

Gobernar con la inteligencia artificial

Panorama actual y hoja de ruta en las funciones centrales de gobierno



Gobernar con la inteligencia artificial

PANORAMA ACTUAL Y HOJA DE RUTA EN LAS
FUNCIONES CENTRALES DE GOBIERNO



El presente trabajo fue aprobado y desclasificado por el Comité de Gobernanza Pública el 09/05/2025.

Tanto este documento, así como cualquier dato y cualquier mapa que se incluya en él, se entenderán sin perjuicio respecto al estatus o la soberanía de cualquier territorio, a la delimitación de fronteras y límites internacionales, ni al nombre de cualquier territorio, ciudad o área.

Los datos estadísticos para Israel son suministrados por y bajo la responsabilidad de las autoridades israelíes competentes. El uso de estos datos por la OCDE es sin perjuicio del estatuto de los Altos del Golán, Jerusalén Este y los asentamientos israelíes en Cisjordania bajo los términos del derecho internacional.

Nota de Turquía

La información del presente documento en relación con “Chipre” se refiere a la parte sur de la Isla. No existe una sola autoridad que represente en conjunto a las comunidades turcochipriota y grecocipriota de la Isla. Turquía reconoce a la República Turca del Norte de Chipre (RTNC). Mientras no haya una solución duradera y equitativa en el marco de las Naciones Unidas, Turquía mantendrá su postura frente al “tema de Chipre”.

Nota de todos los Estados Miembros de la Unión Europea que pertenecen a la OCDE y de la Unión Europea
Todos los miembros de las Naciones Unidas, con excepción de Turquía, reconocen a la República de Chipre. La información contenida en el presente documento se refiere a la zona sobre la cual el Gobierno de la República de Chipre tiene control efectivo.

Por favor, cite esta publicación de la siguiente manera:

OECD (2025), *Gobernar con la inteligencia artificial: Panorama actual y hoja de ruta en las funciones centrales de gobierno*, OECD Publishing, Paris, <https://doi.org/10.1787/dc00e56a-es>.

ISBN 978-92-64-55690-4 (impresa)

ISBN 978-92-64-72718-2 (PDF)

ISBN 978-92-64-94671-2 (HTML)

Publicado originalmente por la OCDE con el título: OECD (2025), *Governing with Artificial Intelligence*, OECD Publishing, Paris, <https://doi.org/10.1787/795de142-en>.

Imágenes: Portada © cofotoisme/Getty images.

Las erratas de las publicaciones se encuentran en línea en: <https://www.oecd.org/en/publications/support/corrigenda.html>.

© OCDE 2025



Atribución/Reconocimiento 4.0 Internacional (CC BY 4.0)

Este trabajo está disponible bajo la licencia Creative Commons Attribution/Reconocimiento 4.0 International. Al utilizar este trabajo, acepta estar sujeto a los términos de esta licencia (<https://creativecommons.org/licenses/by/4.0/>).

Atribución – debe citar el trabajo.

Traducciones – debe citar la obra original, identificar cambios al original y agregar el siguiente texto: *En caso de discrepancia entre la obra original y la traducción, solo se considerará válido el texto de la obra original*.

Adaptaciones – debe citar el trabajo original y agregar el siguiente texto: *Esta es una adaptación de un trabajo original de la OCDE. Las opiniones expresadas y los argumentos utilizados en esta adaptación no deben considerarse representativos de los puntos de vista oficiales de la OCDE o de sus Países miembros*.

Material de terceros – la licencia no se aplica al material de terceros en la obra. Si utiliza dicho material, usted es responsable de obtener el permiso del tercero y de cualquier reclamación por infracción.

No debe utilizar el logotipo, la identidad visual o la imagen de portada de la OCDE sin permiso expreso ni sugerir que la OCDE respalda su uso del trabajo.

Cualquier disputa que surja bajo esta licencia se resolverá mediante arbitraje de conformidad con el Reglamento de Arbitraje de la Corte Permanente de Arbitraje (PCA) de 2012. El lugar del arbitraje será París (Francia). El número de árbitros será de uno.

Prólogo

En los últimos años, los gobiernos de todo el mundo han dado pasos importantes en la digitalización del sector público, acelerada por los desafíos planteados por la pandemia del COVID-19. A pesar de estos avances, siguen existiendo muchos obstáculos. Hoy en día, la atención se centra en aprovechar las tecnologías digitales para satisfacer mejor las necesidades de la ciudadanía, creando eficiencias que aporten más valor a los contribuyentes, prestando servicios públicos adaptados, accesibles e inclusivos, así como mejorando la participación ciudadana.

No se puede subestimar la importancia de estos esfuerzos. Si bien mejorar la eficiencia y la efectividad del gobierno es una prioridad, también resulta fundamental fortalecer la confianza ciudadana en sus instituciones puesto que ésta es decisiva para el éxito de los proyectos digitales públicos. De acuerdo con los datos de la OCDE de 2023, solamente el 39% de las personas tienen una confianza moderadamente alta o mayor en el gobierno nacional. Sin embargo, unos servicios públicos fiables, receptivos y justos pueden aumentar la confianza en el gobierno.

La inteligencia artificial (IA) se está convirtiendo en un componente importante del proceso de gobierno digital, ya que ofrece importantes beneficios en diversas áreas, como la automatización, la detección de anomalías y la mejora en la toma de decisiones. Por ejemplo, asistentes conversacionales basados en IA se están utilizando para responder a las consultas de la ciudadanía y ayudarles a diligenciar formularios. En la gestión de desastres, la IA está ayudando a anticipar desastres naturales y acelerar los esfuerzos de respuesta. En la administración tributaria, la IA se utiliza para la detección de fraudes. En general, el uso de la IA en la administración puede mejorar la productividad, capacidad de respuesta y responsabilidad de la administración pública.

Sin embargo, la adopción de la IA en la administración pública está por detrás de la del sector privado. Los gobiernos se enfrentan a contextos y desafíos únicos que dificultan la rápida adopción de la IA, como la escasez de competencias, los sistemas informáticos heredados y obsoletos, la poca disponibilidad de datos y un entorno con limitaciones financieras, así como requisitos más estrictos en materia de privacidad, transparencia y representación de diversos grupos.

El aprendizaje basado en la experiencia es la mejor manera de progresar en este entorno digital tan dinámico. Si bien la madurez de la IA todavía no es generalizada en los gobiernos, hay muchos ejemplos de aplicaciones de IA de las que se pueden extraer lecciones. La OCDE se ha comprometido en apoyar a los gobiernos en su trayectoria hacia la IA facilitando el intercambio de experiencias y perspectivas. Este informe, basado en el análisis de decenas de enfoques de gobernanza y 200 casos de uso concretos de IA, además de una amplia investigación y discusiones con gobiernos, sienta las bases de una línea de trabajo en curso sobre "Gobernar con IA" (<https://oecd.ai/gov>). Su objetivo es proporcionar a los gobiernos los elementos necesarios para un uso eficaz de la IA e identificar las áreas en las que la IA está teniendo impacto y en las que siguen existiendo deficiencias. Los esfuerzos futuros se basarán en la creciente evidencia empírica de políticas y casos prácticos, tratando de seguir ayudando a los gobiernos a establecer habilitadores, salvaguardas y mecanismos de participación necesarios para un enfoque estratégico y confiable de la IA.

La OCDE reconoce los esfuerzos realizados por muchos países para acelerar la adopción de la IA en la administración pública, además de los datos y la información analizados para el presente informe. Sus ejemplos han sido fundamentales para dar forma a nuestra comprensión del papel de la IA en la transformación del gobierno.

Este informe forma parte del Proyecto Horizontal de la OCDE, Prosperando con la IA: Empoderando Economías y Sociedades.

Agradecimientos

El presente informe ha sido elaborado por la Dirección de Gobernanza Pública de la OCDE (GOV), bajo la dirección de Elsa Pilichowski, Directora, y Gillian Dorner, Directora Adjunta. El informe es fruto del esfuerzo conjunto de GOV, coordinado por la División de Gobierno Innovador y Digital y Abierto (INDIGO), bajo la dirección de Carlos Santiso, Jefe de División, y la supervisión de Marco Daglio, Jefe Interino de la Unidad de Gobierno Digital (DGU). Jamie Berryhill, Barbara Ubaldi y Ricardo Zapata dirigieron el proyecto. Simone Maria Parazzoli y Paula Ayala, consultores, prestaron apoyo analítico y de redacción a lo largo del informe.

Los Capítulos 1-3 fueron redactados por Jamie Berryhill y Ricardo Zapata. El Capítulo 4 fue redactado por Jamie Berryhill, Emma Cantera, Cecilia Emilsson, Felipe González, Julian Olsen, Seong Ju Park, María Pascual-Dapena, Arturo Rivera, Alex Seeman, Barbara Ubaldi y Ricardo Zapata (INDIGO); Daniel Gerson de la División de Gestión Pública y Presupuestos (PMB); Erika Bozzay, Simon Cox y András Hlacs de la División de Infraestructura y Contratación Pública (IPP); y Paula Ayala y Hana Murr, consultoras.

El Capítulo 5 fue un esfuerzo colaborativo, con secciones de:

- **Administración tributaria**, por Paul Marsh y Sofie Stenbøg, de la División de Administración Tributaria e IVA (TAV) del Centro de Política y Administración Tributaria (CTP).
- **Gestión financiera pública**, por Delphine Moretti (GOV/PMB) y Nicolas Botton, consultor.
- **Diseño e implementación de la regulación**, por Anna Pietikäinen, Giuseppa Ottimofiore, Becky King, James Drummond, Margarita Escobar, Carola Bertone, Alexander Roberts, Martha Baxter y Shemsije Jashari, de la División de Política Regulatoria (REG) del gobierno y Ekaterina Zakharyan y Vincent Van Langen, consultores.
- **Reforma de la función pública**, por Daniel Gerson, Donald Mulligan, Alana Baker, Francois Villeneuve y Natalia Nolan-Flecha (GOV/PMB).
- **Contratación pública**, por Erika Bozzay, Andras Hlacs, Simon Cox, Kenza Khachani (GOV/IPP) y Matthieu Cahen (GOV/DO).
- **Lucha contra la corrupción y promoción de la integridad pública**, por Maria Eugenia Heyaca y Helene Wells, de la División Anticorrupción e Integridad en el Gobierno (GOV/ACIG).
- **Evaluación de políticas públicas**, por Stéphane Jacobzone, Silvia Picalarga (GOV/PMB), y Guste Papaurelyte, consultora.
- **Participación ciudadana y gobierno abierto**, por David Goessmann, Giulia Cibrario y Mauricio Mejía Galvan (GOV/INDIGO).
- **Diseño y prestación de servicios públicos**, por Piret Tõnurist y Angela Hanson (GOV/INDIGO). También contribuyeron Oscar Huerta, del Centro de la OCDE para el Emprendimiento, las Pymes, Regiones y Ciudades (CFE) y Alibhe Brioscu, Valerie Frey, y Anne Lauringson, de la Dirección de Empleo, Trabajo y Asuntos Sociales (ELS) de la OCDE.
- **Aplicación de la ley y gestión del riesgo de desastres**, por Jack Radisch y Néstor Alfonzo Santamaría (GOV/IPP).

- **Administración de justicia y acceso a la justicia**, por Tatyana Teplova, Mariane de Langen y Mariane Piccinini Barbieri, de la División de Asociaciones Globales, Inclusión y Justicia (GPIJ) del Gobierno de Guyana, y Paola Gálvez Callirgos, consultora. También hicieron contribuciones Celine Cojocar (GOV/GPIJ), Aitor Cubo, Fabiola Solino Díaz, Alejandro Fernández Muñoz (España) y Tomislav Boršić (Croacia).

Además de estos colaboradores, otros colegas de la OCDE revisaron el informe y aportaron valiosos comentarios, entre ellos, Mia Drazilova, Marianna Kartunen, Harris Khan, Cameron Knott, Bruno Monteiro y Chiara Varazzani (GOV/INDIGO); Celine Caira, Julia Carro, Gallia Daor, Christy Dentler, Sergi Gálvez Duran, Limor Shmerling Magazanik, Tana McCauley, Karine Perset, Lucia Russo y Kasumi Sugimoto, de la Dirección de Ciencia, Tecnología e Innovación (STI); Stefano Marta de CFE; y Theodora Xenogiani y Angelica Salvi Del Pero (ELS).

El informe se benefició significativamente de la experiencia del Grupo de Trabajo de Altos Funcionarios de Gobierno Digital de la OCDE (E-Leaders), presidido por Frank Leyman (Bélgica). También se benefició de las revisiones y comentarios proporcionados por los delegados del Comité de Gobernanza Pública (PGC), el Comité de Política de Desarrollo Regional (RDPC) y el Comité de Política Digital (DPC) y su proyecto, la Alianza Global sobre Inteligencia Artificial (GPAI). El equipo también agradece a Cindy García y Deborah Fernández por su apoyo administrativo; a Sara Sreberny-Mohammadi, Justin Kavanaugh y Thibaut Gigou por su apoyo en las comunicaciones; a Andrea Uhrhammer, Rebecca Bonthius y Jeff Israely por su apoyo editorial; y a Claudio Toledo, Paul Graves, María Stephanía Guzmán Serna, María Elena Hernández Sánchez, Lucía Luque López, Ana Ríos Camarena, Paulina Rosales, Nicolás Penagos y María Roa Vicens por la revisión editorial de la traducción al español.

Índice

Prólogo	3
Agradecimientos	5
Resumen ejecutivo	10
Acrónimos y abreviaturas	13
1 Cómo la inteligencia artificial está acelerando la transformación digital de la administración pública	15
El proceso de transformación digital de la administración pública	16
Entender el potencial transformador de la IA	19
Uso de IA en la administración pública, un contexto único	22
Áreas de oportunidad clave y beneficios de la IA en la administración pública	23
Para aprovechar los beneficios de la IA, es necesario gestionar sus riesgos	34
Hacer realidad un futuro positivo para la IA en el gobierno	48
Bibliografía	49
Notas	61
2 Tendencias y primeras lecciones del uso de IA en las distintas funciones de la administración pública	63
Análisis de la OCDE de 200 casos de uso de 11 funciones gubernamentales	64
La IA es una prioridad, pero los esfuerzos no son sistémicos	65
Tendencias generales en los casos de uso de la IA por parte de las administraciones públicas	67
Los casos de uso podrían plantear riesgos si no se implementan de manera confiable	80
Bibliografía	83
Notas	85
3 Desafíos de implementación que dificultan el uso estratégico de la IA en la administración pública	87
La mayoría de las iniciativas gubernamentales en materia de IA se desarrollan en fases exploratorias o piloto, con una escalada y documentación limitadas	88
Desafíos comunes a las funciones centrales del gobierno	92
Desafíos que son un poco menos comunes o varían entre las funciones gubernamentales	98
Bibliografía	107
Notas	112

4 Habilitadores, salvaguardas y participación para desbloquear una IA confiable	113
Medidas de política para liberar el potencial de la IA	114
Fortalecimiento de habilitadores para apoyar la adopción de una IA confiable	116
Establecer salvaguardas para orientar la IA de forma estratégica y responsable	145
La participación para dar forma a una IA estratégica y responsable	157
Un marco de trabajo para una IA confiable en el sector público	161
Futuro trabajo de la OCDE sobre estas problemáticas	163
Bibliografía	163
Notas	174
5 Inmersión profunda: La situación actual y el futuro potencial de la IA en el sector público	181
IA en la administración tributaria	183
Situación actual	183
IA en la gestión de finanzas públicas	191
IA en el diseño e implementación de la regulación	200
IA en la reforma de la función pública	213
IA en la contratación pública	223
IA en la lucha contra la corrupción y promoción de la integridad pública	233
IA en la evaluación de políticas públicas	242
IA en la participación cívica y gobierno abierto	250
IA en el diseño e implementación de servicios públicos	261
IA en la aplicación de la ley y la gestión del riesgo de desastres	280
IA en la administración y acceso a la justicia	293
Bibliografía	308
Notas	330

GRÁFICOS

Gráfico 1.1. Índice de gobierno digital de la OCDE 2023, resultados compuestos por país	18
Gráfico 1.2. Las inversiones de capital de riesgo en IA han aumentado con el paso de los años	21
Gráfico 1.3. IA en cada fase del ciclo de formulación de políticas	24
Gráfico 1.4. Más de la mitad (59%) de los conjuntos de datos de entrenamiento de IA de código abierto están en inglés	44
Gráfico 1.5. Los incidentes de IA han tendido en general al alza desde finales de 2022	47
Gráfico 2.1. Los casos de uso están más presentes en los servicios públicos, la participación cívica y las funciones de justicia	67
Gráfico 2.2. La UE y ALC siguen una tendencia similar con la muestra de casos de uso de IA recopilada para este informe	69
Gráfico 2.3. Casos de uso de IA por función esencial y actividad de la administración pública	70
Gráfico 2.4. Los servicios públicos y la participación representan una parte importante de los casos de uso en los procesos gubernamentales de la UE y ALC	71
Gráfico 2.5. Beneficios potenciales de los casos de uso de IA en distintas funciones de la administración pública	72
Gráfico 2.6. Beneficios específicos de casos de uso de IA	73
Gráfico 2.7. El potencial de riesgos operativos es el más representado en todas las funciones gubernamentales	81
Gráfico 3.1. La mayoría de los casos de uso de IA en la Unión Europea (UE) se encuentran en fases piloto o de desarrollo	90
Gráfico 4.1. Índice de datos abiertos, útiles y reutilizables (OURdata) de la OCDE, 2023	127
Gráfico 4.2. Gobernanza de datos en el sector público	128
Gráfico 4.3. Considerar el nivel de alfabetización sobre IA necesario para los diferentes grupos de usuarios en la organización	134

Gráfico 4.4. Marco de trabajo de políticas GovTech de la OCDE	145
Gráfico 4.5. Pasos clave para entender a los usuarios y sus necesidades en los desarrollos de IA gubernamental	160
Gráfico 4.6. Marco de la OCDE para una inteligencia artificial confiable en el sector público	162
Gráfico 5.1. Despliegues de IA en los miembros de la OCDE que utilizan IA en la administración tributaria	183
Gráfico 5.2. Objetivos para las actualizaciones del FMIS en países de la OCDE, 2022	197
Gráfico 5.3. Posible uso de IA y análisis de datos a lo largo del ciclo de la contratación pública	223

TABLAS

Tabla 1.1. Entender el uso de la IA en la administración pública	23
Tabla 2.1. Análisis de las funciones gubernamentales para el presente reporte	65
Tabla 2.2. Principios de la IA de la OCDE	65
Tabla 2.3. Ejemplos de IA para procesos y servicios automatizados, racionalizados y adaptados	74
Tabla 2.4. Ejemplos de IA para mejorar la toma de decisiones, interpretación y pronósticos	76
Tabla 2.5. Ejemplos de IA para mejorar la rendición de cuentas y la detección de anomalías	77
Tabla 2.6. Ejemplos de cómo la IA puede brindar oportunidades a partes interesadas externas como un bien para todos	78
Tabla 3.1. Costo de 1 millón de tokens en modelos comunes de IA generativa (USD)	102
Tabla 4.1. Preguntas y medidas de política pública que sustentan el marco para una IA confiable en el sector público	162

Resumen ejecutivo

La inteligencia artificial (IA) es una de las fuerzas más transformadoras del siglo XXI y se está convirtiendo en una parte integral de la transformación digital de los gobiernos de todo el mundo. El uso de la IA por parte de la administración pública puede facilitar procesos internos y servicios públicos automatizados y adaptados, fomentar una mejor toma de decisiones y pronósticos, mejorar la detección del fraude y la calidad del empleo y el aprendizaje de los servidores públicos, todo ello con efectos tangibles. Por ejemplo, el Instituto Alan Turing estima que la IA podría automatizar el 84% de las transacciones repetitivas del servicio público en el Reino Unido, ahorrando el equivalente a 1 200 años-persona de trabajo al año. A pesar de su potencial, el uso de la IA por parte de los gobiernos sigue rezagado respecto al del sector privado.

Principales hallazgos: Cómo puede servir la IA a la ciudadanía

La OCDE realizó una investigación exhaustiva sobre el uso de la IA en 11 funciones centrales de la administración pública, analizando 200 casos prácticos. Los resultados indican que la IA se utiliza con mayor frecuencia en la prestación de servicios públicos, justicia y participación cívica, mientras que su presencia es relativamente menor en áreas como la evaluación de políticas, la administración tributaria y la reforma de la función pública. A medio camino se encuentran la contratación pública, la gestión financiera, la lucha contra la corrupción y la promoción de la integridad pública, y el diseño e implementación de la regulación. Una posible explicación de esta distribución es que algunas funciones abarcan una variedad más amplia de usos (servicios públicos), mientras que otras son más limitadas (reforma de la función pública, administración tributaria). Además, algunas se enfrentan a más limitaciones normativas (por ejemplo, la administración tributaria, dada la normativa sobre el uso de datos fiscales), mientras que otras enfrentan menos desafíos de implementación y pueden madurar más rápido (participación cívica). En algunas funciones, como la administración de justicia, las demandas públicas y los crecientes retrasos en las transacciones podrían precipitar la adopción de la IA como una oportunidad para abordar desafíos urgentes.

El uso de la IA es más frecuente en las operaciones internas y la prestación de servicios públicos, pero menos prominente en la supervisión gubernamental. También se observa un menor uso en la formulación de políticas, consistente con análisis anteriores de la OCDE. Los casos de uso suelen basarse en enfoques clásicos basados en reglas o técnicas de aprendizaje automático (ML) ya establecidas, siendo menos comunes la IA generativa (GenAI), que incluye los modelos extensos de lenguaje (LLM). En cuanto a las ventajas, la mayor parte de los casos se destinan a promover procesos y servicios automatizados, racionalizados y adaptados, seguidos de una mejor toma de decisiones y pronósticos, y de una mayor rendición de cuentas y detección de anomalías. Algunos casos tratan de desbloquear nuevas oportunidades para partes interesadas externas (por ejemplo, la ciudadanía o empresas) mediante el acceso a sistemas de IA proporcionados por el gobierno, pero es posible que se requieran más esfuerzos.

Riesgos del uso de la IA en la administración pública

La adopción de la IA sin riesgos no existe. Liberar los beneficios de la IA requiere mitigar sus riesgos. Los algoritmos sesgados pueden dar lugar a resultados negativos; el uso indebido de la IA puede infringir o impedir el libre ejercicio de los derechos humanos; la insuficiente transparencia, explicabilidad y comprensión pública de la IA puede erosionar la responsabilidad y causar resistencia pública; y la dependencia excesiva de la IA puede ampliar las brechas digitales y permitir que se propaguen errores sistémicos, debilitando la confianza de la ciudadanía en el gobierno. Estos riesgos pueden verse amplificados en países que carecen de los mecanismos necesarios para garantizar el ejercicio, la protección y la promoción de los derechos humanos, o pueden resultar en un uso indebido de la IA por parte de funcionarios públicos a título individual. También podría ocurrir un desplazamiento de la fuerza laboral de la función pública si los gobiernos buscan reemplazar en lugar de aumentar las capacidades de los funcionarios públicos.

El hecho de que los gobiernos no aprovechen la IA también representa un riesgo, que se traduce en oportunidades perdidas para obtener beneficios y en un aumento de la brecha entre las capacidades del sector público y el privado. Para responder eficazmente a las crecientes expectativas ciudadanas y fortalecer la confianza pública, será necesario que los gobiernos implementen soluciones basadas en IA. Ignorar la transformación de la IA, o esperar a que se resuelvan todas las incógnitas, relega al gobierno a ser un tomador de la tecnología en lugar de un creador de opciones, incurriendo en costos y desventajas significativas. Si los gobiernos no refuerzan pronto las capacidades internas de IA, podrían tener dificultades para ponerse al día.

Los gobiernos también enfrentan desafíos en la implementación de la IA

Los desafíos que plantea ampliar con éxito las aplicaciones de IA hacen que las iniciativas gubernamentales de IA sigan a menudo en fases piloto. Las brechas de habilidades y dificultades para obtener y compartir datos de calidad se encuentran a lo largo de todas las funciones gubernamentales. Más aún, aunque las estrategias de IA en el gobierno son comunes, la falta de directrices concretas dificulta su puesta en práctica. Estos factores agravan la aversión al riesgo, dificultando la capacidad de los gobiernos para innovar con IA. Además, la insuficiencia de mecanismos de supervisión y evaluación limita su capacidad para medir los avances, detectar riesgos y demostrar el rendimiento de la inversión. Los costos financieros también son un desafío común.

Algunos desafíos son más frecuentes en algunas funciones que en otras. Por ejemplo, la administración tributaria enfrenta leyes y normas complejas en torno a los procesos y datos fiscales, mientras que la contratación pública enfrenta la falta de normas establecidas en torno a la IA. Por último, el uso de la IA en funciones como la gestión de las finanzas públicas está limitado por infraestructuras tecnológicas heredadas y obsoletas que no son adecuadas para el desarrollo o uso de la IA.

Cómo pueden las administraciones públicas garantizar un uso confiable de la IA

Para cosechar los beneficios de la IA y, al mismo tiempo, mitigar sus riesgos y superar los desafíos de implementación, las instituciones públicas deben poner en marcha:

- **Habilitadores** para promover una adopción confiable, como la gobernanza, los datos, la infraestructura digital, las competencias digitales, las inversiones financieras, la agilidad de los procesos de contratación y la capacidad para asociarse con agentes no gubernamentales.
- **Salvaguardas** para orientar el uso de la IA, incluidas normas y políticas, directrices y marcos, mecanismos de transparencia y rendición de cuentas que abarquen el ciclo de vida del sistema de IA, y órganos de supervisión y asesoramiento para orientar y evaluar su implementación.

- **Participación** para impulsar políticas y casos de uso receptivos y centrados en el usuario, que contemplen mecanismos para interactuar con las principales partes interesadas, incluido el público, la sociedad civil y las empresas.

Se necesita más acción para invertir en y adoptar una IA confiable en la administración pública, pero los enfoques existentes ofrecen lecciones e inspiración

En la medida de lo posible, la OCDE incentiva a los gobiernos a dar prioridad a las aplicaciones de IA de alto beneficio y bajo riesgo, sobre todo al construir un nivel inicial de madurez. Sin embargo, la mayoría carece de los procesos para la medición holística de resultados potenciales o logrados, como la eficiencia del gasto, la calidad del servicio o los daños potenciales, que les permitan hacer estas determinaciones. Abordar esto, debería ser una prioridad, garantizando que las implementaciones de la IA sean transparentes, justas y seguras.

Muchos esfuerzos gubernamentales en materia de IA están en sus inicios, pero algunos están generando valiosas lecciones. La OCDE está comprometida con la recopilación y el análisis de evidencia para ampliar la base de conocimiento empírico sobre las políticas y experiencias que funcionan, centrándose en cómo los gobiernos pueden aprovechar una IA confiable para generar valor público.

Acrónimos y abreviaturas

AIG	Gobernanza anticipatoria de la innovación
AIM	Monitor de incidentes de IA de la OCDE
API	Interfaz de programación de aplicaciones
APP	Alianza público-privada
APS	Función Pública Australiana
ATI	Acceso a la información
BLOOM	Modelo Multilingüístico Extenso de Ciencia-abierta u acceso-aberto BigScience
CAIO	Persona titular de la Dirección de IA
CdE	Consejo de Europa
CE	Comisión Europea
COMPAS	Gestión de perfiles penitenciarios para la aplicación de sanciones alternativas (Estados Unidos)
DGI	Índice de gobierno digital de la OCDE
DSIT	Departamento de Ciencia, Innovación y Tecnología (Reino Unido)
EFS	Entidad fiscalizadora superior
EIA	Evaluación de impacto algorítmico
EIN	Evaluación del impacto normativo
E-Leaders	Grupo de Trabajo de la OCDE de Altos Funcionarios de Gobierno Digital
ETC	Equivalente a tiempo completo
FMIS	Sistema de información de gestión financiera
GenAI	Inteligencia artificial generativa
GFP	Gestión financiera pública
GPAI	Alianza Mundial sobre Inteligencia Artificial (GPAI)
i.AI	Incubadora para IA (Reino Unido)
I+D	Investigación y desarrollo
IA	Inteligencia artificial
IAG	Inteligencia artificial general
IAPR	Autoridad Independiente de Ingresos Públicos (Grecia)
IPD	Infraestructura pública digital
ITTI	Inventario de Iniciativas de Tecnológica Tributaria (OCDE)
JRC	Centro Común de Investigación (Comisión Europea)
LLM	Modelo extenso de lenguaje
ML	Aprendizaje automático

NER	Red de Reguladores Económicos de la OCDE
NIST	Instituto Nacional de Estándares y Tecnología (Estados Unidos)
NLP	Procesamiento del lenguaje natural
OCDE	Organización para la Cooperación y el Desarrollo Económicos
OCR	Reconocimiento óptico de caracteres
ODS	Objetivo de Desarrollo Sostenible
OGD	Datos abiertos gubernamentales
OMB	Oficina de Gestión y Presupuesto (Estados Unidos)
ONU	Organización de las Naciones Unidas
OSC	Organización de la sociedad civil
PAC	Comisión de Cuentas Públicas (Parlamento del Reino Unido)
Palkeet	Government Shared Services Centre for Finance and Human Resources (Finlandia)
PET	Tecnología para mejorar la privacidad
PIB	Producto Interior Bruto
RaC	Reglas como código
RAG	Generación con recuperación aumentada
RCT	Ensayo controlado aleatorizado
RGPD	Reglamento General De Protección De Datos (Unión Europea)
RPA	Automatización robótica de procesos
RR.HH.	Recursos humanos
RSI	Retorno sobre la Inversión
TCU	Tribunal Federal de Cuentas (Brasil)
UA	Unión Africana
UE	Unión Europea

Nota: por lo general, los términos solo se incluyen si la abreviatura o el acrónimo se mencionan fuera del contexto en el que se definieron originalmente. No se incluyen las abreviaturas de los nombres de países.

1

Cómo la inteligencia artificial está acelerando la transformación digital de la administración pública

Este capítulo explica cómo la inteligencia artificial (IA) puede acelerar la transformación digital de la administración pública. Sitúa al gobierno como desarrollador y usuario de la IA, más allá de sus funciones tradicionales de inversor y regulador. El capítulo agrupa las oportunidades - productividad (eficiencia y efectividad), capacidad de respuesta y rendición de cuentas - a lo largo del ciclo de políticas, y destaca los requisitos previos en la gestión de datos e información. También describe los riesgos específicos en contextos gubernamentales, incluyendo los riesgos ante la inacción, dentro de los enfoques regulatorios emergentes, y concluye con una visión de una IA confiable en el sector público.

Mensajes clave

- La inteligencia artificial (IA) tiene el potencial de remodelar industrias, economías, gobiernos y sociedades. Sin embargo, su integración en el sector público ha sido limitada.
- La IA puede ayudar a los gobiernos en tres áreas de oportunidad clave: productividad, capacidad de respuesta y rendición de cuentas.
- En cada etapa del ciclo de formulación de políticas, la IA puede aportar beneficios altamente complementarios:
 - Automatizar tareas rutinarias y repetitivas
 - Mejorar la productividad en tareas analíticas o creativas
 - Adaptar los servicios para abordar necesidades personalizadas de la ciudadanía
 - Adaptar enfoques para fortalecer la función pública
 - Mejorar la toma de decisiones y la interpretación del presente
 - Mejorar los pronósticos del futuro
 - Mejorar la gestión y accesibilidad de la información
 - Detectar transacciones inadecuadas y evaluar riesgos para la integridad
 - Permitir a los actores no gubernamentales comprender y colaborar con la administración pública y promover la rendición de cuentas
 - Desbloquear oportunidades para partes interesadas externas, a través de la IA como un bien para todos
- Estos beneficios no son mutuamente excluyentes y se pueden clasificar en cuatro grandes áreas: Procesos y servicios automatizados, optimizados y adaptados
 - Una mejor toma de decisiones, interpretación y pronósticos
 - Mejor rendición de cuentas y detección de anomalías
 - Desbloquear oportunidades para las partes interesadas externas
 - Los gobiernos deben gestionar los riesgos específicos de su uso, tales como: riesgos éticos, riesgos operativos, riesgos de exclusión, riesgos de resistencia pública y riesgos de inacción.

Está empezando a surgir una visión de futuro en la que las entidades gubernamentales desarrollan y adoptan con éxito una IA confiable para la transformación sistemática de los procesos y servicios públicos.

El proceso de transformación digital de la administración pública

La digitalización de la administración pública es esencial para transformar los procesos y servicios de formas que mejoren la capacidad de respuesta y la fiabilidad del sector público, y acerquen a las administraciones públicas a su ciudadanía. Desde la adopción de la Recomendación de la OCDE sobre Estrategias de Gobierno Digital (2014^[1]), la OCDE ha promovido el gobierno digital en los países miembros y otros territorios, apoyándoles en sus esfuerzos por alcanzar la madurez digital del gobierno. Los gobiernos con madurez digital reconocen que la tecnología es un motor estratégico no solo para mejorar la eficiencia, sino también para que las políticas sean más efectivas, y que los gobiernos son más abiertos, responsables, innovadores, participativos y fiables.

La pandemia de COVID-19 subrayó la importancia de las tecnologías y los datos digitales para generar resiliencia económica y social a través de enfoques gubernamentales estratégicos, ágiles e innovadores. Si bien la pandemia y la crisis multidimensional que ésta provocó trastornaron a las administraciones públicas, también brindaron la oportunidad de revisar enfoques estratégicos sobre el uso de herramientas y datos digitales, para mejorar la prestación de servicios públicos. Frente a la falta de alternativas, los gobiernos comprimieron años de avances tecnológicos en semanas y meses. La implementación de soluciones tecnológicas a escala les permitió continuar operando en tiempos de crisis y aseguró la prestación oportuna de servicios a la ciudadanía y empresas (OECD, 2020^[2]; ^[3]). Cuando las tecnologías o los datos digitales no se utilizaron de manera estratégica o efectiva, la crisis destacó las deficiencias y agravó los desafíos, sobre los que los gobiernos están trabajando hasta la fecha.

Hoy en día, los gobiernos de todo el mundo enfrentan niveles decrecientes de confianza pública (OECD, 2024^[4]), al mismo tiempo que experimenta cambios crecientes y rápidamente acelerados provocados por la era digital. En estos tiempos de disruptión acelerada - rápida evolución tecnológica, necesidades sociales cambiantes, crisis inesperadas - es crucial que sean capaces y estén equipados para utilizar las tecnologías digitales y los datos con el fin de aumentar la productividad y la resiliencia de sus administraciones públicas y mejorar la calidad de los servicios públicos.

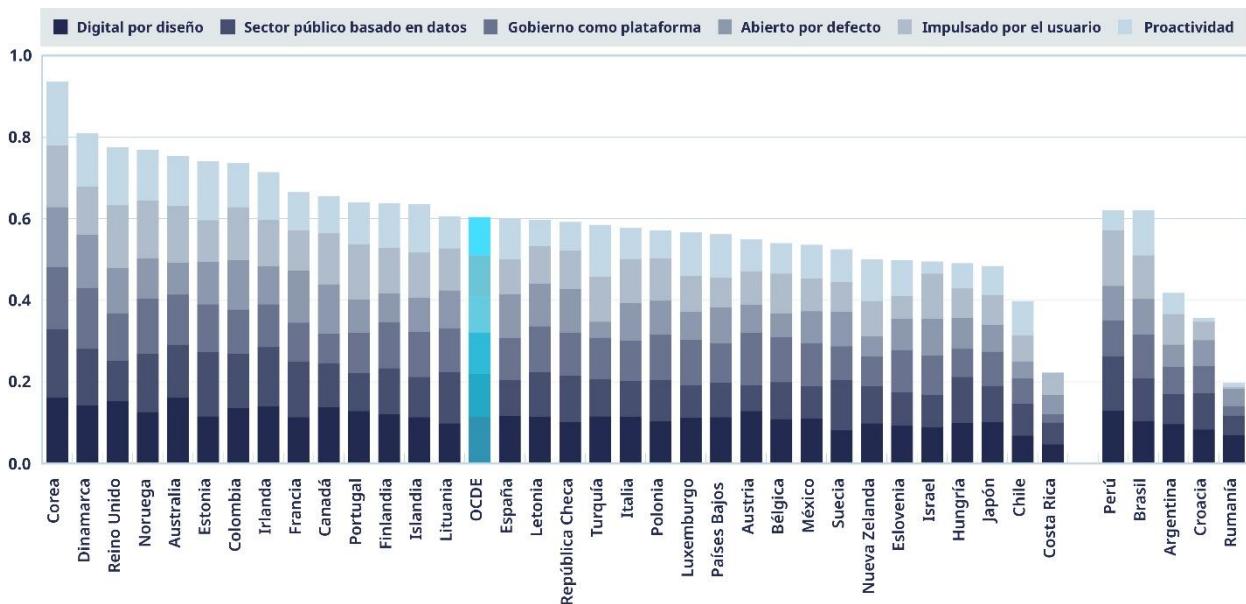
Institucionalizando el Gobierno digital con diversos niveles de madurez

Para aprovechar el potencial del gobierno digital, es fundamental establecer mecanismos institucionales, de coordinación y de políticas, para sostener las transformaciones a largo plazo necesarias y superar las prioridades políticas cambiantes. El marco de políticas de gobierno digital de la OCDE (2020^[3]) establece seis dimensiones críticas para establecer un Gobierno digital:

1. **Digital por diseño:** diseño de políticas que permitan al sector público utilizar herramientas y datos digitales de forma coherente a la hora de formular políticas o transformar servicios públicos.
2. **Impulsado por datos:** desarrollar la gobernanza y los habilitadores necesarios para el acceso, el intercambio y la reutilización de datos en el sector público.
3. **Gobierno como plataforma:** desplegar elementos básicos comunes, como directrices, herramientas, datos, identidad digital y programas informáticos, para avanzar en una transformación coherente de procesos y servicios públicos en todo el sector público.
4. **Abierto por defecto:** apertura más allá de la publicación de datos abiertos, incluidos los esfuerzos por fomentar el uso de tecnologías y datos para comunicarse y relacionarse con diferentes actores.
5. **Orientado a los usuarios:** situar las necesidades de los usuarios en el centro del diseño y la prestación de políticas y servicios públicos, mediante mecanismos de participación con los usuarios y la medición de parámetros para evaluar impacto y satisfacción.
6. **Proactividad:** anticipar las necesidades de usuarios y proveedores de servicios para prestar servicios públicos de forma proactiva.

El Índice de Gobierno Digital (DGI) de la OCDE analiza la madurez de los gobiernos en estas dimensiones (Gráfico 1.1). En este gráfico, es evidente que algunos países están más avanzados que otros en su camino hacia la madurez del Gobierno digital, y la amplia variedad de análisis de la OCDE sobre gobierno digital pone de manifiesto que cada país se enfrenta a sus propios desafíos al hacerlo¹.

Gráfico 1.1. Índice de gobierno digital de la OCDE 2023, resultados compuestos por país



Nota: no se incluyen los datos de Alemania, Grecia, Eslovaquia, Suiza y Estados Unidos (EEUU).

Fuente: (OECD, 2024^[5]).

El creciente papel de la IA en el gobierno digital

La OCDE define un sistema de IA como:

"un sistema basado en máquinas que, para objetivos explícitos o implícitos, infiere, a partir de los datos de entrada que recibe, cómo generar datos de salida como predicciones, contenidos, recomendaciones o decisiones que pueden influir en entornos físicos o virtuales." – Véase la [Exposición de motivos de la](#) OCDE para más aclaraciones (2024^[6]; ^[7]).

Los debates mundiales sobre la IA se centran principalmente en los gobiernos como reguladores o inversores en esta materia, pero existen oportunidades significativas para actuar como desarrolladores y usuarios de IA. Los gobiernos no solo establecen prioridades nacionales, inversiones y regulaciones para la IA, sino que la utilizan cada vez más para diseñar e implementar políticas y servicios. Aunque la publicidad en torno a la IA ha aumentado en los últimos años, el sector público no es nuevo en el uso de la IA: hay miles de proyectos gubernamentales de IA en curso en todo el mundo ².

Desde 2019, la OCDE trabaja para comprender mejor los usos e implicaciones de la IA en el contexto específico del sector público. Esto incluye elaborar documentos fundamentales sobre los fundamentos técnicos de la IA y su uso e implicaciones por y para la administración pública (2019^[8]; ^[9]); análisis focalizado en países específicos (2022^[10]; 2024^[11]; ^[12]); sacar a la luz las tendencias de innovación de la administración, que a menudo implican IA;³ y el establecimiento de un marco preliminar para la IA en el ámbito gubernamental (2024^[13]). La OCDE también ha recopilado detalles sobre cientos de iniciativas públicas relacionadas con la IA⁴.

El DGI 2023 pone de relieve que, aunque algunos gobiernos han puesto en marcha una amplia gama de iniciativas para mejorar su capacidad de uso de IA, su implementación sigue planteando un desafío para la mayoría. En el momento de la publicación del DGI, el 70% de los países había utilizado la IA para mejorar procesos gubernamentales internos, mientras que solo el 33% había utilizado la IA para mejorar el diseño y la implementación de las políticas. Aunque su uso está aumentando, el uso de la IA en la administración pública todavía no ha tenido un impacto transformador. El próximo DGI 2025 incluirá cifras

actualizadas y un análisis comparativo más detallado. También incorporará evidencias cualitativas complementarias para informar aún más sobre cómo los gobiernos pueden implementar habilitadores, salvaguardas, mitigación de riesgos y mecanismos de participación adecuados para adoptar una IA confiable mientras se monitorean los efectos adversos.

Entender el potencial transformador de la IA

La IA es una de las fuerzas más transformadoras del siglo XXI. Está transformando industrias, economías, gobiernos y sociedades a un ritmo sin precedentes. Si los gobiernos y otros actores de la IA logran aprovechar sus beneficios, mientras mitigan sus riesgos, los expertos e investigadores en IA prevén un futuro en el que la IA contribuya a los avances científicos y médicos, como lo son el descubrimiento de nuevos tratamientos contra el cáncer; el crecimiento de la productividad, que pasará de un aumento del 1-7% del producto interno bruto (PIB) mundial para 2033 a un aumento de 10 veces en las próximas décadas; a la eliminación de la pobreza y la reducción de la desigualdad; y a abordar los impactos relacionados con el clima y los desastres naturales (OECD, 2024^[14]).

Si bien la IA ha recibido una intensa atención mundial en los últimos años, la investigación y el desarrollo de la IA se ha llevado a cabo durante más de 70 años. Antes de analizar a profundidad el uso de la IA en la administración pública, resulta útil comprender algunos de los antecedentes de la IA y por qué se ha convertido recientemente en un tema de debate en los hogares, como se analiza en el Recuadro 1.1.

Recuadro 1.1. La evolución de la IA

El panorama de la IA ha evolucionado significativamente desde la década de 1950, cuando el matemático británico Alan Turing planteó por primera vez la cuestión de si las máquinas pueden pensar. Durante décadas, los **sistemas de IA "basados en reglas" o "simbólicos"** dominaron la investigación, utilizando una serie de afirmaciones de "Si-entonces" (Si una condición, entonces una acción) que, cuando se toman en conjunto, darían la apariencia de una acción inteligente. Estos sistemas son limitados y requieren un conocimiento humano significativo para programar las reglas. Todavía se usan hoy en día, como en los robots de software de automatización robótica de procesos (RPA por sus siglas en inglés) que automatizan tareas programadas por humanos. Debido a sus limitaciones, algunos sostienen que los sistemas basados en reglas y la RPA no deberían considerarse IA en absoluto.

El siglo XXI vio grandes avances en la rama de la IA llamada **Aprendizaje Automático (ML por sus siglas en inglés)** que mejoró la capacidad de las máquinas para hacer predicciones a partir de datos históricos. El aprendizaje automático se centra en el desarrollo de sistemas que puedan aprender y adaptarse sin seguir instrucciones explícitas, que imiten la forma en que los humanos aprenden, mejorando gradualmente su precisión, utilizando algoritmos y modelos estadísticos para analizar y extraer inferencias de los patrones de los datos. El proceso de "aprendizaje" mediante técnicas de aprendizaje automático se conoce como "entrenamiento".

La aplicación de técnicas de ML, la disponibilidad de grandes conjuntos de datos y un hardware informático más rápido y potente han convergido, aumentando drásticamente las capacidades, el impacto y la disponibilidad de modelos y sistemas de IA. Inspiradas en el cerebro humano, las **redes neuronales** están formadas por capas de "neuronas", conocidas como "nodos", que procesan entradas con pesos y sesgos para dar salidas específicas. Un subconjunto de algoritmos en el área de las redes neuronales, llamadas redes neuronales profundas (en el campo de estudio y conjunto de técnicas llamadas **aprendizaje profundo**), permite a los sistemas basados en máquinas "aprender" de los ejemplos para hacer predicciones de "inferencias" basadas en una gran cantidad de datos

procesados durante su fase de entrenamiento. Debido a su complejidad, puede resultar difícil entender cómo funcionan o producen un determinado resultado.

Avances conceptuales recientes

En 2017, los investigadores de Google introdujeron un tipo de arquitectura de red neuronal llamada "**transformadores**", que aprenden a detectar cómo los elementos de datos, como las palabras de esta oración, influyen y dependen unos de otros. A diferencia de las redes neuronales anteriores, los transformadores pueden procesar entradas de una secuencia, como palabras de texto, en paralelo. Esto generó un gran progreso al permitir a los desarrolladores de IA diseñar modelos de lenguaje a mayor escala con más parámetros y mayor eficiencia. Esto contribuyó en gran medida a los avances de la **IA generativa** (comúnmente llamada GenAI en inglés), **incluidos los modelos extensos de lenguaje (LLM por sus siglas en inglés)**, que pueden generar contenido novedoso y permitir aplicaciones orientadas al consumidor como chatbots avanzados al alcance de la mano.

Para muchos, la IA se volvió "real" en 2022, el año en que ChatGPT (transformador generativo preentrenado para chat) de OpenAI se convirtió en la aplicación de consumo de más rápido crecimiento en la historia. Los transformadores también contribuyeron a la aparición de **modelos fundacionales**, que se entrena en grandes cantidades de datos que se pueden adaptar (es decir, ajustar) y desarrollar para realizar una amplia gama de tareas posteriores. Aunque a menudo se discuten los transformadores, existen otros enfoques, especialmente para la generación de no texto (por ejemplo, imágenes, video, audio), como las redes generativas adversariales (GANs) y modelos de difusión.

La mayor parte de la IA actual es "estrecha", pero algunos sostienen que están surgiendo formas más "generales" de IA

La mayoría de la IA actual puede considerarse "**estrecha**" (diseñada para realizar una tarea específica), pero algunos expertos sostienen que los modelos de base son una forma temprana de IA más "general". Esto incluye el progreso hacia el hipotético advenimiento de la **inteligencia artificial general (IAG)**, un concepto controvertido que puede describirse como máquinas con nivel humano o mayor inteligencia en un amplio espectro de contextos. Existe un debate sustancial e incertidumbre entre los expertos sobre cuándo o si se pudiera desarrollar la IAG, y qué oportunidades y desafíos potenciales puede brindar.

Si bien algunos expertos creen que la IA se desarrollará en algún momento, las primeras formas emergentes de **sistemas de IA "con agentes"**, que pueden operar de manera un tanto autónoma sin la necesidad constante de orientación humana, apuntan al potencial de sistemas futuros que pueden manejar tareas más generales con una mínima contribución humana. Por ejemplo, los "agentes" basados en LLM ya se han desarrollado para buscar de forma autónoma en Internet e interpretar lo que encuentran en nombre del usuario. Estos sistemas se implantan en una etapa muy temprana, y comprenden muchas limitaciones y riesgos, pero los avances posteriores pueden generar oportunidades para todos los sectores.

A medida que los sistemas de IA se vuelven cada vez más capaces, muchos sostienen que los humanos no deberían diferir las decisiones a las máquinas, sino más bien trabajar en tandem, o en **colaboración máquina-humano**, utilizando la IA para ayudar en la toma de decisiones.

Nota: El documento de la OCDE [Hello World: Artificial intelligence and its use in the public sector](#) (Berryhill et al., 2019^[15]) detalla los fundamentos técnicos y las posibles implicaciones de la IA.

Fuente: (OECD, 2024^[16]), (OECD, 2024^[17]), (OECD.AI, 2023^[18]), (Lorenz, Perset and Berryhill, 2023^[19]), (Berryhill et al., 2019^[15]) (Cognitus, 2024^[20]), (Purdy, 2024^[21]), (NIST, 2025^[22]), (Horvitz, 2014^[23]), (Brizuela et al., 2025^[24]), <https://playbooks.aip.gov.sg/agentic-ai-primer>.

A pesar de la publicidad, el progreso de la IA es limitado

Los datos del Observatorio de Políticas de IA de la OCDE muestran el auge del interés por la IA en los últimos años. Por ejemplo, el Gráfico 1.2 muestra un crecimiento significativo de las inversiones de capital de riesgo en IA a lo largo del tiempo. Si bien el interés es alto, un último estudio de la firma consultora Gartner, coloca a la IA generativa justo más allá de su "máximo de expectativas infladas", e inicia un descenso hacia la "depresión de la desilusión", "a medida que el enfoque empresarial continúa cambiando de entusiasmo en torno a los modelos de cimientos para usar casos que impulsan el RSI" (Gartner, 2024^[25]). Aun así, espera que la IA generativa y algunas otras aplicaciones, como la supercomputación basada en IA y su uso para respaldar y hacer cumplir políticas de gobernanza, confianza, riesgo y seguridad de esta tecnología, alcancen una productividad más plena en tan solo dos o cinco años.

Gráfico 1.2. Las inversiones de capital de riesgo en IA han aumentado con el paso de los años

Inversión de capital global de riesgo en IA en millones de USD por país a partir de 2012



Nota: Puede encontrarse una nota metodológica con más información en <https://oecd.ai/p/methodology>. El aumento de las inversiones en 2021 se debió en parte a un aumento de las inversiones en "atención sanitaria, medicamentos y biotecnología" con IA durante la pandemia de COVID-19. También se observó un repunte significativo ese año en el ámbito de la "Movilidad y vehículos autónomos".

Fuente: OECD.AI (2025), visualizaciones impulsadas por JSI con datos de Preqin, última actualización del 3 de junio de 2025, consultado el 16 de junio de 2025, www.oecd.ai.

Aunque algunos expertos en IA predicen importantes beneficios económicos gracias a la IA, los estudios de la OCDE (2024^[26]) indican un aumento más discreto, estimando que el crecimiento anual de la productividad debido a la IA oscilará entre 0,25 y 0,6 puntos porcentuales en los próximos 10 años en los países más preparados para la IA. La investigación muestra que la IA mejora la productividad individual de los trabajadores (OECD, 2023^[27]; Bengio et al., 2025^[28]). Sin embargo, la evidencia que lo vincula con los beneficios organizacionales y económicos más amplios es débil. Esto se debe, en parte, a que la IA todavía no puede llevar a cabo algunas tareas y no todas las organizaciones o trabajadores están dispuestos a adoptarla. Algunas evidencias sugieren que las empresas que adoptan IA son más productivas y crecen más rápido que las que no lo hacen (Calvino and Fontanelli, 2023^[29]; Hampole et al., 2025^[30]). Sin embargo, esto no debe interpretarse como causalidad. Por ahora, las limitaciones persisten. Según las estadísticas de la Oficina del Censo de Estados Unidos, solo el 5-6% de las empresas estadounidenses utilizan IA para producir bienes y servicios, y solo el 7% planea adoptar IA en los

próximos meses (Williams, 2025^[31]). En una encuesta más global, se estima que solo un 26% de las empresas tienen las capacidades necesarias para obtener valor real de la IA, y solo un 4% han logrado generar un valor significativo (BCG, 2024^[32])⁵.

Más allá de los beneficios económicos, el potencial transformador de la IA para lograr resultados sociales positivos está empezando a mostrar signos prometedores. Sin embargo, todavía no se ha logrado su pleno impacto. Por ejemplo, la IA en la ciencia ha contribuido a un progreso real en robótica, fusión nuclear, descubrimiento de fármacos, generación de anticuerpos y plegamiento de proteínas (OECD, 2023^[33]). A pesar de estos primeros éxitos, muchos usos siguen siendo localizados o experimentales, y el cambio sistémico a escala global todavía está por llegar. La contribución de la IA a la ciencia apenas está comenzando, y en algunas áreas, la tecnología puede haber logrado menos de lo previsto. Por ejemplo, algunos descubrieron que la IA contribuyó poco a la investigación durante la pandemia de COVID-19 (OECD, 2023^[33]). Hasta ahora, la IA ha contribuido en gran medida a avances en un conjunto limitado de ciencias naturales y físicas. Transformaciones similares en otras disciplinas, como las ciencias sociales, han progresado menos a pesar de las grandes expectativas (Manning, Zhu and Horton, 2024^[34]). Por lo tanto, aunque están surgiendo beneficios sociales de la IA, todavía se está desarrollando todo su potencial transformador.

Uso de IA en la administración pública, un contexto único

Además de regular la IA para la sociedad, estableciendo las condiciones y normativas para su uso confiable, los gobiernos se están esforzando por integrar la tecnología para gobernar mejor con IA. Al igual que en el sector privado, el uso de la IA en el sector público promete enormes beneficios, al tiempo que plantea una serie de riesgos y desafíos. Una encuesta de Deloitte (2024^[35]), realizada a 2.770 altos ejecutivos de 14 países, concluyó que los líderes del sector público eran dos veces más propensos que los líderes del sector privado a prever una transformación impulsada por IA en sus organizaciones a corto plazo, pero se mostraban más cautelosos y menos optimistas sobre el hecho de que ello se tradujera en un aumento de la productividad. No obstante, su interés apenas ha comenzado a ocupar un lugar central en la literatura sobre gestión pública y en muchos gobiernos (Mergel et al., 2023^[36]; Mellouli, Janssen and Ojo, 2024^[37]). Esto se debe a una combinación de factores, entre los que se incluyen los recientes avances tecnológicos que han dado lugar a aplicaciones de IA más prácticas y efectivas para el uso gubernamental (véase el Recuadro 1.1); el acceso de los gobiernos a una gran cantidad de datos que pueden utilizarse como insumos para los sistemas de IA; y las presiones fiscales en curso que hacen que la IA resulte atractiva como forma de racionalizar operaciones y reducir costos.

Aunque los esfuerzos del sector pueden aportar algunas lecciones aprendidas y factores de éxito (Santos et al., 2024^[38]), el propósito y contexto dentro del sector público son únicos y presentan una serie de oportunidades y desafíos específicos. Además, el ámbito de la IA es complejo, avanza rápidamente y presenta una pronunciada curva de aprendizaje para servidores públicos y responsables de la formulación de políticas. Sin embargo, si tiene éxito, la aplicación de IA en la administración pública promete tener un impacto significativo en la economía y la sociedad en general al mejorar la calidad y los resultados de servicios públicos, políticas y operaciones gubernamentales (Berglind, Fadia and Isherwood, 2022^[39]).

Los gobiernos tienen una enorme influencia e impacto en la vida de las personas, trayendo consigo un deber de cuidado del bien público, que va más allá del de las empresas (OECD, 2023^[40]; Santiso, 2023^[41]). Por lo tanto, tienen la responsabilidad especial de implementar la IA de forma que se minimice el daño y se priorice el bienestar de las personas y las comunidades. Esto es especialmente cierto cuando se despliega la IA en ámbitos políticos sensibles como la aplicación de la ley, el control migratorio, las prestaciones sociales y la prevención del fraude (OECD, 2024^[13]).

Las administraciones públicas también operan con un mandato único: sirven al interés público y se financian con recursos públicos. Como tales, sus acciones, en particular las que involucran datos y

tecnologías digitales, deben guiarse por principios que defiendan los valores democráticos, los derechos individuales y el estado de derecho. A diferencia de las entidades privadas, que pueden priorizar la eficiencia o el beneficio, se espera que los gobiernos actúen de manera transparente y con la debida consideración por el bien público en mayor medida que las empresas.

Áreas de oportunidad clave y beneficios de la IA en la administración pública

Entre las oportunidades para los gobiernos, en su calidad de desarrolladores y usuarios de IA, se incluyen la posibilidad de transformar la prestación de servicios, la formulación de políticas, las operaciones internas y la supervisión. Este es un momento crucial para los gobiernos de todo el mundo. Al lidiar con los rápidos avances de las tecnologías de IA, están tratando de aprovechar las oportunidades que brinda esta para innovar y modernizar la administración pública, al tiempo que gestiona y mitiga los riesgos asociados, que se exponen a continuación, y los desafíos en materia de implementación, los cuales se exponen en el Capítulo 3.

Incorporar IA en la administración pública abre nuevas posibilidades. A través de años de investigación sobre el tema y trabajando con gobiernos de todo el mundo, la OCDE (2024^[13]) ha identificado tres áreas de oportunidad concretas para el uso gubernamental de la IA:

- **Productividad** con operaciones internas más eficientes y un diseño de políticas, toma de decisiones y prestación de servicios más efectivos. Por ejemplo, el uso de sistemas de IA predictiva para una planificación más efectiva de políticas, la automatización de procesos para acelerar la prestación de servicios y el impulso del desempeño al permitir que los funcionarios se centren menos en tareas rutinarias y más en actividades de misión crítica.
- **Capacidad de respuesta** de las políticas y los servicios públicos mediante la mejora del diseño y los enfoques de prestación de servicios que satisfacen mejor las necesidades cambiantes de la ciudadanía y las comunidades específicas, y mediante la mejora de los mecanismos de participación cívica. Esto incluye ofrecer servicios públicos más personalizados y de forma más proactiva.
- **Rendición de cuentas** mediante el aumento de la capacidad de supervisión y transparencia, por ejemplo, mediante la supervisión en tiempo real. Este cambio puede impulsar la satisfacción pública general y mejorar la percepción de que la administración pública es competente, justa y receptiva, fortaleciendo así la confianza pública en su capacidad de innovación y transformación.

La Tabla 1.1 muestra cómo diversas tareas de IA pueden contribuir a las actividades de la administración públicas, apoyando así estas áreas de oportunidad.

Tabla 1.1. Entender el uso de la IA en la administración pública

Tareas de IA	Actividad gubernamental	Área de oportunidad
<ul style="list-style-type: none"> - Reconocimiento - Detección de eventos - Predicción - Personalización - Apoyo a la interacción - Optimización orientada a objetivos - Generación de contenido - Razonar con estructuras de conocimiento 	Operaciones internas	Productividad (eficiencia y efectividad)
	Formulación de políticas	
	Prestación de servicios	
	Supervisión interna y externa	Rendición de cuentas

Nota: La columna de tareas de IA se ha adaptado a partir de las "Tareas relacionadas con los sistemas de IA" del Marco de la OCDE para la Clasificación de los Sistemas de IA (2022^[42]).

Fuente: (OECD, 2024^[13]).

Principales beneficios de la IA en la administración pública

Para orientar las decisiones de inversión, es crucial que los servidores públicos, especialmente los que toman decisiones en puestos de liderazgo, comprendan los beneficios que puede ofrecer la IA. Un estudio de la Comisión Europea (CE) (2024^[43]), que encuestó a 576 gerentes públicos en siete países, encontró que los beneficios percibidos de la IA influyen significativamente en su adopción. La IA tiene potencial para mejorar la toma de decisiones en diversas etapas del ciclo de formulación de políticas (Gráfico 1.3)⁶. En las secciones siguientes se describen las principales ventajas de utilizar IA en la administración. Estos beneficios no son mutuamente excluyentes y, de hecho, se complementan en gran medida con cierto solapamiento en cuatro conceptos: procesos y servicios automatizados, racionalizados y adaptados; mejor toma de decisiones, interpretación y pronósticos; mayor rendición de cuentas y detección de anomalías; y oportunidades para las partes interesadas externas a través de la IA como un bien para todos. Sin embargo, es importante señalar que el uso de la IA en la administración pública también puede plantear riesgos. Varios de estos riesgos pueden ser la contrapartida de los posibles beneficios o socavarlos. En la siguiente sección se ofrece un análisis específico de los riesgos.

Gráfico 1.3. IA en cada fase del ciclo de formulación de políticas



Fuente: (Pencheva, Esteve and Mikhaylov, 2018^[44]), adaptado a la terminología de la OCDE que figura a continuación.

Procesos y servicios automatizados, racionalizados y adaptados

La automatización basada en IA puede ayudar a automatizar directamente los procesos y servicios existentes, o contribuir a una reformulación completa de la forma en que los gobiernos trabajan, tanto en operaciones internas como en servicios públicos. Al aprovechar sus vastos activos de datos, los gobiernos también pueden usar la IA para desarrollar servicios personalizados específicamente diseñados para individuos y grupos específicos. Estos beneficios no solo hacen que la administración pública sea más eficiente, efectiva y receptiva, sino que también pueden mejorar la calidad del empleo y el bienestar de los servidores públicos al permitirles dedicar más tiempo a trabajos más valiosos y significativos. Se ha demostrado que esto mejora el bienestar de los trabajadores. Casi dos tercios de los trabajadores encuestados por la OCDE (2023^[27]) afirmaron que la IA había mejorado su satisfacción en el trabajo, y algunos estudios muestran que esto puede mejorar el bienestar de los trabajadores (Brougham and Haar, 2017^[45]; Xu, Xue and Zhao, 2023^[46]). Sin embargo, como se analiza más adelante, algunos usos de la IA pueden reducir la calidad del empleo y, potencialmente, provocar el desplazamiento de la fuerza laboral de la función pública.

Automatización de tareas rutinarias y repetitivas

Los gobiernos pueden utilizar IA para mejorar la eficiencia de sus operaciones internas y sus actividades de prestación de servicios, reduciendo el tiempo que los funcionarios públicos dedican a tareas monótonas (OECD, 2024^[13]). Por lo general, estas operaciones internas son repetitivas y no requieren un razonamiento analítico extenso ni el uso del criterio humano. Mediante la automatización de estas tareas, la IA puede agilizar flujos de trabajo, reducir errores, optimizar la asignación de recursos y liberar recursos humanos para actividades más complejas que requieren un mayor uso del criterio humano. En última instancia, esto redundaría en una prestación más eficiente de servicios públicos de mayor calidad (OECD/UNESCO, 2024^[12]).

Tareas repetitivas y que consumen mucho tiempo incluyen:

- **Introducción de datos:** introducción manual de datos en diversos sistemas y bases de datos.
- **Procesamiento de nóminas:** cálculo y procesamiento de salarios y prestaciones de los empleados.
- **Consultas básicas de clientes:** atender consultas rutinarias y proporcionar información al público.
- **Verificación de información:** comprobación y verificación de la autenticidad de documentos.
- **Procesamiento de formularios:** manejo y procesamiento de varios formularios de solicitud.
- **Correo electrónico y correspondencia:** clasificar, responder y gestionar correos electrónicos y la correspondencia oficial.

Los gobiernos pueden utilizar diversos sistemas de IA para este tipo de tareas, desde sistemas simplistas basados en reglas hasta sistemas de aprendizaje automático más avanzados, como los asistentes conversacionales habilitados para LLM. Estos sistemas cuentan con amplias capacidades que van desde la gestión de consultas simples y rutinarias (tanto con la ciudadanía como con los servidores públicos) hasta la generación de contenido totalmente nuevo y personalizado y la optimización de la asignación de recursos (Lorenz, Perset and Berryhill, 2023^[19]; Sapci and Sapci, 2019^[47]).

Por ejemplo, Argentina está automatizando las tareas repetitivas y agilizando la tramitación de casos en la administración de justicia a través de su sistema Prometea, basado en la IA (véase el Capítulo 5, Recuadro 5.62). Sin embargo, como se analiza más adelante, la automatización basada en IA puede plantear riesgos —en ámbitos como la administración de justicia— que los gobiernos deben tener en cuenta. En el caso de Prometea, por ejemplo, Argentina trata de limitar estos riesgos mediante el control humano de cómo se utilizan los sistemas de IA y sus resultados (Corvalán and Le Fevre Cervini, 2020^[48]).

Mejorar la productividad en tareas analíticas o creativas

La IA contribuye además a la productividad mediante el descubrimiento de nuevas ideas y nuevos medios eficientes y efectivos para llevar a cabo el trabajo (Jones, 2022^[49]). Uno de los beneficios más notables de la IA es su capacidad para ayudar a las oficinas gubernamentales a gestionar el análisis y la síntesis de documentación extensa (OECD/UNESCO, 2024^[12]). Mientras una variedad de herramientas basadas en IA puede ser útil, los LLM en particular pueden servir como asistentes de gran alcance para los funcionarios gubernamentales en este sentido, ayudando en tareas como investigación, síntesis y resumen de contenidos (Berglind, Fadia and Isherwood, 2022^[39]). La investigación llevada a cabo con profesionales del conocimiento del sector privado demostró que el uso de la IA puede mejorar el rendimiento de personas y equipos, y también romper los silos funcionales (Dell'Acqua et al., 2025^[50]). Los gobiernos pueden utilizar la IA para reducir la carga de trabajo de los servidores públicos y mejorar el acceso a la información tanto para la ciudadanía como para los servidores públicos.

Algunos usos relevantes incluyen:

- **Procesamiento y clasificación de la información textual:** las herramientas de IA pueden analizar de forma rápida y precisa textos extensos y documentos no estructurados, destacando puntos clave y resumiendo la información. Esto mejora la eficiencia en los departamentos que tratan con grandes cantidades de información, como asuntos jurídicos y procesos administrativos (OECD/UNESCO, 2024^[12]). En consecuencia, se aceleran los flujos de trabajo y se reduce el riesgo de errores humanos, lo que da lugar a resultados más precisos y fiables en operaciones internas y en actividades de prestación de servicios.
- **Redacción de documentos y textos jurídicos:** Los sistemas de IA pueden elaborar borradores preliminares de diversos tipos de documentos utilizando plantillas y la legislación existente. Este proceso ayuda a garantizar el cumplimiento de las normas, mientras ahorra tiempo y recursos. Además, la IA puede hacer referencia cruzada a nuevos borradores con la legislación vigente, identificando posibles conflictos y minimizando errores humanos. Para la redacción de informes, las herramientas de IA pueden ofrecer sugerencias automáticas de estructuras más claras y concisas. Además, puede mejorar la comunicación de informes extensos resumiéndolos en formatos más breves para su difusión a los responsables de la toma de decisiones o al público.
- **Dar sentido a entradas no estructuradas:** la IA puede analizar y sintetizar grandes cantidades de información de procesos participativos, como lo son los comentarios de servicios públicos y consultas, convirtiéndolos en recomendaciones viables. Puede identificar temas recurrentes, agrupar opiniones, detectar valores atípicos, realizar análisis de sentimientos y clasificar opciones de política en función de las preferencias. Este uso puede ayudar a identificar problemas emergentes, a tener mejor en cuenta las preocupaciones de las partes interesadas y a abordar posibles impactos de las políticas.

