e imágenes.
— "Una iniciativa que promueva las capacidades digitales del país debería contar con más recursos". Si bien se ha contado con apoyo para realizar el procesamiento computacional, también son necesarios recursos que permitan la recolección de datos empíricos.

¿Qué se puede esperar en el futuro?
Se pretende seguir con el desarrollo del Regresa UNAM, y lo siguiente que se pretende alcanzar es:

— Incluir más parámetros en el entrenamiento que ayuden a una mejor caracterización del problema y que permitan identificar las condiciones de riesgo relacionadas con la problemática. Se piensa especialmente en parámetros sociodemográficos.
— Aplicar el desarrollo con personas reales, y con ello identificar características y focos rojos regionales, así como sus causas.
— Que el sistema sea utilizado para proporcionar información científica sobre la caracterización del problema a los tomadores de decisiones.
— También se está pensando en formas de crear vinculación con otras instituciones.

Diagnóstico en salud: identificación de riesgos de salud en fetos

"IA para bebés" es un sistema desarrollado por investigadores de la UNAM, la UAM y el INP.[8] El proyecto utiliza técnicas de aprendizaje profundo aplicadas a ultrasonidos fetales para identificar señales de riesgo en el primer y segundo trimestre de gestación, con el propósito de facilitar diagnósticos tempranos y apoyar a médicos y a familias en la prevención de complicaciones durante el embarazo. De acuerdo con la tipología del capítulo, IA para bebés corresponde a la tarea de diagnóstico de salud, ya que convierte imágenes médicas en insumos clínicos que permiten la detección oportuna de problemas en el desarrollo fetal.

[8] La información de este caso se obtuvo en fuentes de acceso público y en una entrevista con Fernando Arámbula Cosío, responsable del desarrollo. Además, Cosío, F. A., Díaz, O. C., Pérez-González, J., Cristerna, R. V., Bañuelos, V. M., Marín, L. C., y Huerta, M. G. (2024). Deep Learning Applied to Automatic Fetometry. En *Congreso Nacional de Ingeniería Biomédica* (pp. 329-334). Cham: Springer.

El flujo del uso de este sistema en general es el siguiente: un médico desde cualquier parte del mundo realiza un ultrasonido del feto de una mujer que se encuentra en el primer o segundo trimestre de gestación. El médico envía el vídeo del ultrasonido para que el sistema de IA lo procese. A partir del análisis de las imágenes del fémur, de la circunferencia y diámetro de la cabeza y la circunferencia del abdomen del feto, el sistema determinará si existe una señal de alarma, a fin de que se tomen cartas en el asunto. Así pues, el sistema está pensado para 1) que cualquier médico lo pueda utilizar, 2) para que pueda ser utilizado desde cualquier clínica, 3) para optimizar los recursos, pues actualmente esta valoración debe ser realizada por un especialista.

Para lograr lo anterior, el sistema utiliza técnicas de *deep learning*, en especial el uso y entrenamiento de redes convolucionales (CNN) con una infraestructura tipo YOLO V2 *(You Only Look Once)* (Cosío *et al.*, 2024). Esto permite que se puedan procesar e identificar distintos elementos de imágenes. El primer paso es conseguir un banco de datos con imágenes para realizar el entrenamiento. Para ello, se ha utilizado un banco público de ultrasonidos de fetos en el primer y segundo trimestre de gestación, y se han recopilado imágenes similares de pacientes del INP.

El segundo paso es entrenar a las CNN para que, a partir de un video, se identifiquen las imágenes (*frames)* que tienen una calidad y posicionamiento óptimos que permitan y faciliten la identificación de características de riesgo. El tercer paso es el entrenamiento y la utilización de otra red que pueda identificar el rasgo que se desea y así predecir si es adecuado o no. Como se mencionó, se trabaja con tres rasgos, y para cada uno de ellos se hizo un procedimiento de entrenamiento distinto. Para el rasgo del abdomen, se usó un conjunto de 255 imágenes de entrenamiento, para la cabeza 375 y para el fémur 247 (Cosío *et al.*, 2024).

Finalmente, el cuarto paso será desarrollar una infraestructura para hacer utilizable el sistema, mismo que se está realizando con lenguaje Python, usando librerías especializadas con Tensonflow y PyTorch.

Motivación

La principal motivación del desarrollo del proyecto fue que, generalmente, este tipo de diagnósticos requiere la intervención de personal altamente especializado. Sin embargo, el acceso a estos especialistas se complica para ciertos grupos de personas, como, por ejemplo, las que viven en lugares remotos. Por tal causa es probable que las evaluaciones del desarrollo fetal para este tipo de mujeres no sea el adecuado, lo cual puede llevar a problemas de salud que se podrían evitar con un diagnóstico acertado. En este sentido, se pretende usar la potencia de la IA para hacer accesible, desde prácticamente cualquier parte, un análisis de fetometría que ayude a identificar y/o prevenir problemas en el desarrollo de los fetos.

Resultados hasta ahora

Se espera que el sistema empiece a ser usado dentro de tres años, sin embargo, se han realizado pruebas satisfactorias en bases de datos públicas, en las que ha sido capaz de identificar las tres características de interés de manera automática.

Lecciones aprendidas

Los principales aprendizajes organizacionales que dejó el desarrollo y la implementación del proyecto de IA para bebés son:

— "Se necesitan recursos". Para el desarrollo del proyecto han participado estudiantes de posgrado, tanto de maestría como de doctorado. Y si bien esto tiene la ventaja de que se aprovecha su conocimiento, la desventaja es que el sistema no es su prioridad. Por otro lado, el tiempo en el que ellos pueden cooperar es limitado, pues los programas de maestría duran dos años y al finalizar, los estudiantes se preocupan por ingresar al mundo laboral. Sería un gran apoyo contar con los recursos necesarios para poder contratarlos para que se dediquen de tiempo completo al desarrollo del sistema.
— "La visibilización del proyecto es clave". Una de las posibles formas en las que se espera obtener más recursos para el desarrollo

del proyecto, es dándolo a conocer a diversas instituciones que pudieran interesarse en él. Por ello, iniciativas como "Algoritmos CIDE" son importantes, pues dan visibilidad a este tipo de proyectos.

¿Qué se puede esperar en el futuro?

— Se trabajará en mejorar el rendimiento de identificación automatizada de las características de interés.
— Se realizarán pruebas con más modelos de IA, una continua optimización de los parámatelos y con datos de pacientes reales en diferentes lugares.
— Desarrollo de la infraestructura para compartir el sistema de IA.

VI. La gestión de los sistemas de IA en el servicio público: diagnóstico y recomendaciones

El propósito de este capítulo es plantear un diagnóstico de las condiciones del sector público para incorporar en su servicio los sistemas de IA, así como recomendaciones de política y gestión públicas para lograr dicho propósito. Reconociendo que los datos obtenidos mediante la estrategia de mapeo de sistemas de IA no son suficientes para un diagnóstico de este tipo, en esta parte del libro se incorporan datos y marcos de análisis complementarios,[1] lo que posibilita vislumbrar un mejor diagnóstico sobre el sector y elaborar recomendaciones desde una perspectiva de políticas públicas.

En los siguientes apartados se presenta el estado del arte de la incorporación de sistemas de IA en el servicio público desde distintos ángulos; de allí se desprenden recomendaciones de política pública.

La IA como instrumento del gobierno para mejorar su relación con la ciudadanía

A pesar de la diversidad de funciones potenciales de la IA en el sector público, nuestro mapeo muestra que la mayoría de los sistemas de

[1] El análisis expuesto en este capítulo se nutre, además del mapeo de sistemas de IA, de otras dos fuentes de información empírica, a saber: un conjunto de entrevistas semiestructuradas dirigidas a conocer los factores que moldean la perspectiva de los gestores públicos y actores relevantes respecto a la adopción de sistemas de IA; y se implementó una encuesta a los directores de Tecnologías de la Información y la Comunicación de agencias federales en México. Véase el apéndice D.

IA se orientan a informar a la ciudadanía: 33 casos (22 identificados en el sector público y 11 en el ámbito público no gubernamental, o 22.7 y 50%, respectivamente). En correspondencia con ello, el tipo de sistema de IA más utilizado es el asistente conversacional, con un total de 56 (44 del sector público y 12 en el ámbito público no gubernamental), es decir, el 47% de los sistemas identificados. Sin embargo, aunque existe correlación entre las funciones de estos sistemas y sus tipos, los números no coinciden de forma exacta, dado que un mismo tipo de sistema puede ser utilizado en otras funciones del ámbito público.

También es importante resaltar que, en correspondencia con lo anterior, muchas de estas aplicaciones se enfocan en mejorar la percepción ciudadana sobre las instituciones públicas. De hecho, desde la perspectiva de a cuáles sectores inciden, se constata que en su mayor parte son los de transparencia e integridad, y seguridad ciudadana y justicia para las entidades del sector público, con 27 y 23 sistemas, respectivamente, que representan el 27.8 y el 23.7% del total de sistemas identificados en este ámbito. Esta tendencia no se aprecia en las entidades no gubernamentales del sector público donde el uso se identifica mayormente en género y diversidad e innovación, ciencia y tecnología con un total de seis sistemas cada uno, que representan el 27.3%, respectivamente.

La tendencia orientada a informar a la ciudadanía contrasta con los patrones observados en otras regiones, donde la adopción de sistemas conversacionales es significativamente menor. Por ejemplo, en la Unión Europea, estos asistentes representan el 23% de los casos de uso de la IA, mientras que la función principal de los sistemas de IA es el monitoreo de la implementación de políticas con 37%, y la detección de problemáticas sociales 35% (Van Noordt y Misuraca, 2022). En América Latina, dichos asistentes constituyen el 20% de los casos de uso documentados (Muñoz-Cadena *et al.*, 2025).

La literatura atribuye estos patrones de elección tecnológica (es decir, los tipos de aplicaciones de IA que se prefieren o evitan) a limitaciones en la capacidad organizacional (Madan y Ashok, 2023; Neumann *et al.*, 2024; Selten y Klievink, 2024). Las organizaciones

con capacidades sólidas pueden concebir la IA como un medio para mejorar, profundizar o transformar sus prácticas de gestión de datos, involucrándose así en aplicaciones de IA más avanzadas. En contraste, las organizaciones que carecen de datos suficientes o de personal calificado tienden con mayor probabilidad a adoptar herramientas más simples, restringiendo en la práctica sus opciones a chatbots genéricos o automatizaciones básicas de procesos.

Esta perspectiva da una explicación plausible para la alta prevalencia de chatbots en el caso mexicano, donde las limitaciones de recursos e infraestructura son comunes en varios niveles de gobierno (Peeters *et al.*, 2023). Sin embargo, tal explicación corre el riesgo de simplificar en exceso las diversas capacidades y niveles de desempeño de las organizaciones públicas dentro de un mismo país al reducirlas a una sola categoría, es decir, a la baja capacidad de información. Además, una explicación basada solo en limitaciones de capacidad no aclara del todo por qué las decisiones tecnológicas tienden a enfatizar la participación ciudadana, la transparencia o el fortalecimiento institucional, por encima de otros propósitos.

Una visión complementaria se sustenta con el argumento de que las decisiones tecnológicas también están moldeadas por estrategias organizacionales orientadas a enfrentar desafíos sociales y políticos en curso. Una proposición clave en los estudios sobre tecnologías de la información en organizaciones es que las etapas iniciales de las que son emergentes (como la IA) se caracterizan por una flexibilidad interpretativa. Este concepto se refiere a la capacidad de una tecnología para ser comprendida y utilizada de múltiples maneras, a medida que distintos grupos atribuyen significados divergentes a sus funciones y utilidad (Doherty *et al.*, 2006). Esta proposición ofrece así una lente útil para examinar cómo las prioridades políticas, las tradiciones burocráticas y las lógicas institucionales influyen en la selección y uso de aplicaciones de IA en las organizaciones públicas.

Esta perspectiva también ofrece una explicación plausible para la alta prevalencia de chatbots en el caso mexicano, ya que es un país donde los gobiernos suelen enfrentar percepciones generalizadas de corrupción y una baja confianza pública en las instituciones

(Dussauge-Laguna, 2022). En tales contextos, el uso de la IA puede enmarcarse como una medida a favor de la transparencia o de mejora en la atención ciudadana, favoreciendo así el uso de chatbots o aplicaciones similares por encima de otras formas de IA. Aunque se trata de una explicación aún especulativa que requiere mayor validación empírica, varios estudios sugieren una narrativa predominante entre los funcionarios públicos mexicanos que conciben las tecnologías de la información y la IA como herramientas para fortalecer la transparencia, la rendición de cuentas y el gobierno abierto (Millan Vargas y Sandoval-Almazán, 2024; Pansera *et al*., 2023).

Este planteamiento invita a una reflexión más amplia sobre el papel de la IA en el contexto de las agendas de reforma administrativa orientadas a recuperar la confianza ciudadana en las instituciones públicas. Los gobiernos siguen enfrentando una creciente pérdida de confianza de los ciudadanos en las instituciones públicas (Dalton, 2008; Kettl, 2017). En sociedades cada vez más interconectadas e informadas, la exigencia social sobre el desempeño del gobierno no solo es creciente, sino compleja y diversificada. La incapacidad para adaptarse a entornos cambiantes es un problema que encuentra sus síntomas en la pérdida de legitimidad de los gobiernos, la caída de la confianza en el gobierno (Dalton, 2008), el resurgimiento de regímenes nacionalistas, autoritarios y populistas (Norris e Inglehart, 2018), el surgimiento de movimientos sociales disruptivos (Milkman, 2017) y en los reclamos por renovar las instituciones públicas. Estas circunstancias forjan una percepción generalizada de que las instituciones públicas están estancadas en paradigmas de gobierno o de gobernanza incapaces de generar soluciones efectivas para retos contemporáneos (Pollitt, 2016; Michel, 2024).

En respuesta a ello, la actual agenda de reforma del ámbito público internacional busca revertir esta tendencia mediante la transición hacia gobiernos más abiertos y colaborativos. Por ejemplo, el paradigma de la nueva gobernanza pública, que promueve la integración de diversos actores relevantes en la acción pública, donde la integración de la perspectiva de la ciudadanía es un insumo esencial (Osborne, 2006). Por otro lado, la agenda de gobierno abierto promueve

una noción de apertura que, en cierta medida, resulta de una combinación de principios diversos pero convergentes, como la innovación abierta, la participación ciudadana y la transparencia (Mergel, 2015). Por tanto, es en este contexto de innovación tecnológica y crisis de legitimidad donde los sistemas de IA emergen como instrumentos para mejorar su relación con la ciudadanía.

Las entrevistas semiestructuradas ofrecen evidencia que sugiere la plausibilidad de esta explicación, pues en ellas se identificó que los funcionarios públicos asocian el valor de la IA con su capacidad para reducir cargas administrativas, facilitar trámites y responder a demandas ciudadanas inmediatas, lo que desplaza la atención de usos internos o de monitoreo hacia aplicaciones visibles para el público. Así, mientras que las limitaciones financieras, de capital humano y de infraestructura generan fuertes barreras para proyectos de análisis complejo o de integración de datos, los entrevistados perciben a los asistentes conversacionales y a los sistemas dirigidos a transparencia y justicia como soluciones más viables y políticamente atractivas. Estas percepciones se traducen en una lógica de adopción "orientada a problemas", donde la IA se implementa principalmente en dominios de alta visibilidad social para reforzar la legitimidad y confianza en las entidades públicas.[2]

No obstante, es una ventana de oportunidad que se debe mirar con cautela, pues no es la primera vez que la administración pública se encuentra frente a tecnologías emergentes a las que se les atribuye el potencial de revolucionar la función pública. En su momento, el gobierno digital (Moon, 2002), la era de la gobernanza digital (Dunleavy y Margetts, 2006, 2010), la gobernanza adaptativa (Janssen y Van der Voort, 2016) y el gobierno como plataforma (Kim *et al.*, 2022) han suscitado, mediante marcos teóricos normativos, la esperanza de encontrar en las tecnologías nuevas soluciones a viejos problemas.

[2] Véase esquema de códigos en el apéndice D.

Todos conocemos esos agrios desencantos: ninguna de estas olas de reforma administrativa siquiera se ha acercado, en logros o resultados, a las expectativas que alguna vez causaron. Dichas experiencias sugerirían una aparente inocuidad de las tecnologías emergentes por catalizar transformaciones profundas en la administración pública, lo que tiene algo de cierto. Como se explica en el marco teórico de la construcción social de la tecnología (Orlikowski y Scott, 2008), acaso el marco teórico más reconocido y recomendado en el campo académico de gestión de sistemas de información. En lugar de una posición de *determinismo tecnológico*, donde se asume que la tecnología es capaz de generar cambios en los comportamientos individuales o interacciones sociales, este marco teórico sostiene que, por el contrario, el desarrollo tecnológico de un artefacto (p. ej., un sistema de IA) es el resultado de las tensiones, negociaciones y coaliciones entre individuos, grupos, estructuras organizacionales y otras tecnologías. En su ya clásico texto, Winner (2007) explica cómo los individuos tienen la posibilidad de inscribir, en el diseño de artefacto tecnológico, sus valores, sus intereses o formas preestablecidas de interacción o de relación de poder. De tal manera que los artefactos —como resultado de estos procesos sociales— pueden reproducir o agudizar dinámicas sociales preestablecidas, en lugar de modificarlas o revertirlas.

Con base en esta perspectiva, se esperaría que las aparentes expectativas puestas en los sistemas de IA como instrumentos para fortalecer la relación del gobierno con la ciudadanía no se materialicen. Por el contrario, es más probable que los sistemas de IA contribuyan a eficientizar procesos administrativos actuales (es decir, reforzar los patrones actuales al interior de una entidad pública) y menos probable que sean útiles para afrontar problemas estructurales, como la pérdida de confianza en las instituciones públicas. En todo caso, la ecuación es a la inversa: para su florecimiento y potenciación, los sistemas de IA precisan de instituciones públicas con recursos humanos e infraestructura sólidos, una flexibilidad y capacidad adaptativa ante los cambios en su entorno y una capacidad de generar datos de calidad para el monitoreo y la toma de decisiones, es decir, de instituciones públicas confiables.

Implicaciones de la brecha en alfabetización en IA en la adopción de sistemas de IA

Los datos obtenidos en el levantamiento revelaron una marcada brecha entre el conocimiento que tienen los servidores públicos sobre la IA y su comprensión real de la misma. Por ejemplo, el 62% de las respuestas a nuestras solicitudes de transparencia clasificaron erróneamente *software* convencional (p. ej., Kaspersky o Microsoft Office) como sistemas de IA. Además, en la encuesta aplicada a directores de Tecnología de la Información y la Comunicación, se observó que el acceso a información sobre casos de uso de IA fue el factor más influyente en la adopción de esta tecnología —hasta 2.6 veces más determinante que otros aspectos, como la calidad de los datos, el presupuesto o la infraestructura.[3]

Estas cifras pueden estar relacionadas con una tendencia más amplia observada en otros países: una alta exposición al discurso sobre IA acompañada de una comprensión limitada. A pesar de la creciente familiaridad con la IA en el debate público y su uso generalizado en la vida cotidiana de las personas, la evidencia empírica indica que una mayor interacción con la IA no conduce necesariamente a niveles más altos de alfabetización. Por ejemplo, aunque en 2023 el 90% de los estadounidenses declaró haber escuchado sobre la IA, solo alrededor del 30% pudo identificar con precisión sistemas de IA (Tyson y Kikuchi, 2023). De manera similar, una encuesta europea encontró que, aunque el 90% de los encuestados estaba al tanto de la IA, solo el 15% se sentía con la confianza suficiente para utilizarla en la vida cotidiana (*El País*, 2024). También ponen de manifiesto la brecha entre el conocimiento público —y, de alguna manera, genérico— sobre la IA y una verdadera alfabetización en IA, que se podría caracterizar como la distinción entre una comprensión "conceptual" amplia de su funcionamiento y capacidades, y la habilidad "procedimental" para emplearla eficazmente en el trabajo o en la vida diaria.

[3] Véase el apéndice D.

Cuando las personas carecen de alfabetización en IA (es decir, conocimiento sobre su mecánica y sus implicaciones), se vuelven más vulnerables a narrativas sensacionalistas en los medios o con carga política (Arnesen *et al.*, 2024; Li *et al.*, 2024). Esto, a su vez, puede derivar en confianza o escepticismo injustificados respecto al uso de la IA en los espacios de trabajo. En tales contextos, las personas suelen recurrir a asociaciones emocionales o intuitivas para evaluar los posibles beneficios o riesgos de la IA (Brauner *et al.*, 2023). Estos atajos cognitivos pueden alimentar respuestas públicas polarizadas, que van del entusiasmo acrítico por las promesas de la IA a los temores alarmistas sobre sus posibles consecuencias sociales catastróficas (Neri y Cozman, 2020).

La evidencia empírica obtenida mediante entrevistas semiestructuradas respalda el argumento de que la falta de alfabetización en IA expone a los servidores públicos a narrativas de temor y escepticismo que dificultan la adopción tecnológica.[4] Los testimonios de las personas entrevistadas revelaron tres dimensiones perceptuales que funcionan como barreras de adopción: miedo, escepticismo y preocupación por la privacidad de los datos. En primer lugar, el miedo surge de la incertidumbre sobre las consecuencias de la IA, expresándose en la ansiedad por la posible sustitución laboral o por verse obligados a abandonar rutinas consolidadas. Por ejemplo, un directivo en el área de tecnologías de la información señaló que existe "resistencia al cambio, falta de conocimiento técnico y miedo a ser reemplazados." En segundo lugar, el escepticismo se manifiesta en la percepción de que la IA carece de utilidad práctica o de valor a largo plazo; al respecto, algunos directivos sostuvieron que sus procesos "requieren intervención del personal" o que la tecnología "no tiene un rango de predictibilidad mayor a cinco o diez años". Finalmente, la preocupación por la privacidad de los datos refleja una carencia de marcos normativos claros, alimentando dudas éticas y legales sobre el uso de información sensible, como lo expresó un directivo en

[4] Véase el apéndice D.

el área de tecnologías de la información: "La inteligencia artificial puede generar preocupaciones éticas y de privacidad, especialmente en lo referente al uso de datos personales y la toma de decisiones automatizadas."

Algunos académicos enmarcan esta condición como un problema de información insuficiente, argumentando que mejorar la alfabetización en IA entre la población puede ayudar a enfrentar percepciones erróneas de riesgo (Arnesen *et al.*, 2024; Cantens, 2024; Zhang y Dafoe, 2019). Por ejemplo, Zhang y Dafoe (2019) proponen que la confianza y el involucramiento público con la IA pueden fortalecerse mediante diseños más transparentes y fáciles de usar, y a través de la difusión estratégica de información y de mecanismos de gobernanza participativa.

Sin embargo, las actitudes públicas hacia la IA no están determinadas únicamente por la transmisión de conocimientos técnicos o prácticos, sino también por las formas en que el discurso público enmarca cómo las personas piensan sobre la IA. La literatura sobre la percepción de riesgos tecnológicos enfatiza que las actitudes hacia las tecnologías emergentes se ven influidas por el nivel de conocimiento individual y por los valores culturales y las narrativas que circulan en la esfera pública (Renn y Benighaus, 2013). La cobertura mediática, los debates políticos y los comentarios de expertos median colectivamente si la imaginación pública está dominada por los posibles daños de la IA o por sus beneficios (Li *et al.*, 2024; Neri y Cozman, 2020; O'Shaughnessy *et al.*, 2023); esta es una explicación posible de la amplia variación de preocupaciones entre países. Por ejemplo, los australianos muestran inquietud por la vigilancia y el uso indebido de datos, mientras que en Reino Unido las conversaciones se centran en el desplazamiento laboral y en el sesgo algorítmico en programas de bienestar (Gerlich, 2024; Seth, 2024).