Estudios muestran que los sistemas de IA generativa reducen el tiempo que las personas dedican a realizar tareas y mejoran la calidad de los resultados. También muestran que estas herramientas tienen un mayor impacto en la productividad de los trabajadores menos calificados, como los empleados subalternos, lo que les permite ponerse al día con sus colegas más antiguos (Noy and Zhang, 2023^[51]; Peng et al., 2023^[52]). Esto puede producir un efecto de coordinación e igualación, e impulsar el aumento de la productividad al ayudar a los trabajadores a realizar muchas de las tareas que normalmente realizan los expertos en la materia (OECD, 2023^[27]). Sin embargo, algunas investigaciones también indican que la IA podría contribuir a aumentar las divisiones entre los trabajadores con mayor y menor nivel de calificación, y que, aunque los sistemas de IA existentes pueden aumentar la productividad de los trabajadores, todavía no pueden realizar muchas tareas que los humanos sí pueden (The Economist, 2025^[53]; Dell'Acqua et al., 2023^[54]; Bengio et al., 2025^[28]). Se necesita más investigación sobre este tema, incluidas las ventajas específicas de la IA y los posibles inconvenientes para los servidores públicos.

A modo de ejemplo, la autoridad tributaria del Reino Unido está utilizando la IA para redactar las descripciones de los puestos de trabajo y para analizar y evaluar las calificaciones de las solicitudes de empleo, con el fin de acelerar la contratación (Recuadro 5.20). Como se analiza más adelante, el uso de la IA puede plantear riesgos si no se hace de manera confiable.

La IA también puede actuar como catalizador de la creatividad y la innovación entre los servidores públicos, así como de la forma en que diseñan e implementan los procesos internos y las políticas y servicios públicos. La IA generativa, por ejemplo, puede respaldar la exploración de alternativas de política, simulaciones de escenarios, redacción de textos legislativos y creación de prototipos de servicios, fomentando una administración pública más imaginativa y experimental. Por ejemplo, el Servicio de Comunicación del Gobierno del Reino Unido está desarrollando una herramienta conversacional impulsada por IA para generar borradores de textos, planes e ideas estratégicas, integrando directrices de comunicación e información sobre la audiencia para garantizar resultados de alta calidad y conformes (Recuadro 5.39). Diseñado como un asistente colaborativo, impulsa la creatividad, reduce la carga de trabajo rutinaria y se está implementando gradualmente después de una prueba piloto exitosa y un refinamiento iterativo impulsado por IA.

Adaptación de servicios a las necesidades personalizadas de la ciudadanía

La IA puede ayudar a los gobiernos a comprender mejor las necesidades y comportamientos de las personas y facilitar la entrega de información y servicios específicos y personalizados a nivel individual (Huang and Rust, 2021^[55]; Flavián and Casaló, 2021^[56]; OECD, 2020^[3]). Esto puede incluir el desarrollo de perfiles de ciudadanos individualizados, la generación y entrega de información personalizada y la configuración de ofertas de servicios en función de necesidades únicas (UN, 2022^[57]). Al mejorar la capacidad de respuesta, estos esfuerzos pueden hacer que los servicios sean más eficientes, efectivos y centrados en la ciudadanía, lo que se traduce en una mayor satisfacción, mejores resultados y un enfoque más ágil y proactivo para satisfacer las necesidades públicas.

Esto permite una mejor respuesta a las necesidades de subgrupos de usuarios, incluidos los grupos vulnerables y desfavorecidos, que pueden tener necesidades específicas y dependientes del contexto (Giest, 2017^[58]). El uso de herramientas de IA para la personalización se ha ampliado a varios sectores de la función pública, a menudo asociados a la prestación de servicios basados en eventos vitales, como los servicios ofrecidos de forma proactiva para el nacimiento de un hijo, las nuevas actividades educativas o el matrimonio (Kopponen et al., 2024^[59]).

Los gobiernos también pueden aprovechar la capacidad de la IA para analizar grandes conjuntos de datos sobre el comportamiento con el fin de comprender mejor la heterogeneidad individual e incorporar factores cognitivos y contextuales, como el momento, la ubicación y las preferencias personales, en el diseño de servicios. Esto permite intervenciones de servicio más adaptables y equitativas que se ajustan a las circunstancias únicas de la ciudadanía, salvaguardando al mismo tiempo la autonomía y toma de decisiones informadas (Mills, Costa and Sunstein, 2023^[60]).

Además de mejorar los propios servicios, la IA puede ayudar a optimizar las comunicaciones sobre la disponibilidad y, por tanto, la adopción de servicios. Utilizando los datos administrativos existentes, la IA puede simplificar los procesos al precompletar (o eliminar) formularios con información conocida y adaptar las preguntas a las personas, reduciendo así el tiempo y el esfuerzo necesarios para llevar a cabo las tareas burocráticas. Esto ya se está haciendo de manera limitada con datos administrativos y algoritmos programados (humanos) en el área de programas sociales (OECD, 2024^[61]). Este enfoque específico puede ayudar a garantizar que la ciudadanía reciba la información que necesita de manera eficiente y que las interacciones con los servicios públicos sean ágiles y fáciles de usar.

Ejemplos de este beneficio en acción incluyen una amplia variedad de asistentes virtuales que pueden responder a consultas únicas de la ciudadanía con información adaptada. Por ejemplo, la autoridad

tributaria de Singapur ofrece un chatbot orientado al público que puede proporcionar información y servicios adaptados a las necesidades individuales (Recuadro 5.4). Otro ejemplo es el de los sistemas de protección social que utilizan la IA para una divulgación proactiva con el fin de promover la aceptación de los servicios entre quienes exigen calidad (Recuadro 5.49).

Sin embargo, es importante señalar que estos servicios suelen requerir una gran cantidad de recopilación y análisis de datos para determinar las características y necesidades individuales. Los gobiernos deben llevar a cabo dichas recopilaciones de datos y utilizar la IA de forma confiable. De lo contrario, como se analiza más adelante, estos esfuerzos corren el riesgo de infringir el libre ejercicio de los derechos humanos, por ejemplo, mediante violaciones de la privacidad.

Adaptación de enfoques para fortalecer la función pública

La IA también puede fortalecer la función pública mediante procesos de contratación más efectivos e inclusivos y programas personalizados de desarrollo continuo. Por ejemplo, en lo que respecta a la gestión de recursos humanos, la IA puede ayudar a los gobiernos a optimizar las decisiones de contratación al ayudar a encontrar a los mejores candidatos para el puesto de trabajo y mejorar la inclusión al controlar los sesgos humanos.

La IA también puede empoderar a los servidores públicos, contribuir al desarrollo del aprendizaje, mejorar la creación de conocimiento y optimizar las plataformas de aprendizaje para mejorar la capacitación de los servidores públicos. Esto incluye la elaboración de estrategias de desarrollo de competencias, el diseño y la realización de cursos de formación adaptados y la aplicación de herramientas para mejorar el acceso a la información. De este modo, la IA puede garantizar que los servidores públicos estén equipados con los últimos conocimientos y competencias necesarios para satisfacer las cambiantes demandas de sus funciones, fomentando una administración pública más receptiva.

Algunas aplicaciones relevantes de la IA son:

- **Desarrollo de material didáctico para funcionarios:** la IA puede crear contenido didáctico (como módulos y materiales de cursos) a partir de documentos originales e integrar información diversa en recursos efectivos. También puede diseñar, estructurar e implementar cursos de aprendizaje en línea. Las herramientas impulsadas por IA pueden actualizar y perfeccionar continuamente estos recursos a medida que se dispone de nueva información, mejorando la educación continua y el desarrollo de competencias de la fuerza laboral del gobierno.
- **Personalización del material y las rutas de aprendizaje para los funcionarios:** la IA puede adaptar el contenido educativo y las vías de aprendizaje para satisfacer las necesidades, preferencias y progresos específicos de cada servidor público, garantizando una experiencia de aprendizaje más efectiva, atractiva y adaptable. Al dotar a los servidores públicos de las competencias adecuadas de manera más eficiente, esta personalización también mejora la capacidad de respuesta del gobierno.
- **Identificación y catalogación de los recursos de aprendizaje:** las herramientas de IA pueden identificar, describir y catalogar múltiples recursos de aprendizaje, de modo que sean fácilmente localizables y compartibles, simplificando así el descubrimiento de recursos y la identificación de contenido relevante para los alumnos. Por ejemplo, la integración de IA en las plataformas digitales puede mejorar las funciones de organización, catalogación y búsqueda.

Por ejemplo, la Comisión Australiana de Administración Pública (APSC) ha llevado a cabo un proyecto piloto de seis semanas para ver si la IA puede diseñar, estructurar e implementar un curso de aprendizaje en línea sobre competencias digitales para el liderazgo (véase el Recuadro 5.22). Para más información sobre este tema, véase el Capítulo 5, sección "Reforma de la función pública".

Mejora de la toma de decisiones, interpretación y pronósticos

Los expertos en IA han identificado la mejora en la toma de decisiones, la interpretación y los pronósticos como los beneficios más importantes de la IA en general, y recomiendan medidas más enérgicas y más inversiones por parte de los gobiernos para lograr estos objetivos (OECD, 2024^[14]). Como se analiza más adelante, los gobiernos deben adoptar estos beneficios con cautela para evitar una dependencia excesiva de la IA, ya que las deficiencias de los sistemas pueden ser difíciles de identificar y, diferir excesivamente el juicio humano a las máquinas, podría permitir la propagación sistemática de errores y causar daños.

Mejorar la toma de decisiones y la interpretación del presente

Al proporcionar información procesable basada en datos, la IA puede ayudar a los gobiernos a mejorar la efectividad y la eficiencia de las acciones focalizadas, asignar recursos e identificar problemas y soluciones en materia de políticas. Por lo tanto, los gobiernos pueden responder de manera más efectiva a los nuevos problemas, pueden garantizar la formulación de políticas informadas, aumentar su capacidad general de respuesta y su rendición de cuentas y, en última instancia, promover el bienestar social.

Algunas formas específicas en las que la IA puede resultar beneficiosa a lo largo del ciclo de formulación de políticas (Gráfico 1.3) incluir:

- **Establecimiento de la agenda y formulación de políticas:** la IA puede desempeñar un papel fundamental para llamar la atención de los responsables de la formulación de políticas y del proceso de establecimiento de la agenda al enmarcar los problemas sociales de manera que respondan mejor a las necesidades sociales reales. Por ejemplo, la IA permite a los gobiernos monitorear y dar sentido a temas emergentes en tiempo real a partir de conjuntos de datos amplios y representativos, mejorando la precisión y la velocidad de la fijación de la agenda (Valle-Cruz et al., 2020^[62]; Kolkman, 2020^[63]). Al detectar los desafíos potenciales con mayor precisión y rapidez, la IA facilita respuestas políticas más rápidas antes de que los problemas se agraven (OECD/UNESCO, 2024^[12]; Höchtl, Parycek and Schöllhammer, 2015^[64]). Mediante la formulación de políticas, la IA puede influir en el proceso de toma de decisiones, al poner en primer plano datos e información importantes sobre problemas (Valle-Cruz et al., 2020^[62]). El análisis basado en la IA puede proporcionar información que estime no solo los impactos probables de las políticas, sino que también identifique a las poblaciones destinatarias y realice diagnósticos económicos y sociales, ayudando así a los responsables de la formulación de políticas a tomar decisiones informadas (Wirjo et al., 2022^[65]; Ubaldi et al., 2019^[9]). La IA también puede ayudar a concebir alternativas políticas, proporcionando una evaluación *ex ante* más profunda de las políticas (Desouza and Jacob, 2014^[66]).
- **Implementación de políticas.** A medida que las políticas pasan a la fase de implementación, la automatización impulsada por la IA, el rápido procesamiento de datos y el análisis en tiempo real mejoran la calidad, velocidad y eficiencia de la implementación de políticas. Los análisis de IA refuerzan y agilizan especialmente la adquisición de datos e información, lo que respalda las mejoras continuas. El análisis de datos en tiempo real puede facilitar mejoras a gran escala y, en última instancia, mejorar la prestación de servicios durante la implementación de políticas (Valle-Cruz et al., 2020^[62]; OECD/UNESCO, 2024^[12]).
- **Supervisión y evaluación.** La IA puede monitorear las intervenciones de política en tiempo real, facilitando una mejor comprensión del proceso de formulación de políticas, contribuyendo con evaluaciones oportunas y precisas de datos permitiendo ajustes rápidos de las políticas cuando sea necesario (OECD, 2019^[67]; OECD/UNESCO, 2024^[12]).

Los datos accesibles y el aumento de la potencia informática están proporcionando a la IA una ventaja competitiva sobre los seres humanos en lo que respecta a la toma de decisiones en algunos casos (Green, 2022^[68]). Por ejemplo, los programas de aprendizaje permanente pueden respaldar el

razonamiento individual y las pruebas demuestran los beneficios reales de la toma de decisiones asistida por IA (Brynjolfsson, Danielle and Raymond, 2023^[69]). Los sistemas de IA pueden mejorar la toma de decisiones al mitigar los errores de razonamiento y cognitivos, ayudando a los humanos a filtrar el "ruido", la variabilidad no deseada en la toma de decisiones humana, y las influencias irrelevantes que pueden conducir a decisiones incoherentes e inexactas (Du, 2023^[70])⁷. El potencial de los sistemas de IA para tomar decisiones basadas en datos está llevando a su adopción en una serie de sectores, incluido el sector público. La IA puede identificar y abordar elementos que distorsionan el juicio humano en diversas aplicaciones en el gobierno (Mills, Costa and Sunstein, 2023^[60]).

Además, el ruido puede ocultar información clave sobre el comportamiento humano, que la IA puede descubrir y cuantificar, contribuyendo así a una formulación de políticas más precisa (Aonghusa and Michie, 2020^[71]). Los algoritmos pueden reducir significativamente el ruido al garantizar la coherencia de los resultados independientemente de los factores contextuales, como el estado de ánimo o la hora del día. Al descubrir patrones de comportamiento previamente no detectados u ocultos, la IA puede permitir a los responsables de la formulación de políticas comprender mejor las tendencias sistémicas y las incoherencias en la toma de decisiones (Ludwig and Mullainathan, 2022^[72]). Si bien la eliminación del ruido no aborda todos los errores, mejora la fiabilidad y reduce las disparidades arbitrarias en las decisiones relativas a funciones gubernamentales como la administración de justicia y la prestación de servicios públicos (Sunstein, 2023^[73]).

A medida que la IA se integra más en la administración pública, es cada vez más importante comprender los procesos psicológicos y cognitivos que moldean las interacciones humanas con estas tecnologías. La administración pública, basada en el comportamiento, puede aportar información valiosa adicional sobre los desafíos que pueden surgir, ofreciendo estrategias para mitigarlos y mejorar la gobernanza y la toma de decisiones (Alon-Barkat and Busuioc, 2024^[74])⁸. Un ejemplo de ello es el uso que hacen los gobiernos de "Polis", una herramienta de participación cívica de código abierto para comprender las opiniones de la ciudadanía, así como los ámbitos de consenso y desacuerdo en materia de políticas públicas (Recuadro 5.36). En el ámbito de la gestión de las finanzas públicas, Corea desarrolló dBrain+, un sistema de información que aprovecha la IA para analizar datos económicos, fiscales y financieros en tiempo real con el fin de optimizar la evaluación de riesgos y la toma de decisiones en las finanzas públicas (Recuadro 5.8).

Mejor pronóstico del futuro

Los sistemas de IA pueden procesar grandes volúmenes de datos de múltiples fuentes, incluidos datos no estructurados, e identificar patrones complejos y señales débiles —señales tempranas de posibles cambios emergentes, amenazas y oportunidades— que no son fácilmente detectables con los métodos existentes. Esto puede mejorar la precisión y puntualidad de las predicciones y puede ser muy beneficioso para otras actividades de prospección estratégica (Fitkov-Norris and Kocheva, 2025^[75]). La analítica predictiva y el pronóstico de IA en el gobierno implican el uso de algoritmos para anticipar tendencias y riesgos futuros. Puede aplicarse ampliamente en diversos ámbitos, como los pronósticos de resultados macroeconómicos y fiscales (por ejemplo, el PIB de "predicción a corto plazo", como se analiza en el Capítulo 5, sección "La IA en la gestión de las finanzas públicas").

Al proporcionar información precisa y oportuna con visión de futuro, la IA mejora la toma de decisiones, la asignación de recursos y la efectividad general de las operaciones gubernamentales. Algunas aplicaciones relevantes de la IA son:

- **Predecir las necesidades de servicios futuros:** La IA tiene un importante potencial en la generación de análisis y previsiones predictivas, lo que permite a los servicios públicos anticiparse a las necesidades y ser proactivos. La IA puede ayudar a predecir las necesidades de servicios futuros, optimizar la asignación de recursos y mejorar la capacidad de respuesta en diversos ámbitos públicos, al permitir el análisis de datos históricos y tendencias.

- **Pronósticos regulatorios:** los reguladores pueden usar IA para descubrir tendencias emergentes y cambios en diversas industrias y planificar proactivamente las respuestas regulatorias. Mediante el seguimiento y análisis continuo de datos de múltiples fuentes, como informes de mercado, redes sociales y artículos de noticias, la IA puede identificar nuevos avances tecnológicos que pueden afectar al panorama regulatorio.
- **Gestión del riesgo de catástrofes:** la IA también puede ayudar a predecir catástrofes naturales mediante el análisis de datos históricos y tendencias actuales. Por ejemplo, los sistemas de IA pueden analizar imágenes satelitales y otros datos para predecir la probabilidad de desastres naturales como incendios forestales y terremotos, lo que permite adoptar medidas proactivas para minimizar daños y mejorar la seguridad pública (Sun, Bocchini and Davison, 2020^[76]; Gupta and Roy, 2024^[77]). Los sistemas de IA pueden ofrecer alertas tempranas e información útil para ayudar a los gobiernos a responder rápidamente a este tipo de eventos y mitigar sus efectos.
- **Anticipación ante riesgos de corrupción y fraude:** los sistemas de IA predictivos están ayudando a los actores de integridad a priorizar los casos para un examen humano más detallado. Aunque la transferencia de estos enfoques de la investigación a los gobiernos es todavía limitada, está creciendo de manera constante. Por ejemplo, los sistemas de IA pueden utilizarse para priorizar casos de riesgo y agilizar procesos de auditoría. También pueden apoyar la focalización de las políticas anticorrupción al proporcionar sistemas de alerta temprana que predicen la corrupción pública basándose en datos como factores económicos y políticos. Las técnicas predictivas también son clave para varias actividades de rendición de cuentas y supervisión gubernamental basada en IA, como la detección del fraude basada en el riesgo, como se analizará más adelante.

Como ejemplo de predicción de las futuras necesidades de servicio público, el sistema PrevocupAI de Portugal pretende predecir las enfermedades relacionadas con el trabajo y los factores de riesgo conexos en la administración pública para minimizar las interrupciones (véase el análisis del Capítulo 5, sección "La IA en el diseño y la prestación de los servicios públicos").

Mejorar la gestión y la accesibilidad de la información

Disponer de datos fiables y de calidad es un requisito previo subyacente para disfrutar de las ventajas de la IA. La IA puede ayudar a maximizar la calidad y la utilidad de los datos, así como la capacidad de los seres humanos y las máquinas para procesarlos y analizarlos (Jarrahi et al., 2023^[78]). Por ejemplo, los sistemas de IA permiten nuevas formas de recopilación de datos, como la detección e identificación automática de elementos en imágenes, grabaciones de audio o vídeos. Las capacidades y la prevalencia de los dispositivos de detección basados en IA han progresado rápidamente, lo que permite la transcripción automática del habla, la detección de movimiento, el reconocimiento de imágenes en vivo y una amplia gama de tareas que antes requerían mano de obra humana (Zhang, Wang and Lee, 2023^[79]; OECD, 2023^[80]). La IA también puede mejorar la forma de almacenar, difundir y aplicar la información. Esto es especialmente evidente con la integración de IA en los sistemas de gestión del conocimiento (Sanzogni, Guzman and Busch, 2017^[80]).

Mejorar la gestión interna de la información dentro de los gobiernos también puede ayudar a estos a proporcionar información y datos abiertos al público. El uso de la IA puede ayudar a minimizar los errores en la gestión de datos, por ejemplo, reduciendo el esfuerzo manual, y cuando se utiliza en combinación con tecnologías que mejoran la privacidad (PET, por sus siglas en inglés) puede también ayudar a mejorar la privacidad y protección de datos personales y la información sensible (OECD, 2025^[81]). Esto, a su vez, permite una publicación más amplia y el acceso a datos abiertos. Por ejemplo, en la administración de justicia, un motor de anonimización impulsado por IA puede identificar y proteger automáticamente los datos personales en el marco de decisiones judiciales (Recuadro 5.64), preparándolos para su publicación como parte de una iniciativa de datos abiertos. La OCDE (2023^[82]) también está explorando el uso de

PET, que son soluciones digitales que permiten recopilar, procesar, analizar y compartir información, al mismo tiempo que se protege la confidencialidad y la privacidad de los datos⁹.

Los asistentes virtuales internos del gobierno para servidores públicos son un buen ejemplo, ya que Albert, de Francia, y Caddy, del Reino Unido, proporcionan una gran cantidad de información histórica e intergubernamental a manos de los servidores públicos para fundamentar decisiones y ayudar a responder a las consultas de la ciudadanía (Recuadro 5.46).

Responsabilidad mejorada y detección de anomalías

Uno de los usos más antiguos de la IA en los gobiernos es detectar anomalías existentes, como el fraude, o prever futuros riesgos para la integridad, mejorando así la rendición de cuentas, la responsabilidad en la gestión pública y la integridad de los programas estatales. Se debe tener especial cuidado en este uso de la IA para evitar resultados posiblemente perjudiciales en los resultados.

Detección de transacciones indebidas y evaluación de riesgos para la integridad

Este tipo de actividades son comunes en una variedad de funciones gubernamentales, entre ellas la contratación pública, la administración tributaria y la gestión de las finanzas públicas. El fraude y los pagos indebidos en los programas gubernamentales pueden ser fuentes importantes de fuga financiera. Por ejemplo, solo en Estados Unidos, el gobierno federal pierde entre 233 y 521 mil millones de dólares estadounidenses al año por fraude (US GAO, 2024^[83]).

Los algoritmos ML son particularmente efectivos en el reconocimiento de patrones, ya que permiten el análisis de grandes conjuntos de datos y detectan valores atípicos, relaciones ocultas (por ejemplo, que indican colusión) y otras anomalías que requieren una investigación humana adicional. Sin la capacidad de analizar datos a escala o de identificar patrones ocultos, estas irregularidades pueden pasar desapercibidas sin IA. Este uso de IA puede mejorar la capacidad de las organizaciones gubernamentales para mantener la integridad y la rendición de cuentas. Por ejemplo, la autoridad tributaria francesa utiliza la IA para analizar fotografías aéreas e identificar bienes inmuebles no declarados (Recuadro 5.1).

Los Gobiernos pueden utilizar la IA para identificar, evaluar, predecir y responder mejor a los posibles riesgos para la integridad, lo que permite una mejor gestión, mitigación e intervenciones oportunas. Por ejemplo, en materia de cumplimiento normativo, los inspectores utilizan cada vez más la IA para evaluar el riesgo que plantean los operadores privados. Esto mejora la focalización de las inspecciones, la protección de los intereses públicos y el uso eficiente de los recursos (OECD, 2018^[84]; OECD, 2021^[85]). La IA ayuda a los inspectores detectando patrones indicativos de posibles incumplimientos, lo que permite una evaluación de riesgos más precisa. Este uso de la IA no solo racionaliza las inspecciones, sino que también mejora la efectividad general de los marcos regulatorios al garantizar que los recursos se destinan a donde más se necesitan.

Permitir a los actores no gubernamentales comprender y colaborar con la administración pública y promover la rendición de cuentas

El sector público puede utilizar la IA para ser más transparente y permitir nuevas formas y canales de interacción entre la ciudadanía, las organizaciones de la sociedad civil (OSC) y las instituciones públicas. De hecho, los expertos en IA sugieren que empoderar a la ciudadanía, las OSC y los interlocutores sociales (por ejemplo, los sindicatos) es uno de los 10 beneficios más importantes que puede reportar la IA. Esto se sustenta en la transparencia gubernamental y el uso de la IA (OECD, 2024^[14]), es decir:

- Ofrecer herramientas basadas en IA que aprovechen los datos abiertos gubernamentales directamente a la ciudadanía para ayudarles a desenvolverse y dar sentido a los procesos y acciones de la administración.

- Permitir la supervisión y el escrutinio por terceros de las operaciones gubernamentales por parte de las OCS y otros agentes no gubernamentales.
- La oferta de oportunidades de participación y de canales a través de los cuales el público pueda dar su opinión y plantear posibles cuestiones sobre el desempeño y las decisiones del gobierno.

Si se ejecuta bien, este uso de la IA puede promover la rendición de cuentas y la integridad pública, reforzar la formulación de políticas y aumentar la confianza de la ciudadanía en la administración pública. Este beneficio, incluidos ejemplos reales y el debate sobre las medidas que los gobiernos deben adoptar para lograrlo, se analiza en profundidad en el Capítulo 5 de la sección "Participación cívica y gobierno abierto".

Desbloquear oportunidades para partes interesadas externas a través de la IA como un común

Un beneficio final del uso de la IA en el sector público es la creación de nuevas oportunidades para las partes interesadas externas, como las empresas y la ciudadanía, proporcionándoles acceso a sistemas de IA desarrollados por la administración pública. Este beneficio de la IA puede no ser tan directamente observable en las actividades gubernamentales inmediatas como los demás beneficios. Sin embargo, tiene el potencial de mejorar la gobernanza pública a través de una mayor confianza en el gobierno, una ciudadanía más informada y capacitada, o incluso a través del crecimiento económico. Los países han puesto en marcha programas de desarrollo sostenible no solo para promover la transparencia y la rendición de cuentas, sino también para promover el espíritu empresarial, el crecimiento económico y la creación de valor que puede que no se hayan entendido o previsto necesariamente desde su inicio (OECD, 2018^[86]). Por ejemplo, los datos de imágenes satelitales Landsat publicados libremente por el Servicio Geológico de los Estados Unidos desde 2008 generan ahora un valor económico de más de 25.000 millones de dólares anuales (USGS, 2024^[87]). Cuando el movimiento de datos abiertos comenzó a finales de la década de 2000 y principios de 2010 (Chignard, 2013^[88]), no se sabía del todo que los datos disponibles servirían como vastos recursos para la capacitación de sistemas de IA una década después.

La naturaleza de la IA difiere claramente de la de los datos en que no es un subproducto natural de una serie de funciones y actividades existentes, y no puede servir como entrada bruta para otros procesos. (Por ejemplo, los datos se han comparado con el combustible, la electricidad o el agua potable para la IA). Sin embargo, el uso de la IA por parte de los gobiernos tiene el potencial de generar bien público al empoderar a las partes interesadas para que mejoren sus capacidades y el acceso a la información con el fin de obtener nuevo valor.

A diferencia de otras aplicaciones —en las que la IA facilita las interacciones entre el gobierno y la ciudadanía, o mejora el suministro de información y la accesibilidad—, determinados usos específicos de la IA en la gobernanza pública pueden aumentar las capacidades de estos agentes, permitiéndoles alcanzar sus misiones y objetivos de manera más efectiva. Por ejemplo, el Gobierno de la India ha puesto a disposición de los agricultores sistemas de IA para ayudarles a garantizar la salud de los cultivos y mitigar los desafíos relacionados con las plagas (Jeevanandam, 2024^[89]). Otro ejemplo es la iniciativa alemana de desarrollo FAIR Forward (Inteligencia Artificial para Todos) que apoya la creación abierta y el uso responsable de sistemas de IA en ámbitos como la agricultura, la protección del clima y la participación ciudadana (OECD, 2023^[90]). Tales enfoques pueden ser particularmente útiles en mercados que aún no cuentan con soluciones del sector privado o que tienen poco atractivo para ellas. Los gobiernos pueden tener los recursos para invertir en campos poco explorados y pueden dar los primeros pasos para construir nuevos mercados, asumiendo riesgos donde otros pueden no estar listos o no estar dispuestos a hacerlo.

El potencial de oportunidades externas positivas se amplificaría aún más si los gobiernos proporcionaran acceso a los facilitadores de IA, como la infraestructura digital en forma de potencia informática (Ho,

2023^[91]) (véase el Capítulo 4, sección "Desarrollo de la infraestructura digital"). En algunos casos, es posible que estos facilitadores ya existan para uso gubernamental, por lo que solo requieren algunos ajustes y una ampliación para ponerlos a disposición de un público más amplio. En otros casos, podrían desarrollarse y suministrarse facilitadores dirigidos a agentes externos. Esto puede democratizar el valor potencial de la IA (OECD, 2024^[14]).

Las entidades no gubernamentales también pueden participar en la creación y el despliegue de herramientas de IA, desempeñando un papel destacado en el fomento de la participación basada en la tecnología (OECD, 2025^[92]).

Para aprovechar los beneficios de la IA, es necesario gestionar sus riesgos

La adopción mundial de IA en todos los sectores plantea cuestiones sobre confianza, equidad, privacidad, seguridad y rendición de cuentas, entre otras. La consideración de estas cuestiones y la gestión de los riesgos de la IA pueden incidir en su adopción y en la materialización de sus beneficios (Tse and Karimov, 2022^[93]). La IA plantea cientos de riesgos¹⁰ y los expertos identifican algunos de los más importantes para la sociedad como (OECD, 2024^[94]; [14]; [95]; 2022^[96]):

- Posibles efectos adversos para algunos grupos o individuos si los sistemas de IA están respaldados por datos inadecuados o sesgados.
- Los sistemas de IA que carecen de suficiente transparencia y explicabilidad erosionan la rendición de cuentas.
- Algunos usos de la IA suscitan preocupación en cuanto a la protección de los datos, la privacidad y la vigilancia.
- La IA puede facilitar amenazas ciberneticas cada vez más sofisticadas.
- En sistemas críticos podrían producirse incidentes y desastres de IA de poca importancia a graves.
- La IA podría contribuir a las perturbaciones del mercado laboral.
- La potencia computacional de la IA requiere un consumo energético muy significativo.

Se trata de cuestiones complejas que están considerando gobiernos, empresas, CSOs y otras partes interesadas pertinentes. Esfuerzos como la adopción de la Ley de IA de la Unión Europea (UE) (2024^[97]) y una variedad de iniciativas de gobernanza a nivel nacional ilustran cómo los gobiernos están asumiendo un papel activo en la gobernanza de la IA para la sociedad¹¹.

Riesgos específicos del uso de IA en la administración pública

Quizás todos los riesgos de la IA podrían implicar a al sector público de alguna manera, como la necesidad de procesos de gobernanza o medidas de mitigación. Sin embargo, definir subconjuntos más reducidos de riesgos es más relevante para los responsables de la formulación de políticas y los servidores públicos a la hora de perseguir el uso estratégico y responsable de la IA en el gobierno. Como se ha mencionado anteriormente, los gobiernos tienen una influencia enorme en la vida de las personas, por lo que el uso que hacen de la IA puede tener un mayor impacto tanto positivo como negativo en la opinión pública. Por lo tanto, las expectativas de rendición de cuentas son mayores para el uso gubernamental de IA. Esto puede verse explícitamente en la Ley de IA de la UE, que clasifica muchos casos de uso de la IA en el sector público como de "alto riesgo", y algunos como de "riesgo inaceptable", por lo que están prohibidos en la UE (Recuadro 1.2). Estados Unidos adopta un enfoque diferente, considerando algunos casos de uso como de "alto impacto" y, por lo tanto, exigiendo ciertas prácticas de gestión de riesgos (Recuadro 1.3). Ley básica de Corea sobre el desarrollo de la IA y el establecimiento de una base de confianza ("Ley básica de la IA") (2024^[98]), que entrará en vigor en enero de 2026 y se aplica a las organizaciones que desarrollan o utilizan IA en el mercado coreano, también designa algunos usos como

de "alto impacto", por lo que requiere medidas reforzadas para garantizar la seguridad y fiabilidad de la IA. La Ley Básica de IA también incluye una designación separada para la IA generativa que incluye requisitos específicos de transparencia y divulgación.

Recuadro 1.2. Niveles de riesgo de la Ley de IA de la UE en relación con la IA en la administración pública

La Ley de IA de la Unión Europea (UE) es un reglamento sobre IA que entró en vigor en agosto de 2024. El reglamento establece obligaciones para la IA en función de sus posibles riesgos y nivel de impacto. La ley identifica cuatro niveles diferentes de riesgos que son relevantes para el uso de la IA por parte del gobierno.

- **Riesgo inaceptable.** Los usos de IA en esta categoría están prohibidos. Algunos ejemplos son la actuación policial predictiva, la identificación biométrica remota en tiempo real (incluido el reconocimiento facial) en espacios públicos para la aplicación de la ley, la puntuación social y la evaluación del riesgo de que una persona cometa delitos. La aplicación de la ley y la administración de justicia se encuentran entre las funciones del gobierno más afectadas por esta categoría. Sin embargo, se aplican algunas excepciones, como los casos de uso relacionados con la seguridad nacional y los que siguen estando sujetos a supervisión judicial.
- **De alto riesgo.** Los usos de IA en esta categoría están permitidos pero regulados debido a su importante potencial de perjuicio para la salud, la seguridad, los derechos fundamentales, el medio ambiente, la democracia y el Estado de Derecho. Debido a su posible impacto en estos aspectos, muchos usos gubernamentales de la IA podrían entrar en esta categoría. Algunos ejemplos son los sistemas utilizados para influir en el resultado de las elecciones y en el comportamiento de los votantes, el tratamiento automatizado de datos personales para evaluar diversos aspectos de la vida de una persona, la evaluación de la idoneidad para acceder a prestaciones y servicios, y los componentes de seguridad utilizados en la gestión y el funcionamiento de infraestructuras críticas. Entre las obligaciones se incluyen el establecimiento de un sistema de gestión de riesgos, la realización de la gobernanza de datos, el establecimiento de documentación técnica para demostrar el cumplimiento y la evaluación obligatoria del impacto en los derechos fundamentales.
- **Riesgo limitado.** Estos usos pueden incluir sistemas destinados a comunicarse con personas, como chatbots, así como sistemas que generan contenido como texto e imágenes. Las obligaciones de transparencia exigen que los desarrolladores e implementadores se aseguren de que los usuarios finales sepan que interactúan con IA.
- **Riesgo mínimo.** Estos sistemas no están regulados, pero se sugiere un código de conducta. Algunos ejemplos son los videojuegos y los filtros de spam.

Fuente: (EU, 2024^[97]).

Recuadro 1.3. El concepto de IA de "alto impacto" en la política de Estados Unidos (EE.UU.)

En Estados Unidos, la política *M-25-21 Aceleramiento del uso federal de IA mediante la innovación, gobernanza y confianza pública* establece un enfoque dualista para los casos de uso de IA en el gobierno: o bien el caso de uso de IA se considera de "alto impacto", o bien no lo es.

La IA de alto impacto es aquella cuya producción sirve de base principal para decisiones o acciones con efectos jurídicos, materiales, vinculantes o significativos sobre:

- Los derechos civiles, las libertades civiles o la privacidad de una persona o entidad; el acceso a la educación, vivienda, seguros, créditos, empleo, recursos o servicios gubernamentales críticos; u otros programas
- Salud y seguridad humana
- Infraestructura crítica o seguridad pública
- Activos o recursos estratégicos, incluidos los bienes de alto valor y la información etiquetada como sensible o clasificada por el gobierno federal.

Los organismos federales son responsables de revisar sus casos de uso de IA y de determinar la aplicabilidad de la definición de alto impacto. La política ofrece ejemplos de 15 categorías de casos de uso de IA que se supone tienen un alto impacto.

Las agencias deben implementar **prácticas mínimas de gestión de riesgos** para los casos de uso de IA de alto impacto. Entre las prácticas mínimas de gestión de riesgos se incluyen las siguientes: realizar pruebas previas al despliegue y una evaluación del impacto de la IA; supervisar de forma continua el rendimiento y los posibles impactos negativos; garantizar una formación humana adecuada, la evaluación, supervisión humana, intervención y rendición de cuentas; ofrecer soluciones y recursos coherentes; y consultar e incorporar comentarios de los usuarios finales y del público. Las agencias deben tener un plan para interrumpir el uso de cualquier sistema de IA de alto impacto que no funcione a un nivel adecuado de conformidad con la política, hasta que se adopten medidas para lograr el cumplimiento.

Los programas piloto limitados deben seguir prácticas mínimas de gestión de riesgo cuando sea posible. Sin embargo, si la persona titular de la Dirección de IA (CAIO por sus siglas en inglés) certifica y se cumplen otros criterios, tal como se detalla en el documento fuente, los programas piloto estarán exentos de las prácticas mínimas de gestión de riesgo. Las CAIO de las agencias también pueden eximir de una o más prácticas mínimas de gestión de riesgo en determinadas circunstancias para un caso de uso específico, aunque deben certificar la validez continua de cada determinación y eximir anualmente, realizar un seguimiento centralizado de ellas y publicar un resumen de cada una.

Fuente: <https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-21-Accelerating-Federal-Use-of-AI-through-Innovation-Governance-and-Public-Trust.pdf>.

En su afán por desarrollar y utilizar la IA, los gobiernos se enfrentan a riesgos que incluyen posibles peligros y amenazas que podrían causar graves problemas a las personas y la sociedad (Valle-Cruz, García-Contreras and Gil-García, 2023^[99]), lo que podría socavar la confianza pública, la legitimidad del uso de IA por parte del gobierno e incluso los valores democráticos. Para abordar estas preocupaciones, es importante identificar y gestionar estos riesgos, tener en cuenta cómo los sistemas de IA pueden afectar de manera diferente a la ciudadanía o a las poblaciones marginadas, ayudar a garantizar la distribución equitativa de los beneficios de la IA y mitigar posibles daños. La consideración continua de los riesgos potenciales es importante porque los riesgos conocidos pueden evolucionar y pueden surgir nuevos riesgos, incluidos los que anteriormente se consideraban fuera del ámbito de la posibilidad.

Este informe identifica cinco tipos generales de riesgo relacionados con el uso de IA en la administración pública, que se presentan a continuación. Más allá de lidiar con estos riesgos, los gobiernos también se enfrentan a una serie de desafíos de implementación al intentar desarrollar y utilizar la IA. Estos desafíos de implementación se analizan en el Capítulo 3.

- **Riesgos éticos:** Entre ellos se incluyen los usos de la IA que socavan el libre ejercicio de los derechos humanos y las libertades, incluida la privacidad, y que pueden infringir, de forma deliberada o inadvertida, valores centrados en la persona. Los algoritmos de IA pueden introducir riesgos éticos desde el ámbito digital al mundo físico a través de algoritmos sesgados y comportamientos poco éticos como la vigilancia invasiva. Entre las principales preocupaciones se incluyen las amenazas a la confianza, la equidad, la libertad, la dignidad, la autonomía individual y los derechos laborales.
- **Riesgos operativos:** Entre ellos se incluyen fallas técnicas y operativas que podrían afectar a la privacidad de los datos, la calidad de los resultados de la IA y las operaciones internas de los gobiernos debido a ciberamenazas, consecuencias no deseadas, alucinaciones, errores sistemáticos y la dependencia excesiva de los sistemas de IA.
- **Riesgos de exclusión:** Estos riesgos están relacionados con las carencias que surgen cuando la ciudadanía que no tiene acceso a la tecnología, o la alfabetización digital, puede quedarse atrás y no poder beneficiarse de los avances de la IA en los servicios públicos.
- **Riesgos de resistencia pública:** Entre ellos, la resistencia pública al uso de IA por parte de las administraciones públicas. Esto puede deberse a la desconfianza en los sistemas o procesos de IA de los gobiernos, o a la difusión de información falsa o engañosa sobre cómo se implementa la IA en el ámbito público y sus posibles impactos.
- **Riesgos de inacción:** Aunque a menudo se pasa por alto, este riesgo incluye los retrasos del gobierno en el uso de la IA para obtener beneficios positivos. Esto puede generar importantes costos financieros y no financieros, que de otro modo se habrían evitado con la adopción exitosa de IA, y una brecha cada vez mayor entre las capacidades del sector público y privado.

Riesgos éticos

Datos inadecuados o sesgados en sistemas de IA

Los sistemas de IA tienen el potencial de perpetuar o generar resultados adversos o perjudiciales, derivados de datos incompletos o inadecuados, así como de la intersección del uso de la IA con prácticas institucionales y sociales de naturaleza sistémica o centrada en el ser humano.

Es importante reconocer que los algoritmos no funcionan de forma autónoma; están moldeados por elecciones humanas en cada etapa, desde la selección del modelo y los datos de entrenamiento hasta el ajuste fino y los ajustes de parámetros. Puesto que los sistemas de IA suelen aprender de los datos generados por el ser humano, reflejan inevitablemente las perspectivas y los comportamientos sociales existentes. En el sector público, los algoritmos rara vez toman decisiones de manera independiente, por lo general sirven como herramientas que informan e influyen en la toma de decisiones humanas en lugar de reemplazarla. Incluso cuando las políticas tienen por objeto la normalización, su correcta aplicación depende en gran medida del contexto local y del "complicado compromiso de múltiples agentes con conocimientos diversos" (Davies, Nutley and Walter, 2008^[100]). Esta interacción entre la IA y el juicio humano significa que los errores pueden persistir no solo en los resultados algorítmicos, sino también en la forma en que se interpretan y aplican las recomendaciones generadas por la IA. Por lo tanto, es fundamental comprender cómo procesan y responden los responsables de la toma de decisiones a la información generada por la IA y garantizar que cuentan con las competencias necesarias para utilizar la IA de forma confiable. Por un lado, la subjetividad cognitiva sigue siendo relevante para dar forma a las interacciones humano-IA a pesar de la neutralidad percibida por la IA, mientras que, por otro lado, las

máquinas carecen de empatía. Por lo tanto, es importante que los servidores públicos puedan utilizar su juicio profesional y su sabiduría práctica para garantizar la equidad mediante el ejercicio de una discrecionalidad juiciosa.

Los sistemas de IA también son muy sensibles a la calidad de los datos de entrenamiento, lo que los hace susceptibles a patrones de sobreajuste¹², a aprender correlaciones espurias y a amplificar errores o sesgos integrados en conjuntos de datos generados por el ser humano. Si bien las llamadas de IA ayudan a eliminar el ruido, como se ha comentado anteriormente como una ventaja, su naturaleza de búsqueda de patrones también puede exacerbarlo, especialmente cuando existen incoherencias sistémicas en los datos de entrenamiento. Sin una supervisión y mitigación rigurosas, existe el riesgo de que la IA refuerce las distorsiones en lugar de adoptar decisiones objetivas y fiables, lo que pone de relieve la necesidad crítica de una cuidadosa conservación y validación de los datos, pruebas del sistema, seguimiento y evaluación del impacto anticipatorios y retrospectivos, y transparencia algorítmica en la toma de decisiones impulsada por IA (Shane, 2019^[101]).

Debido a su potencial para generar impactos sistémicos, disponer de datos y algoritmos insuficientes o sesgados, también puede exacerbar otros tipos de riesgos, como los que se analizan a continuación.

Uso indebido o cuestionable de la IA que genera problemas de vigilancia y privacidad

El uso indebido o la aplicación cuestionable de IA en el gobierno podría perjudicar el libre ejercicio de las libertades y los derechos individuales. Una preocupación destacada es el uso de la IA en la prestación de servicios de seguridad pública, donde permite una identificación y seguimiento eficiente a través de datos biométricos y un monitoreo en tiempo real. Aunque estas herramientas pueden ser útiles para la aplicación de la ley y la prevención del delito, también suscitan preocupación en cuanto a la privacidad de los datos y la vigilancia y el control social por parte de las administraciones públicas. Por ejemplo, durante la pandemia de COVID-19, se utilizó IA para realizar un seguimiento de los movimientos de las personas y garantizar así el cumplimiento de los mandatos de autoaislamiento (Saheb, 2022^[102]). Aunque esta medida tenía por objeto controlar la propagación del virus, suscitó preocupación entre el público en cuanto a su privacidad (OECD, 2020^[103]). De hecho, la vigilancia impulsada por la IA es cada vez más generalizada, tal y como revela el Fondo Carnegie para la Paz Internacional (Feldstein, 2022^[104]) que ha descubierto que 97 de 179 (54 %) países analizados están utilizando tecnologías de IA para la vigilancia pública¹³. El índice identificaba a los Estados no democráticos como uno de los principales motores de la vigilancia de la IA, entre otras cosas, mediante la venta de productos a otros países. Y otros, incluidas las democracias liberales, también son grandes usuarios de la vigilancia con IA (Saheb, 2022^[102]). Como se observa en el Recuadro 1.2, la Ley de IA de la UE regula cuándo es admisible el uso de IA para la vigilancia e identifica casos de uso que constituyen un "riesgo inaceptable" y están prohibidos en toda la UE.

Del mismo modo, la prestación de servicios personalizados suele requerir un procesamiento algorítmico de datos dispersos entre múltiples fuentes de datos públicos y, en ocasiones, privados (Nikiforova et al., 2023^[105]). Esto plantea problemas de privacidad y pone de relieve la necesidad de que los gobiernos persigan tanto una prestación de servicios eficiente como una sólida protección de los derechos individuales a la privacidad. Por ejemplo, para proporcionar servicios sociales personalizados, los gobiernos pueden agregar datos de registros de salud, logros educativos, historial de empleo e incluso actividad de redes sociales. Si bien el objetivo es ofrecer apoyo personalizado e intervenciones oportunas, el procesamiento de cantidades tan grandes de información personal debe hacerse con cuidado y de manera deliberada, con controles establecidos para mitigar las preocupaciones de la vigilancia. Dichas directrices pueden encontrarse en (OECD, 2017^[106]) y (OECD, 1980^[107]). Se necesitan políticas e infraestructuras sólidas para considerar las concesiones entre capacidad de respuesta, transparencia y protección de la información sensible (OECD, 2024^[108]; [61]).

Otro caso de uso indebido es la puntuación social en la prestación de servicios o la formulación de políticas, una práctica en la que las personas se clasifican en función de su comportamiento o

características personales. Los algoritmos de IA analizan datos de fuentes como las redes sociales, las transacciones financieras y los registros públicos para asignar puntuaciones. Estas puntuaciones pueden afectar al acceso a servicios, préstamos y oportunidades de empleo, dando lugar a un trato injusto. El uso gubernamental de estos sistemas está prohibido en la UE como un uso inaceptable (EU, 2024^[97]).

En una encuesta reciente a cientos de expertos en diferentes campos, el 79% dijo que la IA tendrá un impacto negativo en la privacidad de las personas para 2040, una preocupación compartida por el público en general (Rainie and Anderson, 2024^[109]; Fazlioglu, 2024^[110]). En respuesta, los gobiernos tendrán que garantizar que su uso de la IA es confiable para disipar estas preocupaciones.

El posible uso indebido de las herramientas de IA para la vigilancia ciudadana por parte de las autoridades puede dar lugar a extralimitaciones y abusos de poder. La vigilancia continua y la recopilación de datos pueden crear un clima de miedo y desconfianza, especialmente entre las comunidades que ya pueden sentirse desproporcionadamente afectadas por la aplicación de la ley (UN, 2024^[111]). Los gobiernos también pueden usar la IA para fortalecer el poder político, facilitando potencialmente la subyugación y el autoritarismo a gran escala (OECD, 2024^[14]), especialmente por parte de gobiernos no democráticos o de aquellos que no dan prioridad a la protección de los derechos humanos. Algunos expertos sostienen que la IA podría usarse —y en algunos casos ya se está usando— para rastrear y monitorear a la ciudadanía y residentes a escala, utilizando algoritmos y análisis de comportamiento, identificar y suprimir la oposición y perpetuar régímenes totalitarios (OECD, 2022^[112]; Tegmark, 2017^[113]; Clarke and Whittlestone, 2022^[114]; Byler, 2021^[115]).

Por último, la manipulación algorítmica, en la que los sistemas de IA y sus resultados se alteran para producir resultados específicos, es otro posible uso indebido y comportamiento poco ético en el uso gubernamental de la IA (Valle-Cruz, Garcia-Contreras and Gil-Garcia, 2023^[99]). Esta manipulación puede deberse a que los servidores públicos, los responsables de la toma de decisiones o los desarrolladores de sistemas de IA modifican los resultados del sistema de forma intencionada para beneficiar o perjudicar a determinados grupos o personas, o para impulsar un programa específico. La complejidad inherente a la IA también puede dificultar el seguimiento y la comprensión del funcionamiento interno de los algoritmos.

Falta de transparencia y explicabilidad

Los sistemas basados en el aprendizaje profundo son "cajas negras", lo que significa que es difícil describir cómo producen un resultado determinado. Estos resultados se generan indirectamente a partir de la formación en aprendizaje profundo, ya que los ingenieros ajustan continuamente los parámetros hasta que el modelo obtiene una puntuación alta en los objetivos de formación (Clarke and Whittlestone, 2022^[114]). Incluso los científicos que trabajan en modelos avanzados de aprendizaje profundo no entienden el funcionamiento interno de sus sistemas, y les resulta difícil rastrear estos resultados y probar la confiabilidad de estos sistemas a través de métodos tradicionales (OECD, 2022^[112]).

Esto dificulta la detección y mitigación de resultados perjudiciales y plantea desafíos a la hora de determinar la responsabilidad cuando surgen problemas. A medida que los sistemas de IA se van integrando cada vez más en las funciones de gobierno, los sistemas de caja negra podrían dificultar la explicación a la ciudadanía de los motivos que justifican las decisiones respaldadas por IA. También puede exacerbar otros riesgos. Por ejemplo, puede ser más difícil identificar el sesgo algorítmico y sus causas fundamentales en sistemas opacos. Los servidores públicos también podrían tener cada vez más un falso sentido de confianza en sistemas de IA aparentemente eficientes pero defectuosos, ya que estos defectos pueden ser inobservables (lo que contribuye al sesgo de automatización, que se analiza más adelante) (OECD, 2024^[14]; Russell, 2019^[116]). Esto puede erosionar la rendición de cuentas de los gobiernos y desempoderar a la ciudadanía al limitar su capacidad para tomar decisiones informadas o, potencialmente, al someterlos a decisiones opacas e imperfectas basadas en la IA (Lima et al., 2022^[117]).

Riesgos operativos

"Sesgo de automatización": dependencia excesiva de la IA

Muchas personas perciben que los sistemas de IA y sus decisiones son neutrales e imparciales, lo que lleva a los usuarios a aceptar resultados sin escrutinio. Los estudios sobre la toma de decisiones asistida por IA han identificado una tendencia a sobreponer las recomendaciones algorítmicas, a menudo asumiendo que su predicción es más confiable que el juicio humano, incluso cuando el propio sistema de IA tiene limitaciones (Alon-Barkat and Busuioc, 2024^[74]). Este "sesgo de automatización" (la propensión de las personas a confiar en los resultados de la IA porque parecen racionales y neutrales) puede llevar a la aplicación de la toma de decisiones en materia de IA a un número cada vez mayor de desafíos sociales (Horowitz, 2023^[118]; Alon-Barkat and Busuioc, 2022^[119]) — tal vez para evitar conversaciones y decisiones difíciles sobre enfoques humanos a estas cuestiones. Algunos expertos afirman que este hábito está creando una "fe ciega" en la tecnología, un fenómeno problemático que puede reforzar los problemas sistémicos existentes contra ciertos grupos o individuos y contribuir al abandono del sufrimiento humano y la erosión de la empatía (Goldman, 2023^[120]; Olson, 2023^[121]).

El llamado "sesgo de automatización" en el gobierno se produce cuando las organizaciones públicas o los funcionarios dependen demasiado de los sistemas de IA para la toma de decisiones o la ejecución de tareas. Por ejemplo, si los profesionales sanitarios confían demasiado en las sugerencias automatizadas por IA sin realizar comprobaciones cruzadas, podrían pasar por alto información crítica o hacer diagnósticos incorrectos. Esta dependencia excesiva puede hacer que los usuarios no reconozcan los errores, acepten resultados incorrectos de IA y reduzcan la supervisión y el juicio humano (Passi and Vorvoreanu, 2022^[122]; Klingbeil, Grützner and Schrec, 2024^[123]). La IA podría adoptarse sistemáticamente sin evaluar plenamente la precisión y las posibles consecuencias, lo que daría lugar a la dependencia de los sistemas de IA y a la propagación de errores compuestos por todos los sistemas (Valle-Cruz, Garcia-Contreras and Gil-Garcia, 2023^[99]).

Uno de los factores que contribuyen a estos problemas son las alucinaciones, que ocurren cuando los sistemas de IA generativa presentan hechos de una manera creíble, a menudo cuando no se encuentra una respuesta correcta en los datos de entrenamiento. Esto puede ser perjudicial en contextos como la toma de decisiones por parte de los gobiernos, en los que puede dar lugar a decisiones o acciones equivocadas (OECD, 2024^[94]; Beltran, Ruiz Mondragon and Han, 2024^[124]). Por ejemplo, un sistema de IA diseñado para responder a las consultas del público podría proporcionar datos erróneos sobre servicios públicos o enumerar servicios que no existen. Si no se mitigan cuidadosamente, estas consecuencias no deseadas y los resultados incorrectos pueden ampliarse rápidamente, afectando a grandes poblaciones o a decisiones gubernamentales internas críticas.

Los sistemas de IA también pueden ser propensos a errores y fallas de funcionamiento sencillos, lo que puede dar lugar a desviaciones sistemáticas e imprecisión en los resultados algorítmicos. Si estos sistemas no se implementan cuidadosamente o no son técnicamente fiables, se corre el riesgo de que disminuya la confianza pública, incluida la de los servidores públicos. La ciudadanía puede experimentar frustraciones frecuentes debido a la tramitación incorrecta de sus solicitudes o a los retrasos en la prestación de servicios. Por ejemplo, los sistemas automatizados que gestionan los envíos de formularios, las consultas de servicio al cliente y las solicitudes de pago deben operar con gran precisión; de lo contrario, los errores o las fallas técnicas pueden conducir a una percepción de incompetencia y falta de fiabilidad en los servicios públicos. Los problemas técnicos o los errores en las interacciones de los chatbots, como las respuestas incorrectas, la incapacidad de comprender las consultas o las interrupciones del sistema, pueden erosionar aún más la confianza, haciendo que la ciudadanía sienta que los sistemas de IA no son fiables ni efectivos.

La aversión al riesgo también influye en el "sesgo de automatización". Los servidores públicos pueden tener miedo de asumir la responsabilidad personal por las decisiones por miedo a meterse en problemas

más tarde. Un ejemplo es cuando un ser humano decide en contra del consejo de un sistema de IA, y resulta ser la decisión equivocada. En tales casos, el funcionario parecerá doblemente responsable. Como se observa en el Recuadro 5.43, algunos sistemas públicos de IA se han diseñado incluso para hacer que los servidores públicos justifiquen su revisión si van en contra de los consejos de un sistema de IA, lo que aumenta la carga de su trabajo y refuerza los incentivos para seguir las recomendaciones del sistema.

La investigación también destaca el concepto de adherencia selectiva, donde los tomadores de decisiones son más propensos a seguir recomendaciones algorítmicas cuando se alinean con sus creencias preexistentes o estereotipos sociales (Alon-Barkat and Busuioc, 2022^[119]). Esto puede dar lugar a una toma de decisiones públicas distorsionada, ya que los funcionarios pueden utilizar inconscientemente los resultados de la IA para justificar sesgos preexistentes en lugar de evaluarlos de forma crítica.

Algunas investigaciones sugieren que la dependencia excesiva de la IA puede contribuir a una disminución de las habilidades cognitivas humanas, reduciendo la exploración, creatividad y el pensamiento independiente, a medida que las personas se acostumbran a las soluciones generadas por la IA. Los estudios sugieren que la dependencia frecuente de los sistemas de apoyo a la toma de decisiones impulsados por la IA puede conducir a una descarga cognitiva, reduciendo el compromiso de las personas con el pensamiento crítico y la resolución independiente de problemas (Gerlich, 2025^[125]). La toma de decisiones impulsada por IA también entraña el riesgo de promover la homogeneización del comportamiento, ya que sus resultados suelen reflejar una diversidad limitada de perspectivas. Este estrechamiento de perspectivas podría dificultar el pensamiento adaptativo y reducir la capacidad de las sociedades y los gobiernos para navegar por la incertidumbre y el riesgo (Meng, 2024^[126]). Aunque el exceso de confianza en la IA puede plantear riesgos, las herramientas de IA también pueden ayudar a los operadores humanos a interpretar y cuestionar decisiones complejas en materia de IA, desalentando la dependencia excesiva (OECD, 2024^[94]).

Reducción de la calidad del empleo de servidores públicos

Si bien la IA puede mejorar la calidad del empleo y el bienestar de los servidores públicos, algunos usos pueden tener el efecto contrario. El uso de herramientas de gestión algorítmica está aumentando significativamente, alcanzando una tasa de adopción del 90 % en las empresas estadounidenses y del 79 % en la UE (Milanez, Lemmens and Ruggiu, 2025^[127]). Si bien no se han realizado estudios específicos del sector público, se han planteado preocupaciones tangibles sobre los impactos negativos existentes de la IA y las herramientas algorítmicas en la calidad del empleo, incluida la intensificación del trabajo, el aumento del estrés, la percepción de reducción de equidad y la vigilancia en el lugar de trabajo (OECD, 2023^[27]). Por ejemplo, la IA podría hacer que los trabajos fueran menos gratificantes y más estresantes al incentivar nuevos tipos de vigilancia en el lugar de trabajo, o nuevas formas de "taylorismo digital" hipereficiente pero agotador, en el que el trabajo esté sujeto a una mayor vigilancia y regulación, incluso a través de la gestión algorítmica (UC Berkeley, 2021^[128]; EC, 2025^[129])¹⁴. Investigaciones posteriores han demostrado que esta vigilancia de la IA puede perjudicar la salud mental (APA, 2023^[130]) y que la gestión de tareas con IA puede erosionar la autonomía y la voz de los trabajadores, reduciendo la comprensión humana sobre cómo se gestiona el trabajo (Gmyrek, Berg and Bescond, 2023^[131]). Sin embargo, si se usan bien, también se ha demostrado que estas herramientas mejoran la seguridad y el bienestar de los trabajadores (por ejemplo, alertando a los trabajadores sobre los peligros y riesgos, o identificando el agotamiento) (EC, 2025^[129]). La gestión algorítmica y sus impactos ya se están viendo y estudiando con respecto a su uso en el sector privado (OECD, 2023^[27]). La investigación sobre la gestión algorítmica en el sector público es escasa, aunque su uso se está expandiendo rápidamente (EC, 2025^[129]).

Tensiones en materia de privacidad y gobernanza de datos

El desarrollo y despliegue de sistemas de IA plantea desafíos en materia de privacidad y gobernanza de datos a lo largo del ciclo de vida de la IA (OECD, 2024^[132]; forthcoming^[133]). En la fase de entrenamiento

de la IA, muchos desarrolladores dependen de fuentes de acceso público para crear conjuntos de datos de entrenamiento de la IA, que pueden incluir, de forma intencionada o involuntaria, datos personales o información sujeta a derechos de propiedad intelectual. Sin embargo, el hecho de que los datos sean accesibles a través de Internet no significa automáticamente que sean gratuitos para su recopilación y uso en la formación de modelos de IA. Además, es posible que se hayan compartido datos personales consintiendo otros usos, lo que no incluye necesariamente la formación sobre modelos de IA (ICO, 2023^[134]). La recopilación de datos personales para la formación en sistemas de IA, al igual que cualquier actividad de procesamiento de datos, está sujeta a los principios de privacidad comúnmente reconocidos establecidos en la Recomendación de la OCDE relativa a las Directrices que Regulan la Protección de la Privacidad y los Flujos Transfronterizos de Datos Personales (Directrices de Privacidad de la OCDE) (1980^[107]). Estos principios exigen que los datos personales se obtengan a través de medios legales y justos, con el conocimiento del interesado, y que cualquier otro uso de los datos no sea incompatible con los fines originales.

Otro aspecto importante a tener en cuenta es la capacidad de los modelos de IA para memorizar datos personales dentro de sus parámetros durante la fase de entrenamiento. En consecuencia, los grandes modelos de lenguaje que respaldan las herramientas de IA generativa basadas en textos plantean un riesgo particular de acceso y uso no autorizados de datos personales de terceros sin el conocimiento de las personas afectadas (Brown et al., 2022^[135]). Algunas investigaciones también muestran que los modelos generativos de IA pueden inferir atributos personales del sujeto de datos a partir de texto con alta precisión, pero a bajo costo (Staab et al., 2023^[136]). Esto plantea problemas de privacidad no solo porque estas inferencias pueden revelar información personal o características personales, especialmente cuando dichos rasgos no estaban destinados a compartirse.

Durante la fase de despliegue, los sistemas de IA también pueden estar en tensión con los derechos de las personas a acceder, corregir y, cuando sea necesario, eliminar sus datos personales (también conocidos como el "principio de participación individual" en las Directrices de Privacidad de la OCDE). Por ejemplo, el cumplimiento de los derechos de las personas a que se eliminen o corrijan sus datos puede ser técnicamente complejo y exigir muchos recursos, ya que podría requerir la identificación de puntos de datos específicos relacionados con una persona dentro de conjuntos de datos no estructurados o, en algunos casos, volver a entrenar el modelo de IA. Además, las investigaciones realizadas antes del uso generalizado de modelos generativos de IA sugieren que, en ciertos casos, es posible reconstruir o desanonymizar los datos originales de entrenamiento analizando el comportamiento de un modelo que incluye esos datos (Salem et al., 2018^[137]). Para hacer frente a estas tensiones, promover una mayor cooperación internacional entre las comunidades de privacidad de datos y de IA puede contribuir a armonizar las prácticas en materia de datos con el desarrollo y el uso de la IA. Por ejemplo, el Grupo de Expertos de la OCDE en IA, Datos y Privacidad¹⁵ está explorando respuestas políticas sobre gobernanza de datos y privacidad en el contexto de la IA, con la participación de expertos de múltiples sectores y disciplinas de todo el mundo.

Amenazas informáticas

Los sistemas de IA incorporados a los procesos gubernamentales pueden ser vulnerables a las amenazas informáticas, que pueden dar lugar a filtraciones de datos, violaciones de privacidad y pérdida de funcionalidad. La amplia recopilación y análisis de datos personales para aplicaciones de IA puede provocar la pérdida, alteración o divulgación no autorizada de estos datos, lo que infringe los derechos de privacidad individuales (Beltran, Ruiz Mondragon and Han, 2024^[124]). El acceso no autorizado y las violaciones de datos pueden comprometer la integridad operativa y los datos personales, lo que podría resultar en robo de identidad, fraude financiero y otras violaciones de la privacidad, socavando la confianza pública en las instituciones gubernamentales y dando lugar a consecuencias legales. Además, los ciberactores malintencionados podrían manipular los sistemas de IA, alterando sus resultados y causando decisiones o acciones erróneas (Brundage et al., 2018^[138]; Gopireddy, 2024^[139]). Los riesgos cibernéticos

pueden provenir de malos actores externos o de amenazas internas dentro del gobierno (Eshelby, 2025^[140]). En general, la ciberseguridad puede verse como una función horizontal del gobierno en sí misma, y de hecho representa una de las áreas de mayor adopción de IA en el gobierno para mejorar la seguridad de los sistemas informáticos gubernamentales (Mariani, Kishnani and Alibage, 2025^[141]). Sin embargo, esta función es muy especializada y no está incluida en el ámbito principal de este informe, por lo que no es un tema de análisis y debate exhaustivos. Sin embargo, la OCDE ha puesto en marcha una línea de trabajo sobre seguridad digital que ha llevado a cabo una labor pertinente¹⁶.

Riesgos de exclusión

Exacerbación de las brechas digitales

El riesgo de omitir a las personas al utilizar la IA en la administración pública está estrechamente relacionado con las brechas digitales (UN, 2024^[111]), es decir, la brecha entre quienes pueden acceder y utilizar las tecnologías de la información y la comunicación y quienes no pueden hacerlo. Esto es especialmente evidente en las actividades de prestación de servicios públicos y formulación de políticas, especialmente en el uso de la IA para el análisis predictivo, la pronósticos y la personalización de servicios.

La brecha digital podría dificultar el acceso a los beneficios y el valor público que ofrece la IA. Este es el caso, en particular, de poblaciones que carecen de la infraestructura y la alfabetización digital necesarias para acceder a servicios públicos impulsados por la IA (OECD, 2024^[14]; ITU, 2023^[142]). La IA en el ámbito gubernamental puede ofrecer ventajas, como información adaptada, tiempos de respuesta más rápidos y una mejor prestación de servicios, pero puede ser inaccesible para quienes no tienen acceso a Internet o conocimientos digitales. La transición de los gobiernos hacia la digitalización puede intensificar las barreras a los servicios digitales públicos entre determinados segmentos de la población. Las investigaciones realizadas en Canadá revelaron que los usuarios de zonas rurales, las mujeres y las niñas y los hogares de bajos ingresos se vieron negativamente afectados por el impulso hacia los servicios digitales en primer lugar debido a la pandemia de COVID-19 (Singh and Chobotaru, 2022^[143]). En Noruega, la digitalización y automatización del sistema de concesión de prestaciones por hijos a cargo permitió que la mayoría de los beneficiarios recibieran la prestación automáticamente. Al mismo tiempo, hizo necesario que otros beneficiarios lo solicitaran manualmente, una carga que afectaba desproporcionadamente a los segmentos de bajos ingresos (Larsson, 2021^[144]).

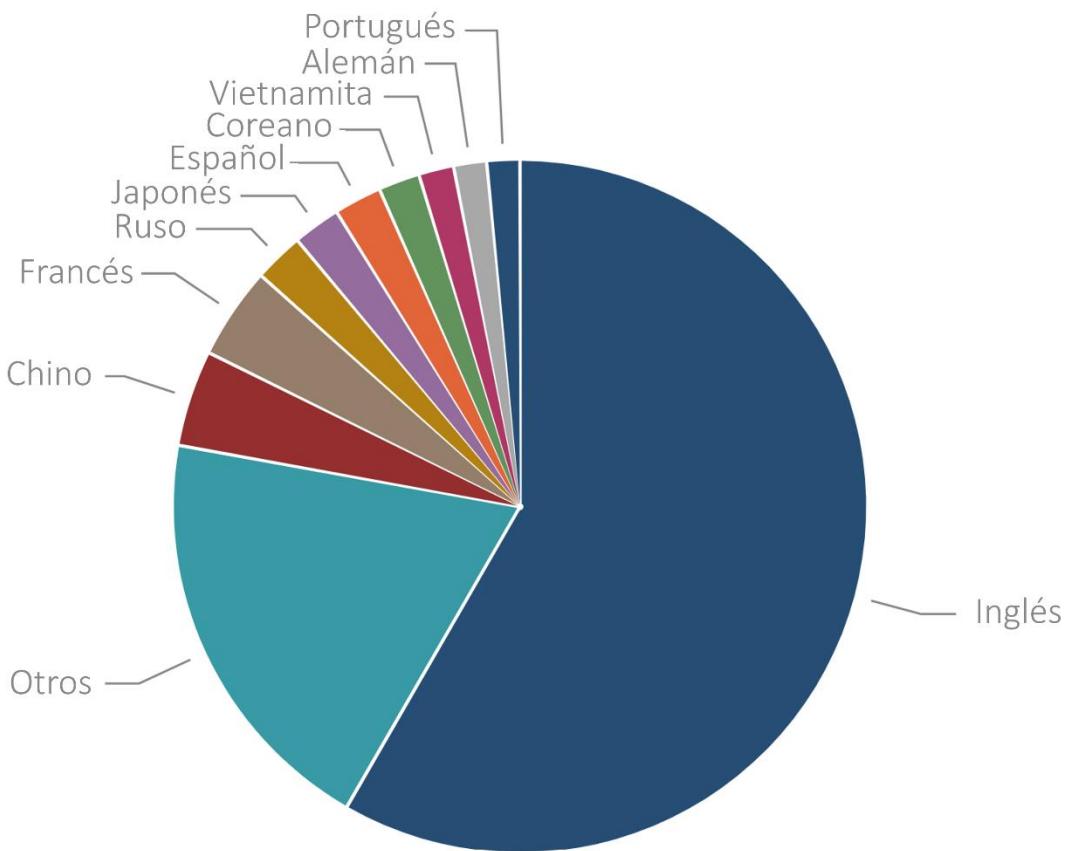
Las brechas de datos son otra forma de exclusión causada por la IA que está vinculada a datos no representativos, como consecuencia de las brechas digitales existentes. Las personas sin acceso a Internet suelen estar ausentes de los datos utilizados para desarrollar algoritmos. Dado que la mayoría de los datos se recopilan principalmente a través de Internet y a través de interacciones en línea, el uso generalizado de IA puede excluir a estas personas, ya que los algoritmos utilizados para fundamentar la formulación de políticas carecen de datos representativos. Estas deficiencias dificultan que los gobiernos respondan adecuadamente a las necesidades de toda la ciudadanía, lo que da lugar a conjuntos de datos insuficientes o inadecuados en los procesos de toma de decisiones impulsados por la IA y en la prestación de servicios. Por ejemplo, las brechas de datos limitan el potencial de beneficios de la IA, como los servicios de IA personalizados, por lo que solo resultan útiles y precisos para poblaciones ricas en datos (UNESCO, 2019^[145]; Perry and Turner Lee, 2019^[146]; Dieterle, Dede and Walker, 2022^[147]).

Otra forma de brecha digital es la subrepresentación de las lenguas (Röttger et al., 2024^[148]; Peixoto, Canuto and Jordan, 2024^[149]). Los conjuntos de datos de formación tienden a representar en exceso los idiomas ampliamente utilizados, como se observa en el Gráfico 1.4. Este desequilibrio puede dar lugar a que los sistemas de IA no presten un servicio efectivo a los grupos lingüísticos no dominantes. Los esfuerzos de preservación de idiomas, como Donate a Speech de Estonia, reducen las actuales limitaciones de la tecnología del habla, que favorece los idiomas más ampliamente reconocidos, y mejoran

la prestación de servicios (OECD, 2023^[40]). En la sección "Crear fundamentos sólidos para los datos" del Capítulo 4 se tratan consideraciones y ejemplos adicionales.

Gráfico 1.4. Más de la mitad (59%) de los conjuntos de datos de entrenamiento de IA de código abierto están en inglés

Desglose porcentual de los idiomas para los conjuntos de datos de entrenamiento de IA de código abierto sobre Hugging Face



Nota: Este gráfico representa la distribución de idioma de todos los conjuntos de datos. Los conjuntos de datos multilingües y de traducción en Hugging Face contienen más de un idioma y, por lo tanto, se cuentan por partida doble. Más información metodológica disponible en: <https://oecd.ai/huggingface>.

Fuente: OECD.AI (2024), visualizaciones impulsadas por JSI con datos de Hugging Face. Última actualización: 5 de junio de 2025 (consultado el 16 de junio de 2025).

Desplazamiento de la fuerza laboral de los funcionarios públicos

La OCDE (2023^[27]) concluyó que, si bien la IA es capaz de automatizar tareas no rutinarias, sus impactos futuros sobre la mano de obra son ambiguos; dependen del equilibrio entre el desplazamiento de la mano de obra humana por parte de la IA, el aumento de la demanda laboral debido a la mayor productividad generada por la IA y la creación de nuevos puestos de trabajo causados por la adopción de IA. La IA puede mejorar la productividad pública al automatizar tareas, y también tiene el potencial de reducir la necesidad de mano de obra humana, lo que da lugar a la necesidad de recalificar a los servidores públicos para que asuman tareas más significativas (Peixoto, Canuto and Jordan, 2024^[149]).

Al mismo tiempo, el despliegue de la IA está impulsando una creciente demanda de competencias relacionadas con IA en toda la economía, y el sector público no es la excepción. A medida que aumentan

las funciones relacionadas con la IA, pueden disminuir los esfuerzos de contratación y retención para puestos no relacionados con la IA o tradicionales, lo que pone de relieve la necesidad de una mano de obra calificada en IA y tecnologías relacionadas para satisfacer la evolución de la demanda (Acemoglu et al., 2022^[150]).

Las tecnologías de IA tienen repercusiones en una amplia gama de profesiones y sectores, afectan a trabajadores de todos los niveles de calificación e influyen en los mercados laborales (OECD, 2023^[27]). Mientras que algunos trabajadores públicos pueden adaptarse a la IA e incluso ver mejorado su trabajo, otros, como los trabajadores de más edad y menos calificados que realizan tareas que son fáciles de automatizar, se enfrentan a importantes riesgos. Esto sugiere que los beneficios de la IA no se reparten de manera uniforme entre los servidores públicos (Milanez, 2023^[151]). Por ejemplo, los chatbots basados en IA se utilizan ahora de forma generalizada para la prestación de servicios y la comunicación centrada en la ciudadanía con el fin de responder preguntas básicas y proporcionar información, lo que reduce la demanda de representantes gubernamentales de servicios humanos al cliente (Acemoglu, 2024^[152]).

Durante muchos años, las preocupaciones sobre la automatización del empleo se centraron en la mano de obra poco calificada, pero la IA está ayudando cada vez más a las tareas cognitivas de alto rendimiento, lo que puede afectar a los servidores públicos, como a los analistas de políticas. La IA generativa puede producir textos significativos, realizar análisis de datos e incluso proponer estrategias políticas para abordar desafíos complejos. Muchos funcionarios tendrán que recibir formación para colaborar efectivamente con la IA y centrarse en el pensamiento estratégico de mayor nivel y en la toma de decisiones.

Riesgos de resistencia pública

La ciudadanía puede aceptar selectivamente los resultados basados en IA, contribuyendo potencialmente a cometer errores

La comprensión de la interacción humano-IA es cada vez más considerada como un desafío clave para la administración pública, con importantes implicaciones para la confianza y la legitimidad. A medida que los gobiernos integran la IA en la toma de decisiones, la aceptación pública de estos sistemas se vuelve crítica. Las investigaciones sugieren que las decisiones algorítmicas, a pesar de su promesa de neutralidad, no siempre se perciben como justas o legítimas, particularmente cuando contradicen las expectativas de los individuos o carecen de transparencia (Alon-Barkat and Busuioc, 2022^[119]).

La ciudadanía procesa las decisiones algorítmicas a través de percepciones cognitivas y creencias previas; a menudo aceptan selectivamente recomendaciones generadas por la IA que se alinean con sus expectativas mientras se resisten a las que dicen lo contrario (Alon-Barkat and Busuioc, 2022^[119]). Esta adherencia selectiva puede reforzar los errores en la toma de decisiones del gobierno, ya que las personas pueden ser más propensas a confiar en las predicciones algorítmicas cuando confirman estereotipos preexistentes o conocimientos previos.

Falta de empoderamiento público y comprensión del uso que hace el gobierno de la IA

A menudo hay una falta de conocimiento y comprensión entre el público sobre la IA en general y, más específicamente, sobre cómo la utilizan los gobiernos (Arnesen et al., 2024^[153]). Esto puede dar lugar a ideas erróneas y temores sobre sus capacidades y sobre si los gobiernos las están utilizando de manera confiable. Esto podría potencialmente resultar en líneas de rumores sobre el uso de IA por parte del gobierno en la formulación de políticas o en la prestación de servicios como préstamos, acceso a la justicia o prestaciones sociales. Esto puede ocurrir porque la complejidad de la IA a menudo se malinterpreta, lo que da lugar a supuestos inexactos sobre cómo se toman las decisiones, la equidad de los resultados y la capacidad de exigir responsabilidades a estos sistemas, o debido a la escasa transparencia sobre cómo

las autoridades públicas están utilizando la IA. Esta transparencia limitada o esta brecha en la comprensión hace que el público sea más susceptible al escepticismo sobre las acciones e intenciones de los gobiernos, lo que resulta en resistencia a las soluciones impulsadas por la IA en los servicios públicos (Valle-Cruz, Garcia-Contreras and Gil-Garcia, 2023^[99]).

Las decisiones sobre IA pueden **percibirse como excesivamente rígidas o irresponsables**, especialmente en ausencia de una supervisión humana o vías de reparación. Las investigaciones sugieren que cuando la ciudadanía se siente desempoderada en las interacciones con los sistemas automatizados, como cuando no pueden impugnar decisiones incorrectas basadas en IA, experimentan mayores costos psicológicos y de cumplimiento normativo (Alon-Barkat and Busuioc, 2024^[74]).

Por el contrario, unos niveles más altos de empoderamiento público y transparencia sobre el uso, el conocimiento y la comprensión de la IA se asocian a una mayor confianza en la capacidad del gobierno para utilizar la IA de forma responsable (Lahusen, Maggetti and Slavkovik, 2024^[154]; KPMG, 2025^[155]; Alessandro et al., 2021^[156]). Una ciudadanía bien informada es más propensa a reconocer las salvaguardias y consideraciones éticas implementadas por el gobierno, fomentando la confianza en su competencia e integridad en el despliegue de las tecnologías de IA.

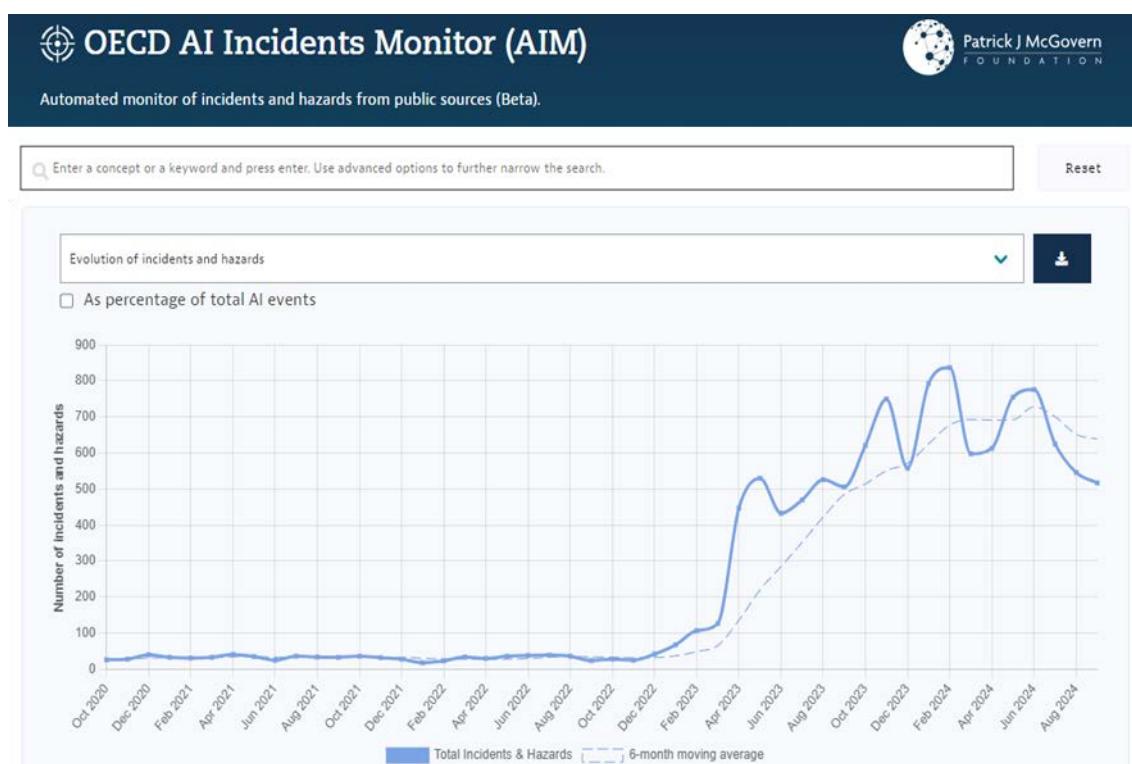
El uso indebido de la IA y las controversias pueden socavar la confianza y contribuir a la resistencia pública

El uso de la IA en el gobierno ha dado lugar a varias controversias y casos de daños en la vida real. Esto pone de relieve los altos costos de reputación del uso indebido de la IA, un problema que hace años todavía residía en el discurso público. Un fallo notorio de un sistema de IA puede erosionar la confianza en una gama más amplia de aplicaciones gubernamentales (Longoni, Cian and Kyung, 2022^[157]). La reacción negativa de la opinión pública incluso ha provocado que la gente se retire del intercambio de datos o ha obstaculizado el uso de las herramientas existentes (Ada Lovelace Institute, 2025^[158]).

La reacción pública es especialmente probable cuando los errores de IA generan daños visibles, como una denegación injusta de beneficios o un trato injusto basado en predicciones sesgadas.

Las fallas notorias pueden reforzar el escepticismo de la opinión pública y suscitar preocupación porque las decisiones impulsadas por la IA pueden socavar la justicia procesal y la rendición de cuentas democrática (Alon-Barkat and Busuioc, 2022^[119]). Los datos del Monitor de Incidentes de IA (AIM por sus siglas en inglés) de la OCDE muestran el crecimiento de los incidentes y peligros de la IA notificados por fuentes de medios de comunicación acreditadas en los últimos años (Gráfico 1.5)¹⁷. Hasta abril de 2025, 3.816 de los 14.981 incidentes enumerados (25%) estaban relacionados con "gobierno, seguridad y defensa", lo que demuestra que los gobiernos deben garantizar que mitigan los riesgos para garantizar la confianza de la ciudadanía.

Gráfico 1.5. Los incidentes de IA han tendido en general al alza desde finales de 2022



Nota: Puede consultarse un resumen de la metodología en <https://oecd.ai/incidents-methodology>.

Fuente: Monitor de Incidentes de IA de la OCDE (AIM) – <https://oecd.ai/incidents>.

La ciudadanía también puede mostrar aversión algorítmica, a menudo resistiéndose a la toma de decisiones algorítmica debido a una preferencia por la agencia y el control personales, incluso cuando los sistemas de IA superan a los humanos. Esta reticencia se ve reforzada por una mayor tolerancia a los errores humanos en comparación con los errores algorítmicos; las personas tienden a perder la confianza en la IA después de observar un solo fallo, mientras que son más tolerantes con los errores similares cometidos por los humanos. La aversión algorítmica puede ser más común en algunas funciones (por ejemplo, seguridad pública) que en otras (por ejemplo, administración general), lo que puede contribuir a diferentes desafíos y diferentes niveles de madurez entre las funciones gubernamentales (véase el Capítulo 3) (Zehnle, Hildebrand and Valenzuela, 2025^[159]). Ofrecer a los usuarios información sobre el razonamiento que subyace a la recomendación de una IA u ofrecer un control limitado para modificar su resultado puede mejorar significativamente la aceptación, aumentando la probabilidad de que los usuarios adopten consejos basados en IA (Sunstein, 2023^[73]). Además de contribuir a los riesgos de resistencia pública, la aversión algorítmica de los servidores públicos puede obstaculizar la capacidad de los gobiernos para aprovechar los beneficios de la IA, como se comenta en el Capítulo 3.

Riesgos de la inacción

Los riesgos más debatidos de la IA tienen que ver con las implicaciones de su despliegue y adopción. Sin embargo, un riesgo menos discutido implica demoras en el aprovechamiento de la IA para generar beneficios positivos en el mundo real, incluso en el gobierno y los servicios públicos. El estudio de la Universidad de Stanford, *cien años sobre inteligencia artificial (AI100)* (2014^[23]), señaló que "los numerosos avances en IA pueden reducir los costos, introducir nuevas eficiencias y aumentar la calidad de vida [...] Sin embargo, los métodos no se han utilizado ampliamente. La lenta traducción de estas

tecnologías al mundo se traduce en muertes y costos innecesarios. Existe una necesidad urgente de entender mejor cómo podemos traducir más rápidamente las valiosas competencias y avances existentes en IA en la práctica real". Si bien este estudio tiene más de una década, la conclusión sigue siendo la misma, especialmente en el gobierno. Más allá de las oportunidades perdidas, el riesgo de la inacción en materia de IA sirve para ampliar la brecha entre el sector público y la capacidad del sector privado (Pahlka, 2024^[160]). Esto no solo significa que los gobiernos podrían quedarse atrás en su capacidad para utilizar la IA, sino también en su capacidad para regular la tecnología. Los expertos en IA sugieren que uno de los riesgos más críticos de la IA es la incapacidad de los mecanismos de gobernanza y las instituciones para mantenerse al día con la rápida evolución de la IA (OECD, 2024^[14]).

Algunas investigaciones ponen de relieve que la exageración negativa y el miedo en torno a la IA pueden contribuir a este riesgo (Laplante et al., 2020^[161]). Estas investigaciones señalan que otras limitaciones pueden incluir la falta de datos adecuados, la confusión en torno a los problemas y complejidades de la privacidad y el tratamiento de sistemas informáticos obsoletos. Los expertos también han atribuido la cuestión de la "parálisis del análisis"—el miedo a equivocarse con la IA— como una posibilidad que podría paralizar la implementación de esfuerzos incluso de bajo riesgo, renunciando a beneficios significativos (OECD.AI, 2023^[162]). Esto ha demostrado ser cierto en los gobiernos. La última iniciativa transversal de la OCDE sobre la IA en el gobierno (2024^[13]) descubrió que estos últimos necesitan promover y habilitar mejor los aspectos positivos del uso de la IA, en lugar de centrarse de manera desproporcionada en la prevención de los negativos. Esta atención a los riesgos podría disuadir del despliegue de usos de la IA de alto beneficio y bajo riesgo para mejorar las políticas y los servicios públicos.

Hacer realidad un futuro positivo para la IA en el gobierno

Como se ha señalado anteriormente, los gobiernos están tratando de utilizar la IA para aumentar la productividad, con operaciones internas más eficientes y políticas y servicios más efectivos; capacidad de respuesta, a través de enfoques adaptados que satisfagan las necesidades cambiantes de la ciudadanía y las empresas; y rendición de cuentas, mejorando su capacidad de supervisión. Sin embargo, existe un enorme potencial sin explotar para que los gobiernos utilicen las tecnologías de IA de hoy y se preparen para desbloquear las oportunidades que presenta la IA del mañana. Aun así, está empezando a surgir una visión de futuro en la que los gobiernos desarrollen y adopten con éxito una IA confiable.

Consideramos que la IA no es una oportunidad para automatizar el sector público, sino para reimaginarlo. Acogemos con satisfacción una visión a largo plazo para la transformación de los servicios públicos en la que la IA siga, en lugar de liderar, una visión que se base en la legitimidad pública y profesional. Los líderes del sector público deberían ver la implementación de la IA como una oportunidad para reimaginar el estado, en lugar de centrarse únicamente en las ganancias inmediatas de eficiencia o en automatizar el statu quo. La IA debe considerarse un catalizador para el rediseño de servicios fundamentales, que sitúe a la ciudadanía en el centro de la prestación de servicios públicos. – Instituto Ada Lovelace (2025^[158])

Esta visión se refleja en el análisis del Capítulo 5 sobre la IA en las funciones centrales de la administración. Al integrar la IA en sus itinerarios digitales, los gobiernos pueden transformar, en lugar de limitarse a optimizar, la forma en que cumplen sus misiones, generar valor público y promover el bienestar social.

Las capacidades y los usos futuros de la IA pueden presentar beneficios y cambios que actualmente son imposibles o incluso inconcebibles. Esto también es cierto en el caso de sus riesgos potenciales. El contenido del Capítulo 5 representa lo que se sabe actualmente sobre IA en el gobierno y, por extensión, lo que se puede imaginar sobre el futuro. Los gobiernos y la OCDE deben mantenerse alerta a la hora de evaluar cómo la evolución de las tecnologías y aplicaciones de la IA puede afectar a las instituciones públicas, los funcionarios y la sociedad en general, garantizando una evaluación continua y la adaptación al servicio del bien público.

A través del análisis y la síntesis de las funciones gubernamentales en el Capítulo 5 y otras investigaciones, y para profundizar en el debate sobre la IA en el gobierno, la OCDE ha llevado a cabo una investigación transversal sobre las tendencias actuales del uso de la IA en las funciones gubernamentales y ha identificado las primeras lecciones de estos casos de uso. Esta investigación y sus hallazgos se discuten en el siguiente capítulo.

Bibliografía

- Acemoglu, D. (2024), "The Simple Macroeconomics of AI", *NBER, Working Paper 32487*, [152]
<http://www.nber.org/papers/w32487>.
- Acemoglu, D. et al. (2022), "Artificial Intelligence and Jobs: Evidence from Online Vacancies", [150]
Journal of Labor Economics, Vol. 40/S1, pp. S293-S340, <https://doi.org/10.1086/718327>.
- Ada Lovelace Institute (2025), *Learn fast and build things: Lessons from six years of studying AI in the public sector*, Ada Lovelace Institute, <https://www.adalovelaceinstitute.org/policy-briefing/public-sector-ai/>. [158]
- Alessandro, M. et al. (2021), "Transparency and Trust in Government. Evidence from a Survey Experiment", *World Development*, Vol. 138, p. 105223, [156]
<https://doi.org/10.1016/j.worlddev.2020.105223>.
- Alon-Barkat, S. and M. Busuioc (2024), "Public administration meets artificial intelligence: Towards a meaningful behavioral research agenda on algorithmic decision-making in government", *Journal of Behavioral Public Administration*, Vol. 7, [74]
<https://doi.org/10.30636/jbpa.71.261>.
- Alon-Barkat, S. and M. Busuioc (2022), "Human–AI Interactions in Public Sector Decision Making: “Automation Bias” and “Selective Adherence” to Algorithmic Advice", *Journal of Public Administration Research and Theory*, Vol. 33/1, pp. 153-169, [119]
<https://doi.org/10.1093/jopart/muac007>.
- Aonghusa, P. and S. Michie (2020), "Artificial intelligence and behavioral science through the looking glass: Challenges for real-world application.", *Annals of Behavioural Medicine*, [71]
pp. 942-947, <https://doi.org/10.1093/abm/kaaa095>.
- APA (2023), *Worries about artificial intelligence, surveillance at work may be connected to poor mental health*, <https://www.apa.org/news/press/releases/2023/09/artificial-intelligence-poor-mental-health>. [130]
- Arnesen, S. et al. (2024), "Knowledge and support for AI in the public sector: a deliberative poll experiment", *AI & SOCIETY*, <https://doi.org/10.1007/s00146-024-02104-w>. [153]
- Austin, T. et al. (2024), *A snapshot of how public sector leaders feel about generative AI*, [35]
<https://www2.deloitte.com/us/en/insights/industry/public-sector/ai-adoption-in-public-sector.html>.
- BCG (2024), *Where's the Value in AI?*, <https://www.bcg.com/publications/2024/wheres-value-in-ai>. [32]

- Beltran, M., M. Ruiz Mondragon and S. Han (2024), "Comparative Analysis of Generative AI Risks in the Public Sector", *Proceedings of the 25th Annual International Conference on Digital Government Research*, pp. 610-617, <https://doi.org/10.1145/3657054.3657125>. [124]
- Bengio, Y. et al. (2025), *International AI Safety Report*, DSIT 2025/001, 2025, <https://www.gov.uk/government/publications/international-ai-safety-report-2025>. [28]
- Berglind, N., A. Fadia and T. Isherwood (2022), *The potential value of AI—and how governments could look to capture it*, <https://www.mckinsey.com/industries/public-sector/our-insights/the-potential-value-of-ai-and-how-governments-could-look-to-capture-it> (accessed on July 2024). [39]
- Berryhill, J. et al. (2019), "Hello, World: Artificial intelligence and its use in the public sector", *OECD Working Papers on Public Governance*, No. 36, OECD Publishing, Paris, <https://doi.org/10.1787/726fd39d-en>. [8]
- Berryhill, J. et al. (2019), "Hello, World: Artificial intelligence and its use in the public sector", *OECD Working Papers on Public Governance*, No. 36, OECD Publishing, Paris, <https://doi.org/10.1787/726fd39d-en>. [15]
- Brizuela, A. et al. (2025), *Analysis of the generative AI landscape in the European public sector*, European Commission, <https://op.europa.eu/s/z4XY>. [24]
- Brougham, D. and J. Haar (2017), "Smart Technology, Artificial Intelligence, Robotics, and Algorithms (STARA): Employees' perceptions of our future workplace", *Journal of Management & Organization*, Vol. 24/2, pp. 239-257, <https://doi.org/10.1017/jmo.2016.55>. [45]
- Brown, H. et al. (2022), "What Does it Mean for a Language Model to Preserve Privacy?", 2022 ACM Conference on Fairness, Accountability, and Transparency, pp. 2280-2292, <https://doi.org/10.1145/3531146.3534642>. [135]
- Brundage, M. et al. (2018), *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*, <https://arxiv.org/abs/1802.07228>. [138]
- Brynjolfsson, E., L. Danielle and L. Raymond (2023), *Generative AI at Work*, National Bureau of Economic Research, <https://doi.org/10.3386/w31161>. [69]
- Byler, D. (2021), *In the Camps: China's High-Tech Penal Colony*, Columbia Global Reports, <https://www.jstor.org/stable/j.ctv2dzzqqm>. [115]
- Calvino, F. and L. Fontanelli (2023), "A portrait of AI adopters across countries: Firm characteristics, assets' complementarities and productivity", *OECD Science, Technology and Industry Working Papers*, No. 2023/02, OECD Publishing, Paris, <https://doi.org/10.1787/0fb79bb9-en>. [29]
- Chignard, S. (2013), *A brief history of Open Data*, <https://www.paristechreview.com/2013/03/29/brief-history-open-data>. [88]
- Clarke, S. and J. Whittlestone (2022), *A Survey of the Potential Long-term Impacts of AI - How AI Could Lead to Long-term Changes in Science, Cooperation, Power, Epistemics and Values*, <https://dl.acm.org/doi/abs/10.1145/3514094.3534131>. [114]
- Cognitus, A. (2024), *9 Agentic AI Examples: Real-World Use Cases and Applications*, <https://integral.ai/blog/agentic-ai-examples>. [20]

- Corvalán, J. and E. Le Fevre Cervini (2020), *Prometea experience. Using AI to optimize public institutions*, <https://ceridap.eu/prometea-experience-using-ai-to-optimize-public-institutions>. [48]
- DataHeroes (2023), *Noise in Machine Learning*, <https://dataheroes.ai/glossary/noise-in-machine-learning>. [165]
- Davies, H., S. Nutley and I. Walter (2008), “Why ‘knowledge transfer’ is misconceived for applied social research”, *Journal of Health Services Research & Policy*, Vol. 13/3, pp. 188-190, <https://doi.org/10.1258/jhsrp.2008.008055>. [100]
- Dell'Acqua, F. et al. (2025), *The Cybernetic Teammate: A Field Experiment on Generative AI Reshaping Teamwork and Expertise*, Elsevier BV, <https://doi.org/10.2139/ssrn.5188231>. [50]
- Dell'Acqua, F. et al. (2023), “Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality”, *SSRN Electronic Journal*, <https://doi.org/10.2139/ssrn.4573321>. [54]
- Desouza, K. and B. Jacob (2014), “Big Data in the Public Sector: Lessons for Practitioners and Scholars”, *Administration & Society*, Vol. 49/7, pp. 1043-1064, <https://doi.org/10.1177/0095399714555751>. [66]
- Dieterle, E., C. Dede and M. Walker (2022), “The cyclical ethical effects of using artificial intelligence in education”, *AI & SOCIETY*, Vol. 39/2, pp. 633-643, <https://doi.org/10.1007/s00146-022-01497-w>. [147]
- Du, M. (2023), “Machine vs. human, who makes a better judgment on innovation? Take GPT-4 for example”, *Frontiers in Artificial Intelligence*, Vol. 6, <https://doi.org/10.3389/frai.2023.1206516>. [70]
- EC (2025), *Study exploring the context, challenges, opportunities, and trends in algorithmic management*, European Commission, https://employment-social-affairs.ec.europa.eu/study-exploring-context-challenges-opportunities-and-trends-algorithmic-management_en. [129]
- EC (2024), *What factors influence perceived artificial intelligence adoption by public managers*, <https://publications.jrc.ec.europa.eu/repository/handle/JRC138684>. [43]
- Eshelby, L. (2025), *Addressing insider threats in the public sector*, <https://www.openaccessgovernment.org/addressing-insider-threats-in-the-public-sector/187801/>. [140]
- EU (2024), *Regulation (EU) 2024/1689 laying down harmonised rules on artificial*, European Union, <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>. [97]
- Fazlioglu, M. (2024), *Consumer Perspectives of Privacy and Artificial Intelligence*, <https://iapp.org/resources/article/consumer-perspectives-of-privacy-and-ai/>. [110]
- Feldstein, S. (2022), *AI & Big Data Global Surveillance Index (2022 updated)*, <https://doi.org/10.17632/gjhf5y4xp.4>. [104]
- Filippucci, F., P. Gal and M. Schief (2024), “Miracle or Myth? Assessing the macroeconomic productivity gains from Artificial Intelligence”, *OECD Artificial Intelligence Papers*, No. 29, OECD Publishing, Paris, <https://doi.org/10.1787/b524a072-en>. [26]

- Fitkov-Norris, E. and N. Kocheva (2025), "Leveraging AI for strategic foresight: Unveiling future horizons", in *Improving and Enhancing Scenario Planning*, Edward Elgar Publishing, [75]
<https://doi.org/10.4337/9781035310586.00023>.
- Flavián, C. and L. Casaló (2021), "Artificial intelligence in services: current trends, benefits and challenges", *The Service Industries Journal*, Vol. 41/13-14, pp. 853–859, [56]
<https://doi.org/10.1080/02642069.2021.1989177>.
- Gartner (2024), *Gartner 2024 Hype Cycle for Emerging Technologies Highlights Developer Productivity, Total Experience, AI and Security*, <https://www.gartner.com/en/newsroom/press-releases/2024-08-21-gartner-2024-hype-cycle-for-emerging-technologies-highlights-developer-productivity-total-experience-ai-and-security>. [25]
- Gerlich, M. (2025), "AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking", *Societies*, Vol. 15/1, p. 6, <https://doi.org/10.3390/soc15010006>. [125]
- Giest, S. (2017), "Big data for policymaking: fad or fasttrack?", *Policy Sciences*, Vol. 50/3, pp. 367-382, [58]
<https://doi.org/10.1007/s11077-017-9293-1>.
- Gmyrek, P., J. Berg and D. Bescond (2023), *Generative AI and Jobs: A global analysis of potential effects on job quantity and quality*, ILO, <https://doi.org/10.54394/FHEM8239>. [131]
- Goldman, S. (2023), *OpenAI has grand 'plans' for AGI. Here's another way to read its manifesto*, [120]
<https://venturebeat.com/ai/openai-has-grand-plans-foragi-heres-another-way-to-read-its-manifesto-the-ai-beat/>.
- Gopireddy, R. (2024), *Securing AI Systems: Protecting Against Adversarial Attacks and Data Poisoning*, <https://jsaer.com/download/vol-11-iss-5-2024/JSAER2024-11-5-276-281.pdf>. [139]
- Government of Korea (2024), *A New Chapter in the Age of AI: Basic Act on AI Passed at the National Assembly's Plenary Session*, [98]
<https://www.msit.go.kr/eng/bbs/view.do?bbsSeqNo=42&mld=4&mPid=2&nttSeqNo=1071>.
- Green, B. (2022), "The flaws of policies requiring human oversight of government algorithms", [68]
Computer Law & Security Review, Vol. 45, p. 105681,
<https://doi.org/10.1016/j.clsr.2022.105681>.
- Grzegorzek, J. (2024), *Digital Taylorism: The Use of Data to Monitor Employees*, [163]
<https://medium.com/%40JerryGrzegorzek/digital-taylorism-the-use-of-data-to-monitor-employees-582b331d970a>.
- Gupta, T. and S. Roy (2024), "Applications of Artificial Intelligence in Disaster Management", [77]
Proceedings of the 2024 10th International Conference on Computing and Artificial Intelligence, pp. 313-318, <https://doi.org/10.1145/3669754.3669802>.
- Hampole, M. et al. (2025), *Artificial Intelligence and the Labor Market*, National Bureau of Economic Research, Cambridge, MA, <https://doi.org/10.3386/w33509>. [30]
- Höchtl, J., P. Parycek and R. Schöllhammer (2015), "Big data in the policy cycle: Policy decision making in the digital era", *Journal of Organizational Computing and Electronic Commerce*, Vol. 26/1-2, pp. 147-169, <https://doi.org/10.1080/10919392.2015.1125187>. [64]

- Ho, D. (2023), *Opportunities and Risks of Artificial Intelligence in the Public Sector*, [91]
<https://law.stanford.edu/2023/05/25/opportunities-and-risks-of-artificial-intelligence-in-the-public-sector/>.
- Horowitz, M. (2023), *Bending the Automation Bias Curve: A Study of Human and AI-based Decision Making in National Security Contexts*, [118]
<https://arxiv.org/abs/2306.16507>.
- Horvitz, E. (2014), *Reflections and Framing: One-Hundred Year Study on Artificial Intelligence: Reflections and Framing*, [23]
<https://ai100.stanford.edu/reflections-and-framing>.
- Huang, M. and R. Rust (2021), “Engaged to a Robot? The Role of AI in Service”, *Journal of Service Research*, Vol. 24/1, pp. 30–41, [55]
<https://doi.org/10.1177/1094670520902266>.
- ICO (2023), *Joint statement on data scraping and the protection of privacy*, [134]
<https://ico.org.uk/about-the-ico/media-centre/news-and-blogs/2023/08/joint-statement-on-data-scraping-and-data-protection/>.
- ITU (2023), *Measuring digital development: Facts and Figures 2023*, [142]
<https://www.itu.int/en/ITU-D/Statistics/Pages/facts/default.aspx> (accessed on July 2024).
- Jarrahi, M. et al. (2023), “Artificial intelligence and knowledge management: A partnership between human and AI”, *Business Horizons*, Vol. 66/1, pp. 87-99, [78]
<https://doi.org/10.1016/j.bushor.2022.03.002>.
- Jeevanandam, N. (2024), *AI in agriculture in 2025: Transforming Indian farms for a sustainable future*, [89]
<https://indiaai.gov.in/article/ai-in-agriculture-in-2025-transforming-indian-farms-for-a-sustainable-future>.
- Jones, C. (2022), “The Past and Future of Economic Growth: A Semi-Endogenous Perspective”, [49]
Annual Review of Economics, Vol. 14/1, pp. 125-152, <https://doi.org/10.1146/annurev-economics-080521-012458>.
- Klingbeil, A., C. Grützner and P. Schrec (2024), *Trust and reliance on AI — An experimental study on the extent and costs of overreliance on AI*, [123]
<https://doi.org/10.1016/j.chb.2024.108352> (accessed on September 2024).
- Kolkman, D. (2020), “The usefulness of algorithmic models in policy making”, *Government Information Quarterly*, Vol. 37/3, p. 101488, [63]
<https://doi.org/10.1016/j.giq.2020.101488>.
- Kopponen, A. et al. (2024), “Personalised public services powered by AI: the citizen digital twin approach”, in *Research Handbook on Public Management and Artificial Intelligence*, Edward Elgar Publishing, [59]
<https://doi.org/10.4337/9781802207347.00020>.
- KPMG (2025), *Trust in artificial intelligence: global insights 2025*, [155]
<https://kpmg.com/au/en/home/insights/2025/04/trust-in-ai-global-insights-2025.html>.
- Lahusen, C., M. Maggetti and M. Slavkovik (2024), “Trust, trustworthiness and AI governance”, [154]
Scientific Reports, Vol. 14/1, <https://doi.org/10.1038/s41598-024-71761-0>.
- Laplante, P. et al. (2020), “Artificial Intelligence and Critical Systems: From Hype to Reality”, [161]
Computer, Vol. 53/11, pp. 45-52, <https://doi.org/10.1109/mc.2020.3006177>.

- Larsson, K. (2021), "Digitization or equality: When government automation covers some, but not all citizens", *Government Information Quarterly*, Vol. 38/1, p. 101547, [144] <https://doi.org/10.1016/j.giq.2020.101547>.
- Lima, G. et al. (2022), "The Conflict Between Explainable and Accountable Decision-Making Algorithms", *2022 ACM Conference on Fairness, Accountability, and Transparency*, pp. 2103-2113, [117] <https://doi.org/10.1145/3531146.3534628>.
- Longoni, C., L. Cian and E. Kyung (2022), "Algorithmic Transference: People Overgeneralize Failures of AI in the Government", *Journal of Marketing Research*, Vol. 60/1, pp. 170-188, [157] <https://doi.org/10.1177/00222437221110139>.
- Lorenz, P., K. Perset and J. Berryhill (2023), "Initial policy considerations for generative artificial intelligence", *OECD Artificial Intelligence Papers*, No. 1, OECD Publishing, Paris, [19] <https://doi.org/10.1787/fae2d1e6-en>.
- Ludwig, J. and S. Mullainathan (2022), "Algorithmic Behavioral Science: Machine Learning as a Tool for Scientific Discovery", *SSRN Electronic Journal*, [72] <https://doi.org/10.2139/ssrn.4164272>.
- Manning, B., K. Zhu and J. Horton (2024), *Automated Social Science: Language Models as Scientist and Subjects*, [34] <https://arxiv.org/abs/2404.11794>.
- Mariani, J., P. Kishnani and A. Alibage (2025), *Government's less trodden path to scaling generative AI*, [141] <https://www2.deloitte.com/us/en/insights/industry/public-sector/government-faces-challenges-with-generative-ai-adoption.html>.
- Mellouli, S., M. Janssen and A. Ojo (2024), "Introduction to the Issue on Artificial Intelligence in the Public Sector: Risks and Benefits of AI for Governments", *Digital Government: Research and Practice*, Vol. 5/1, pp. 1-6, [37] <https://doi.org/10.1145/3636550>.
- Meng, J. (2024), *AI emerges as the frontier in behavioral science*, [126] <https://doi.org/10.1073/pnas.2401336121>.
- Mergel, I. et al. (2023), "Implementing AI in the public sector", *Public Management Review*, [36] pp. 1-14, <https://doi.org/10.1080/14719037.2023.2231950>.
- Milanez, A. (2023), "The impact of AI on the workplace: Evidence from OECD case studies of AI implementation", *OECD Social, Employment and Migration Working Papers*, No. 289, OECD Publishing, Paris, [151] <https://doi.org/10.1787/2247ce58-en>.
- Milanez, A., A. Lemmens and C. Ruggiu (2025), "Algorithmic management in the workplace: New evidence from an OECD employer survey", *OECD Artificial Intelligence Papers*, No. 31, OECD Publishing, Paris, [127] <https://doi.org/10.1787/287c13c4-en>.
- Mills, S., S. Costa and C. Sunstein (2023), "AI, Behavioural Science, and Consumer Welfare", *J Consum Policy*, Vol. 46, pp. 387–400, [60] <https://doi.org/10.1007/s10603-023-09547-6>.
- Nikiforova, A. et al. (2023), "Towards High-Value Datasets Determination for Data-Driven Development: A Systematic Literature Review", in *Lecture Notes in Computer Science, Electronic Government*, Springer Nature Switzerland, Cham, [105] https://doi.org/10.1007/978-3-031-41138-0_14.

- NIST (2025), *Technical Blog: Strengthening AI Agent Hijacking Evaluations*, [22]
<https://www.nist.gov/news-events/news/2025/01/technical-blog-strengthening-ai-agent-hijacking-evaluations>.
- Noy, S. and W. Zhang (2023), “Experimental evidence on the productivity effects of generative artificial intelligence”, *Science*, Vol. 381/6654, pp. 187-192, [51]
<https://doi.org/10.1126/science.adh2586>.
- OECD (2025), *How Innovation Ecosystems Foster Citizen Participation Using Emerging Technologies in Portugal, Spain and the Netherlands*, OECD Public Governance Reviews, [92]
OECD Publishing, Paris, <https://doi.org/10.1787/2cb37a30-en>.
- OECD (2025), “Sharing trustworthy AI models with privacy-enhancing technologies”, *OECD Artificial Intelligence Papers*, No. 38, OECD Publishing, Paris, [81]
<https://doi.org/10.1787/a266160b-en>.
- OECD (2025), “Towards a common reporting framework for AI incidents”, *OECD Artificial Intelligence Papers*, No. 34, OECD Publishing, Paris, [164]
<https://doi.org/10.1787/f326d4ac-en>.
- OECD (2024), “2023 OECD Digital Government Index: Results and key findings”, *OECD Public Governance Policy Papers*, No. 44, OECD Publishing, Paris, [5]
<https://doi.org/10.1787/1a89ed5e-en>.
- OECD (2024), “AI, data governance and privacy: Synergies and areas of international co-operation”, *OECD Artificial Intelligence Papers*, No. 22, OECD Publishing, Paris, [132]
<https://doi.org/10.1787/2476b1a4-en>.
- OECD (2024), “Assessing potential future artificial intelligence risks, benefits and policy imperatives”, *OECD Artificial Intelligence Papers*, No. 27, OECD Publishing, Paris, [14]
<https://doi.org/10.1787/3f4e3dfb-en>.
- OECD (2024), “Assessing potential future artificial intelligence risks, benefits and policy imperatives”, *OECD Artificial Intelligence Papers*, No. 27, OECD Publishing, Paris, [17]
<https://doi.org/10.1787/3f4e3dfb-en>.
- OECD (2024), “Defining AI incidents and related terms”, *OECD Artificial Intelligence Papers*, [166]
No. 16, OECD Publishing, Paris, <https://doi.org/10.1787/d1a8d965-en>.
- OECD (2024), “Explanatory memorandum on the updated OECD definition of an AI system”, [7]
OECD Artificial Intelligence Papers, No. 8, OECD Publishing, Paris,
<https://doi.org/10.1787/623da898-en>.
- OECD (2024), *Facts not Fakes: Tackling Disinformation, Strengthening Information Integrity*, [95]
OECD Publishing, Paris, <https://doi.org/10.1787/d909ff7a-en>.
- OECD (2024), *Global Trends in Government Innovation 2024: Fostering Human-Centred Public Services*, OECD Public Governance Reviews, OECD Publishing, Paris, [108]
<https://doi.org/10.1787/c1bc19c3-en>.
- OECD (2024), “Governing with Artificial Intelligence: Are governments ready?”, *OECD Artificial Intelligence Papers*, No. 20, OECD Publishing, Paris, [13]
<https://doi.org/10.1787/26324bc2-en>.
- OECD (2024), *Modernising Access to Social Protection: Strategies, Technologies and Data Advances in OECD Countries*, OECD Publishing, Paris, [61]
<https://doi.org/10.1787/af31746d-en>.

- OECD (2024), *OECD Artificial Intelligence Review of Germany*, OECD Publishing, Paris, [11]
<https://doi.org/10.1787/609808d6-en>.
- OECD (2024), *OECD Digital Economy Outlook 2024 (Volume 1): Embracing the Technology Frontier*, OECD Publishing, Paris, [94]
<https://doi.org/10.1787/a1689dc5-en>.
- OECD (2024), *OECD Digital Economy Outlook 2024 (Volume 1): Embracing the Technology Frontier*, OECD Publishing, Paris, [16]
<https://doi.org/10.1787/a1689dc5-en>.
- OECD (2024), *OECD Survey on Drivers of Trust in Public Institutions – 2024 Results: Building Trust in a Complex Policy Environment*, OECD Publishing, Paris, [4]
<https://doi.org/10.1787/9a20554b-en>.
- OECD (2024), *Recommendation of the Council on Artificial Intelligence*, [6]
<https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.
- OECD (2023), *Artificial Intelligence in Science: Challenges, Opportunities and the Future of Research*, OECD Publishing, Paris, [33]
<https://doi.org/10.1787/a8d820bd-en>.
- OECD (2023), “Emerging privacy-enhancing technologies: Current regulatory and policy approaches”, *OECD Digital Economy Papers*, No. 351, OECD Publishing, Paris, [82]
<https://doi.org/10.1787/bf121be4-en>.
- OECD (2023), *Global Trends in Government Innovation 2023*, OECD Public Governance Reviews, OECD Publishing, Paris, [40]
<https://doi.org/10.1787/0655b570-en>.
- OECD (2023), *OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market*, OECD Publishing, Paris, [27]
<https://doi.org/10.1787/08785bba-en>.
- OECD (2023), “The state of implementation of the OECD AI Principles four years on”, *OECD Artificial Intelligence Papers*, No. 3, OECD Publishing, Paris, [90]
<https://doi.org/10.1787/835641c9-en>.
- OECD (2022), *Emerging future AI risks*, OECD, [112]
<https://wp.oecd.ai/app/uploads/2023/03/OECD-Foresight-workshop-notes-1.pdf>.
- OECD (2022), “Measuring the environmental impacts of artificial intelligence compute and applications: The AI footprint”, *OECD Digital Economy Papers*, No. 341, OECD Publishing, Paris, [96]
<https://doi.org/10.1787/7babf571-en>.
- OECD (2022), “OECD Framework for the Classification of AI systems”, *OECD Digital Economy Papers*, No. 323, OECD Publishing, Paris, [42]
<https://doi.org/10.1787/cb6d9eca-en>.
- OECD (2021), *Data-Driven, Information-Enabled Regulatory Delivery*, OECD Publishing, Paris, [85]
<https://doi.org/10.1787/8f99ec8c-en>.
- OECD (2020), “Dealing with digital security risk during the Coronavirus (COVID-19) crisis”, *OECD Policy Responses to Coronavirus (COVID-19)*, OECD Publishing, Paris, [103]
<https://doi.org/10.1787/c9d3fe8e-en>.
- OECD (2020), *Embracing Innovation in Government - Global Trends 2020: Innovative responses to the COVID-19 crisis*, OECD Publishing, [2]
<https://trends.oecd-opsi.org/trend-reports/innovative-covid-19-solutions/>.

- OECD (2020), "The OECD Digital Government Policy Framework: Six dimensions of a Digital Government", *OECD Public Governance Policy Papers*, No. 02, OECD Publishing, Paris, [\[3\]](https://doi.org/10.1787/f64fed2a-en)
- OECD (2019), *The Path to Becoming a Data-Driven Public Sector*, OECD Digital Government Studies, OECD Publishing, Paris, [\[67\]](https://doi.org/10.1787/059814a7-en)
- OECD (2018), *OECD Regulatory Enforcement and Inspections Toolkit*, OECD Publishing, Paris, [\[84\]](https://doi.org/10.1787/9789264303959-en)
- OECD (2018), *Open Government Data Report: Enhancing Policy Maturity for Sustainable Impact*, OECD Digital Government Studies, OECD Publishing, Paris, [\[86\]](https://doi.org/10.1787/9789264305847-en)
- OECD (2017), *Recommendation of the Council on Health Data Governance*, [\[106\]](https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0433)
- OECD (2014), *Recommendation of the Council on Digital Government Strategies*, OECD Publishing, [\[1\]](https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0406)
- OECD (1980), *Recommendation of the Council concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data*, [\[107\]](https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0188)
- OECD (forthcoming), *Mapping Relevant Data Collection Mechanisms for AI Training*, OECD Publishing. [\[133\]](#)
- OECD.AI (2023), *OECD Expert Group on AI Futures – Meeting 2 (18th & 20th September 2023)*, [\[18\]](https://wp.oecd.ai/app/uploads/2024/01/Expert-Group-on-AI-Futures-Meeting-2-Summary.pdf)
- OECD.AI (2023), *What do you see as the most significant potential benefits and risks of AI 10+ years from now?*, [\[162\]](https://oecd.ai/en/network-of-experts/ai-futures/discussions/future-benefits-risks)
- OECD/CAF (2022), *The Strategic and Responsible Use of Artificial Intelligence in the Public Sector of Latin America and the Caribbean*, OECD Public Governance Reviews, OECD Publishing, Paris, [\[10\]](https://doi.org/10.1787/1f334543-en)
- OECD/UNESCO (2024), *G7 Toolkit for Artificial Intelligence in the Public Sector*, OECD Publishing, Paris, [\[12\]](https://doi.org/10.1787/421c1244-en)
- Olson, P. (2023), *Don't Go Down That AI Longtermism Rabbit Hole*, [\[121\]](https://www.bloomberg.com/opinion/articles/2023-05-19/ai-longtermism-alarmists-are-dragging-us-all-down-existential-rabbit-hole)
- Pahlka, J. (2024), *AI meets the cascade of rigidity*, [\[160\]](https://www.niskanencenter.org/ai-meets-the-cascade-of-rigidity/)
- Passi, S. and M. Vorvoreanu (2022), *Overreliance on AI: Literature review*, [\[122\]](https://www.microsoft.com/en-us/research/publication/overreliance-on-ai-literature-review/)
(accessed on September 2024).

- Peixoto, T., O. Canuto and L. Jordan (2024), *AI and the Future of Government: Unexpected Effects and Critical Challenges*, <https://www.policycenter.ma/publications/ai-and-future-government-unexpected-effects-and-critical-challenges>. [149]
- Pencheva, I., M. Esteve and S. Mikhaylov (2018), “Big Data and AI – A transformational shift for government: So, what next for research?”, *Public Policy and Administration*, Vol. 35/1, pp. 24-44, <https://doi.org/10.1177/0952076718780537>. [44]
- Peng, S. et al. (2023), *The Impact of AI on Developer Productivity: Evidence from GitHub Copilot*, <https://arxiv.org/abs/2302.06590>. [52]
- Perry, A. and N. Turner Lee (2019), *AI is coming to schools, and if we're not careful, so will its biases*, <https://www.brookings.edu/articles/ai-is-coming-to-schools-and-if-were-not-careful-so-will-its-biases/>. [146]
- Purdy, M. (2024), *What Is Agentic AI, and How Will It Change Work?*, <https://hbr.org/2024/12/what-is-agentic-ai-and-how-will-it-change-work>. [21]
- Rainie, L. and J. Anderson (2024), *Experts Imagine the Impact of Artificial Intelligence by 2040*, <https://imaginingthefuture.org/wp-content/uploads/2024/02/AI2040-FINAL-White-Paper-2-29.24.pdf>. [109]
- Röttger, P. et al. (2024), “SafetyPrompts: a Systematic Review of Open Datasets for Evaluating and Improving Large Language Model Safety”, *arXiv.org*, <https://doi.org/10.48550/arXiv.2404.05399>. [148]
- Russell, S. (2019), *Human Compatible: Artificial Intelligence and the Problem of Control*, Viking. [116]
- Sahib, T. (2022), “Ethically contentious aspects of artificial intelligence surveillance: a social science perspective”, *AI and Ethics*, Vol. 3/2, pp. 369-379, <https://doi.org/10.1007/s43681-022-00196-y>. [102]
- Salem, A. et al. (2018), *ML-Leaks: Model and Data Independent Membership Inference Attacks and Defenses on Machine Learning Models*, <https://arxiv.org/abs/1806.01246>. [137]
- Santiso, C. (2023), *Public Governance in the Age of Artificial Intelligence*, <https://www.chandlerinstitute.org/governancematters/public-governance-in-the-age-of-artificial-intelligence>. [41]
- Santos, R. et al. (2024), “The use of AI in government and its risks: lessons from the private sector”, *Transforming Government: People, Process and Policy*, <https://doi.org/10.1108/tg-02-2024-0038>. [38]
- Sanzogni, L., G. Guzman and P. Busch (2017), “Artificial intelligence and knowledge management: questioning the tacit dimension”, *Prometheus*, Vol. 35, pp. 37-56, <https://doi.org/10.1080/08109028.2017.1364547>. [80]
- Sapci, A. and H. Sapci (2019), “Innovative Assisted Living Tools, Remote Monitoring Technologies, Artificial Intelligence-Driven Solutions, and Robotic Systems for Aging Societies: Systematic Review”, *JMIR Aging*, Vol. 2/2, p. e15429, <https://doi.org/10.2196/15429>. [47]

- Shane, J. (2019), *You Look Like a Thing and I Love You: How Artificial Intelligence Works and Why It's Making the World a Weirder Place*, [101]
<https://www.hachettebookgroup.com/titles/janelle-shane/you-look-like-a-thing-and-i-love-you/9780316525220/>.
- Singh, V. and J. Chobotaru (2022), “Digital Divide: Barriers to Accessing Online Government Services in Canada”, *Administrative Sciences*, Vol. 12/3, p. 112, [143]
<https://doi.org/10.3390/admsci12030112>.
- Staab, R. et al. (2023), *Beyond Memorization: Violating Privacy Via Inference with Large Language*, [136]
<https://arxiv.org/abs/2310.07298>.
- Sunstein, C. (2023), “The use of algorithms in society”, *The Review of Austrian Economics*, [73]
Vol. 37/4, pp. 399-420, <https://doi.org/10.1007/s11138-023-00625-z>.
- Sun, W., P. Bocchini and B. Davison (2020), “Applications of artificial intelligence for disaster management”, *Natural Hazards*, Vol. 103/3, pp. 2631-2689, [76]
<https://doi.org/10.1007/s11069-020-04124-3>.
- Tegmark, M. (2017), *Life 3.0: Being Human in the Age of Artificial Intelligence*, Penguin, [113]
<https://mitpressbookstore.mit.edu/book/9781101970317>.
- The Economist (2025), *How AI will divide the best from the rest*, [53]
<https://www.economist.com/finance-and-economics/2025/02/13/how-ai-will-divide-the-best-from-the-rest>.
- Tse, T. and S. Karimov (2022), *Decision-making risks slow down the use of artificial intelligence in business*, [93]
<https://blogs.lse.ac.uk/businessreview/2022/05/18/decision-making-risks-slow-down-the-use-of-artificial-intelligence-in-business-1/>.
- Ubaldi, B. et al. (2019), “State of the art in the use of emerging technologies in the public sector”, [9]
OECD Working Papers on Public Governance, No. 31, OECD Publishing, Paris,
<https://doi.org/10.1787/932780bc-en>.
- UC Berkeley (2021), *Positive AI Economic Futures*, World Economic Forum, [128]
<https://www.weforum.org/reports/positive-ai-economic-futures>.
- UN (2024), *Governing AI for Humanity*, [111]
https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf.
- UN (2022), *E-Government Survey 2022: The Future of Digital Government*, United Nations, [57]
<https://desapublications.un.org/sites/default/files/publications/2022-09/Web%20version%20E-Government%202022.pdf>.
- UNESCO (2019), *Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development*, [145]
<https://www.gcedclearinghouse.org/sites/default/files/resources/190175eng.pdf>.
- US GAO (2024), *Fraud Risk Management: 2018-2022 Data Show Federal Government Loses an Estimated \$233 Billion to \$521 Billion Annually to Fraud, Based on Various Risk Environments*, [83]
<https://www.gao.gov/products/gao-24-105833>.

- USGS (2024), *Landsat's Economic Value increased to \$25.6 Billion in 2023*, [87]
<https://www.usgs.gov/news/featured-story/landsats-economic-value-increases-256-billion-2023?>
- Valle-Cruz, D. et al. (2020), "Assessing the public policy-cycle framework in the age of artificial intelligence: From agenda-setting to policy evaluation", *Government Information Quarterly*, Vol. 37/4, p. 101509, <https://doi.org/10.1016/j.giq.2020.101509>. [62]
- Valle-Cruz, D., R. Garcia-Contreras and R. Gil-Garcia (2023), "Exploring the negative impacts of artificial intelligence in government: the dark side of intelligent algorithms and cognitive machines", *International Review of Administrative Sciences*, <https://doi.org/10.1177/002085232311870>. [99]
- Williams, C. (2025), *There will be no immediate productivity boost from AI*, [31]
<https://www.economist.com/the-world-ahead/2024/11/20/there-will-be-no-immediate-productivity-boost-from-ai>.
- Wirjo, A. et al. (2022), *Artificial Intelligence in Economic Policymaking*, Policy Brief No. 52, [65]
https://www.apec.org/docs/default-source/publications/2022/11/artificial-intelligence-in-economic-policymaking/222_psu_artificial-intelligence-in-economic-policymaking.pdf.
- Xu, G., M. Xue and J. Zhao (2023), "The Relationship of Artificial Intelligence Opportunity Perception and Employee Workplace Well-Being: A Moderated Mediation Model", *International Journal of Environmental Research and Public Health*, Vol. 20/3, p. 1974, <https://doi.org/10.3390/ijerph20031974>. [46]
- Zehnle, M., C. Hildebrand and A. Valenzuela (2025), "Not all AI is created equal: A meta-analysis revealing drivers of AI resistance across markets, methods, and time", *International Journal of Research in Marketing*, <https://doi.org/10.1016/j.ijresmar.2025.02.005>. [159]
- Zhang, Z., L. Wang and C. Lee (2023), "Recent Advances in Artificial Intelligence Sensors", [79]
Advanced Sensor Research, Vol. 2/8, p. 2200072, <https://doi.org/10.1002/adsr.202200072>.

Notas

¹ <https://www.oecd.org/en/topics/policy-issues/digital-government/>.

² Solo en Estados Unidos se han catalogado más de 2 000 casos prácticos en organismos civiles del Gobierno federal (<https://github.com/ombegov/2024-Federal-AI-Use-Case-Inventory>). Del mismo modo, la Comisión Europea tiene un catálogo de más de 1 300 (<https://interoperable-europe.ec.europa.eu/collection/public-sector-tech-watch/cases>). En los Gobiernos de América Latina y el Caribe (ALC) se han catalogado más de 700 casos prácticos (<https://sistemaspublicos.tech/sistemas-de-ia-en-america-latina>).

³ Véase <https://trends.oecd-opsi.org>, <https://cross-border.oecd-opsi.org>, (OECD, 2023^[40]), y (OECD, 2024^[108]).

⁴ Véase <https://oecd-opsi.org/innovation-tag/artificial-intelligence-ai> y https://oecd.ai/en/dashboards/policy-instruments/AI_use_cases_in_the_public_sector.

⁵ Estas conclusiones se basan en una encuesta realizada a 1 000 altos directivos de 20 sectores de 59 países de Asia, Europa y América del Norte.

⁶ A menos que se indique lo contrario, las secciones siguientes sobre los beneficios de la IA, en el sector público, se basan en el análisis de las funciones de los gobiernos y en los casos prácticos presentados en el capítulo 5 del presente informe.

⁷ "Ruido" es también un término utilizado en ML para referirse a "fluctuaciones aleatorias o impredecibles en los datos que interrumpen la capacidad de identificar patrones o relaciones objetivo". El resultado es una menor precisión o fiabilidad de las predicciones o resultados de un modelo". (DataHeroes, 2023^[165]). Este no es el concepto de ruido discutido en este capítulo, que se centra en los factores que pueden influir en los seres humanos.

⁸ Para obtener más información sobre este tema, consulte <https://www.oecd.org/en/topics/behavioural-science> y <https://oecd-opsi.org/work-areas/behavioural-insights>.

⁹ Véase también <https://www.oecd.org/en/topics/sub-issues/privacy-enhancing-technologies>.

¹⁰ Véase el Repositorio de riesgos de IA del MIT, una base de datos viva de más de 1.000 riesgos de IA (<https://airisk.mit.edu>). El Monitor de Incidentes de IA (AIM) de la OCDE documenta los incidentes y peligros de la IA para ayudar a los responsables de la formulación de políticas, a los profesionales de la IA y a todas las partes interesadas de todo el mundo a obtener información valiosa sobre los riesgos y daños de los sistemas de IA (<https://oecd.ai/incidents>).

¹¹ Véase <https://oecd.ai/dashboards> y (OECD, 2024^[14]).

¹² El sobreajuste se refiere a los casos en los que el algoritmo es demasiado específico en la medida en que captura y se centra demasiado en el ruido y las anomalías (Berryhill et al., 2019^[8]). Durante la fase de entrenamiento, un modelo de sobreajuste puede lograr un alto nivel de precisión y los problemas pueden pasar desapercibidos. Sin embargo, una vez que el modelo entrenado se expone a nuevos datos, la precisión puede caer gravemente.

¹³ El índice no distingue entre usos legítimos e ilegítimos de las técnicas de vigilancia con IA. Más bien, el propósito de la investigación es mostrar cómo las nuevas capacidades de vigilancia están transformando la capacidad de los Gobiernos para monitorear y rastrear a individuos o grupos.

¹⁴ Taylorismo digital se refiere a la adaptación moderna de los principios de gestión científica de Frederick Winslow Taylor, utilizando tecnologías digitales para monitorear y controlar las actividades de los empleados con el objetivo de mejorar la eficiencia y la productividad. Este enfoque implica desglosar tareas complejas en componentes más simples, estandarizar los flujos de trabajo y emplear métodos basados en datos para supervisar y evaluar el rendimiento de los trabajadores. Si bien su objetivo es optimizar las operaciones organizacionales, los críticos argumentan que puede conducir a una menor autonomía de los trabajadores y una mayor vigilancia en el lugar de trabajo (Grzegorzek, 2024^[163]).

¹⁵ <https://oecd.ai/site/data-privacy>.

¹⁶ La Dirección de Ciencia, Tecnología e Innovación (STI) de la OCDE tiene una línea de trabajo dedicada a la seguridad digital. Consulte <https://www.oecd.org/en/topics/policy-issues/digital-security> para obtener más información.

¹⁷ Un **incidente de IA** es un evento, circunstancia o serie de eventos en los que el desarrollo, uso o mal funcionamiento de uno o más sistemas de IA conduce directa o indirectamente a cualquiera de los siguientes daños: (a) lesiones o daños a la salud de una persona o grupos de personas; (b) interrupción de la gestión y funcionamiento de infraestructuras críticas; (c) violaciones de los derechos humanos o incumplimiento de las obligaciones en virtud de la ley aplicable destinada a proteger los derechos fundamentales, laborales y de propiedad intelectual; o (d) daños a la propiedad, las comunidades o el medio ambiente. Un **peligro para la IA** es un evento, circunstancia o serie de eventos en los que el desarrollo, uso o mal funcionamiento de uno o más sistemas de IA podría dar lugar de manera plausible a un incidente de IA, es decir, a cualquiera de los siguientes daños: (a) lesiones o daños a la salud de una persona o grupos de personas; (b) trastornos en la gestión y funcionamiento de infraestructuras críticas; (c) violaciones de los derechos humanos o incumplimiento de las obligaciones en virtud de la legislación aplicable destinada a proteger los derechos fundamentales, laborales y de propiedad intelectual; o (d) daños a la propiedad, las comunidades o el medio ambiente. Para obtener más información, consulte <https://oecd.ai/incidents-methodology> y (OECD, 2025^[164]; 2024^[166]).

2

Tendencias y primeras lecciones del uso de IA en las distintas funciones de la administración pública

Este capítulo sintetiza 200 casos de uso de IA en 11 funciones gubernamentales. Concluye que el uso de la IA es una prioridad para los gobiernos, pero que su adopción es fragmentada y desigual. Su uso se concentra en los servicios de atención al público y las operaciones internas, con menos ejemplos en la formulación de políticas. Los gobiernos buscan aumentar la productividad, la capacidad de respuesta y la rendición de cuentas, pero los esfuerzos por empoderar a los actores externos son limitados. La madurez varía según la función y la tecnología, con sistemas tradicionales basados en reglas, aprendizaje automático selectivo e inteligencia artificial generativa limitada. Cada caso de uso puede plantear riesgos operativos, éticos, de resistencia o de exclusión si no es confiable, lo que subraya la necesidad de contar con fundamentos sólidos para los datos y una gobernanza coherente.

Mensajes clave

- La OCDE analizó 200 casos de uso de 11 funciones gubernamentales. Concluyó que, si bien la IA es una prioridad para la mayoría de los gobiernos, los esfuerzos no son sistemáticos.
- Este análisis ayudó a comprender mejor el estado actual de la IA en el sector público e identificar tendencias generales. Estas tendencias son las siguientes:
 - La IA está distribuida de forma desigual entre las distintas funciones gubernamentales
 - La IA se utiliza sobre todo en actividades de servicio público de atención al público y operaciones internas
 - Los gobiernos están utilizando la IA para perseguir una serie de potenciales beneficios
 - Algunas funciones gubernamentales son más maduras en lo que respecta a la gobernanza y adopción de IA
 - Las diferentes funciones gubernamentales tienen diferentes contextos y necesidades.
- La OCDE concluyó que cada uno de los 200 casos de uso analizados podría entrañar uno o más tipos de riesgo (operativo, ético, de resistencia pública y de exclusión) si no se diseñan y utilizan de forma confiable.
 - Estos riesgos varían en función del uso que se dé a cada caso y de las funciones de la administración; por lo tanto, es importante reconocer los principales factores de riesgo en cada uso y ámbito.
 - Aunque los gobiernos están atentos a varios riesgos de la IA, algunos reciben menos atención.

Análisis de la OCDE de 200 casos de uso de 11 funciones gubernamentales

En su último trabajo transversal sobre la IA en el sector público, la OCDE (2024^[1]) constató la necesidad de recopilar, documentar y analizar sistemáticamente diferentes casos de uso de la IA para hacer un seguimiento de las tendencias de las opciones de políticas en los distintos países. La OCDE también concluyó que una mayor cantidad y una mejor evidencia del impacto de la IA en los gobiernos contribuirán a garantizar que la tecnología se utilice para lograr un impacto óptimo. La facilidad de acceso a dicha evidencia, así como la información sobre políticas, prácticas y uso de la IA en las administraciones públicas, podría promover avances en la adopción de una IA confiable, un diálogo estructurado e intercambios entre países. En general, se necesita un enfoque holístico de sistemas para maximizar el valor de la IA en las administraciones públicas, inclusive el establecimiento de habilitadores, salvaguardas y mecanismos de participación.

Para ayudar a abordar estas necesidades, este capítulo se basa en investigaciones de la OCDE y otras fuentes relevantes para comprender mejor el estado actual del uso de la IA en la administración pública y poner de relieve las tendencias generales. En este capítulo se analizan y sintetizan 200 casos de uso de la IA que abarcan 11 funciones gubernamentales, enumerados en la Tabla 2.1 y que se analizan en profundidad en el Capítulo 5.¹ Estos casos de uso se identificaron a través de una investigación documental, reuniones de la OCDE y conversaciones con funcionarios públicos en grupos de trabajo y redes pertinentes de la OCDE, y en las recopilaciones de datos en curso del Observatorio de Innovación del Sector Público (OPSI) de la OCDE y del Observatorio de Políticas de IA de la OCDE². Las conclusiones analizadas en este capítulo se basaron en dicha investigación.