Por tanto, la alfabetización en IA es en particular relevante, considerando que esta tecnología es en especial sensible al encuadre que recibe en la esfera pública. La percepción social sobre la IA se mantiene ambivalente, pues es vista simultáneamente como beneficiosa y potencialmente dañina (Brauner *et al.*, 2023). Por ejemplo, Bozkurt

y Gursoy (2025) reportan que, en Turquía, el 63% de los encuestados consideró que la IA facilitaría la vida, mientras que el 52% expresó preocupación por el desempleo y la desigualdad social que podría generar. Esta ambivalencia se observa en la literatura académica, donde se ponderan tanto beneficios como riesgos en el uso de los sistemas de IA en el sector público (Wirtz *et al.*, 2020). Más aún, esta ambivalencia también queda de manifiesto entre los practicantes de política pública y el ámbito de toma de decisiones en el sector público, donde en los últimos años ha surgido un conjunto amplio de reportes institucionales con principios, recomendaciones o regulaciones que pretenden guiar la adopción efectiva y responsable de la IA en el sector público (World Bank, 2020; Santiso, 2024; Soler Garrido *et al.*, 2023).

Esta situación crea una considerable incertidumbre respecto a las implicaciones sociales e individuales de la IA, dejando a las personas altamente susceptibles a las señales externas provenientes de narrativas mediáticas, discursos políticos y comentarios de expertos. Además, la IA sigue siendo relativamente abstracta y poco familiar para muchos; aparte de herramientas ampliamente reconocidas como ChatGPT, la mayoría de las personas carece de experiencia directa con sistemas de IA, lo que limita su comprensión sustantiva de cómo funcionan estas tecnologías y su capacidad "procedimental" para utilizarlas. Esta falta de contacto directo incrementa la dependencia de fuentes externas para formarse una opinión sobre la IA (Brauner *et al.*, 2023; Renn y Benighaus, 2013).

Las narrativas de los expertos desempeñan un papel importante en la configuración de la percepción pública acerca de una tecnología que conlleva riesgos (Neri y Cozman, 2020). Al tener conocimiento o experiencia limitada en la interacción con la tecnología, las personas tienden a confiar en expertos para conocer e interpretar los beneficios y los riesgos de su uso (Renn y Benighaus, 2013). Por ello, los expertos actúan como transmisores principales que elaboran, codifican y difunden narrativas que resuenan con las inquietudes públicas sobre los riesgos existenciales asociados a la IA (Neri y Cozman, 2020). Por ejemplo, Neri y Cozman (2020) observan que

los tuits vinculados a discusiones relevantes de expertos y posteriores a 2014 —por ejemplo, las advertencias de figuras prominentes como Stephen Hawking o Elon Musk— generaron una preocupación pública significativa respecto a los riesgos de la IA, pese a la escasez de sucesos catastróficos tangibles. De manera similar, O'Shaughnessy *et al*. (2023) demuestran que los expertos que abogan por marcos de gobernanza de IA inclusivos y participativos influyen de forma indirecta en el escepticismo público hacia los sistemas de IA no regulados, al enfatizar sus implicaciones éticas y sociales.

Todas estas condiciones podrían traducirse en un reto de opacidad o accesibilidad al uso de estas herramientas en el servicio público, lo cual incide en la percepción ciudadana de transparencia de las decisiones apoyadas con sistemas de IA. Cuando los ciudadanos perciben que los sistemas de IA utilizados para decisiones gubernamentales no son transparentes ni explicables, la confianza en las instituciones públicas tiende a disminuir (Grimmelikhuijsen y Meijer, 2022). Por lo tanto, todas estas condiciones hacen que la adopción de la IA en las organizaciones públicas sea altamente contingente a cómo esta tecnología es enmarcada, narrada y comprendida en la esfera pública.

Con base en las observaciones empíricas y el planteamiento teórico arriba expuesto, aquí se argumenta que la combinación entre alto nivel de conocimiento de la IA, pero bajo nivel de alfabetización en IA, conllevan a una postura (de funcionarios públicos y ciudadanos) basada en la percepción de riesgo tecnológico que las personas tienen sobre la IA y que, en gran medida, esta percepción depende de la información que se permea hacia ellos por diversos medios. Las noticias, reportajes, debates públicos y opinión de expertos juegan un papel fundamental en la aceptación (o no) de la ciudadanía y los servidores públicos de la incorporación de la IA como instrumento de trabajo en la función pública.

No todas las sociedades comparten la misma percepción sobre la IA, puesto que la percepción pública sobre esta tecnología se moldea en el discurso público (Bozkurt y Gursoy, 2025). Neri y Cozman (2020) han demostrado que el discurso público en diferentes países se mueve en función del contraste entre los beneficios de la IA (p. ej.,

eficiencia y aumento de capacidades humanas) y las amenazas o resultados negativos que expusimos arriba. Por ejemplo, en Australia las preocupaciones de la sociedad respecto a la IA están relacionadas con la pérdida de privacidad, la vigilancia o el uso indebido de datos (Selwyn y Gallo Cordoba, 2022), mientras que en Reino Unido las personas manifiestan intranquilidad ante el desplazamiento laboral y los riesgos para la privacidad de datos (Gerlich, 2024).

Como lo han demostrado diversos estudios, cuando los ciudadanos o funcionarios públicos ponderan más riesgos que beneficios, estos no apoyarán su uso en el sector público, no asumirán como legítimas las decisiones o acciones públicas basadas en sistemas de IA o no las adoptarán para su organización (Schiff *et al.*, 2023; Wang *et al.*, 2023; Grimmelikhuijsen, 2023; Sidhu *et al.*, 2024; Whitford y Whitford, 2024). Por tanto, un área crítica en la adopción de sistemas de IA se enmarca en la dinámica de las discusiones en la esfera pública, y en los mecanismos de diseminación de información o alfabetización en IA.

El proceso de adopción, en una escala macro, pasa por un proceso de generación y difusión de información por parte de expertos de los riesgos y beneficios de la IA que permiten ponderar mejor la deseabilidad (o no) de su uso. Es decir, la percepción de riesgo tecnológico de la ciudadanía sobre la IA es un determinante importante de la legitimidad del uso de los sistemas de IA en el sector público. No se trata de apelar a la generación de información que minimice los riesgos en aras de promover su uso, sino de promover mecanismos de diálogo y difusión que formen en la ciudadanía y en los funcionarios públicos una percepción más ponderada y equilibrada sobre lo que es (y no es) IA; cuándo puede presentar riesgos y cuándo beneficios.

Por lo tanto, el proceso de adopción de la IA en el sector público requiere de intensificar los esfuerzos de difusión y socialización de los riesgos y beneficios de la IA en la población en general. Los riesgos vinculados a esta tecnología requieren esfuerzos para que la ciudadanía no solo los conozca, sino que además tenga la posibilidad de discutir y deliberar sobre la pertinencia de su uso en distintos contextos del sector público. Algunos organismos internacionales han sentado bases útiles en la consecución de este propósito.

Tal es el caso del reporte *Artificial Intelligence in the Public Sector* del Banco Mundial, en donde, entre otras recomendaciones, sugiere que los hacedores de política pública deben buscar que los gobiernos adopten políticas y marcos de gobernanza que promuevan una IA centrada en el ser humano, junto con inversiones en capital humano e infraestructura digital para fomentar la investigación en IA, el desarrollo de habilidades digitales y el emprendimiento en tecnología. Asimismo, identificar y gestionar los riesgos mediante marcos de evaluación del impacto de los sistemas de IA, adaptados al contexto de cada país, en lugar de evitarlos (World Bank, 2021).

De igual forma, la OCDE, a través de Governing with Artificial Intelligence, señala que, con las salvaguardas necesarias, el uso de sistemas IA puede transformar positivamente la operación gubernamental, la formulación de políticas y la entrega de servicios públicos. Sugiere que es necesario comprender y habilitar los aspectos positivos del uso de IA en lugar de solo enfocarse en la prevención de los aspectos negativos dado que este enfoque puede desalentar el uso de esa tecnología. En ese sentido, establece que el uso confiable de la IA surge de un esfuerzo colectivo a través de la participación de múltiples actores que buscan compartir sus aprendizajes y experiencias conjuntamente (Santiso, 2024).

Por su parte, el Joint Research Centre de la Comisión Europea, en el reporte técnico *AI Watch: Artificial Intelligence Standardization Landscape Update*, hace énfasis en tres elementos que considera clave: 1) La estandarización de la evaluación y mitigación del riesgo, resaltando la importancia de contar con estándares técnicos para identificar, medir y mitigar sesgos dentro de la IA. 2) La transparencia y explicabilidad, para efecto de evaluar qué tan comprensible es un sistema para distintos usuarios. Y 3) la gestión de riesgos y procesos de diseño ético, a fin de identificar riesgos en la fase de diseño y de generar controles técnicos para mitigar impactos negativos (Soler Garrido *et al.*, 2023).

Con base en esta discusión, la primera recomendación derivada de esta obra es promover la difusión de información sobre IA y la deliberación con la ciudadanía.

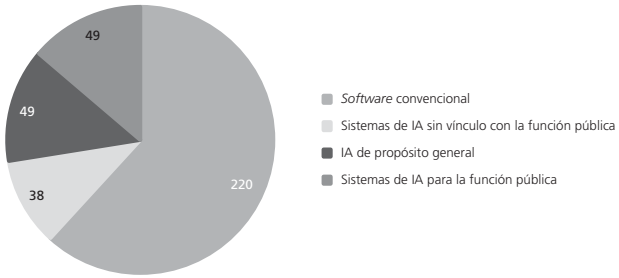
Recomendación 1. Promover la difusión de información sobre la IA y sus aplicaciones entre la ciudadanía.

Los niveles de información y conocimiento sobre la IA y sus aplicaciones entre la ciudadanía y, en general, entre los diferentes actores del ecosistema de IA es variable y depende de cuestiones como niveles de alfabetización y competencias digitales. Por ende, es necesario establecer distintos niveles de acceso a la información dirigido a:

— **Alfabetización en IA.** Con el fin de que las personas comprendan los beneficios, riesgos, salvaguardas, derechos y obligaciones en relación con el uso de un sistema de IA, tomando como base el artículo 4 del Reglamento de IA de la Unión Europea.
— **Conocer cuándo se interactúa con un sistema de IA.** Lo que incluye la prestación de servicios directamente o con ayuda de un sistema de IA y la identificación de los contenidos que han sido generados y/o modificados por sistemas de IA.
— **Poseer información significativa sobre los sistemas de IA.** La complejidad de estos sistemas no puede ser una excusa para no ofrecer información a las personas interesadas, sobre todo cuando estas interactúan con un sistema de IA. El principio debe ser el de brindar información significativa sobre la lógica aplicada, el cual permite obtener información concisa, accesible y entender, con un lenguaje claro y sencillo, sobre el método y los criterios utilizados en cada decisión. La información debe ser suficientemente completa y contextualizada para permitir al interesado verificar su exactitud, coherencia y una relación de causalidad entre el método y los criterios utilizados y el resultado al que ha llegado la decisión automatizada (TJUE, C-203/22). Poder determinar con exactitud la cantidad y calidad de la información que es necesario revelar para poder cumplir con los estándares de transparencia algorítmica de la ciudadanía no es claro, comprender las capacidades y limitaciones de los sistemas está asociado a temas de índole tecnológico y a cuestiones de seguridad, contextos, conocimientos y habilidades técnicas, derechos de terceros y de expectativas. Por un lado, se debe reconocer en todos los supuestos que existe la obligación de los proveedores y de la administración pública de proporcionar dicha información de forma significativa; y por el otro, que existe un derecho de los ciudadanos a saber.

La desinformación sobre los beneficios y riesgos de la IA no solo permea a la ciudadanía en general, sino también a la administración pública. A pesar de los avances en el campo y la emergencia de sistemas de IA en el sector público, prevalece la desinformación sobre sus potenciales usos en las entidades públicas. Los hallazgos del mapeo sugieren que la falta de alfabetización en IA es un reto por atender de inmediato. Por ejemplo —y solo como sugerencia—, los resultados de las solicitudes de transparencia muestran un nivel importante de desinformación en la función pública sobre qué tipo de *software* está basado en IA.[5] De los 356 sistemas de IA reportados en las respuestas, el 62% refirió como sistemas de IA el *software* tradicional como Contpaq, Adobe Creative Cloud, Karspersky o Microsoft Office (gráfico VI.1). Esto denota cierta desinformación en la función pública sobre el *software* que tiene o no componentes de IA.

Gráfico VI.1. Número de sistemas informáticos (de IA y otros) reportados mediante solicitud de transparencia, por tipo de sistema informático



Notas: (a) Véase definiciones de *software* convencional, IA de propósito general y sistemas de IA para la función pública en la sección "Casos excluidos" del capítulo III; (b) los datos de este gráfico se reportan en el cuadro III.3.

Fuente: Elaboración propia.

5 En las solicitudes de transparencia se definió a los sistemas de IA como "cualquier instrumento de *software* que haga uso de sistemas de IA, herramientas tecnológicas o equivalente".

Por otro lado, los hallazgos de la encuesta realizada a las direcciones generales de Tecnologías de la Información del gobierno federal indican que los determinantes relacionados con los recursos humanos, su formación y disposición son importantes al momento de decidir sobre la adopción o no de la tecnología, pero también en las etapas posteriores relacionadas con su contratación y/o desarrollo interno, y despliegue. De hecho, el predictor más importante para adoptar o contratar sistemas de IA es el conocimiento de los adoptados por otras agencias públicas, a nivel nacional e internacional, lo cual implica un conocimiento sobre en qué y cómo se puede aprovechar la tecnología.[6] Sin embargo, este conocimiento debe ser analizado en conjunto con otros factores, como la falta de personal especializado en temas relacionados con IA, capacitación, conocimientos técnicos en los equipos para el desarrollo de herramientas de IA, la resistencia al cambio por parte de usuarios y empleados, y las modificaciones que se introducen en la forma en que se realiza el trabajo.

Con base en esta evidencia, se emite la recomendación 2, que enfatiza la importancia de no solo conocer sobre aplicaciones y usos, sino también generar las capacidades en recursos humanos necesarios para garantizar una adopción y despliegue responsable.

Recomendación 2. Promover la difusión de información y formación de los funcionarios públicos en temas de IA.

Esta promoción debe ser entendida desde dos perspectivas: desde aquellas personas que intervienen en el desarrollo y despliegue de soluciones tecnológicas, y desde los funcionarios públicos que utilizan las soluciones de IA en el ejercicio de sus funciones. Tanto en uno como en otro caso es necesario que existan programas y acciones destinadas a la formación en temas de IA, y a la interlocución entre los diversos actores.

— **Promover conocimientos para la replicación y compartición de sistemas de IA en los diferentes niveles de gobierno.** Garantizar

Continúa

6 Véase la figura 1 en el apéndice D.

conocimientos técnicos, legales y éticos a fin de que la tecnología se pueda replicar y compartir en los diferentes niveles de gobierno de forma ética, responsable y en favor de la ciudadanía y los procesos de la administración pública federal (APF), para lograr los resultados esperados. Esta alfabetización es en especial relevante para quienes trabajan en procesos de acompañamiento técnico e implementación de la tecnología de IA en los diferentes niveles de la APF y en el desarrollo de productos tecnológicos públicos (lo que incluye la IA).

- **Alfabetización en IA.** Desarrollar programas de formación en materia de IA, con la finalidad de que se conozcan las capacidades y riesgos que se le asocian, en sus etapas de diseño, despliegue e implementación. Ello incluye desde saber cómo funciona, sus características generales, beneficios de su utilización en la administración pública, riesgos, y formas de mitigarlos, entre otros aspectos.[7] Esto es muy importante para los funcionarios públicos expertos en las problemáticas y/o procesos donde la IA impactará y/o será utilizada.

— **Promover el trabajo conjunto entre desarrolladores y usuarios de sistemas de IA, a fin de minimizar los riesgos asociados a su uso.** Lograr que los funcionarios públicos trabajen en conjunto con diseñadores y desarrolladores de IA en la evaluación de los riesgos de incorporarla en contextos críticos o en la prestación de servicios, así como la detección de sus riesgos y minimización de sus efectos. Esta participación podría incluir aspectos tan disímiles como gestión de programas, investigación, desarrollo, pruebas, análisis, evaluación y supervisar cómo la IA podría incidir en los servicios públicos, en especial en cuestiones relacionadas con riesgos, oportunidades y brechas de áreas esenciales como pueden ser temas financieros, educativos y sanitarios, entre otros. En este tenor, también se deben conformar mecanismos de coordinación conjunta y equipos multidisciplinarios compuestos por personal técnico y académicos, sector privado, público e instituciones de la sociedad civil que sean expertas en temas para identificar los casos en los que el uso de los sistemas de IA podría resultar inadecuado en el sector, y que participen activamente en su diseño, puesta en funcionamiento de mecanismos de supervisión y evaluación, así como prácticas de seguridad y protección.

---

7 Un ejemplo es Artificial Intelligence Training for the Acquisition Workforce Act or the "AI Training Act". Public Law, (117-207) (10/17/2022), donde se establece un programa de formación en IA para el personal de adquisiciones, y para otros fines.

Finalmente, es importante reseñar que no solo se necesita fortalecer las actividades de difusión de información, sino también las de generación de conocimiento sobre sistemas de IA en el sector público. Actualmente, existe poca información sobre los sistemas en uso, carencia que los esfuerzos de conformación de repositorios buscan mitigar alrededor del mundo. En Europa, destacan los registros de IA en Ámsterdam (Gemeente Amsterdam, s. f.), Helsinki (City of Helsinki, s. f.), Nueva York (City of New York, 2023) y Nantes (Nantes Métropole, s. f.), así como el Estándar de Transparencia Algorítmica del Foro Digital de Eurocities (Algorithm Register, s. f.; European Commission, s. f.). En Reino Unido, el *Algorithmic Transparency Recording Standard* centraliza la información sobre estos sistemas. Asimismo, en América Latina tienen lugar iniciativas como el Repositorio de Algoritmos Públicos en Chile (Gobierno de Chile, s. f.), la base de datos sobre sistemas de decisión automatizada en Colombia (Gobierno de Colombia, s. f.), y el Observatorio de IA en el Estado en Uruguay (Agesic-Gobierno de Uruguay, s. f.). En América del Norte, el gobierno de los Estados Unidos exige a los organismos federales publicar un inventario de sus sistemas de IA (U. S. Government, s. f.), mientras que Canadá ha implementado la Directiva sobre Toma de Decisiones Automatizada para centralizar esta información. La presente obra es, en parte, un producto derivado de la implementación del "Repositorio de sistemas de IA en el ámbito público", promovido por el programa de investigación Algoritmos CIDE para establecer un inventario de sistemas de IA en el ámbito público mexicano.[8]

Los repositorios de sistemas de IA generan información valiosa en varios sentidos, por ejemplo, dan origen a una base de casos de éxito o de estudio de buenas prácticas, estableciendo criterios uniformes para la publicación de información sobre IA; mapean los tipos de funciones que los sistemas de IA están cumpliendo en las entidades públicas; mapean a las entidades públicas que están avanzando en su desarrollo o uso y sientan las bases para generar documentos téc-

8 Véase el sitio https://algoritmoscide.org/

nicos (*white papers*) para facilitar la compartición de desarrollos propios (GPAI, 2024a, 2024b). De igual forma, proveen información que permite a legisladores, jueces y *policymakers* tomar mejores decisiones y diseñar mejores políticas sobre el uso de la IA en la administración pública.

Sin embargo, se debe reconocer que incluso este tipo de proyectos aún carecen de información esencial particularmente relacionada con la transparencia algorítmica. Esta es parte de la denominada gobernanza técnica de la IA, su implementación depende, en gran medida, de encontrar las herramientas y mecanismos necesarios para garantizar el acceso a información significativa del sistema de IA. Su principal reto se encuentra en determinar cuánta información es suficiente o necesaria, para ello es importante tener en cuenta el contexto en el cuál será suministrada, sus fines y sus destinatarios. La transparencia debe ser suficiente tanto para los ciudadanos como para que la propia APF pueda comprender el funcionamiento de la tecnología, dentro de lo razonable.

Recomendación 3. Promover políticas y herramientas de transparencia algorítmica.

La transparencia algorítmica es un medio para el ejercicio de otros derechos y obligaciones en cuanto a la gobernanza de un sistema de IA. Se materializa con el establecimiento de medidas jurídicas y tecnológicas durante todo el ciclo de vida de un sistema de IA, y son distintas para los diversos agentes que intervienen en dicho ciclo.

Algunas de estas medidas pueden ser:

— Informar cuando un usuario interactúa con un sistema de IA, y de las capacidades y limitaciones del mismo a fin de poder comprender y cuestionar las decisiones.
— Informar cuando un producto o servicio se proporciona directamente o con ayuda de un sistema de IA.
— Documentar la información pertinente, lo cual incluye documentación técnica y no técnica a lo largo del ciclo de vida de un sistema de

Continúa

IA, información sobre decisiones de diseño, datos de entrenamiento, estructura del modelo, decisiones en relación con el despliegue y uso de cada sistema, modelos, arquitectura, datos, rendimiento, limitaciones, uso adecuado y pruebas, etcétera.
— La obligación anterior incluye conservación de registros para su uso por parte de los proveedores posteriores, su introducción en el mercado y facilitarla a las autoridades nacionales competentes.
— Identificar los contenidos que han sido generados y/o modificados por la IA.
— Resumen de los contenidos utilizados para el entrenamiento de los modelos de IA de propósito general.
— Obligaciones relativas a la detección y etiquetado de contenidos generados o manipulados de manera artificial.
— Obligación de proporcionar acceso al código fuente y/o a los conjuntos de datos.
— Exigir a los proveedores, contratistas y beneficiarios de subvenciones públicas la adopción de prácticas sólidas de gobernanza y gestión de la IA, en particular en relación con los modelos de los que se trate la contratación o subvención (NTIA, 2024).
— Establecer protocolos claros y ejecutables de acceso a la información, durante todo el ciclo de vida de los sistemas a partir de su estandarización y el establecimiento de formatos accesibles en un lenguaje sencillo, como tarjetas de modelos o sistemas u hojas de datos de IA.
— Repositorios de sistemas de IA en el sector público. Prácticas de cocreación a partir de la implicación de la ciudadanía en procesos de diseños, para garantizar la aceptación pública (European Union, 2024).

Debilidades en capacidades organizacionales dificultan la adopción de sistemas de IA

La construcción de capacidades organizacionales constituye un requisito fundamental para la adopción efectiva de sistemas de IA en el sector público. La literatura ha mostrado que estas capacidades abarcan recursos financieros, infraestructura tecnológica, calidad de la información y capital humano especializado, los cuales determinan en gran medida la viabilidad de proyectos de innovación digital (Madan

y Ashok, 2023; Pumplun *et al.*, 2019; Tomazevic *et al.*, 2024). Sin embargo, estudios recientes subrayan que no basta con disponer de infraestructura técnica, pues la adopción de IA depende también de capacidades institucionales más amplias, como los arreglos de gobernanza de datos y la gestión del cambio organizacional (Neumann *et al.*, 2024; Selten y Klievink, 2024).