Con base en esta metodología, los hallazgos de este capítulo no se pueden generalizar de cara al universo más amplio de esfuerzos de IA en el sector público. Además, la adopción de la IA por parte de las administraciones públicas varía de un país a otro, en función de su situación a escala nacional y de los niveles de preparación en materia de IA. Sin embargo, los hallazgos sí proporcionan observaciones basadas en la práctica del mundo real, las últimas investigaciones y los puntos de vista actuales de los responsables de la creación de políticas. Para ello, el capítulo trata de medir las actividades actuales y sus características, así como de las posibles lagunas en las que puede haber un potencial sin explotar o la necesidad de realizar más investigaciones.

En los próximos meses, la OCDE creará un repositorio mundial dinámico de iniciativas y estudios de casos relevantes en el marco del Observatorio de Políticas de IA de la OCDE.

Tabla 2.1. Análisis de las funciones gubernamentales para el presente reporte

Categoría	Función	Ámbito de análisis
Funciones de política gubernamental	Administración tributaria	Especialistas de la OCDE en diversas áreas gubernamentales recurrieron tanto a estudios internos como externos, y examinaron 200 casos de uso para establecer:
	Gestión financiera pública	
	Diseño e implementación de la regulación	
Procesos gubernamentales clave	Reforma de la función pública	<ul style="list-style-type: none"> - la situación actual, incluidos las acciones y beneficios que el gobierno busca conseguir, - los desafíos relevantes y cómo los están gestionando los gobiernos, y - el potencial sin explotar y el camino que habría que seguir.
	Contratación pública	
	Lucha contra la corrupción y promoción de la integridad pública	
	Evaluación de políticas públicas	
	Participación ciudadana y gobierno abierto	
Servicios gubernamentales y funciones de justicia	Diseño y prestación de servicios públicos	
	Aplicación de la ley y gestión del riesgo de desastres	
	Administración de justicia y acceso a la justicia	

La IA es una prioridad, pero los esfuerzos no son sistémicos

En total, 48 países y la Unión Europea (UE) han adherido a los Principios de la IA de la OCDE (Tabla 2.2), comprometiéndose a promover el diseño, desarrollo, despliegue y uso confiables de la IA, también en el sector público. La OCDE está realizando un seguimiento de la aplicación de estos principios e informa al respecto a lo largo del tiempo (2023^[2]; 2021^[3]). Los resultados sugieren que los gobiernos están menos centrados en el uso de la IA que en sus esfuerzos por promover una adopción confiable de la IA en la economía y la sociedad en general.

Tabla 2.2. Principios de la IA de la OCDE

	Principio	Descripción
Principios basados en valores	Crecimiento inclusivo, desarrollo sostenible y bienestar (Principio 1.1)	Destaca el potencial de una IA confiable para contribuir al crecimiento y a la prosperidad generales para todos —personas, sociedad y planeta— y promover los objetivos de desarrollo mundial.
	Respeto del Estado de Derecho, los derechos humanos y los valores democráticos , incluidas la equidad y la privacidad (Principio 1.2)	Los sistemas de IA deben diseñarse de forma que respeten el Estado de Derecho, los derechos humanos, los valores democráticos y la diversidad, y deben incluir las salvaguardias adecuadas para garantizar una sociedad justa y equitativa.
	Transparencia y explicabilidad (Principio 1.3)	Transparencia y divulgación responsable en torno a los sistemas de IA para garantizar que las personas comprendan cuándo interactúan con ellos y puedan impugnar los resultados.
	Solidez, seguridad y protección (Principio 1.4)	Los sistemas de IA deben funcionar de manera sólida, segura y protegida a lo largo de su vida útil.

	1.4)	toda su vida útil, y los riesgos potenciales deben evaluarse y gestionarse de forma continua.
	Responsabilidad (Principio 1.5)	Las organizaciones y las personas que desarrollan, despliegan u operan sistemas de IA deben rendir cuentas de su correcto funcionamiento de acuerdo con los principios basados en valores de la OCDE para la IA.
Recomendaciones para los responsables de formular políticas	Invertir en investigación y desarrollo de la IA (Principio 2.1)	Los gobiernos deben facilitar la inversión pública y privada en investigación y desarrollo para estimular la innovación en IA confiable.
	Fomentar un ecosistema inclusivo que propicie la IA (Principio 2.2)	Los gobiernos deben fomentar ecosistemas de IA accesibles dotados de una infraestructura y de tecnologías digitales, así como mecanismos para compartir datos y conocimientos, y garantizar la calidad de dicha información.
	Dar forma y habilitar un entorno político y de gobernanza interoperable para la IA (Principio 2.3)	Los gobiernos deben crear un entorno de políticas que abra el camino al despliegue de sistemas de IA confiables.
	Construir capacidad humana y prepararse para la transformación del mercado laboral (Principio 2.4)	Los gobiernos deben dotar a las personas de las competencias necesarias para desarrollar la IA y apoyar a los trabajadores para garantizar una transición justa.
	Cooperación internacional en aras de una IA confiable (Principio 2.5)	Los gobiernos deben cooperar más allá de las fronteras y los sectores para compartir información, desarrollar normas y trabajar en pos de una administración responsable en materia de IA.

Fuente: <https://oecd.ai/en/ai-principles>.

Los gobiernos están haciendo realidad el potencial de la IA en las administraciones públicas y convirtiéndola en una prioridad estratégica. **Casi todos los países de la OCDE han puesto en marcha estrategias y programas para la IA que establecen una visión y un enfoque de alto nivel para su uso en el gobierno.** Estos están integrados principalmente en estrategias nacionales con respecto de la IA más amplias. Sin embargo, países como Canadá (2025^[4]), Suiza (2025^[5]) y Uruguay (2021^[6]) han desarrollado estrategias específicas. Muchos gobiernos también han intentado poner en práctica sus estrategias mediante políticas e iniciativas transversales o específicas de cada ámbito. Estas iniciativas se analizan más a fondo en el Capítulo 4. Además, muchos gobiernos han adoptado la IA a través del uso práctico y, cada vez más, desde el desarrollo propio.

Se han realizado avances desde que la OCDE comenzó a explorar la IA en el gobierno en 2019; sin embargo, los **esfuerzos realizados hasta la fecha son limitados y no sistemáticos**. Las oportunidades potenciales de la IA en el gobierno son significativas, pero no resultan fáciles de alcanzar. Los gobiernos de dos tercios de los países de la OCDE han empezado a explorar el uso de la IA para aumentar la eficiencia interna mediante la mejora de los procesos. Sin embargo, se necesitan más avances no solo en el uso de la IA para otros fines, como la mejora de las políticas, sino también en la construcción de los componentes fundamentales necesarios para que la IA en el gobierno florezca (OECD, 2024^[1]). Además, la revisión de los casos de uso sugiere una proliferación de herramientas de IA, cuya implementación a menudo se produce de forma gradual. Estos esfuerzos suelen llevarse a cabo sin contar con mecanismos de gobernanza generales que orienten las iniciativas en los distintos sectores o en el conjunto del gobierno, o que permitan extraer enseñanzas de su aplicación. Como resultado, el potencial para el aprendizaje coordinado, la escala y el impacto sigue siendo limitado. Establecer marcos de gobernanza sólidos podría ayudar a garantizar que los sistemas de IA se desplieguen de forma cohesionada, eficiente y responsable, en consonancia con las prioridades estratégicas y los valores públicos.

En las siguientes secciones se pretende dar a conocer la situación actual del uso de la IA en la administración pública, incluidos los patrones y tendencias clave de los primeros en adoptarla. De este modo, estas secciones tratan de identificar en qué medida los gobiernos van más allá de los principios para tomar medidas en el uso de la IA, qué resultados están logrando y a qué limitaciones se enfrentan.

Tendencias generales en los casos de uso de la IA por parte de las administraciones públicas

Distribución desigual del uso de la IA entre las funciones gubernamentales

El Índice de Gobierno Digital (DGI, por sus siglas en inglés) de la OCDE de 2023 ([\[7\]](#)) muestra que, si bien algunos países han puesto en marcha una amplia gama de iniciativas para mejorar su capacidad de utilizar la IA en la administración pública, su aplicación sigue planteando dificultades en la mayoría de los países. Al profundizar en los casos de uso analizados para este informe, el análisis de la OCDE sugiere que los esfuerzos gubernamentales en materia de IA pueden agruparse en torno al diseño y la prestación de servicios públicos, la participación cívica y el gobierno abierto, y la administración de justicia y el acceso a la justicia. Por el contrario, solo se identificaron algunas iniciativas en funciones como la evaluación de políticas (Gráfico 2.1).

Gráfico 2.1. Los casos de uso están más presentes en los servicios públicos, la participación cívica y las funciones de justicia



Fuente: Análisis de la OCDE de los casos de uso identificados.

Existen varias explicaciones potenciales para esta distribución:

- El diseño y prestación de los servicios públicos se extienden horizontalmente a través de muchos tipos diferentes de organizaciones y áreas temáticas, lo que los hace más frecuentes en términos de casos de uso totales que en el caso de funciones más verticales del gobierno, como la administración tributaria.
- La prevalencia de casos de uso en participación cívica y gobierno abierto podría atribuirse a que estas áreas no están tan expuestas a muchos de los riesgos (Capítulo 1) y desafíos (Capítulo 3) que afectan a otras funciones. Por ejemplo, las preocupaciones en torno al acceso a los datos y la seguridad no son aplicables en gran medida porque el objetivo de estas funciones suele ser recopilar datos sobre problemas y cuestiones de carácter público. Además, los equipos

gubernamentales que participan en la participación cívica suelen estar entre los grupos más innovadores de la función pública y, por lo tanto, tal vez sean más propensos a adoptar nuevos enfoques tecnológicos.

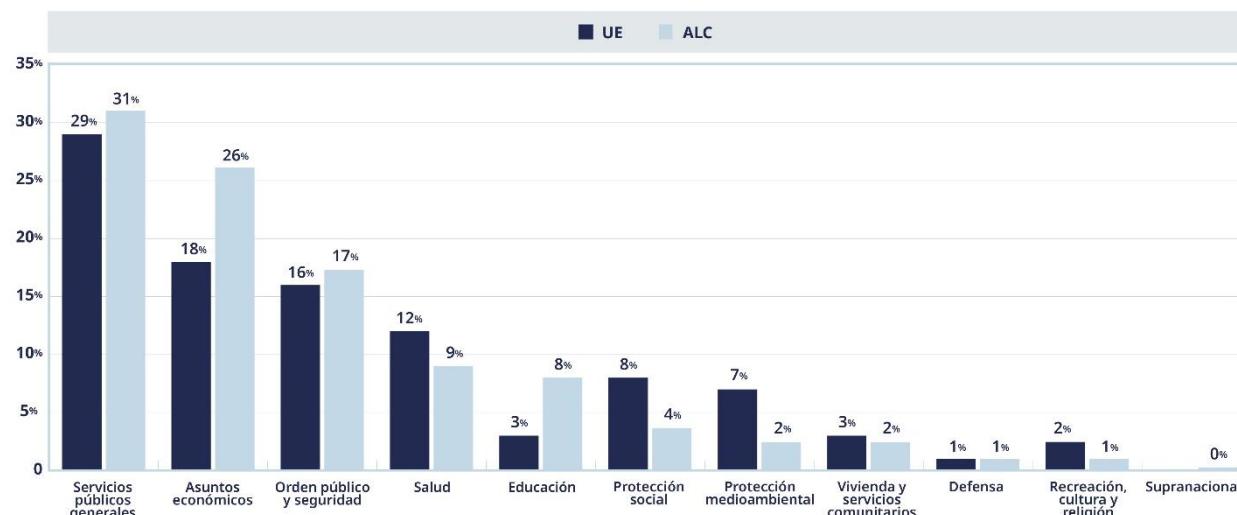
- Las funciones de políticas más representadas suelen ser públicas, lo que podría indicar que se centran en áreas de visibilidad inmediata para los ciudadanos. Algunos factores podrían ser el mayor número de demandas de los ciudadanos, así como un deseo entre el gobierno y los líderes políticos de demostrar visiblemente su valor.
- Algunas funciones se enfrentan a obstáculos o complejidades, como normas estrictas sobre el acceso y el intercambio de datos en la administración tributaria y requisitos para llevar a cabo auditorías exhaustivas de la integridad pública.
- Algunas funciones parecen estar más maduras que otras relacionadas con la preparación para la IA, incluidos sus fundamentos subyacentes para la IA, como datos suficientes y de calidad, como se analiza a continuación.
- Algunas funciones pueden tener estructuras y procesos preexistentes que no pueden sustituirse o complementarse fácilmente con sistemas de IA.

La prevalencia del uso de la IA en la justicia y la función conexa de la aplicación de las leyes son especialmente interesantes. En general, la OCDE (2024^[1]) ha alentado a los gobiernos a apuntar a resultados fáciles de alcanzar en su iniciación en el uso de la IA, centrándose en áreas que representan usos de la IA de alto beneficio y bajo riesgo. El uso de la IA en la justicia y en la aplicación de la ley puede reportar grandes beneficios, pero también un alto riesgo. Una razón para la prevalencia de casos de uso en algunas de estas áreas puede ser la prevalencia de datos más completos y estructurados. Otra razón puede estar en línea con el gran volumen de tareas necesarias para algunas de estas funciones. En el caso de la administración de justicia y el acceso a la justicia, los sistemas judiciales de todo el mundo suelen funcionar con recursos limitados, incluido un presupuesto y un personal judicial limitados, aun teniendo en cuenta que el volumen de casos sigue creciendo (Harvard Kennedy School, 2023^[8]; Columbia University, 2020^[9]). Este desajuste ha provocado retrasos crónicos en los procesos en muchas jurisdicciones, lo que ha creado una intensa presión sobre los administradores judiciales para que exploren tecnologías que puedan impulsar la productividad y mitigar el problema de los retrasos. Esto podría reflejarse en el mayor número de casos de uso de la IA en la justicia relacionados con operaciones internas, en comparación con otras funciones gubernamentales (Gráfico 2.3).

Un motivo adicional de esta distribución puede ser la disponibilidad de datos para su examen, ya que la OCDE tiende a identificar (o los gobiernos tienden a presentar) información sobre iniciativas que desempeñan algunas funciones más que otras. Esto parece atenuado en cierta medida por la validación en debates y revisiones de la OCDE por parte del Grupo de Trabajo de la OCDE sobre Altos Funcionarios de Gobierno Digital (E-Leaders), así como por las comparaciones con los resultados de bases de datos más grandes. En cuanto a esta última, la recopilación de datos para el presente informe se ajusta a las tendencias observadas en la UE y en América Latina y el Caribe (ALC), según lo registrado por el Observatorio Tech Watch del Sector Público de la Comisión Europea (CE) (2025^[10]) y la base de datos "Sistemas de IA en el sector público de América Latina y el Caribe" (Muñoz-Cadena et al., 2025^[11]) (Gráfico 2.2). En la UE, las tres principales funciones identificadas en su repositorio de casi 1 500 casos de uso de la IA, a 31 de marzo de 2025, son los servicios públicos generales, los asuntos económicos y el orden público y la seguridad³. La categoría de orden público y seguridad incluye los casos de uso tanto en la aplicación de la ley como en la administración de justicia. La categoría de asuntos económicos está representada principalmente por casos de uso en sectores como el transporte, la agricultura y la energía. Alrededor del 70 % de los casos de uso de la IA en asuntos económicos está relacionado con la regulación y servicios públicos específicos y la participación. ALC presenta una tendencia comparable en las tres principales funciones del repositorio revisado de aproximadamente 700 sistemas de IA, actualizado el 19 de marzo de 2025.

Gráfico 2.2. La UE y ALC siguen una tendencia similar con la muestra de casos de uso de IA recopilada para este informe

Porcentaje de casos de uso categorizados según la Clasificación de Funciones de Gobierno de las Naciones Unidas



Fuente: Datos analizados por la OCDE de (EC, 2025^[10]; Muñoz-Cadena et al., 2025^[11]),

Un último factor que también puede influir en los resultados es una variación del "efecto IA", según el cual "tan pronto como los investigadores [en el campo de la IA] logran un hito que durante mucho tiempo se pensó que significaba el logro de la verdadera inteligencia artificial, por ejemplo, vencer a un humano en el ajedrez, repentinamente deja de considerarse como auténtica IA" (Bailey, 2016^[12]). En las discusiones con el grupo de trabajo E-Leaders, los delegados han sugerido que las aplicaciones estrechas y tradicionales de la IA pueden haberse vuelto tan integradas o comunes que ya no activan informes externos ni una respuesta a los esfuerzos de recopilación de datos. Esto podría ocurrir sobre todo en ámbitos en los que se hace un uso prolongado de estos sistemas, como la administración tributaria, lo que da lugar a una menor representación en las iniciativas examinadas. Es difícil determinar en qué medida puede ocurrir esto; sin embargo, en el análisis del presente informe se identificaron e incluyeron muchos de esos casos de uso.

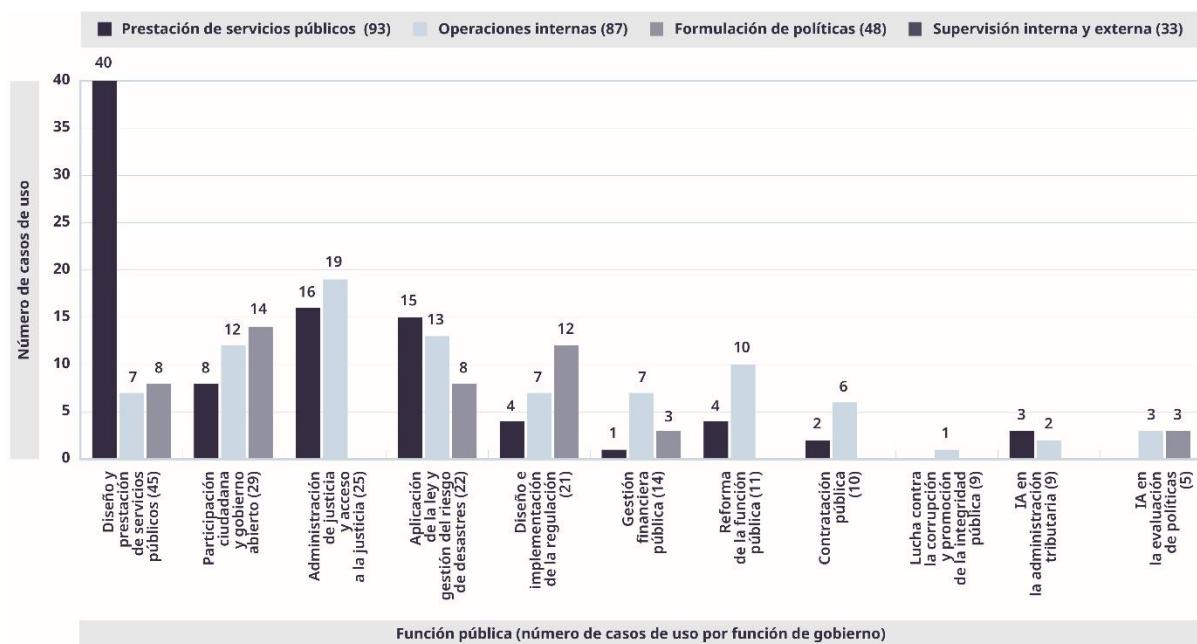
La IA se utiliza sobre todo en actividades de servicio público de atención al público y operaciones internas

En las 11 funciones centrales cubiertas por este informe, los gobiernos están utilizando la IA en cuatro actividades generales: prestación de servicios de atención al público, operaciones internas, actividades de supervisión interna y externa y asistencia en la formulación de políticas. El uso de la IA es más común en la prestación de servicios de atención al público. Las actividades relacionadas con las operaciones internas no se quedan atrás (Gráfico 2.3). Las actividades de supervisión interna y externa y la asistencia para la formulación de políticas no son tan frecuentes. Esto no es inesperado, ya que la prestación de servicios y las operaciones internas constituyen la mayoría de lo que hacen las organizaciones gubernamentales. Las actividades de supervisión son importantes, aunque con mayor frecuencia se limitan a determinadas oficinas o equipos. El uso de IA en las actividades de formulación de políticas tampoco es tan frecuente entre los gobiernos. Esta conclusión concuerda con mediciones anteriores del DGI de la OCDE (2024^[7]) que también indicaba que los gobiernos podían hacer más esfuerzos en este

ámbito. Muchos siguen siendo cautelosos o carecen de las habilidades necesarias para incorporar la IA en los procesos de toma de decisiones.

La prevalencia de la IA en las actividades gubernamentales varía en función de la naturaleza de cada función esencial. En la función de diseño y prestación de servicios públicos, la mayoría de las aplicaciones de IA implican, naturalmente, interacciones de atención al público, y algunos casos de uso también abordan operaciones internas con respecto a cómo se diseñan o prestan los servicios públicos. El importante número de casos de uso relacionados con los servicios en el sector de la justicia, que comprende casi tres cuartas partes (16 de 25 casos, señalando que un caso puede referirse a más de una actividad) de las instancias documentadas en esa función, indica que esta función ha dado prioridad a la capacidad de respuesta a los ciudadanos, junto con la mejora de la eficiencia en sus operaciones internas. Este enfoque puede verse influido por una mayor demanda social, la presión para reducir los retrasos en los casos y gestionar los recursos escasos. Funciones como la participación cívica y el diseño e implementación de la regulación abarcan la mayoría de los casos de uso relacionados con las actividades de formulación de políticas. Entre los usos relacionados se incluyen el apoyo al procesamiento de evidencia y las contribuciones de las partes interesadas para la formulación de políticas, y diversas aplicaciones que ayudan a la toma de decisiones mediante análisis, simulaciones o previsiones.

Gráfico 2.3. Casos de uso de IA por función esencial y actividad de la administración pública



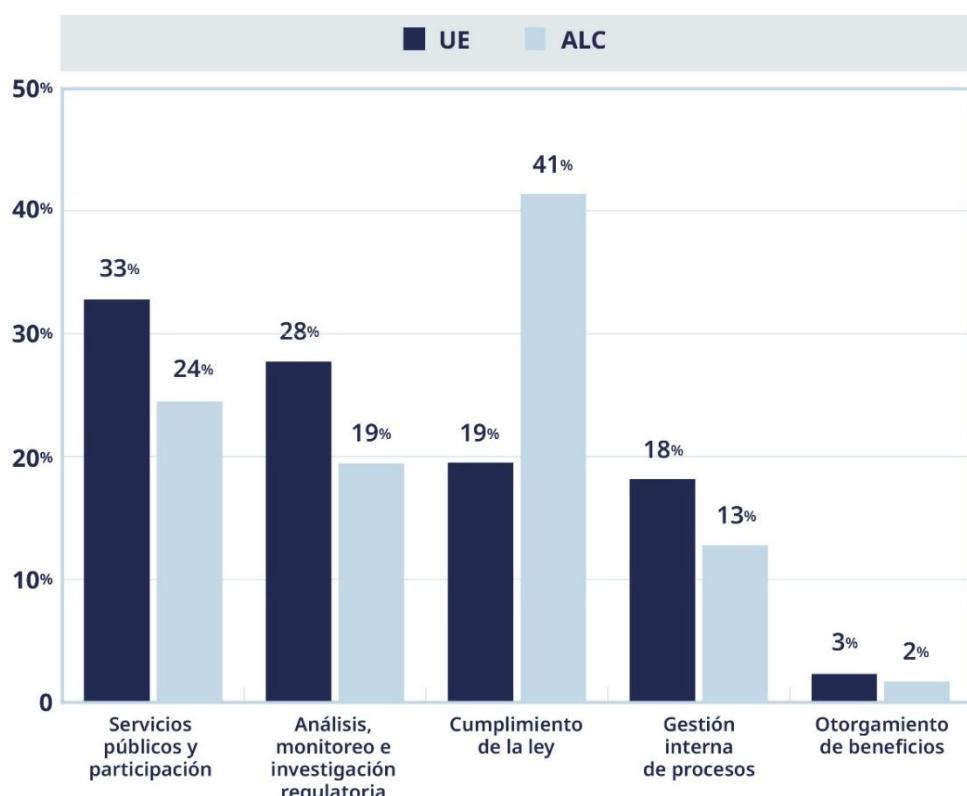
Nota: Las cuatro actividades de este gráfico no son mutuamente excluyentes (por ejemplo, un caso práctico de IA podría tratar de mejorar tanto las operaciones internas como la prestación de servicios). Por lo tanto, la suma de actividades es mayor que el número total de casos de uso.
Fuente: Análisis de la OCDE de los casos de uso identificados.

Por lo general, estos resultados coinciden con las conclusiones de la UE (2025^[10]) y la base de datos de ALC (2025^[11]). Si bien no existe una metodología uniforme para clasificar las actividades gubernamentales, estas bases de datos y las de la OCDE tienden a mostrar una tendencia común, particularmente en torno a los servicios. Tanto la base de datos de la UE como la de ALC incluyen una clasificación de casos de uso en el marco de actividades gubernamentales (Gráfico 2.4), donde los "servicios públicos y la participación" representan una parte significativa de todos los casos de uso, siendo los más destacados de la UE y los segundos de ALC. Sin embargo, en ALC, la mayoría de los casos de uso se clasifican como parte de las actividades de "aplicación efectiva", que incluyen la aplicación

predictiva, el registro y la certificación de datos, el reconocimiento inteligente y las inspecciones de apoyo, entre otros. Por lo general, estos procesos coincidirían con las operaciones internas y las categorías de supervisión (internas y externas) de la OCDE. Esto significa que la categoría de "gestión interna" en la UE y ALC, que generalmente es menos representativa de estas bases de datos, no es la única que contiene casos de uso que la OCDE podría clasificar como "operaciones internas". Por lo tanto, no es posible concluir si el conjunto de datos de la OCDE contiene una mayor proporción de casos de uso pertenecientes a operaciones internas de la administración, en comparación con la UE y ALC. Por último, cabe destacar que los procesos incluidos en la categoría "análisis, supervisión e investigación regulatoria" en la UE y ALC suelen coincidir con la categoría "formulación de políticas" de la OCDE y presentan una proporción similar (alrededor del 20 - 30 % de todos los casos de uso).

Gráfico 2.4. Los servicios públicos y la participación representan una parte importante de los casos de uso en los procesos gubernamentales de la UE y ALC

Porcentaje de casos de uso clasificados por el modo en que la tecnología respalda la toma de decisiones y la implementación por parte de la administración pública



Fuente: Datos analizados por la OCDE de (EC, 2025^[10]; Muñoz-Cadena et al., 2025^[11]).

Los gobiernos están utilizando la IA en busca de una variedad de beneficios potenciales

Los casos de uso analizados por la OCDE tienen el potencial de abordar todos los beneficios de la IA descritos en el Capítulo 1, y la mayoría de ellos tienen el potencial de generar múltiples beneficios. Sin embargo, algunos beneficios reciben más atención de los gobiernos que otros (Gráfico 2.5):

- Alrededor de seis de cada diez casos de uso examinados tienen por objeto contribuir a la automatización, la racionalización y la adaptación y personalización de procesos y servicios, en

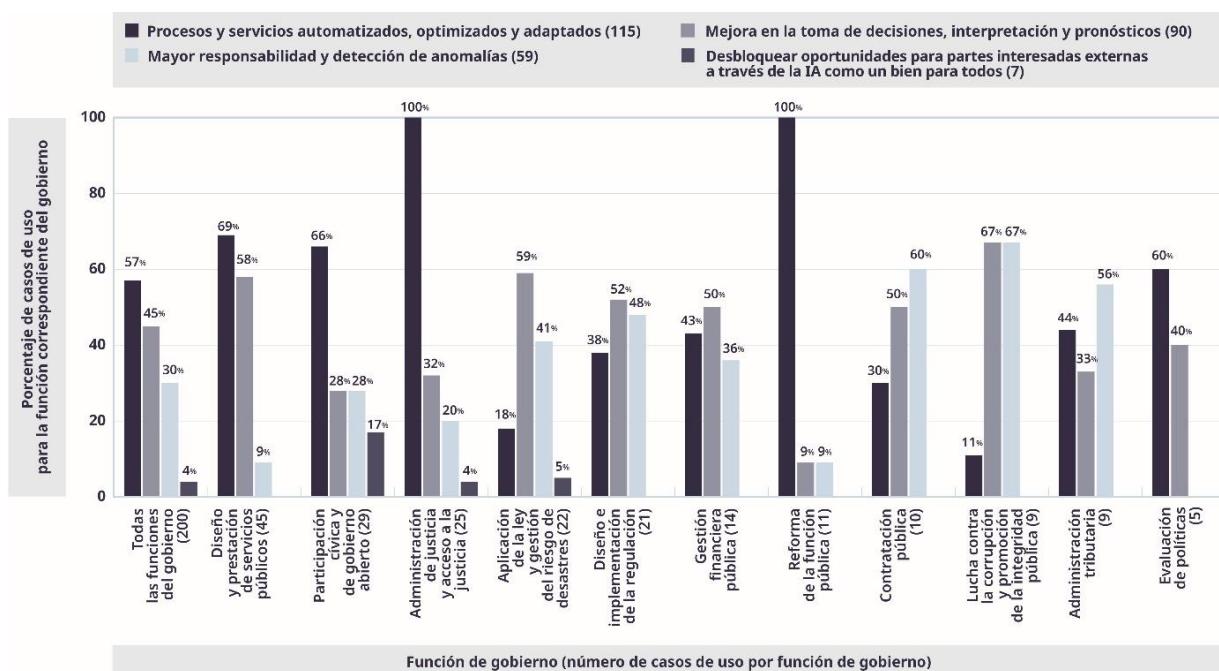
particular en el ámbito de justicia, servicios públicos, participación ciudadana y el diseño e implementación de la regulación.

- Casi la mitad de los casos de uso tienen por objeto mejorar la toma de decisiones, interpretación y pronósticos, y la mayoría de ellos se concentran en los servicios públicos, regulación y participación ciudadana.
- Alrededor de un tercio de los casos de uso tienen el potencial de mejorar la rendición de cuentas y detección de anomalías, principalmente en el marco de la aplicación de la ley y gestión del riesgo de desastres, la participación cívica, la lucha contra la corrupción y la promoción de la integridad pública, contratación pública y regulación.
- Una pequeña proporción de casos de uso tiene el potencial de brindar oportunidades a partes interesadas externas, como ciudadanos, OSC y empresas, especialmente en materia de participación cívica, acceso a la justicia y gestión del riesgo de desastres.

La falta de énfasis en desbloquear oportunidades para las partes interesadas externas, a través de la IA como un bien para todos, destaca como una brecha potencial. Los expertos en IA han sugerido que este tipo de empoderamiento es importante y que los gobiernos podrían tomar más medidas para aprovecharlo (OECD, 2024^[13]). Sin embargo, estos esfuerzos podrían ser más frecuentes en áreas no cubiertas por este informe (por ejemplo, sectores específicos, como la agricultura o la educación). Si bien los resultados de este beneficio pueden ser percibidos de manera menos directa por los gobiernos, pueden pagar dividendos a través de una mayor confianza en el gobierno o incluso ganancias económicas.

Gráfico 2.5. Beneficios potenciales de los casos de uso de IA en distintas funciones de la administración pública

Porcentaje de casos de uso para la función correspondiente del gobierno

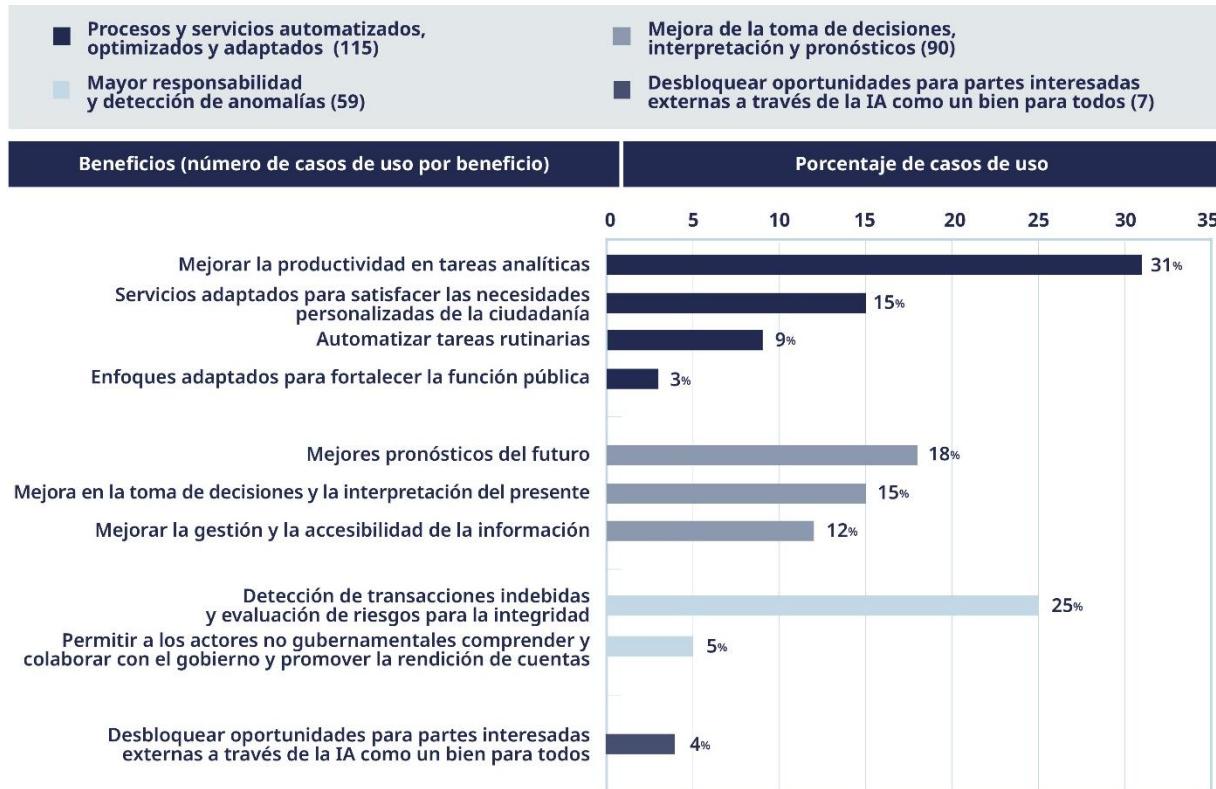


Nota: Los beneficios potenciales en este gráfico no son mutuamente excluyentes (es decir, un caso práctico puede ofrecer más de un tipo de beneficio). Por lo tanto, la suma de los beneficios potenciales observados es mayor que el número total de casos de uso.

Fuente: Análisis de la OCDE de los casos de uso identificados.

Al examinar los beneficios específicos dentro de las cuatro categorías generales de actividad mencionadas anteriormente, se puede obtener información más detallada sobre los beneficios directos que los gobiernos aspiran a lograr mediante el uso de IA (como se muestra en Gráfico 2.6). En las secciones siguientes se describen con más detalle estas ventajas y se ofrecen ejemplos de algunos de los casos de uso que sirvieron de base para estas tendencias.

Gráfico 2.6. Beneficios específicos de casos de uso de IA



Nota: Los beneficios potenciales en este gráfico no son mutuamente excluyentes (es decir, un caso práctico puede ofrecer más de un tipo de beneficio). Por lo tanto, la suma de los beneficios potenciales observados es mayor que el número total de casos de uso.

Fuente: Análisis de la OCDE de los casos de uso identificados.

Procesos y servicios automatizados, racionalizados y adaptados

Alrededor de un tercio (31 %) de los casos de uso analizados tiene como objetivo mejorar la productividad en las tareas analíticas. En menor medida, el 15 % de los casos de uso representan esfuerzos gubernamentales para utilizar IA con el fin de adaptar los servicios a las necesidades personalizadas de los ciudadanos. Esta adopción relativamente menor de IA para la personalización podría atribuirse en parte a las limitaciones o restricciones de la gobernanza de datos debido al gran volumen de datos personales necesarios para este tipo de aplicaciones (véase el Capítulo 3 para consultar los desafíos de implementación). Esta parece ser una laguna que merece un análisis más detallado. Curiosamente, la automatización de tareas rutinarias llega al 9 % de los casos de uso analizados. Esto es contrario a las expectativas convencionales de que la IA se utilice principalmente para automatizar tareas repetitivas que requieren poca consideración analítica (Gráfico 2.6). Si bien una metodología de revisión de casos no es totalmente generalizable para el universo de la IA en el gobierno, esto sugiere un posible cambio en el uso de IA para mejorar los procesos de toma de decisiones más complejos y apoyar el trabajo más especializado de servidores públicos y responsables de la formulación de políticas. Sin embargo, también

podría sugerir que los gobiernos no están capitalizando plenamente la IA con respecto a las tareas repetitivas, en las que los servidores públicos dedican una cantidad significativa de tiempo y para las que se pueden lograr enormes eficiencias a través de la IA (The Alan Turing Institute, 2024^[14]; Berryhill et al., 2019^[15]).

La Tabla 2.3 presenta una serie de ejemplos de cómo se está utilizando la IA para estos fines. Los casos de uso destinados a mejorar la productividad en las tareas analíticas incluyen usos como la estimación de costos de cumplimiento en evaluaciones del impacto normativo (Alemania), el análisis y la puntuación de las respuestas registradas de los candidatos en determinados procesos de contratación (Reino Unido) o el apoyo al personal gubernamental con consultas comunes sobre contratación (Carolina del Norte, Estados Unidos). La automatización en tareas repetitivas que requieren un menor compromiso intelectual engloba diversos ámbitos, como las tareas judiciales repetitivas o los procesos financieros y de recursos humanos. Este es el caso de Prometea en Argentina, el Proyecto de resolución de litigios en materia de IA en Brasil o el uso por Finlandia de la RPA y la IA en la gestión financiera. Los usos destinados a adaptar los servicios y la personalización pueden observarse en funciones como los servicios públicos, los impuestos, la regulación o la justicia. Por ejemplo, el Servicio Público de Empleo de Suecia utiliza BÄR para adaptar el apoyo a la búsqueda de empleo, optimizando la asignación de recursos a través de recomendaciones personalizadas de formación y orientación. En el caso de Singapur, la autoridad tributaria desarrolló un chatbot para mejorar el autoservicio al ayudar a los contribuyentes con sus consultas y pagos. Por último, la IA se está utilizando para reforzar programas de contratación y desarrollo profesional de la función pública, como el ensayo de la Comisión Australiana de Administración Pública para utilizar la IA con el fin de agilizar el diseño, estructuración y despliegue de la formación en competencias digitales; o el uso del Instituto Nacional español de Administración Pública para transformar la forma en que los funcionarios acceden y utilizan los recursos de aprendizaje mejorando la capacidad de búsqueda y las recomendaciones de los materiales pertinentes.

Tabla 2.3. Ejemplos de IA para procesos y servicios automatizados, racionalizados y adaptados

País	Iniciativa	Descripción	Sub-beneficio	Función
Argentina	Prometea y ChatGPT en el sector justicia	El Ministerio Público de Buenos Aires adoptó Prometea en 2017 para automatizar las tareas judiciales repetitivas y agilizar los procedimientos. En 2024, también comenzó a explorar el uso de ChatGPT para analizar casos legales y proyectos de decisiones. Esta herramienta de IA reduce el tiempo de redacción de las sentencias de 60 a 10 minutos, lo que aumenta la eficiencia en la gestión de casos.	Automatizar tareas rutinarias (y recientemente, analíticas)	Justicia (Recuadro 5.62)
Australia	La IA para generar un aprendizaje en línea	La Comisión Australiana de Administración Pública (APSC) probó acelerar la creación de cursos para servidores públicos utilizando IA para diseñar, estructurar e implementar la capacitación en competencias digitales en cuestión de minutos en lugar de semanas. El piloto probó la alimentación en materiales controlados para generar esquemas de curso y cuestionarios y refinar el contenido a través de circuitos de retroalimentación.	Enfoques adaptados para fortalecer la función pública	Reforma de la función pública (Recuadro 5.22)
Brasil	Proyecto de litigios de IA	Los tribunales tributarios de Brasil utilizan IA para agrupar casos similares de apelación fiscal, asignándolos a los mismos funcionarios para una tramitación más rápida. Los ensayos iniciales demostraron una alta precisión, reduciendo significativamente la acumulación de casos y mejorando la velocidad de decisión.	Automatizar tareas rutinarias	Administración tributaria (Recuadro 5.2)
Finlandia	RPA e IA en Gestión Financiera	Finlandia aprovecha una herramienta que automatiza los procesos financieros y de recursos humanos a través de RPA e IA, optimizando tareas como el procesamiento de facturas. Su estrategia de automatización estructurada mejora la escalabilidad y eficiencia.	Automatizar tareas rutinarias	Finanzas públicas
Alemania	IA para evaluaciones de impacto normativo	La Oficina Federal de Estadística de Alemania está explorando el uso de IA para estimar los costos de cumplimiento en las evaluaciones de impacto regulatorio. La IA identifica pasajes pertinentes del texto jurídico y predice las implicaciones en materia de costos, lo que permite a los funcionarios centrar los recursos en casos complejos.	Mejorar la productividad en tareas analíticas	Reglamento (Recuadro 5.13)

Singapur	Chatbot para servicios al contribuyente	La autoridad tributaria de Singapur desarrolló un chatbot que utiliza IA y NLP para ayudar a los contribuyentes con sus consultas y pagos. El sistema mejora las opciones de autoservicio, reduciendo la carga de trabajo administrativo y mejorando la satisfacción del usuario.	Servicios a medida para satisfacer necesidades personalizadas	Administración tributaria (Recuadro 5.4)
España	Gráfico de conocimientos	El gráfico de conocimientos mejorados por IA del Instituto Nacional de Administración Pública (INAP) transforma la forma en que los funcionarios acceden y utilizan los ingentes recursos de aprendizaje. Al crear un "banco de recursos" que mejora la capacidad de búsqueda y recomienda materiales pertinentes, el INAP permite a los funcionarios públicos encontrar y aplicar eficazmente conocimientos críticos.	Enfoques adaptados para fortalecer la función pública	Reforma de la función pública
Suecia	BÄR	El Servicio Público de Empleo utiliza BÄR, una herramienta de IA incluida en el programa Prepare and Match, para adaptar el apoyo a la búsqueda de empleo. Mediante el análisis de los perfiles de los demandantes de empleo y la predicción de las posibilidades de empleo, orienta las decisiones y optimiza la asignación de recursos a través de recomendaciones personalizadas de formación y orientación.	Servicios a medida para satisfacer necesidades personalizadas	Reforma de la función pública (5.43)
Reino Unido	Outmatch	La autoridad tributaria del Reino Unido (HMRC) emplea Outmatch para automatizar la contratación de personal subalterno mediante el análisis y la puntuación de las respuestas registradas de los candidatos. Esto agiliza la contratación de gran volumen y, al mismo tiempo, garantiza la coherencia en la evaluación.	Mejorar la productividad en tareas analíticas	Reforma de la función pública (Recuadro 5.20)
Estados Unidos	Chatbot de apoyo a la contratación pública	El departamento de TI de Carolina del Norte introdujo un chatbot 24/7 impulsado por IA para apoyar al personal del gobierno con consultas comunes de adquisiciones. Proporciona respuestas instantáneas, agiliza los procesos y reduce los tiempos de espera.	Mejorar la productividad en tareas analíticas	Contratación pública

Mejora en la toma de decisiones, interpretación y pronósticos

El uso de la IA para mejorar la toma de decisiones y la interpretación del presente se midió en el 18% de los casos, mientras que el 15 % de los casos de uso tenían como objetivo mejorar los pronósticos del futuro y el 12 % la gestión de información y accesibilidad para respaldar estas actividades (Gráfico 2.6). Estos usos de la IA no solo respaldan los procesos de formulación de políticas, que de hecho son una minoría en lo que respecta a los esfuerzos de los gobiernos en materia de IA (OECD, 2024^[7]), sino que también contribuyen a una aplicación más inteligente de las políticas y al funcionamiento interno, así como a una mejor calidad y pertinencia del diseño y prestación de servicios.

La Tabla 2.4 presenta algunos ejemplos de cómo se está utilizando la IA para estos fines. Los gobiernos están utilizando herramientas de código abierto como Polis para darle una mejor interpretación al presente, específicamente en ejercicios deliberativos, donde se agrupa opiniones públicas e identifica áreas de consenso. Otros usos permiten a los gobiernos optimizar mejor la toma de decisiones. Un ejemplo es dBrain+ de Corea, que analiza datos de gestión financiera en tiempo real e integra herramientas de evaluación de riesgos, gestión presupuestaria y evaluación del rendimiento. La mayoría de los casos de uso de pronósticos tienen como objetivo predecir ciertas condiciones para tomar decisiones por adelantado y preposicionar recursos, como predecir condiciones resbaladizas para el mantenimiento de carreteras en invierno en Bélgica o predecir la probabilidad de incendios forestales en Canadá, para mejorar la mitigación de riesgos y los tiempos de respuesta de las autoridades. Otros usos de la predicción tienen como objetivo simular escenarios alternativos. Un ejemplo es el uso que hace Helsinki (Finlandia) de UrbanistAI para generar visualizaciones de escenarios de planificación urbana alternativos con el fin de apoyar la creación de consenso entre las partes interesadas. Por último, el uso de la IA para mejorar la gestión de la información y la accesibilidad puede adoptar la forma de herramientas a disposición de las partes interesadas, tanto dentro como fuera del gobierno, para acceder a grandes cantidades de datos. Algunos ejemplos son la herramienta de búsqueda del Parlamento Europeo, que permite a los usuarios analizar más de veinte años de documentos parlamentarios, o la plataforma piloto IA generativa del Tribunal de Cuentas neerlandés para analizar los informes de auditoría pública. También podría adoptar la forma de herramientas de asistencia para recuperar rápidamente

información gubernamental precisa y garantizar respuestas confiables en los servicios de asistencia al cliente, como los asistentes virtuales *Caddy* en el Reino Unido y *Albert* en Francia. Si bien esos asistentes abarcan diferentes dominios, algunos se centran en áreas específicas, como el agente conversacional *Sofia* de Francia para la información ecológica.

Tabla 2.4. Ejemplos de IA para mejorar la toma de decisiones, interpretación y pronósticos

País	Iniciativa	Descripción	Sub-beneficio	Función
Bélgica	IA para la seguridad vial	Bélgica predice condiciones resbaladizas y optimiza la asignación de recursos para el mantenimiento de carreteras en invierno. Al analizar los datos meteorológicos y de tráfico, la herramienta ayuda a las autoridades a aplicar de forma proactiva medidas de deshielo, mejorando la seguridad vial y reduciendo los accidentes.	Mejores pronósticos del futuro	Servicios públicos; relacionados con la gestión de catástrofes
Canadá	Anticipación de los riesgos de incendios forestales	El sistema de predicción de incendios forestales con IA de Alberta pronostica la probabilidad de incendios forestales en los bosques protegidos de la provincia utilizando datos históricos de incendios, meteorológicos y ecológicos. La herramienta ayuda a las autoridades a preponer recursos, mejorar los tiempos de respuesta y mitigar los riesgos.	Mejores pronósticos del futuro	Aplicación de la ley y gestión de desastres (Recuadro 5.54)
Unión Europea	IA para examinar documentos parlamentarios	La herramienta de búsqueda del Parlamento Europeo permite a los ciudadanos y a los responsables políticos analizar eficazmente más de veinte años de documentos parlamentarios, incluidas 38 000 propuestas de resolución y preguntas parlamentarias. Al automatizar la recuperación de información, el sistema de IA mejora la accesibilidad y facilita la toma de decisiones informadas.	Gestión/accesibilidad de la información mejorada	Participación ciudadana y gobierno abierto (Recuadro 5.35)
Finlandia	Inteligencia artificial urbana	La ciudad de Helsinki utilizó UrbanistAI para generar visualizaciones de escenarios alternativos de planificación urbana, ayudando a los ciudadanos y a las empresas locales a participar en debates sobre la peatonalización de calles clave. Gracias a los renderizados generados por IA, la herramienta propició la creación de consenso entre las partes interesadas.	Mejores pronósticos del futuro	Participación ciudadana (Recuadro 5.38)
Francia	Albert y Sofia	Albert es una herramienta IA generativa desarrollada para ayudar a los empleados de las administraciones públicas a responder a las consultas de los ciudadanos. La herramienta ayuda a los funcionarios a buscar regulaciones, resumir información y redactar respuestas, mientras que los agentes humanos verifican el resultado final. Sofia es una agente conversacional que facilita el acceso al conocimiento científico y técnico del Ministerio de Transición Ecológica.	Mejora de la accesibilidad de la gestión de la información	Servicios públicos (Recuadro 5.46)
Internacional	Polis	Polis es una plataforma de código abierto impulsada por IA diseñada para facilitar procesos deliberativos a gran escala mediante la agrupación de opiniones públicas y la identificación de declaraciones de consenso. Se ha utilizado en múltiples países para fundamentar la política climática, los debates de referéndum, la toma de decisiones municipales y las plataformas de partidos políticos.	Mejora de la toma de decisiones e interpretación del presente	Participación ciudadana (Recuadro 5.36)
Corea	dBrain+	dBrain+ es un sistema de gestión financiera basado en IA que analiza datos económicos, fiscales y financieros en tiempo real para optimizar la toma de decisiones de finanzas públicas. Integra la evaluación de riesgos, la gestión presupuestaria y la evaluación del desempeño.	Mejora de la toma de decisiones e interpretación del presente	Finanzas públicas (Recuadro 5.8)
Países Bajos	Plataforma de IA generativa sobre el trabajo de auditoría pública	El Tribunal de Cuentas de los Países Bajos, una plataforma pública de la Dirección General de Inteligencia, está actualmente en fase piloto para permitir a los ciudadanos y otras partes interesadas desplazarse por los informes públicos y encontrar respuestas y fuentes a sus preguntas sobre el trabajo de auditoría pública.	Gestión/accesibilidad de la información mejorada	Lucha contra la corrupción y promoción de la integridad
Reino Unido	Caddy	Caddy, un asistente impulsado por IA desarrollado en el Reino Unido, apoya a los agentes de servicio al cliente recuperando rápidamente información gubernamental precisa. Gracias a su sistema de validación “in the loop”, Caddy garantiza respuestas confiables, al tiempo que mejora la eficiencia en la gestión de las consultas de los ciudadanos.	Gestión/accesibilidad mejorada de la información	Servicios públicos

Rendición de cuentas mejorada y detección de anomalías

En cuanto a los usos para mejorar la rendición de cuentas y la detección de anomalías, el 25 % de los casos de uso analizados se centraron en detectar transacciones inadecuadas y evaluar los riesgos para la integridad, y el 5 % en mejorar la capacidad de los gobiernos para implicar a los actores no gubernamentales y promover la rendición de cuentas (Gráfico 2.6). El primero abarca casos de uso relacionados con la supervisión, los controles preventivos y la evaluación y gestión de riesgos. Por lo general, estos están vinculados a los mandatos fundamentales de algunas funciones específicas del gobierno, como los responsables de luchar contra la corrupción y promover la integridad pública o hacer cumplir la regulación. Los casos de uso que conectan mejor al gobierno con los actores no gubernamentales, contribuyendo en última instancia a una mayor rendición de cuentas y capacidad de respuesta, a menudo están relacionados con la participación cívica y las prácticas de transparencia utilizadas en diversas funciones y organizaciones gubernamentales.

La Tabla 2.5 presenta algunos ejemplos de cómo se está utilizando la IA para lograr este beneficio. Algunos usos se centran en priorizar acciones de precaución basadas en el análisis de patrones y anomalías estadísticas. Por ejemplo, el Tribunal de Cuentas de Portugal utiliza IA para detectar casos críticos y prioritarios en la contratación pública que podrían requerir una concentración de los esfuerzos de auditoría. En Chile, la IA se utiliza en la plataforma de contratación pública del país para detectar irregularidades y mejorar la supervisión del cumplimiento normativo. Otros usos se destinan a detectar lagunas o salvaguardias insuficientes en la formulación de políticas, como la asistencia a los funcionarios encargados de la prevención de la corrupción en la evaluación de la legislación para evaluar los factores de riesgo de corrupción en textos jurídicos de Lituania. La IA también puede respaldar la conexión de los gobiernos con el público para reforzar la rendición de cuentas. Este es el caso de algunas plataformas y herramientas en línea, el asistente virtual Chatico de Bogotá (Colombia), que cuenta con un módulo de gobierno abierto que facilita la participación en campañas públicas y procesos de toma de decisiones.

Tabla 2.5. Ejemplos de IA para mejorar la rendición de cuentas y la detección de anomalías

País	Iniciativa	Descripción	Sub-beneficio	Función
Chile	ChileCompra	El Observatorio de Contratación Pública de ChileCompra utiliza los modelos de aprendizaje permanente para analizar los datos de las licitaciones en busca de irregularidades y mejorar el seguimiento del cumplimiento normativo, lo que permite una supervisión más eficiente y promueve normas éticas en la contratación pública.	Detección de transacciones indebidas y evaluación de riesgos para la integridad	Contratación pública (Recuadro 5.24)
Colombia	Chatico	La ciudad de Bogotá lanzó Chatico, un asistente virtual basado en IA, para facilitar las interacciones entre los ciudadanos y la administración local. A través de su sitio web y sus interfaces de WhatsApp, el chatbot facilita la participación ciudadana en campañas públicas y procesos de toma de decisiones, y ofrece además una mayor accesibilidad a los servicios públicos.	Permitir a los actores no gubernamentales comprender y colaborar con el gobierno y promover la rendición de cuentas	Participación ciudadana (Recuadro 5.41)
Unión Europea	DATA CROS	DATA CROS utiliza IA para detectar anomalías en las estructuras de propiedad de las empresas que puedan indicar corrupción o lavado de dinero. El sistema analiza datos de más de 70 millones de empresas de 44 países europeos, señalando patrones ocultos y posibles actividades ilícitas.	Detección de transacciones indebidas y evaluación de riesgos para la integridad	Lucha contra la corrupción y promoción de la integridad (Recuadro 5.27)
Lituania	IA para ayudar a los funcionarios encargados de prevenir la corrupción	Lituania está desarrollando una herramienta impulsada por IA que utiliza grandes modelos lingüísticos para ayudar a los funcionarios encargados de la prevención de la corrupción a evaluar los factores de riesgo de corrupción en los textos jurídicos, como las lagunas o las salvaguardias insuficientes.	Detección de transacciones indebidas y evaluación de riesgos para la integridad	Lucha contra la corrupción y promoción de la integridad (Recuadro 5.30)
Portugal	Evaluación de los riesgos de contratación pública	El Tribunal de Cuentas de Portugal está aplicando métodos de evaluación de riesgos basados en IA para mejorar sus auditorías, garantizando que se dé prioridad a los casos más críticos en la contratación pública. Esta iniciativa optimiza los recursos y refuerza la rendición de cuentas en la contratación pública.	Detección de transacciones indebidas y evaluación de riesgos para la integridad	Contratación pública

Desbloquear oportunidades para partes interesadas externas a través de la IA como un bien para todos

Por último, una pequeña minoría de casos de uso (4 %) tiene el potencial de desbloquear oportunidades para partes interesadas externas, como ciudadanos, organizaciones de la sociedad civil y empresas. Esas actividades son diferentes de la prestación de servicios generales o de las nuevas formas y canales de interacción a efectos de participación y rendición de cuentas. En este sentido, la IA puede empoderar a agentes externos mejorando sus capacidades y su acceso a la información, así como utilizar sistemas de IA apoyados por el gobierno para lograr sus misiones y objetivos de manera más eficaz. Estos usos siguen siendo marginales y representan una brecha potencial en los esfuerzos gubernamentales que justifica una mayor investigación y acción. Sin embargo, como se ha señalado anteriormente, estos casos de uso podrían ser más frecuentes en áreas no cubiertas por el presente informe.

La Tabla 2.6 presenta algunos ejemplos de cómo se está utilizando la IA para estos fines. Por lo general, los casos de uso incluyen herramientas en plataformas participativas que pueden utilizarse en función de las necesidades y objetivos de los usuarios. Por ejemplo, la plataforma participativa Decide Madrid (España) experimentó con herramientas de IA para ayudar a los ciudadanos a agregar y desarrollar sus propias propuestas de acción. Del mismo modo, la herramienta de IA MAPLE ayuda a los ciudadanos resumiendo los proyectos de textos jurídicos y permitiéndoles presentar insumos y comentarios sobre la legislación pendiente. En el ámbito de la gestión del riesgo de desastres, los gobiernos también pueden contribuir a generar bienes que empoderen a los ciudadanos para una mayor resiliencia ante los desastres naturales. Por ejemplo, el Bot Bencana en Indonesia incita a los residentes a informar sobre las inundaciones a través de las redes sociales, generando mapas en línea en tiempo real que, combinados con datos oficiales, pueden ser reutilizados por partes interesadas no gubernamentales. Aunque los ejemplos que figuran en este documento ayudan a abordar las ventajas relevantes, lo hacen de forma un tanto tangencial, ya que los elementos externos a veces se benefician como un efecto indirecto positivo de las actividades gubernamentales. Podrían resultar oportunidades emergentes para que los gobiernos abran o proporcionen sistemas de IA de forma más directa a partes interesadas externas, por ejemplo, en áreas que actualmente no cuentan con servicios o que no resultan atractivas para el sector privado.

Tabla 2.6. Ejemplos de cómo la IA puede brindar oportunidades a partes interesadas externas como un bien para todos

País	Iniciativa	Descripción	Función
Grecia	DidaktorikaAI	La plataforma griega DidaktorikaAI, puesta en marcha por el Centro Nacional de Documentación (EKT), mejora la accesibilidad del conocimiento académico y científico para los responsables de la formulación de políticas y para la sociedad en general gracias a una biblioteca en línea impulsada por IA que reúne más de 50 000 publicaciones.	Participación ciudadana y gobierno abierto
Indonesia	Bencana Bot	En Yakarta (Indonesia), el chatbot Bencana Bot impulsado por IA incita a los residentes a informar de las inundaciones a través de las redes sociales, generando mapas en línea en tiempo real y de libre acceso en PetaBencana.id. Combinada con datos oficiales proporcionados por la Agencia de Mitigación de Desastres de Yakarta, la plataforma ayuda a los residentes a mantenerse seguros durante emergencias, con un aumento del uso del 2 000 % durante grandes inundaciones.	Aplicación de la ley y gestión del riesgo de desastres
España	Decide Madrid	En 2021, la plataforma participativa Decide Madrid (España), basada en el software de código abierto Consul, experimentó con un sistema de Procesamiento del Lenguaje Natural (NLP) para ayudar a los ciudadanos a agregar y desarrollar propuestas.	Participación ciudadana y gobierno abierto
Estados Unidos	MAPLE	La herramienta impulsada por la IA MAPLE (Plataforma de Massachusetts para la Participación Legislativa), que permite a los ciudadanos comprender mejor el contexto y los objetivos de los proyectos de textos jurídicos a través de resúmenes generados por la IA y presentar sus insumos y comentarios a la legislación pendiente.	Participación ciudadana y gobierno abierto

Algunas funciones gubernamentales son más maduras en relación a la gobernanza y adopción de IA

Se reconoce el potencial de la IA en todas las funciones de la administración, pero la madurez de su adopción y gobernanza varía significativamente. Algunas funciones ya están implementando iniciativas de IA de manera estructurada, aprendiendo de su implementación y, en algunos casos, ampliando las soluciones exitosas a otras áreas problemáticas y en contextos más amplios. Por ejemplo, en el diseño y la prestación de servicios públicos, la IA se utiliza de forma generalizada para automatizar tareas, recuperar y sintetizar información y mejorar la eficacia y la utilidad de los servicios digitales para los usuarios. En el ámbito de la aplicación de la ley y la gestión del riesgo de desastres, la IA se utiliza de forma generalizada para dar prioridad a los recursos policiales, acelerar las investigaciones y anticipar y recuperarse mejor de las catástrofes. Sin embargo, otras funciones gubernamentales sólo han comenzado a experimentar con prototipos y pilotos limitados y ad-hoc. En ámbitos como la evaluación de políticas, la gestión de las finanzas públicas y el diseño y la implementación de la regulación, la adopción de la IA se encuentra principalmente en proyectos piloto aislados. En algunas funciones, como la administración de justicia y la lucha contra la corrupción y la promoción de la integridad pública, existe una variación geográfica significativa. Algunos países, como Argentina (Recuadro 5.62 de Prometeo) y España (Recuadro 5.67 de Respuesta a la Violencia Doméstica Basada en IA) despliegan activamente sofisticadas soluciones de IA incluso en funciones de alto riesgo donde logran mitigar los riesgos que, de lo contrario, daría lugar a escándalos como ha ocurrido con sistemas similares en otros contextos. Sin embargo, otros países todavía se encuentran en las primeras etapas de la transformación digital. Como la IA sigue apareciendo en otros casos de uso, los gobiernos deben mitigar el riesgo para promover una adopción responsable de la IA.

En términos de madurez técnica, la adopción de IA varía no solo en cuanto a su escala, sino también en cuanto a los tipos de sistemas utilizados. Algunas funciones gubernamentales, como la administración tributaria y las adquisiciones, dependen en gran medida de sistemas basados en reglas, que han sido eficaces en la automatización de procesos estructurados para la toma de decisiones durante muchos años. Otros, como la aplicación de la ley y la gestión de desastres, la lucha contra la corrupción y la promoción de la integridad pública, utilizan sistemas de ML más avanzados para identificar patrones, mejorar las evaluaciones de riesgos y apoyar la toma de decisiones. Sin embargo, en la mayoría de las funciones gubernamentales, hay muy poco uso de los últimos modelos de IA generativa, como los LLM, que ofrecen nuevas capacidades en síntesis de conocimiento y generación de contenido que podrían ser más transformacionales. Esta tendencia también puede observarse en otras bases de datos: solo 61 de los 1 343 (4,5 %) casos de uso de IA del repositorio de Tech Watch del sector público de la CE son IA generativa (Brizuela et al., 2025^[16]). Una encuesta de Deloitte (2024^[17]) puso de manifiesto también un nivel desigual de preparación para la adopción de IA en los diferentes sectores de la administración, aunque sus categorías sectoriales no se corresponden directamente con las funciones gubernamentales de este informe. El uso de sistemas de IA generativa puede verse en algunas funciones gubernamentales del Capítulo 5, como el diseño y la implementación de regulaciones y servicios públicos, y en los esfuerzos de participación cívica, aunque muchos de estos esfuerzos parecen esporádicos o experimentales. Algunos de los usos más avanzados adoptan la forma de chatbots, que pueden tener un gran impacto, pero no explotar plenamente el potencial de la tecnología para la síntesis a gran escala, la generación de contenido personalizado o la creación de servicios más proactivos o personalizados. La lenta adopción de estos sistemas avanzados en muchos ámbitos, sumada a la relativamente reciente tecnología y a los casos de uso pertinentes de IA, sugiere que, si bien la experimentación con IA es generalizada, la transición hacia sistemas más sofisticados y de alto impacto sigue siendo desigual entre las funciones gubernamentales y los países. Esto no quiere decir que los gobiernos deban abandonar todos los esfuerzos por utilizar formas más consolidadas de IA en pos de la IA general, ya que perseguir la última "tecnología de moda" en áreas donde no es adecuada para el problema que se quiere resolver puede contribuir al fracaso de los proyectos de IA (Ryseff and Narayanan, 2025^[18]).

Un factor clave que determina la adopción de IA en todas las funciones de la administración es la disponibilidad y calidad de los datos. En la administración tributaria, por ejemplo, la IA se ha implementado ampliamente debido a la abundancia de datos estructurados, lo que ha permitido la automatización y la evaluación de riesgos durante años. Sin embargo, debido a la complejidad del panorama jurídico en esta función y en lo que respecta a los datos de los contribuyentes, la mayoría de los esfuerzos se basan en sistemas clásicos basados en normas, que plantean desafíos a la hora de perseguir sistemas de ML más modernos que puedan generar aumentos de productividad, incluso aprovechando datos no estructurados. Por el contrario, aunque la IA se utiliza cada vez más en las funciones de gestión de recursos humanos del sector privado, su aplicación en la reforma de la función pública sigue siendo limitada debido a la insuficiencia de datos exhaustivos sobre la fuerza laboral, que abarcan las competencias de los empleados, las demandas de empleo y los indicadores de rendimiento. La abundancia de datos en los gobiernos no se traduce necesariamente en la disponibilidad de datos preparados para la IA (véase el Capítulo 3 sobre los desafíos de implementación).

Más allá de los datos, otros ámbitos podrían enfrentarse a desafíos en cuanto a la madurez de la IA porque la naturaleza de sus tareas requiere capacidades técnicas que solo los sistemas de IA avanzados más recientes podrían proporcionar. Esto sugiere la posibilidad de una rápida aceleración en los próximos meses. Por ejemplo, la adopción de IA en el diseño y distribución de la regulación parece haber aumentado y acelerado gracias al uso de LLMs para ayudar en tareas analíticas que otros sistemas no podrían haber realizado (por ejemplo, véase el recuadro 5.12 para consultar las ayudas legales y reglamentarias y las ayudas a la redacción). En otro ejemplo, la IA está posibilitando iniciativas de participación cívica deliberativa en masa a escalas nunca antes viables.

Las diferentes funciones gubernamentales tienen diferentes contextos y necesidades

La adopción de IA en el gobierno suele discutirse en términos generales, pero su impacto y riesgos varían significativamente entre las distintas funciones. Las diferentes funciones presentan desafíos únicos, limitaciones normativas y niveles de preparación para la IA. Por ejemplo, cuando se utiliza la IA en los servicios públicos para mejorar la asistencia sanitaria, las aplicaciones deben regirse por estrictas regulaciones de privacidad de datos y preocupaciones éticas en torno a la toma de decisiones médicas, mientras que la IA en otros ámbitos, como la participación cívica, puede ser más experimental, utilizando la optimización en tiempo real con menos posibilidades de infringir los derechos de las personas. Los actores de una función de gobierno podrían usar una solución de IA similar a la de otra, con resultados, impactos e implicaciones enormemente diferentes. Por lo tanto, la discusión sobre la función del gobierno en esta sección no puede verse como una comparación similar, y se necesitan más investigaciones para comprender las diferencias, incluso con un alcance de análisis más amplio.