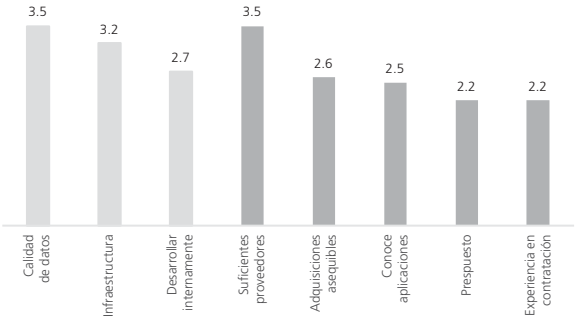
En este sentido, la evidencia empírica confirma que la falta de personal con competencias analíticas, la fragmentación de los sistemas de información y la rigidez de los marcos normativos representan barreras decisivas para la integración de la IA en los procesos públicos (Millan Vargas y Sandoval-Almazán, 2024). En consecuencia, las capacidades organizacionales no solo deben entenderse como un factor estructural de preparación tecnológica, sino como un conjunto de condiciones dinámicas que habilitan la absorción de la innovación, el manejo de riesgos percibidos y la legitimidad de la adopción de IA dentro de las instituciones gubernamentales.

Los datos obtenidos mediante la encuesta a directivos en el área de tecnologías de la información y de las entrevistas semiestructuradas muestran que las capacidades organizacionales son un determinante crítico para la adopción de sistemas de IA en el sector público, pero también que estas capacidades deben ser desagregadas y entendidas de manera diferenciada. Se observó que las entidades públicas se autoperciben con un nivel bajo de capacidades de información. Los resultados de la encuesta indican que, en promedio, las entidades públicas se autocalifican con 3.2 en infraestructura en sistemas de información, 3.5 en capacidad de datos y 2.7 en capacidad para el desarrollo interno. Las tres calificaciones están por debajo del promedio de la escala, que es de 4.0 (gráfico VI.2).

El déficit en la autopercepción de capacidades también se extiende a otras de tipo organizacional y que son claves para la adopción de la IA: el presupuesto asignado para el desarrollo o contratación de tecnologías de la información; la experiencia previa en la contratación de sistemas de IA o tecnologías similares; la capacidad institucional para realizar adquisiciones flexibles y su nivel de conocimiento de este sector, de proveedores de o servicios relacionados, así como el nivel de

conocimiento de sistemas de IA que estén siendo utilizadas en el ámbito público. Algunos de estos aspectos han sido previamente debatidos, y están relacionados con las capacidades de gestión de proyectos de IA. En el rubro de capacidades organizacionales, los resultados de la encuesta muestran un déficit de capacidades (autovaloraciones por debajo del promedio del rango) en los cinco rubros analizados, de los que destacan los déficits en presupuesto y experiencia en contratación de sistemas de IA (gráfico VI.2).

Gráfico VI.2. Autoevaluación en capacidades organizacionales clave para la adopción de sistemas de IA



Notas: Barras en tono claro denotan las capacidades de información. Barras en tono oscuro denotan otras capacidades organizacionales relacionadas con la adopción de IA. La escala de las preguntas aplicadas en el cuestionario es de 1 a 7, donde 1 significa bajo nivel de preparación y 7 significa alto nivel de preparación.

Fuente: Elaboración propia.

Los hallazgos emergentes de las entrevistas muestran que la falta de capacidades organizacionales constituye uno de los principales obstáculos para la adopción de sistemas de IA en el sector público. También se observan limitaciones estructurales en cuatro ámbitos críticos: recursos financieros, capital humano, infraestructura tecnológica, y calidad de la información. La insuficiencia presupuestaria aparece como una restricción transversal, impidiendo la inversión en

programas de capacitación, infraestructura y acceso a tecnologías digitales, lo que se traduce en carencias persistentes incluso en agencias con ingresos propios. A ello se suma la escasez de personal con competencias técnicas en procesamiento de datos, modelado analítico y desarrollo algorítmico, así como la sobrecarga de trabajo que deja sin margen a los funcionarios para incorporar nuevas tecnologías. En paralelo, los déficits de infraestructura básica —*hardware*, *software* y plataformas digitales— se entrelazan con vacíos en la gobernanza de datos y protocolos de seguridad.

En este marco, la literatura sobre capacidades de información aporta una perspectiva clave: la capacidad del Estado para recolectar, procesar e integrar información de calidad sobre ciudadanos y territorio determina la efectividad de las políticas y la legitimidad de las instituciones (Peeters *et al.*, 2023). La capacidad de información es un concepto relativamente reciente para describir los sistemas de información que un gobierno necesita para recolectar y procesar la que tiene sobre sus ciudadanos y su territorio (Peeters *et al.*, 2023). Para algunos, la idea de capacidad de información parte del trabajo seminal de Scott (1998) sobre capacidad del Estado, de tal manera que la primera se posiciona como una dimensión de la segunda para establecer el orden social e instrumentar procesos de ingeniería social (Lee y Zhang, 2017; Brambor *et al.*, 2020; D'Arcy y Nistotskaya, 2017).

Aunque su tratamiento en la literatura está más relacionado con el estudio de la información estadística nacional o la capacidad coercitiva del Estado, la idea de capacidades de información desde la perspectiva de la gestión pública tiene una relevancia fundamental para conocer las posibilidades prácticas de adoptar sistemas de IA en las entidades públicas. La capacidad de los gobiernos para recopilar, procesar y utilizar información en la toma de decisiones es crucial para comprender las posibilidades prácticas de adoptar sistemas de IA, pues la efectividad de algunos de esos sistemas —o la no maleficencia en su uso— requiere que los gobiernos generen información de calidad. La capacidad de información significa, en términos simples, que los gobiernos cuenten con información completa y precisa sobre la

población que sería afectada por las decisiones o acciones públicas asistidas por un sistema de IA.

Por tanto, un aspecto central del fortalecimiento de la capacidad de información son los sistemas de información y las tecnologías de registro de un gobierno (Bowles, 2020). Otros pilares de la capacidad de información de un gobierno son sus capacidades de generación de datos de calidad y de realizar o administrar sistemas de IA, que son resultado de la inversión, no solo en infraestructura, sino en la capacitación de funcionarios de primera línea en la captura y llenado de registros, la adquisición de recursos humanos expertos en ingeniería de datos, ciencia de datos, o ciencias de la información.

De ahí que, en el caso de la IA, el fortalecimiento de las capacidades organizacionales no pueda reducirse a la adquisición de infraestructura tecnológica, sino que implica desarrollar sistemas de información robustos, competencias técnicas entre el personal y marcos institucionales que aseguren el uso responsable y sostenible de la tecnología en la gestión pública.

Con base en este análisis, se considera importante fortalecer la capacidad de información y de ciencia de datos de los gobiernos, como se expone a continuación.

Recomendación 4. Fortalecer la capacidad de información y de ciencia de datos de los gobiernos.

— Fortalecer la capacidad de los gobiernos en procesos de generación, procesamiento y uso de la información de calidad, durante todas las etapas del ciclo de la información: generación, integración, procesamiento y uso. Las agencias públicas deben establecer estrategias y protocolos de cómo fortalecer de manera eficiente sus capacidades para producir información de calidad, incluyendo lo relacionado con la seguridad de la información, protección de los datos personales, estándares de ciberseguridad e interoperabilidad.
— Formar a funcionarios públicos encargados de sistemas de información en aspectos técnicos de la IA, metodologías ágiles de desarrollo, meca-

Continúa

nismos de gobernanza de datos y su interpretación y análisis para ser utilizados en la toma de decisiones y diseño de políticas públicas. Ello incluye la formación de los funcionarios en la ética del uso de los sistemas de IA, lo cual es fundamental para que tengan las herramientas necesarias y puedan gestionar el desarrollo o contratación de sistemas de IA, o plantear las salvaguardas necesarias en protección y privacidad de datos personales, procedimientos de recogida de datos personales y no personales; preparación de los datos para operaciones de tratamiento, titularidad de los conjuntos de datos de las entidades públicas, de los proveedores y de terceros, transferencia de datos, auditoría, prevención, detección y mitigación de sesgos presentes en los conjuntos de datos.

Los sistemas de IA dependen fundamentalmente de la calidad de los datos disponibles. Generarlos con calidad es un reto para todo el sector público, no solo porque permite implementar tecnologías que dependen de ellos, como la IA, sino para fortalecer la planificación, la operación de programas y la toma de decisiones estratégicas.

En consecución de ello, la cultura de datos y de transparencia son condiciones necesarias, porque alientan una dinámica organizacional más abierta a la correcta gestión de sistemas de información y a la compartición de datos. La cultura de datos y la cultura de transparencia aseguran insumos confiables y gobernados —calidad, integridad, interoperabilidad y seguridad de los datos— porque institucionalizan la trazabilidad del ciclo de vida algorítmico (documentación, explicabilidad, auditoría), reduciendo la opacidad y sosteniendo la legitimidad pública de su uso. La literatura sobre laboratorios de innovación en políticas públicas sitúa datos abiertos y gobierno abierto como capacidades nucleares que catalizan innovación y confianza, al tiempo que refuerzan decisiones basadas en evidencia y mecanismos de rendición de cuentas, lo que conecta directamente con el desempeño técnico y la aceptación social de sistemas de IA en el sector público (Renteria y Gómez-Alvarez, 2024).

La creación de la Dirección General de Inteligencia de Datos dentro de la Agencia de Transformación Digital y Telecomunicaciones se perfila como un paso importante en tal sentido. Entre las diversas funciones que tiene esta dirección se pueden mencionar: a) diseñar,

supervisar y evaluar la política de gestión de datos de la APF; b) coordinar la estrategia de datos abiertos de la APF; c) crear un inventario de los datos que las dependencias, sus órganos administrativos desconcentrados, y entidades de la APF y otras autoridades generen y guarden; d) diseñar, implementar, gestionar y modernizar un sistema de información unificado para uso de las dependencias, sus órganos administrativos desconcentrados, y entidades de la APF y otras autoridades; y e) procesar, analizar y visualizar los sistemas de información y los conjuntos de datos de las dependencias, sus órganos administrativos desconcentrados, y entidades de la APF y otras autoridades, entre otras (Presidencia de la República, 2025). Este reconocimiento legal constituye el primer paso para poder consolidar una cultura de transparencia y de datos de las agencias públicas; sin embargo, además de este reconocimiento es necesario trazar una estrategia que se materialice en acciones concretas y viables para identificar áreas de oportunidad de mejora de las operaciones actuales; el fortalecimiento de capacidades es consolidar las culturas de transparencia y de datos en las organizaciones públicas y la producción de datos de calidad.

Por este motivo, se recomienda la promoción de una cultura de datos y una cultura de transparencia en las instituciones públicas.

Recomendación 5. Fomentar una cultura de datos y de transparencia en las entidades públicas.

Revalorizar los datos en la administración pública, pensándolos como un activo, conociendo y dominando su ciclo de vida. Identificar las características que diferencian los buenos de los malos datos y gerenciar su gobernanza. En el ámbito público, la consideración del dato como activo importante para la toma de decisiones, así como para el diseño y evaluación de políticas públicas va aparejado de la consolidación de una cultura de transparencia de datos de calidad dentro de las agencias públicas. Los funcionarios públicos no solo deben reconocer el valor del dato, sino que también deben gestionar su utilización de manera comprometida con su puesta a disposición y compartición de forma pública y transparente. Solo así nos encontraríamos en mejores condiciones para generar datos de calidad.

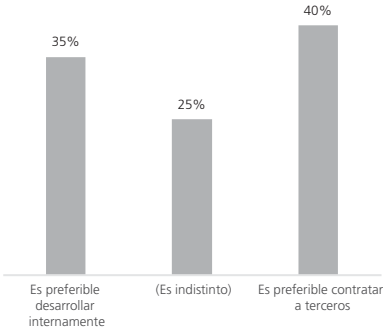
Desarrollo y compartición de sistemas de IA propios en el sector público

El caso mexicano revela signos de desarrollo interno como una de las principales formas de incorporación de sistemas de IA en el servicio público. Es decir, sistemas que fueron desarrollados por el equipo técnico de la misma entidad pública que utiliza el sistema de IA, por ejemplo, el clasificador de precios de Banxico, que clasifica el precio de una canasta definida o Progen del Instituto Nacional de Electricidad y Energías Limpias, que realiza el pronóstico de generación eólica a partir de aprendizaje profundo. El 33.3% o 18 de los 54 sistemas de IA sobre las que tenemos datos de forma de adquisición fueron desarrollados al interior de la dependencia que los utiliza: 15 en el sector público (31.25%) y 3 en el ámbito público no gubernamental (50%) (véase el gráfico IV.9, en el capítulo IV). A su vez, los resultados de la encuesta realizada a las direcciones de Tecnologías de la Información del gobierno federal indican que el 35% de los respondientes consideran que es preferible desarrollar los sistemas de IA internamente (gráfico VI.3), cifra que es cercana al 40% de respondientes que consideraron preferible contratar a terceros para desarrollar sistemas de IA. Estos resultados son notables, sobre todo considerando que las entidades públicas se autoperciben con bajos niveles de capacidades para la adopción de sistemas de IA, como se observa en el gráfico VI.3.

Estos patrones contrastan con la práctica global tradicional de la subcontratación de tecnologías de la información (TI), que típicamente se ha justificado por la eficiencia en costos, el acceso a recursos externos, la experiencia técnica ofrecida por el sector privado o la influencia de las normas de subcontratación establecidas en la gestión del sector privado (Cox *et al.*, 2012; Haugen y Klungseth, 2017; Moon *et al.*, 2016). Los gobiernos están modificando sus políticas de aprovisionamiento a medida que comienzan a considerar las tecnologías digitales —como las bases de datos ciudadanas y los sistemas de identidad digital— como componentes esenciales y estratégicamente sensibles de las operaciones gubernamentales (Moon *et al.*, 2016; Thunes y Kempton, 2025). Sin embargo, no existe consenso respecto a por qué

los desarrollos internos representan una política adecuada para la gestión de infraestructuras críticas. Algunos académicos enfatizan razones técnicas o pragmáticas. Por ejemplo, Mergel (2019) y Clarke (2020) sostienen que los equipos de servicios digitales (una forma de desarrollo interno) surgieron como respuesta a ineficiencias históricas en la infraestructura de TI gubernamental y al fracaso de los modelos de gobernanza de TI para ofrecer servicios digitales ágiles y centrados en el usuario. Desde una perspectiva más amplia, el marco de la *digital era governance* (DEG) —un modelo normativo de administración pública— interpreta iniciativas como el desarrollo interno de *software* como parte de un cambio más amplio orientado a reintegrar capacidades digitales, revalorizar la experiencia interna y establecer infraestructura digital de propiedad estatal para mantener el control soberano sobre servicios públicos críticos (Clarke, 2020; Kempeneer y Heylen, 2023).

Gráfico VI.3. Preferencia de direcciones de Tecnologías de Información sobre el desarrollo o la contratación de sistemas de IA



Nota: La pregunta del cuestionario fue: "En estos proyectos, existe la disyuntiva sobre si un sistema de IA debiese desarrollarse internamente o contratarse a terceros. En una escala del 1 al 5, donde 1 significa 'Es preferible desarrollar internamente' y 5 significa 'Es preferible contratar a terceros', ¿cuál considera que es la mejor opción para su organización?".

Fuente: Elaboración propia.

Las políticas de aprovisionamiento de los gobiernos están moldeadas no solo por consideraciones técnicas o pragmáticas, sino tam-

bién por narrativas de política en competencia que circulan en la esfera pública. En algunos países, los cambios hacia modelos alternativos de aprovisionamiento son impulsados por narrativas que presentan la subcontratación como problemática debido a la dependencia de proveedores. En esta narrativa, la subcontratación se percibe como una práctica que encierra a los gobiernos en contratos costosos y restrictivos, debilitando su capacidad operativa (Cox *et al*., 2012; Thunes y Kempton, 2025). Por el contrario, la narrativa de la soberanía tecnológica está ganando protagonismo en varios países, particularmente como respuesta a los desafíos recientes vinculados con la dependencia de estructuras de gobernanza global, acuerdos de defensa militar y cadenas de suministro extranjeras (Edler *et al.,* 2023; Luitse, 2024). Esta narrativa de soberanía prioriza la capacidad nacional para reducir la dependencia tecnológica respecto a otros países, en especial en sectores económicos críticos (Lee *et al*., 2024). Por ejemplo, la Comisión Europea asignó recientemente mil trescientos millones de euros a inversiones en IA, lo que señala un alejamiento de la narrativa previamente dominante de la "IA confiable", que enfatizaba un enfoque cauteloso priorizando preocupaciones éticas, hacia una centrada en la soberanía tecnológica. Esta nueva postura subraya la importancia de alcanzar a los competidores globales para evitar que Europa se convierta en una "colonia digital" (European Commission, 2025; Verhofstadt, 2025).

Diversos casos muestran que el desarrollo interno de servicios tecnológicos es una tendencia internacional. Por ejemplo, en los Estados Unidos el U.S. Digital Service (USDS) es un equipo dentro del gobierno federal que desarrolla soluciones digitales para mejorar la eficiencia y accesibilidad de los servicios gubernamentales (United States Digital Service, s. f.). En el Reino Unido, el Government Digital Service (GDS) busca transformar la provisión de servicios públicos mediante el uso de tecnología digital y ha liderado iniciativas como GOV.UK, la plataforma unificada para acceder a servicios gubernamentales en línea. En Canadá, el Canadian Digital Service (CDS) colabora con departamentos federales para diseñar y ofrecer servicios digitales centrados en las necesidades de los usuarios. En Australia, la Digital

Transformation Agency (DTA) lidera las iniciativas digitales del gobierno australiano, con el objetivo de hacer que los servicios sean más simples, claros y rápidos. La DTA proporciona orientación, herramientas y apoyo a las agencias gubernamentales para mejorar la prestación de servicios digitales.

Esta tendencia de la preferencia de las entidades públicas por el desarrollo interno de sistemas de IA tiene varias explicaciones. En primer lugar, se alinea con una creciente cultura de desarrollo de *software* en código abierto y de un interés de las entidades públicas por compartir sus desarrollos con otras entidades públicas. Tal es el caso de la Agencia Digital de Innovación Pública (ADIP) de Ciudad de México en 2018, que impulsó el desarrollo de *software* interno como un mecanismo para lograr austeridad para la eficiencia del gasto público, eliminar dependencias con privados, generar capacidades públicas y alcanzar la soberanía tecnológica (Merino y Muñoz del Olmo, 2025).

La agenda de transformación de la ADIP se sustenta, principalmente, en una narrativa de búsqueda de *soberanía tecnológica*, que se entiende como la independencia de las entidades públicas de proveedores en el desarrollo, gestión, operación y mantenimiento de sus plataformas digitales y sistemas de información críticos (Merino y Muñoz del Olmo, 2025). Más allá de las necesidades funcionales inmediatas, los servicios tecnológicos son considerados como un activo estratégico que otorga ventajas económicas importantes en el largo plazo (Mergel, 2017). El nacionalismo tecnológico busca la autosuficiencia en tecnologías digitales y la consolidación de ecosistemas locales de empresas y expertos capaces de brindar una mayor autonomía al gobierno y a los sectores industriales frente a la potencial escasez de proveedores o fluctuaciones de los mercados externos (Margetts y Dorobantu, 2019).

Más allá de las necesidades funcionales inmediatas, los gobiernos cada vez más ven en la IA un activo estratégico que otorga importantes ventajas económicas y geopolíticas (Margetts y Dorobantu, 2019). El nacionalismo tecnológico, que a menudo se manifiesta en políticas orientadas a la autosuficiencia en tecnologías digita-

les, impulsa ecosistemas nacionales de IA que respaldan la industria local y el desarrollo de la fuerza laboral. Al fomentar capacidades internas en IA, los gobiernos buscan mayor autonomía, y protegerse de las fluctuaciones del mercado externo y de vulnerabilidades estratégicas.

La ADIP instrumentó esta visión mediante varias líneas de acción. Por un lado, invirtió en el fortalecimiento de capacidades de información, tanto de infraestructura como de recursos humanos. Por ejemplo, diseñó programas de capacitación y desarrollo profesional para funcionarios públicos en el uso de nuevas tecnologías y en la gestión de datos a fin de asegurar que las herramientas digitales se utilicen de manera efectiva y eficiente. Asimismo, el gobierno de la Ciudad de México aumentó su capacidad en infraestructura tecnológica, pasando de 180 procesadores y 37 TB de almacenamiento en 2018 a 520 procesadores y 113 TB de almacenamiento para 2019. E invirtió —en conjunto con la iniciativa privada y la academia— en un centro de datos con certificación de estándares en calidad y seguridad. El interés de adquirir y robustecer la infraestructura tecnológica propia también responde al objetivo de implementar medidas de seguridad más estrictas y mayor resiliencia en los sistemas de información y adaptar la infraestructura a las necesidades específicas del gobierno de la Ciudad de México (Merino y Muñoz del Olmo, 2025); es decir, la ADIP parte de la premisa de que es fundamental que el gobierno tenga la propiedad de la infraestructura tecnológica crítica para sus operaciones y para la seguridad tecnológica.

En línea con el argumento de este apartado, la ADIP ha invertido en sus propios desarrollos tecnológicos, argumentando que los de tipo interno promueven la eficiencia de costos, una reducción (en el largo plazo) de la dependencia de proveedores de servicios tecnológicos, y un mejor control en la seguridad de la información sensible (Merino y Muñoz del Olmo, 2025). La producción de sus propios desarrollos está a cargo de la "Fábrica de Software", la cual reporta la generación de 380 desarrollos con 40 programadores, con un ahorro estimado de 2500 millones de pesos que se hubieran gastado de ha-

ber contratado a desarrolladores externos de *software* (Merino y Muñoz del Olmo, 2025).

Al igual que otras entidades públicas que han generado desarrollos tecnológicos propios, la ADIP ha mostrado interés en compartir su *know how* y sus desarrollos tecnológicos con otras entidades públicas. Muestra de ello ha sido el establecimiento de convenios de colaboración con los estados de Veracruz, Tlaxcala, Baja California, Michoacán, Campeche y Colima.[9] Y, más recientemente, la visión de compartición del desarrollo interno de *software* ha sido incorporada en la agenda del gobierno federal para el sexenio 2018-2024, mediante la recién creada Agencia de Transformación Digital y Telecomunicaciones. Como parte de la creación de capacidades tecnológicas propias del país también se ha anunciado la creación del Centro Nacional de Tecnología Pública (CNTP), cuyo fin es impulsar la adopción de tecnologías digitales en los tres niveles de gobierno. Entre los componentes de este centro se encuentran el "Repositorio de códigos", con un catálogo de productos digitales, acompañamiento técnico y normativo para la implementación de la tecnología, y la Escuela Pública de Códigos (Agencia de Transformación Digital y Telecomunicaciones, 2025). La Escuela Pública de Códigos prepara actualmente a más de cuatrocientas personas servidoras públicas de distintos estados y municipios del país en desarrollo de *software*, programación y nube pública. Aunque no es exclusivo en materia de IA, en principio nada obsta para que así pudiera llegar a ser, al propio tiempo que se brinda la posibilidad de formar recursos humanos altamente capacitados (Agencia de Transformación Digital y Telecomunicaciones, 2025).