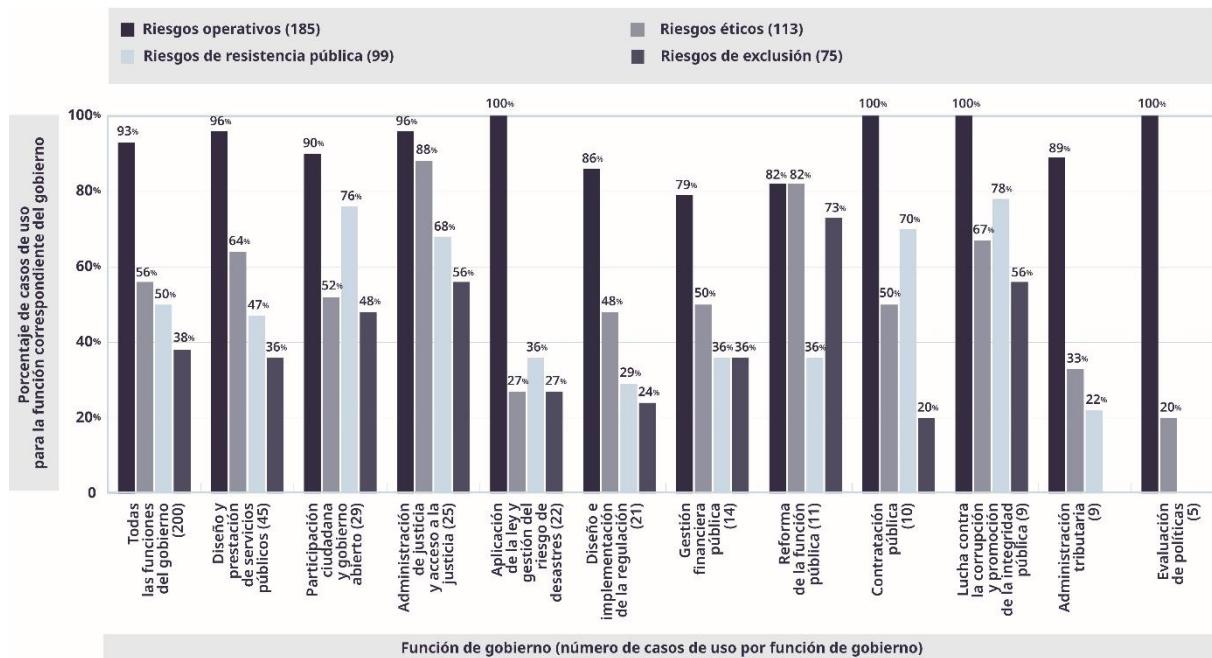
Los casos de uso podrían plantear riesgos si no se implementan de manera confiable

Como se analiza en el Capítulo 1, este informe clasifica cinco tipos de riesgos a los que se enfrentan los gobiernos al adoptar la IA. Además de los riesgos que se muestran en el Gráfico 2.7, el riesgo de inacción implica oportunidades perdidas y la creciente brecha de capacidad entre el sector público y privado. Este riesgo, que posiblemente se materialice al *no desarrollar* capacidades para la IA y utilizarla en el gobierno, no se puede visualizar y puede ser difícil o imposible de medir con precisión.

La OCDE concluyó que cada uno de los 200 casos de uso analizados para este informe podría entrañar uno o más tipos de riesgo si no se diseñan y utilizan de forma confiable.

Gráfico 2.7. El potencial de riesgos operativos es el más representado en todas las funciones gubernamentales

Número de casos de uso en distintas funciones gubernamentales, clasificados según los tipos de riesgo seleccionados



Nota: Entre paréntesis, el número de incidencias de los tipos de riesgo. Los casos de uso pueden implicar más de un tipo de riesgo. Por lo tanto, el número total de posibles casos de riesgo es mayor que el número total de casos de uso.

Fuente: Análisis de la OCDE de los casos de uso identificados.

Los riesgos operacionales potenciales son los más prevalentes entre los casos analizados (93 %). Un ejemplo de ello es el plan Robodebt de Australia, en el que las investigaciones revelaron que las insuficiencias en el diseño algorítmico desempeñaron un papel importante en sus fallos y que, en última instancia, sus cálculos se consideraron ilícitos (Recuadro 5.11). En concreto, la simplificación excesiva del algoritmo y la falta de salvaguardas dieron lugar a la emisión de 470 000 notificaciones de deuda incorrectas sin verificación humana. Esto pone de relieve los riesgos operacionales de automatizar sistemas sociales complejos sin una supervisión humana suficiente o pruebas rigurosas y, además, presentaba riesgos éticos que daban lugar a daños en el mundo real⁴.

Los riesgos éticos potenciales fueron los segundos más frecuentes entre los casos de uso analizados (56 %), como en las herramientas de revisión de solicitantes apoyadas por IA utilizadas por las oficinas de recursos humanos. En realidad, los riesgos éticos pueden estar presentes en la gran mayoría de los casos de uso de la IA. Sin embargo, la menor presencia de este riesgo en los casos de uso analizados podría deberse al alcance limitado y específico, y a la forma de aplicación de los casos de uso identificados para este informe. Un ejemplo de riesgo ético que provocó daños en la vida real es el Toeslagenaffaire (escándalo de prestaciones por cuidado infantil) de los Países Bajos, en el que un sistema de IA acusó injustamente a 26 000 familias de solicitar fraudulentamente prestaciones por cuidado infantil debido a un algoritmo sesgado que se dirigía a familias con doble nacionalidad u origen inmigrante; obligó a muchas a pagar deudas indebidas (Recuadro 5.6). Este caso ilustra cómo los riesgos éticos pueden causar daños si no se mitigan adecuadamente.

Los posibles riesgos de resistencia pública fueron frecuentes en el 50 % de los casos de uso analizados. Los fallos previos en el despliegue de IA han afectado significativamente a la reputación y

erosionado la confianza pública en la capacidad del gobierno para utilizar la IA de forma responsable. Estos casos ponen de relieve la necesidad de que los gobiernos tomen medidas para prevenir riesgos y abordar rápidamente posibles fallos en el uso de IA, fomentando la confianza pública. Esto puede lograrse con barreras de seguridad adecuadas, incluidos mecanismos sólidos de rendición de cuentas y reparación, una supervisión y un control continuo, y una gestión eficaz de riesgos.

Finalmente, se **identificó** el potencial de **riesgos de exclusión en el 38 % de los casos analizados**. Por ejemplo, las plataformas de participación ciudadana que han desplegado herramientas de IA para ayudar a los ciudadanos a agregar y desarrollar propuestas pueden plantear desafíos para las personas que carecen de competencias digitales. Esto podría beneficiar involuntariamente a personas ya favorecidas, mejorando así su capacidad para promover aún más sus ideas (Duberry et al., 2021^[19]; Wang et al., 2024^[20]). Los casos de uso de IA también ayudan a los funcionarios a procesar de manera eficiente grandes cantidades de aportes ciudadanos y a facilitar mejor los procesos participativos. Sin embargo, el éxito de estas aplicaciones depende de la diversidad incorporada en los datos de formación del sistema de IA empleado; existe el riesgo de que estas herramientas no captén adecuadamente la diversidad de la opinión pública (ECNL, 2024^[21]). Muchas herramientas de traducción asistida por IA en las plataformas de participación también pueden estar sujetas a no captar los matices y comprender los diferentes contextos culturales para las lenguas de las minorías, ya que generalmente se entran con datos del inglés y otras lenguas dominantes (ECNL, 2024^[21])⁵.

También es de interés la mayor o menor presencia de ciertos riesgos en las funciones gubernamentales. Por ejemplo, los casos observados en la reforma de la administración de justicia y la administración pública parecen presentar de manera más prominente riesgos éticos debido a la posibilidad de resultados negativos en algunos de sus casos. En el ámbito de la administración tributaria, los controles y las regulaciones existentes podrían crear salvaguardias frente a los riesgos éticos y de exclusión, haciendo que los riesgos operativos sean más frecuentes. En la participación ciudadana, los riesgos de resistencia pública adquieren mayor prevalencia, en comparación con otras funciones. Esto es probable porque la mayoría de sus casos de uso incluyen interacciones entre gobiernos y ciudadanos y, por lo tanto, exhiben una mayor dependencia de la aceptación pública. Esto es similar en el caso de la contratación pública, donde la percepción de los proveedores y la confianza pública en los sistemas de IA desempeñan un papel clave en su éxito, por lo que suelen figurar más en los riesgos de resistencia. Este análisis comparativo muestra la importancia de reconocer los principales factores que impulsan los riesgos potenciales en cada campo, lo que puede servir de base para mitigarlos y gestionarlos.

Los gobiernos están atentos a varios riesgos de la IA, aunque algunos pueden pasarse por alto

En general, al realizar análisis e interactuar con los gobiernos en la elaboración de los Capítulos 4 y 5, parece que muchos gobiernos nacionales están bien informados y han implementado procesos para gestionar el riesgo asociado a los **datos, la falta de transparencia y explicabilidad, y el uso indebido de la IA** —ya sea intencional o involuntario— y la posibilidad de que se produzcan daños o infracciones a la privacidad. Esto es positivo y no es sorprendente, ya que los expertos en IA y la investigación dentro y fuera del sector público suelen plantear estos riesgos (OECD, 2024^[13])⁶. En cierta medida, el riesgo de una dependencia excesiva de las tecnologías de IA también parece tenerse en cuenta en los esfuerzos de los gobiernos. En general, sin embargo, parece haber menos énfasis en garantizar que el uso gubernamental de la IA no agrave aún más las brechas digitales. Cuando no se dispone de canales de servicio complementarios, las ambiciones gubernamentales de automatizar y agilizar los procesos podrían dar lugar a una reducción de las oportunidades de servicio para las comunidades con menos acceso a los servicios digitales o con preferencia por los enfoques no digitales (Welby and Hui Yan Tan, 2022^[22]). Además, si bien los gobiernos son claramente conscientes de cómo la IA podría contribuir a la productividad y ayudar a cambiar los esfuerzos de los servidores públicos de tareas repetitivas a trabajos

más significativos, aparentemente hay menos reconocimiento del potencial de la IA para reducir la calidad del empleo (por ejemplo, a través de una gestión algorítmica invasiva) o para el desplazamiento laboral (Peixoto, Canuto and Jordan, 2024^[23]). Incluso si la adopción de la IA se vuelve sistemática en las administraciones públicas, los gobiernos tendrán que garantizar que todos los ciudadanos reciban un servicio adecuado y que se tengan en cuenta las preocupaciones y los derechos de los funcionarios públicos, con el fin de promover la adopción de la IA.

Por último, es necesario que los gobiernos consideren mejor el **riesgo de la inacción**. A lo largo de los Capítulos 4 y 5, está claro que muchos gobiernos están trabajando en pos de objetivos específicos con sus esfuerzos en el ámbito de la IA, y que son conscientes de los diversos riesgos de la IA y están tratando de mitigarlos de diferentes maneras para desbloquear el potencial de la IA. Sin embargo, el análisis realizado para el presente informe y las conversaciones con los funcionarios gubernamentales pertinentes indican que puede haber un escaso conocimiento de las oportunidades perdidas debido a la lenta adopción de IA o a las consecuencias del aumento de la brecha en las capacidades de IA entre los sectores público y privado.

Pocos gobiernos han evaluado en qué medida la IA podría tener un impacto en sus operaciones internas y en sus servicios públicos. Además, la mayoría no ha articulado plenamente sus ambiciones en materia de IA en el gobierno, no ha determinado las deficiencias existentes ni ha propuesto hojas de ruta aclaradas para colmar esas deficiencias y alcanzar los objetivos. Los gobiernos deben explorar la IA no solo para mejorar el diseño y la aplicación de las políticas y los servicios públicos, sino también para garantizar que disponen de los conocimientos y la capacidad necesarios para regular el desarrollo y el uso de la IA en la administración y fuera de ella. En una investigación del Parlamento de Australia (2025^[24]) el Comité Mixto de Cuentas Públicas y Auditoría expresó su "grave preocupación" por el hecho de que la IA pronto supere la capacidad del gobierno para regularla. Estas conclusiones del comité no son necesariamente representativas de las opiniones del gobierno australiano. Los expertos en IA también han señalado que la incapacidad de los mecanismos de gobernanza y las instituciones para mantenerse al día con la rápida evolución de la IA es uno de los riesgos más críticos asociados a esta tecnología (OECD, 2024^[13]).

Además de estos riesgos críticos asociados al uso de la IA en la administración pública, el análisis de la OCDE de 200 casos de uso ha identificado una serie de desafíos a los que se pueden enfrentar los gobiernos al adoptar la tecnología. Estos desafíos de implementación son comunes a todas las funciones, mientras que otros son más frecuentes en ciertos campos. Pueden traducirse en cuestiones más amplias y obstaculizar el uso estratégico de la IA en el gobierno. Estas cuestiones se tratan en el capítulo siguiente.

Bibliografía

- Austin, T. et al. (2024), *A snapshot of how public sector leaders feel about generative AI*, [17] <https://www2.deloitte.com/us/en/insights/industry/public-sector/ai-adoption-in-public-sector.html>.
- Bailey, K. (2016), *Reframing the “AI Effect”*, <https://medium.com/@katherinebailey/reframing-the-ai-effect-c445f87ea98b>. [12]
- Berryhill, J. et al. (2019), “Hello, World: Artificial intelligence and its use in the public sector”, *OECD Working Papers on Public Governance*, No. 36, OECD Publishing, Paris, <https://doi.org/10.1787/726fd39d-en>. [15]
- Brizuela, A. et al. (2025), *Analysis of the generative AI landscape in the European public sector*, [16] European Commission, <https://op.europa.eu/s/z4XY>.

- Columbia University (2020), *The Future of AI in the Brazilian Judicial System*, [9]
<https://www.sipa.columbia.edu/aidriven-innovations-brazilian-judiciary>.
- Duberry, J. et al. (2021), “Artificial intelligence and civil society participation in policy-making processes: Thinking about AI and participation.”, *SSRN Electronic Journal*, [19]
<https://doi.org/10.2139/ssrn.3817666>.
- EC (2025), *Public Sector Tech Watch latest dataset of selected cases*, [10]
<http://data.europa.eu/89h/e8e7bdd-8510-4936-9fa6-7e1b399cbd92> (accessed on 4 April 2025).
- ECNL (2024), *Can AI tools and platforms make public engagement truly meaningful and inclusive?*, <https://ecnl.org/news/ai-public-participation-hope-or-hype> (accessed on 17 March 2025). [21]
- Government of Canada (2025), *AI Strategy for the Federal Public Service*, [4]
<https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/gc-ai-strategy-overview.html>.
- Government of Switzerland (2025), *Strategy Use of AI systems in the Federal Administration*, [5]
<https://www.bk.admin.ch/bk/en/home/digitale-transformation-ikt-lenkung/ikt-vorgaben/strategien-teilstrategien/sb021-strategie-einsatz-von-ki-systemen-in-der-bundesverwaltung.html>.
- Government of Uruguay (2021), *AI Strategy for the Digital Government*, [6]
<https://www.gub.uy/agencia-gobierno-electronico-sociedad-informacion-conocimiento/comunicacion/publicaciones/ia-strategy-english-version/ia-strategy-english-version/ai-strategy-for>.
- Harvard Kennedy School (2023), *AI, judges and judgement: setting the scene*, [8]
<https://www.hks.harvard.edu/centers/mrcbg/publications/awp/awp220>.
- Muñoz-Cadena, S. et al. (2025), *Sistemas de IA en el sector público de América Latina y el Caribe (Versión V2)*, <https://sistemaspublicos.tech/sistemas-de-ia-en-america-latina/> (accessed on 29 April 2025). [11]
- Netherlands Court of Audit (2024), *Central government often does not assess risks of AI*, [26]
<https://english.rekenkamer.nl/latest/news/2024/10/16/central-government-often-does-not-assess-risks-of-ai>.
- OECD (2024), “2023 OECD Digital Government Index: Results and key findings”, *OECD Public Governance Policy Papers*, No. 44, OECD Publishing, Paris, [7]
<https://doi.org/10.1787/1a89ed5e-en>.
- OECD (2024), “Assessing potential future artificial intelligence risks, benefits and policy imperatives”, *OECD Artificial Intelligence Papers*, No. 27, OECD Publishing, Paris, [13]
<https://doi.org/10.1787/3f4e3dfb-en>.
- OECD (2024), “Governing with Artificial Intelligence: Are governments ready?”, *OECD Artificial Intelligence Papers*, No. 20, OECD Publishing, Paris, <https://doi.org/10.1787/26324bc2-en>. [1]
- OECD (2023), “The state of implementation of the OECD AI Principles four years on”, *OECD Artificial Intelligence Papers*, No. 3, OECD Publishing, Paris, [2]
<https://doi.org/10.1787/835641c9-en>.

- OECD (2021), “State of implementation of the OECD AI Principles: Insights from national AI policies”, *OECD Digital Economy Papers*, No. 311, OECD Publishing, Paris, [3]
<https://doi.org/10.1787/1cd40c44-en>.
- Parliament of Australia (2025), *Report 510: Inquiry into the use and governance of artificial intelligence systems by public sector entities - 'Proceed with Caution'*, [24]
https://parlinfo.aph.gov.au/parlInfo/download/committees/reportJnt/RB000567/toc_pdf/Report510Inquiryintotheuseandgovernanceofartificialintelligencesystemsbypublicsectorentities-'ProceedwithCaution'.pdf.
- Peixoto, T., O. Canuto and L. Jordan (2024), “AI and the Future of Government: Unexpected Effects and Critical Challenges”, *Policy briefs on Economic Trends and Policies*, Vol. 2408, [23]
https://ideas.repec.org/p/ocp/pbecon/pb_10-24.html.
- Ryseff, J. and A. Narayanan (2025), *Why AI Projects Fail*, [18]
<https://www.rand.org/pubs/presentations/PTA2680-1.html>.
- The Alan Turing Institute (2024), *AI for bureaucratic productivity: Measuring the potential of AI to help automate 143 million UK government transactions*, [14]
<https://www.turing.ac.uk/news/publications/ai-bureaucratic-productivity-measuring-potential-ai-help-automate-143-million-uk>.
- Wang, C. et al. (2024), “The artificial intelligence divide: Who is the most vulnerable?”, *New Media & Society*, [20]
<https://doi.org/10.1177/14614448241232345>.
- Welby, B. and E. Hui Yan Tan (2022), “Designing and delivering public services in the digital age”, *OECD Going Digital Toolkit Notes*, No. 22, OECD Publishing, Paris, [22]
<https://doi.org/10.1787/e056ef99-en>.
- Yigitcanlar, T. et al. (2024), *Local governments are using AI without clear rules or policies, and the public has no idea*, <https://theconversation.com/local-governments-are-using-ai-without-clear-rules-or-policies-and-the-public-has-no-idea-244647>. [25]

Notas

¹ No todos los casos de uso analizados aparecen en este informe. La OCDE seleccionó los casos para el informe que mejor ilustraban los diversos temas y conclusiones presentados.

² Véase <https://oecd-opsi.org/innovation-tag/artificial-intelligence-ai> y https://oecd.ai/dashboards/policy-instruments/AI_use_cases_in_the_public_sector, respectivamente. La colección del OPSI incluyó un ejercicio global de crowdsourcing abierto llamado a la innovación centrado en las innovaciones en los servicios públicos en 2024.

³ Según se clasifica en la *Clasificación de las Funciones de Gobierno* del observatorio. Véase <https://ec.europa.eu/eurostat/statistics->

explained/index.php?title=Glossary%3AClassification_of_the_functions_of_government_%28COFOG%29.

⁴ El esquema Robodebt aprovechó la coincidencia automatizada de datos, el promedio de ingresos y el cálculo de pagos en exceso. Como se expone en el Recuadro 1.1 de este informe, muchos sostienen que estos sistemas no deben considerarse IA en absoluto. Por lo tanto, el esquema Robodebt podría describirse mejor como un sistema automatizado de toma de decisiones. Sin embargo, ayuda a ilustrar problemas de gobernanza, supervisión ética y diseño algorítmico.

⁵ Véase también el análisis relacionado con los riesgos de exclusión en el Capítulo 1.

⁶ Si bien es posible que se informe a los gobiernos sobre estas y otras cuestiones, y como se indica en el Capítulo 4, muchos países cuentan con marcos y procesos para mitigar los riesgos, deben tomar medidas y cumplir con sus marcos y procesos para gestionar los riesgos. Hay casos en los que los gobiernos pueden no cumplir plenamente los requisitos internos o pueden enfrentarse a incentivos para considerar los sistemas de bajo riesgo, reduciendo así los requisitos de rendición de cuentas (Netherlands Court of Audit, 2024^[26]). Además, algunas investigaciones sugieren que los gobiernos locales pueden no estar tan informados o ser tan propensos a llevar a cabo actividades de mitigación de riesgos como los gobiernos nacionales (Yigitcanlar et al., 2024^[25]). Por último, este informe no pretende evaluar la calidad y eficacia de los procesos y mecanismos gubernamentales, aunque sí pretende poner de relieve los que parecen sólidos y están surgiendo como mejores prácticas.

3

Desafíos de implementación que dificultan el uso estratégico de la IA en la administración pública

Este capítulo examina los desafíos observados en los casos de uso analizados y en investigaciones y análisis más amplios. Concluye que muchas iniciativas gubernamentales en materia de IA siguen siendo proyectos piloto. Destaca las barreras sistémicas comunes: la falta de competencias, las dificultades para acceder y compartir datos de alta calidad, la escasa orientación práctica, la aversión al riesgo y la débil medición de resultados y rendimiento de la inversión. Otros desafíos varían según la función, como la regulación inflexible o ambigua, los costos elevados o inciertos y los sistemas informáticos obsoletos.

Mensajes clave

- El examen de la OCDE de los casos de uso de la IA en la administración pública indica una elevada presencia de iniciativas en fase inicial, como experimentos y proyectos piloto. Esto indica:
 - Posibles desafíos en la transición de la experimentación a la implementación
 - Necesidad de mayor monitoreo e intercambio de información
 - Posible necesidad de medidas de política pública para fomentar la implementación y la escalabilidad.
- Los desafíos en la implementación que comparten todas las funciones gubernamentales son:
 - Brechas en habilidades
 - Dificultades para obtener o compartir datos de calidad
 - Falta de marcos viables y de directrices sobre el uso de IA, especialmente en funciones gubernamentales específicas
 - Aversión al riesgo
 - Demostrar resultados y rendimiento de la inversión (RSI).
- Las funciones gubernamentales se enfrentan a una variedad de otros desafíos, algunos más frecuentes en algunas funciones que en otras:
 - Entornos legales y regulatorios inflexibles u obsoletos
 - Costos elevados o inciertos de la adopción y escalado de la IA
 - Sistemas de tecnología de la información (TI) obsoletos.

El sector público enfrenta diversos desafíos de implementación en la adopción de la IA. Algunos abarcan todas las funciones gubernamentales, mientras que otros parecen más marcados en determinadas funciones. Por ejemplo, en casi todas las funciones existen carencias en materia de competencias y dificultades para acceder y compartir datos de calidad. La implementación de la IA debe tener en cuenta distintos entornos regulatorios en diferentes funciones, y los requisitos de cumplimiento varían significativamente entre funciones, como la contratación pública, la aplicación de la ley y la administración tributaria. Además, algunas funciones se enfrentan a más desafíos que otras en cuanto a garantizar el financiamiento o con sistemas obsoletos, ya que solo ahora estamos asistiendo a la aparición de infraestructuras digitales que comparten servicios con una mayor interoperabilidad e integración. Estos desafíos pueden traducirse en problemas más amplios, como las dificultades para ampliar el alcance de las soluciones y la aversión al riesgo que dificulta la innovación. En este capítulo se analizan los diversos desafíos en materia de implementación a los que se enfrentan las funciones gubernamentales, tal y como se analiza con más detalle en relación con cada función en el Capítulo 5. Muchos de estos desafíos reflejan el análisis regional específico sobre IA en el gobierno (OECD/CAF, 2022^[1]; Brizuela et al., 2025^[2]).

La mayoría de las iniciativas gubernamentales en materia de IA se desarrollan en fases exploratorias o piloto, con una escalada y documentación limitadas

Possibles dificultades para pasar de la experimentación a la aplicación

La revisión de la OCDE sobre los usos de IA en la administración pública indica una elevada presencia de iniciativas en fase inicial, como experimentos y proyectos piloto. Esto concuerda con los debates

mantenidos entre los grupos de trabajo y las redes pertinentes de la OCDE, en los que servidores públicos de ámbito nacional y local informan que los gobiernos se encuentran en las primeras fases de utilización de IA, buscando aprender a través de iniciativas a pequeña escala y probando diferentes enfoques.

En general, este enfoque es bueno. La OCDE lleva mucho tiempo alentando a los gobiernos tanto a nivel nacional como local a experimentar con nuevos enfoques de manera controlada e iterativa para minimizar los riesgos y costos y promover que los fracasos, que son inevitables, se produzcan rápidamente y generen lecciones aprendidas para los esfuerzos futuros (OECD, 2017^[3]; 2024^[4]). Esto es fundamental, sobre todo cuando se acostumbran por primera vez al uso de IA, ya que algunas estimaciones sugieren que más del 80 % de los proyectos de IA fracasan, el doble de la tasa de proyectos no relacionados con IA (Ryseff, De Bruhl and Newberry, 2024^[5]). Por ejemplo, para que la adopción de la IA tenga éxito, puede ser necesario incorporar IA en áreas y procesos de bajo riesgo y utilizar datos generados internamente o abiertos para demostrar valor y obtener resultados rápidamente. Un ejemplo de este caso es el Centro de Servicios Compartidos del Gobierno de Finlandia para Finanzas y Recursos Humanos (Palkeet), que comenzó con aplicaciones modestas de RPA, allanando el camino hacia la "hiperautomatización" con aprendizaje automático (ML) (véase el Capítulo 5, "IA en la gestión financiera pública").

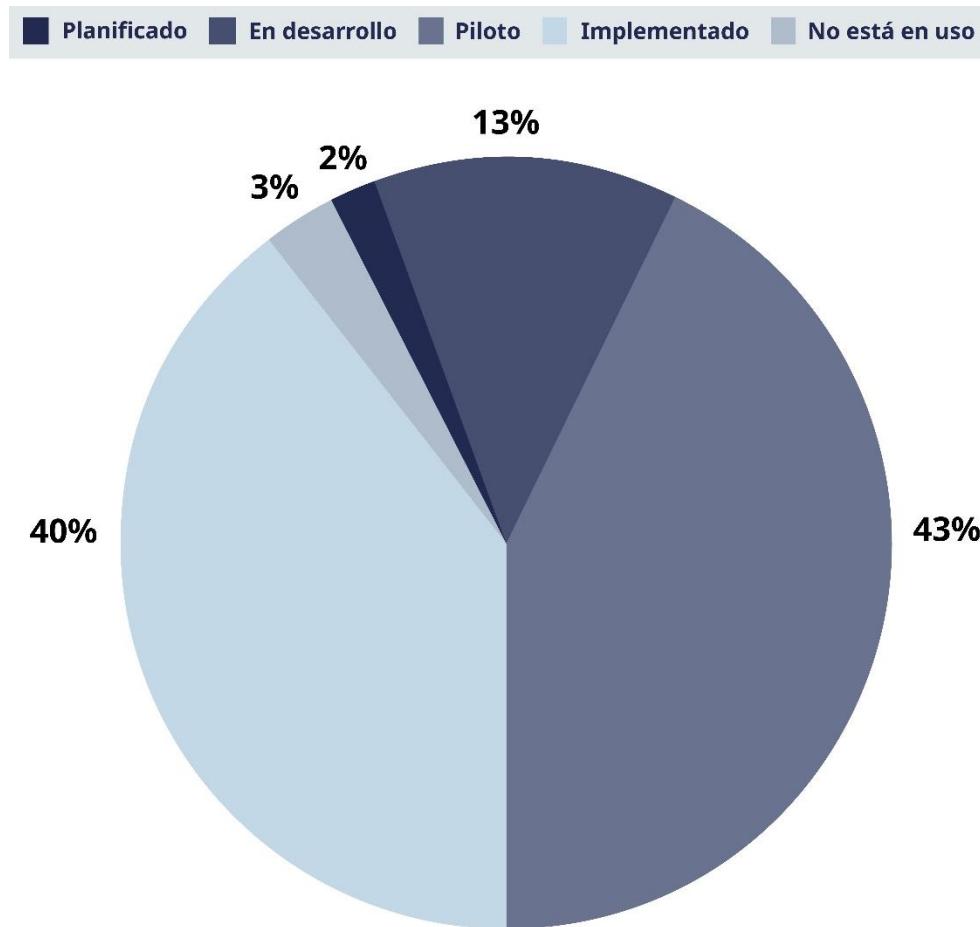
Sin embargo, el objetivo final de la mayoría de los proyectos de IA es la eventual implementación y, según corresponda, la ampliación de soluciones exitosas. En muchos países y funciones de la administración, los casos de uso de IA se encuentran en gran medida en la fase exploratoria (por ejemplo, pruebas de concepto o proyectos piloto) y todavía no se han implementado de forma más amplia ni se han ampliado más allá de su uso limitado. Por ejemplo, en el Reino Unido, uno de los gobiernos más avanzados en lo que respecta al uso de IA, un informe de la Comisión de Cuentas Públicas (PAC) del Parlamento (2025^[6]) concluyó que el Gobierno "no cuenta con un mecanismo sistemático para reunir las enseñanzas extraídas de los proyectos piloto y que hay pocos ejemplos de adopción exitosa a gran escala en todo el Gobierno". Los resultados de una encuesta de Deloitte (2024^[7]) realizada en 14 países refuerzan los desafíos relacionados con la ampliación de la GenAI, señalando también que este desafío no es exclusivo de los gobiernos. Según la encuesta, "un número significativo de encuestados tanto comerciales como gubernamentales han trasladado menos de un tercio de sus experimentos de GenAI a la producción completa". Esto se atribuye a otros desafíos analizados en este capítulo: la falta de conocimientos técnicos y la dificultad para medir el valor de la GenAI.

Las fuentes de datos existentes ofrecen información sobre el estado actual de la adopción de IA en el gobierno. Por ejemplo, los casos de uso analizados para este informe muestran que la mayoría han ido más allá de los pequeños proyectos piloto para aplicarse de alguna manera de forma más completa. De estos casos implementados, la mayoría no se han ampliado más allá de sus contextos originales (por ejemplo, en determinadas oficinas o para determinados procesos) para abordar otras necesidades. Esto no siempre es un objetivo p los usos de la IA en algunas funciones gubernamentales a menudo no son apropiados para otros. Sin embargo, los casos de uso de IA en el diseño y la prestación de servicios públicos que se exponen en el Capítulo 5 demuestran que, en efecto, es posible ampliar el alcance de enfoques que han tenido éxito. En este informe, la preponderancia de casos implementados que van más allá de los programas piloto se debe en parte a la tendencia de la OCDE a seleccionar más casos de uso implementados porque hay más información pública disponible sobre ellos, y es más probable que los gobiernos informen sobre ellos como parte de los ejercicios de recopilación de datos de la OCDE.

Otras fuentes de datos ofrecen perspectivas complementarias. El Observatorio de Tech Watch del Sector Público de la CE ha recopilado sistemáticamente casos de usos de IA durante varios años. Sus datos sobre casi 1 500 casos de uso de IA indican que la mayoría de las soluciones de IA se encuentran todavía en la fase prevista, piloto o en desarrollo (58 %), lo que sugiere que en el sector público de la UE la mayoría de los casos siguen siendo experimentales o no se han aplicado plenamente (Gráfico 3.1). Aunque el paso de los proyectos piloto a la producción parece ser un desafío —como se ha visto reforzado en los debates de la OCDE con los gobiernos—, la proporción de proyectos en fase de implementación

ha aumentado en las últimas recopilaciones de datos. Esto sugiere que las administraciones podrían estar transitando sus iniciativas de las pruebas iniciales a una implementación más completa (EC, 2024^[8]). Estos datos no tienen en cuenta en qué medida los proyectos ejecutados han ido más allá de su contexto inicial.

Gráfico 3.1. La mayoría de los casos de uso de IA en la Unión Europea (UE) se encuentran en fases piloto o de desarrollo



Fuente: Datos analizados por la OCDE de (EC, 2024^[9]).

La base de datos "Sistemas de IA en el sector público de América Latina y el Caribe" indica una proporción mucho mayor de casos de uso implementados (70 %). (Muñoz-Cadena et al., 2025^[10]). Sin embargo, es probable que esto se deba a la recopilación de datos para esta base de datos, que fue elaborada por investigadores utilizando información de dominio público. Es menos probable que la información de dominio público incluya detalles sobre casos de uso planificados o pilotados en comparación con las comunicaciones de los gobiernos, que en gran medida sirvieron de base para las bases de datos de la OCDE y la UE.

En general, de acuerdo con el trabajo de la OCDE sobre IA en los gobiernos desde 2019, incluso directamente con muchos gobiernos, parece que el estado esperado de implementación debería ser mayor en relación con las primeras etapas de las pruebas. La elevada presencia de casos en etapas iniciales se explicaría fácilmente si los casos de uso de los gobiernos hicieran especial hincapié en aprovechar las últimas tecnologías de IA, como los modelos de bases de IA generativa. Sin embargo, las

investigaciones de la OCDE para este informe indican que, en gran medida, no es el caso; los gobiernos todavía tienden a adoptar enfoques más duraderos. Estas observaciones sugieren que es necesario seguir trabajando en este ámbito para comprender mejor los caminos que toman los gobiernos de la planificación a la implementación, así como para extraer lecciones aprendidas más específicas y factores para el éxito.

Tal como se analiza en el Capítulo 2, puede que algunas aplicaciones estrechas pero tradicionales de la IA —que es más probable que se implementen o incluso se amplíen— se hayan integrado o sean tan habituales que ya no den lugar a informes externos ni a una respuesta a los esfuerzos de recopilación de datos. Esto podría sesgar las cifras por la infrarrepresentación de las iniciativas implementadas que no dependen de sistemas y enfoques de IA más nuevos.

La falta de evidencia sobre la continuidad exige mayor monitoreo

Se necesita información adicional para entender mejor el estado y evolución de los casos de uso de IA gubernamental. Investigaciones futuras deberían indagar en esto más a fondo monitoreando el progreso del desarrollo de soluciones a lo largo del tiempo, lo que podría generar lecciones tanto de éxitos como de fracasos. La OCDE ha intentado explorar los avances conseguidos en los casos de uso de IA pertinentes analizados en productos anteriores; sin embargo, por lo general, hay pocos informes actualizados sobre los resultados de los proyectos piloto o el estado de los casos de uso implementados, lo que dificulta su seguimiento longitudinal. La mayor parte de la investigación realizada para determinar si un caso de uso es operativo se basa en la capacidad de acceder a un producto funcional (si es público), o la presencia de comunicados de prensa del gobierno, cobertura de noticias, artículos de blog o presentaciones públicas. Otra fuente son los casos de uso compartidos a través de informes públicos periódicos o archivos oficiales, incluidos los de la OCDE. Los inventarios o catálogos actuales son a menudo estáticos; representan proyectos como una fotografía en el tiempo, sin proporcionar información sobre su desarrollo y evolución, y a menudo no se actualizan. La recopilación de datos primarios por parte de los investigadores (es decir, realizar encuestas directamente a organizaciones gubernamentales para identificar nuevos esfuerzos y obtener actualizaciones y lecciones aprendidas sobre iniciativas conocidas) sería útil para futuras investigaciones, aunque requiere más recursos y tiempo.

Estas dificultades para acceder a información actualizada sobre casos de uso individuales ponen de relieve la importancia de que los gobiernos supervisen los casos de uso de IA, junto con un intercambio sistemático y regular de información, lo que reforzaría aún más los Principios de la OCDE en materia de transparencia y rendición de cuentas en el ámbito de IA. Esto no solo sería valioso para audiencias externas; documentar y difundir métodos y casos de uso exitosos (y no exitosos) puede ayudar a las organizaciones gubernamentales a replicar y escalar proyectos de IA de manera más efectiva. Este enfoque ayuda a evitar errores comunes, contribuye a garantizar la coherencia y acelera la adopción de tecnologías de IA en diversas entidades gubernamentales (OECD/UNESCO, 2024^[11]). La supervisión inadecuada o inexistente también puede afectar futuros casos de uso, ya que las innovaciones en IA potencialmente eficaces pueden pasarse por alto o, por el contrario, los enfoques refutados pueden ampliarse de forma inadecuada. En el Capítulo 4, "Promoción de la transparencia en la forma en que el gobierno utiliza la IA", se ofrecen más debates y ejemplos de cómo lo están haciendo algunos gobiernos.

Podría ser necesario adoptar medidas de política para fomentar la implementación y escalabilidad

El hecho de que muchas iniciativas de IA se encuentren en las fases de planificación y piloto, o no estén claras en cuanto a su progreso, sugiere que los gobiernos deben mejorar sus capacidades de implementación para avanzar en proyectos más allá de las etapas de prueba iniciales, garantizar un despliegue exitoso y mantener el impacto a largo plazo (EC, 2024^[8]). Para ello, es necesario establecer elementos básicos, como garantizar el acceso a conjuntos de datos, recursos informáticos y la experiencia

necesaria para desarrollar y ampliar proyectos de IA. Estos factores se analizan en profundidad en el Capítulo 4. También es necesario superar otras dificultades de implementación, como se analiza más adelante.

Desafíos comunes a las funciones centrales del gobierno

Brecha de competencias, el desafío más común

Una reciente encuesta realizada por Salesforce en cinco países (2024^[12]) señaló que la principal barrera para la adopción gubernamental de IA era la brecha de competencias internas para utilizarla, y el 60 % de los encuestados del sector público se refirieron a este desafío¹. Los encuestados del sector público tenían un tercio más de probabilidades de indicar una brecha de competencias en su organización en comparación con el promedio de la industria. Las revisiones a nivel nacional arrojan resultados comparables: el 70 % de los organismos gubernamentales del Reino Unido afirman que las competencias son un obstáculo para la adopción de IA (UK NAO, 2024^[13]). En un estudio sindical nacional realizado entre agosto y octubre de 2024 a 2000 empleados de la Función Pública Australiana (APS), el 92 % dijeron que no habían recibido formación sobre el uso de IA, y solo el 16 % dijo que se sentían equipados para usar la tecnología². Desde octubre de 2024, el Gobierno australiano ha publicado y puesto a disposición de todos los empleados de la APS un módulo de formación sobre los fundamentos de la IA en la administración pública y una serie de sesiones magistrales sobre IA impartidas por profesionales. El Gobierno australiano cuenta con una serie de programas para mejorar las capacidades en materia de IA. La iniciativa AI CoLab proporciona un marco para la colaboración intersectorial, el diseño conjunto y la celebración de eventos periódicos. El acceso a una herramienta gubernamental de IA a través de una prueba beta cerrada de GovAI comenzó el 5 de mayo de 2025 y se amplió a todos los empleados de la APS el 31 de agosto de 2025.

Las brechas en materia de competencias también suponen un reto importante cuando se centran específicamente en los gobiernos subnacionales (UN Habitat, 2024^[14]). Por ejemplo, las encuestas realizadas en Estados Unidos (EE.UU.) a los directores de sistemas estatales y a los ejecutivos de TI locales revelaron que solo el 20 % de los directores de sistemas estatales y el 25 % de los encuestados locales confiaban incluso ligeramente en que su personal de tecnología contaba con los conocimientos técnicos necesarios para responder a la llegada de la IA generativa (NASCIO, 2024^[15]; PTI, 2024^[16]). Más allá del desafío comúnmente discutido de los gobiernos, que compiten con el sector privado por el talento, en algunos casos, los gobiernos nacionales y subnacionales compiten entre sí por el mismo grupo limitado de talento. Las ciudades más pequeñas también pueden sufrir una "fuga de cerebros", ya que los jóvenes talentos se trasladan a ciudades más grandes que ofrecen más posibilidades profesionales (de Mello and Ter-Minassian, 2020^[17]).

Los desafíos relacionados con la brecha de competencias se pueden observar en casi todas las funciones gubernamentales analizadas en el Capítulo 5. Los desafíos en materia de competencias limitan la capacidad de los gobiernos para aprovechar los últimos avances en IA y pueden contribuir a la reticencia de los servidores públicos a aceptar el uso de la IA en general. En varias funciones, los gobiernos están luchando por determinar exactamente qué tipos de habilidades se necesitan y para quién.

La falta de competencias puede exacerbar otros riesgos y desafíos (Trajkovski, 2024^[18]). Por ejemplo, pueden dar lugar a malos resultados, a una sobreestimación y a una confianza indebida en las capacidades y sistemas de IA, a un uso indebido involuntario y al incumplimiento general de las leyes y otras normas. Los niveles inconsistentes de habilidades en los gobiernos pueden conducir a nichos de innovación, pero con poca capacidad para expandirse más allá de ellos.

Además, la falta de competencias internas en las administraciones públicas puede dar lugar a una dependencia excesiva de la externalización a través de la contratación pública (Mitchell, 2025^[19]; Autio,

Communigs and Elliott, 2023^[20]). Si bien la contratación pública es un aspecto importante y normal de la obtención de bienes y servicios relacionados con IA, depender excesivamente de la contratación pública para desarrollar las capacidades internas puede provocar el vaciamiento de las capacidades gubernamentales (Trajkovski, 2024^[18]). Esto puede formar un círculo vicioso, en el que los gobiernos carecen de las competencias adecuadas para diseñar programas de mejora de las competencias, comprender para qué competencias contratar y comprender plenamente las ofertas de los proveedores para adquirir los bienes y servicios adecuados a un precio justo. En general, sin un desarrollo proactivo de las habilidades, "las agencias públicas se encontrarán simplemente reaccionando a los cambios tecnológicos en lugar de dirigir estas tecnologías emergentes para servir eficazmente a los intereses de la sociedad" (Trajkovski, 2024^[18]). Si el gobierno no puede demostrar un uso eficiente y eficaz de la IA y el autocontrol de esta, es poco probable que pueda regular la tecnología. Esto también contribuye al desafío de los altos costos de adopción de la IA, como se analiza a continuación; contratar personal externo puede costar de tres a cuatro veces más por persona que los empleados gubernamentales (UK DSIT, 2025^[21]).

Varios gobiernos han puesto en marcha programas de mejora de competencias y de contratación específica (véase el Capítulo 4, "Fomento de las competencias y el talento"), y algunos incluso han utilizado la IA como herramienta para lograr estos objetivos (véase el Capítulo 5, "La IA en la reforma de la función pública").

Falta de datos de alta calidad y de capacidad para compartirlos

A través de todos los niveles de la administración y de casi todas las funciones gubernamentales analizadas en profundidad en el Capítulo 5, los desafíos en materia de datos constituyen un obstáculo para el desarrollo y uso de IA en la administración. Un reciente estudio realizado por RAND (2024^[5]; 2025^[22]) constató que los problemas relacionados con los datos eran uno de los principales factores que impulsaban los proyectos de IA fallidos, incluida la falta de datos adecuados, y señaló la importancia del trabajo, que a menudo se percibía como de bajo "prestigio de actividad" (por ejemplo, limpieza de datos).

En algunas funciones, los datos necesarios pueden simplemente no existir o no haberse digitalizado nunca del papel (por ejemplo, como se observa a menudo en la administración de justicia), o la calidad de los datos disponibles es deficiente (por ejemplo, registros mal estructurados, incompletos o mal escritos, discrepancias en los formatos de los datos). Esto puede deberse a diversas razones, como un control deficiente de los procesos de entrada de datos o incluso la falta de previsión previa de que dichos datos podrían llegar a ser importantes algún día. Pese a ser demorosos y engorrosos, estos problemas de calidad a menudo pueden superarse, por ejemplo, mediante procesos de digitalización, limpieza de datos y validación.

El énfasis reiterado en la falta de datos suficientes y de calidad puede parecer contradictorio, ya que los gobiernos poseen enormes cantidades de datos y, a menudo, los ponen a disposición como datos abiertos de gobierno (OGD) para, entre otras cosas, ayudar a servir como insumos para la formación de sistemas de IA (OECD, 2023^[23]). Sin embargo, en muchos casos, lo que supone un desafío es la capacidad de los organismos gubernamentales para compartir datos entre sí. Esto puede deberse a normas que impiden el intercambio o no son claras para los servidores públicos, procesos de aprobación de meses de duración o, con menos frecuencia, a que los gobiernos supriman los derechos de reutilización de datos en los contratos con las empresas. Estas cuestiones también se reflejan al compartir datos entre jurisdicciones y niveles de gobierno, lo que añade desafíos a la adopción de IA en los gobiernos subnacionales. Algunos países están tratando de cumplir con las normas de protección y gestión de datos, como el Reglamento General de Protección de Datos (RGPD) de la UE, aunque estas cuestiones de gestión de datos generalmente existían mucho antes de que dichas normas entraran en vigor. En otros casos, existe una falta de protocolos técnicos o políticos para el intercambio, o una falta de interoperabilidad entre los sistemas informáticos y los formatos de datos. En la mayoría de los casos, estos problemas son síntomas

de un problema más sistémico —gobernanza de datos inadecuada que da lugar a una recopilación y gestión de datos no estratégica, esporádica y fragmentada— y normas conexas en todo el gobierno. Las normas anticuadas u onerosas sobre el intercambio de datos también contribuyen y es posible que deban reconsiderarse para tener en cuenta los avances tecnológicos y, al mismo tiempo, proteger la privacidad.

En general, es fundamental que los gobiernos establezcan actividades sólidas de gobernanza y gestión de datos para tener éxito en la adopción de IA, aunque solo el 59 % de los países de la OCDE cuentan con una estrategia de datos para el sector público, y son aún menos los que proporcionan una guía práctica para su aplicación (OECD, 2024^[24]). Sin una sólida gobernanza de datos, los gobiernos se arriesgan a desarrollar e implementar sistemas de IA que utilicen datos de baja calidad, lo que da lugar a resultados que van desde simples inexactitudes hasta sesgos sistémicos y resultados injustos para los ciudadanos. Sin una sólida gobernanza de datos entre organizaciones y niveles de gobierno, las ambiciones de los gobiernos en materia de IA tendrían que limitarse en gran parte a pequeños experimentos y proyectos piloto en entornos limitados que son la norma hoy en día, como se ha comentado anteriormente.

Fomentar el desarrollo de sectores públicos basados en datos ha sido durante mucho tiempo un objetivo central de la OCDE (2019^[25]; 2023^[23]). Además, varios gobiernos están consolidando los fundamentos necesarios para que los datos puedan ser utilizados por los gobiernos al desplegar sistemas de IA (véase en el Capítulo 4, "Crear fundamentos sólidos para los datos").

Falta de marcos viables y directrices sobre el uso de IA

Las estrategias nacionales para la IA en el gobierno, ya sea las estrategias específicas o las integradas en instrumentos más amplios, son ahora comunes e importantes para definir una visión del éxito de la IA. Sin embargo, por lo general solo proporcionan detalles de alto nivel sobre compromisos y aspiraciones, ofreciendo orientaciones concretas limitadas para facilitar la materialización de los beneficios de la IA y, al mismo tiempo, proteger frente a sus riesgos. Además, a menudo no abordan las consideraciones operativas clave que podrían hacerlas efectivas. Por ejemplo, a menudo se pasan por alto la gestión de la inversión en tecnología y la contratación pública, a pesar de ser cruciales para la IA en el gobierno (van Noordt, Medaglia and Tangi, 2023^[26]; Monteiro, Hlacs and Boéchat, 2024^[27]). Para cerrar esta brecha, los gobiernos necesitan una guía práctica que esté alineada con las estrategias y proporcione a sus instituciones una dirección y garantías tangibles. La orientación también es importante para los gobiernos subnacionales, como las ciudades. Esto puede ser importante tanto para garantizar la alineación con los enfoques nacionales —ya que los gobiernos subnacionales a menudo siguen o se inspiran en los esfuerzos nacionales— como para ayudar a los gobiernos subnacionales a cumplir sus propias ambiciones digitales y de IA, así como las necesidades de sus ciudadanos y residentes.

Dichas orientaciones pueden ser de delimitación, abordando cuestiones de todo el sistema para promover una adopción fiable de IA, o específicas, centradas en funciones y aplicaciones gubernamentales particulares. Las orientaciones transversales aportan claridad y orientación sobre elementos fundamentales como la gobernanza de datos, el desarrollo del talento y la inversión. Por el contrario, la orientación vertical ayuda a los servidores públicos a navegar por las oportunidades y riesgos de la IA a través de medios efectivamente adaptados a los diferentes ámbitos de políticas. Estos enfoques no son exclusivos y ambos pueden aplicarse de forma complementaria.

En general, no existen directrices transversales y verticales concretas para la IA en el gobierno. Por ejemplo, una investigación del Comité Mixto de Cuentas Públicas y Auditoría del Parlamento de Australia (2025^[28]) — que no necesariamente representa opiniones del Gobierno Australiano — constató que, si bien algunas entidades gubernamentales están empezando a adoptar IA, carecen de orientación al respecto. Recomendó que se creara un grupo de trabajo integrado por todos los sectores del gobierno para estudiar las normas y los marcos de gobernanza necesarios para los sistemas de IA en toda la administración pública. Sin embargo, existen excepciones notables, como el "Manual de IA para el

Gobierno del Reino Unido" (Capítulo 4, Recuadro 4.2) y el enfoque estructurado de Francia para integrar la IA en los RR.HH. (Recuadro 5.19). Los gobiernos deben ir más allá de las declaraciones de intenciones y ofrecer orientaciones claras y prácticas sobre la adopción de la IA, la inversión, la gobernanza de datos, la contratación pública y el desarrollo de la fuerza laboral (Morley et al., 2019^[29]).

En el Capítulo 5 se cita esta brecha como un desafío en muchas funciones gubernamentales. En general, la falta de orientación puede contribuir a la aversión al riesgo, ya que brindarla puede aclarar la incertidumbre y reducir las dudas entre los funcionarios. La necesidad de orientación se ha planteado específicamente en las funciones de diseño regulatorio, contratación pública, lucha contra la corrupción y promoción de la integridad pública, y administración tributaria para abordar las incertidumbres jurídicas y de gobernanza que dejan a los servidores públicos con incertidumbre sobre si pueden usar la IA en estas funciones y de qué manera. En algunos casos, se necesita orientación para interpretar claramente las leyes y reglamentos, y su aplicación práctica. En otros, aportan claridad en situaciones en las que puede que todavía no existan normas formales. Dadas las variaciones sectoriales, los enfoques transversales de la IA deben complementarse con orientaciones especializadas que tengan en cuenta los desafíos políticos, los perfiles de riesgo y el panorama de datos. Los gobiernos también deberán tener en cuenta que la IA es un ámbito que evoluciona rápidamente, y que dicha orientación debe ser flexible y es probable que necesite adaptación e iteración para mantenerse al día con el ritmo del cambio.

En el Capítulo 4, "Establecimiento de mecanismos y procesos clave de gobernanza" y "Uso de instrumentos de política para orientar una IA fiable", se puede encontrar más información sobre las medidas gubernamentales para superar este desafío.

Impulsar la innovación mitigando los riesgos

Un estudio elaborado por la OCDE (2024^[30]) reveló que, en lo que respecta a la IA en la administración pública, los gobiernos tienden a centrarse en los aspectos negativos (es decir, los riesgos de la IA y cómo mitigarlos), pero no tanto en los positivos (es decir, los beneficios de la IA y cómo aprovecharlos). Esto no es exclusivo de los enfoques de los gobiernos con respecto a las actividades internas de IA. En lo que respecta a la economía y la sociedad en general, los expertos en IA han constatado que los debates e iniciativas de política pública suelen reconocer que la IA puede reportar beneficios significativos, pero las acciones públicas a menudo no se centran explícitamente en la consecución de dichos beneficios. Por el contrario, las abordan indirectamente a través de efectos positivos indirectos al tratar de mitigar los riesgos (OECD, 2024^[31]). Estos expertos han instado a los gobiernos a tomar medidas más directas para aprovechar las oportunidades que presenta la IA.

Gran parte de este enfoque orientado al riesgo puede atribuirse a la aversión al riesgo, que durante mucho tiempo ha obstaculizado los esfuerzos de innovación del gobierno digital y el sector público, fomentando una cultura resistente al cambio en la que se debe evitar el fracaso a toda costa, incluso en lo relativo a la IA (OECD, 2021^[32]; 2017^[33]; 2019^[33]; Desouza, 2018^[34]; SAS, 2020^[35]; Richter, 2024^[36]). Una encuesta llevada a cabo por Deloitte (2024^[7]) mostró que el 63 % de los líderes del sector público creían que la GenIA erosionaría el nivel general de confianza en las instituciones nacionales y mundiales. Esta cautela contribuyó a una adopción más lenta de la IA en el gobierno que en la industria. En el Capítulo 5 se presentan ejemplos de aversión al riesgo en varias funciones gubernamentales, especialmente en la contratación pública, la lucha contra la corrupción y la promoción de la integridad pública. Por ejemplo, las instituciones de integridad pueden ser reacias al riesgo por miedo a cometer errores en el proceso de adopción de IA, ya que las directrices gubernamentales hacen hincapié en *lo que no se debe hacer* en lugar de ofrecer directrices prácticas sobre *cómo* adoptar la IA de manera fiable. También surgen problemas como resultado de una corrección excesiva en respuesta a incidentes de IA, especialmente los que se cubren en los medios de comunicación. Ejemplos de ello son los casos en los que chatbots públicos proporcionaron información errónea o fueron pirateados (Hodges, 2024^[37]; Fagan, 2024^[38]). La aversión

al riesgo también es un tema de conversación frecuente cuando se debate sobre IA en las reuniones de grupos de trabajo y foros pertinentes de la OCDE.

Los gobiernos deben aplicar una de las mejores prácticas generalmente aceptadas para promover el desarrollo de la IA, por ejemplo, considerando el nivel de riesgo o impacto potencial que podría tener un sistema de IA y desarrollando medidas adaptadas y proporcionales para superar posibles problemas adversos. Sin embargo, los gobiernos a menudo parecen tratar la mayoría de los esfuerzos de IA como si fueran de alto nivel de riesgo o impacto, exigiendo requisitos estrictos en todos los ámbitos e imponiendo engorrosos requisitos burocráticos, desalentando a los servidores públicos que buscan innovar. Esto puede servir como desincentivo suficiente para evitar el proceso de exploración y puede verse en un estudio de Deloitte (2023^[39]), que analizó todas las iniciativas de política en la base de datos OECD.AI de políticas y estrategias nacionales³. En él se concluyó que las políticas ponderadas por riesgo, cuyo objetivo es pasar de un enfoque único para todos a un enfoque basado en datos y riesgo, eran poco frecuentes y representaban el 2 % de las iniciativas incluidas en la base de datos. El problema también se reconoce en el memorándum de Estados Unidos de abril de 2025 sobre "Aceleración del uso federal de IA mediante innovación, gobernanza y confianza pública" (^[40]), que señala a las agencias gubernamentales de "eliminar los requisitos burocráticos innecesarios que inhiben la innovación y la adopción responsable".

Los servidores públicos también pueden desarrollar "aversión algorítmica". Esto es algo opuesto al "sesgo de automatización" discutido en el Capítulo 1, donde "los humanos son reacios a usar algoritmos a pesar de su rendimiento superior" (Cheng and Chouldechova, 2023^[41]; Sunstein and Gaffe, 2024^[42]), a menudo después de ver errores en los resultados de la IA. Esto sugiere posibles problemas de competencias relacionados con la comprensión de la IA, sus puntos fuertes y débiles relativos y cómo utilizar óptimamente sus resultados. También sugiere una falta de confianza en sus habilidades para la colaboración hombre-máquina, así como una falta de entornos controlados para probar y experimentar de manera segura con IA. Estos sesgos, que pueden distorsionar la percepción de la fiabilidad de la IA, pueden mitigarse mediante intervenciones estructuradas, como formaciones sobre las fortalezas y limitaciones de la IA, tal y como se analiza más adelante en el Capítulo 4 (Featherson, Shlonsky and Lewis, 2019^[43]). Los trabajadores también necesitan sentir que tienen voz sobre los insumos utilizados para un sistema de IA y que pueden utilizar su criterio profesional en la forma en que se utilizan los resultados (Dietvorst, Simmons and Massey, 2018^[44]; Cheng and Chouldechova, 2023^[41]).

Existe cierta evidencia de que la aversión al riesgo por la IA en el gobierno puede estar disminuyendo a medida que los gobiernos se familiarizan más con la tecnología. Un estudio reciente realizado por Google Public Sector concluyó que las preocupaciones sobre IA de los líderes de TI pública de EE.UU., en relación con cuestiones como la privacidad y la seguridad, están desapareciendo (Teale, 2025^[45]). Sin embargo, los gobiernos tendrán que ser más activos para superar este enfoque orientado al riesgo y considerar mejor las concesiones que apuntan mejor a las oportunidades. Los diversos habilitadores y salvaguardias de la IA que se analizan en el Capítulo 4 pueden ayudar a desviar la atención de una cultura de aversión al riesgo hacia una adopción más controlada y una gestión del riesgo informada.

Demostrar resultados y rendimiento de la inversión

Los gobiernos han dado pasos importantes en la implementación de soluciones de IA en diversos ámbitos públicos, demostrando beneficios tangibles en eficiencia, precisión y prestación de servicios. Sin embargo, el seguimiento de los avances y la evaluación retrospectiva exhaustiva del impacto siguen siendo aspectos subdesarrollados de la implementación de la IA gubernamental. Si bien algunos casos aislados de éxito están bien documentados, como se observa más adelante, a menudo faltan esfuerzos integrales para evaluar la contribución de la IA a la creación de valor público. Esto se puede observar en el caso del Reino Unido, donde "solo el 8 % de los proyectos de IA muestran beneficios medibles y solo el 16 % muestran costos de pronóstico, lo que dificulta su evaluación en comparación con un análisis de costo-

beneficio". (UK DSIT, 2025^[21]). Una encuesta de Deloitte específica para la IA generativa (Austin et al., 2024^[7]) en 14 países muestra que, a pesar de prever un aumento de las inversiones en IA, el 78 % de los líderes gubernamentales encuestados afirman tener dificultades para medir los impactos de la IA general, un porcentaje significativamente mayor que el de otros sectores, lo que supone un obstáculo para la adopción y la ampliación de la IA incluso cuando se resuelven otros desafíos, como las brechas de talento.

Múltiples soluciones de IA han demostrado resultados concretos y medibles que ilustran el potencial de la tecnología para transformar la prestación de servicios públicos. Estos resultados cuantificables proporcionan puntos de referencia valiosos para comprender el impacto directo de la IA en la eficiencia operativa y la calidad del servicio:

- El sistema Amauta Pro AI del Perú ha transformado la velocidad a la que los tribunales pueden responder a las víctimas de violencia doméstica. Este sistema impulsado por IA ha reducido el tiempo necesario para redactar proyectos de resolución sobre medidas de protección de 3 horas a 40 segundos (véase el Recuadro 5.63).
- En la UE, el proyecto DATA CROS desarrolló una herramienta para detectar anomalías en las estructuras de propiedad de las empresas que puedan indicar riesgos de corrupción, blanqueo de capitales y otros delitos financieros. En 2021, la herramienta de predicción detectó correctamente el 83 % de las empresas sancionadas y el 88 % de las empresas con propietarios sancionados (véase el Recuadro 5.27).
- La Agencia Federal para la Gestión de Emergencias (FEMA) de los Estados Unidos desarrolló un sistema de IA para evaluar los daños estructurales en las zonas afectadas por el huracán Ian, que redujo el número de estructuras que requieren revisión humana de más de un millón a solo 77 000. En un plazo de 72 horas tras la llegada del huracán a tierra en 2022, la FEMA tuvo conocimiento de la magnitud de los daños en las regiones afectadas, lo que permitió acelerar la asignación de recursos y la planificación de la recuperación (véase el Recuadro 5.58).

Son especialmente notables los casos en los que los sistemas de IA se han comparado explícitamente con el rendimiento humano, lo que pone de relieve mejoras significativas en la velocidad, escala y utilización de recursos que superan la capacidad humana. En Singapur, las agencias gubernamentales transformaron la contratación con herramientas de IA disponibles en el mercado, lo que permitió a una agencia procesar más de 3 000 solicitudes para su programa de gestión asociada de manera eficiente, ahorrando 44 000 EUR (equivalente) y más de 150 días de productividad del personal⁴. Las comparaciones con el desempeño humano son importantes porque se centran en el contraste clave que se necesita para la toma de decisiones basada en evidencia. Además, promueven una comprensión más profunda del desempeño humano, lo que permite revelar supuestos y sesgos implícitos que afectan el rendimiento humano.

Más allá de los casos de uso individuales, algunos gobiernos han comenzado a documentar el impacto de la IA a nivel organizativo y nacional, revelando importantes beneficios financieros y mejoras operativas. Estas evaluaciones más amplias ayudan a establecer el valor acumulado de las inversiones en IA en todas las funciones gubernamentales. La Agencia Tributaria australiana, por ejemplo, ha informado de que su enfoque de IA, que combina análisis en tiempo real, formularios prellenados y sistemas de detección de anomalías, ayudó a proteger unos 78,9 millones de dólares australianos de ingresos en más de 636 000 interacciones con usuarios en 2023-2024 (Recuadro 5.5). Del mismo modo, se observaron resultados sustanciales en Austria gracias a la actividad del Centro de Competencia en Analítica Predictiva (PACC) del Ministerio Federal de Finanzas, que permitió analizar 6,5 millones de casos en los sectores del impuesto sobre la renta, impuesto de sociedades y valor agregado, así como en las transacciones aduaneras en 2023 (Recuadro 5.3). Estos análisis detectaron casos de declaración falsa en las liquidaciones de impuestos de los empleados e identificaron actividades fraudulentas, que dieron lugar a unos ingresos fiscales adicionales de aproximadamente 185 millones EUR. Mirando al futuro, un estudio reciente del Instituto Alan Turing (2024^[46]) sobre los servicios públicos del Reino Unido indican que la IA

podría ayudar a automatizar el 84 % de las transacciones relacionadas con los servicios del gobierno central, ahorrando un equivalente de aproximadamente 1 200 personas-año de trabajo cada año.

A pesar de estos éxitos, este tipo de consideraciones son raras, y los gobiernos se enfrentan a importantes desafíos a la hora de monitorear sistemáticamente el progreso e impacto de la IA. Una barrera clave es la falta de marcos de medición y evaluación bien definidos que permitan evaluar las contribuciones de la IA de forma estandarizada. Muchas aplicaciones de IA se integran en complejos procesos administrativos, lo que dificulta el aislamiento y la medición de sus efectos específicos. Además, el desafío de comparar la IA con el rendimiento humano se agrava por el hecho de que muchas tareas basadas en IA serían inviables o prohibitivamente engorrosas si no existiera la automatización. También existe un conocimiento limitado del impacto a largo plazo del uso de los LLM en la cognición humana, y de si su uso coherente afecta a la creatividad, las habilidades de pensamiento crítico y la productividad de quienes los utilizan.

Una consideración final es que diferentes contextos pueden requerir diferentes metodologías. Por ejemplo, un tema de debate en la última edición de la Mesa Redonda de la OCDE sobre Ciudades Inteligentes (2024^[41]) señaló que las ciudades deben explorar diferentes metodologías para medir y evaluar el éxito que se ajusten a sus propios objetivos y les permitan establecer metas medibles. Están surgiendo algunos esfuerzos iniciales de los gobiernos para abordar estos problemas, como la publicación del Gobierno del Reino Unido sobre las mejores prácticas para evaluar el impacto de los métodos de evaluación de IA (Frontier economics, 2024^[47]). Los EE.UU. (2025^[48]) también han publicado recientemente una política de adquisiciones de IA que reconoce que los organismos gubernamentales deben "salvaguardar el dinero de los contribuyentes mediante el seguimiento del rendimiento de IA y gestión de riesgos". Sin mecanismos de monitoreo sólidos, los gobiernos corren el riesgo de estimar erróneamente el valor y los riesgos potenciales de la IA, así como perder oportunidades de mejora.

Establecer marcos eficaces de medición del impacto es crucial para garantizar que las inversiones en IA generen valor real para las administraciones públicas y los ciudadanos. A medida que los gobiernos asignen más recursos al desarrollo y despliegue de la IA, será imprescindible demostrar un claro retorno de la inversión (RSI). Disponer de mecanismos fiables de evaluación retrospectiva del impacto puede ayudar a los responsables de la formulación de políticas a tomar decisiones informadas sobre la ampliación de las soluciones de IA, la optimización de la asignación de recursos y la justificación de un financiamiento adicional. Además, la evaluación de impacto proporciona información esencial para refinar los sistemas y enfoques de IA, lo que permite ciclos de mejora continua. Los productos documentados también facilitan el intercambio de conocimientos entre entidades gubernamentales, lo que ayuda a ampliar los enfoques exitosos y a evitar que se repitan los infructuosos. Tal vez lo más importante sea la presentación de informes transparentes sobre los impactos de la IA, tanto positivos como negativos, esencial para mantener la confianza y la rendición de cuentas a medida que estas tecnologías se integran más profundamente en las funciones y actividades centrales del gobierno. Los diferentes métodos de evaluación son adecuados para cada contexto, pero los gobiernos deben intentar comparar la implementación de la IA con la situación en la que no existe. La OCDE ha elaborado directrices para elegir un enfoque de evaluación basado en diversas consideraciones clave (Varazzani et al., 2023^[49]).

Desafíos que son un poco menos comunes o varían entre las funciones gubernamentales

Entornos legales y regulatorios inflexibles u obsoletos

Los entornos regulatorios inflexibles, obsoletos o inadecuados (por ejemplo, excesivos o deficientes) plantean muchos desafíos. Muchas funciones se enfrentan a restricciones regulatorias o legales en el acceso e intercambio de datos, como se ha comentado anteriormente. Además, puede haber confusión sobre la precisión de la IA y sobre si los errores involuntarios introducidos por el uso de IA podrían dar

lugar a un incumplimiento de regulaciones y otras normas, por ejemplo, en informes fiscales. La complejidad de las regulaciones también es un factor por considerar. Por ejemplo, los funcionarios de las administraciones tributarias se enfrentan a leyes muy complejas en torno a procesos fiscales, lo que contribuye a que se basen en gran medida en enfoques clásicos basados en normas. Estos desafíos son tan comunes a nivel local como en los gobiernos nacionales (OECD, 2024^[4]).

A veces, el reto está en las lagunas regulatorias que conducen a confusión sobre lo que es aceptable con IA. Esta confusión puede contribuir a otros desafíos, como la aversión al riesgo o la preferencia por el mantenimiento de la propia situación existente (Samuelson and Zeckhauser, 1988^[50]). Por ejemplo, como en muchos países no se aborda específicamente, los funcionarios encargados de la contratación pública a menudo no tienen claro si la IA puede utilizarse en procesos de contratación, por temor a que al hacerlo puedan verse expuestos a desafíos por parte de los licitadores infructuosos u otros que cuestionen la equidad del proceso. Esto proporciona una falta general de incentivos para el cambio. También existe confusión sobre si los sistemas avanzados de IA, que son muy capaces pero funcionan de forma opaca, pueden cumplir con estándares regulatorios como las Normas Internacionales de Auditoría o las reglas probatorias del sistema judicial penal. Como resultado, algunas personas podrían optar por no utilizar la IA con el fin de evitar estos riesgos, lo que a su vez limitaría los posibles beneficios derivados de su implementación.

Los entornos regulatorios plantean un desafío único en cuanto a diseño e implementación de la regulación. Más allá de las normas que restringen el uso de IA, los reguladores también deberían ser cautos y evitar hacer cambios frecuentes en las regulaciones, y en la forma en que se implementan y se hacen cumplir. Las entidades reguladas necesitan un nivel de claridad y previsibilidad para cumplir con las regulaciones de una manera que cause una interrupción mínima en las operaciones comerciales. Los frecuentes cambios regulatorios, aunque se basen en información de calidad basada en la IA, pueden dar lugar a un entorno regulatorio volátil, lo que dificulta a las empresas la planificación de estrategias a largo plazo y a que el público se mantenga informado sobre la legislación vigente.