Más allá de la cultura del *software* abierto o la narrativa de la autonomía tecnológica, hay otras razones por las que los gobiernos podrían optar por desarrollar en casa servicios tecnológicos, en particular de sistemas de IA. Un reto es la gestión de la privacidad de

[9] Gobierno de la Ciudad de México. (2022, enero); *El Economista* (2024, 19 de agosto); Gobierno del Estado de Baja California. (s. f.), y Agencia Digital de Innovación Pública de la Ciudad de México (s. f.).

datos y la soberanía sobre su resguardo. Al optar por desarrollos internos, los gobiernos buscan salvaguardar los datos privados e información sensible del gobierno y no exponerlos a terceros contratistas o a espacios en la nube (OCDE, 2021). La dependencia de proveedores externos plantea preguntas sobre el cumplimiento de las normativas de privacidad, especialmente en jurisdicciones con leyes estrictas de protección de datos, como el Reglamento General de Protección de Datos (GDPR) en la Unión Europea (Comisión Europea, 2016). Al mantener el control sobre el diseño, la implementación y el mantenimiento de los sistemas de IA, los gobiernos pueden garantizar una supervisión más estricta y mitigar los riesgos asociados con el uso no autorizado de datos.

Conceptos como desarrollos internos, compartir y autonomía tecnológica no están desligados del fortalecimiento del ecosistema de innovación en IA en el país, todo lo contrario. La innovación pública tecnológica de la IA debe convertirse en un mecanismo de tracción del resto de las partes de ese ecosistema, no solo para poder determinar las necesidades a las cuales deben responder, sino también al cómo debe ser dicha respuesta. La adopción responsable, efectiva y ética de la IA en el sector público solo es posible con un ecosistema nacional de IA consolidado. Por un lado, amplía, fortalece y vincula a las entidades que forman parte del ecosistema, como emprendimientos y empresas GovTech que proveen soluciones y/o requerimientos para los desarrollos propios de las entidades públicas, al propio tiempo que las universidades y/o centros de investigación contribuyen a la formación de capital humano, la investigación aplicada y a la vinculación de los nodos del ecosistema. Sin embargo, actualmente se observa una desarticulación entre los diferentes miembros de este ecosistema y un desbalance entre la discusión en los foros públicos sobre el tema y los desarrollos que realmente existen en México.

Por lo tanto, se recomienda fortalecer el ecosistema de IA en el sector público, de tal manera que el *know how* y los desarrollos propios no solo sean conocidos por otras entidades públicas, sino también incorporados.

Recomendación 6. Fortalecimiento del ecosistema de IA para compartir *know how* y desarrollos propios.

El ecosistema de IA nacional debe vincularse y ser parte fundamental de los esfuerzos de transformación digital para la innovación pública en el país. En este sentido, es importante:

— Reconocer el papel de la IA como parte integral de la política nacional de transformación digital, y el rol de los diversos actores del ecosistema de IA en ese proceso. La adopción de sistemas de IA en el sector público no es posible sin la vinculación de este sector con el privado, la academia y la sociedad civil. Ello debe realizarse conforme a los principios y particularidades de la política nacional de transformación digital, incluyendo el desarrollo interno, el intercambio de conocimiento y la autonomía tecnológica.
— Vincular el ecosistema de IA con la innovación pública a partir de los principios tecnológicos que asume la APF, en especial compartir tecnología; la generación de insumos para la creación de lineamientos y regulaciones en el uso de los sistemas de IA o tecnologías complementarias (p. ej., cómputo en la nube), entre otros. Esto implica el diálogo y la colaboración responsable de los actores de este ecosistema para hallar soluciones socialmente responsables conforme a los valores públicos.
— Establecer mecanismos y equipos multidisciplinarios para operar la gobernanza de los sistemas de IA en el ámbito público; y el reconocimiento de entidades, instituciones y organizaciones dedicadas al monitoreo de sistemas de IA para evaluar su desempeño técnico o de cumplimiento de la ética en el uso de los sistemas de IA en el ámbito público.

Los hallazgos de este proyecto de investigación sugieren que el desarrollo de sistemas de IA en la APF ocurre de forma fragmentada y poco sistematizada, sin una entidad que concentre su documentación, lo que dificulta el aprendizaje entre pares y la evaluación comparada (p. ej., desempeño, sesgos, seguridad). En este contexto, los repositorios de sistemas de IA y las entidades que los operan cumplen una función crítica como infraestructura de conocimiento: estandarizan la recolección de información; fomentan la mejora

continua a partir de casos de éxito, y reducen costos al promover la adopción de buenas prácticas. Además, su comunicación y difusión entre la ciudadanía favorece el acceso a información clara y actualizada sobre beneficios y riesgos, lo que refuerza la rendición de cuentas. La evidencia empírica de la encuesta realizada a directivos en el área de tecnologías de la información muestra que el conocimiento de sistemas adoptados por otras agencias —nacionales o internacionales— es el predictor más importante de adopción, por encima de las capacidades internas, lo que subraya que documentar y socializar experiencias exitosas acelera decisiones informadas de adopción. Finalmente, la literatura respalda que la transparencia algorítmica (como la que promueven los registros y repositorios públicos) mejora la confianza percibida en la toma de decisiones automatizadas, reforzando la legitimidad y el desempeño de estos sistemas (GPAI, 2024a).

Por lo tanto, se recomienda documentar y difundir casos de éxito, lo que no solo corrige la desarticulación actual, sino que activa mecanismos comprobados de aprendizaje interorganizacional, confianza pública y adopción responsable.

Recomendación 7. Documentar y difundir casos de éxito de sistemas de IA.

— Documentar y difundir casos de éxito de sistemas de IA en la APE. La socialización de casos de éxito es crucial para que las agencias públicas aterricen, mediante ejemplos prácticos, los problemas de la gestión pública que los sistemas de IA pueden o no ayudar a solucionar. Dado que la información de los desarrollos sobre estas tecnologías está fragmentada y dispersa, resulta necesaria estandarizarla para facilitar su recogida y sistematización, al tiempo que permita a las agencias públicas conocer los problemas que los sistemas pueden resolver y cómo los pueden adaptar en su organización.
— Comunicar los casos de éxito entre las entidades públicas y la ciudadanía. Además, es importante que las personas tengan acceso a información comprensible y actualizada sobre estos sistemas, y que parti-

Continúa

cipen en los procesos de construcción y difusión de estos casos, tanto para determinar cuánta información es necesaria suministrar, como para la rendición de cuentas y evaluación de resultados de dichos sistemas, conforme a las necesidades y expectativas que se posean sobre el uso y aplicaciones de la IA en el ámbito público.

El desarrollo interno busca atender preocupaciones propias de la ética del servicio público, como la transparencia y la rendición de cuentas. También brinda a los gobiernos una vía directa para supervisar el diseño y la implementación de sistemas de IA de acuerdo con los valores del sector público. Cuando los sistemas de IA se desarrollan y prueban dentro de un marco gubernamental, las directrices oficiales sobre equidad, transparencia y mitigación de sesgos pueden integrarse desde el inicio (OCDE, 2019). Esto es especialmente relevante en los procesos de toma de decisiones relacionados con el bienestar social, la salud o la justicia penal, donde el sesgo discriminatorio de los sistemas de IA podría tener profundas consecuencias sociales.

Por otro lado, las estrategias nacionales de desarrollo interno de *software* pueden, en el largo plazo, convertirse en mecanismos de ahorro en costos para el gobierno. Como ya se mencionó, la ADIP reportó un ahorro estimado de 2500 millones de pesos producido por el desarrollo interno en lugar de usar los servicios externos de *software* (Merino y Muñoz del Olmo, 2025). Si bien la externalización de la IA puede resultar rentable a corto plazo, los costos a largo plazo suelen incluir tarifas recurrentes de licencias y dependencia de plataformas propietarias (World Bank, 2022). El desarrollo interno requiere una inversión inicial en talento, capacitación e infraestructura, pero estos gastos pueden compensarse con el tiempo mediante la reducción de costos en licencias y servicios de terceros. Además, desarrollar experiencia en IA dentro del sector público permite a los gobiernos adaptar e iterar soluciones en respuesta a la evolución de las políticas sin depender de proveedores externos para cada actualización.

La innovación en IA depende del sector y, por supuesto, del contexto. Aun cuando una de las líneas de acción del Plan Nacional de

Desarrollo es precisamente generar productos tecnológicos públicos, en los que exista autonomía tecnológica y soluciones propias que reduzcan la dependencia del exterior,[10] lo cierto es que esto no se puede lograr sin comprender la necesidad de establecer sinergias entre las capacidades de la administración pública y el ecosistema de innovación en IA del país conformado tanto por el sector privado de la tecnología, como por otros actores que son el académico, organizaciones de la sociedad civil y centros públicos de investigación, entre otros.

Con base en esta discusión, se recomienda establecer canales y vías comunicantes entre las entidades públicas que han desarrollado sus propios sistemas de IA y las entidades públicas que puedan estar interesadas en incorporarlos.

Recomendación 8. Compartir los desarrollos internos (o propios) de sistemas de IA.

Compartir los sistemas de IA desarrollados dentro del sector público. Se han logrado múltiples desarrollos tecnológicos al interior del sector público mexicano, muchos de los cuales tienen el potencial de ser adaptados en otras entidades públicas. Y como en el programa de transformación digital se ha implementado el Centro Nacional de Tecnología Pública, consideramos que este es un importante instrumento para compartir los desarrollos tecnológicos obtenidos en materia de IA responsable, incluyendo buenas prácticas éticas. Que estos desarrollos se realicen con inversión pública y bajo licencias de código abierto sienta las bases para su compartición. Es necesario que el Repositorio de Código, como componente del Centro Público de Tecnología Pública, cuente con mecanismos claros y efectivos para garantizar la compartición de la tecnología entre las diversas entidades, así como la definición de los parámetros de diseño, implementación, adopción y evaluación de dicho proceso con toda la complejidad que ello implica (p. ej., monitoreo de desarrollos, adaptación de la tecnología a otros contextos o necesidades, mantenimiento de los sistemas, cumplimiento de éticas y normativas, entre otros).

---

[10] Constituyen parte de las líneas de acción del Eje Transversal 2: "Innovación Pública para el Desarrollo Tecnológico Nacional" del Plan Nacional de Desarrollo.

Complejidad en los procesos de adquisición y despliegue de sistemas de IA

La negociación y contratación de la tecnología es compleja y diversa; los sistemas de adquisición, arrendamiento y prestación de servicios son distintos al de un servicio tecnológico tradicional, no solo cuando se adquiere de terceros, sino también cuando se desarrolla internamente. No existe tampoco una forma única de adquirir la tecnología. Durante la realización de este proyecto hemos comprobado que existen formas diversas de adquisición, incluso no tradicionales o al menos contempladas en el régimen de contratación de la APF, como la donación entre instituciones, el establecimiento de convenios de colaboración, financiamiento proveniente de organismos internacionales e, incluso, proyectos donde se diferencia la institución financiadora de quien desarrolla y estos de la institución que despliega la tecnología. Es el caso del Sistema de Inteligencia de Datos para la Inspección Laboral (SIDIL), el cual permite responder a la demanda de selección inteligente de centros de trabajo a inspeccionar basada en la predicción de riesgos de incumplimiento.

A nivel federal los procedimientos de contrataciones de tecnologías y seguridad de la información se encuentran regulados en el acuerdo por el que se emiten las políticas y disposiciones para impulsar el uso y aprovechamiento de la informática, el gobierno digital, las tecnologías de la información y comunicación, y la seguridad de la información en la APF. Dicho acuerdo no contempla ninguna disposición sobre el tema de cómo se puede adquirir y adoptar tecnología de IA en el sector, aunque algunas de sus disposiciones podrían ser aplicadas, en particular las relacionadas con aplicativos de cómputo, como las políticas de privacidad y protección de datos personales (Coordinación de la Estrategia Digital Nacional, 2021). Sin embargo, otros aspectos de la contratación no se han contemplado, como los mecanismos de supervisión, transparencia algorítmica y explicabilidad.

La experiencia en contratación y adquisiciones flexibles constituyen parte de los principales ítems considerados para valorar el ni-

vel de preparación de las organizaciones para adoptar sistemas de IA por parte de los funcionarios encuestados, nos referimos a los ítems 2.2 y 2.6, respectivamente. Otros elementos deben tenerse en cuenta, como el hecho de que no siempre la entidad tiene la capacidad o todos los elementos para el desarrollo de la tecnología deseada, o que inclusive los proyectos pueden fracasar o no obtener los resultados esperados y, por lo tanto, necesitan ser evaluados en función de su eficacia. Los costos iniciales, al menos en etapas tempranas, pudieran ser elevados o superiores a los inicialmente presupuestados. Asimismo, los periodos para su implementación, pruebas y supervisión no siempre coinciden con los legalmente previstos en los acuerdos contractuales, así como tampoco con las limitaciones que habitualmente suele tener la APF en cuanto al cierre del año fiscal y necesidades de contratación.

A esto se le suma el esquema de innovación que ha asumido el país de compartir tecnología en el ámbito público, lo que implica que no solo sea necesario atender mecanismos de adquisición, arrendamiento y prestación de servicios al interior de la propia APF, es decir, cómo se otorgarán licencias de los desarrollos propios y de terceros en las diferentes dependencias, órganos administrativos desconcentrados y entidades de la APF, si de forma gratuita u onerosa.

No se trata solo de inversión en el desarrollo de la tecnología, sino también en cuestiones conexas a ella para garantizar su despliegue en términos éticos, como puede ser inversión en infraestructura, *hardware*, talento especializado y en el cumplimiento de estándares éticos y de normatividad. Todo ello pudiera atraer costos no previstos. El despliegue e implementación de la tecnología no está supeditado solo al momento en que se adopta la tecnología, sino que se relaciona también con el contexto en el que se aplica, y con la introducción de mejoras, la puesta en práctica de mecanismos de supervisión, mitigación de riesgos, cumplimiento de obligaciones normativas y éticas, y la determinación de responsabilidades de actores y agentes que intervienen durante todo su ciclo de vida.

Recomendación 9. Promover esquemas de adquisición ágiles, flexibles y responsables.

Dada la complejidad en los procesos de contratación en materia de IA, es necesario reducir la incertidumbre en ellos estableciendo marcos normativos claros para una contratación o desarrollo que dé certeza jurídica a los funcionarios públicos en la adopción de estas herramientas de forma ética y responsable. Esto incluye:

— Abrir espacios de diálogo sobre nuevos esquemas de contratación que se adapten a las necesidades de desarrollo, adopción y aplicación de IA, en particular los relacionados con la compartimentación del código fuente; el uso de mecanismos y herramientas como entornos controlados de prueba y contratación transparente; lineamientos básicos de ética al usar sistemas de IA; medidas de seguridad; protocolos para gestionar la factibilidad de sistemas, entre otros.
— Adoptar un marco jurídico claro que actualice y/o modifique la base normativa para la adquisición y aplicación ética de la IA en el ámbito de la APF de forma responsable, partiendo de la regulación jurídica existente, así como de la tecnología, usos y aplicaciones. El marco jurídico de contratación de la tecnología debe ser flexible y ágil dadas las diversas formas de adquisición y el nivel de incertidumbre de sus resultados; y debe establecer garantías para eliminar prácticas de corrupción.
— Llevar a cabo procesos de contratación en los que se tome en cuenta la tecnología, así como las características o elementos particulares del contexto y funcionalidades en donde se pretende utilizar el sistema de IA.
— Contemplar en la contratación disposiciones específicas sobre los sistemas de IA, como gestión de datos adecuada al contexto de uso, incluidos los relacionados con el entrenamiento, la validación y prueba; preparación de los datos para las operaciones de tratamiento, incluyendo la anotación, etiquetado, depuración, enriquecimiento y agregación; medidas adecuadas para detectar, prevenir y mitigar posibles sesgos, obligaciones relacionadas con la conservación y entrega de registros; medidas adecuadas para garantizar supervisión humana, nivel de seguridad y ciberseguridad, titularidad de los conjuntos de datos del organismo de administración pública, de terceros y del propio proveedor, evaluaciones de conformidad y medidas correctoras, obligaciones relacionadas con el cumplimiento y auditoría.

Los retos de la rendición de cuentas en el creciente uso informal de IA generativa

Una observación sutil del mapeo que casi pasa desapercibida, es que el 14% de los funcionarios que respondieron a la solicitud de transparencia indicaron la incorporación de sistemas de IA de propósito general (gráfico VI.1). Recapitulando la definición brindada en el capítulo III, la IA de propósito general se refiere a sistemas de IA que pueden realizar, o ser adaptados para realizar, un rango diverso de tareas, incluidas algunas para las que el sistema no fue diseñado. En este marco, se identificó que los funcionarios públicos utilizan sistemas por suscripción como ChatGPT, DALL-E, Midjourney o Firefly. Este tipo de uso de sistemas de IA se caracteriza por ser a título personal —con cuentas de usuarios personales— e informales. El uso informal de la IA generativa no es un patrón aislado ni menor. Todo lo contrario, se trata de una tendencia internacional de uso de sistemas de IA en la función pública. Por ejemplo, Bright *et al*. (2024) reportaron que el 44% de los servidores públicos en Reino Unido y el 24% en Canadá ya usaban IA generativa de manera informal; un informe del IBM Center for the Business of Government (2025) reporta que el 70% de los funcionarios públicos coreanos usan o tienen la intención de usar sistemas de IA generativa en su trabajo; la encuesta de Ernst y Young (2024) en agencias federales estadounidenses revela que el 60% de los funcionaros públicos reportan un uso no oficial de IA generativa para tareas rutinarias como redacción, síntesis de información y análisis preliminar de datos.

La emergente evidencia sobre este patrón de uso señala que los funcionarios públicos utilizan la IA generativa principalmente para realizar síntesis de texto, análisis de datos, exploración de ideas, redacción o corrección de documentos, aprender sobre un tema o realizar consultas. Además de los potenciales beneficios, este tipo de uso abre la puerta para nuevos retos de la rendición de cuentas del servicio público. Al respecto, un caso ilustrativo ocurrió en Brasil, un juez dictó una sentencia fundamentada en jurisprudencia inexistente, pero atribuida al Supremo Tribunal de Justicia de Brasil

(*La Nación*, 2023). De acuerdo con la nota periodística, el juez habría utilizado ChatGPT para redactar la sentencia y —se argumenta— no revisó todas las citas de jurisprudencia que el sistema de IA habría generado.

Los retos de la rendición de cuentas en este tipo de casos van más allá de la falla individual de verificación de la información, pues también entran en juego condiciones subyacentes de la propia tecnología (Rentería y Martínez, en prensa): la ausencia de transparencia y rastreabilidad epistémica, que impide identificar el origen de las citas; la falta de explicabilidad, puesto que los modelos producen secuencias lingüísticas plausibles pero no razonamientos jurídicos; la reproducción de sesgos narrativos y de representatividad, que puede inducir a la interpretación de la información bajo narrativas ajenas al contexto local; y, lo que caso en el caso brasileño ocurrió de forma evidente, la generación de información no confiable. Aunque la naturaleza tecnológica de la IA generativa difiere sustancialmente de la IA predictiva, los retos de la rendición de cuentas no desaparecen, sino que se reformulan y adquieren nuevas dimensiones (Rentería y Martínez, en prensa). Los tres desafíos ya conocidos —transparencia, explicabilidad y sesgos— resurgen bajo modalidades distintas y a ellos se suman otros retos: la reproducibilidad de los insumos de información para tomar decisiones, la generación de información no confiable y la dependencia de los usuarios en herramientas de IA generativa (cuadro VI.1).

Cuadro VI.1. Retos de la rendición de cuentas en el uso informal de la IA generativa

Reto	*Descripción*
Transparencia	Incapacidad de conocer o rastrear las fuentes de datos que respaldan la respuesta de un sistema. El escrutinio de los datos de entrenamiento de la IA generativa puede requerir infraestructura informática avanzada. Los derechos de propiedad intelectual sobre los modelos de IA generativa y datos de entrenamiento restringen el acceso a las fuentes de datos y a los parámetros entrenados.
Explicabilidad	Las respuestas de la IA generativa se basan en inferencias probabilísticas, más que en esquemas de razonamiento lógico-deductivo. Un modelo de IA generativa puede no ser capaz de generar respuestas consistentes, previsibles o confiables.

Cuadro VI.1 (*Continuación*)

Reto	*Descripción*
Reproducibilidad	La IA generativa produce respuestas con un grado de aleatoriedad.
Sesgo discriminatorio	Los datos de entrenamiento pueden no ser representativos de ciertas poblaciones. Reproducción de contenido socialmente inaceptable; o reproducción de narrativas culturalmente dominantes; o invisibilización de las cosmovisiones de grupos marginados.
Información no confiable	El modelo genera afirmaciones falsas, inexistentes o no verificables. El modelo genera una respuesta que no corresponde al objetivo informativo buscado por el usuario. El modelo genera respuestas desactualizadas, incompletas o superficiales. Incapacidad del modelo para recuperar y analizar la totalidad del corpus de datos disponible.
Dependencia respecto a la herramienta	Delegación impropia de decisiones de juicio y agencia profesional a los modelos de lenguaje. Deterioro de la experticia organizacional y capacidad deliberativa de los funcionarios públicos.

Fuente: Rentería y Martínez (en prensa).

A continuación, se presenta una breve descripción de cada uno de los retos.

Transparencia. Una diferencia sustancial de la IA generativa es su ausencia de rastreabilidad epistémica, entendida como la incapacidad del sistema para conocer o rastrear las fuentes de datos que respaldan una respuesta. Mientras que en la IA predictiva los revisores podían vincular el resultado de un modelo con un conjunto de datos concreto (conocido y auditable), los modelos generativos se entrenan con grandes repositorios de datos, heterogéneos y no estructurados (compuestos por texto, imágenes o video) (Mokander *et al.*, 2024; Zhao *et al.*, 2024). El reto de la transparencia muta en la IA generativa hacia lo que Zhao *et al*. (2024) describen como *opacidad epistémica*: la escala y diversidad de los conjuntos de datos de entrenamiento y de los parámetros, sumadas a las arquitecturas de inferencia de los modelos de lenguaje, hacen imposible rastrear las respuestas hasta fuentes de datos, causas o esquemas de razonamiento específicos. Los autores distinguen este fenómeno del ampliamente conocido problema de la caja negra. Desde una visión mecanicista de la explicabilidad, este problema se refiere a la incapacidad humana para comprender las operaciones internas de un sistema, es decir, la función que transforma entradas en salidas (Zhao *et al.*, 2024). Esta forma de explicabilidad es

aplicable a sistemas en los que el entrenamiento del modelo puede rastrearse en conjuntos de datos bien definidos y donde los expertos podrían conocer cómo los parámetros influyen en las inferencias estadísticas. Sin embargo, los modelos de IA generativa no permiten esta rastreabilidad retrospectiva de una afirmación a una fuente de datos concreta, pues los modelos de lenguaje generan respuestas mediante inferencias estadísticas (a partir de una vasta diversidad de datos), lo que hace epistemológicamente imposible ofrecer explicaciones causales o referenciales (Raza *et al.*, 2025). A ello se suma que gran parte de los datos de entrenamiento de los modelos comerciales más usados son de carácter propietario, lo que restringe el acceso de terceros (Esposito y Tse, 2024; Taeihagh, 2025).