Los gobiernos pueden superar estos desafíos garantizando que las regulaciones y otras reglas formales estén actualizadas, sean ágiles y minimicen la ambigüedad. Hay evidencia de que los gobiernos están avanzando en esta dirección. Por ejemplo, Deloitte (2023^[39]) concluyó que la regulación adaptativa, que de un modo "regular y olvidar" a un enfoque receptivo e interactivo, es uno de los tipos de políticas más comunes monitoreadas por el Observatorio de Políticas OECD.AI. También es importante contar con directrices adicionales que ayuden a contextualizar las regulaciones en funciones particulares del gobierno. El análisis de cómo lo están haciendo los gobiernos se encuentra en el Capítulo 4, "Establecimiento de mecanismos y procesos clave de gobernanza" y "Uso de instrumentos de política para orientar una IA confiable".

Costos elevados o inciertos de la adopción y escalada de la IA

Si bien la adopción de la IA tiene el potencial de reducir costos a través de una mayor productividad y eficiencia, muchas organizaciones gubernamentales tienen dificultades para realizar las inversiones financieras iniciales para comenzar sus viajes de adopción de IA, o para ampliar los casos de uso que resultan exitosos. Estos costos pueden variar en ocasiones desde el pago de tarifas de licencia por empleado para ofertas de IA basadas en servicios, como ChatGPT o Microsoft Copilot, hasta amplios costos de desarrollo, personalización y soporte para soluciones internas más personalizadas (Shark, 2025^[51]; Barrett and Greene, 2024^[52]). En el Reino Unido, una encuesta realizada por SAS a servidores públicos (2025^[53]) destacó que el principal desafío eran las restricciones presupuestarias y los costos (planteadas por el 67 % de los encuestados), seguidas de cerca por la falta de competencias internas (63 %). A pesar de la naturaleza crítica del financiamiento de IA, el Índice de Gobierno Digital (DGI) de la OCDE (2024^[24]) destaca que solo el 15 % de los países de la OCDE cuentan con un marco de inversión en IA pública.

En el Capítulo 5 se mencionan los desafíos financieros que plantea la adopción de IA para varias funciones gubernamentales, como el diseño y la implementación de la regulación, los servicios públicos, la administración tributaria, la lucha contra la corrupción y el fomento de la integridad pública, así como la participación ciudadana. En algunos casos, los desafíos financieros también están relacionados con los costos de contratación o adquisición de talento calificado, y las carencias en materia de competencias se han analizado anteriormente como un desafío aparte. Funciones como la administración tributaria también ha indicado que el proceso para asegurar los presupuestos en el gobierno es un desafío.

Este desafío puede contribuir a la conclusión de la OCDE de que los gobiernos a menudo parecen estancados en las fases exploratoria y piloto, con una escalada limitada de soluciones exitosas. Por ejemplo, las autoridades tributarias han dicho a la OCDE que llevar a cabo pequeñas pruebas piloto es barato y fácil, incluso con sistemas avanzados obtenidos del sector privado. Sin embargo, estos costos pueden aumentar exponencialmente a medida que las ofertas de IA se implementan de forma más amplia dentro de las organizaciones o se amplían a otras partes del gobierno. Los costos son particularmente altos para las soluciones específicamente diseñadas, ya que un grupo de 10 países se centró en la IA en aras del interés público y afirmó que "se ha asumido que la barrera para escalar los modelos de IA es principalmente la falta de disponibilidad y asequibilidad de los ordenadores" (France Élysée, 2025^[54]).

Los gobiernos deben reconocer que la inversión insuficiente en tecnología aumenta los costos a largo plazo y los costos totales de propiedad (UK DSIT, 2025^[21]). Algunos gobiernos están tratando de abordar estos problemas mediante inversiones específicas, así como la prestación de servicios centrales que ayudan a aliviar la necesidad de que cada organismo cree o compre sus propias soluciones. Pueden verse en el Capítulo 4, en los apartados "Invertir con propósito", "Crear infraestructuras digitales" y "Crear espacios para experimentar". Algunos también utilizan modelos de código abierto o exploran modelos más pequeños que pueden diseñarse para responder a necesidades sociales y comunitarias específicas, lo que requiere menos poder computacional y datos (France Élysée, 2025^[54]).

Aunque los costos señalados por los gobiernos en el Capítulo 5 tienden a centrarse en los financieros, es útil que los gobiernos tengan en cuenta que no solo los costos monetarios afectan a la adopción y la escalada de la IA. Los costos psicológicos relacionados con el uso de la IA también pueden afectar al grado en que las personas utilizan las herramientas de IA en su trabajo diario, incluso si las inversiones se realizan para ponerlas a disposición. Estos costos pueden incluir los costos de búsqueda, que ocurren cuando las personas buscan información pero encuentran información obsoleta, lenguaje confuso o requisitos confusos, o los costos cognitivos, los recursos mentales que las personas gastan en comprender información compleja (Shahab and Lades, 2021^[55]). Las "auditorías de fricción" (también conocidas como 'sludge audits' en inglés) son evaluaciones estructuradas del comportamiento de un proceso de toma de decisiones cuyo objetivo es identificar, prevenir y reducir fricciones innecesarias y costos psicológicos, que impiden que las personas tomen acciones que de otro modo tomarían (OECD, 2024^[56]). Al realizar auditorías de fricción sobre el uso de herramientas de IA, los gobiernos pueden comprender y abordar las barreras a la aceptación que pueden limitar la adopción y escalada de la IA.

Los gobiernos pueden carecer de claridad sobre cuánto puede o debe costar el desarrollo y uso de la IA

Los desafíos analizados anteriormente son más relevantes cuando se conocen dichos costos. Los gobiernos han informado a la OCDE de que, con frecuencia, existe incertidumbre o confusión sobre el costo que podría o debería suponer el desarrollo o el uso de diferentes tipos de sistemas de IA. Esto dificulta a las instituciones públicas en su planificación eficaz y evaluación de ofertas de proveedores a la hora de considerar la contratación pública para buscar soluciones. Aclarar estos costos puede ayudar a garantizar que los gobiernos estén preparados para adoptar sistemas de IA de forma estratégica y sostenible. Sin embargo, la OCDE no pudo identificar ninguna investigación que abordara el costo del desarrollo o el uso de diferentes tipos de sistemas de IA en el gobierno. Esto sugiere un área óptima para

futuras investigaciones y análisis. Aun así, comprender los costos en un sentido más amplio o a partir de proyectos gubernamentales específicos de IA puede ayudar a los gobiernos a llegar a estimaciones con fines de planificación. Esta sección pretende dar los primeros pasos para ayudar a los gobiernos a hacerlo, con la posibilidad de profundizar en el trabajo de la OCDE sobre este tema en el futuro.

El costo de adoptar IA puede variar significativamente dependiendo del tipo de sistema y su escalada de uso. Por ejemplo, los gobiernos pueden optar por diversas opciones a la hora de adoptar la IA, como las que se mencionan a continuación.

Herramientas de concesión de licencias del sector privado con precios fijos por usuario o licencia

Las empresas que ofrecen herramientas de IA suelen cobrar por usuario. Este es el caso de servicios como Microsoft 365 Copilot, ChatGPT de OpenAI o Claude de Anthropic. Las licencias para estos servicios se pueden adquirir a nivel de empresa, con precios que oscilan entre 30 y 100 USD por usuario y mes. Por ejemplo, la versión empresarial de Microsoft 365 Copilot cuesta 30 USD al mes⁵. Por lo tanto, un gobierno ejecutando un proyecto piloto comparable al piloto de seis meses de Australia con 7.600 empleados en 60 organismos (2024^[57]) podría estimarse en torno a los 1,37 millones de dólares solo en el caso de las licencias; sin incluir adiciones como tiempo de personal para la administración del programa piloto y la presentación de informes sobre resultados, así como otros gastos generales. Para el proyecto piloto de Australia, el Tesoro australiano (2025^[58]) estimó que la licencia podría pagarse por sí misma para un personal gubernamental de nivel medio si la IA permitiera ahorrar 13 minutos de su tiempo semanal y destinarlo a tareas de mayor valor. Los costos de ChatGPT Enterprise no se publican en el sitio web de OpenAI, aunque cuando se le preguntó, ChatGPT sugirió un costo de 60-100 USD por usuario al mes, que varía según el volumen y características. Los precios del Plan Empresarial Claude de Anthropic también dependen de las necesidades y características de las empresas. Si bien los precios no están disponibles en el sitio web de Anthropic, los sitios web de terceros calculan un costo de 60 USD por persona al mes⁶.

Uso de sistemas de IA generativa del sector privado con precios basados en el volumen, como tokens

Los sistemas GenAI suelen ofrecerse a través de precios basados en el volumen, basados en el número de llamadas a la API. Mientras que la prestación a través de licencias es más pertinente cuando los servidores públicos utilizan directamente las herramientas de IA (por ejemplo, al utilizar Copilot para ayudar a redactar documentos), la tarificación basada en el volumen es más pertinente cuando los gobiernos crean servicios internos o de cara al público que se conectan con un modelo propietario, o cuando quieren personalizar (es decir, ajustar) el contenido que considera el modelo o cómo produce los resultados. Las principales funciones del precio incluyen:

- Tokens de entrada: tokens incluidos en un prompt, o indicación con instrucciones, contexto o datos enviados al modelo.
- Tokens de salida: tokens generados por el modelo en respuesta a una entrada.
- Tokens de entrenamiento: datos (por ejemplo, fragmentos de texto) de los que aprende un modelo de IA durante la formación.

Los modelos ven los datos como tokens, no como oraciones o párrafos. El costo de los tokens depende de la empresa y puede basarse en el nivel de complejidad y recursos necesarios para un modelo individual. Por ejemplo, los costos de algunos modelos de uso común se indican en la Tabla 3.1.

Tabla 3.1. Costo de 1 millón de tokens en modelos comunes de IA generativa (USD)

Un millón de tokens representan aproximadamente 750 000 palabras, 100 000 líneas de código, 11 horas de audio transcrita o 1 hora de vídeo transcrita

Modelo	Tokens de entrada	Salida	Entrenamiento (si procede)
OpenAI GPT-4o	\$2,50 USD	10 \$ USD	
OpenAI GPT-4o (si se ajusta para personalizarlo)	\$3,75 USD	15 \$ USD	25 \$ USD
OpenAI GPT-3.5-turbo	\$0,50 USD	1,50 \$ USD	8 \$ USD
Google Gemini 2.5 Pro	\$1,25 USD	10 \$ USD	
Flash de Google Gemini 2.0	\$0,10 USD	0,40 \$ USD	
Mistral Grande 24,11	\$2 USD	6 \$ USD	9 \$ USD
Mistral NeMo	\$0,15 USD	0,15 \$ USD	1 \$ USD

Nota: A 10 de abril de 2025. Los costos de inferencia para utilizar un modelo concreto o su equivalente tienden a disminuir con el tiempo (Stanford HAI, 2025^[59]).

Fuente: <https://openai.com/api/pricing>, <https://ai.google.dev/gemini-api/docs/pricing>, <https://mistral.ai/products/la-plateforme>, <https://help.openai.com/en/articles/4936856-what-are-tokens-and-how-to-count-them>, <https://prompt.16x.engineer/blog/code-to-tokens-conversion>, <https://blog.google/technology/ai/google-gemini-next-generation-model-february-2024>.

Además, algunas empresas están ofreciendo modelos base a través de niveles específicos dedicados a organismos gubernamentales⁷. Estos servicios buscan cumplir con los estrictos estándares de seguridad de los gobiernos. Además, su objetivo es adaptarse a las necesidades de los gobiernos, proporcionando soluciones que faciliten la gestión de sus propios requisitos de seguridad, privacidad y cumplimiento, así como permitirles utilizar los servicios para actividades que puedan estar fuera de las políticas de uso estándar. Las experiencias de un laboratorio de IA del gobierno central en la aplicación de este enfoque se analizan en el Recuadro 3.1. Los modelos privados también se pueden utilizar junto con los modelos de código abierto, como se indica a continuación.

Recuadro 3.1. La experiencia del laboratorio de IA gubernamental con IA privada

Operaciones y gastos

Un laboratorio de IA del gobierno central en un país sigue un enfoque gradual para explorar, probar y ampliar los proyectos de IA para su uso en el sector público por parte de servidores públicos. En total, tiene alrededor de 100 proyectos en fase de estudio, unos 15 en fase de pruebas limitadas y entre cinco y siete son accesibles para usuarios reales mediante un despliegue piloto o completo.

El laboratorio utiliza alojamiento en la nube e IA de Azure OpenAI, Vertex AI (Google) y Amazon Web Services (AWS). Cuenta con un presupuesto de unos 17,5 millones de euros. Todo el trabajo se realiza internamente. La mayoría de sus gastos son de personal, que incluyen aproximadamente:

- 15 empleados a tiempo completo equivalentes (ETC) para talento técnico (siete para desarrollo e ingeniería de IA, cuatro para diseño e investigación de usuarios y cuatro para nube/infraestructura).
- 6 ETC para la gestión de la prestación de servicios, que es fundamental para garantizar que el talento técnico pueda centrarse en desafíos técnicos, mientras que los gestores de la prestación de servicios se centran en abordar desafíos relacionados con las políticas y la burocracia.

- 6 ETC para análisis de impacto que utilizan la ciencia de datos para estudiar resultados e impacto de los proyectos.

Su proyecto más grande cuenta con unos 4 000 usuarios y alrededor de 8 ETC que trabajan en él. Otros proyectos son más pequeños, y algunos cuentan con uno o dos ETC. En total, los productos del laboratorio tienen unos 10 000 usuarios mensuales. Los costos totales de los servicios en la nube de IA, incluidos los tokens, rondan los 3 500 EUR al mes.

Lecciones aprendidas

- Los primeros proyectos son, por mucho, los más caros y lentos, con importantes inversiones en la configuración de infraestructuras en la nube y plantillas de despliegue que pueden reutilizarse fácilmente para futuros proyectos. Los despliegues que duraron tres semanas cada uno para los primeros proyectos ahora tardan 30 minutos.
- Es importante contar con una sólida infraestructura en la nube y compartirla de manera óptima entre proyectos para promover sinergias.
- El laboratorio tuvo en cuenta las ventajas y desventajas de utilizar modelos de IA privados frente a las implementaciones personalizadas de modelos de código abierto (por ejemplo, Llama de Meta). Determinó que los modelos privados de precio por volumen eran más efectivos porque los servidores públicos tienden a usar los sistemas de IA de 9:00 a 18:00. Para una implementación personalizada, tendrían que pagar por el uso de la GPU todo el día, incluso cuando los modelos no se están utilizando. En general, los precios basados en tokens eran menos costosos para sus necesidades. Además, este enfoque permite que el laboratorio genere nuevas instancias de modelos de forma más rápida y sencilla. Por ejemplo, puede implementar un nuevo modelo GPT-4.1 en unos cinco minutos, mientras que la implementación personalizada de un modelo de código abierto podría llevar semanas de trabajo de infraestructura.
- En general, el laboratorio estimó que habría costado 9 300 EUR al mes alojar por sí mismo un modelo Llama, donde actualmente gastan alrededor de 3 500 EUR en tokens.
- A medida que ha crecido el uso de herramientas de IA del laboratorio, este está llegando a los límites en que los proveedores de nube están dispuestos a proporcionar en términos de precios de pago por uso. A medida que su uso continúa aumentando, se enfrentan a la opción de 1) comprar capacidad de GPU de los proveedores de la nube, o 2) modelos de código abierto con alojamiento propio y pagar por el acceso a la GPU directamente (como se mencionó en el punto anterior). Para el laboratorio, la opción 1 puede ser óptima porque todavía permite un despliegue rápido.
- El costo del desarrollo de IA, tanto en términos de recursos técnicos (por ejemplo, servicios en la nube o tokens) como de recursos humanos, está disminuyendo rápidamente. El laboratorio está descubriendo que cada vez más pueden usar IA para construir IA, lo que puede cambiar las demandas laborales. Aún no se han determinado las implicaciones para esto.

Fuente: Entrevista de la OCDE a funcionarios de un país no revelado el 18 de abril de 2025. La OCDE no publica el nombre del país o del laboratorio debido a la naturaleza preliminar de las estimaciones y los análisis.

Desarrollo de aplicaciones de ML personalizadas y específicas (tanto internas como adquiridas)

Los sistemas de IA restringidos y adaptados a tareas específicas del sector público pueden variar desde gastos relativamente pequeños hasta proyectos multimillonarios. Estos sistemas implican desarrollos de ML para casos de uso específicos, como la detección de fraudes, la optimización del tráfico o la

clasificación de documentos. Los pilotos simples podrían construirse por unos pocos miles de dólares, mientras que los sistemas nacionales complejos pueden costar millones o decenas de millones (USD), especialmente si se exploran aplicaciones de defensa (Barnett, 2020^[60]). Por ejemplo, el gobierno de Australia Meridional está probando cuatro cámaras con IA para reducir la congestión del tráfico mediante el análisis de congestión y ajuste de los ciclos de los semáforos a un costo de 218 000 USD (equivalente) (Jackson, 2025^[61]).

Sin embargo, los costos varían mucho según la complejidad y el contexto, y es posible que se necesiten costos adicionales para la preparación de datos, la infraestructura y el monitoreo y mantenimiento continuos. Dado que los enfoques y costos asociados varían de forma considerable según el caso de uso, es difícil proporcionar estimaciones más allá de estos ejemplos. Podría justificarse un análisis más detallado para considerar los diferentes aspectos de estos casos de uso y lo que los distintos gobiernos de todo el mundo han pagado por ellos.

Desarrollo de modelos de GenAI personalizados diseñados y entrenados desde cero

En comparación con el desarrollo y la formación personalizados de un modelo de GenAI, que se analiza más adelante, los modelos de IA de código abierto previamente entrenados (como los modelos Llama de Meta) pueden ofrecer costos reducidos, ya que el gobierno sigue siendo capaz de personalizar en gran medida el modelo para satisfacer sus necesidades⁸. Los modelos de código abierto pueden alojarse por sí solos en la nube o en las instalaciones, lo que ofrece a los gobiernos un mayor control sobre sus datos y una eficiencia de costos a largo plazo. Aunque el uso de un modelo de código abierto previamente entrenado puede reducir los costos de ajustes a los que se generan con los ajustes, el resto de los costos pueden ser considerables. Por ejemplo, el autoalojamiento elimina las tarifas recurrentes de tokens o licencias, pero requiere una inversión inicial significativa en hardware e infraestructura, recursos de nube, consumo de energía y costos de mantenimiento y soporte.

A pesar de una mayor inversión inicial en comparación con los precios basados en licencias o volumen, algunos gobiernos han descubierto que el autoalojamiento puede ser rentable a gran escala y desbloquear usos no viables a través de API comerciales (por ejemplo, tareas de inteligencia sensible, servicios locales siempre activos u operaciones sin conexión en infraestructuras críticas). Por ejemplo, China Taipéi invirtió 7,4 millones de USD en el desarrollo de su propio modelo básico, denominado Motor de diálogo con IA confiable (TAIDE por sus siglas en inglés), que utiliza los modelos de código abierto Llama de Meta (Creery, 2024^[62])⁹.

El uso gubernamental de la plataforma de código abierto Polis (Recuadro 5.36) representa un uso más restringido de IA de código abierto que los esfuerzos de TAIDE para desarrollar un modelo base. Una organización gubernamental que creó una versión personalizada y autodistribuida de Polis para una campaña de participación pública a gran escala, compuesta por 33 Polis polacos regionales y nacionales con 30 000 participaciones, incurrió en unos costos globales de alrededor de 422 500 EUR durante un período de 14 meses¹⁰. Estos gastos consistieron en 195 000 EUR para el desarrollo web experto, servicios en la nube y un sprint de diseño conjunto de experiencia de usuario (UX) subcontratado, y 227 500 EUR en tiempo de personal para implementar Polis en los flujos de trabajo existentes de la organización y para coordinar y llevar a cabo actividades de participación cívica. Desde entonces, la misma organización ha invertido 200 000 EUR en seguir mejorando su aplicación de Polis, dividido 50/50 entre costos técnicos y de personal. Este trabajo de mejora incluyó el diseño e implementación de la interfaz de usuario (UI) y otros trabajos de desarrollo técnico, que han sido de código abierto para otros usuarios de Polis. En general, la organización estima que basta con dos empleados a tiempo completo (ETC) para gestionar el trabajo, con la combinación de conocimientos técnicos necesarios (como desarrollo técnico, gestión de proyectos, diseño de UX y participación digital).

Los modelos de código abierto también se pueden utilizar junto con modelos privados. Por ejemplo, uno de los asistentes virtuales analizados en el Capítulo 5 utiliza tanto Gemini 1.5 Flash de Google como un

modelo Llama de Meta, que utiliza la interfaz de chat y la orquestación de los dos modelos desarrollados internamente mediante tecnologías de código abierto¹¹. El sistema se está pilotando con 18 000 usuarios, y los principales costos están asociados con el uso del LLM de Gemini y el alojamiento web en la nube para la aplicación de chat. Si bien la plataforma LLM le cuesta al gobierno alrededor de 18 000 EUR al mes, esperan una reducción sustancial ya que entienden cómo usar el modelo de manera más eficiente. El alojamiento web cuesta alrededor de 2 300 EUR al mes. En conjunto, estiman sus costos en 0,93-1,55 euros por usuario y mes. Los funcionarios están en conversaciones iniciales para ampliar el programa piloto en otros departamentos. El equipo de desarrollo y coordinación está formado por aproximadamente 10 ETC.

Desarrollo de modelos GenAI personalizados construidos y entrenados desde cero

El desarrollo de modelos GenAI personalizados y entrenados suele ser la opción más cara (para un rendimiento comparable) debido a la alta inversión inicial y la complejidad operativa. Los costos de capacitación de un LLM dependen del tamaño del modelo (los modelos más grandes con más parámetros requieren más potencia computacional y consumen más energía), la calidad y cantidad de datos de entrenamiento (lo que influye en el costo de adquisición y curación de los datos), las opciones de infraestructura (ya sea que la capacitación se realice en instalaciones o en la nube) y la eficiencia de los algoritmos de entrenamiento utilizados.

Las empresas de IA a menudo no divulgan públicamente los costos de formación asociados a sus modelos, aunque los investigadores estiman que los modelos populares actuales cuestan entre 41 y 192 millones de USD (Stanford HAI, 2025^[59]). Dado que el costo de la formación de los modelos más avanzados se multiplica por dos o por tres cada año, algunos estudios estiman que la formación de los modelos más grandes podría costar más de 1 000 millones EUR de aquí a 2027 (Cottier et al., 2024^[63]). Si bien estos costos pueden parecer altos, pueden palidecer en comparación con las importantes inversiones en investigación y desarrollo, los costos de personal y los esfuerzos de recopilación de datos necesarios para lograr los últimos modelos de base (Stanford HAI, 2025^[59]). Las empresas líderes en IA también tienen otros requisitos previos para desarrollar tales modelos, como un profundo talento técnico y, a menudo, asociaciones estratégicas con otras empresas.

Sin embargo, los gobiernos no necesariamente tienen que embarcarse en la construcción de sistemas tan extensos y poderosos que busquen vencer a los competidores del mercado. Por ejemplo, la formación desde cero de LLM financiados por el gobierno puede requerir menos personal y recursos de costos, especialmente para aquellos con menos parámetros o que buscan maximizar la relevancia para un país, región o idioma específicos. Por ejemplo, OpenEuroLLM tiene un presupuesto total de 37,4 millones de euros, lo que implica que una fracción de esta suma se dedicará al entrenamiento de su modelo de fundación con otra fracción para el personal (EC, 2025^[64])¹². En otro ejemplo, un país europeo desarrolló y entrenó un LLM en la lengua nacional, cuyo costo total, incluido el personal, ascendió a unos 500 000 EUR, de los cuales 300 000 EUR se destinaron al uso de la GPU.¹³ Otro ejemplo es GPT-NL en Países Bajos, que está invirtiendo unos 13,5 millones EUR proporcionados por el Ministerio de Asuntos Económicos y Clima de los Países Bajos (EZK) para formar un modelo (2023^[65]). Se han emprendido otros esfuerzos en Japón, Singapur, España, Suecia y los Emiratos Árabes Unidos (Chavez, 2024^[66]). El modelo de lenguaje multilingüe de acceso abierto de BigScience (BLOOM)¹⁴ desarrollado de forma colaborativa, contó con importantes contribuciones de organismos públicos. El proyecto fue liderado principalmente por Hugging Face y el Centro Nacional de Investigación Científica (CNRS) de Francia, con el apoyo de una subvención pública para computación en el supercomputador público francés «Jean Zay». El costo estimado del entrenamiento es de 2 a 5 millones de USD.

Infraestructuras tecnológicas heredadas y obsoletas

Las ambiciones de muchos gobiernos en materia de IA tardan en materializarse debido a que los sistemas informáticos heredados están obsoletos y no son adecuados para el desarrollo o uso de la IA, o son inadecuados para gestionar e intercambiar grandes cantidades de datos interoperables y de calidad (Irani et al., 2023^[67]). Estos sistemas pueden dar lugar a importantes oportunidades perdidas. Por ejemplo, solo en un área de atención en el Reino Unido, "los servicios financiados por los contribuyentes del NHS a los consejos locales se están perdiendo 45 000 millones de libras esterlinas en ahorros de productividad, más que suficiente para pagar por cada escuela primaria en el Reino Unido durante un año completo, porque con demasiada frecuencia dependen de tecnología antigua y obsoleta" (UK DSIT, 2025^[68]). El gobierno (2025^[21]) estima que el 28 % de los sistemas informáticos del gobierno central están obsoletos, alcanzando el 70 % en algunas organizaciones, y el 57 % de los funcionarios del gobierno del Reino Unido encuestados por la empresa de software SAS (2025^[53]) citó los sistemas heredados como una barrera para la adopción de IA. La cuestión también ha sido planteada por el UK PAC (2025^[6]) como un impedimento para el uso de IA en el gobierno.

En el Capítulo 5 se analiza cómo la tecnología obsoleta heredada afecta a la adopción de IA. Por ejemplo, el potencial de IA en la gestión financiera pública está limitado por los sistemas de información de gestión financiera obsoletos de los gobiernos de todo el mundo, que tienen más de una década de antigüedad en la mayoría de los países de la OCDE (Rivero del Paso et al., 2023^[69]; OECD, 2024^[70]). A pesar de la importancia del desafío, las consideraciones y el análisis de los efectos adversos de la tecnología heredada en la adopción de IA parecen ser menores en la mayoría de los países y las funciones gubernamentales. Si bien el párrafo anterior incluye detalles significativos sobre la escala del desafío en el Reino Unido para ilustrar el punto, esto se debe en gran medida a que la mayoría de los demás gobiernos no han realizado el análisis necesario para articular el problema de tal manera.

Este desafío depende de otros, como los importantes costos de financiar la reparación de los sistemas heredados. La tecnología obsoleta heredada también contribuye a otros desafíos, como los problemas de datos y una "dependencia excesiva de los contratistas para aumentar los costos", incluso para mantener sistemas obsoletos, con "el mantenimiento de los sistemas heredados que a menudo cuesta de tres a cuatro veces el de las alternativas modernas" (UK DSIT, 2025^[68]). Estos gastos podrían asignarse mejor en esfuerzos de innovación y modernización.

Los gobiernos están adoptando diversas medidas para modernizar sus sistemas con el fin de que estén más preparados para la IA. En un caso novedoso, el Departamento de Defensa de Estados Unidos está usando IA para modernizar el código heredado (Harper, 2024^[71]). Más tradicionalmente, algunos gobiernos están proporcionando financiamiento específico para los esfuerzos de modernización (véase el Capítulo 4, "Financiamiento de la IA y apoyo a inversiones coherentes en todo el gobierno").

Para superar los desafíos de implementación descritos en este capítulo, así como para mitigar los riesgos descritos en el Capítulo 1, los gobiernos pueden recurrir a las políticas públicas. De hecho, algunos gobiernos ya lo están haciendo. El siguiente capítulo examina las medidas de política que los gobiernos pueden adoptar —y algunas que ya están adoptando— para ofrecer una IA fiable y aprovechar todo su potencial.

Bibliografía

- Austin, T. et al. (2024), *A snapshot of how public sector leaders feel about generative AI*, [7]
<https://www2.deloitte.com/us/en/insights/industry/public-sector/ai-adoption-in-public-sector.html>.
- Australia DTA (2024), *Evaluation of whole-of-government trial into generative AI: Now available*, [57]
<https://www.dta.gov.au/blogs/evaluation-whole-government-trial-generative-ai-now-available>.
- Australia Treasury (2025), *Evaluation of a trial of generative AI (Copilot) in The Treasury*, [58]
<https://evaluation.treasury.gov.au/publications/evaluation-generative-artificial-intelligence>.
- Autio, C., K. Communig and B. Elliott (2023), *A Snapshot of Artificial Intelligence Procurement Challenges*, <https://files.thegovlab.org/a-snapshot-of-ai-procurement-challenges-june2023.pdf>. [20]
- Barnett, J. (2020), *JAIC awards biggest contract yet to buy AI for the battlefield*, [60]
<https://fedscoop.com/battlefield-ai-jaic-booz-allen>.
- Barrett, K. and R. Greene (2024), *The Future of AI For the Public Sector: The Challenges and Solutions*, [52]
<https://www.businessofgovernment.org/blog/future-ai-public-sector-challenges-and-solutions>.
- Brizuela, A. et al. (2025), *Analysis of the generative AI landscape in the European public sector*, [2]
European Commission, <https://op.europa.eu/s/z4XY>.
- Chavez, P. (2024), *Sovereign AI in a Hybrid World: National Strategies and Policy Responses*, [66]
<https://www.lawfaremedia.org/article/sovereign-ai-in-a-hybrid-world--national-strategies-and-policy-responses>.
- Cheng, L. and A. Chouldechova (2023), “Overcoming Algorithm Aversion: A Comparison between Process and Outcome Control”, *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1-27, <https://doi.org/10.1145/3544548.3581253>. [41]
- Cottier, B. et al. (2024), *The rising costs of training frontier AI models*, [63]
<https://arxiv.org/abs/2405.21015>.
- Creery, J. (2024), *Taiwan Builds Own AI Language Model to Counter China’s Influence*, [62]
<https://www.bloomberg.com/news/articles/2024-01-25/taiwan-builds-own-ai-language-model-to-counter-china-s-influence?>.
- de Mello, L. and T. Ter-Minassian (2020), “Digitalisation challenges and opportunities for subnational governments”, *OECD Working Papers on Fiscal Federalism*, No. 31, OECD Publishing, Paris, <https://doi.org/10.1787/9582594a-en>. [17]
- Desouza, K. (2018), *Delivering Artificial Intelligence in Government: Challenges and opportunities*, [34]
<https://www.businessofgovernment.org/sites/default/files/Delivering%20Artificial%20Intelligence%20in%20Government.pdf>.
- Dietvorst, B., J. Simmons and C. Massey (2018), “Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them”, *Management Science*, Vol. 64/3, pp. 1155-1170, <https://doi.org/10.1287/mnsc.2016.2643>. [44]

- EC (2025), *A pioneering AI project awarded for opening Large Language Models to European languages*, <https://digital-strategy.ec.europa.eu/en/news/pioneering-ai-project-awarded-opening-large-language-models-european-languages>. [64]
- EC (2024), *Adoption of AI, blockchain and other emerging technologies within the European public sector – A public sector Tech Watch report*, Publications Office of the European Union, <https://data.europa.eu/doi/10.2799/3438251>. [8]
- EC (2024), *Public Sector Tech Watch latest dataset of selected cases*, <http://data.europa.eu/89h/e8e7bdd-8510-4936-9fa6-7e1b399cbd92> (accessed on 4 March 2025). [9]
- Fagan, M. (2024), *AI for the People: Use Cases for Government*, https://www.hks.harvard.edu/sites/default/files/centers/mrcbg/working_papers/M-RCBG%20Working%20Paper%202024-02_AI%20for%20the%20People.pdf. [38]
- Featherson, R., A. Shlonsky and C. Lewis (2019), “Interventions to Mitigate Bias in Social Work Decision-Making: A Systematic Review”, *Research on Social Work Practice*, Vol. 29/7, <https://doi.org/10.1177/1049731518819160>. [43]
- France Élysée (2025), *The Paris Charter on Artificial Intelligence in the Public Interest*, <https://www.elysee.fr/en/emmanuel-macron/2025/02/11/the-paris-charter-on-artificial-intelligence-in-the-public-interest>. [54]
- Frontier economics (2024), *Guidance on the impact of AI interventions*, <https://www.frontier-economics.com/uk/en/news-and-insights/news/news-article-i21121-analysing-the-impact-of-ai-interventions-in-government>. [47]
- Government of the Netherlands (2023), *The Netherlands is building its own open language model GPT-NL*, <https://www.digitaleoverheid.nl/nieuws/nederland-bouwt-eigen-open-taalmodel-gpt-nl/>. [65]
- Harper, J. (2024), *Pentagon using AI to modernize legacy code*, <https://defensescoop.com/2024/09/12/pentagon-artificial-intelligence-modernize-legacy-code-john-hale/>. [71]
- Hodges, D. (2024), *Fumbles can't kill the government's AI appetite*, <https://www.themandarin.com.au/249756-red-faces-and-fumbles-cant-kill-governments-ai-appetite/>. [37]
- Irani, Z. et al. (2023), “The impact of legacy systems on digital transformation in European public administration: Lesson learned from a multi case analysis”, *Government Information Quarterly*, Vol. 40/1, p. 101784, <https://doi.org/10.1016/j.giq.2022.101784>. [67]
- Jackson, B. (2025), *South Australian drivers to be monitored by AI cameras*, <https://www.news.com.au/technology/south-australian-drivers-to-be-monitored-by-ai-cameras/news-story/8a63be3e80d4bba60735d58ec8c473db>. [61]
- Mariani, J., W. Eggers and P. Kishnani (2023), *The AI regulations that aren't being talked about*, <https://www2.deloitte.com/us/en/insights/industry/public-sector/ai-regulations-around-the-world.html>. [39]
- Mitchell, S. (2025), *Skills gap in public sector IT fuels outsourcing reliance*, <https://itbrief.co.uk/story/skills-gap-in-public-sector-it-fuels-outsourcing-reliance>. [19]

- Monteiro, B., A. Hlacs and P. Boéchat (2024), "Public procurement for public sector innovation: Facilitating innovators' access to innovation procurement", *OECD Working Papers on Public Governance*, No. 80, OECD Publishing, Paris, <https://doi.org/10.1787/9aad76b7-en>. [27]
- Morley, J. et al. (2019), "From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices", *Science and Engineering Ethics*, Vol. 26/4, pp. 2141-2168, <https://doi.org/10.1007/s11948-019-00165-5>. [29]
- Muñoz-Cadena, S. et al. (2025), *Sistemas de IA en el sector público de América Latina y el Caribe (Versión V2)*, <https://sistemaspublicos.tech/sistemas-de-ia-en-america-latina/> (accessed on 29 April 2025). [10]
- NASCIO (2024), *Generative Artificial Intelligence and its Impact on State Government IT Workforces*, National Association of State Chief Information Officers, <https://www.nascio.org/resource-center/resources/generative-artificial-intelligence-and-its-impact-on-state-government-it-workforces/>. [15]
- OECD (2024), "2023 OECD Digital Government Index: Results and key findings", *OECD Public Governance Policy Papers*, No. 44, OECD Publishing, Paris, <https://doi.org/10.1787/1a89ed5e-en>. [24]
- OECD (2024), "Assessing potential future artificial intelligence risks, benefits and policy imperatives", *OECD Artificial Intelligence Papers*, No. 27, OECD Publishing, Paris, <https://doi.org/10.1787/3f4e3dfb-en>. [31]
- OECD (2024), "Financial Management Information Systems in OECD countries", *OECD Papers on Budgeting*, No. 2024/02, OECD Publishing, Paris, <https://doi.org/10.1787/ce8367cd-en>. [70]
- OECD (2024), *Fixing Frictions: 'Sludge audits' around the world*, OECD Publishing, <https://doi.org/10.1787/14e1c5e8-en-fr>. [56]
- OECD (2024), "Governing with Artificial Intelligence: Are governments ready?", *OECD Artificial Intelligence Papers*, No. 20, OECD Publishing, Paris, <https://doi.org/10.1787/26324bc2-en>. [30]
- OECD (2024), *Shaping smart cities of all sizes*, OECD Publishing, https://www.oecd.org/content/dam/oecd/en/about/programmes/cfe/the-oecd-programme-on-smart-cities-and-inclusive-growth/Proceedings-4th-Roundtable-Smart-Cities-Inclusive-Growth.pdf/_jcr_content/renditions/original./Proceedings-4th-Roundtable-Smart-Cities-In. [4]
- OECD (2023), "2023 OECD Open, Useful and Re-usable data (OURdata) Index: Results and key findings", *OECD Public Governance Policy Papers*, No. 43, OECD Publishing, Paris, <https://doi.org/10.1787/a37f51c3-en>. [23]
- OECD (2021), "The OECD Framework for digital talent and skills in the public sector", *OECD Working Papers on Public Governance*, No. 45, OECD Publishing, Paris, <https://doi.org/10.1787/4e7c3f58-en>. [32]
- OECD (2019), *The Path to Becoming a Data-Driven Public Sector*, OECD Digital Government Studies, OECD Publishing, Paris, <https://doi.org/10.1787/059814a7-en>. [25]
- OECD (2017), *Fostering Innovation in the Public Sector*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264270879-en>. [3]

- OECD/CAF (2022), *The Strategic and Responsible Use of Artificial Intelligence in the Public Sector of Latin America and the Caribbean*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/1f334543-en>. [1]
- OECD/UNESCO (2024), *G7 Toolkit for Artificial Intelligence in the Public Sector*, OECD Publishing, Paris, <https://doi.org/10.1787/421c1244-en>. [11]
- Parliament of Australia (2025), *Report 510: Inquiry into the use and governance of artificial intelligence systems by public sector entities - 'Proceed with Caution'*, https://parlinfo.aph.gov.au/parlInfo/download/committees/reportJnt/RB000567/toc_pdf/Report5_10Inquiryintotheuseandgovernanceofartificialintelligencesystemsbypublicsectorentities-'ProceedwithCaution'.pdf. [28]
- PTI (2024), *AI and City/County Government: Survey Results*, Public Technology Institute, <https://fusionlp.org/wp-content/uploads/2024/11/AI-Survey-City-and-County-Final-2024.pdf>. [16]
- Richter, A. (2024), *Navigating Generative AI in Government*, <https://businessofgovernment.org/report/navigating-generative-ai-government>. [36]
- Rivero del Paso, L. et al. (2023), *Digital Solutions Guidelines for Public Financial Management*, <https://www.imf.org/en/Publications/TNM/Issues/2023/10/06/Digital-Solutions-Guidelines-for-Public-Financial-Management-537781>. [69]
- Ryseff, J., B. De Bruhl and S. Newberry (2024), *The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed*, RAND, https://www.rand.org/pubs/research_reports/RRA2680-1.html. [5]
- Ryseff, J. and A. Narayanan (2025), *Why AI Projects Fail*, <https://www.rand.org/pubs/presentations/PTA2680-1.html>. [22]
- Salesforce (2024), *6 in 10 IT Workers Report Shortage of AI Skills in the Public Sector*, <https://www.salesforce.com/news/stories/public-sector-ai-statistics/>. [12]
- Samuleson, W. and W. Zeckhauser (1988), "Status quo bias in decision making", *Journal of Risk and Uncertainty*, Vol. 1, pp. 7-59, <https://doi.org/10.1007/bf00055564>. [50]
- SAS (2025), *Slow uptake of AI in government hindering strategic goals, new research finds*, https://www.sas.com/en_gb/news/press-releases/2024/september/slow-uptake-of-ai-in-government-hindering-strategic-goals.html. [53]
- SAS (2020), *AI in government: The path to adoption and deployment*, https://www.sas.com/en_sa/insights/articles/analytics/ai-in-government.html. [35]
- Shahab, S. and L. Lades (2021), "Sludge and transaction costs", *Behavioral Public Policy*, Vol. 1/22, <https://doi.org/10.1017/bpp.2021.12>. [55]
- Shark, A. (2025), *What the Rising Costs of AI Means for Government*, <https://statetechmagazine.com/article/2025/01/what-rising-costs-ai-means-government>. [51]
- Stanford HAI (2025), *Artificial Intelligence Index Report 2025*, https://hai-production.s3.amazonaws.com/files/hai_ai_index_report_2025.pdf. [59]
- Sunstein, C. and J. Gaffe (2024), *An Anatomy of Algorithm Aversion*, Elsevier BV, <https://doi.org/10.2139/ssrn.4865492>. [42]

- Teale, C. (2025), *Public-sector concerns over AI are lessening, survey says*, <https://www.route-fifty.com/artificial-intelligence/2025/02/public-sector-concerns-over-ai-are-lessening-survey-says/403328>. [45]
- The Alan Turing Institute (2024), *AI for bureaucratic productivity: Measuring the potential of AI to help automate 143 million UK government transactions*, <https://www.turing.ac.uk/news/publications/ai-bureaucratic-productivity-measuring-potential-ai-help-automate-143-million-uk>. [46]
- Trajkovski, G. (2024), “Bridging the public administration-AI divide: A skills perspective”, *Public Administration and Development*, Vol. 44/5, pp. 412-426, <https://doi.org/10.1002/pad.2061>. [18]
- Ubaldi, B. et al. (2019), “State of the art in the use of emerging technologies in the public sector”, *OECD Working Papers on Public Governance*, No. 31, OECD Publishing, Paris, <https://doi.org/10.1787/932780bc-en>. [33]
- UK Committee of Public Accounts (2025), *Use of AI in Government*, <https://committees.parliament.uk/publications/47199/documents/244683/default/>. [6]
- UK DSIT (2025), *Archaic tech sees public sector miss £45 billion annual savings*, <https://www.gov.uk/government/news/archaic-tech-sees-public-sector-miss-45-billion-annual-savings>. [68]
- UK DSIT (2025), *State of digital government review*, <https://www.gov.uk/government/publications/state-of-digital-government-review/state-of-digital-government-review>. [21]
- UK NAO (2024), *Use of artificial intelligence in government*, National Audit Office, <https://www.nao.org.uk/wp-content/uploads/2024/03/use-of-artificial-intelligence-in-government.pdf>. [13]
- UN Habitat (2024), *Global assessment of Responsible AI in cities*, https://unhabitat.org/sites/default/files/2024/08/global_assessment_of_responsible_ai_in_cities_21082024.pdf. [14]
- US OMB (2025), *Accelerating Federal Use of AI through Innovation, Governance, and Public Trust*, <https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-21-Accelerating-Federal-Use-of-AI-through-Innovation-Governance-and-Public-Trust.pdf>. [40]
- US OMB (2025), *Driving Efficient Acquisition of Artificial Intelligence in Government*, White House Office of Management and Budget, <https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-22-Driving-Efficient-Acquisition-of-Artificial-Intelligence-in-Government.pdf>. [48]
- van Noordt, C., R. Medaglia and L. Tangi (2023), “Policy initiatives for Artificial Intelligence-enabled government: An analysis of national strategies in Europe”, *Public Policy and Administration*, <https://doi.org/10.1177/09520767231198411>. [26]
- Varazzani, C. et al. (2023), “Seven routes to experimentation in policymaking: A guide to applied behavioural science methods”, *OECD Working Papers on Public Governance*, No. 64, OECD Publishing, Paris, <https://doi.org/10.1787/918b6a04-en>. [49]
- Waters, R. (2024), *Meta under fire for ‘polluting’ open-source*, <https://www.ft.com/content/397c50d8-8796-4042-a814-0ac2c068361f>. [72]

Notas

¹ "Salesforce realizó una encuesta doble anónima a 600 profesionales de TI (200 líderes de TI y 400 colaboradores individuales de TI) en Australia, Francia, Alemania, el Reino Unido y Estados Unidos. Los encuestados trabajan en diferentes sectores, como el tecnológico, el de servicios financieros, el de medios de comunicación y entretenimiento, el de manufactura, el minorista, el sanitario, el sector público, etc. La encuesta se presentó en diciembre de 2023 y enero de 2024". (2024^[12]).

² La fuente de esta frase es un informe del Comité Mixto de Cuentas Públicas y Auditoría, Parlamento de Australia (2025^[28]). Las conclusiones del Comité y de la Comisión no son necesariamente representativas de las opiniones del Gobierno australiano.

³ <https://oecd.ai/dashboards>.

⁴ <https://impress.ai/case-studies/publicsector>.

⁵ Basado en el sitio web orientado a Estados Unidos de Microsoft (<https://www.microsoft.com/en-us/microsoft-365/copilot>) a 10 de abril de 2024. Precios de suscripción anual.

⁶ <https://team-gpt.com/blog/clause-pricing>.

⁷ Véanse, por ejemplo, <https://openai.com/global-affairs/introducing-chatgpt-gov> y <https://www.anthropic.com/news/expanding-access-to-claude-for-government>.

⁸ El uso de modelos de "código abierto" para este informe no implica que dichos modelos se lancen bajo una licencia de código abierto aprobada por la Open Source Initiative (OSI), una organización sin fines de lucro administradora de The Open Source Definition (<https://opensource.org/osd>). OSI ha criticado a algunas compañías que llaman a sus modelos código abierto porque solo proporcionan los pesos para el modelo, y no otros elementos, como los datos de entrenamiento, código y prácticas de entrenamiento (Waters, 2024^[72]). Algunos argumentan que estos modelos deberían llamarse "peso abierto" en lugar de "código abierto".

⁹ <https://en.taide.tw>.

¹⁰ Información proporcionada a la OCDE por un país que no ha sido divulgado. La OCDE no publica el nombre del país o proyecto debido a la naturaleza preliminar de las estimaciones y los análisis.

¹¹ Información proporcionada a la OCDE por un país que no ha sido divulgado. La OCDE no publica el nombre del país o proyecto debido a la naturaleza preliminar de las estimaciones y los análisis.

¹² <https://openeurollm.eu>.

¹³ Cifras comunicadas a la OCDE por un país no divulgado.

¹⁴ <https://huggingface.co/bigscience/bloom>.

4

Habilitadores, salvaguardas y participación para desbloquear una IA confiable

Este capítulo expone cómo los gobiernos pueden aprovechar el potencial de la IA al tiempo que gestionan los riesgos. Establece tres pilares — habilitadores, salvaguardas y participación— que juntos conforman el Marco de la OCDE para una IA confiable en el sector público. Los habilitadores incluyen la gobernanza, los datos, la infraestructura digital, las capacidades y el talento, las inversiones con objetivos definidos, la contratación pública y las alianzas; las salvaguardas abarcan instrumentos no vinculantes y vinculantes, la transparencia y la gestión de riesgos, y la supervisión; la participación comprende a la ciudadanía, los servidores públicos y la colaboración transfronteriza. El capítulo aboga por un enfoque sistémico, una aplicación proporcionada y basada en el riesgo de las medidas, y mecanismos prácticos como la experimentación, la evaluación de impacto y las auditorías para apoyar una adopción confiable a gran escala.

Mensajes clave

- Las iniciativas en materia de gobernanza y política pública pueden ayudar a las administraciones públicas a aprovechar al máximo el potencial de la IA y abordar sus diversos riesgos y desafíos de implementación. Los gobiernos deben adoptar un enfoque sistémico y procurar anticiparse a los cambios futuros. Las medidas propuestas deben fortalecer los habilitadores, establecer salvaguardas y promover la participación con actores clave.
- Los habilitadores incluyen mecanismos y procesos clave de gobernanza; una comprensión clara del papel de los datos como base de la IA; el desarrollo de infraestructuras digitales; el fortalecimiento de capacidades y talento; inversiones en tecnología orientadas a objetivos definidos; el uso eficaz de la contratación pública; y la ampliación del potencial de la IA mediante alianzas.
- Las salvaguardas pueden ser instrumentos de política no vinculantes y vinculantes; procesos de transparencia; y mecanismos de rendición de cuentas.
- La participación con actores clave puede adoptar la forma de asambleas ciudadanas, interacción con servidores públicos, involucrar a las personas usuarias en el desarrollo de la IA y colaboración transfronteriza.
- En conjunto, estas medidas de política conforman un marco para una IA confiable que puede ayudarle al sector público a alinear sus acciones con los Principios de IA de la OCDE. La labor futura de la OCDE abordará elementos del marco con mayor profundidad.

Medidas de política para liberar el potencial de la IA

Mientras que el resto de este informe analiza las oportunidades y los desafíos de los gobiernos a la hora de *gobernar con IA*, este capítulo se centra en cómo *gobernar la IA* exitosamente en el sector público. Para aprovechar al máximo el potencial de la IA y, al mismo tiempo, mitigar los riesgos, los gobiernos deben adoptar un enfoque intencionado, estratégico y responsable. Este enfoque debería estar en consonancia con los Principios de IA de la OCDE, pero ser específico al contexto y adecuado para el desarrollo y uso de la IA por y para las administraciones públicas. En particular, las administraciones deben seguir tres líneas de acción:

- Fortalecer los **habilitadores** (por ejemplo, datos de calidad, capacidades digitales y de IA, financiamiento e infraestructura digital) para superar los principales desafíos en la aplicación y obtener los resultados esperados
- Establecer **salvaguardas** (por ejemplo, herramientas de transparencia, rendición de cuentas y gestión de riesgos) para anticipar y gestionar los riesgos asociados, y
- Fomentar la **participación** con las actores clave (incluida la ciudadanía) para desarrollar sistemas de IA que tengan en cuenta sus necesidades.

El objetivo de estas acciones es aprovechar las oportunidades y abordar los diversos riesgos y desafíos de implementación asociados a la IA mediante políticas públicas específicas. Por ejemplo, las cuestiones relacionadas con la necesidad de contar con datos suficientes y de calidad se abordan a través de habilitadores centrados en construir una infraestructura y gobernanza de datos sólidas. En otro ejemplo, la orientación insuficiente y las regulaciones obsoletas pueden abordarse a través de salvaguardas como instrumentos de política vinculantes y no vinculantes, incluidos instrumentos regulatorios ágiles, que ayuden a garantizar que la IA funcione dentro de límites éticos, operativos y legales claramente establecidos. La participación activa de las actores clave a lo largo del ciclo de vida del sistema de IA

(desarrollo, despliegue y uso de tecnologías de IA) y del ciclo de políticas (para diseñar, implementar y evaluar la gobernanza y las políticas de IA) puede enriquecer la comprensión, las actitudes y los comportamientos de las partes interesadas, incluida la ciudadanía, y alinear los avances tecnológicos y de gobernanza con las necesidades de la sociedad.

Las secciones siguientes abordan los principales puntos de atención para que los gobiernos creen un entorno que permita el uso estratégico y responsable de la IA en todos los sistemas y funciones gubernamentales. Proporcionan un análisis exhaustivo de acciones y prioridades específicas en los tres ámbitos. En cada sección se describen las principales opciones de política pública que los gobiernos deben tener en cuenta para adoptar un enfoque coherente y sostenible. Los gobiernos deben considerar estas opciones a la luz de su propio contexto, incluida la madurez del gobierno digital. Por ejemplo, en lugar de intentar poner en marcha todos los temas debatidos a la vez, los gobiernos podrían considerar hojas de ruta progresivas de gobernanza y adopción de la IA basadas en capacidades institucionales, culturales y tecnológicas.

Un enfoque holístico basado en sistemas puede maximizar el valor de la IA en el sector público

Al establecer los habilitadores, salvaguardas y mecanismos de participación analizados en este capítulo, los gobiernos deben adoptar un enfoque sistémico de la IA y tratar de comprender y abordar los problemas públicos como parte de un sistema más amplio e interconectado, en lugar de considerarlos de forma aislada (OECD, 2017^[1]). Tradicionalmente, los responsables de las políticas públicas han abordado los problemas sociales a través de intervenciones discretas superpuestas, construyendo una relación de "causa y efecto". Sin embargo, estas intervenciones pueden trasladar las consecuencias de una parte del sistema a otra, simplemente abordando los síntomas e ignorando las causas. La IA representa una oportunidad para reimaginar el funcionamiento del gobierno, tanto en operaciones internas como de servicios públicos. Los gobiernos deben pensar más allá de cómo la IA puede encajar en los sistemas y estructuras gubernamentales existentes; deben pensar en cómo la IA puede contribuir a repensar por completo los procesos y sistemas; de lo contrario, corren el riesgo de simplemente automatizar la ineficiencia y reforzar aún más los incentivos y los enfoques de gobernanza desalineados. Las prácticas emergentes pueden ayudar, como las "auditorías de fricción", que son evaluaciones conductuales estructuradas para identificar fricciones que reducen la probabilidad de que las personas completen un proceso o realicen un esfuerzo psicológico indebido (OECD, 2024^[2]). La OCDE ha creado una línea de trabajo dedicada a enfoques sistémicos que pueden ser de gran ayuda¹.

Reconocer el potencial de los cambios futuros y tratar de anticiparse a ellos

Todavía hay mucho que aprender sobre la IA y mucho se desconoce sobre su evolución en curso. El establecimiento de habilitadores, salvaguardas y mecanismos de participación que se tratan en este capítulo solo puede tener en cuenta los conocimientos actuales, lo que se sabe sobre los usos y las implicancias de la IA actual y sobre el potencial para el futuro. Sin embargo, existen importantes incógnitas que solo se resolverán con el tiempo a medida que la tecnología se desarrolle y se exploren usos potenciales. La IA en el gobierno será un proceso continuo de descubrimiento, con desarrollos bienvenidos y no bienvenidos, y evoluciones inesperadas o no deseadas. Los gobiernos deben adoptar un enfoque ágil y adaptable para adaptarse a las nuevas oportunidades y a los cambios de comportamiento. Muchas tareas que hoy no puede abordar la IA probablemente se vuelvan factibles en el futuro. Las estrategias y los marcos de IA deben ser lo suficientemente flexibles como para evolucionar con las capacidades y los contextos cambiantes. Los gobiernos deben mejorar su participación temprana con señales débiles que indiquen cómo puede transcurrir el futuro. Esto les permitirá comprender mejor dónde y cuándo intervenir, sin esperar a que los procesos y las tendencias se consoliden y, por lo tanto, resulten costosos y difíciles de modificar.

Estas consideraciones son pertinentes no solo para la gobernanza de la IA, sino también para su uso en la administración pública. Los casos de uso analizados en este informe (que se analizan con mayor profundidad en el Capítulo 5 y se sintetizan en el Capítulo 2) suelen representar mejoras graduales y aumentos de productividad. Sin embargo, esto no debería ocultar que los usos emergentes o futuros de la IA podrían ser completamente nuevos o permitir tareas antes imposibles, poco prácticas o incluso impensables, generando nuevas oportunidades y riesgos para el gobierno. Esto puede dar lugar a casos de uso y enfoques de gobernanza que tendrán que crearse, reinventarse o detenerse. Los esfuerzos de la OCDE en materia de gobernanza anticipatoria de la innovación (AIG) y prospectiva estratégica pueden ayudar a los gobiernos a comprender y configurar mejor los futuros potenciales de la IA².

Fortalecimiento de habilitadores para apoyar la adopción de una IA confiable

Los habilitadores son elementos y recursos fundamentales necesarios para la implementación de la IA en el gobierno³. Crean un entorno en el que los servidores públicos cualificados pueden diseñar e implementar la IA de forma eficaz y confiable. Su apoyo práctico permite a las instituciones gubernamentales aprovechar plenamente el potencial de la IA. En las siguientes secciones se examinan siete habilitadores claves: **gobernanza, datos (incluidos datos abiertos gubernamentales), infraestructura digital, competencias y talento, inversiones en IA, contratación pública y alianzas con actores no gubernamentales**. Estos fueron definidos inicialmente por la OCDE en 2024 ([3]) y se desarrollan en mayor medida en las secciones siguientes. Cada sección analiza las opciones de política que pueden adoptar los gobiernos para implementar estos habilitadores en sus contextos, basándose en las mejores prácticas internacionales.

Establecimiento de mecanismos y procesos clave de gobernanza

Los gobiernos están acelerando la adopción de la IA y, en la mayoría de los casos, han incluido sus objetivos en las estrategias nacionales de IA. Sin embargo, en la práctica, esta adopción suele ser fragmentada a nivel estratégico, sin establecer mecanismos de gobernanza integrales y sólidos que garanticen el uso responsable, el impacto a largo plazo y la sostenibilidad de la IA. A medida que la IA se vaya integrando cada vez más en las operaciones gubernamentales, deben establecerse a nivel nacional mecanismos de gobernanza sólidos para garantizar un uso confiable, sostenible y eficaz de la IA. También pueden ofrecer una visión clara de los beneficios de la IA para recabar apoyo dentro y fuera del gobierno.

Un liderazgo sólido para lograr una visión cohesionada

Un liderazgo sólido es un factor crítico para lograr la adopción de la IA en el gobierno. Es fundamental para establecer el tono correcto por parte de los más altos niveles de gobierno y comunicar activamente los beneficios potenciales de la IA (Berryhill et al., 2019[4]). Si bien es fundamental establecer estrategias y principios que ayuden a garantizar una adopción confiable de la IA, un liderazgo sólido y eficaz puede crear una visión coherente de IA y establecer un "tono en la cima" que genere confianza en la IA, tanto dentro como fuera del gobierno. Por ejemplo, en el Reino Unido, los últimos dos primeros ministros, Rishi Sunak y Keir Starmer, han defendido la adopción de la IA tanto dentro como fuera del gobierno. Bajo el liderazgo de Sunak, el Reino Unido catalizó la atención mundial y la colaboración internacional en materia de IA al convocar la primera Cumbre Mundial sobre IA en noviembre de 2023, con eventos posteriores organizados por Corea y Francia. Cada una de ellas ha concluido con una declaración firmada por muchos gobiernos en la que se describen los compromisos relacionados con la IA⁴. Durante el mandato de Sunak, también se puso en marcha la Incubadora de IA (i.AI) del Reino Unido, más enfocada en el gobierno y cuyo objetivo es mejorar vidas, impulsar el crecimiento y ofrecer mejores servicios públicos. Bajo el actual liderazgo del primer ministro Starmer, el Reino Unido (2025[5]) lanzó el manual de IA para el Gobierno del Reino Unido (Recuadro 4.2) y presentó un plan ambicioso para usar la IA en "remodelar el estado para

que funcione para los trabajadores", incluso mediante la creación de 2 000 aprendices de tecnología de IA en el gobierno. Por último, con el fin de garantizar un liderazgo sólido en todas las organizaciones gubernamentales del Reino Unido, (2025^[6]) "el plan para un gobierno digital moderno" exige que todas las organizaciones del sector público incluyan a un líder digital en su comité ejecutivo para 2026.

Quienes están en la cima tienen el poder de establecer una dirección estratégica que pueda permear los niveles inferiores, ayudando a enmarcar el uso de la IA dentro de la cultura en general. Como se afirma en el *Marco de la OCDE para el talento y las competencias digitales en el sector público* (2021^[7]). Por último, "el liderazgo que cree un entorno propicio para fomentar la transformación digital transmitirá una visión clara del gobierno digital y promoverá activamente sus beneficios. [Estos] líderes estarán comprometidos, serán visibles y accesibles, y empoderarán a sus equipos a través de la descentralización de la toma de decisiones"⁵. Un estudio de la Comisión Europea (CE) (2024^[8]), basado en una encuesta a 576 gestores públicos de siete países, concluyó que el liderazgo puede influir especialmente en la adopción de la IA ofreciendo incentivos o recursos financieros sólidos para poner en marcha iniciativas de IA, y que, en general, los encuestados consideran que el estado actual de estas es insatisfactorio.

Un liderazgo sólido en IA puede fomentar un enfoque "orientado a misiones" para su innovación. Este enfoque hace hincapié en la resolución de problemas, en la que las intervenciones en materia de políticas están diseñadas para movilizar recursos, coordinar a los actores clave y estimular la innovación y la colaboración entre gobiernos y sectores para abordar los desafíos identificados y cumplir con los objetivos establecidos de la misión. Las políticas orientadas a misiones suelen conllevar una combinación de medidas regulatorias, incentivos financieros, financiamiento de la investigación e inversiones específicas para impulsar el progreso hacia la misión. Los líderes juegan un papel crítico al proporcionar una dirección descendente y galvanizar el apoyo para alinear todas las piezas del gobierno para moverse al unísono hacia el mismo objetivo (OECD, 2021^[9])⁶. Algunos gobiernos y organizaciones intergubernamentales han adoptado un enfoque de política de IA orientado a misiones, aunque por lo general su objetivo es catalizar el crecimiento económico en el sector privado (UCL IIPP, 2019^[10]; Vinnova, 2022^[11]).

Adoptar un enfoque estratégico y dirigido

Los gobiernos pueden implementar estrategias y directrices gubernamentales holísticas sobre IA para identificar y priorizar su uso y desarrollo coherentes, en consonancia con los valores y objetivos generales del gobierno. Por ejemplo, la estrategia nacional de IA de República Dominicana de 2023 (2024^[12]) se centra principalmente en la IA en el gobierno. Canadá (2025^[13]), Suiza (2025^[14]) y Uruguay (2021^[15]) han desarrollado una estrategia específica para la IA en el gobierno, y otra está en desarrollo en el Reino Unido (2024^[16]). El memorándum de abril de 2025 de Estados Unidos (EE.UU.) sobre "Aceleración del uso federal de la IA a través de la innovación, la gobernanza y la confianza pública", aunque no se denomina "estrategia", se ajusta a muchas de las características de una estrategia que busca impulsar el cambio en todo el gobierno federal y tiene el beneficio añadido de ser de carácter vinculante (Recuadro 4.1). Muchas otras iniciativas tienen objetivos sustantivos para que la IA en el gobierno se integre en estrategias nacionales más amplias. En general, estas estrategias de alto nivel tienden a afectar a muchos de los elementos habilitadores, las salvaguardas, los procesos de participación y los tipos de casos de uso analizados en este informe. También existen estrategias específicas para orientar los esfuerzos de IA en determinadas funciones gubernamentales. Por ejemplo, Francia ha desarrollado una estrategia para el uso de la IA en los RR.HH (véase el Recuadro 5.19 del Capítulo 5), y la política estadounidense mencionada anteriormente concede a los organismos gubernamentales 180 días para desarrollar y publicar su propia estrategia de IA.

Recuadro 4.1. Acelerar el uso de IA por parte del gobierno federal en Estados Unidos

En Estados Unidos, de conformidad con los requisitos de su política sobre IA de 2020, el 3 de abril de 2025, la Oficina de Gestión y Presupuesto de la Casa Blanca (OMB) emitió el memorándum *M-25-21 Accelerating Federal Use of AI through Innovation, Governance, and Public Trust* (Acelerar el uso federal de la IA a través de la innovación, la gobernanza y la confianza pública). Este último promueve la innovación en IA, la adopción y el uso responsable de la IA y la salvaguardar las protecciones estadounidenses sobre la privacidad, los derechos civiles y las libertades civiles. Según el memorándum entre otras cosas, las agencias deben:

- Designar un director de IA (CAIO) que actúe como asesor senior, promueva los objetivos de la IA de la agencia, coordine los esfuerzos de IA dentro de su agencia y represente a la agencia ante los órganos de coordinación y los foros externos; la OMB se compromete a convocar un Consejo Interinstitucional de la CAIO para apoyar la coordinación con el fin de maximizar la eficiencia.
- Rendir cuentas cumpliendo los requisitos de presentación de informes, incluida la actualización de un inventario de casos de uso de IA al menos una vez al año.
- Implementar prácticas mínimas de gestión de riesgos para los sistemas de IA que puedan tener impactos significativos cuando se desplieguen ("IA de alto impacto") de una manera proporcional al riesgo anticipado de su uso previsto, como se describe con más detalle (Recuadro 1.3).
- Publicar estrategias de IA de las agencias para identificar y eliminar barreras al uso responsable de la IA y lograr mejoras en toda la empresa en la madurez de sus aplicaciones (CFO Act agencies).
 - Las estrategias deben incluir una evaluación del estado actual de madurez de la IA de la agencia y un plan para lograr sus objetivos de madurez, abordando, entre otras cosas, planes y procesos para desarrollar infraestructuras que permitan la IA (por ejemplo, infraestructuras informáticas de alto rendimiento) a lo largo del ciclo de vida de la IA; garantizar el acceso a datos de calidad; desarrollar la capacidad de la empresa para la innovación en IA; reclutar, contratar, formar, retener y empoderar a una organización lista para la IA y lograr la alfabetización en IA para los no profesionales que participan en la IA; y desarrollar las operaciones, la gobernanza y la infraestructura necesarias para gestionar los riesgos derivados del uso de la IA.
- Desarrollar una política de IA generativa que establezca las condiciones para el uso aceptable de la IA generativa en sus misiones y establezca salvaguardias y mecanismos de supervisión adecuados que permitan a la IA generativa ser utilizada en la agencia sin plantear riesgos indebidos.
- Compartir proactivamente datos y activos de IA en todo el gobierno federal, incluido el código desarrollado a medida, incluidos los modelos, tanto si la agencia los ha desarrollado como si los ha adquirido, y, en la medida de lo posible, publicar y mantener el código como software de código abierto en un archivo público (se aplican algunas excepciones).
- Desarrollar y publicar planes de cumplimiento de la agencia para lograr la coherencia con M-25-21 (que se actualizarán cada dos años).
- Revisar y actualizar, cuando sea necesario, las políticas internas sobre infraestructura de TI, datos, ciberseguridad y privacidad para alinearlas con el M-25-21 y otras órdenes ejecutivas y leyes relevantes.
- Adquirir responsablemente capacidades de IA (Recuadro 4.7).

Nota: La directiva se aplica generalmente a todos los departamentos y agencias del Poder Ejecutivo, incluidas las agencias regulatorias independientes. Algunas partes del memorándum se aplican únicamente a las agencias de la "CFO Act" (Chief Financial Officers Act de 1990). Los organismos nacionales de inteligencia y el Departamento de Defensa también están excluidos de algunos requisitos. El documento de origen incluye detalles más específicos sobre la aplicabilidad.

Fuente: (US OMB, 2025^[17]), <https://www.congress.gov/bill/116th-congress/house-bill/2575/text>, <https://www.congress.gov/bill/116th-congress/house-bill/133>.

En un intento por ir más allá de la estrategia, algunos gobiernos han desarrollado una guía integral. Por ejemplo, además del manual de IA para el Gobierno del Reino Unido (Recuadro 4.2), Nueva Zelanda (2025^[18]) ha establecido un marco de IA de servicio público, e Irlanda (2025^[19]) ha publicado Directrices para el Uso Responsable de la IA en la Función Pública. Estas orientaciones son útiles no solo para implementar estrategias, sino también para ayudar a superar la aversión al riesgo al implementar leyes y reglamentos más formales (véase el análisis sobre las salvaguardas más adelante), al eliminar la necesidad de que cada organización o equipo haga sus propias interpretaciones. Como afirmó el cofundador de Code for America, "la legislación bien intencionada y bien escrita se origina en la cima de una jerarquía muy alta, y a medida que desciende, la flexibilidad que sus autores pretendían degradada. Las leyes a menudo tienen un efecto completamente diferente de lo que los legisladores pretendían debido a esta cascada de rigidez" (Pahlka, 2024^[20]).

Para contrarrestar aún más la aversión al riesgo, estas directrices podrían promover el uso del criterio y la discrecionalidad por parte de servidores públicos experimentados. También podría reconocer que, al igual que con cualquier acción humana, el aprovechamiento de la IA no puede estar libre de riesgos. Se pueden utilizar técnicas como la ciencia del comportamiento para diseñar orientaciones y comunicaciones que ayuden a garantizar el efecto conductual deseado en los usuarios, maximizando el valor de la adopción de la IA y mitigando los riesgos, aumentando la probabilidad de cambios significativos en el comportamiento y la adopción responsable de la IA (OECD, 2021^[21])⁷.

Recuadro 4.2. Manual de inteligencia artificial para el Gobierno del Reino Unido

El Servicio Digital del Gobierno del Reino Unido (GDS), con el apoyo de diversos departamentos del gobierno central, empresas tecnológicas del sector privado, instituciones académicas y usuarios, publicó el manual en febrero de 2025. Incluye 10 principios, junto con orientaciones sobre varios aspectos.

10 principios que guían la IA en las organizaciones gubernamentales

1. Sabes lo que es la IA y cuáles son sus limitaciones.
2. Utilizas la IA de forma legal, ética y responsable.
3. Sabes cómo usar la IA de forma segura.
4. Tienes un control humano significativo en el momento adecuado.
5. Entiendes cómo gestionar el ciclo de vida de la IA.
6. Utilizas la herramienta adecuada para cada tarea.
7. Eres abierto y colaborativo.
8. Trabajas con compañeros del ámbito comercial desde el principio.
9. Dispones de las habilidades y la experiencia necesarias para implementar y utilizar la IA.
10. Utilizas estos principios junto con las directrices de tu organización y cuentas con las garantías adecuadas.

Explicaciones y directrices

El manual incluye material informativo sobre lo que es la IA, los ámbitos de la IA, las aplicaciones de la IA en el gobierno, las limitaciones, la creación de soluciones de IA, la creación de un equipo, la definición del objetivo, la compra de IA, el uso de la IA de forma segura y responsable, la ética, las consideraciones jurídicas, la protección de datos y la privacidad, la seguridad y la gobernanza. También incluye una serie de casos de uso de IA que ilustran los esfuerzos del mundo real en el gobierno del Reino Unido.

Fuente: (UK Government Digital Service, 2025^[22]).

Determinar si la IA es la mejor solución

Deben existir directrices que se centren en determinar si la IA es la mejor solución para un problema determinado, alejándose así del enfoque en la IA. Si bien la IA tiene enormes capacidades, no siempre es la mejor solución y, en muchos casos, es inviable. Un hallazgo clave de un informe reciente del Instituto Ada Lovelace (2025^[23]) señala que "existe una sorprendente falta de pruebas sobre la efectividad y el impacto de las herramientas de IA, incluso desde un punto de vista puramente técnico. Evaluar las intervenciones de IA en contexto es crucial para determinar su rendimiento y valor en comparación con los métodos manuales o tradicionales existentes". El trabajo de RAND concluyó que una comprensión inadecuada del problema que debe resolverse y los proyectos que utilizan la IA innecesariamente son dos de los principales factores que impulsan el fracaso de los proyectos de IA (2024^[24]; 2025^[25]).

Un problema común con la IA es que las personas comienzan con soluciones y luego buscan problemas para que la tecnología los resuelva. En general, los gobiernos deben tratar de comprender los resultados que tanto ellos como sus ciudadanos pretenden alcanzar y los problemas que lo impiden, y centrarse en ellos. Armados con este conocimiento y prioridades, pueden identificar si la IA (o algo más) es la mejor solución para ayudar a lograr estos objetivos (Berryhill et al., 2019^[4]; Mulgan, 2019^[26]). En consecuencia, los gobiernos necesitan capacidades para identificar y comprender los problemas. También tendrán que aprovechar otros habilitadores presentados en este capítulo, como las competencias de la organización para entender las fortalezas y debilidades de la IA en relación con otras tecnologías y procesos, con el fin de involucrar a los usuarios y entender sus necesidades.

Algunos gobiernos han incorporado consideraciones al respecto en sus directrices para todo el gobierno. Por ejemplo, el manual de IA para el Gobierno del Reino Unido indica que los servidores públicos deben "estar abiertos a la conclusión de que, a veces, la IA no es la mejor solución para su problema: puede resolverse más fácilmente con tecnologías más establecidas". Las directrices para la adjudicación de contratos públicos relacionados con la IA del Reino Unido (2020^[27]) recomiendan a los responsables de adquisición de IA que "empiecen con [una] declaración del problema" y articulen "por qué consideran que la IA es relevante para el problema y están abiertos a soluciones alternativas". En Estados Unidos, la guía de IA para el gobierno ([28]) incluye componentes para "centrarse en el problema raíz" y considerar "¿Es la mejor opción para resolver este problema en particular? ¿Se han evaluado soluciones alternativas?"

Un proceso para realizar estas determinaciones podría integrarse como parte de una evaluación de impacto ex ante (véase la sección sobre salvaguardas más adelante en el apartado "Evaluaciones de impacto"), o podría establecerse como un proceso independiente antes de entrar en la cartera de proyectos de IA.

Definir funciones y responsabilidades claras

Los gobiernos deben definir funciones y responsabilidades claras para facilitar el desarrollo coherente, el uso y la posible ampliación de la IA. Estas funciones y responsabilidades deben definirse y acordarse con las partes interesadas pertinentes y asignarse a las instituciones gubernamentales, incorporándolas a los mandatos de las instituciones. Esto permite disponer de una estructura institucional sólida para apoyar la aplicación de las estrategias nacionales dentro de las instituciones y entre ellas, y facilita la rendición de cuentas y la supervisión en toda la administración pública. Varios países europeos están ampliando los mandatos de los ministerios o agencias existentes para garantizar un despliegue coherente de la IA. Algunos ejemplos son el Ministerio de Digitalización y Gobernanza Pública de Noruega y la Secretaría de Estado de Digitalización e IA de España adscrita al Ministerio para la Transformación Digital y de la Función Pública (OECD, 2024^[31]). El Reino Unido ha consolidado la mayoría de las funciones y responsabilidades en materia de IA de diferentes sectores en su Departamento de Ciencia, Innovación y Tecnología (DSIT) (OECD/UNESCO, 2024^[29]). A nivel individual, el memorándum estadounidense en el Recuadro 4.1 exige que los organismos del gobierno federal designen un CAIO para dirigir sus esfuerzos en materia de IA. Los gobiernos también pueden tener que analizar diferentes funciones y responsabilidades para lograr diversos objetivos estratégicos. Por ejemplo, en Chile y Colombia, las estrategias nacionales de IA definen actores responsables vinculados a cada compromiso tal como se define en la estrategia. Colombia también detalla los plazos en los que los actores responsables deben cumplirlos, así como los indicadores presupuestarios y de seguimiento (OECD/CAF, 2022^[30]; CONPES, 2025^[31]). Al establecer sus funciones y responsabilidades, los gobiernos deben dejar claro qué entidad o entidades tienen autoridad para establecer políticas sobre el uso de la IA en la administración. De hecho, la labor de la OCDE con los países ha revelado confusión en torno a quién es el responsable de establecer las normas, lo que podría obstaculizar los esfuerzos de transformación digital.

Coordinar los esfuerzos dentro y fuera de la administración pública

Los gobiernos pueden reforzar los esfuerzos de coordinación y colaboración para garantizar un enfoque holístico de la adopción y la gobernanza de la IA. El establecimiento de grupos de trabajo o comités interministeriales puede facilitar la toma de decisiones, la comunicación y la colaboración entre las diferentes instituciones. Estos mecanismos permiten a todos los agentes participar en el establecimiento de objetivos generales y trabajar juntos para alcanzarlos. Por ejemplo, el memorándum de Estados Unidos analizado en el Recuadro 4.1 exige la creación de un Consejo intergubernamental de CAIO. En otro ejemplo, Australia creó un grupo de trabajo temporal sobre IA en el gobierno (septiembre de 2023 a junio de 2024) para desarrollar políticas, normas y directrices que permitan un uso seguro, ético y responsable de la IA en la función pública (OECD, 2024^[3])⁸. A nivel subnacional, en el gobierno de Dubai, en los Emiratos Árabes Unidos, 22 CAIO de todo el gobierno se encargan de dirigir y coordinar los esfuerzos en materia de IA (WAM, 2024^[32]).

La coordinación también es importante de forma transversal, en todos los niveles de la administración. Muchas aplicaciones de IA tienen un impacto local significativo, sobre todo en la prestación de servicios públicos y el bienestar social, ya que los gobiernos locales están más cerca de la ciudadanía y los residentes. Sin embargo, sin coordinación, puede producirse una fragmentación de los enfoques de IA. Cuando los municipios y los gobiernos regionales desarrollan soluciones de IA de forma aislada, pueden surgir ineficiencias, duplicaciones de esfuerzos e incoherencias en los marcos de gobernanza y la experiencia del usuario (Verhulst and Sloane, 2020^[33]). Del mismo modo, el diseño de estrategias y políticas nacionales sin tener en cuenta las necesidades locales puede dar lugar a enfoques que no se correspondan con las realidades sobre el terreno o que no sean viables. Dinamarca ha tomado medidas para abordar este desafío a través de su nuevo Grupo de Trabajo Digital sobre IA, que está trabajando para ampliar la adopción de la IA en todos los niveles de la administración, garantizando la alineación de las prioridades, los estándares y los enfoques de gobernanza ⁹. Del mismo modo, en Suecia, AI Sweden

y Vinnova pusieron en marcha la iniciativa Colaboración para la IA en los municipios y la sociedad civil (Kraftsamlingen) para ayudar a los municipios y las organizaciones de la sociedad civil a integrar la IA en sus operaciones. Desde 2022, esta iniciativa ha prestado un apoyo personalizado, que incluye orientaciones sobre la adopción de la IA y oportunidades de financiamiento para proyectos concretos, fomentando un ecosistema de IA más coordinado y eficaz entre los gobiernos locales.

Creación de espacios para experimentar

Los gobiernos deben asignar tiempo y espacio para explorar el uso de la IA, ya que tanto la experimentación como el aprendizaje iterativo son cruciales para desarrollar la capacidad en este ámbito (OECD/CAF, 2022^[30]). Además de ayudar a promover el aprendizaje y a identificar nuevas posibilidades y enfoques, los entornos controlados para la experimentación y las pruebas de IA facilitan la identificación oportuna de posibles deficiencias técnicas, sesgos de comportamiento tanto de los sistemas de IA como de las personas que los utilizan, y desafíos de gobernanza asociados. Además, a través de la experimentación, los sistemas de IA pueden incubarse hasta que las soluciones sean técnicamente lo suficientemente robustas como para ampliarse. Al hacerlo, también pueden poner de relieve las preocupaciones del público, especialmente mediante pruebas en condiciones casi reales (OECD, 2019^[34]). Estos enfoques implican la participación de las partes interesadas durante la fase de desarrollo, la evaluación de las necesidades de los usuarios, la evaluación de la disponibilidad y la calidad de los datos y el seguimiento continuo del progreso de las fases de prototipado y piloto (OECD/UNESCO, 2024^[29]). Dichos entornos incluyen centros de innovación, laboratorios y "sandboxes" (o entornos de prueba). Los experimentos pueden operar en "modo de arranque", por lo que se despliegan, evalúan y modifican, luego se escalan o reducen, o se abandonan rápidamente (OECD, 2023^[35]). Más allá de la experimentación interna, los gobiernos también pueden trabajar con actores no gubernamentales, como las startups de GovTech, para diseñar y ejecutar experimentos de IA (véase "Recurrir a las startups de GovTech" más adelante).

Además, estos entornos fomentan la colaboración entre el gobierno, el mundo académico y la industria, promoviendo el intercambio de ideas y acelerando el desarrollo de tecnologías de IA. Al simular las condiciones del mundo real, estas instalaciones permiten una validación rigurosa de los sistemas de IA, garantizando que sean robustos, fiables y seguros antes de su despliegue. Este enfoque de las pruebas y la experimentación no solo mejora la efectividad de las soluciones de IA, sino que también genera confianza pública al validar estas tecnologías mediante la identificación temprana y al abordar posibles riesgos, sesgos o ineficiencias antes de su amplio despliegue.

Por ejemplo, la UE, en colaboración con sus Estados miembros, ha puesto en marcha una red de instalaciones permanentes de ensayo y experimentación (TEF), entre ellas CitCom.ai, que se centra en ciudades y comunidades inteligentes. La iniciativa acelera el desarrollo de una IA confiable en la Unión Europea proporcionando a los innovadores —tanto empresas como organismos públicos— acceso a pruebas y experimentos con productos basados en la IA en condiciones reales. Otros ejemplos son:

- En Estados Unidos, la Corporación Mitre, un centro de investigación y desarrollo (I+D) financiado por el gobierno, está desarrollando una supercomputadora de IA para alimentar un nuevo entorno de prueba de IA, que será capaz de entrenar nuevos sistemas avanzados de IA específicos del gobierno.¹⁰
- En el Reino Unido, la Incubadora de IA (i.AI) promueve la experimentación con IA y, en última instancia, su ampliación a través de cuatro enfoques clave: 1) la creación de prototipos para probar y evaluar rápidamente ideas para aplicaciones de IA; 2) la entrega de prototipos exitosos a equipos gubernamentales pertinentes donde puedan tener un impacto; 3) la modularización para compartir el trabajo técnico entre gobiernos, incluido el código de fuente abierta; y 4) la convocatoria y el asesoría para identificar áreas para compartir el aprendizaje y los productos.¹¹

- En Francia, ALLiance, una incubadora interministerial de IA puesta en marcha por la Dirección Interministerial de Asuntos Digitales (DINUM) en julio de 2023, es una iniciativa dirigida por el gobierno para estructurar la experimentación y la ampliación de la IA según las funciones centrales del gobierno.¹² La incubadora ALLiance aplica una metodología de "modo de producto" ágil y centrada en el usuario, desarrollada originalmente por beta.gouv.fr, que se centra en la iteración rápida, la retroalimentación del usuario y el impacto medible, acelerando así la adopción de la IA en el gobierno. Los criterios de selección estructurados de ALLiaNCE, basados en el impacto, el potencial de mutualización, la participación del usuario y el cumplimiento ético, demuestran un enfoque riguroso para probar y escalar proyectos de IA responsables.
- En Australia (2024^[36]), más de 60 agencias del Servicio Público Australiano (APS) llevaron a cabo una prueba de seis meses de Microsoft Copilot. Más de 7 700 servidores públicos participaron en el ensayo, cuyos resultados fueron mixtos; sin embargo, a nivel agregado revelaron que los usuarios percibían mejoras en la eficiencia y la calidad de la IA al resumir y preparar un primer borrador de documentos. A pesar de ello, la adopción de la IA requiere esfuerzos concertados para abordar las barreras técnicas, culturales y de capacidad.
- Portugal y España se han comprometido con organizaciones de GovTech para promover la experimentación con tecnologías digitales, incluida la IA, en la administración de justicia.¹³ También con GovTech pero con un enfoque más amplio, el GovTech Lab de España es un caso de uso de IA que identifica escenarios donde la IA generativa puede tener un gran impacto en las administraciones públicas – ya sea alcanzando una mayor eficiencia en la provisión de servicios públicos, reduciendo las cargas de trabajo o mejorando el servicio cívico¹⁴. 20 de 300 casos de uso serán piloteados en áreas como la clasificación de documentos, asistentes de IA y preparación de licitaciones y subvenciones. Aquellos que sean exitosos serán escalados y ofrecidos como un servicio a la totalidad de la administración.

Al poner en marcha pequeños experimentos y proyectos piloto de IA, los gobiernos deben tener en cuenta su definición de éxito y establecer mediciones y un marco de evaluación para determinar si un proyecto ha tenido éxito, determinar qué ha funcionado y qué no, y ayudar a recopilar y difundir las lecciones aprendidas. La Guía del Reino Unido para la Evaluación del Impacto de las Intervenciones en el ámbito de la IA es un buen ejemplo, ya que ofrece consideraciones sobre proyectos de IA desde las pruebas a pequeña escala hasta su plena aplicación.¹⁵

La OCDE está elaborando actualmente un informe específico sobre la experimentación con IA en la administración pública con el fin de revisar las prácticas actuales y extraer enseñanzas clave que ayuden a fundamentar y orientar a los responsables políticos a la hora de establecer sus propias directrices de experimentación para sus organizaciones (OECD, forthcoming^[37]).

Crear fundamentos sólidos para los datos

Los datos son el activo fundamental que impulsa la capacidad de la IA para funcionar, evolucionar y crear valor público. Basándose en el concepto de "entrada y salida de basura", el rendimiento de la IA se correlaciona directamente con la calidad y la representatividad de los insumos con los que se entrena. Los sistemas de IA a menudo requieren grandes cantidades de datos a lo largo del ciclo del sistema de IA para ofrecer resultados valiosos.

Los datos pertinentes pueden obtenerse de administraciones públicas, el sector privado u otras fuentes. Esta sección se centra principalmente en los datos gubernamentales, aunque la OCDE (2025^[38]) ha llevado a cabo un trabajo que sistematiza las fuentes a partir de las cuales los desarrolladores de IA obtienen datos para la formación en IA y destaca sus principales atributos.

El acceso y el intercambio de datos gubernamentales para la IA plantean complejos desafíos en materia de gobernanza de datos. Los gobiernos se enfrentan a obstáculos regulatorios y operacionales, desde la

protección de la privacidad, los resultados no sesgados y la seguridad de los datos hasta la navegación por los marcos políticos y jurídicos que rigen el intercambio y el uso de los datos y los derechos de propiedad intelectual. Las organizaciones del sector público también tienen que abordar cuestiones técnicas, desde garantizar la interoperabilidad entre los sistemas de datos hasta desarrollar la capacidad técnica para gestionar los datos de manera efectiva.

Garantizar la privacidad, la seguridad y los derechos de propiedad intelectual

Los gobiernos están creando marcos, directrices y mecanismos para promover una sólida gobernanza de datos que salvaguarde la privacidad, la propiedad intelectual y la seguridad. Suelen ser el resultado de colaboraciones entre organismos reguladores, partes interesadas del sector y la sociedad civil. Por ejemplo

- La Comisión de Protección de la Información Personal de Corea (2023^[39]) ha publicado una guía sobre el procesamiento de la información personal y el desarrollo de la IA. Esta guía describe la base jurídica para el tratamiento de datos personales, establece normas de seguridad y sugiere medidas para proteger los derechos de las personas en los sistemas de IA.
- El Marco de IA para el Servicio Público de Nueva Zelanda (2025^[18]) y el Manual de IA para el Gobierno del Reino Unido (2025^[22]) recogen los principios para un uso de la IA en la administración pública que sea seguro y respetuoso con la privacidad.
- Al conectar los conceptos de experimentación (analizados anteriormente) con la protección de datos personales, Francia ha creado un entorno de pruebas de datos (o "sandbox" como comúnmente se le conoce en inglés) para proporcionar un ambiente propicio para la experimentación segura, junto con formación y apoyo práctico en la gestión de datos personales y la garantía del cumplimiento de la regulación (Recuadro 4.3).
- Algunos gobiernos están explorando tecnologías que mejoren la privacidad (PET, por su sigla en inglés), como la anonimización de los datos, en relación con los datos confidenciales utilizados para la formación en IA. A su vez, estas tecnologías pueden mejorarse con IA (OECD, 2024^[40])¹⁶.

Recuadro 4.3. Un entorno de pruebas en Francia sobre datos personales aplicados a la IA en servicios públicos

En 2023, la autoridad francesa de protección de datos (CNIL) puso en marcha un entorno de pruebas para apoyar la innovación en IA para servicios públicos. Este entorno ofrece a determinadas organizaciones orientación experta para ayudarles a gestionar las regulaciones de datos personales en una fase temprana del desarrollo de sus proyectos. Si bien no elimina ningún requisito legal, ayuda a identificar soluciones a los desafíos de cumplimiento normativo. Se seleccionaron cuatro proyectos para recibir orientación del nuevo departamento de IA de la CNIL:

- Albert (DINUM): ayuda a los servidores públicos con un modelo lingüístico para mejorar las respuestas a las consultas de los usuarios, mediante proyectos piloto en los centros "France Services".
- Job Intelligence (Pôle Emploi): ofrece orientación personalizada para la búsqueda de empleo, utilizando datos profesionales que relacionan a los solicitantes de empleo con servicios adaptados.
- Ekonom AI (Nantes Métropole): ofrece información sobre el consumo de agua y recomendaciones para los residentes, apoyando los objetivos ecológicos y potencialmente adaptables a otras políticas públicas.

- Proyecto de vídeo de la RATP: desarrolla IA para detectar eventos mediante la captura de datos matriciales, garantizando la privacidad desde el diseño sin recopilación de datos personales.

Fuente: (CNIL, 2023^[41]; ^[42]).

Garantizar la representatividad de los datos

Garantizar que los sistemas de IA reciban formación sobre datos representativos es crucial para obtener resultados precisos y pertinentes. En algunos países, las diferentes poblaciones tienen idiomas y tradiciones únicos. En otros, diferentes factores demográficos u otros factores contextuales determinan los datos necesarios para que la IA sea eficaz. Tal como se describe en los Principios de Buenas Prácticas de la OCDE sobre Ética de los Datos en el Sector Público (2021^[43]). Sin embargo, el uso de datos que no son representativos para formar a la IA puede dar lugar a problemas importantes, sobre todo en el caso de aplicaciones gubernamentales que requieren políticas y decisiones justas y precisas que puedan tener un impacto tangible en la población objetivo. Entre estas cuestiones se incluyen los algoritmos y las decisiones sesgados, así como la incapacidad de desarrollar servicios y políticas adaptados para grupos infrarrepresentados en los datos, tal como se analiza en el Capítulo 1.

Los gobiernos están tomando medidas para abordar este problema. Por ejemplo, diversos países han tomado acciones para invertir recursos en la promoción de la representatividad lingüística (OECD, 2023^[44]; Peixoto, Canuto and Jordan, 2024^[45])¹⁷. Los ejemplos incluyen:

- Para promover la representación lingüística, la iniciativa de Recursos Comunes y la Plataforma Danesa de Recursos, ambas en danés, dirigida por la Agencia Danesa de Gobierno Digital (2024^[46]), recopilar, desarrollar y mostrar datos sobre el idioma y otras herramientas que puedan apoyar el desarrollo de soluciones de IA danesas.
- Grecia (2024^[47]) Prosigue el desarrollo de un espacio de datos sobre la lengua y la cultura griegas, centrado en la integración del patrimonio lingüístico y cultural de Grecia en las aplicaciones de la IA.¹⁸ También relacionado con el idioma griego, el desarrollo de los LLM "Meltemi" y "Llama-Krikri" representan pasos prometedores en esta dirección, destacando la importancia del acceso abierto y los esfuerzos colaborativos en la expansión de los recursos lingüísticos disponibles.¹⁹
- La plataforma india Bhashini, lanzada en el marco de la Misión Nacional de Traducción de Idiomas (NLTM), es el proyecto insignia de infraestructura lingüística de India liderado por la IA. Admite la traducción en tiempo real en 22 lenguas oficiales indias y docenas de dialectos, y proporciona soporte para asistentes de voz multilingües e interfaces de entrega de servicios de IA.²⁰ Gracias a un ejercicio masivo de participación ciudadana, fue posible crear conjuntos de datos multilingües que mejoran la representatividad de los datos, con modelos y API disponibles como código abierto.
- En Arabia Saudita, la Autoridad Saudita de Datos e Inteligencia Artificial (SDAIA, por su sigla en inglés) ha lanzado ALLaM, un LLM desarrollado con 500 mil millones de tokens y más de 300 000 textos árabes, que incluyen encyclopedias, investigación científica y trabajos históricos (M Saiful Bari, 2024^[48]). ALLaM pretende reflejar la riqueza lingüística y cultural de la lengua árabe.
- España está trabajando en una familia de modelos de IA llamada ALIA, que son entrenados exhaustivamente en español nativo y otros datos de lenguaje oficial y estará disponible como código abierto²¹. El primer caso incluye un asistente para el diagnóstico de fallo al corazón en el sector de salud pública y un asistente para facilitar la respuesta a la ciudadanía por parte de funcionarios de impuestos.
- También se han desarrollado modelos lingüísticos para lenguas americanas nativas. Aunque no fue desarrollado por el gobierno, los investigadores han desarrollado LakotaBERT para apoyar los

esfuerzos de revitalización del idioma para "Lakota, un idioma en peligro crítico del pueblo sioux en América del Norte (Parankusham, Rizk and Santosh, 2025^[49]).

Permitir que el acceso e intercambio de datos sea efectivo y confiable

Como se analiza en el Capítulo 3, los gobiernos suelen enfrentarse a una escasez significativa de datos de alta calidad, pertinentes y fácilmente disponibles, necesarios para formar eficazmente a los sistemas de IA (OECD, 2025^[50]). Para subsanar esta carencia es necesario centrar los esfuerzos en mejorar el acceso a los datos, por ejemplo, mediante la recopilación colaborativa de datos y acuerdos abiertos (Recuadro 4.4).

Recuadro 4.4. Recopilación colaborativa de datos en Suecia para Svea

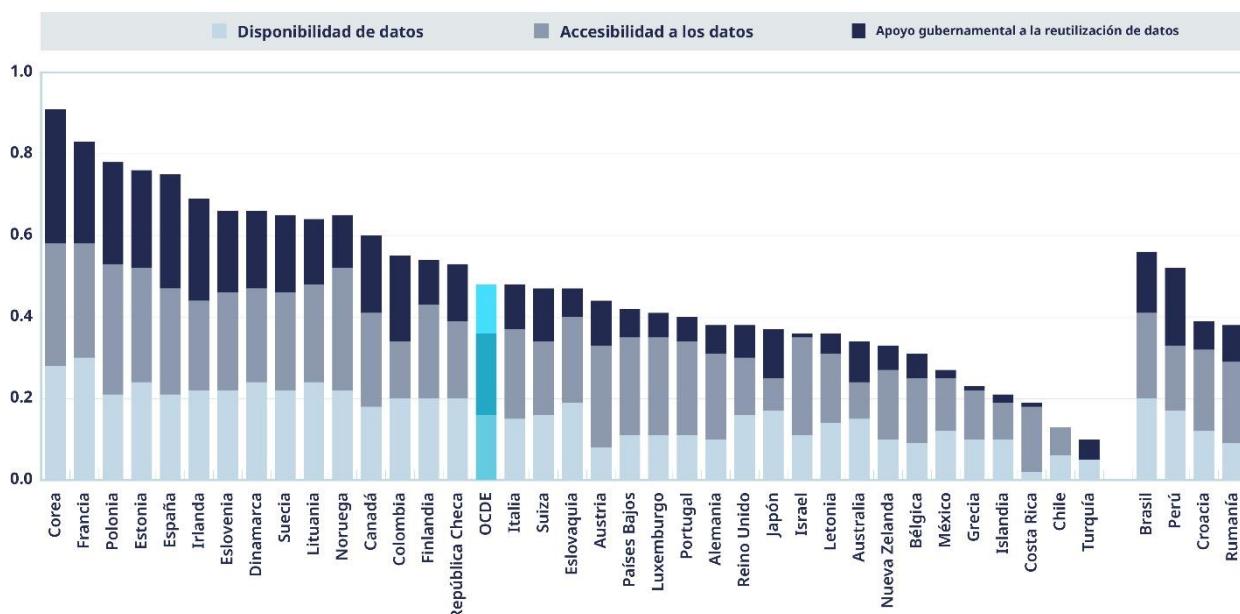
Svea es una iniciativa sueca coordinada por AI Sweden que reúne a organismos públicos, municipios, regiones y el sector para hacer frente a los desafíos que plantea la creación de soluciones de IA para los servicios públicos. El objetivo principal es aunar recursos para recopilar datos en sueco que reflejen las necesidades únicas del gobierno, una tarea demasiado grande para una sola organización.

Mediante la colaboración, las organizaciones gubernamentales pueden compartir la carga de trabajo de la recopilación de datos esencial para desarrollar un asistente de IA útil. En la primera fase, los participantes identificaron necesidades específicas y comenzaron a generar datos desde sus organizaciones para formar el sistema. En la próxima fase, tendrán acceso a bases de datos compartidas de información nacional pertinente para seguir informando al asistente de IA.

Fuente: (AI Sweden, 2024^[51]).

Otra iniciativa clave son los datos abiertos de organismos públicos. En promedio, solo el 46% de los conjuntos de datos gubernamentales de alto valor están disponibles como datos abiertos en la OCDE (2023^[52]), frente a más del 80% en Francia y Corea. Sigue habiendo desafíos por lo que respecta al fomento de la reutilización de datos abiertos por parte de los agentes y la integración de estos conjuntos de datos en los sistemas de IA. Los resultados de la edición de 2023 del índice de datos abiertos, útiles y reutilizables (OURdata) muestran que los países obtienen mejores resultados en disponibilidad y accesibilidad de los datos que el apoyo gubernamental a la reutilización de datos (Gráfico 4.1).

Gráfico 4.1. Índice de datos abiertos, útiles y reutilizables (OURdata) de la OCDE, 2023



Fuente: (OECD, 2023^[52]).

Una cuestión clave es cómo aumentar el valor de los datos abiertos de organismos públicos para los sistemas de IA desde su diseño y, por tanto, su accesibilidad y su preparación para la IA. Por un lado, la estandarización (por ejemplo, en términos de estructura y formatos) de los datos abiertos de organismos públicos puede reducir el tiempo que los desarrolladores de IA necesitan invertir en la preparación de datos para formar sistemas de IA. Por otro lado, el mayor uso de herramientas, como las API, también puede favorecer la integración de datos con sistemas basados en IA al proporcionar un método estandarizado para compartir y acceder a los datos de forma automática, directamente desde su origen. En la actualidad, solo el 47% de los conjuntos de datos de alto valor se publican con API (OECD, 2023^[52]).

Más allá de los datos abiertos, otras iniciativas relevantes incluyen aumentar el acceso a grandes conjuntos de datos públicos o financiados con fondos públicos, como el Buscador de Datos del Centro de IA de Corea del Sur (2024^[53]), que proporciona acceso a conjuntos de datos de texto, imágenes, vídeo, audio y sensores relevantes para la formación en IA, en ámbitos como la sanidad, el transporte público y la seguridad y gestión de catástrofes.

Uso de datos del sector privado

Tanto los datos del sector público como los del privado pueden desempeñar un papel fundamental en el desarrollo de aplicaciones de IA para la administración. Si bien los datos gubernamentales proporcionan información esencial sobre la demografía y los servicios públicos, los datos del sector privado, como los patrones de movilidad, el comportamiento del consumidor y las tendencias financieras, pueden mejorar esta información. Por ejemplo, los sistemas de IA para la planificación urbana pueden beneficiarse de los datos de telecomunicaciones para analizar el flujo de tráfico, mientras que la IA sanitaria puede aprovechar los datos anonimizados de los pacientes de las clínicas privadas para mejorar la predicción de enfermedades. Al combinar ambas fuentes de forma responsable y ética, las aplicaciones de IA pueden ser más precisas, eficientes y responder a las necesidades públicas. Un ejemplo de puesta en común y combinación de datos de los sectores público y privado son los espacios europeos comunes de datos. El objetivo de los espacios de datos es poner a disposición más datos para su acceso y reutilización en toda

la Unión Europea en un entorno confiable y seguro en beneficio de las empresas y la ciudadanía europeos (OECD, 2024^[54]; EC, 2025^[55]).

Crear un entorno propicio con gobernanza de datos

Desde facilitar el acceso e intercambio de datos hasta sentar las bases necesarias para que la IA sea una posibilidad en el gobierno, los gobiernos deben desarrollar mecanismos sólidos de "gobernanza de datos", que puedan integrarse en estrategias y políticas de IA más amplias (OECD, 2024^[3]).

La gobernanza de datos se refiere a "diversos mecanismos, incluidas disposiciones técnicas, normativas, regulatorias e institucionales, que afectan a los datos y a su creación, recopilación, almacenamiento, uso, protección, acceso, intercambio y eliminación, incluso entre ámbitos de actuación y fronteras organizativas y nacionales" (OECD, 2022^[56]).

El marco de la OCDE para la gobernanza de datos en el sector público, desarrollado originalmente para explorar los mecanismos específicos que deberían existir para permitir el acceso y el intercambio de datos, puede aplicarse al contexto de los datos para los sistemas de IA (Gráfico 4.2). Los gobiernos pueden desarrollar capacidades de gobernanza de datos en el sector público dando prioridad al desarrollo de estrategias integrales de datos, definiendo funciones de liderazgo y estableciendo una visión para gestionar y gobernar los datos a un nivel más técnico para materializar los beneficios y resultados previstos de la IA.

Gráfico 4.2. Gobernanza de datos en el sector público



Fuente: (OECD, 2019^[57]).

Por ejemplo, la Estrategia de Datos para la Función Pública Federal de 2023-2026 de Canadá describe los resultados deseados y los principios rectores para avanzar en una gobernanza de datos sólida en todo el gobierno federal en su conjunto, junto con las expectativas en cuanto a funciones y responsabilidades²². Otros ejemplos son la Estrategia Nacional de Datos del Reino Unido y la Estrategia de Datos y Gobierno Digital de Australia²³. Grecia (2024^[47]) está llevando a cabo un programa emblemático sobre gobernanza de datos y coordinación de estrategias de IA para establecer los elementos necesarios para apoyar a un sector público y privado preparado para la IA. El memorándum de Estados Unidos analizado en el

Recuadro 4.1 establece varios requisitos para mejorar la gobernanza de los datos. También es importante integrar la participación de las partes interesadas con los titulares de los derechos de datos, como la ciudadanía, las empresas y los representantes de la sociedad civil, que podrían verse afectados por el uso de los datos para la IA, ya sea en el contexto de los derechos de propiedad intelectual o de los derechos de los datos personales (OECD, 2022^[58]), o por el posible uso de datos inadecuados o sesgados.

Los gobiernos deben apoyar su visión y estrategia sobre los datos para la IA con una capacidad adecuada para una implementación coherente en toda la administración pública, junto con directrices y marcos jurídicos para garantizar la efectividad (OECD, 2019^[59]). Esto puede incluir numerosas áreas, como la mejora de la alfabetización en materia de datos y las habilidades para la IA, con ejemplos como la iniciativa²⁴ CAPACITA GOV.BR de Brasil y el Programa Nacional para Mejorar la Protección de los Datos Personales de Argentina. Esto también puede incluir esfuerzos para impulsar la coordinación y la colaboración institucional, con ejemplos como el Centro Noruego de Recursos para el Intercambio y el Uso de Datos (Digdir, 2024^[60]). Por último, es esencial reconocer el papel facilitador de los marcos jurídicos que orquestan y aceleran la integración y el intercambio de datos entre instituciones públicas, salvaguardando al mismo tiempo los derechos individuales y la privacidad. Es importante disponer de una legislación clara, moderna y aplicable sobre gobernanza de datos y protección de los datos personales para desplegar sistemas de IA fiables a gran escala. En Chile, se está debatiendo actualmente un proyecto de ley de gobernanza de datos, en el que se definen los principios, funciones y mecanismos de coordinación interinstitucional para la gobernanza de datos. Adicionalmente, está entrando en vigor una nueva Ley de Protección de Datos Personales, acorde con las normas internacionales. Otros ejemplos son la Ley de Intercambio y Gobernanza de datos de Irlanda de 2019 y el Reglamento General de Protección de Datos (RGPD) de la UE, la Ley de Datos, la Ley de Gobernanza de Datos, la Directiva sobre datos abiertos y los marcos de interoperabilidad de datos²⁵.

Por último, la porción de entrega de gobernanza de datos se refiere a los procesos, mecanismos y herramientas que permiten la implementación operativa de la gobernanza de datos a nivel organizativo y de equipo, garantizando que se implementan prácticas sólidas de gobernanza y gestión de datos e integradas en el ciclo de valor de los datos de IA (OECD, 2019^[59]). Un ejemplo es la evaluación realizada por Estados Unidos (2024^[61]) sobre la madurez de las operaciones de datos (DataOps) en los organismos federales, como parte de su Guía de IA para el gobierno. Este marco evalúa lo bien que las organizaciones pueden descubrir, acceder y utilizar datos para respaldar el desarrollo de la IA a lo largo del ciclo de vida del valor de los datos. Entre los componentes clave se incluyen la garantía de un catálogo completo de activos de datos, métodos flexibles de acceso a los datos y herramientas que faciliten los experimentos documentados con IA. Otros componentes clave de la gobernanza de datos son las competencias técnicas y los perfiles de trabajo necesarios, entre ellos, los científicos de datos, los expertos en el ámbito, los ingenieros de datos y los proveedores de datos, que participan en la recopilación y el procesamiento de datos para la IA (OECD, 2022^[58]). En la siguiente sección se tratan cuestiones relacionadas con la infraestructura de datos.

Construcción de infraestructuras digitales

Garantizar la disponibilidad de una infraestructura digital confiable y escalable puede ayudar a respaldar y ampliar la IA en el sector público. Además de los propios datos, como se ha comentado en el apartado anterior, la infraestructura de datos, las plataformas informáticas escalables, los modelos fundacionales de IA y las herramientas comunes de IA son importantes componentes básicos de la IA en el sector público. También existen otras formas de infraestructura digital, que se analizan en el trabajo de la OCDE (2024^[62]). Esta sección se centra en las más relevantes para la IA.

En sus intentos de perseguir sistemas de IA confiables y escalables, los gobiernos se enfrentan a decisiones estratégicas con respecto a los sistemas de IA. Por una parte, el desarrollo de la capacidad

nacional a través del desarrollo de la infraestructura digital nacional puede ayudar a un país a aplicar sus propias normas de protección de datos y privacidad. Varios países están tratando de desarrollar esa capacidad (Letzing, 2024^[63]; France Élysée, 2025^[64]; African Union, 2024^[65]; Ray, 2025^[66]; EC, 2025^[67]; Brizuela et al., 2025^[68]). Por otra parte, esto también podría contribuir a la fragmentación tecnológica y al cierre de los ecosistemas que limitan la colaboración internacional (Komaitis, Ponce de León and Thibaut, 2024^[69]; Frazier, 2025^[70]). Los gobiernos deben considerar varias opciones para determinar un equilibrio que les resulte adecuado para desarrollar soluciones internamente en comparación con la colaboración con el sector privado y con otros países.

La infraestructura digital no solo es relevante para los gobiernos nacionales, sino que también puede ser formativa en la adopción de IA en gobiernos subnacionales, como las ciudades. El desarrollo de herramientas digitales compartidas y reutilizables puede ayudar a los gobiernos locales a superar los costos de entrada debidos a las economías de escala subyacentes y permitir que las soluciones de IA se adapten a las necesidades y contextos locales.

Potencia de cómputo e infraestructura de datos

El acceso a los recursos de infraestructura informática puede ser clave para el desarrollo y uso eficaz de la IA en la administración pública (OECD, 2022^[71]). La elección entre soluciones locales y soluciones en la nube para el despliegue de la IA depende de las necesidades específicas, las opciones políticas, los requisitos reglamentarios, las limitaciones presupuestarias y los objetivos a largo plazo.

Las soluciones locales ofrecen un mayor control, personalización y seguridad, lo que las hace adecuadas para aplicaciones muy sensibles y para cumplir con las leyes de localización de datos (Redapt, 2023^[72]). Las soluciones en la nube, por otro lado, proporcionan una escalabilidad, rentabilidad y acceso inigualables a tecnologías de IA de vanguardia, lo que las convierte en ideales para proyectos dinámicos y en rápida evolución y más prácticas que las soluciones locales para proyectos pequeños o nuevos o pequeños participantes en el desarrollo de IA (Dombo, 2023^[73]). Más de la mitad de los países de la OCDE cuentan con iniciativas de tecnología en la nube, incluidas capacidades de almacenamiento e informática (Infraestructura como servicio, o IaaS). En particular, el acceso a las tecnologías en la nube depende de soluciones tanto públicas como privadas (48% frente a 52% respectivamente), y varios países están desarrollando tecnologías en la nube dirigidas por el sector público (OECD, 2024^[74]). En muchos casos, un enfoque híbrido que combine recursos locales o infraestructura dedicada de otro tipo y nube pública (compartida, de terceros) ("nube híbrida") puede ofrecer una solución equilibrada, aprovechando las fortalezas de cada uno. En particular, un estudio realizado por RAND (2024^[24]; 2025^[25]) destacó que, entre las empresas, aquellas que podían utilizar soluciones en la nube no solían tener problemas para garantizar una capacidad informática adecuada, pero las que no podían transferir datos a la nube se enfrentaban a desafíos importantes que contribuían al fracaso de los proyectos de IA.

La demanda mundial de capacidad de centros de datos preparados para IA podría triplicarse de aquí a 2030, lo que demuestra el creciente uso de la IA (McKinsey, 2024^[75]). Se estima que las emisiones de carbono generadas por los centros de datos y las redes de transmisión de datos ya representan el 1% de todas las emisiones relacionadas con la energía, pero en general han crecido solo de forma modesta a pesar del rápido crecimiento de la demanda de servicios digitales, debido en parte al aumento de la eficiencia de los equipos y modelos a lo largo del tiempo (OECD, 2022^[71]; IEA, 2023^[76]). Un Análisis reciente de la Agencia Internacional de la Energía (AIE) (2025^[77]) reveló que los centros de datos se encuentran entre las fuentes de emisiones de más rápido crecimiento y que dichas emisiones podrían aumentar significativamente en los próximos diez años, pero también que "la adopción generalizada de las aplicaciones de IA existentes podría conducir a reducciones de emisiones que son mucho mayores que las emisiones de los centros de datos". Sin embargo, las considerables emisiones de carbono vinculadas a datos ponen de relieve la necesidad de adoptar prácticas para gestionar los requisitos energéticos de la IA. También es importante reconocer las crecientes cantidades de agua necesarias para

enfriar los centros de datos (Metz et al., 2025^[78]). Las investigaciones recientes y los avances del sector indican una tendencia creciente de la IA hacia la adopción de modelos más pequeños o especializados, como los modelos de lenguaje pequeño (MLS) que consumen menos recursos, requieren menos datos y son menos costosos (Hassani et al., 2022^[79]; Jones, 2025^[80]).

Recuadro 4.5. Centros de datos compartidos y nube gubernamental de Corea

El Servicio Nacional de Recursos de Información (NIRS) de Corea ha estado trabajando con el Ministerio del Interior y Seguridad para actualizar el hardware, las redes y las herramientas de gestión clave para ayudar a modernizar la tecnología de Corea y permitir la migración a la nube. Una parte fundamental de esta labor ha sido la construcción de nuevos centros de datos gubernamentales, que pueden ayudar a garantizar el cumplimiento de los requisitos gubernamentales, la eficiencia de costos con una huella tecnológica reducida y la creación de empleo y la inversión local en las zonas objetivo. Estos centros de datos también se han puesto a disposición de los principales socios del gobierno en el sector privado, lo que ayuda a garantizar que las empresas que poseen o manejan datos confidenciales lo hagan en un entorno que cumple con los requisitos gubernamentales de seguridad, copia de seguridad y redundancia, entre otros. Con medidas en torno a la sostenibilidad y las energías renovables, los centros de datos ayudan a reducir el impacto ambiental del gobierno digital de Corea, sobre todo cuando se prepara para hacer un mayor uso de las soluciones de IA.

Fuente: (OECD, 2025^[81]).

Al tratar de adoptar un enfoque holístico, algunos países están creando infraestructuras informáticas y de datos como parte de un paquete de medidas, incluidos enfoques sobre la infraestructura pública digital (IPD)²⁶ para apoyar el intercambio seguro y significativo de datos entre gobiernos, respaldado por una sólida integración de datos y capacidades de análisis. Estas infraestructuras no solo ayudan a ampliar el uso de la IA, sino que también pueden fomentar la colaboración interinstitucional y la generación de valor público a partir de los datos. Tomemos, por ejemplo, la Infraestructura Nacional de Datos (IND) de Brasil²⁷. Esta iniciativa estratégica establece un conjunto de políticas, normas, tecnologías y mecanismos de gobernanza para organizar, compartir y gestionar los datos del sector público de forma segura y eficiente. Su objetivo principal es hacer que los datos gubernamentales sean localizables, accesibles, interoperables y reutilizables (principios FAIR), promoviendo la transparencia, la mejora de los servicios públicos, la eficiencia administrativa y la toma de decisiones basada en evidencia en toda la administración, sirviendo como base para la transformación digital y la innovación. La plataforma gov.br del país sirve como centro para integrar el acceso a casi 5 000 servicios públicos digitales e incluye una plataforma Conecta gov.br como capa de interoperabilidad de datos en todo el gobierno. Estas plataformas y otras infraestructuras digitales son la base de la IA en el gobierno. En otro ejemplo, la plataforma nacional en la nube de Arabia Saudí, Deem Cloud, desarrollada por la Autoridad Saudita de Datos e Inteligencia Artificial (SDAIA), consolida la infraestructura digital en más de 190 entidades públicas y más de 260 centros de datos (SDAIA, 2025^[82]). Proporciona un conjunto de servicios en la nube para respaldar operaciones digitales seguras y eficientes, y ha contribuido al ahorro energético y de costos, como parte de los esfuerzos más amplios para modernizar la infraestructura del sector público y apoyar las estrategias digitales nacionales. Más centrado en informática, Grecia (2025^[83]) está construyendo DAEDALUS, que será uno de los superordenadores más potentes de Europa y será accesible para las instituciones públicas.

Desarrollo de modelos fundacionales de IA

Los modelos de base son una forma de modelos de IA formados en grandes cantidades de datos, que generalmente utilizan la autosupervisión a escala, y que pueden adaptarse a una amplia gama de tareas posteriores (OECD, 2024^[84]). Los gobiernos pueden desarrollar sus propios modelos fundacionales o basarse en los existentes para crear enfoques adaptados al contexto específico de un país y/o su público.

Los modelos base pueden "perfeccionarse" mediante una formación adicional sobre conjuntos de datos más estrechos relacionados con una tarea o un dominio concretos, mejorando su rendimiento para ese contexto específico (Montgomery, Rossi and New, 2023^[85]).

Generalmente, la construcción de un modelo base desde cero se considera costosa y normalmente requiere importantes recursos de datos y energía. Entre los ejemplos de modelos de base privados y privados se incluyen Mistral Large y los que impulsan a Claude Anthropic (por ejemplo, el soneto Claude 3.7), Google Gemini (por ejemplo, Gemini Ultra) y ChatGPT de OpenAI (por ejemplo, GPT-5o).

Aunque pueden ser costosas y requieren una gran cantidad de datos y gran uso de energía, los gobiernos pueden, de hecho, formar y construir sus propios modelos base, como se comenta en el Capítulo 3. Los gobiernos también pueden perfeccionar y adaptar un modelo de base propietario para que se ajuste mejor a su propio contexto, lo que puede reducir significativamente los costos financieros y de tiempo asociados con el despliegue de la IA para tareas específicas. Un ejemplo de este enfoque es el asistente público virtual habilitado para ChatGPT en Portugal para los servicios públicos (Recuadro 5.46).

Los gobiernos también pueden utilizar modelos de código abierto "previamente formados". Se trata de modelos base que han sido formados por una empresa u otra organización que ha hecho que su "arquitectura modelo y sus ponderaciones sean de acceso libre y público para que cualquier persona pueda modificarlos, estudiarlos, construirlos y utilizarlos" (Seger et al., 2024^[86])²⁸. La mayoría de los modelos de código abierto son creados por grandes empresas tecnológicas, como la serie Llama de Meta, aunque se han desarrollado modelos más orgánicos, impulsados por la comunidad de código abierto, como el Gran modelo lingüístico multilingüe de ciencia abierta y acceso abierto de BigScience (BLOOM por sus siglas en inglés)²⁹. Un ejemplo de gobierno que aprovecha modelos de IA de código abierto es el asistente virtual Albert de Francia para servidores públicos (Recuadro 5.46).

Un modelo base que se adapte a un contexto nacional y gubernamental, por ejemplo, mediante la puesta a punto de un modelo propietario o la personalización de un modelo de código abierto, puede reducir significativamente el costo de adopción para los equipos que desean implementar la IA. Sin embargo, el desarrollo y uso de modelos base conlleva riesgos que los gobiernos deben tener en cuenta, como se analiza en el Capítulo 1.

Los gobiernos están demostrando cada vez más interés en invertir en modelos de base nacionales o regionales para mejorar la soberanía tecnológica y reflejar mejor los diferentes idiomas y culturas. Por ejemplo, más de 30 instituciones latinoamericanas están desarrollando Latam-GPT, dirigido por Chile, para crear un modelo basado en datos regionales (Gob.cl, 2025^[87]). OpenEuroLLM es una iniciativa financiada por la UE para construir modelos de código abierto que cubran todas las lenguas oficiales europeas (EC, 2025^[88]). En el Sudeste Asiático han surgido modelos adaptados al contexto y al idioma (Noor and Kanitroj, 2025^[89]). Del mismo modo, el proyecto italiano Minerva ha desarrollado el primer Máster en Lengua Italiana desde cero y es uno de los pocos ejemplos en este contexto de modelos de cimentación a medida que se desarrollan desde cero (University of Rome Sapienza, 2024^[90]).

Herramientas comunes de IA

Las herramientas comunes de IA que pueden utilizarse en todo un gobierno y adaptarse a las necesidades gubernamentales específicas pueden servir como una forma de IPD que habilita y mejora otros servicios. A veces construidas con modelos base, otras veces proporcionados como otro tipo de IPD, estas herramientas proporcionan una capa de servicio compartida y pueden soportar la automatización de tareas rutinarias, mejorar la interacción del usuario y mejorar la prestación del servicio.

Por ejemplo, los chatbots pueden manejar un gran volumen de consultas de la ciudadanía, proporcionando respuestas instantáneas a preguntas comunes y renunciando a recursos humanos para tareas más complejas. Esto no solo mejora la eficiencia, sino que ayuda a garantizar que los servicios públicos sean más accesibles y respondan mejor a las necesidades de la comunidad. Para que se

consideren IPD, estas herramientas de IA deben resolver una necesidad básica común y, por lo tanto, ser utilizables en una amplia gama de organizaciones del sector público. Un ejemplo de este enfoque es el Asistente Virtual Inteligente de Conversación (VICA, por su sigla en inglés) de Singapur, que se ofrece como servicio compartido y que utilizan más de 60 organismos públicos para crear más de 100 chatbots³⁰.

Herramientas comunes de apoyo a la IA no siempre recurren a ella. En otros ejemplos:

- La herramienta gubernamental Whole of Government Application Analytics (WOGAA, por su sigla en inglés) (Aplicación de Análisis del Gobierno en su Totalidad) de Singapur se ha desarrollado como una herramienta gubernamental para supervisar el rendimiento de los sitios web y servicios digitales públicos, incluidos los habilitados por la IA, que proporciona un panel central para realizar un seguimiento del tráfico de los sitios web, informes automatizados con métricas clave, rendimiento de referencia en comparación con otros sitios web gubernamentales, etc.
- En Estonia, para aplicar el principio de "solo una vez", todas las bases de datos públicas se describen obligatoriamente en el catálogo de recursos de interoperabilidad (RIHA), que sirve de registro nacional de sistemas, componentes, servicios, modelos de datos, activos semánticos y otros, garantizando una gestión transparente, equilibrada y eficiente de los sistemas de información pública.
- La incubadora francesa de IA de ALLiaNCE promueve el desarrollo de productos de IA reutilizables en todas las administraciones, con el objetivo de mutualizar esfuerzos y reducir la duplicación. ALLiaNCE ofrece una oferta de servicios de varios niveles, incluidas las herramientas habilitadas para IA integradas en el paquete digital de Francia (La Suite Numérique), así como un nivel fundacional, Albert API, para admitir la reutilización intersectorial de los sistemas GenAI. Como norma digital común, la API de Albert proporciona sistemas de GenAI abiertos y reutilizables, reduciendo el umbral de adopción para los organismos públicos y contribuyendo a la creación de una IPD nacional para la IA. La incubadora ALLIANCE también da prioridad a la soberanía de los datos y a las soluciones digitales abiertas y soberanas, lo que refleja los esfuerzos gubernamentales para mitigar los riesgos de dependencia y garantizar la confianza en los sistemas de IA.

El Catálogo de herramientas y métricas de IA de OECD.AI para una IA confiable incluye una variedad de herramientas de dentro y fuera de los gobiernos que pueden ayudarles a fundamentar la determinación de sus propias necesidades³¹.

Fomentar las competencias y el talento

Como se analiza en el Capítulo 3, las carencias en materia de competencias son uno de los desafíos más significativos para la adopción de una IA confiable en el gobierno. Por lo tanto, los gobiernos deben adoptar medidas adecuadas para fomentar la competencia y la capacidad internas.

Los gobiernos deben dotar a los servidores públicos de las competencias adecuadas para maximizar la efectividad de la IA, garantizando al mismo tiempo un uso seguro y confiable. Un servicio público preparado para la IA es fundamental para el desarrollo y el despliegue de soluciones de IA, así como para el uso eficaz de herramientas impulsadas por la IA con el fin de mejorar las tareas diarias y la formulación de políticas. Un enfoque estratégico y coordinado de las competencias y el talento en IA puede ayudar a dirigirse a diferentes grupos de la organización, identificar carencias de competencias, desarrollar las competencias adecuadas y atraer y retener talento más especializado en IA. Asimismo, el fortalecimiento de las competencias internas puede contribuir al desarrollo de la capacidad nacional, un tema tratado en la sección anterior.

Esto puede incluir la contratación de personas con las competencias necesarias para trabajar con IA, así como el perfeccionamiento profesional de las funciones existentes. Los gobiernos tendrán que prever que, a medida que la IA evolucione, también lo harán las competencias necesarias, lo que exige un aprendizaje

continuo. Un enfoque sólido requerirá una evaluación de las necesidades para mapear el nivel actual de capacidad de datos y la IA de los actuales servidores públicos, identificar las principales carencias y fundamentar la estrategia para abordar estas necesidades. Esto significaría poder tomar decisiones informadas entre la contratación, la retención y el desarrollo del talento digital. Esto también podría servir de base para programas de formación adaptados que aborden las carencias en materia de competencias y para un enfoque que permita gestionar y formar (mediante la calificación, el perfeccionamiento o la recalificación) las funciones más afectadas por la integración de la IA.

Evaluación de necesidades de distintos grupos de usuarios

Las instituciones gubernamentales deben evaluar las necesidades de los diferentes grupos de usuarios para adoptar la IA y utilizarla de manera efectiva. Una organización preparada para la IA va desde los usuarios generales de sistemas de IA hasta los líderes institucionales, los profesionales de datos y digitales, y las funciones más especializadas. Como se ilustra en el Gráfico 4.3, los grupos de usuarios se vuelven más estrechos y especializados más abajo en la pirámide.

Gráfico 4.3. Considerar el nivel de alfabetización sobre IA necesario para los diferentes grupos de usuarios en la organización



En general, los servidores públicos no especializados son fundamentales para la adopción y el uso eficaz de la IA. Su formación debe centrarse en la alfabetización general en torno a los datos y las tecnologías de IA, su uso efectivo con respecto a determinadas tareas y la consideración ética y jurídica para su uso.

Los líderes serán parte integral de la adopción de la IA, siendo la capa en la que la tecnología se encuentra con el negocio del gobierno. Ayudan a concienciar a los usuarios, impulsan la adopción de la IA y promueven oportunidades de aprendizaje y desarrollo en la administración pública. Este grupo de usuarios necesita una visión estratégica y una comprensión a nivel ejecutivo de lo que pueden hacer las tecnologías de IA, su impacto y cómo abordar los riesgos, el cumplimiento, el financiamiento y la gestión de la organización.

Los profesionales de datos y digitales lideran y facilitan el diseño, desarrollo e implementación de servicios específicos. Este grupo de usuarios necesita un mayor grado de alfabetización en IA para comprender cómo debe implementarse la IA para lograr los objetivos previstos para los servicios de los que son responsables. Junto con los especialistas, este grupo puede ser responsable de la adquisición de IA a través de procesos de contratación pública. Por lo tanto, el conjunto adecuado de competencias puede empoderarlos en las negociaciones con los proveedores que desean vender productos y servicios de IA.

Aunque son una parte menor de la organización, los **especialistas en IA** son fundamentales para el desarrollo, la implementación, la gestión y el uso de los sistemas de IA. Van más allá del desarrollo directo de la IA, e incluyen funciones en materia de contratación pública y gestión jurídica y de proyectos. Para desarrollar este grupo de usuarios, los gobiernos deben centrarse en la atracción, retención y aprendizaje y desarrollo. Además, para superar la posible escasez de competencias en el mercado, las organizaciones gubernamentales también pueden considerar aprovechar las capacidades externas a través de la contratación pública y las alianzas, como se analiza a continuación.

Preparación de los usuarios de la función pública

Una administración pública preparada para la IA requerirá una combinación de competencias digitales básicas y una alfabetización más específica en materia de datos e IA. El Marco de *políticas de la OCDE para el talento y las competencias digitales en el sector público* (OECD, 2021^[7]) describe las diversas competencias digitales fundamentales aplicables a todos los servidores públicos, esenciales para apoyar la transformación digital:

- Comprender el potencial de la transformación digital.
- Comprender a los usuarios y sus necesidades.
- Colaborar abiertamente para la entrega iterativa.
- Uso confiable de los datos y la tecnología.
- Capacidades que permitan un sector público basado en datos.
- Habilidades socio-emocionales del gobierno digital.
- Habilidades de liderazgo del Gobierno Digital.

Para que los trabajadores estén preparados para la IA, es necesario partir de estas bases con la alfabetización necesaria para que "las personas evalúen de forma crítica las tecnologías de IA, se comuniquen y colaboren eficazmente con la IA y utilicen la IA como herramienta en línea, en el hogar y en el lugar de trabajo" (Long and Magerko, 2020^[91]). Esto incluye entender los sistemas de IA, el manejo y la gestión de datos y la ética. Un enfoque, por ejemplo, es el marco de políticas de competencias empresariales basado en las competencias de IA desarrollado por el Instituto Alan Turing en el Reino Unido (2023^[92]), que tiene cinco dimensiones:

- **Privacidad y protección:** para mitigar los riesgos relacionados con la seguridad y protección de los datos, especialmente en lo que respecta a consideraciones legales, reglamentarias y éticas.
- **Especificación, adquisición, ingeniería, arquitectura, almacenamiento y conservación:** para el manejo y gestión de datos que permitan un uso más efectivo y ético de los sistemas de IA.
- **Definición y comunicación de problemas:** identificar, definir y comunicar los "problemas" que más se benefician de la aplicación de las soluciones de IA.
- **Resolución de problemas, análisis, modelización y visualización:** incluye una serie de herramientas y métodos que pueden utilizarse para el análisis, la aplicación de la IA y la comunicación.
- **Evaluación y reflexión:** para comprender el impacto del trabajo en la IA, evaluar la eficiencia y la efectividad de los proyectos de IA e identificar oportunidades de mejora.

El nivel de competencia y la medida en que cada uno de ellos es necesario varía en función del grupo de usuarios, la capacidad existente o el nivel de capacidad requerido para sus funciones. Sin embargo, esta combinación de bases digitales y competencias en materia de IA debería propiciar un uso más eficaz de la IA en la administración pública.

Desarrollo de competencias y talento

Para desarrollar competencias y talento en materia de IA en la función pública, los gobiernos deben tener en cuenta tanto las prácticas de desarrollo interno como la contratación más amplia de talento clave. Entre los principales mecanismos para desarrollar el talento existente se incluyen los siguientes elementos, con ejemplos adicionales facilitados en la Guía del G7 para la IA en el sector público de la OCDE/UNESCO (2024^[29]).

- **Los marcos de competencias** describen las principales habilidades y vías de aprendizaje de la organización, que deben adaptarse en función de las evaluaciones de necesidades y los grupos de usuarios descritos anteriormente. Esto también podría contribuir a la profesionalización de funciones clave de la IA para el gobierno. Por ejemplo, el Centro Común de Investigación (JRC, por sus siglas en inglés) de la CE ha desarrollado un marco global de competencias (Recuadro 4.6). India también ha desarrollado un marco de competencias específico para preparar a los servidores públicos a liderar la transformación de la IA de manera responsable, adaptando el contenido a los diferentes tipos de funcionarios³².
- **El aprendizaje formal** puede incluir cursos, talleres y módulos en línea. Por ejemplo, Irlanda ofrece varios cursos relevantes, entre ellos formación sobre la IA en la administración pública y sobre las directrices del país para el uso responsable de la IA en la administración pública, y el Ministerio del Interior griego, en colaboración con Google, ha creado cursos de formación en IA para servidores públicos³³. Los elementos del curso gratuito y abierto disponible en todo el mundo sobre IA pueden ayudar a mejorar la alfabetización fundacional en IA tanto de los servidores públicos como de la ciudadanía³⁴.
- **El aprendizaje informal**, incluso a través de comunidades de práctica, tutoría o rotaciones de empleo, entre otros, puede ayudar a crear conciencia e impulsar la adopción. Estos enfoques se analizan más adelante.

Recuadro 4.6. Marco de competencias de la Unión Europea para la IA en la administración pública

La UE ha desarrollado un marco global de competencias para orientar a los servidores públicos en la adopción y gestión eficaces de la IA. Este marco, elaborado por el Centro Común de Investigación (JRC, por su sigla en inglés), se basa en investigaciones empíricas, que incluyen revisiones bibliográficas, talleres de expertos y estudios de casos, e identifica las competencias clave necesarias para la integración de la IA en la administración pública.

El marco clasifica las competencias en tres dimensiones principales:

- Competencias técnicas: abarcar conocimientos y competencias relacionados con la gestión de datos, el aprendizaje automático y la implementación de sistemas de IA.
- competencias de gestión: abordar la propiedad de los proyectos, la intermediación en el intercambio de conocimientos y la toma de decisiones en iniciativas relacionadas con la IA.
- Competencias políticas, jurídicas y éticas: conocimientos sobre contratación pública en IA, auditoría y colaboración con expertos en el ámbito de la IA para garantizar el cumplimiento normativo y las consideraciones éticas.

Además, las competencias se agrupan en tres grupos transversales relacionados con: (i) competencias actitudinales (saber por qué), que se refieren a mentalidades y disposiciones que respaldan la adopción de la IA, como la curiosidad tecnológica y una cultura orientada a los datos; (ii) competencias operativas (saber hacer), que abarcan habilidades prácticas para la implementación de la IA, incluida la gestión de bases de datos, la formación en algoritmos y los procesos de toma de decisiones; y (iii)

competencias de alfabetización (saber qué), relacionadas con el conocimiento basado en hechos de los conceptos de la IA, los marcos regulatorios y los principios fundamentales del aprendizaje automático.

Además de desarrollar el marco, el trabajo del JRC contiene tres recomendaciones clave sobre competencias:

- Desarrollar programas de formación interdisciplinarios y centrados en las competencias en IA.
- Promover la investigación interdisciplinaria aplicada sobre competencias en IA.
- Establecer procesos de contratación específicos y destinar recursos adicionales a atraer a especialistas con competencias en IA.

Fuente: <https://publications.jrc.ec.europa.eu/repository/handle/JRC138702>.

El desarrollo interno de competencias relacionadas con la IA debe complementarse con estrategias externas para atraer talento de primer orden y retener a servidores públicos cualificados. El marco de competencias de la Unión Europea para la IA en el gobierno pone de relieve la necesidad de un enfoque estructurado para el desarrollo de competencias en IA, que haga hincapié en las competencias técnicas, de gestión y relacionadas con las políticas. La *Estrategia del Talento Digital* de Canadá se ajusta a estos principios y reconoce que la adopción de la IA en el sector público requiere un equilibrio de competencias actitudinales (saber por qué), operativas (saber hacer) y de alfabetización (saber qué) para formar una organización lista para la IA³⁵.

Para subsanar las principales carencias de competencias, es esencial llevar a cabo iniciativas de contratación específicas y procesos de contratación específicos. El modelo de competencias y orientación para la contratación basado en las competencias de la Oficina de Gestión del Personal de los Estados Unidos (OPM, por su sigla en inglés) en materia de IA ofrece un enfoque estructurado para definir y evaluar las clasificaciones de empleo de la IA³⁶, mientras que el marco de la UE subraya la importancia de los programas de formación interdisciplinaria y la investigación aplicada. Los gobiernos también deben centrarse en una remuneración competitiva, trayectorias profesionales claras y flexibilidad en el lugar de trabajo para atraer y retener talento en IA.

Cuando la capacidad interna sigue siendo limitada, las alianzas con la industria y el mundo académico, así como la contratación estratégica de expertos externos, pueden proporcionar el apoyo necesario. En las siguientes secciones, dedicadas a la contratación pública y las alianzas en el ámbito de la IA, se analizan más a fondo estas estrategias, ofreciendo ejemplos en materia de políticas. Además, es fundamental realizar una evaluación continua de la alineación de la organización, los avances en la reducción de las carencias en materia de competencias en IA y la efectividad de las iniciativas de aprendizaje. Al integrar la información basada en competencias procedente de las mejores prácticas mundiales, los gobiernos pueden garantizar que su fuerza laboral digital evolucione junto con los avances tecnológicos, las consideraciones éticas y los cambios en las demandas del mercado laboral.

Facilitar las conexiones y el intercambio de conocimientos

Las comunidades de práctica y las redes permiten la colaboración, aprendizaje, intercambio de conocimientos, a través de las fronteras de la organización, y la identificación de problemas colectivos o comunes. También pueden servir como un conducto útil para solicitar la opinión de los usuarios sobre los sistemas y servicios internos de IA. El Marco de competencias del JRC de la CE (Recuadro 4.6) señala que este tipo de actividades pueden ser fundamentales para ayudar a los servidores públicos a adquirir conocimientos sobre IA y para ayudar a superar los desafíos que plantea su adopción temprana. Este marco establece tres puntos de acción para desarrollar comunidades de práctica: 1) crear alianzas con las partes interesadas pertinentes, 2) desplegar plataformas digitales para la comunicación y la

colaboración de las entidades implicadas, y 3) financiar subvenciones por sinergia para la colaboración público-privada y el intercambio de conocimientos, que pueden ayudar a reducir las asimetrías de conocimiento y generar empresas conjuntas.