Explicabilidad. Se debe partir del hecho de que las respuestas de los sistemas de IA generativa se basan en inferencias probabilísticas y no en esquemas de razonamiento lógico-deductivo. Esta condición limita fuertemente la capacidad de los modelos de lenguaje de fundamentar sus respuestas en un principio racional-legal, que se espera de los servidores públicos como justificación para sus acciones o decisiones administrativas. El reto central radica en la "ilusión conversacional" que los modelos de lenguaje generan en sus usuarios; los usuarios tienden a antropomorfizar a los sistemas de IA y atribuirles capacidades de comprensión, razonamiento o intención (Shanahan, 2024), cuando estos modelos no están diseñados tecnológicamente para razonar o interpretar intenciones, sino recrear estadísticamente patrones conversacionales. En consecuencia, para su uso en la función pública, conviene caracterizar a los modelos de lenguaje como herramientas estadísticas con rasgos conversacionales, pero no como agentes que comprenden o razonan en sentido humano (Shanahan, 2024).

Reproducibilidad. En la IA predictiva, la reproducibilidad de los resultados constituye un elemento esencial de la explicabilidad: si un modelo arroja el mismo resultado para los mismos valores ingresados, se asume que produce resultados consistentes, rastreables y procedimentalmente justos. En contraste, la IA generativa produce respuestas con

un grado de aleatoriedad introducido a través, por ejemplo, del parámetro de "temperatura". Esposito y Tse (2024) y Ulnicane y Erkkilä (2023) señalan que estas condiciones plantean un reto de rendición de cuentas no solo respecto de la equidad procedimental o la obligación de responder, sino también en términos de su confiabilidad, es decir, la capacidad del modelo para generar respuestas previsibles y consistentes.

Sesgo discriminatorio. En la IA predictiva, el desafío del sesgo discriminatorio se caracteriza en la literatura como una inequidad estructural cuantificable entre grupos (por ejemplo, disparidades en el acceso a beneficios sociales), lo que genera preocupación por decisiones discriminatorias que afectan directamente a la ciudadanía. Este es, quizá, el reto de rendición de cuentas más visibilizado en los medios de comunicación: entre una multitud de casos de decisiones discriminatorias asistidas por IA, se puede mencionar el escándalo de subsidios a la niñez en los Países Bajos, donde el sistema acusó erróneamente a madres solteras, migrantes y hogares de bajos ingresos de cometer fraude con subsidios para cuidado infantil (Arts y Van den Berg, 2025), o el caso de la Caja Nacional de Asignaciones Familiares en Francia, donde el uso indebido de sistemas de puntuación de riesgo tuvo efectos discriminatorios sobre poblaciones vulnerables (Amnesty International, 2024), por citar un par de ejemplos.

En contraste, el sesgo discriminatorio en la IA generativa no necesariamente se manifiesta en afectaciones directas sobre los ciudadanos, sino más bien en resultados indirectos que provienen de narrativas no representativas o sesgadas. Se trata de formas cualitativas y muy sutiles de sesgo —apenas perceptibles para algunos usuarios— en las que los modelos de lenguaje producen respuestas impregnadas de estructuras discursivas o de poder subyacentes, estadísticamente inferidas de los datos de entrenamiento (Sabour *et al.*, 2025; Satyanarayan y Jones, 2024). La falta de representatividad de los datos de entrenamiento ha sido señalada como una posible fuente de sesgo discriminatorio. Estudios recientes han demostrado que los modelos de lenguaje comerciales más utilizados se entrenan con grandes

repositorios de texto que presentan asimetrías estructurales en la representación de regiones geográficas y minorías (Manvi *et al.*, 2024; Mirza *et al.*, 2025). De manera específica, esta falta de representatividad se relaciona con un posible énfasis (o invisibilización) de ciertos temas o perspectivas en las respuestas de un modelo. Por ejemplo, en Estados Unidos, varias investigaciones periodísticas han señalado que los modelos de lenguaje presentan un menor desempeño en español (conocido como sesgo por lenguaje) y tienden a proporcionar menos información o información menos confiable a votantes latinos (conocido como sesgo de representación étnica) (Proof News, 2024).

La literatura reciente ha destacado una forma particular de sesgo en la IA generativa conocida como *sesgo narrativo*, un fenómeno que va más allá de las inexactitudes fácticas o de la desproporción representacional. El sesgo narrativo abarca tendencias sistemáticas en el tono, el estilo discursivo, la estructura argumentativa y las orientaciones de valor subyacentes que los modelos de lenguaje exhiben en sus respuestas (Ferrara, 2023; Jakesch *et al.*, 2023; Sabour *et al.*, 2025; Stahl y Eke, 2024). Este tipo de sesgo refleja cómo la IA generativa no solo responde a las consultas, sino que también enmarca los problemas, construye significados y se posiciona epistémicamente de maneras que pueden alinearse con supuestos ideológicos, culturales o normativos incrustados durante el entrenamiento del modelo. Algunos estudios han evidenciado este sesgo en modelos de lenguaje; por ejemplo, un informe de la UNESCO muestra que algunos representan a las mujeres en roles domésticos y reproducen estereotipos de género o raciales (UNESCO, 2024). Otro estudio encontró que algunos modelos estereotipan negativamente a las personas de ascendencia africana al asociarlas con empleos de menor prestigio o con mayor probabilidad de actividad delictiva (Hofmann *et al.*, 2024).

El concepto de sesgo narrativo amplía nuestra comprensión de la subjetividad de la IA generativa al señalar formas cualitativas y sutiles de sesgo discriminatorio que configuran no solo lo que se dice en la respuesta de un modelo, sino también cómo y desde qué perspectiva se dice. Aunque diferente en su naturaleza a la noción de sesgo algorítmi-

co, las consecuencias de los sesgos representacionales y narrativos son equivalentes: reproducción de estereotipos, invisibilización de experiencias o priorización de ciertos discursos culturales. Sin embargo, a diferencia del sesgo representacional o algorítmico (que puede diagnosticarse a través de análisis estadísticos o de disparidades demográficas), el sesgo narrativo requiere un escrutinio cualitativo del lenguaje, la retórica y el posicionamiento discursivo, lo que dificulta su detección y auditoría mediante las herramientas estándar de supervisión de IA.

Información no confiable. Otro de los retos centrales de la IA generativa —y quizá el más visible— es la producción de información no confiable. Aunque los sesgos narrativos y de representatividad pueden considerarse precursores de este fenómeno, la falta de fiabilidad responde también a limitaciones intrínsecas de los modelos de lenguaje. En la literatura, este problema suele denominarse *alucinación*, entendida como la generación de contenido ficticio, incorrecto o no respaldado por fuentes verificables (Banerjee *et al*., 2024; Maleki *et al*., 2024). No obstante, el desafío va más allá de la simple invención de hechos: incluye interpretaciones erróneas de instrucciones, respuestas débiles que resultan plausibles en apariencia pero analíticamente deficientes, o fallos en la recuperación de información disponible (Mokander *et al*., 2024; Zhao *et al*., 2024). Estas limitaciones son especialmente riesgosas en el sector público, donde la calidad de la información es condición para la toma de decisiones legítimas.

Dependencia respecto a la herramienta. Los sistemas de IA abren la puerta al deterioro de las capacidades y la agencia humana en funciones administrativas, donde la automatización desplaza el juicio y la discrecionalidad de los funcionarios públicos (Mokander, 2024; Raza *et al*., 2025). Además del riesgo de delegar agencia o discrecionalidad, debe reconocerse que los modelos de lenguaje no son herramientas pasivas, sino actores epistémicos que influyen sobre sus usuarios, ya que proporcionan marcos interpretativos, como se expuso arriba. Al mismo tiempo, los usuarios también influyen sobre el modelo a través de sus propios marcos interpretativos. Por tanto, debe tomarse en

cuenta que los modelos de lenguaje son entidades dinámicas que se adaptan a los patrones de interacción de sus usuarios, tanto a nivel individual como colectivo.

De esta manera, los modelos de lenguaje constituyen marcos de coconstrucción del conocimiento, donde la interacción iterativa entre usuario y modelo produce una interpretación de la realidad, en la que el modelo emerge como una entidad que orienta, estructura y convalida los prejuicios o las intenciones del usuario (Selten y Klievink, 2024). Una herramienta de trabajo con estas características conlleva implicaciones críticas para el sector público, ya que en muchas funciones administrativas (por ejemplo, la redacción de normas, manuales, resoluciones o actos administrativos), el modelo no solo asiste al funcionario público, sino que estructura su forma de pensar y, en consecuencia, las formas escritas del quehacer administrativo (Cantens, 2024). Incluso, los modelos de lenguaje pueden convertirse en un elemento clave en la configuración de la cultura organizacional, con un potencial significativo para influir en las formas de razonamiento profesional, institucional y ético (Selten y Klievink, 2024).

Estudios recientes han identificado al menos dos desafíos de rendición de cuentas derivados de este fenómeno (Cantens, 2024; Selten y Klievink, 2024): *a*) la delegación impropia de funciones y *b*) la desprofesionalización del servicio público. El primero se refiere a los casos en que un funcionario público delega al modelo tareas que implican juicio, interpretación o agencia profesional, sin mecanismos de supervisión humana validados institucionalmente. Esto deriva en un problema de legitimidad de las acciones administrativas, ya que se adoptan o aceptan acríticamente decisiones emitidas por entidades de automatización cognitiva.

El segundo desafío se relaciona con la desprofesionalización del servicio público, que está vinculada al deterioro paulatino del conocimiento experto. Representa una amenaza para la profesionalización y la preservación del conocimiento organizacional en la administración pública. Este riesgo ha sido discutido en algunos ámbitos como pérdida de competencias profesionales, desprofesionalización o *deskilling*, en referencia a la pérdida de los procesos de-

liberativos que constituyen el núcleo del conocimiento experto en la administración pública. La preocupación de fondo es que, si los funcionarios públicos ya no pueden explicar racionalmente los fundamentos de sus decisiones —porque estas son mediadas o sustituidas por la IA—, se erosiona de forma fundamental el principio de rendición de cuentas.

En suma, estos planteamientos teóricos muestran que el uso informal de la IA generativa en el sector público representa un desafío creciente para la legitimidad de su uso como herramienta de trabajo en sus servicios. La combinación de opacidad estructural, ausencia de explicabilidad, sesgos narrativos y la generación de información no confiable, convierte a estas herramientas en fuentes potenciales de errores o discriminación en la toma de decisiones. Al mismo tiempo, la evidente masificación de su uso en el espacio de trabajo (y los beneficios de su uso) exige que las entidades públicas canalicen esfuerzos para su incorporación regulada o modulada como herramienta de trabajo en el servicio público. Por un lado, este patrón de uso precisa de una estrategia organizacional de formalización de su uso, acompañada de políticas de uso y de transparencia algorítmica (Recomendación 4). Por otro, requiere de un trabajo de formación de capital humano en el uso responsable de la IA generativa como herramienta de trabajo (Recomendación 2). Los gobiernos deben dedicar esfuerzos en la provisión de información profesionalizante sobre la IA, es decir, desmitificar a la IA y sus usos, así como información que permita a los funcionarios identificar en qué actividades o tareas puede un sistema de IA agregar valor. Asimismo, los gobiernos deben fortalecer las capacidades de sus funcionarios públicos en el uso efectivo y responsable de la IA de propósitos generales, tomando en consideración los retos expuestos arriba.

Conclusiones

Este libro es útil para comprender y sistematizar el uso de sistemas de IA en el sector público mexicano, en un contexto donde la transparencia algorítmica y la gobernanza de estas tecnologías son temas fundamentales para garantizar su implementación ética y efectiva. La obra presenta un análisis sobre la falta de información clara y estandarizada sobre los sistemas de IA en las entidades gubernamentales, lo que permite una mayor comprensión de su impacto, riesgos y potencial para generar valor público. En un entorno donde el uso de sistemas de IA en la administración pública se expande rápidamente, se requiere de herramientas que permitan monitorearlos, evaluarlos y transparentarlos para fortalecer la confianza ciudadana y la rendición de cuentas. Además, el estudio presenta evidencia empírica que podría informar a un diseño de políticas públicas más efectivo en la adopción de IA, identificando buenas prácticas y desafíos estructurales para su implementación en México.

Esta obra se inserta dentro de un creciente cuerpo de literatura sobre transparencia algorítmica, una disciplina que ha ganado relevancia a medida que los gobiernos adoptan tecnologías de IA para optimizar la gestión pública. La revisión de la literatura contenida en el capítulo I se enfoca en la delimitación de las distintas dimensiones del concepto de sistemas de IA, y si bien no hay una definición precisa del término, se considera que es importante determinar su utilización en este trabajo desde una perspectiva sociotécnica, que

incluya no solo la dimensión tecnológica, sino también su impacto en el ámbito público que es donde se despliega. En este sentido se reconoce que, aunque la IA puede mejorar la eficiencia gubernamental y la prestación de servicios, presenta desafíos significativos en cuanto a opacidad, sesgos discriminatorios de sistemas de IA y falta de mecanismos de control adecuados, por lo cual es imprescindible comenzar a diseñarlos e implementarlos en los diversos niveles en los que estos sistemas se utilizan.

En el capítulo II se realiza una revisión de literatura sobre la utilización de los repositorios de sistemas de IA como herramienta de transparencia algorítmica en el sector público. Allí se hace una revisión de las experiencias internacionales y el uso de repositorios de sistemas de IA en ciudades como Ámsterdam y Helsinki, donde se ha demostrado que la documentación y monitoreo de tales sistemas permite a los gobiernos mitigar riesgos y garantizar que sean evaluados bajo criterios éticos por la ciudadanía. Sin embargo, en México, la falta de normas claras sobre el uso de la IA en el sector público ha dificultado la implementación de esquemas de transparencia algorítmica, explicabilidad y rendición de cuentas efectivos. Esta obra contribuye a la literatura sobre la materia al ofrecer una primera aproximación sistemática sobre cómo se están utilizando los sistemas de IA en la administración pública mexicana, al tiempo que se genera un diagnóstico que puede ser tomado como referencia para futuras reformas.

El estudio se basó en un diseño metodológico mixto que combinó la recolección de datos a partir de solicitudes de transparencia, monitoreo de medios, referencias directas y encuestas a funcionarios públicos responsables de las tecnologías en el gobierno federal. Esta estrategia permitió obtener una panorámica amplia y representativa del uso de los sistemas de IA en la administración pública mexicana, brindando información sobre sus características, distribución por nivel de gobierno y sectores donde se aplica. La metodología adoptada, que incluyó técnicas cualitativas y cuantitativas, permitió capturar tendencias en la adopción de la IA, y evaluar su impacto en distintas áreas del ámbito público en México.

Los datos ofrecidos y las conclusiones que se presentan deben ser analizados desde las limitaciones que enfrenta el estudio. En primer lugar, la falta de estandarización en la información proporcionada por las entidades públicas dificultó la comparación entre distintos desarrollos tecnológicos. Asimismo, la baja tasa de respuesta a las solicitudes de transparencia impidió obtener una visión completa del ecosistema de IA en el país, dejando fuera potenciales sistemas que podrían estar en uso. Otro desafío importante fue la dificultad para acceder a detalles técnicos sobre los sistemas de IA utilizados, lo que impide evaluar su nivel de autonomía, el grado de supervisión humana involucrado y su posible impacto en la toma de decisiones públicas. Estas limitaciones subrayan la necesidad de promover iniciativas de transparencia algorítmica y fortalecer la obligación de los organismos públicos de documentar y divulgar el uso de IA en sus procesos administrativos.

En el capítulo IV se revela información de valor para caracterizar el uso de estos sistemas en el ámbito público mexicano. Acerca de ello se puede resaltar que la IA se está utilizando principalmente en el poder ejecutivo y en los gobiernos estatales. Los sistemas de IA identificados se caracterizan por no ser sistemas complejos e impactan fundamentalmente en sectores como transparencia e integridad y seguridad ciudadana y justicia, particularmente en el caso de las entidades del sector público. No obstante, resalta que la adopción de IA a nivel municipal es muy baja, y que persisten disparidades significativas entre entidades federativas donde la adopción es heterogénea. Estas brechas reflejan diferencias significativas en la capacidad institucional, la disponibilidad de recursos financieros y la existencia de estrategias digitales dentro de los gobiernos locales.

Un hallazgo relevante es la tendencia hacia el desarrollo interno de sistemas de IA dentro del sector público. Aunque en el ámbito internacional es común que las administraciones públicas contraten proveedores privados para el desarrollo de soluciones tecnológicas, en México, de acuerdo a la información a la que se tuvo acceso, una cantidad considerable de sistemas ha sido creada por equipos técnicos gubernamentales. Este fenómeno responde a una cultura de

soberanía tecnológica y al interés de reducir costos a largo plazo, aunque también enfrenta desafíos como la falta de personal capacitado y la necesidad de mejorar las condiciones para la retención de talento especializado en IA dentro del sector público.

El capítulo V aporta lecciones valiosas para los tomadores de decisiones en el sector público sobre las tareas y los rubros en que los sistemas de IA pueden generar valor público. Sistemas de IA como el SIDIL, Siempre Seguras y Regresa UNAM muestran que el uso de la IA puede generar un impacto positivo en áreas como la fiscalización laboral, la seguridad ciudadana y la búsqueda de personas desaparecidas. Sin embargo, estos casos subrayan la importancia de garantizar la transparencia en el diseño y operación de estas tecnologías, asegurando que los ciudadanos comprendan cómo funcionan y que existen mecanismos de control y supervisión adecuados. La replicabilidad de estas soluciones en otros contextos gubernamentales dependerá de la existencia de marcos normativos claros y de la capacidad de las instituciones para adaptar la tecnología a sus necesidades específicas.

Las recomendaciones de política pública establecidas en el capítulo VI son fundamentales para orientar una adopción más responsable y efectiva de los sistemas de IA en el sector público. Entre estas recomendaciones destacan la necesidad de promover la difusión de información sobre IA tanto entre la ciudadanía como en el funcionariado público, el fortalecimiento de capacidades institucionales en ciencia de datos y transparencia algorítmica, y la creación de un ecosistema de IA en la innovación pública. También se enfatiza la importancia de documentar y compartir casos de éxito para facilitar la replicabilidad de soluciones tecnológicas en distintas entidades gubernamentales.

Un aspecto clave de las recomendaciones es la necesidad de establecer marcos regulatorios claros que definan criterios de transparencia, supervisión y auditoría de los sistemas de IA utilizados en el sector público. Sin estos mecanismos la opacidad en el uso de IA puede generar riesgos significativos, desde el sesgo discriminatorio de los sistemas de IA hasta la falta de rendición de cuentas en decisiones automatizadas. Por ello, se debe considerar la creación de me-

canismos de transparencia algorítmica y supervisión para evaluar el impacto de la IA en la administración pública y garantizar su uso responsable.

Esta obra constituye un primer esfuerzo sistemático para mapear y analizar el uso de los sistemas de IA en el sector público mexicano, proporcionando una base empírica para la formulación de políticas públicas en la materia. Sus hallazgos y recomendaciones pueden contribuir a mejorar la gobernanza de la IA y fortalecer los principios de transparencia y rendición de cuentas en la administración pública mexicana. La incorporación de mecanismos de supervisión y regulación, así como la generación de capacidades técnicas y normativas, será clave para garantizar que la IA se utilice de manera ética y eficaz en beneficio de la sociedad.

Referencias

Ada Lovelace Institute, AI Now Institute, y Open Government Partnership. (2021). *Algorithmic Accountability for the Public Sector*. https://www.opengovpartnership.org/documents/algorithmic-accountability-public-sector/

Agencia de Gobierno Electrónico, y Sociedad de la Información y el Conocimiento. (AGESIC). (s. f.). *Mirador de Gobierno Abierto*. https://miradordegobiernoabierto.agesic.gub.uy/SigesVisualizador/gu/o/GA/p/2264

Agencia de Transformación Digital y Telecomunicaciones (ATDT). (2025). Centro Nacional de Tecnología Pública. Soberanía tecnológica nacional. https://www.gob.mx/atdt/agendas/centro-nacional-de-tecnologia-publica

Agencia Digital de Innovación Pública de la Ciudad de México. (s. f.). Colaboraciones. https://www.adip.cdmx.gob.mx/oficina-de-cooperacion/colaboraciones

Agesic-Gobierno de Uruguay. (s. f.). *Mirador de Gobierno Abierto*. https://miradordegobiernoabierto.agesic.gub.uy/SigesVisualizador/gu/o/GA/p/2264

AI Now Institute. (2020). *Government AI accountability strategies*. AI Now Institute.

Algorithm Register. (s. f.). Algorithm Register Europe. https://www.algorithmregister.org/

Algoritmos Públicos Chile. (s. f.). Repositorio de algoritmos públicos en Chile. https://www.algoritmospublicos.cl/repositorio

Alishani, A., Homburg, V., y Velsberg, O. (2025). Public encounters and government chatbots: When servers talk to citizens. *Public Administration Review*.

American Institute for Research (AIR). (2025). *SIDIL: Manual del procesamiento y explotación de la información*. Recuperado de https://capacitacionstps.com/wp-content/uploads/2025/01/Sistema-de-Inteligencia-de-Datos-para-la-Inspeccion-Laboral.pdf.

Amnesty International. (2024). France: Discriminatory algorithm used by the social security agency must be stopped. *Amnesty International*. https://www.amnesty.org/en/latest/news/2024/10/france-discriminatory-algorithm-used-by-the-social-security-agency-must-be-stopped/

Alon-Barkat, S., y Busuioc, M. (2022). Human-AI interactions in public sector decision-making: "Automation bias" and "selective adherence" to algorithmic advice. *Journal of Public Administration Research and Theory*, *33*(1).

Arnesen, S., Broderstad, T. S., Fishkin, J. S., Johannesson, M. P., y Siu, A. (2024). Knowledge and support for AI in the public sector: A deliberative poll experiment. *AI y Society*. https://doi.org/10.1007/s00146-024-02104-w

Arts, J., y Van den Berg, M. (2025). What the Dutch benefits scandal and policy's focus on 'fraud' can teach us about the endurance of empire. *Critical Social Policy*, *45*(1), 177-187.

Association for Computing Machinery (ACM). (2017). *Statement on Algorithmic Transparency and Accountability*. ACM U.S. Public Policy Council y ACM Europe Policy Committee. https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf

Banerjee, S., Agarwal, A., y Singla, S. (2024). *LLMs will always hallucinate, and we need to live with this*. En arXiv [stat.ML]. arXiv. http://arxiv.org/abs/2409.05746

Banco Internacional de Desarrollo (BID). (2020). *Documentos de Marco Sectorial (SFD)*. División de Innovación para servir al ciudadano. https://idbinvest.org/es/sectores#:~:text=Los%20Documentos%20de%20Marco%20Sectorial,de%20en%20un%20sector%20espec%C3%ADfico

Bowles, J. (2020). *The limits of legibility: How distributive conflicts constrain state-building*. Work.

Bozkurt, V., y Gursoy, D. (2025). The artificial intelligence paradox: Opportunity or threat for humanity? *International Journal of Human–Computer Interaction*, *41*(1), 174-187.

Brambor, T., Goenaga, A., Lindvall, J., y Teorell, J. (2020). The lay of the land: Information capacity and the modern state. *Comparative Political Studies*, *53*(2), 175-213.

Brauner, P., Hick, A., Philipsen, R., y Ziefle, M. (2023). What does the public think about artificial intelligence? -A criticality map to understand bias in the public perception of AI. *Frontiers in Computer Science*, (5), 1113903.

Bright, J., Enock, F. E., Esnaashari, S., Francis, J., Hashem, Y., y Morgan, D. (2024). *Generative AI is already widespread in the public sector*. En arXiv [cs. CY]. arXiv. http://arxiv.org/abs/2401.01291

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, *3*(1). https://doi.org/10.1177/2053951715622512

Cantens, T. (2024). How will the state think with ChatGPT? The challenges of generative artificial intelligence for public administrations. *AI y Society*, (40). https://doi.org/10.1007/s00146-023-01840-9

Centro Latinoamericano de Administración para el Desarrollo (CLAD). (2023). *Carta Iberoamericana de Inteligencia Artificial en la Administración Pública*. https://clad.org/wp-content/uploads/2023/10/Borrador-CIIA-V1-ES-08-2023.pdf

Centro México Digital. (2024). *Principios para el desarrollo confiable, responsable y seguro de la inteligencia artificial en México*. Centro México Digital. https://centromexico.digital/wp-content/uploads/2024/06/DesarrolloResponsabledeIA.pdf

City of Amsterdam. (s. f.). *Algorithm Register*. https://algoritmeregister.amsterdam.nl/

City of Helsinki. (s. f.). *Artificial Intelligence Register*. https://ai.hel.fi/

City of New York. (2023). *Algorithmic Tools Reporting*. https://www.nyc.gov/assets/oti/downloads/pdf/reports/2023-algorithmic-tools-reporting-updated.pdf

Clarke, A. (2020). Digital government units: What are they, and what do they mean for digital era public management renewal? *International Public Management Journal*, *23*(3), 358-379.

Cox, M., Roberts, M., y Walton, J. (2012). IT outsourcing in the public sector local government: experiences of the management and selection of IT service providers. *Electronic Journal of Information Systems Evaluation*, *15*(3), 231-243.

Comisión Económica para América Latina y el Caribe (CEPAL). (2024). *Superar las trampas del desarrollo de América Latina y el Caribe en la era digital: El potencial transformador de las tecnologías digitales y la inteligencia artificial*. (LC/CMSI.9/3). CEPAL.

Comisión Europea. (2025, 29 de julio). *Comunicación de la Comisión relativa a las directrices sobre la definición de sistema de inteligencia artificial establecida en el Reglamento (UE) 2024/1689 (Reglamento de IA) (C(2025) 5053 final)*. https://digital-strategy.ec.europa.eu/es/library/commission-publishes-guidelines-ai-system-definition-facilitate-first-ai-acts-rules-application

Comisión Europea. (2024). *Public sector tech watch: Adoption of AI, blockchain and other emerging technologies within the European public sector*. Luxemburgo: Publications Office of the European Union. https://doi.org/10.2799/4393

Comisión Europea. (2016). Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos. *Diario Oficial de la Unión Europea*. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A32016R0679

Congreso de la Unión. (2025a, 3 de marzo). Ley Federal de Protección de Datos Personales en Posesión de los Particulares. *Diario Oficial de la Federación*.

Congreso de la Unión. (2025c, 3 de marzo). Ley General de Transparencia y Acceso a la Información Pública. *Diario Oficial de la Federación*.

Congreso de la Unión (2025d, 16 de abril). Ley de adquisiciones, arrendamientos y servicios del sector público. *Diario Oficial de la Federación*.

Congreso de la Unión. (2024a, 26 de enero). Ley General de Mecanismos Alternativos de Solución de Controversias, Nueva Ley publicada en el *Diario Oficial de la Federación*.

Congreso de la Unión. (2024b, 24 de enero). Decreto por el que se adicionan diversas disposiciones de la Ley Federal del Trabajo, en materia de Plataformas Digitales. *Diario Oficial de la Federación*.

Congreso de la Unión. (2000, 4 de enero). Ley de Adquisiciones, arrendamientos y servicios del sector público. *Diario Oficial de la Federación*. Última reforma publicada en el DOF publicada DOF 20-05-2021.

Connolly, P. (2025). Conversations with chatbots. En P. Connolly, S. Goldberg y J. Saul, *Conversations online: Explorations in philosophy of language* (pp. 1-35). Oxford University Press.

Consejo de Europa. (2024). *Convención Marco sobre Inteligencia Artificial y Derechos Humanos, Democracia y Estado de Derecho*. https://www.coe.int/en/web/conventions/full-list?module=treaty-detailytreatynum=225

Consejo para la Transparencia de Chile. (2024). *Informe de recomendaciones al Congreso*. https://www.consejotransparencia.cl/wp-content/uploads/instruccion/2024/09/Informe-recomendaciones-congreso-CPLT.pdf

Constitución Política de los Estados Unidos Mexicanos. (2025). Publicada en el *Diario Oficial de la Federación* el 5 de febrero de 1917. Última reforma publicada en el DOF 17-01-2025.

Coordinación de la Estrategia Digital Nacional. (2021, 6 de septiembre). Acuerdo por el que se emiten las políticas y disposiciones para impulsar el uso y aprovechamiento de la informática, el gobierno digital, las tecnologías de la información y comunicación, y la seguridad de la información en la Administración Pública Federal. *Diario Oficial de la Federación*.

Criado, J. I., Alcaide-Muñoz, L., y Liarte, I. (2025). Two decades of public sector innovation: Building an analytical framework from a systematic literature review of types, strategies, conditions, and results. *Public Management Review*, *27*(3), 623-652.

Criado, J. I., Sandoval-Almazán, R., y Gil-Garcia, J. R. (2024). Artificial intelligence and public administration: Understanding actors, governance, and policy from micro, meso, and macro perspectives. *Public Policy and Administration*, *40*(2), 173-184. https://doi.org/10.1177/09520767241272921

Cuenta de X (@CONSAR_mx). (2024, 2 de mayo). [Publicación en X]. X. https://x.com/CONSAR_mx/status/1786175927694598434

Cuenta de X (@SESNAOficial). (2024, 2 de julio). [Publicación en X]. X. https://x.com/SESNAOficial/status/1808263737809236228

Cuenta de X (@LaAgenciaCDMX). (2024, 30 de julio). [Publicación en X]. X. https://x.com/LaAgenciaCDMX/status/1818340145247957126

Cuenta de X (@PJEdomex). (2024, 1 de diciembre). [Publicación en X]. X. https://x.com/PJEdomex/status/1863230260143071591

Cumbre de seguridad de la IA. (2023). *Declaración de Bletchley. 1 y 2 de noviembre de 2023*. https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/dbc58681-1b68-47e0-8e3f-f91435fdf8ce

Cumbre Ministerial y de Altas Autoridades de América Latina y el Caribe Santiago de Chile. (2023). *Declaración de Santiago, "Para promover una inteligencia artificial ética en América Latina y el Caribe"*. Santiago de Chile, 23-24 de octubre de 2023. https://minciencia.gob.cl/uploads/filer_public/40/2a/402a35a0-1222-4dab-b090-5c81bbf34237/declaracion_de_santiago.pdf

D'Arcy, M., y Nistotskaya, M. (2017). State first, then democracy: Using cadastral records to explain governmental performance in public goods provision. *Governance*, *30*(2), 193-209.

Dalton, R. J. (2008). Citizenship norms and the expansion of political participation. *Political Studies*, *56*(1), 76-98.

Del Pozo, C., Nuno Gomes de Andrade, N., y Rojas Arroyo, D. (2023). *Prototipo de Políticas Públicas sobre Transparencia y Explicabilidad de Sistemas de Inteligencia Artificial*. https://openloop.org/reports/2023/08/public-policy-prototype-on-the-transparency-and-explainability-of-ai-systems-es.pdf

Dunleavy, P., y Margetts, H. Z. (2010). The second wave of digital was governance. En *APSA 2010 Annual Meeting Paper*.

Dunleavy, P., Margetts, H., Bastow, S., y Tinkler, J. (2006). New public management is dead-long live digital-era governance. *Journal of Public Administration Research and Theory*, *16*(3), 467-494.

Edler, J., Blind, K., Kroll, H., y Schubert, T. (2023). Technology sovereignty as an emerging frame for innovation policy. Defining rationales, ends and means. *Research Policy*, *52*(6), 104765.

El Economista. (2024, 19 de agosto). CDMX e Hidalgo fortalecen cooperación regional con enfoque en la digitalización. *El Economista*. https://www.eleconomista.com.mx/estados/CDMX-e-Hidalgo-fortalecen-cooperacion-regional-con-enfoque-en-la-digitalizacion-20240819-0074.html

El País. (2024, 9 de octubre). El 85% de los europeos no sabe cómo utilizar la inteligencia artificial en la vida cotidiana. *El País*. https://elpais.com/tecnologia/2024-10-09/el-85-de-los-europeos-no-sabe-como-utilizar-la-inteligencia-artificial-en-la-vida-cotidiana.html

Esposito, M., y Tse, T. (2024). Mitigating the risks of generative AI in government through algorithmic governance. En *Proceedings of the 25th Annual International Conference on Digital Government Research* (pp. 605-609). Nueva York: Association for Computing Machinery.

European Commission. (2025, 26 de marzo). *Daily news: 26 March 2025*. https://ec.europa.eu/commission/presscorner/detail/en/mex_25_912

European Commission. (s. f.). *Public sector tech watch*. https://joinup.ec.europa.eu/collection/public-sector-tech-watch

European Union. (2024). *AI and GenAI Adoption by Local and Regional Administrations*. Publications Office of the European Union. https://op.europa.eu/en/publication-detail/-/publication/40363d58-bdc8-11ef-91ed-01aa75ed71a1/language-en

Ernst y Young. (2024). *EY Pulse Survey: Insights into the integration of AI in government*. https://www.ey.com/en_us/industries/government-public-sector/insights-into-the-integration-of-ai-in-government

Ferrara, E. (2023). *Should ChatGPT be biased? Challenges and risks of bias in large language models*. En arXiv [cs.CY]. arXiv. http://arxiv.org/abs/2304.03738

Fisher, S., y Rosella, L. C. (2022). Priorities for successful use of artificial intelligence by public health organizations: A literature review. *BMC Public Health*, *22*(1), 2146.

G7 Italia. (2024). *G7 Toolkit for Artificial Intelligence in the Public Sector*. Report Prepared for the 2024 Italian G7 Presidency and the G7 Digital and Tech Working Group.

Gemeente Amsterdam. (s. f.). *Algorithm Register Amsterdam*. https://algoritmeregister.amsterdam.nl/

Gerlich, M. (2024). Public anxieties about AI: Implications for corporate strategy and societal impact. *Administrative Sciences*, *14*(11), 288.

Global Partnership on Artificial Intelligence (GPAI). (2024a). *Algorithmic transparency in the public sector: Recommendations for Governments to Enhance the Transparency of Public Algorithms*. Report, November 2024, Global Partnership on Artificial Intelligence.

Gobierno de Chile. (s. f.). Repositorio de Algoritmos Públicos. https://www.algoritmospublicos.cl/repositorio

Gobierno de Colombia. (s. f.). Repositorio de Algoritmos Estatales. https://www.algoritmosestatales.org/

Gobierno de Jalisco. (2024, 22 de julio). Acuerdo del Gobernador Constitucional del Estado de Jalisco, mediante el cual se emiten las Políticas y Disposiciones Generales en materia de IA y Gobernanza de Datos de la Administración Pública del Gobierno del Estado de Jalisco. *Estado de Jalisco Periódico Oficial*.

Gobierno de la Ciudad de México. (2022, enero). *Convenio con el Estado de Veracruz*. Gobierno de la Ciudad de México. https://gobierno.cdmx.gob.mx/noticias/convenio-con-el-estado-de-veracruz/

Gobierno del Estado de Baja California. (s. f.). *Fortalecen lazos de colaboración la Agencia Digital de BC y la Agencia Digital de Innovación Pública de la CDMX*. https://www.bajacalifornia.gob.mx/Prensa/Noticia/17845

Government Digital Service. (2024). *State of digital government review*. GOV.UK. https://www.gov.uk/government/publications/state-of-digital-government-review/state-of-digital-government-review

Government Digital Service. (2025). *Same name, new ambitions*. [Blog]. https://gds.blog.gov.uk/2025/01/27/same-name-new-ambitions/

Government Digital Service. (s. f.). *State of digital government review*. https://www.gov.uk/government/publications/state-of-digital-government-review/state-of-digital-government-review

Government Digital Service, Department for Science, Innovation and Technology, y Central Digital and Data Office (CDDO). (2024). *Algorithmic Transparency Recording Standard Hub*. https://www.gov.uk/government/collections/algorithmic-transparency-recording-standard-hub

Government of Canada. (2021). *Directive on Automated Decision-Making*. Treasury Board of Canada Secretariat.

Global Partnership on Artificial Intelligence (GPAI). (2024b). *Algorithmic Transparency in the Public Sector: Case studies of repositories of public algorithms in Chile, the EU and the UK*. Report, November 2024, Global Partnership on Artificial Intelligence.

Grimmelikhuijsen, S. (2023). Explaining why the computer says no: Algorithmic transparency affects the perceived trustworthiness of automated decision-making. *Public Administration Review*, *83*(2), 241-262.

Grimmelikhuijsen, S., y Meijer, A. (2022). Legitimacy of algorithmic decision-making: Six threats and the need for a calibrated institutional response. *Perspectives on Public Management and Governance*.

Grimmelikhuijsen, S., y Tangi, L. (2024). *What factors influence perceived artificial intelligence adoption by public managers?* (No. JRC138684). Joint Research Centre.

Gutiérrez, J. D. (2024). *Consultation paper on AI regulation: Emerging approaches across the world*. UNESCO.

Gutierrez, C. I., Aguirre, A., Uuk, R., Boine, C. C., y Franklin, M. (2023). A proposal for a definition of general purpose artificial intelligence systems. *Digital Society*, *2*(3), 36.

Gutiérrez, J. D., Muñoz-Cadena, S., y Castellanos-Sánchez, M. (2023). *Sistemas de decisión automatizada en el sector público colombiano v. 1.0.xlsx*. Universidad del Rosario. https://doi.org/10.34848/YN1CRT/8OHRT0

Haataja, M., Van de Fliert, L., y Rautio, P. (2020). *Public AI Registers: Realising AI Transparency and Civic Participation in Government Use of AI*. https://ai.hel.fi/wp-content/uploads/White-Paper.pdf

Haugen, T. B., y Klungseth, N. J. (2017). In-house or outsourcing FM services in the public sector: A review of 25 years research and development. *Journal of Facilities Management*, *15*(3), 262-284.

Hernández, G. (2024). STPS usará IA para definir inspecciones de trabajo. *El Economista*. https://www.eleconomista.com.mx/capital-humano/stps-usara-ia-definir-inspecciones-trabajo-20241202-736437.html

Hofmann, V., Kalluri, P. R., Jurafsky, D., y King, S. (2024). AI generates covertly racist decisions about people based on their dialect. *Nature*, *633*(8028), 147-154.

IA2030Mx. (2020). Agenda Nacional Mexicana de Inteligencia Artificial. https://wp.oecd.ai/app/uploads/2022/01/Mexico_Agenda_Nacional_Mexicana_de_IA_2030.pdf

Instituto Nacional de Estadística y Geografía (INEGI). (2024). *Producto Interno Bruto por entidad federativa PIB 2023*. https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2024/PIBEF/PIBEF2023.pdf

IBM Center for the Business of Government. (2025). *AI works: How public sector agencies are using AI today to deliver better government*. https://www.ibm.com/downloads/cas/Z6ENJQY1

Info CDMX. (2024). *Acuerdo mediante el cual se aprueba la iniciativa con proyecto de decreto para expedir la Ley para el uso de Inteligencia Artificial y el tratamiento de datos personales por sujetos obligados en la Ciudad de México. Acuerdo 5038/SO/03-10/2024*. Instituto de Transparencia, Acceso a la Información Pública, Protección de Datos Personales y Rendición de Cuentas de la Ciudad de México. https://documentos.infocdmx.org.mx/acuerdos/2024/A121Fr01_2024-T04_Acdo_2024-03-10-5038.pdf

International Association of Privacy Professionals. (IAPP). (2024). *AI Governance in Practice Report 2024*. https://iapp.org/media/pdf/resource_center/ai_governance_in_practice_report_2024.pdf

International Organization for Standardization, y International Electrotechnical Commission. (2022). *ISO/IEC 22989:2022 – Information technology – Artificial intelligence – Artificial intelligence concepts and terminology*. ISO. https://www.iso.org/standard/74296.html

Jakesch, M., Bhat, A., Buschek, D., Zalmanson, L., y Naaman, M. (2023). *Co-writing with opinionated language models affects users' views*. Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, 1-15.

Janssen, M., y Van Der Voort, H. (2016). Adaptive governance: Towards a stable, accountable and responsive government. *Government Information Quarterly*, *33*(1), 1-5.

Japan AI Safety Institute. (2024). *Guide to Red Teaming Methodology on AI Safety (Version 1.00)*. https://aisi.go.jp/assets/pdf/ai_safety_RT_v1.00_en.pdf

Jobin, A., y Katzenbach, C. (2023). The becoming of AI: A critical perspective on the contingent formation of AI. En S. Lindgren (Ed.), *Handbook of critical studies of artificial intelligence*. Edward Elgar Publishing.

Jung, H., y Camarena, L. (2024). Street-level bureaucrats y AI interactions in public organizations: An identity based framework. *Public Performance y Management Review*, 1-30.

Kempeneer, S., y Heylen, F. (2023). Virtual state, where are you? A literature review, framework and agenda for failed digital transformation. *Big Data y Society*, *10*(1). https://doi.org/10.1177/20539517231160528

Kettl, D. F. (2017). *Can governments earn our trust?* John Wiley y Sons.

Kim, S., Andersen, K. N., y Lee, J. (2022). Platform government in the era of smart technology. *Public Administration Review*, *82*(2), 362-368.

Kimbell, L., y Vesnić-Alujević, L. (2020). After the toolkit: Anticipatory logics and the future of government. *Policy Design and Practice*, *3*(2), 95-108.

Lagunes, A., Martínez, Y., Cárdenas, C., De la Peña, S., Mancilla, D., Xilotl, R., Sánchez, O., Moguel, A., y Cárdenas, J. (2024). *Propuesta de Agenda Nacional de la Inteligencia Artificial para México 2024-2030*. Alianza Nacional de Inteligencia Artificial. https://www.ania.org.mx/_files/ugd/447d95_ae0e425d83644651bd19ab4acdfd9007.pdf

La Nación. (2023, 15 de noviembre). Insólito: investigan si un juez brasileño usó a ChatGPT para redactar una sentencia y la publicó con errores. *La Nación*. https://www.lanacion.com.ar/tecnologia/insolito-investigan-si-un-juez-brasileno-uso-a-chatgpt-para-redactar-una-sentencia-y-la-publico-con-nid15112023/

Lee, M. M., y Zhang, N. (2017). Legibility and the informational foundations of state capacity. *The Journal of Politics*, *79*(1), 118-132.

Lee, J. D., Kim, H., Si, S., y Lee, S. (2024). Techno-nationalism to collaborative technology sovereignty. *Science y Public Policy*, *51*(6), 1185-1190. https://academic.oup.com/spp/article-abstract/51/6/1185/7739935

Li, W., Xu, S., Zheng, X., y Sun, R. (2024). Bridging the knowledge gap in artificial intelligence: The roles of social media exposure and information elaboration. *Science Communication*, *46*(4), 399-430.

Luitse, D. (2024). Platform power in AI: The evolution of cloud infrastructures in the political economy of artificial intelligence. *Internet Policy Review*, *13*(2). https://doi.org/10.14763/2024.2.1768

Madan, R., y Ashok, M. (2023). AI adoption and diffusion in public administration: A systematic literature review and future research agenda. *Government Information Quarterly*, *40*(1), 101774.

Maffei, S., Leoni, F., y Villari, B. (2020). Data-driven anticipatory governance. Emerging scenarios in data for policy practices. *Policy Design and Practice*, *3*(2), 123-134.

Maleki, N., Padmanabhan, B., y Dutta, K. (2024). *AI hallucinations: A misnomer worth clarifying*. En arXiv [cs.CL]. arXiv. http://arxiv.org/abs/2401.06796

Manvi, R., Khanna, S., Burke, M., Lobell, D., y Ermon, S. (2024). *Large Language Models are Geographically Biased*. En arXiv [cs.CL]. arXiv. http://arxiv.org/abs/2402.02680

Margetts, H., y Dorobantu, C. (2019). Rethinking public policy in the era of AI: A comparative analysis of AI policy in the public sector. *Journal of European Public Policy*, *26*(10), 1417-1433.

Medaglia, R., Mikalef, P., y Tangi, L. (2024). *Competences and governance practices for artificial intelligence in the public sector*. Luxemburgo: Publications Office of the European Union.

Meijer, A., Lorenz, L., y Wessels, M. (2021). Algorithmization of Bureaucratic Organizations: Using a Practice Lens to Study How Context Shapes Predictive Policing Systems. *Public Administration Review*, *81*(5), 837-846.

Mergel, I. (2015). Opening government: Designing open innovation processes to collaborate with external problem solvers. *Social Science Computer Review*, *33*(5), 599-612.

Mergel, I. (2017). *Digital service teams: Challenges and recommendations for government*. IBM Center for The Business of Government.

Mergel, I. (2019). Digital service teams in government. *Government Information Quarterly*, *36*(4), 101389.

Mergel, I., Dickinson, H., Stenvall, J., y Gasco, M. (2024). Implementing AI in the public sector. *Public Management Review*, 1-14.

Merino, J., y Muñoz del Olmo, M. (2025). *Transformación digital: el modelo de la Ciudad de México*. Washington, D.C.: Banco Interamericano de Desarrollo.

Michel, G. (2024). Sobrepasar la "crisis de ideas" para un gobierno del siglo XXI. En D. Gómez-Álvarez y C. Rentería (Coords.), *Pensar fuera de la caja burocrática: laboratorios de innovación pública en América Latina* (pp. 149-178). Universidad de Guadalajara/Debate.

Milkman, R. (2017). A new political generation: Millennials and the post-2008 wave of protest. *American Sociological Review*, *82*(1), 1-31.

Millan Vargas, A., y Sandoval-Almazan, R. (2024). *Public managers perception on artificial intelligence: the case of the State of Mexico*. https://scholarspace.manoa.hawaii.edu/items/3acc6722-f1ed-4706-ae1d-9e15bbcbbf4d

Ministerio de Asuntos Económicos y Transformación Digital. (2021). *Carta de Derechos Digitales del Gobierno Español*. https://portal.mineco.gob.es/RecursosNoticia/mineco/prensa/ficheros/Carta_Derechos_Digitales_RedEs.pdf

Mirza, I., Jafari, A. A., Ozcinar, C., y Anbarjafari, G. (2025). Quantifying gender bias in large language models using information-theoretic and statistical analysis. *Information (Basel)*, *16*(5), 358.

Mokander, J., Curl, J., y Kshirsagar, M. (2024). *A blueprint for auditing generative AI*. En arXiv [cs.CY]. arXiv. https://doi.org/10.2139/ssrn.4887768

Moon, M. J. (2002). The evolution of e-government among municipalities: rhetoric or reality? *Public Administration Review*, *62*(4), 424-433.

Moon, J., Choe, Y. C., Chung, M., Jung, G. H., y Swar, B. (2016). IT outsourcing success in the public sector: Lessons from e-government practices in Korea. *Information Development*, *32*(2), 142-160.

Nam, J., y Bell, E. (2024). Efficiency or equity? How public values shape bureaucrats' willingness to use artificial intelligence to reduce administrative burdens. *Public Performance & Management Review*, *48*(6), 1268-1301. https://doi.org/10.1080/15309576.2024.2419132

Nantes Métropole. (s. f.). *Les algorithmes à Nantes Métropole*. https://data.nantesmetropole.fr/pages/algorithmes_nantes_metropole/

Nantes Métropole. (s. f.). *Registre des algorithmes*. https://data.nantesmetropole.fr/pages/algorithmes_nantes_metropole/

National AI Initiative Office. (2020). *National AI Initiative Act*. U.S. Government.

National Institute of Standards and Technology. (2019). *U.S. leadership in AI: A plan for federal engagement in developing technical standards and related tools*. Prepared in response to Executive Order 13859. U.S. Department of Commerce. https://doi.org/10.6028/NIST.AI.100-1

National Telecommunications and Information Administration (NTIA). (2024). *Artificial Intelligence Accountability Policy Report*. https://www.ntia.gov/sites/default/files/publications/ntia-ai-report-final.pdf

Neri, H., y Cozman, F. (2020). The role of experts in the public perception of risk of artificial intelligence. *AI & Society*, *35*(3), 663-673.

Neumann, O., Guirguis, K., y Steiner, R. (2024). Exploring artificial intelligence adoption in public organizations: a comparative case study. *Public Management Review*. https://doi.org/10.1080/14719037.2022.2048685

Nicolás, M. A., y Sampaio, R. C. (2024). Balancing efficiency and public interest: The impact of AI automation on social benefit provision in Brazil. *Internet Policy Review*, *13*(3).

Noonia, A., Beg, R., Patidar, A., Bawaskar, B., Sharma, S., y Rawat, H. (2024). Chatbot vs intelligent virtual assistance (IVA). En *Conversational Artificial Intelligence* (pp. 655-673). Wiley.

Norris, P., y Inglehart, R. F. (2018). *Understanding Brexit: Cultural resentment versus economic grievances*. Paper for presentation at the Panel on 'Populism in Advanced Capitalist Democracies', Thursday, 30, 4-00.

NYC Automated Decision Systems Task Force. (2019). *Automated decision systems task force report*. *NYC Government*. https://www.nyc.gov/assets/adstaskforce/downloads/pdf/ADS-Report-11192019.pdf

NYC Office of Technology and Innovation. (2023). *Algorithmic tools reporting update*. NYC Government. https://www.nyc.gov/assets/oti/downloads/pdf/reports/2023-algorithmic-tools-reporting-updated.pdf

O'Shaughnessy, M. R., Schiff, D. S., Varshney, L. R., Rozell, C. J., y Davenport, M. A. (2023). What governs attitudes toward artificial intelligence adoption and governance? *Science and Public Policy*, *50*(2), 161-176.

Organisation for Economic Co-operation and Development OCDE. (2019). *Recommendation of the Council on Artificial Intelligence*. OECD Legal Instruments.

Organisation for Economic Co-operation and Development OCDE. (2021). *AI in the Public Sector*. OECD AI Policy Observatory.

Organisation for Economic Co-operation and Development OCDE. (2022a). *Uso estratégico y responsable de la inteligencia artificial en el sector público de América Latina y el Caribe*. (Estudios de la OCDE sobre Gobernanza Pública). París: OECD Publishing. https://doi.org/10.1787/5b189cb4-es

Organisation for Economic Co-operation and Development OCDE. (2022b). *OECD framework for the classification of AI systems*. (OECD Digital Economy Papers, núm. 323). (s. l.): OCDE.

Organisation for Economic Co-operation and Development OCDE. (2024a). *Governing with artificial intelligence: Are governments ready?* (s. l.): OECD.

Organisation for Economic Co-operation and Development OCDE. (2024b). *Explanatory memorandum on the updated OECD definition of an AI system*. (OECD Artificial Intelligence Papers, no. 8). (s. l.): OECD.

Office of the Federal Register. (2020). *Guidance for regulation of artificial intelligence applications*. Federal Register. https://www.govinfo.gov/content/pkg/FR-2020-12-08/pdf/2020-27065.pdf

Orlikowski, W. J., y Scott, S.V. (2008). 10 sociomateriality: Challenging the separation of technology, work and organization. *Academy of Management Annals*, *2*(1), 433-474.

Ortiz-Espinoza, Á., y Trujillo, A. E. (2024). "Siempre seguras": observatorio de acoso sexual callejero basado en inteligencia artificial. *Regiones y Desarrollo Sustentable*, *24*(45).

Osborne, S. P. (2006). The New Public Governance? *Public Management Review*, *8*(3), 377-387.

Parlamento Europeo y Consejo. (2024). Reglamento (UE) 2024/1689 de 13 de junio de 2024 por el que se establecen normas armonizadas en materia de inteligencia artificial. *Diario Oficial de la Unión Europea*. https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=OJ:L_202401689

Parlamento Europeo, Consejo y Comisión. (2023). Declaración Europea sobre los Derechos y Principios Digitales para la Década Digital (2023/C 23/01). *Diario Oficial de la Unión Europea*. https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32023C0123(01)

Pasquale, F. (2020). *New laws of robotics: Defending human expertise in the age of AI*. Harvard University Press.

Peeters, R., Rentería, C., y Cejudo, G. M. (2023). How information capacity shapes policy implementation: A comparison of administrative burdens in COVID-19 vaccination programs in the United States, Mexico, and the Netherlands. *Government Information Quarterly*, *40*(4), 101871.

Pitti, I. (2018). Unconventional political participation: An overview. En I. Pitti, *Youth and unconventional political engagement* (pp. 7-21). Palgrave.

Pollitt, C. (2016). *Advanced introduction to public management and administration*. Edward Elgar Publishing.

Presidencia de la República. (2025). *Decreto por el que se expide el Reglamento Interior de la Agencia de Transformación Digital y Telecomunicaciones, DOF 24/01/2025*. https://www.dof.gob.mx/nota_detalle.php?codigo=5747756yfecha=24/01/2025#gsc.tab=0

Proof News. (2024, 29 de abril). *AI models falter answering election questions in Spanish*. https://www.proofnews.org/ai-models-falter-answering-election-questions-in-spanish/

Pumplun, L., Tauchert, C., y Heidt, M. (2019). *A new organizational chassis for artificial intelligence-exploring organizational readiness factors*. En Proceedings of the 27th European Conference on Information Systems (ECIS), Stockholm & Uppsala, Sweden, June 8-14, 2019.

Raza, S., Qureshi, R., Zahid, A., Fioresi, J., Sadak, F., Saeed, M., Sapkota, R., Jain, A., Zafar, A., Hassan, M., Zafar, A., Maqbool, H., Vayani, A., Wu, J., y Shoman, M. (2025). *Who is responsible? The data, models, users or regulations? Responsible Generative AI for a sustainable future*. Authorea Preprints. Recuperado de https://www.techrxiv.org/doi/full/10.36227/techrxiv.173834932.29831105

Rentería, C., y Gómez-Álvarez, D. (2024). Ideas diferentes, pero del mismo vivero. Un esquema para estudiar la diversidad institucional de los laboratorios de innovación. En D. Gómez-Álvarez y C. Rentería (Coords.), *Pensar fuera de la caja burocrática: laboratorios de innovación pública en América Latina* (pp. 217-242). Universidad de Guadalajara/Debate.

Rentería, C. y Martínez, J. (En prensa). Gobernanza algorítmica: desafíos críticos de la IA en la Administración Pública. En L. Montaño, D. Arellano-Gault, y L. Ramos (Eds.), *Estudios críticos de administración contemporánea*. Universidad Autónoma Metropolitana.

Renn, O., y Benighaus, C. (2013). Perception of technological risk: insights from research and lessons for risk communication and management. *Journal of Risk Research*, *16*(3-4), 293-313.

Reuel, A., y Bucknall, B. (2024). *Open Problems in Technical AI Governance*. Centre for the Governance of AI. https://www.governance.ai/research-paper/open-problems-in-technical-ai-governance

Ricaurte, P., Cortés, U., Díaz, M., Jalife, S., Lemaitre, C., Muñoz, W., Revilla, T., Savage, S., y Sordo, M. (2024). *Hacia una política soberana de inteligencia*

artificial centrada en derechos humanos: Oportunidades y recomendaciones. https://zenodo.org/records/13899856

Rhodes, R. A. W. (1996). The new governance: governing without government. *Political Studies*, *44*(4), 652-667.

Roehl, U. B. U. (2023). Automated decision-making and good administration:Views from inside the government machinery. *Government Information Quarterly*, *40*(4), 101864.

Ruvalcaba, E., y García, V. (2023). *La adopción de la inteligencia artificial en los gobiernos estatales de México*. Universidad de Guadalajara/Centro de Estudios Estratégicos para el Desarrollo (CEED). https://editorial.udg.mx/gpd-la-adoptacion-de-la-inteligencia-artificial-en-los-gobiernos-estatales-de-mexico-9786075810218-659c3f318e2b3.html

Sabour, S., Liu, J. M., Liu, S., Yao, C. Z., Cui, S., Zhang, X., Zhang, W., Cao, Y., Bhat, A., Guan, J., Wu, W., Mihalcea, R., Wang, H., Althoff, T., Lee, T. M. C., y Huang, M. (2025). *Human Decision-making is Susceptible to AI-driven Manipulation*. En arXiv [cs.AI]. arXiv. http://arxiv.org/abs/2502.07663

Santiso, C. (2024). *Governing with artificial intelligence: Are governments ready?* París: Organisation for Economic Co-operation and Development.

Satyanarayan, A., y Jones, G. M. (2024). Intelligence as agency: Evaluating the capacity of generative AI to empower or constrain human action. *An MIT Exploration of Generative AI*. https://mit-genai.pubpub.org/pub/94y6e0f8

Schiff, K. J., Schiff, D. S., Adams, I. T., McCrain, J., y Mourtgos, S. M. (2023). Institutional factors driving citizen perceptions of AI in government: Evidence from a survey experiment on policing. *Public Administration Review*.

Schneider, J. (2025). *Generative to agentic ai: Survey, conceptualization, and challenges*. arXiv preprint arXiv:2504.18875.

Scott, J. C. (1998). *Seeing like a state: How certain schemes to improve the human condition have failed*. Yale University Press.

Secretaría de Comunicaciones y Transporte. (2019). *Conclusiones. Inteligencia artificial en las Telecomunicaciones y Radiodifusión*. Mesas temáticas de la Subsecretaría de Comunicaciones. https://www.gob.mx/cms/uploads/attachment/file/479534/conclusiones_inteligencia_artificial_en_telecomunicaciones_y_radiodifusion.pdf

Secretaría de Desarrollo Agrario, Territorial y Urbano (Sedatu). (2024). *Metrópolis de México 2020*. Sedatu/CONAPO/INEGI.

Selten, F., y Bram, K. (2024). Organizing public sector AI adoption: Navigating between separation and integration. *Government Information Quarterly*, *41*(1), 101885.

Selwyn, N., y Gallo Cordoba, B. (2022). Australian public understandings of artificial intelligence. *AI & Society*, *37*(4), 1645-1662.

Sidhu, D., Magistro, B., Stevens, B. A., y Loewen, P. J. (2024). Why do citizens support algorithmic government? *Journal of Public Policy*, *44*(3), 659-677.

Shanahan, M. (2024). Talking about large language models. *Communications of the ACM*, *67*(2), 68-79.

Soler Garrido, J., Tolan, S., Hupont Torres, I., Fernandez Llorca, D., Charisi, V., Gomez Gutierrez, E., Junklewitz, H., Hamon, R., Fano Yela, D., y Panigutti, C. (2023). *AI Watch: Artificial intelligence standardization landscape update*. Publications Office of the European Union. https://doi.org/10.2760/131984

Stahl, B. C., y Eke, D. (2024). The ethics of ChatGPT – Exploring the ethical issues of an emerging technology. *International Journal of Information Management*, *74*(102700), 102700.

Tangi, L., Combetto, M., Martin, B. J., y Rodriguez, M. P. (2023). *Artificial intelligence for interoperability in the European public sector*. European Commission.

Taeihagh, A. (2025). Governance of generative AI. *Policy y Society*. https://doi.org/10.1093/polsoc/puaf001

The Algorithm Register. (s. f.). The Algorithm Register. https://www.algorithmregister.org/

Thunes, J., y Kempton, A. (2025). *Balancing Insourcing and Outsourcing towards Digital Government: Establishing Obligatory Gateways as an Approach for Targeted Insourcing*. https://scholarspace.manoa.hawaii.edu/items/78c2fe81-7c4f-4205-907e-05a226576184

Tomazevic, N., Murko, E., y Aristovnik, A. (2024). Organisational Enablers of Artificial Intelligence Adoption in Public Institutions: A Systematic Literature Review. *Central European Public Administration Review*, *22*, 109.

Treasury Board of Canada Secretariat. (s. f.). *Directive on automated decision-making. Government of Canada*. https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592

Tyson, A., y Kikuchi, E. (2023). *Growing public concern about the role of artificial intelligence in daily life*. Pew Research Center. https://www.pewresearch.org/short-reads/2023/08/28/growing-public-concern-about-the-role-of-artificial-intelligence-in-daily-life/

Ulnicane, I., y Erkkilä, T. (2023). Politics and policy of Artificial Intelligence. *The Review of Policy Research*, *40*(5), 612-625.

UK Parliament. (s. f.). Online Safety Bill. https://bills.parliament.uk/publications/56170/documents/5068

US Government. (s. f.). *AI use cases*. https://ai.gov/ai-use-cases/

Unesco. (2024a). *Evaluación del estadío de preparación de la inteligencia artificial en México*. https://unesdoc.unesco.org/ark:/48223/pf0000390568

Unesco. (2024b). *Generative AI: UNESCO study reveals alarming evidence of regressive gender stereotypes*. https://www.unesco.org/en/articles/generative-ai-unesco-study-reveals-alarming-evidence-regressive-gender-stereotypes

Unesco. (2022). *Recomendación sobre la ética de la inteligencia artificial*. https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

United States Digital Service. (s. f.). *United States Digital Service*. https://www.usds.gov

Valderrama, J., Paz Hermosilla, C., y Garrido, S. (2022). *Algorithmic transparency. Open government partnership*. https://www.opengovpartnership.org/skeptics-guide-to-open-government-2022-edition/#toc_8

Verhofstadt, G. (2025). *The EU must defend its digital sovereignty*. https://foreignpolicy.com/2025/03/31/europe-digital-sovereignty-colony-trump-asml-ai-eurostack/

Wang, Y. F., Chen, Y. C., y Chien, S. Y. (2023). Citizens' intention to follow recommendations from a government-supported AI-enabled system. *Public Policy and Administration*, 09520767231176126.

Washington State Legislature. (2021). *Senate Bill 5116: Artificial intelligence in state government*. https://lawfilesext.leg.wa.gov/biennium/2021-22/Pdf/Bills/Senate%20Bills/5116.pdf

Whitford, A. B., y Whitford, A. M. (2024). The uneven deployment of algorithms as tools in government: evidence from the use of an expert system. *Data y Policy*, *6*, e50.

Winner, L. (2007). Do artifacts have politics? En *Computer ethics* (pp. 177-192). Routledge.

Wirtz, B. W., Langer, P. F., y Fenner, C. (2021). Artificial intelligence in the public sector-a research agenda. *International Journal of Public Administration*, *44*(13), 1103-1128.

Wirtz, B. W., Weyerer, J. C., y Geyer, C. (2019). Artificial Intelligence and the public sector—Applications and challenges. *International Journal of Public Administration*, *42*(7), 596-615.

Wirtz, B. W., Weyerer, J. C., y Sturm, B. J. (2020). The dark sides of artificial intelligence: An integrated AI governance framework for public administration. *International Journal of Public Administration*, *43*(9), 818-829.

World Bank. (2020). *Artificial intelligence in the public sector: Maximizing opportunities, managing risks (EFI Insight-Governance)*. Washington, DC: World Bank.

World Bank. (2021). *GovTech: Putting people first*. World Bank.

World Bank. (2022). *GovTech Maturity Index: The state of public sector digital transformation*. World Bank.

Zhao, H., Chen, H., Yang, F., Liu, N., Deng, H., Cai, H., Wang, S., Yin, D., y Du, M. (2024). Explainability for large language models: A survey. *ACM Transactions on Intelligent Systems and Technology*, *15*(2), 1-38.

Zhang, B., y Dafoe, A. (2019). Artificial intelligence: American attitudes and trends. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.3312874

Zhang, Y., Hsieh, C. W., y Liu, F. (2025). Employee Response to GenAI: Exploring the Effect of ChatGPT Dependence on Work Effort in a Public Service Context. *Public Performance y Management Review*, 1-29.

Zumaya, M., Guerrero, R., Islas, E., Pineda, O., Gershenson, C., Iñiguez, G., y Pineda, C. (2021). *Identifying tax evasion in Mexico with tools from network science and machine learning*. arXiv preprint arXiv:2104.13353.

Apéndices

Apéndice A. Libro de códigos de la base de datos Repositorio de sistemas de IA en el ámbito público mexicano. (Versión 1.2).[1]

En este apartado se proporciona una recopilación completa de los indicadores que conforman el *Repositorio de sistemas de IA en el ámbito público mexicano*, el cual documenta los sistemas de IA implementados por el sector público mexicano. Se incluye información clave sobre cada sistema, como el nombre, la institución que la utiliza, su estatus de implementación, y sus características tecnológicas y funcionales. Además, se describe las características contractuales y de transferencia tecnológica asociadas a los sistemas de IA, proporcionando detalles sobre los procesos de adquisición, contratación y financiamiento.

El libro de códigos se organiza en tres secciones. La primera describe información general sobre el sistema de IA (tabla 1). La segunda se ocupa de las características tecnológicas del sistema y las tareas y funciones que este tiene al interior de la entidad pública (tabla 2). La tercera sección describe las características contractuales y de transferencia tecnológica del sistema de IA (tabla 3).

[1] Forma de citación sugerida: Rentería, C., y Ordelin, J. L. (2025). *Repositorio de sistemas de IA en el ámbito público mexicano. (Versión 1.2)*. [Data set]. https://doi.org/10.17605/OSF.IO/CU6W3

Tabla 1. Información general sobre el sistema de IA

Atributo	*Descripción*	*Categorías*
ID	Número de identificación único en el repositorio de sistemas de IA.	No aplica
Nombre	Nombre del sistema de IA.	No aplica
Entidad pública	Entidad pública donde se utiliza el sistema de IA.	No aplica
Acrónimo	Acrónimo de la entidad pública.	No aplica
Rama	Rama del poder público de la entidad pública.	Judicial, ejecutivo, legislativo, autónomo, otras entidades
Nivel de gobierno	Nivel de gobierno de la entidad pública.	Federal, estatal y municipal
Entidad federativa	Nombre de la entidad federativa en que se encuentra la entidad pública.	Entidades federativas de México
Código INEGI	Clave de registro único de la entidad política, de acuerdo con la clasificación del INEGI.	No aplica
Incorporación	Fecha de incorporación del sistema de IA al repositorio.	[Fecha]
Fuente	Fuente de información mediante la que se documentó el registro del sistema de IA.	Transparencia, monitoreo de medios, entrevistas y concurso.
Enlace	Enlace de internet que provee información adicional sobre el sistema de IA.	[URL]
Estatus	Situación de utilización en la que se encuentra el sistema de IA.	Activo, piloto e inactivo.
Verificación	Fecha de última consulta en las fuentes de información para verificar el estatus del sistema de IA.	[Fecha]

Tabla 2. Descripción de las características tecnológicas y funcionales del sistema de IA

Atributo	*Descripción*	*Categorías*
Descripción	Descripción general del sistema de IA.	No aplica
Tipo de sistema	Tipo de sistema de IA implementado.	Asistente conversacional, reconocimiento de imágenes, reconocimiento del habla, sistema de clasificación y sistema de predicción.
Tarea	Tarea cognitiva en la que se usa el sistema de IA para sustentar actividades rutinarias de la entidad pública.	Análisis de datos, consulta de información, detección de anomalías, diagnóstico de salud, estimación de riesgos, gestión de trámites, reportes o denuncias, identificación o localización de personas, procesamiento de información, transcripción o traducción y otros.
Función	Función que cumple el sistema en la entidad pública.	Entrega de servicios públicos (ESP): informar a la ciudadanía, reducir cargas administrativas a la ciudadanía, facilitar trámites, reportes o denuncias a la ciudadanía. Operaciones internas (OI): operación de programas públicos, reducir carga laboral de servidores públicos, planeación o toma de decisiones. Gestión del desempeño (GD): evaluación de programas, formación de capital humano.

Tabla 2 (*Continuación*)

Atributo	*Descripción*	*Categorías*
Sector	Sector de política pública en el que se enmarca en el sistema de IA.	No aplica
ODS	Objetivo de Desarrollo Sostenible en el que se enmarca el sistema de IA.	No aplica

Tabla 3. Características contractuales y de transferencia tecnológica del sistema de IA

Atributo	Descripción	Valores
Adquisición	*Forma de adquisición del sistema de IA.*	*Adjudicación directa, convenio de colaboración, desarrollo interno, donación, invitación, licitación pública.*
Desarrollador	Entidad que desarrolló el sistema de IA	NA
Ejecutor	Entidad pública que despliega y/o utiliza el sistema de IA.	NA
Financiador	Entidad que financia el desarrollo o implementación del sistema de IA.	NA
Costo	Costo del desarrollo (en USD).	NA

Apéndice B. Información detallada de las solicitudes de transparencia

En este apartado se proporciona un resumen detallado de las solicitudes de transparencia enviadas y las respuestas recibidas desde los niveles de gobierno federal, estatal y municipal. Estos datos permiten analizar la interacción y cumplimiento de las obligaciones en materia de transparencia por parte de los sujetos obligados en las distintas entidades. El envío de estas solicitudes se inició el 1 marzo de 2024, primero a las instituciones de nivel federal, luego a las del estatal y por último las del municipal. Esta etapa concluyó el 20 de marzo de 2024.

Las respuestas del nivel federal tardaron un promedio de 30 días naturales. El nivel estatal mostró inconsistencia al respecto: en algunos casos la respuesta fue casi inmediata o rondando los 30 días naturales, pero hubo una gran cantidad de sujetos obligados que no emitieron respuesta o tardaron hasta tres meses en emitirla. A nivel municipal la respuesta fue más dinámica, la mayoría contestó en aproximadamente 25 días naturales.

Es importante mencionar que la Plataforma Nacional de Transparencia (PNT) brinda a los sujetos obligados 20 días hábiles contados a partir del día siguiente a la presentación de la solicitud para dar respuesta a los usuarios. En dicha plataforma se registraron 747 sujetos obligados a nivel federal. Sin embargo, el estudio excluyó a sindicatos, fondos, consejos, fideicomisos, asociaciones civiles, mandatos, partidos políticos y empresas del Estado. Por lo tanto, el total de sujetos aplicables para este análisis a nivel federal fue de 225 (tabla 4). A continuación, se desglosan los resultados de las solicitudes enviadas, las respuestas recibidas, así como la cantidad de respuestas positivas obtenidas.

Tabla 4. Resumen de solicitudes de transparencia enviadas a y respuestas recibidas de entidades públicas del gobierno federal

Sujetos obligados		*Respuestas recibidas*	*Respuestas positivas*
Descentralizados	103	97	10
Descentralizados no sectorizados	12	12	4
Desconcentrados	56	54	4
Centros de investigación	5	5	0
Administración pública centralizada	2	2	0
Instituciones banca descentralizada	3	3	1
Institución pública universitaria	1	1	0
Dependencia del poder ejecutivo	8	8	0
Dependencia de la administración pública	11	11	1
Entidad paraestatal	4	4	1
Órgano técnico especializado	1	1	0
Órgano regulador	1	1	0
Organismo jurisdiccional	3	3	0
Organismo público autónomo	3	3	2
Órgano del poder judicial	3	3	1
Organismo del poder legislativo	2	2	1
Órgano imponedor de justicia	1	1	0
Órgano público constitucional	2	2	0
Órgano constitutivo autónomo	4	4	2
Total	**225**	**217**	**27**

Nota: La cuarta columna indica las respuestas a las solicitudes de transparencia en las que las entidades públicas confirmaron que contaban con sistemas de IA.

En el caso del levantamiento de solicitudes de transparencia a nivel estatal, se registraron al menos 6275 sujetos obligados en la Plata-

forma Nacional de Transparencia. No obstante, aplicando el criterio de exclusión, solo 1671 fueron seleccionados como sujetos obligados aplicables (tabla 5).

Tabla 5. Resumen de solicitudes de transparencia enviadas a entidades públicas de gobiernos estatales y respuestas recibidas

	Total de sujetos obligados	*Sujetos obligados aplicables*	*Respuestas recibidas*	*Respuestas positivas*
Aguascalientes	115	57	57	7
Baja California Sur	113	43	43	4
Baja California	159	56	54	5
Campeche	153	40	40	3
Coahuila	174	59	55	5
Colima	164	59	50	3
Chiapas	289	66	64	5
Ciudad de México	146	69	52	6
Durango	141	44	44	1
Guanajuato	72	12	12	3
Guerrero	200	64	57	3
Hidalgo	144	9	9	2
Jalisco	488	69	65	9
Estado de México	238	67	4	2
Michoacán	289	33	32	5
Morelos	218	55	54	6
Nayarit	182	52	47	8
Nuevo León	217	64	63	7
Oaxaca	181	77	74	2
Puebla	366	53	53	6
Querétaro	130	42	40	6
Quintana Roo	105	62	61	3
San Luis Potosí	297	66	60	4
Sinaloa	198	53	53	3
Sonora	222	64	59	9
Tabasco	129	56	54	4
Tamaulipas	176	49	2	5
Tlaxcala	148	57	49	8
Veracruz	400	58	46	0
Yucatán	236	60	48	1
Zacatecas	185	56	46	3
Total	**6275**	**1671**	**1401**	**138**

El levantamiento de solicitudes de transparencia a nivel municipal fue diferente, ya que para seleccionar los sujetos obligados aplicables se utilizó la delimitación de las áreas metropolitanas establecidas por la Sedatu. Tal delimitación considera las zonas urbanas de alta concentración poblacional, lo que permitió una identificación más precisa de los sujetos obligados en las principales áreas metropolitanas del país (tabla 6).

Tabla 6. Resumen de solicitudes de transparencia enviadas a entidades públicas de gobiernos municipales, por entidad federativa, y respuestas recibidas

Estado	*Sujetos obligados aplicables*	*Respuestas recibidas*	*Respuestas positivas*
Aguascalientes	3	3	0
Baja California Sur	2	2	0
Baja California	4	4	0
Campeche	1	1	0
Coahuila	13	11	1
Colima	3	2	1
Chiapas	9	8	0
Chihuahua	7	6	0
Ciudad de México	16	16	1
Durango	2	2	0
Guanajuato	8	8	2
Guerrero	3	3	0
Hidalgo	10	10	3
Jalisco	9	7	1
Estado de México	62	0	0
Michoacán	7	6	0
Morelos	14	8	3
Nayarit	2	2	0
Nuevo León	16	15	0
Oaxaca	12	7	0
Puebla	22	20	0
Querétaro	4	4	0
Quintana Roo	4	4	1
San Luis Potosí	2	2	0
Sinaloa	3	3	2
Sonora	8	6	2
Tabasco	2	2	0
Tamaulipas	8	8	0
Tlaxcala	39	25	3

Tabla 6 (*Continuación*)

Estado	*Sujetos obligados aplicables*	*Respuestas recibidas*	*Respuestas positivas*
Veracruz	33	23	3
Yucatán	16	7	1
Zacatecas	5	5	1
Total	**345**	**230**	**25**

A continuación, se presenta un resumen de las *respuestas negativas* (es decir, aquellos sujetos obligados que afirmaron no contar con sistemas de IA) obtenidas en las solicitudes de transparencia enviadas a nivel federal, estatal y municipal. Se obtuvo un total de 198 respuestas negativas a nivel federal: 184 instituciones indicaron la inexistencia de *software* o herramientas tecnológicas que emplearan IA; 1 institución declaró que la información solicitada era reservada; 3 instituciones desecharon la solicitud, argumentando que la petición no era suficientemente clara; 4 no emitieron respuesta; 5 manifestaron incompetencia para atender la solicitud; y, finalmente, una respuesta fue errónea, ya que se adjuntó un archivo correspondiente a otra solicitud.

En cuanto a las solicitudes para el nivel estatal se obtuvieron 1533 respuestas negativas: 171 instituciones no emitieron respuesta; 1249 confirmaron que la información solicitada era inexistente en su organización; 10 desecharon la solicitud; 5 adjuntaron una respuesta correspondiente a una solicitud distinta; 94 indicaron la inexistencia de *software* o herramienta tecnológica que utilizara IA; 4 declararon que la información era reservada, y 94 declararon incompetencia para atender la solicitud.

En el análisis de las solicitudes sin respuesta, destaca el Estado de México, ya que, a pesar de ser uno de los que tienen mayor volumen de solicitudes a nivel estatal, la tasa de respuesta fue considerablemente baja. De 67 instituciones a las que se les pidió información, solo 4 respondieron. De estas, 1 institución rechazó la solicitud, 2 confirmaron el uso de inteligencia artificial en sus procesos, y 1 indicó la inexistencia de la información solicitada. Por ello, se presentaron las correspondientes quejas a través de la misma plataforma de transparencia; sin embargo, hasta la fecha no se ha recibido pronunciamiento de parte de los sujetos obligados.

A nivel municipal se enviaron 345 solicitudes de las que 319 respuestas fueron negativas. Estas instituciones contestaron de la siguiente forma: 4 solicitudes desechadas argumentando falta de información para emitir respuesta; 205 indicaron la inexistencia de *software* o herramienta tecnológica, y 110 no respondieron a la solicitud.

Apéndice C. Estrategia de codificación de la base de datos

Las tablas que siguen presentan un esquema detallado de instrucciones que utilizó el equipo de investigación para clasificar y codificar diversos sistemas de IA utilizados en entidades públicas.

Tabla 7. Esquema de codificación de las categorías descritas en la tabla 1

Atributo	*Estrategia de codificación*
ID	NA
Nombre	Nombre del sistema de IA.
Institución	Entidad pública donde se utiliza el sistema de IA.
Acrónimo	Acrónimo de la entidad pública.
Rama	Se registra la categoría en función de la entidad pública que usa el sistema.
Nivel de gobierno	Se registra la categoría en función de la entidad pública que usa el sistema.
Entidad federativa	Para los gobiernos subnacionales se especifica la entidad federativa o el municipio y su entidad federativa. Para la categoría "Otras entidades" se especifica "Universidades", "OSC" u "Otros".
Código INEGI	Se registra la codificación de cada entidad con la identidad específica del espacio geográfico que ocupa en el país
Incorporación	Se registra la fecha en la que se identificó el sistema por primera vez y se le asignó un ID en el registro.
Fuente	Se registra la forma en la que se recibe la información, ya sea a través de una solicitud en la PNT, información encontrada en la web, una entrevista directa con personal de la institución.
Enlace	Se adjunta la dirección electrónica en donde se puede consultar más información referente al sistema.
Estatus	Los investigadores codifican el estatus según la situación de operación del sistema. Se codifican como piloto aquellos sistemas que los usuarios declaran como en fase de prueba o experimentación.
Verificación	Se registra la fecha en que se realizó la última consulta en las fuentes de información para verificar el estatus del sistema.

Tabla 8. Esquema de codificación de las categorías descritas en la tabla 2

Atributo	*Estrategia de codificación*
Descripción	Se registra la descripción del sistema, la cual se obtiene mediante solicitud de transparencia o consulta de medios de comunicación (u otras fuentes).
Tipo de sistema	Se registra en función de lo reportado por las entidades públicas que lo utilizan.
Tarea	Los investigadores etiquetan en función de la interpretación de la información disponible sobre el sistema.
Función	Los investigadores etiquetan en función de la interpretación de la información disponible sobre el sistema.
Sector	Sector de política pública en el que se enmarca el uso del sistema de IA, de acuerdo con el sistema de clasificación de sectores del Banco Interamericano de Desarrollo
ODS	Objetivo de Desarrollo Sostenible al que potencialmente contribuye el uso del sistema de IA

Tabla 9. Esquema de codificación de las categorías descritas en la tabla 3

Mnemónico	*Estrategia de codificación*
Adquisición	Se registra la forma en la que se recibe la información, ya sea a través de una solicitud en la PNT, información encontrada en la web, o una entrevista directa con personal de la institución.
Desarrollador	Se registra la forma en la que se recibe la información, ya sea a través de una solicitud en la PNT, información encontrada en la web, o con una entrevista directa con personal de la institución.
Ejecutor	Se registra la forma en la que se recibe la información, ya sea a través de una solicitud en la PNT, información encontrada en la web, o en una entrevista directa con personal de la institución.
Financiador	Se registra la forma en la que se recibe la información, ya sea a través de una solicitud en la PNT, información encontrada en la web, o en una entrevista directa con personal de la institución.
Monto	Se registra la forma en la que se recibe la información, ya sea a través de una solicitud en la PNT, información encontrada en la web, o en una entrevista directa con personal de la institución.

Apéndice D. Recolección y análisis de información complementaria

Este apartado reporta las estrategias de recolección de información complementarias al mapeo de sistemas de IA. Se compone de dos estudios complementarios. En el apartado D.1 se describe una estrategia empírica cualitativa, basada en entrevistas semiestructuradas a actores relevantes del ecosistema de IA en México. En el apartado D.2 se presenta la metodología utilizada para el levantamiento de una encuesta a personal directivo en el área de tecnologías de la información en distintas entidades del gobierno federal.

D.1. Recolección y análisis de entrevistas semiestructuradas

Para las entrevistas semiestructuradas, se implementó un muestreo selectivo con actores relevantes: personas involucradas en el desarrollo y uso de sistemas de IA dentro de una entidad pública determinada, como los directores de Tecnologías de la Información, desarrolladores y responsables del mantenimiento del sistema de IA dentro de la entidad, o desarrolladores de empresas contratistas. En la tabla 10, se presentan estadísticas resumidas del tipo de participante en las entrevistas semiestructuradas.

Tabla 10. Descripción de los participantes por tipo de informante

Informantes	*Número*
Usuarios	4
Desarrolladores	6
Contratistas	2

Como parte del proceso de análisis de los datos cualitativos, se desarrolló un esquema de códigos para clasificar las categorías teóricas, su interpretación y la evidencia que las respalda dentro de los datos (tabla 11). El esquema comprende cinco categorías distintas, en las cuales los códigos comparten características y patrones comunes relacionados con las razones o limitaciones para adoptar herramientas de IA en organizaciones del sector público.

Tabla 11. Esquema de códigos derivado del análisis cualitativo

Código Axial	*Descripción*	*Citas ilustrativas*
	Categoría teórica: Capacidades	
Recursos humanos	El conocimiento, habilidades y capacidades del personal para diseñar, desarrollar e implementar tecnología de IA es el factor más frecuente que influye en su adopción.	"Otra limitación observada es que la mayoría de las herramientas de IA llevan poco tiempo en el mercado y hay falta de capacidad y conocimiento dentro de las instituciones" | "No hay personal con estas capacidades técnicas" | "Los funcionarios de primer nivel no tienen sofisticación estadística... necesitan instrucciones concretas".
Financieros	La falta de recursos financieros o presupuestales para capacitar al personal o adquirir infraestructura tecnológica es una barrera importante.	"No hay presupuesto asignado para capacitación en nuevas tecnologías" | "No hay presupuesto para desarrollar ningún proyecto que integre inteligencia artificial" | "Falta de recursos... así como falta de infraestructura y personal especializado".

Tabla 11 (*Continuación*)

Código Axial	*Descripción*	*Citas ilustrativas*
Infraestructura	La infraestructura tecnológica como *hardware* y *software* para el almacenamiento y análisis de datos es fundamental para el desarrollo de IA.	"No se cuenta con la infraestructura necesaria para implementar este algoritmo" | "No se ha pensado en ningún tema donde se pueda incluir... falta de infraestructura tecnológica adecuada".
Calidad de datos	Los conjuntos de datos deben cumplir criterios mínimos de precisión, validez y consistencia para usarse en proyectos de IA.	"Las fuentes alternativas no están diseñadas para fines estadísticos... tienes sesgos, falta de representatividad..." | "Análisis de fuentes de información, que es una carrera de obstáculos... los datos tienen diferentes propietarios...".
	Categoría teórica: Entorno normativo y político	
Barreras administrativas	Procesos administrativos que no contemplan líneas presupuestales para IA o que no facilitan su desarrollo.	"Las limitaciones establecidas para la adquisición de esta tecnología son un impedimento... las TIC están muy reguladas y limitadas" | "...principalmente presupuesto y falta de autorización oportuna por parte de la Unidad de Gobierno Digital".
Marco regulatorio	Falta de normas específicas para el desarrollo e implementación de IA en agencias públicas.	"Alcance actual del Reglamento Interno... no está dentro de las atribuciones conferidas..." | "La IA no está considerada en los estándares técnicos definidos por la Estrategia Digital Nacional...".
Entorno político	Condiciones políticas que afectan a las agencias públicas.	"Cambios de gobierno, sexenios... no saben los resultados de proyectos anteriores" (D 1) "Como administración pública, hay un componente de lógica política... si va mal, no lo compartas; si va bien, anúncialo" (D 4)
	Categoría teórica: Gestión	
Estrategia	Establecer metas claras para el desarrollo de IA, identificando cómo puede mejorar procesos internos.	"El área de tecnología ha identificado procesos que pueden mejorarse con estas tecnologías..." | "Mayor conocimiento de estas tecnologías y una necesidad concreta que puede beneficiarse...".
Liderazgo	Liderar una transición efectiva hacia el uso de IA y habilidades personales para promover su implementación.	"Hay que evangelizar sobre métodos de aprendizaje computacional... la naturaleza del método hace que sean cajas negras" | "Resistencia al cambio... socializamos desde el principio... vamos a explicar el proyecto y que no los va a reemplazar".
	Categoría teórica: Percepción	
Temores	Sentimientos de que la IA amenaza los empleos, resistencia al cambio y miedo a adaptarse a nuevas tecnologías.	"Resistencia al cambio, falta de conocimiento técnico y miedo a ser reemplazados" | "Las fuentes alternativas y métodos modernos están fuera de la tradición de las oficinas estadísticas...".
Escepticismo	Creencias negativas y prejuicios sobre la utilidad de la IA en agencias públicas.	"La mayoría de nuestros procesos son operativos y requieren intervención del personal" | "La entidad no requiere el uso de IA actualmente... además de no contar con suficiencia presupuestal".
Ética y privacidad	Preocupaciones por la privacidad en la recolección y uso de datos, así como cuestiones éticas.	"El uso de IA puede generar preocupaciones éticas y de privacidad, especialmente sobre datos personales y decisiones automatizadas…". | "…necesidad de implementar políticas relacionadas con la seguridad de la información y datos personales".

D.2. Levantamiento y análisis de encuesta

La encuesta fue dirigida a los directores de Tecnologías de la Información y Comunicación de agencias federales en México; se obtuvieron 70 respuestas, lo que representa una tasa de respuesta aproximada del 35%, sobre un universo de 208 agencias públicas convocadas. La encuesta fue distribuida por la Coordinación de la Estrategia Digital Nacional (CEDN) a los directores de Tecnologías de la Información y la Comunicación de las agencias públicas que formaban parte de la Política Nacional de Transformación Digital. La CEDN fue la entidad coordinadora de dicha política durante la administración presidencial 2018-2024. Esta instancia fue disuelta el 1 de diciembre de 2024, al inicio de la administración presidencial 2024-2030. La encuesta fue enviada por medio de comunicaciones institucionales vía correo electrónico por la CEDN y se recolectó entre el 30 de abril y el 16 de junio de 2024.

Se diseñó un cuestionario de encuesta con el propósito de cumplir dos objetivos de investigación. El primero consistió en recopilar datos cualitativos adicionales, provenientes de una muestra más amplia de directores de Tecnologías de la Información para profundizar en el análisis de los temas identificados en la primera etapa. El segundo objetivo fue formular preguntas orientadas a medir la adopción de IA, la intención de adoptarla y los factores que la determinan, con base en los hallazgos obtenidos en la fase inicial. Para atender el primer objetivo, se incluyeron tres preguntas cualitativas abiertas, que permitieron que los encuestados expresaran sus opiniones, a través de campos de texto, sobre el valor de la IA en sus organizaciones, las barreras para su adopción y los desafíos éticos asociados a su uso (tabla 12).

Las preguntas cuantitativas proveen información sobre el nivel de adopción de IA, de intención de adopción de IA, así como de factores determinantes de la adopción de los sistemas de IA (tabla 13).

Tabla 12. Conjunto de preguntas cualitativas del cuestionario

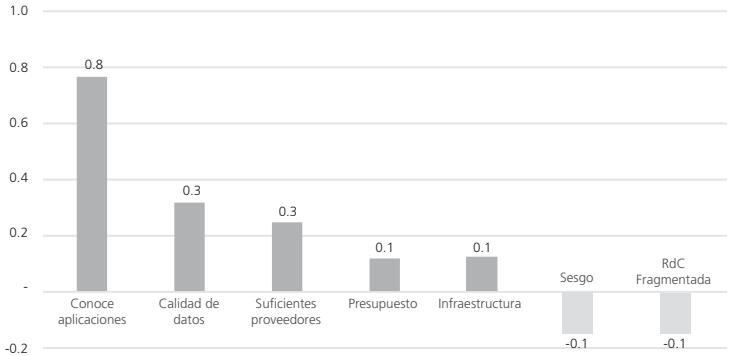
Tema	*Pregunta*
Valor de la IA en la organización	En su opinión, ¿valió la pena incorporar este algoritmo en su organización? ¿En qué otras tareas o funciones podrían ser útiles los sistemas IA? Por favor, comparta sus consideraciones en el recuadro de texto a continuación.
Barreras para la adopción de la IA	En su opinión, ¿qué impide o limita la posibilidad de desarrollar o adoptar sistemas de IA en su organización? Por favor, comparta sus reflexiones en el recuadro de texto a continuación.
Desafíos en el uso ético de la IA	Se habla mucho sobre la ética en el uso de la IA, especialmente en el sector público, donde los funcionarios deben apegarse a los principios éticos del servicio público. En este contexto, ¿cree usted que la ética del servicio público podría estar en riesgo con el uso de la IA? Por favor, comparta sus consideraciones en el recuadro de texto a continuación.

Tabla 13. Conjunto de preguntas cuantitativas del cuestionario

Categoría teórica	Tema	Pregunta de la encuesta
Adopción	Adopción de IA	¿Su departamento utiliza algún algoritmo de inteligencia artificial? (Respuestas permitidas: sí o no).
	Intención de adoptar IA	En una escala del 1 al 7, donde 1 significa "Nada probable" y 7 "Muy probable", ¿qué tan probable es que su organización desarrolle o adopte un algoritmo de IA para una actividad de servicio público?
Capacidad	Recursos financieros	Contamos con presupuesto suficiente para financiar el desarrollo y adopción de algoritmos de IA.
	Recursos humanos	Nuestro personal cuenta con la capacidad para desarrollar o gestionar proyectos de ciencia de datos.
	Infraestructura tecnológica	Contamos con infraestructura tecnológica suficiente para digitalizar registros administrativos y procesar datos.
	Calidad de los datos	Contamos con datos e información de alta calidad necesarios para entrenar algoritmos de IA.
Entorno	Contratación flexible	El proceso de contratación de servicios tecnológicos es suficientemente ágil y flexible en nuestra organización.
Gestión	Alineación estratégica	Conocemos varios sistemas de IA adoptados por otras agencias públicas (nacionales o internacionales) y tenemos claridad sobre cómo podríamos beneficiarnos de ellas.
Percepción	Sesgo discriminatorio	Los algoritmos de IA pueden introducir sesgos en la toma de decisiones.
	Opacidad	Los algoritmos de IA son una "caja negra" y es imposible entender cómo generan sus resultados o predicciones.
	Vulnerabilidad de datos	Los algoritmos de IA comprometen la privacidad y la protección de datos.
	Rendición de cuentas compleja	El desarrollo de IA involucra múltiples proveedores, cada uno responsable de diferentes etapas, lo que complica la rendición de cuentas y la supervisión.
Experiencia	Experiencia en proyectos de IA	Contamos con experiencia en el desarrollo o contratación de servicios tecnológicos basados en IA.

El análisis estadístico principal para esta obra corresponde a un modelo lineal de mínimos cuadrados ordinarios, donde la variable dependiente es la intención de adopción de IA y, como variables independientes, un conjunto de factores determinantes de la adopción. La figura 1 presenta los factores determinantes incorporados en el modelo y sus coeficientes estandarizados.

Figura 1. Factores determinantes de la intención de adopción de IA



Colección Asuntos Públicos

Este libro es un esfuerzo pionero para sistematizar nuestra comprensión y conocimiento sobre la incorporación e implementación de la inteligencia artificial en las instituciones del ámbito público mexicano. Los autores exploran el potencial de esta tecnología, lo mismo que sus complejidades, hasta convertir su análisis e investigación en una pieza de calibre mayor para el cuerpo de un conocimiento científico en pleno crecimiento. Se trata de una obra útil por su diagnóstico y prospectiva.

La investigación encuentra que el uso de la inteligencia artificial en la administración pública busca eficientizar procesos, optimizar la toma de decisiones, mejorar los servicios públicos, reducir la burocracia. Objetivos que generan resistencias, recelo y cuestionamientos. Por ello, es necesario comprender y aprender cómo funcionan, se utilizan y operan los sistemas de inteligencia artificial; cómo ocurre su influencia en la toma de decisiones y, más aún, cuál es su impacto en la vida de las personas, en la política nacional de transformación digital y en la soberanía tecnológica.

FLACSO
MÉXICO