La difusión de métodos, estrategias y casos de uso de éxito a través de dichos métodos puede ayudar a las organizaciones gubernamentales a replicar y ampliar los proyectos de IA de manera más efectiva (OECD/UNESCO, 2024^[29]). Este enfoque ayuda a evitar errores comunes, ayuda a garantizar la coherencia y acelera la adopción de las tecnologías de IA en diversas entidades gubernamentales. Por ejemplo

- El programa de desarrollo de competencias de Estonia incluye una red de expertos en datos con más de 500 participantes, encuentros de IA y eventos de experimentación (por ejemplo, hackatones o competiciones) (OECD/UNESCO, 2024^[29]).
- La Red de Innovadores Públicos de Chile conecta a más de 30 000 servidores públicos de todos los niveles de gobierno y otros actores relevantes para el aprendizaje colectivo, la creación y la experimentación, incluso en materia de IA³⁷.
- La Conferencia de Datos de Canadá sirve como el principal foro para que los servidores públicos y los líderes de datos mejoren la concienciación, compartan conocimientos y avancen en las aplicaciones de datos en todo el gobierno canadiense.³⁸ Además, los grupos de trabajo dirigidos por departamentos sobre temas de IA permiten a los servidores públicos de diversos departamentos compartir experiencias e ideas, fomentando la colaboración y la innovación en la implementación de la IA (OECD/UNESCO, 2024^[29]).
- La Red Internacional de Ciudades Inteligentes, dirigida por Alemania, promueve el intercambio internacional y la transferencia de conocimientos a nivel nacional y local al servir como lugar para el diálogo internacional y el intercambio de ideas y mejores prácticas.³⁹
- Francia creó ALLIANCE y *Communauté des labos*, grupos interministeriales informales para compartir las mejores prácticas en materia de IA⁴⁰.
- En Suiza, la Red de Competencias sobre IA ha promovido comunidades de práctica para comprender los desafíos comunes en la implementación de sistemas de IA, incluida la administración pública⁴¹.

No es necesario que estas comunidades y redes se centren específicamente en la IA; de hecho, los grupos más generales pueden ayudar a sacar a la luz una base más amplia de cuestiones relevantes y considerar mejor enfoques alternativos. Sin embargo, es posible que los gobiernos deseen desarrollar más comunidades y redes centradas en la IA o garantizar que las comunidades y redes generales incluyan a personas con experiencia en IA, con el fin de ayudar a identificar vínculos entre problemas y enfoques de IA que puedan constituir una solución óptima.

Aunar perspectivas y competencias multidisciplinarias

Las competencias digitales y en materia de IA no son las únicas pertinentes para diseñar y utilizar la IA en la administración pública. Algunos gobiernos han intentado crear uno o más equipos multidisciplinarios para garantizar que las iniciativas de IA se beneficien de diversas perspectivas y conocimientos. La sensibilidad y complejidad en torno a la IA exigen la participación de expertos de diversas disciplinas, como la tecnología, la ética, el derecho y las políticas públicas, para establecer un enfoque estratégico del uso de la IA. Dichos equipos pueden aportar una diversidad de perspectivas y conocimientos, facilitando así la identificación de posibles riesgos y garantizando un uso integral e inclusivo de la IA en toda la administración pública (Berryhill et al., 2019^[4]).

Invertir con propósito

Los gobiernos están invirtiendo cada vez más en IA financiando iniciativas gubernamentales de IA. Las estimaciones indican que los gobiernos podrían aumentar su gasto anual en tecnologías relacionadas con la IA un 19% en 2025 y seguir aumentando a partir de entonces (Gartner, 2024^[93]). Es esencial que los gobiernos planifiquen, implementen y supervisen estratégicamente las inversiones en IA en el sector público para garantizar la optimización de los recursos, identificar y mitigar los posibles riesgos de inversión, implementar y desplegar tecnologías de forma oportuna y evaluar si se consiguen los beneficios previstos (OECD, 2025^[94]).

El Índice de Gobierno Digital (DGI) de la OCDE de 2023 (2024^[74]) muestra que los países todavía no han desarrollado capacidades sólidas para gestionar las inversiones digitales en el sector público. Si bien el 88% de los países de la OCDE han adoptado un enfoque estandarizado para desarrollar propuestas de valor, solo el 41% han desarrollado un mecanismo de evaluación de riesgos para las inversiones en gobierno digital, que incluye los riesgos operativos (por ejemplo, de ciberseguridad, de acceso a servicios o de interrupción relacionados con el uso de tecnologías heredadas) y financieros (por ejemplo, incertidumbre sobre el rendimiento de la inversión, sostenibilidad del financiamiento o gastos generales y de mantenimiento). Para avanzar hacia inversiones en IA fiables en el gobierno, los países pueden desarrollar mecanismos de garantía, como:

- Fortalecimiento de la planificación estratégica.
- Apoyar inversiones coherentes en toda la administración.
- Reforzar los mecanismos de supervisión de las inversiones.

La coordinación de todo el sector público en estos tres ámbitos permitirá a los gobiernos invertir en sistemas de IA capaces de cumplir los objetivos de política pública a tiempo y dentro del presupuesto. Las siguientes subsecciones revisan estos ámbitos.

Fortalecer la planificación estratégica para una inversión coherente

Los gobiernos deben coordinarse con las principales partes interesadas para planificar y gestionar las inversiones en IA sobre la base de principios claros. El establecimiento de tales principios puede ayudar a garantizar que las decisiones de inversión sean coherentes con los objetivos estratégicos generales. Por ejemplo, articular un compromiso para desarrollar servicios más proactivos podría dar lugar a un mayor apoyo financiero y de gestión para la implementación de chatbots de IA para la interacción entre el gobierno y la ciudadanía. Además, la coordinación entre las autoridades presupuestarias, las administraciones digitales y las encargadas de la contratación pública puede ayudar a identificar las necesidades de IA y alinearlas con los recursos disponibles y las posibles adquisiciones y alianzas con el sector privado. Alemania ha intentado conseguirlo a través de una declaración de misión sobre IA de 2024 que se complementa con un nuevo Centro de IA en la Administración Pública (BeKI, por su sigla en alemán) como organismo iniciador y coordinador de inversiones y orientación en IA en la administración federal (OECD, 2024^[95])⁴².

Los gobiernos también pueden utilizar las herramientas de gestión existentes, como los mecanismos de propuesta de valor y de evaluación de riesgos de inversión, para reforzar la garantía y garantizar la coherencia de las decisiones de inversión en sistemas de IA. La adaptación de los mecanismos de propuesta de valor a las especificidades de los sistemas de IA permite a los gobiernos reforzar los procesos de garantía para un desarrollo y uso confiable de la IA. Esto incluye la evaluación de aspectos clave de la IA, como el cumplimiento de la regulación y las normas políticas. La propuesta de valor puede incluir una evaluación de riesgos e impacto que mida y evalúe los beneficios y riesgos potenciales de los sistemas de IA en el gobierno, así como los planes para cumplir con las regulaciones. Por ejemplo, Australia ha publicado un Marco Piloto de Aseguramiento de la IA⁴³ para orientar a las agencias en la

alineación de los casos de uso de la IA con los Principios Éticos de la IA de Australia, identificar los impactos y riesgos, y aplicar medidas de mitigación⁴⁴.

Financiar la IA y apoyar a las inversiones coherentes en toda la administración

Aunque a menudo se pasan por alto en las estrategias nacionales de IA, los mecanismos de financiamiento son una consideración importante para las aplicaciones gubernamentales de la IA (van Noordt, Medaglia and Tangi, 2023^[96]). Incluso las iniciativas sencillas necesitan algún nivel de financiamiento y apoyo financiero para pasar de la idea a la realidad, y se necesitan muchos más fondos para ampliar un proyecto exitoso. La disponibilidad y la naturaleza de este financiamiento pueden contribuir en gran medida al éxito final de la innovación basada en la IA. Por el contrario, la falta de financiamiento para el desarrollo y la implementación de la IA es una barrera importante para la adopción de la IA por parte del gobierno (EC, 2024^[97]; UK NAO, 2024^[16]). Disponer de recursos financieros específicos puede respaldar la experimentación y la ampliación de la IA, así como ayudar a reducir la fragmentación de los esfuerzos y la adopción desigual de la IA. La Comisión Europea (CE) (2024^[97]) lo ha señalado recientemente en un estudio de 2024 sobre la adopción estratégica de la IA para los servicios públicos, en el que se recomienda a los gobiernos que aumenten el financiamiento y los recursos destinados a su uso de la IA. Algunos ejemplos de instrumentos de financiamiento específicos son:

- El Fondo de Modernización Tecnológica de Estados Unidos (TMF, por su sigla en inglés) lanzó una convocatoria especial de inversiones en IA para apoyar a las agencias públicas como parte de un mecanismo de financiamiento más amplio diseñado para reemplazar la TI heredada obsoleta⁴⁵.
- En Francia, el Fondo para la Transformación de la Acción Pública ofrece apoyo financiero a las instituciones del sector público que buscan mejorar las políticas y los servicios con IA, en particular para propuestas de proyectos con potencial de escalabilidad y replicabilidad en toda la administración pública⁴⁶.
- El Reino Unido está invirtiendo 110 millones de GBP para acelerar el uso de la IA en el gobierno, lo que incluye añadir capacidad a su incubadora de IA (i.AI), y se ha comprometido a proporcionar 10 millones de GBP para impulsar las capacidades de IA de los reguladores (Cover-Kus, 2024^[98]; UK House of Commons, 2024^[99]).
- En Polonia, se ha pedido a los departamentos gubernamentales que reserven un porcentaje de su presupuesto para adquisiciones de IA (van Noordt, Medaglia and Tangi, 2023^[96]).

Mecanismos de monitoreo para una inversión coherente

Garantizar el desarrollo y despliegue dentro del presupuesto y de los plazos de las inversiones en IA contribuye a la realización de los beneficios y a la obtención de los resultados previstos. En consonancia con las inversiones generales en tecnologías digitales, los gobiernos pueden aprovechar las herramientas de supervisión para supervisar la gestión y el desarrollo de los sistemas de IA en toda la Administración. Estas actividades deben contemplar el desarrollo de indicadores clave de rendimiento (KPI por sus siglas en inglés) y enfoques estructurados para gestionar los avances en curso en los sistemas de IA a través de la gestión de carteras de TI. Estas herramientas de gestión pueden crear o complementar salvaguardas para el desarrollo de la IA centradas en un desarrollo, despliegue y uso confiable de IA en la administración pública. Estos enfoques pueden tener en cuenta mecanismos de control de calidad, vinculándolos con la planificación y el seguimiento de las inversiones digitales para garantizar la coherencia en todo el ciclo de vida del sistema de IA. Los países han desarrollado directrices para integrar acciones de supervisión y medición en el ciclo de inversión en las iniciativas de IA. Por ejemplo, Francia utiliza una herramienta de seguimiento para los principales proyectos digitales estatales, incluidas las iniciativas de IA, que cuestan más de 9 millones EUR⁴⁷. Enumera proyectos estratégicos de TI y ayuda a identificar acciones para el éxito. La herramienta supervisa la distribución del proyecto por ministerio, fase de progreso, área funcional

y costo estimado (OECD/UNESCO, 2024^[29]). Sin embargo, la mayoría de los países siguen enfrentándose al reto de aplicar prácticas de supervisión continuas o ad hoc. Reconociendo la necesidad de desarrollar capacidades y planes específicos para el seguimiento de las políticas de IA, la Oficina del Auditor General (OAG, por su sigla en inglés) de Noruega comenzó a auditar el uso de la IA en el gobierno central como parte de su cartera de nuevas auditorías de rendimiento desde 2023. A nivel ejecutivo, el país está adoptando medidas para reforzar el seguimiento y la supervisión de la cartera de proyectos gubernamentales de IA mediante auditorías internas periódicas, el seguimiento del rendimiento y las evaluaciones de impacto (OECD, 2024^[100]).

Utilización de la contratación pública para obtener productos y servicios de IA y orientar el mercado

Los procesos y mecanismos de contratación pública adecuados son fundamentales para permitir un acceso óptimo y rentable a los sistemas de IA desarrollados por terceros, desde grandes empresas hasta empresas emergentes y emprendedores. Más allá de simplemente comprar soluciones o contratar recursos, la contratación pública sirve a un enfoque estratégico clave, que utiliza la compra como un puente para conectar las misiones y objetivos públicos con las necesidades y valores de la sociedad. Para desempeñar esta función de manera efectiva, los funcionarios de contratación pública deben llevar a cabo una evaluación exhaustiva de la coherencia de la IA con los objetivos internos, el cumplimiento de las normas de equidad y transparencia, la eficiencia de los recursos, la mitigación de riesgos (por ejemplo, sesgos o vulnerabilidades de seguridad), el participación con las partes interesadas y los grupos afectados, y el cumplimiento de los marcos legales y regulatorios pertinentes.

Al promover el despliegue efectivo de las tecnologías de IA, las entidades públicas podrían considerar la adopción de mecanismos de contratación que fomenten la agilidad, la iteración y la innovación. El proceso debe comenzar con una cuidadosa preparación y planificación para lograr procesos de adquisición flexibles y eficientes que fomenten una amplia participación que esté abierta y sea accesible para todos (UK DSIT, 2020^[27]). Esta fase preparatoria debería incluir:

- la creación de un equipo multidisciplinario para apoyar la adquisición de sistemas de IA
- una evaluación de los enfoques actuales en materia de datos y gobernanza para evaluar la preparación y las capacidades y recursos existentes para la formación y el uso efectivos de los sistemas de IA, según proceda
- una evaluación de los riesgos potenciales a lo largo del ciclo de vida de la IA y la identificación de las estrategias de mitigación asociadas.

Métodos de licitación ágiles e innovadores

Los métodos de contratación ágiles e innovadores brindan oportunidades para acelerar la adopción de nuevas tecnologías dentro de los gobiernos y sobre el desarrollo y uso confiable de la IA (Monteiro, Hlacs and Boéchat, 2024^[101]). Entre ellos se incluyen concursos tecnológicos, demostraciones, procesos de contratación basados en desafíos y diálogos competitivos (UK DSIT, 2020^[27]). Además, los acuerdos marco de políticas que establecen normas, prioridades y directrices generales de contratación pública —ya sea específicamente para la IA o con proveedores clave— pueden desempeñar un papel en la realización de negocios con el sector privado, incluso en el caso de países con un conjunto menos diverso de prácticas de contratación pública a su disposición. En Australia, por ejemplo, el gobierno utilizó un acuerdo existente con una empresa multinacional de tecnología para desplegar un proyecto piloto a gran escala de una solución de IA en toda su administración pública (Australia DTA, 2024^[102]). Los acuerdos marco sobre políticas también pueden adoptar la forma de entornos predefinidos que permitan la contratación de IA con arreglo a principios rectores más amplios. Por ejemplo, el programa Adopt AI de la Comisión Europea (2024^[97]) tiene por objeto modernizar la contratación pública de sistemas de IA

fomentando el diálogo entre los contratistas públicos y la industria europea de la IA. Promueve el entendimiento mutuo, impulsa la inversión en la industria y busca crear un espacio de datos para la contratación pública. El diálogo sectorial tiende un puente entre los contratantes que buscan soluciones y los proveedores que necesitan información sobre los planes de la administración pública (OECD/UNESCO, 2024^[29]).

La contratación pública como motor del bien público y la IA confiable

La contratación pública puede ser una herramienta estratégica para dar forma al mercado y garantizar que los sistemas de IA se ajustan a las normas gubernamentales. También desempeña un papel fundamental a la hora de establecer requisitos para los sistemas de IA que reflejen los valores públicos, garantizando la rendición de cuentas, la seguridad y la equidad en la adopción de la IA. Por ejemplo, la Comisión Europea estableció un modelo de cláusulas contractuales para poner a prueba los contratos de IA en 2023, que se actualizaron en 2025 con el fin de adaptarse a los requisitos de la Ley de IA de la UE y proporcionar directrices exhaustivas para las aplicaciones de alto riesgo y opciones personalizables para la IA no de alto riesgo (2023^[103]; 2025^[104])⁴⁸. Australia también ha establecido cláusulas modelo⁴⁹. A nivel subnacional, la ciudad de Barcelona (España), ha introducido cláusulas de contratación pública que hacen hincapié en la soberanía de los datos, garantizando que los datos recopilados de la ciudadanía, incluso por parte de empresas privadas, sigan siendo de propiedad pública (Berryhill et al., 2019^[4]). Las directrices de contratación pública son instrumentos de política que pueden influir en el uso de la IA en todo el mundo. Por ejemplo, ChileCompra, el organismo chileno de contratación pública ha introducido una nueva herramienta para garantizar que los sistemas de IA adquiridos sean responsables y éticos (véase el Recuadro 5.24)⁵⁰. A nivel internacional, iniciativas como el Foro Económico Mundial (WEF, por su sigla en inglés) (2025^[105]) "AI Procurement in a Box" ("Kit de Adquisición de IA") proporciona una guía estructurada para ayudar a los gobiernos a integrar las mejores prácticas de contratación de IA y alinear las adquisiciones de IA con marcos éticos y regulatorios. En abril de 2025, Estados Unidos publicó una nueva política de adquisiciones de IA que parece promover la contratación rápida y la eliminación de la burocracia innecesaria y los procesos de contratación obsoletos (Recuadro 4.7).

Recuadro 4.7. Impulsar adquisiciones eficientes de IA en Estados Unidos

En Estados Unidos, la Oficina de Gestión y Presupuesto de la Casa Blanca (OMB, por su sigla en inglés) emitió el memorándum *M-25-22 Driving Efficient Acquisition of Artificial Intelligence in Government (Impulsar adquisiciones eficientes de IA en Estados Unidos)* el 3 de abril de 2025. Este último incluye una variedad de requisitos y recomendaciones para las agencias federales con respecto a las adquisiciones de IA en seis fases de contratación. Una lista no exhaustiva de estos incluye:

- Identificación de los requisitos:** convocar un equipo multifuncional que informe sobre la adquisición de sistemas de IA y ayude a crear una lista inicial de posibles riesgos que deban evaluarse. En la medida de lo posible, considerar qué usos pueden ser "IA de alto impacto" (Recuadro 1.3).
- Investigación y planificación de mercado:** buscar capacidades de IA de última generación mediante la realización de estudios de mercado exhaustivos, en particular mediante el intercambio de conocimientos entre organismos y la consideración de capacidades novedosas a partir de nuevos participantes. Buscar demostraciones y pruebas detalladas de la IA potencialmente útil para evaluar a los proveedores e identificar los obstáculos a la rentabilidad a largo plazo. Utilizar técnicas basadas en el rendimiento para identificar los requisitos y los términos del contrato a fin de comprender y evaluar las reclamaciones de los proveedores.

3. **Desarrollo de solicitudes:** incluir en las solicitudes requisitos que protejan contra el bloqueo del proveedor y los términos relacionados con los derechos de PI y el uso legítimo de los datos gubernamentales, y cuando sea factible, las agencias deben ser transparentes con respecto a si el uso de la IA podría considerarse de "alto impacto" y lo que esto podría significar para el proveedor.
4. **Selección y adjudicación:** probar las soluciones propuestas para comprender sus capacidades y limitaciones. Por separado, las agencias deben evaluar las propuestas para identificar cualquier nuevo riesgo potencial relacionado con la IA que no se haya identificado previamente. Abordar, en los términos del contrato, cuando corresponda, los derechos de PI y el uso de datos gubernamentales, la privacidad, las protecciones de bloqueo de proveedores y los requisitos de cumplimiento para el memorándum analizado en el Recuadro 4.1, las pruebas y supervisión continuas, los requisitos de rendimiento de los proveedores.
5. **Administración de contratos:** ayudar a garantizar que los sistemas de IA estén autorizados por un funcionario adecuado antes de su despliegue, poner en marcha procesos de supervisión y seguimiento de los contratos para la ejecución de los contratos y para identificar y mitigar los riesgos emergentes. Organizar una evaluación periódica del valor del sistema o servicio de IA para el gobierno, teniendo en cuenta, en la medida de lo posible, la efectividad, la eficiencia, los riesgos, los costos de operaciones y mantenimiento, y las opiniones de las partes interesadas. Considere los criterios de extinción.
6. **Liquidación de contratos:** ayuda a garantizar la protección del bloqueo de proveedores, por ejemplo, garantizando derechos continuos y el acceso a cualquier dato o producto derivado.

Para ayudar a los organismos, un centro de un organismo público publicará una o varias guías de acceso público para ayudar al personal de adquisiciones en la adquisición de sistemas de IA, o bien creará un archivo digital a disposición de los servidores públicos que facilite el intercambio de información, conocimientos y recursos sobre las adquisiciones de IA (por ejemplo, buenas prácticas, herramientas, lenguaje para las cláusulas contractuales o costes negociados).

Fuente: (US OMB, 2025^[106])

Contratación pública para dar forma al mercado

Más allá de garantizar que los gobiernos adquieren IA confiable para su propio uso, la contratación pública puede actuar como una potente palanca para influir en la dinámica más general del mercado, impulsando la innovación y armonizando el desarrollo de sistemas de IA con principios de IA confiable. La contratación pública representa alrededor del 13% del PIB en los países de la OCDE (OECD, 2024^[107]). Al abordar las adquisiciones de manera estratégica, los gobiernos pueden usar el "peso económico del poder adquisitivo del gobierno" para fomentar el desarrollo de soluciones de IA que no solo satisfagan sus necesidades, sino que también promuevan una alineación más amplia con los estándares éticos y regulatorios (World Bank, 2025^[108]).

Ampliar el potencial de la IA mediante las alianzas

Los gobiernos pueden beneficiarse enormemente de las alianzas intersectoriales activas y en curso en las que cada sector tiene un papel y contribuciones concretos (OECD/CAF, 2022^[30]). Estas alianzas pueden facilitar la colaboración entre entidades públicas y especialistas en IA de otros sectores, incluidas empresas del sector privado y agentes no gubernamentales (por ejemplo, instituciones académicas o fundaciones), promoviendo el desarrollo y la aplicación de soluciones de vanguardia. Las asociaciones

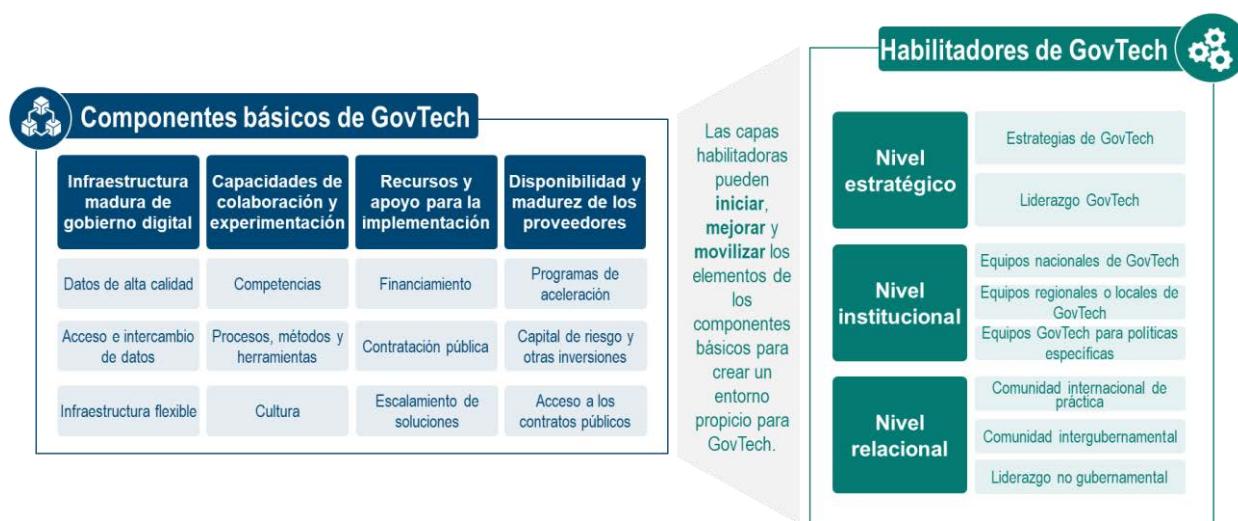
público-privadas (PPP, por su sigla en inglés) son quizás el tipo de acuerdo más común. A continuación, se presentan algunos ejemplos:

- La iniciativa InvestAI de la UE pretende movilizar 200 000 millones EUR en inversión en IA a través de una PPP similar a un CERN para la IA, con el fin de permitir el desarrollo de sistemas de IA punteros en todos los sectores⁵¹.
- Como se anunció en la Cumbre de Acción sobre IA de Francia de febrero de 2025, 10 países están desarrollando una plataforma e incubadora de IA de interés público para apoyar, amplificar y reducir la fragmentación entre las iniciativas públicas y privadas existentes sobre la IA de interés público y abordar las brechas digitales. Apoyará los bienes públicos digitales, la asistencia técnica y el desarrollo de capacidades para fomentar un ecosistema de IA confiable para el interés público en materia de IA⁵².
- El Observatorio de Datos de Chile (DO) es una organización sin fines de lucro liderada conjuntamente por el gobierno, la industria y el mundo académico. Funciona como un centro tecnológico que, a través de la gestión de grandes volúmenes de datos, busca contribuir al bienestar social promoviendo el desarrollo sostenible del país, contribuyendo a la generación de factores propicios para el desarrollo óptimo de la IA y promoviendo la creación de políticas públicas y la toma de decisiones estratégicas basadas en evidencia⁵³.
- El asistente virtual de servicios públicos de Portugal dirigido por ChatGPT (véase el Recuadro 5.46) se creó a través de un APP con el gobierno y varias empresas.
- El nuevo Centro de Inteligencia Artificial de Letonia es una fundación privada cofundada por el gobierno, el mundo académico y la industria, está diseñado para promover la adopción confiable y sostenible de la IA en todos los sectores, con especial atención a la integración de la IA en la administración pública, y para garantizar la incorporación del idioma y la cultura letones en los sistemas de IA⁵⁴.

Recurrir a las startups GovTech

Acercando los conceptos de contratación pública y alianzas, GovTech es la colaboración entre el gobierno y las empresas emergentes, los innovadores, los "intraemprendedores" gubernamentales y el mundo académico en soluciones innovadoras de gobierno digital. Complementa la capacidad gubernamental existente para procesos y servicios ágiles, centrados en el usuario, receptivos y rentables (OECD, 2024[109]). Su objetivo es contribuir a un gobierno ágil y mejorar la madurez del gobierno digital. Esto no solo contribuye a mejorar la efectividad y la eficiencia, sino que también puede fomentar la participación de empresas emergentes y proveedores más recientes en el mercado público. La innovación de GovTech se caracteriza por la cocreación y la experimentación. Estas interacciones colaborativas tienen como objetivo trascender las relaciones tradicionales proveedor-contratista para construir nuevas formas de asociación. En lugar de centrarse en los términos de referencia detallados y las especificaciones técnicas, el enfoque de GovTech se centra en los resultados esperados de la solución y en involucrar a los actores de GovTech en su construcción. Si bien muchas de estas colaboraciones fomentan la contratación pública (como se ha comentado anteriormente), también pueden utilizar subvenciones y premios monetarios para incentivar la creación de soluciones innovadoras (por ejemplo, a través de días de demostración o programas de incubación). La OCDE ha desarrollado un marco de políticas GovTech para describir los factores importantes para maximizar los compromisos de GovTech (Gráfico 4.4).

Gráfico 4.4. Marco de trabajo de políticas GovTech de la OCDE



Fuente: (OECD, 2024^[109]).

Los gobiernos pueden aprovechar las colaboraciones de GovTech para experimentar y desarrollar sistemas de IA con el fin de abordar desafíos gubernamentales y sociales. El Índice de Gobierno Digital (DGI) de la OCDE de 2023 muestra que el 42% de los 33 países de la OCDE encuestados están estableciendo objetivos GovTech para facilitar las pruebas y la adopción de tecnologías emergentes, incluida la IA. Por ejemplo, España (2024^[110]) está utilizando su GovTech Lab para desarrollar proyectos piloto de IA. Un reciente informe de la CE (2024^[111]) titulado *GovTech: Influencing factors, common requirements and recommendations* (*GovTech: Factores de influencia, requerimientos comunes y recomendaciones*) ofrece otros casos de uso y conclusiones.

Establecer salvaguardas para orientar la IA de forma estratégica y responsable

Las salvaguardas ayudan a garantizar un despliegue, desarrollo y uso confiable de la IA en la administración pública. Pueden ser **instrumentos de política vinculantes y no vinculantes, procesos de transparencia y mecanismos de rendición de cuentas, como los órganos de supervisión y control**. Las salvaguardas son esenciales para gestionar los riesgos asociados a la IA e implementar la IA de acuerdo con los límites legales y los valores sociales. En última instancia, esto ayuda a generar confianza en el sector público. En las siguientes secciones se examinan las principales salvaguardas, junto con las opciones de políticas disponibles que los gobiernos pueden considerar poner en marcha en sus propios contextos, basándose en ejemplos de buenas prácticas internacionales.

Sin embargo, es importante tener en cuenta que estas salvaguardas deben verse junto con los habilitadores anteriores. Las salvaguardas y su interpretación pueden ser un importante factor de aversión al riesgo, lo que contribuye a los riesgos de inacción y pérdida de oportunidades (véase el Capítulo 1). Los gobiernos también deberían considerar la importancia de eliminar o revisar las salvaguardas que ya no cumplen su propósito o causan consecuencias negativas que sobrepasan sus beneficios.

Por último, este apartado no pretende sugerir que los gobiernos deban establecer necesariamente todas las salvaguardas analizadas en este apartado, ni que deban aplicarse a todos los usos de la IA. Eso también podría contribuir a la aversión al riesgo y obstaculizar la adopción confiable de la IA en el gobierno. En su lugar, los gobiernos deben determinar qué salvaguardas se ajustan a sus operaciones y contextos

y aplicarlas a los usos de la IA de una manera basada en el riesgo y proporcional a su nivel potencial de riesgo.

Uso de instrumentos de política para orientar una IA confiable

Los instrumentos de política para promover una IA confiable pueden incluir **instrumentos no vinculantes**, como documentos de orientación, marcos éticos, normas técnicas y marcos de gestión de riesgos, así como **instrumentos vinculantes**, como leyes y reglamentos. Estos instrumentos de política trabajan juntos, de la mano con otras salvaguardas para proteger los derechos humanos. Lo hacen a través de enfoques de gestión de riesgos y fomentando el desarrollo y despliegue responsable de la IA. Ayudan a mitigar el uso indebido de la IA, los resultados sesgados, las infracciones de la privacidad y las consecuencias no deseadas, al tiempo que ofrecen herramientas prácticas para aplicar los principios de gobernanza y garantizar un rendimiento coherente en el uso de la IA (UNESCO, 2024^[112]). En general, el desarrollo de instrumentos de política para la IA en el gobierno debe seguir un conjunto de buenas prácticas para garantizar que promueven de manera efectiva una IA confiable. Los instrumentos de política deben:

- Alinearse con los principios éticos y valores sociales para garantizar que la tecnología sirva al bien común. Esto es importante para mantener la confianza pública, salvaguardar el libre ejercicio de los derechos humanos y garantizar que los sistemas de IA funcionen de manera justa y responsable.
- Tener en cuenta tanto la innovación como la gestión de riesgos, ayudando a las organizaciones gubernamentales a navegar de manera responsable por el cambiante panorama de la IA aprovechando el potencial transformador de la IA al tiempo que se abordan los riesgos y desafíos (OECD, 2023^[35]; 2021^[113]).
- Evaluar e identificar de forma continua los posibles riesgos asociados a los sistemas de IA. Las herramientas de evaluación de riesgos ofrecen un enfoque equilibrado, en el que pueden coexistir tanto el uso responsable como los avances de vanguardia, garantizando que la IA beneficie a la sociedad y cumpla con los estándares éticos (UNESCO, 2023^[114]).
- Involucrar a las partes interesadas a través de diferentes acuerdos institucionales para alinear el uso de la IA con las necesidades y los valores de aquellos a los que afecta (véase la sección Participación de este capítulo).

Instrumentos no vinculantes

Uno de los puntos de entrada más comunes para promover la IA confiable es la adopción o el desarrollo de principios o marcos éticos. Estos instrumentos establecen un conjunto de valores y buenas prácticas que guían el uso de la IA de forma transparente, justa y responsable. En todo el mundo se han desarrollado más de 200 instrumentos de este tipo (Corrêa et al., 2023^[115]). Sin embargo, y a menudo, tales principios están integrados en la estrategia nacional de IA de un país. También suelen abordar una amplia gama de preocupaciones éticas en torno a la IA, como los prejuicios, la transparencia, la rendición de cuentas y el impacto de la IA en la sociedad.

Se han logrado avances importantes en la gobernanza mundial de la IA, con varios instrumentos desarrollados por organizaciones intergubernamentales y supranacionales para normalizar y unificar el desarrollo de la IA a escala mundial. Entre ellos se incluyen los Principios de IA de la OCDE (véase la Tabla 2.2 del Capítulo 2), las Directrices Éticas de la CE para una IA confiable y los Principios rectores internacionales del G7 para los sistemas avanzados de IA en el marco del proceso de IA de Hiroshima.⁵⁵ La Unión Africana (UA) también está trabajando en su propia carta fundacional sobre IA confiable (OECD.AI, 2025^[116]). Otros esfuerzos internacionales no vinculantes incluyen la resolución de la Asamblea General de las Naciones Unidas de 2024 sobre la promoción de la "seguridad y la confianza", que fue

respaldada por más de 120 países⁵⁶ así como declaraciones de cumbres internacionales sobre IA, que se han celebrado en el Reino Unido, Corea y Francia⁵⁷. A nivel nacional, los gobiernos también establecen sus propios marcos éticos. Por ejemplo:

- Australia cuenta con los "Principios éticos de inteligencia artificial", diseñados para instar a las organizaciones a considerar el impacto del uso de sistemas basados en la IA y ayudar a las empresas y los gobiernos a aplicar los más altos estándares éticos en el diseño, el desarrollo y la implementación de la IA. Sobre la base de este marco, la Norma Voluntaria de Seguridad en IA (VAISS, por su sigla en inglés) de Australia ofrece orientación práctica a todas las organizaciones australianas sobre cómo utilizar e innovar con IA de forma segura y responsable⁵⁸.
- En Colombia, el "Marco ético para la inteligencia artificial" ofrece un conjunto de principios y una metodología que deben tenerse en cuenta en el diseño, desarrollo e implementación de sistemas de IA (OECD, 2024^[3])⁵⁹.
- Egipto ha elaborado una "Carta Egipcia para la IA Responsable" en torno a principios clave.⁶⁰

Los gobiernos elaboran cada vez más documentos de orientación para el uso de la IA. Estas son comparables a las directrices anteriormente descritas, que sirven como habilitadores, pero se centran más en establecer los parámetros para un uso confiable de la IA. Estos documentos, normalmente dirigidos a servidores públicos responsables del desarrollo de proyectos de IA y de gestionar una amplia recopilación y análisis de datos, pretenden dotar a los servidores públicos de los conocimientos necesarios para dar forma ética a los proyectos de IA y concienciar sobre los posibles riesgos, incluidas, entre otras, las violaciones de los datos personales (OECD/UNESCO, 2024^[29]). Estas directrices son más concretas que los principios, ya que suelen abordar aspectos técnicos que pueden afectar al despliegue de la IA. Por ejemplo:

- En el Reino Unido, el gobierno y el Instituto Alan Turing desarrollaron conjuntamente una guía para "Comprender la ética y la seguridad de la IA" (2019^[117]). En Canadá, la "Guía sobre el uso de la IA generativa" sirve de recurso a las instituciones federales que utilizan tecnologías de IA generativa (OECD/UNESCO, 2024^[29])⁶¹.
- Alemania cuenta con dos conjuntos principales de directrices para garantizar el uso ético de la IA en los servicios públicos: las Directrices para el uso de la IA en los servicios de empleo y protección social; y las Directrices de la IA para la administración federal (OECD/UNESCO, 2024^[29]; Policy Lab Digital, Work & Society within the German Federal Ministry of Labour and Social Affairs, 2024^[118]). Este último ya ha sido publicado, mientras que el primero está en desarrollo⁶².

Es importante señalar que las medidas no vinculantes tienen un alcance limitado. Por ejemplo, en lo que respecta a la IA en el lugar de trabajo, las medidas específicas de la mayoría de los países de la OCDE para promover una IA confiable en el lugar de trabajo no son principalmente vinculantes y dependen de la capacidad de las organizaciones para autorregularse (OECD, 2023^[119]). Debido a su naturaleza no vinculante, este tipo de instrumentos también llamados leyes blandas, puede no ser suficiente para prevenir o reparar los daños relacionados con la IA en el lugar de trabajo. Los gobiernos también deberían considerar medidas vinculantes para superar esta limitación en áreas importantes.

Instrumentos vinculantes

Hasta la fecha, la mayoría de las medidas vinculantes para la IA en el gobierno se han puesto en marcha a nivel nacional y subnacional, como se analiza a continuación. Sin embargo, algunos **mecanismos internacionales** han entrado en acción recientemente. Quizás el ejemplo más notable sea el Reglamento de IA de la UE (Recuadro 1.2). Más recientemente, el Convenio Marco del Consejo de Europa sobre IA y Derechos Humanos, Democracia y Estado de Derecho (2024^[120]) fue aprobado como el primer tratado internacional jurídicamente vinculante sobre la IA. Abierto a la firma en septiembre de 2024, se aplica tanto al sector público como al privado. Su objetivo es garantizar que las actividades dentro del ciclo de

vida de los sistemas de IA sean plenamente coherentes con los derechos humanos, la democracia y el Estado de Derecho, al tiempo que propician el progreso tecnológico y la innovación. En septiembre de 2025, el tratado contaba con 17 signatarios, entre ellos la Unión Europea y el Reino Unido, así como países no europeos, como Canadá, Japón y Estados Unidos.

Las leyes y regulaciones nacionales sobre IA pueden regular las actividades a lo largo del ciclo de vida de la IA y abordar cuestiones como la protección de los datos, la privacidad, el uso indebido y otras preocupaciones. Estas normas ayudan a garantizar que el desarrollo de la IA se ajuste a los valores sociales y a las normas jurídicas, al ofrecer a los desarrolladores una orientación clara sobre el cumplimiento normativo y al establecer límites a través de normas vinculantes en materia de transparencia, rendición de cuentas y equidad (OECD, 2025^[121]). También pueden definir la responsabilidad con respecto a los resultados de la IA y promover la coherencia y cooperación entre jurisdicciones mediante la armonización de las normas. En última instancia, unas leyes y reglamentos adecuados y bien ajustados ayudan a promover la innovación, al tiempo que instauran las protecciones necesarias, garantizando que la IA sirva al interés público. Estos instrumentos vinculantes pueden ser leyes generales de gobernanza de la IA que afecten a todos los sectores, leyes específicas de la IA centradas en el uso gubernamental y leyes transversales relacionadas con la IA pero no específicamente dirigidas a ella.

Sin embargo, los gobiernos deben tener en cuenta la naturaleza dinámica de la IA en el presente y sus muchas trayectorias potenciales en el futuro. Como se analiza en el Capítulo 3, los servidores públicos se enfrentan a desafíos debido a las normas confusas u obsoletas que limitan su capacidad para adoptar la IA. Al elaborar instrumentos nacionales vinculantes, deben tratar de alinearse con la *Recomendación de la OCDE sobre Gobernanza regulatoria ágil para aprovechar la innovación* (2021^[122]). El Reino Unido ha intentado conseguirlo mediante su "enfoque a favor de la innovación en la regulación de la IA"⁶³.

Las leyes generales sobre gobernanza de la IA son normas generales que rigen los sistemas de IA en diversos sectores económicos, ámbitos de actuación y regiones. Se centran en establecer marcos para la gestión de riesgos de la IA, el uso ético y el impacto social. Si bien muchos países cuentan con leyes que pueden influir en la IA (por ejemplo, leyes de protección de datos), pocos cuentan con leyes formales específicas sobre la IA. Las leyes y regulaciones nacionales existentes o propuestas incluyen:

- Ley básica de Corea sobre el desarrollo de la IA y el establecimiento de una base de confianza, o "Ley básica de IA" (2024^[123]), que entrará en vigor en enero de 2026, establece un marco integral para promover la innovación en IA y, al mismo tiempo, garantizar normas éticas, seguridad y confianza pública para todas las organizaciones que utilizan IA en el mercado coreano.
- Bahréin y Omán presentaron un proyecto de legislación sobre IA en 2024, tras lo cual Omán aplicó una consulta pública⁶⁴.
- Gobiernos de América Latina (como Argentina, Brasil, Chile, Colombia, Costa Rica, Ecuador, México, Panamá, Perú y Uruguay) están debatiendo legislación general sobre IA (UNESCO, 2024^[112]).

También pueden elaborarse leyes y reglamentos específicos que rijan el uso de los sistemas de IA en la administración. Suelen centrarse en la transparencia, la rendición de cuentas, la gobernanza de la IA en las organizaciones públicas, las consideraciones éticas y el uso responsable de IA en funciones y servicios. En los Estados Unidos, a nivel subnacional, los Estados han tomado medidas en esta dirección. Por ejemplo, Nueva York introdujo la Ley LOADING en 2024 para limitar el uso de sistemas automatizados de toma de decisiones por parte de los organismos estatales y proporcionar algunas protecciones a los servidores públicos en relación con la IA (Werner, 2024^[124]). Delaware ha reforzado la supervisión de la IA y la IA generativa en el estado (2024^[125]). Sin embargo, la adopción de leyes o reglamentos específicos sobre IA para uso gubernamental sigue siendo limitada, y la mayoría de los marcos surgen como parte de iniciativas más amplias de gobernanza de la IA. Además de leyes y regulaciones estrictas, la orientación formal en materia de políticas también puede proporcionar reglas vinculantes para las organizaciones

gubernamentales. Por ejemplo, el memorándum de Estados Unidos de 2025 sobre "Aceleración del uso federal de la IA a través de la innovación, la gobernanza y la confianza pública" abarca una serie de cuestiones y medidas para garantizar una IA confiable en el gobierno y, al mismo tiempo, eliminar los obstáculos a la innovación (Recuadro 4.1)⁶⁵.

Las leyes transversales son leyes más amplias que, si bien no son exclusivas de la IA, tienen implicancias significativas para su implementación y uso en el gobierno. Entre ellas podrían figurar regulaciones sobre protección de datos, privacidad, riesgos cibernéticos y derechos humanos, que determinan cómo pueden usarse los sistemas de IA en contextos del sector público. Por ejemplo, la legislación en materia de protección de datos y privacidad salvaguarda la información personal estableciendo normas sobre la forma en que los sistemas de IA recopilan, almacenan y procesan los datos, garantizando así el respeto de los derechos de privacidad de las personas. Si bien no suelen estar diseñadas específicamente para la IA, tienen implicancias significativas para su uso. Por ejemplo, el RGPD de la UE establece requisitos específicos para la gestión de datos personales en todos los sectores, que abarcan aspectos como la recopilación de datos, el almacenamiento, el tratamiento y los derechos de las personas.

Fomento de la transparencia en el uso de la IA por parte del sector público

Para ser transparentes sobre el uso de la IA, los gobiernos deben, en la medida en que sea factible y apropiado, hacer que los algoritmos sean abiertos, comprensibles y accesibles para el escrutinio público, así como divulgar los procesos y decisiones a los que contribuyen los sistemas de IA. Esto significa que los gobiernos deben facilitar información clara y adecuada para cada contexto sobre el funcionamiento de sus sistemas de IA, los datos que utilizan y la forma en que llegan a conclusiones y resultados, así como mecanismos para cuestionar los resultados. Dicha transparencia permite a las partes interesadas acceder a la información, tomar decisiones informadas y, si es necesario, buscar reparación por los posibles daños. Diseñar un enfoque coherente, sostenible e impactante sobre la transparencia algorítmica de la IA implica divulgar información clave sobre los sistemas y algoritmos de IA utilizados. Esto incluye activos complementarios como los datos de formación; la colaboración con un conjunto amplio y diverso de partes interesadas para garantizar que se tengan en cuenta sus necesidades e inquietudes (véase la sección "Participación" de este capítulo); la promoción de la alfabetización en IA para empoderar a las comunidades a fin de que tengan una voz informada en las cuestiones que les afectan; y el fortalecimiento de las normas nacionales existentes sobre transparencia y rendición de cuentas para abordar de manera efectiva los desafíos y riesgos planteados por la IA.

La transparencia no solo promueve la confianza y mejora el valor público, sino que también sustenta la rendición de cuentas por parte del gobierno. Como se contempla en los Principios de IA de la OCDE, la transparencia y rendición de cuentas son dos conceptos diferentes pero complementarios. La transparencia permite una supervisión más informada, fomenta la confianza y aumenta la rendición de cuentas de quienes desarrollan o controlan estos sistemas. La transparencia también mejora la equidad de los sistemas impulsados por la IA y ayuda a garantizar que su implementación pueda supervisarse y evaluarse de manera efectiva, especialmente cuando las decisiones tienen un impacto directo en la vida de las personas en ámbitos como la atención sanitaria, las finanzas y la justicia penal.

Los países deben emplear diversas políticas, herramientas, métodos y enfoques en materia de transparencia que se adapten a su público y proporcionen información clara. En términos generales, estos instrumentos pueden clasificarse como **proactivos o reactivos** (GPAI, 2024^[126]).

Instrumentos proactivos de transparencia

En general, "la comprensión del sector público sobre su propio uso de IA es muy deficiente, lo que dificulta tanto la rendición de cuentas democrática como el intercambio interno de conocimientos" (Ada Lovelace Institute, 2025^[23]). Se necesitan instrumentos que permitan a los gobiernos entender su propio uso de la

IA y, por extensión, que les permitan compartir proactivamente información sobre los sistemas de IA utilizados en la administración pública sin que se les solicite (GPAI, 2024^[126]). Entre las opciones de actuación en materia de transparencia proactiva se incluyen los registros públicos de sistemas de IA, la publicación de código fuente y documentación de algoritmos, las publicaciones proactivas impulsadas por el usuario y las respuestas automatizadas activadas por interacciones.

Los registros públicos de sistemas de IA son cada vez más comunes, ya que sirven como repositorios centralizados que consolidan la información sobre los sistemas de IA utilizados actualmente en el gobierno. El objetivo es crear una "ventanilla única" en la que la ciudadanía y las partes interesadas puedan acceder fácilmente a información sobre los sistemas de IA en uso, sus fines, los sectores a los que se aplican y las jurisdicciones a las que afectan. Algunos ejemplos son:

- Conjunto de datos de Colombia sobre sistemas automatizados de decisión en la administración pública colombiana⁶⁶.
- Registros algorítmicos de transparencia del Reino Unido⁶⁷.
- Inventario de casos de uso de IA del gobierno estadounidense, que las agencias federales deben actualizar al menos una vez al año (US OMB, 2025^[17])⁶⁸.
- Inventarios de algoritmos públicos de gobiernos nacionales en Chile, Francia y Países Bajos
Inventario de algoritmos públicos⁶⁹.
- Algoritmo subnacional registra Ámsterdam, Países Bajos, y Helsinki, Finlandia⁷⁰.

El desarrollo de un registro central y público de sistemas de IA en el que se puedan realizar búsquedas es una buena práctica que mejora la transparencia. Sin embargo, hacerlo puede resultar complicado, en parte debido a la rapidez con la que los gobiernos están implantando sistemas de IA en diversos ámbitos. En Chile, las dificultades para hacer transparentes los usos de la IA provocaron que el Consejo de Transparencia chileno, un organismo independiente creado por ley, emitiera recomendaciones sobre la mejora de la transparencia algorítmica en el gobierno (2024^[127]). En los Países Bajos, hasta octubre de 2024, solo alrededor del 5 % de los sistemas de IA se habían publicado en el registro neerlandés de IA (Netherlands Court of Audit^[128]). En el Reino Unido, el Comité de Cuentas Públicas (PAC) (2025^[129]) descubrió que se habían publicado relativamente pocos (33) registros algorítmicos de transparencia, lo que ponía en peligro la confianza pública en la adopción de la IA en el gobierno. Sin embargo, estos registros podrían completarse automáticamente, dependiendo de si el gobierno lleva a cabo evaluaciones de impacto y riesgo y de cómo lo hace (OECD, 2024^[95]). Por ejemplo, la Directiva canadiense sobre la toma de decisiones automatizada exige la realización de una evaluación de impacto algorítmico (EIA) para los sistemas de decisión automatizada. Los resultados de la evaluación del impacto económico deben publicarse como información abierta en el Portal de gobierno abierto de Canadá. Si fuera necesario para todos los sistemas de IA, dicho proceso podría llenar automáticamente un registro público.

La publicación del código fuente y la documentación del algoritmo también fomenta la transparencia. El código de código abierto para algoritmos públicos se considera una buena práctica en transparencia algorítmica y es especialmente valioso para el público técnico y experto (Ada Lovelace Institute, 2021^[130]). Esto permite a quienes poseen las habilidades necesarias examinar, probar y verificar el funcionamiento de estos sistemas, promoviendo la rendición de cuentas y la confianza. Algunos esfuerzos son un paso más allá de la publicación del código fuente completo, pero requieren la publicación de documentación exhaustiva que puede tener un efecto similar.

- En Francia, la ley de la República Digital obliga a los organismos gubernamentales a "poner a disposición del público, en un formato abierto y fácilmente reutilizable, las normas que definen el principal procesamiento algorítmico utilizado en el cumplimiento de su misión, cuando dicho procesamiento es la base de decisiones individuales"⁷¹.

- El Estándar de Registro de Transparencia Algorítmica (ATRS, por su sigla en inglés) del Reino Unido obliga a las organizaciones del sector público a divulgar de forma transparente los detalles sobre su uso de métodos algorítmicos en los procesos de toma de decisiones (OECD, 2023^[131])⁷².
- En Canadá, la Directiva sobre la toma de decisiones automatizada establece explícitamente, de forma detallada, requisitos de explicabilidad para los sistemas de IA diferenciados por niveles de riesgo, determinados por el uso de una herramienta de evaluación algorítmica de impacto (cuyos resultados también deben publicarse)⁷³.
- El memorándum estadounidense analizado en el Recuadro 4.1 exige que los organismos del gobierno federal, cuando sea factible y esté sujeto a algunas exclusiones, publiquen y mantengan código de IA como software de código abierto en un repositorio público.

Los esfuerzos en materia de transparencia también pueden ser iterativos, como se observa en las **publicaciones proactivas impulsadas por los usuarios**. Este tipo de publicaciones proactivas implica que las entidades públicas opten por divulgar información de forma proactiva después de recibir numerosas solicitudes similares. Al publicar esta información de forma proactiva, se evitan futuras solicitudes, lo que ahorra tiempo tanto a los funcionarios como a los solicitantes. Aunque su uso para garantizar la transparencia algorítmica no está bien documentado, este enfoque podría ser un método pertinente y rentable para divulgar la información relacionada con los algoritmos que se solicita con frecuencia (GPAI, 2024^[126]).

Es posible que algunos tipos de divulgaciones solo se realicen para algunos usuarios en situaciones específicas del contexto. **Las respuestas automatizadas desencadenadas por interacciones** ocurren cuando la información sobre un sistema automatizado de toma de decisiones se proporciona automáticamente durante procesos gubernamentales específicos. Por ejemplo, cuando alguien interactúa con el sitio web o la plataforma en línea de un organismo público en relación con un servicio o procedimiento administrativo que implica un sistema automatizado de toma de decisiones, la información pertinente sobre el sistema podría divulgarse automáticamente, sin que el usuario tenga que solicitarlo explícitamente (GPAI, 2024^[126]).

Instrumentos de transparencia reactivos

Los instrumentos de transparencia de la divulgación de información reactiva permiten al gobierno responder a solicitudes específicas de información de personas, grupos o autoridades. A diferencia de la divulgación proactiva, este enfoque se inicia a través de la demanda externa, en lugar de la divulgación proactiva del gobierno (GPAI, 2024^[126]). Para ello, es preciso presentar una solicitud en virtud de la ley de acceso a la información (ATI, por su sigla en inglés) pertinente del país para obtener información sobre un algoritmo o su uso, aprovechando un instrumento de política ya existente ampliamente disponible en la mayoría de los contextos o países⁷⁴. Sin embargo, estos regímenes no están diseñados específicamente para la transparencia algorítmica y pueden ser ineficaces cuando se aplican en este contexto (Valderrama, Hermosilla and Garrido, 2023^[132]). Por ejemplo, es poco probable que solicitar información sobre el código fuente de un algoritmo o sobre su aplicación produzca los resultados deseados debido a problemas con las prácticas de gestión de registros y a excepciones habituales de las leyes de la ATI, como conflictos con las restricciones de propiedad intelectual y los secretos comerciales de proveedores privados de servicios públicos (Fink, 2017^[133]; Brauneis and Goodman, 2017^[134]).

Mayor rendición de cuentas mediante la gestión de riesgos a lo largo del ciclo de vida de los sistemas de IA

Para algunos sistemas de IA gubernamentales, el contexto de su desarrollo o uso puede plantear un mayor riesgo. Esto puede estar relacionado con su escala (gravedad y probabilidad de impacto negativo), alcance (ámbito de aplicación, como el número de personas afectadas) u optionalidad (grado de elección

de si deben estar sujetas a los efectos de un sistema de IA) (OECD, 2022^[58]). Los procedimientos de gestión de riesgos pueden ayudar a identificar qué sistemas o contextos plantean mayores riesgos para mitigarlos (OECD, 2023^[135]).

La gestión de riesgos para los sistemas de IA que pueden conllevar elevados riesgos debe fundamentarse en orientaciones sobre qué niveles de riesgo son aceptables para diferentes usos y contextos. La gestión del riesgo es necesaria tanto antes —por ejemplo, mediante evaluaciones ex ante del impacto y del riesgo— como después del despliegue de los sistemas de IA. Uno de los ejemplos más conocidos es el Marco de gestión de riesgos de la IA del Instituto Nacional de Estándares y Tecnología (NIST, por su sigla en inglés) de los Estados Unidos (2023^[136]). Este marco ayuda a las organizaciones públicas o privadas a identificar los riesgos únicos que plantea la IA generativa y propone acciones para una gestión de riesgos de la IA generativa que se ajusten mejor a sus objetivos y prioridades. Aunque está diseñado para los EE.UU, se ha traducido a varios otros idiomas y se ha utilizado en otros países. El Código de Conducta y los Principios Rectores Internacionales del Proceso de IA de Hiroshima del G7 para los Sistemas Avanzados de IA también establecen normas de referencia para gestionar los riesgos (2023^[137]; ^[138]). EE.UU también exige evaluaciones de riesgos para el uso de la IA y establece prácticas de gestión de riesgos para los usos considerados de "alto impacto" (véase el recuadro 1.3) (US OMB, 2025^[17]). Otro ejemplo nacional es la Oficina de Transformación Digital de Türkiye, que lleva a cabo los estudios "Recomendación sobre la gestión del riesgo en el ámbito de la IA" y "Sello de IA confiable" para supervisar de cerca el uso de la IA en beneficio del público (OECD, 2024^[3]).

Los expertos en IA recomiendan que los gobiernos prioricen el establecimiento o adopción de dichos procesos para mitigar los daños causados por la IA (OECD, 2024^[139]). Sin embargo, la proliferación de marcos puede dificultar que los gobiernos determinen cuál es el más adecuado para seguir. Dado que las solicitudes de desarrollo de marcos de gestión de riesgos siguen creciendo, la interoperabilidad mejoraría la eficiencia y reduciría los costos de cumplimiento y aplicación. La OCDE (2023^[140]) está trabajando activamente para promover la coherencia de las políticas y la interoperabilidad entre estos marcos⁷⁵.

Evaluaciones de impacto

Las evaluaciones de impacto, incluidas las evaluaciones de impacto algorítmico (EIA), pueden ayudar a las organizaciones públicas a anticipar y evaluar cómo puede funcionar un algoritmo en un contexto específico. Se trata de evaluaciones de un sistema de IA que utilizan avisos, talleres, documentos y conversaciones con los desarrolladores de un sistema de IA y otras partes interesadas para explorar cómo el sistema afectará a las personas o a la sociedad de manera positiva o negativa (Valderrama, Hermosilla and Garrido, 2023^[132]). Estos suelen producirse en las primeras fases de desarrollo de un sistema antes de su utilización (ex ante), pero pueden ocurrir después de su implementación (ex post).

El objetivo principal de las EIA ex ante es evaluar los posibles impactos de un sistema algorítmico en las economías y las sociedades y proporcionar un mecanismo de rendición de cuentas (Valderrama, Hermosilla and Garrido, 2023^[132]). Las EIA también ayudan a comprender, clasificar y mitigar mejor los posibles riesgos o daños asociados al algoritmo. Un ejemplo que emplea esa técnica es la "Directiva sobre la toma de decisiones automatizada" de Canadá, que requiere una evaluación del impacto ambiental que tenga en cuenta diversos factores y, a su vez, proporcione una puntuación de riesgo que prescriba determinadas acciones. El enfoque se ha adaptado en otros países, como Uruguay, donde se basó en su Guía para el estudio de impacto algorítmico (OECD/CAF, 2022^[30]). En 2024, el Consejo de Europa publicó la Evaluación de Impacto sobre los Derechos Humanos, la Democracia y el Estado de Derecho (HUDERIA)⁷⁶. Su metodología prevé la creación de una evaluación de riesgos y un plan de mitigación para minimizar o eliminar los riesgos identificados, protegiendo al público de posibles daños.

Las EIA ex ante son el enfoque más común en uso en la actualidad. En general, la OCDE ha apoyado este enfoque porque ayuda a convertir los principios en acciones⁷⁷. Sin embargo, algunos sostienen que evaluar los impactos no es lo mismo que evaluar los daños y que, en algunos casos, hacerlo puede ocultar

los daños reales. Esto se debe en parte a que las métricas utilizadas para las evaluaciones de impacto a menudo no miden el daño emocional o psicológico (Gupta et al., 2021^[141]). Este trabajo también sugiere que EIA AIA a menudo no tienen en cuenta las voces de todos los que pueden verse afectados por un sistema de IA. Otros críticos de las evaluaciones de impacto argumentan que estas evaluaciones no están diseñadas para monitorear continuamente los efectos de los sistemas implementados y adaptarse en consecuencia en función de los comentarios y las ramificaciones del mundo real (Mehta, Rogers and Gilbert, 2023^[142]). Hallazgos recientes del Instituto Ada Lovelace (2025^[23]). Sin embargo, las lecciones aprendidas durante los seis años de estudio de la IA en el gobierno subrayan aún más este punto. Uno de sus principales hallazgos es que "la IA es 'sociotécnica', en la medida en que influye y se ve influenciada por los contextos sociales en los que se despliega, a menudo con efectos dominó no deseados. El éxito y la aceptación de las herramientas de IA dependen de su interacción con los sistemas sociales, los valores y la confianza existentes. Centrarse exclusivamente en criterios técnicos sin tener en cuenta estos factores puede llevar al escepticismo y, en última instancia, dificultar su adopción y uso".

Para garantizar que las EIA sean valiosas, los gobiernos tendrán que realizar evaluaciones sociotécnicas exhaustivas que involucren a las partes interesadas apropiadas e integren diversas perspectivas (Lam et al., 2023^[143]). Sin embargo, las evaluaciones ex ante por sí solas suelen ser insuficientes; deben complementarse con evaluaciones de impacto ex post que se basen en las evaluaciones de impacto ex ante. Esto implica el desarrollo de mecanismos de seguimiento, adaptación y rendición de cuentas continuos, que garanticen que los sistemas de IA evolucionan en respuesta a las evidencias del mundo real. El NIST de EE. UU. ha puesto en marcha la herramienta Evaluación de Riesgos e Impactos de la IA (ARIA, por su sigla en inglés) para avanzar en las pruebas sociotécnicas y la evaluación de la IA⁷⁸.

En lo que respecta a las evaluaciones de impacto ex ante, los gobiernos deben considerar de antemano si la IA es la mejor solución para abordar un problema determinado, como se analiza en la sección de habilitadores sobre "Determinar si la IA es la mejor solución".

Auditorías algorítmicas

Después del despliegue de un sistema de IA, es importante que los gobiernos continúen monitoreando sus comportamientos para determinar qué riesgos esperados o inesperados se pueden estar materializando, y para garantizar que las organizaciones gubernamentales continúen con la implementación de manera responsable. Esto suele adoptar la forma de auditorías algorítmicas. Implican un control independiente, generalmente externo, de un sistema de IA o de los procesos que lo rodean. Pueden ser auditorías técnicas de las entradas o salidas del sistema; auditorías de cumplimiento para determinar si un equipo de desarrollo de IA ha completado los procesos o los requisitos reglamentarios; inspecciones regulatorias por parte de los reguladores para supervisar el comportamiento de un sistema de IA a lo largo del tiempo; o auditorías sociotécnicas que evalúen las formas en que un sistema está afectando a procesos y contextos sociales más amplios en los que está operando. Dado que las auditorías suelen llevarse a cabo después de que un sistema está en uso, sirven como mecanismos de rendición de cuentas para verificar si un sistema se comporta como los desarrolladores pretenden o afirman. En el se pueden ver ejemplos de procesos de auditoría algorítmica en el Recuadro 4.8.

Sin embargo, los gobiernos deben diseñar cuidadosamente sus auditorías. Las deficiencias en la auditoría de la IA podrían crear una falsa confianza y "ocultar problemas con los sistemas algorítmicos y crear una estructura de permisos en torno a una IA mal diseñada o implementada" (Goodman and Trehu, 2022^[144]). Algunos expertos sostienen que una auditoría insuficiente puede no tener sentido o exacerbar los problemas que se pretende abordar, además de utilizarse como "lavado de cuentas" para dar la apariencia de debida diligencia.

Evaluaciones de riesgos de las capacidades de los sistemas de IA

Las evaluaciones de riesgos de las capacidades de los sistemas de IA son similares a las evaluaciones de impacto, pero analizan específicamente la probabilidad de que se produzcan resultados perjudiciales en un sistema de IA debido a las capacidades de un sistema. Estos también suelen ocurrir en las primeras etapas del desarrollo de un sistema antes de que esté en uso, pero pueden ocurrir después de que se haya implementado un sistema. Dichos enfoques deben tener en cuenta los riesgos relacionados con las limitaciones y capacidades de los sistemas de IA, así como los contextos de uso. Por ejemplo, el gobierno de Queensland, Australia, emitió una "directriz fundacional de evaluación de riesgos en materia de IA" para los servidores públicos⁷⁹. La evaluación de riesgos se ha vuelto cada vez más común en el sector privado para los sistemas de IA avanzados, como en el caso de las "políticas de escalado responsable" (RSP, por su sigla en inglés), que se comprometen a adoptar medidas basadas en la evaluación de riesgos de las capacidades de los sistemas de IA (OECD, 2024^[139]). Al identificar capacidades potencialmente peligrosas, los RSP suelen establecer umbrales que desencadenan acciones para ralentizar o detener el desarrollo (METR, 2023^[145])⁸⁰.

Los organismos de ensayo y evaluación y los organismos nacionales, multilaterales o regionales, como los institutos de seguridad y protección de la IA, desempeñan un papel cada vez más importante en la facilitación de la gestión de riesgos, en particular, mediante la creación de ecosistemas de ensayo y evaluación (OECD, 2024^[139]). Por ejemplo, el DSIT del Reino Unido proporcionó orientación sobre la ampliación responsable de las capacidades⁸¹.

En relación con la abundancia de ejemplos de gobiernos que utilizan evaluaciones de impacto o auditorías algorítmicas, son pocos los ejemplos de gobiernos que han desarrollado o utilizado otras evaluaciones de riesgos de IA en su administración. Esto puede deberse a que la mayoría de los sistemas públicos de IA se adquieren a empresas del sector privado, que pueden llevar a cabo evaluaciones de riesgos antes de poner sus productos en el mercado. También puede deberse a que los sistemas de IA gubernamentales, tal como se analiza en el Capítulo 2, a menudo no aprovechan los últimos enfoques y, en cambio, se basan en sistemas basados en reglas o en enfoques de aprendizaje automático más establecidos. Aun así, los gobiernos deben considerar y utilizar estos enfoques de evaluación de riesgos en su intento de utilizar sistemas de IA más avanzados y capaces. Por ejemplo, además de las pruebas estructuradas, los gobiernos podrían llevar a cabo pruebas de confrontación —comúnmente conocidas como *red-teaming*— especialmente para modelos de base complejos utilizados o adquiridos por los gobiernos, con el fin de identificar proactivamente vulnerabilidades, riesgos de uso indebido y actividades o productos nocivos del sistema en aplicaciones gubernamentales sensibles. Las evaluaciones de la capacidad también podrían incluir evaluaciones específicas de cada idioma, en particular para los programas de aprendizaje permanente utilizados en los servicios públicos multilingües. Estas evaluaciones podrían ayudar a garantizar un desempeño justo, abordar los desequilibrios de datos en los modelos de base y ayudar a verificar que los modelos sean adecuados para los idiomas a los que sirven.

Empoderar a los órganos de supervisión y asesoramiento para orientar una IA responsable

Entidades de supervisión

El papel de los órganos de supervisión está evolucionando para hacer frente a los nuevos desafíos que plantea la expansión del uso de la IA en todo el gobierno. Por ejemplo, las Entidades Fiscalizadoras Superiores (EFS)⁸² están obligadas cada vez más a ampliar sus actividades de auditoría para incluir el escrutinio de los algoritmos de IA que sustentan las operaciones gubernamentales y las decisiones públicas. Este cambio requiere una evaluación exhaustiva de los algoritmos de IA no solo por su precisión, seguridad y efectividad, sino también por su transparencia y equidad. El Recuadro 4.8 muestra cómo las

EFS han adaptado su función a la realización de auditorías algorítmicas y han desarrollado los marcos para ello.

Las auditorías sirven para una serie de propósitos, entre los que se incluyen evaluar el rendimiento de los sistemas algorítmicos con respecto a los estándares establecidos, garantizar el cumplimiento de la regulación, detectar la discriminación ilegal, mejorar la transparencia y la explicabilidad, evaluar la seguridad y la solidez, evaluar los impactos sociales y éticos más amplios y responsabilizar a las organizaciones por sus sistemas. Además, las inspecciones pueden servir para detectar fallos sistémicos en el uso de sistemas algorítmicos, ofreciendo información valiosa que puede servir de base para su aplicación en otro contexto (Ada Lovelace Institute, 2021^[130]).

Recuadro 4.8. Enfoques y herramientas de auditoría algorítmica del sector público

Francia

En 2024, la Corte de Cuentas (*Cour des Comptes* en francés) francesa evaluó la integración de la IA en el Ministerio de Economía y Hacienda. Desde 2015, el ministerio ha puesto en marcha 35 programas de IA destinados a detectar riesgos de fraude individuales, identificar dificultades comerciales y ofrecer respuestas más rápidas a los usuarios. Si bien los aspectos tecnológicos están bien gestionados, el informe concluye que las consideraciones éticas, de recursos humanos y ambientales siguen sin explorarse. El Tribunal recomendó una supervisión ministerial sólida para garantizar una IA pública confiable, una mejor evaluación y una asignación transparente de los aumentos de productividad, y una anticipación proactiva del impacto de la IA en las funciones del personal

Países Bajos

El Tribunal de Cuentas neerlandés, órgano independiente, ha auditado el uso de algoritmos por parte del Gobierno Neerlandés. A través de su evaluación, el Tribunal desarrolló un marco de auditoría concebido específicamente para evaluar el uso de algoritmos en la administración. El marco evalúa una amplia gama de aspectos, desde la gobernanza y la rendición de cuentas hasta aspectos técnicos como los sistemas de IA y los datos, la privacidad, los controles generales de TI y las consideraciones éticas. El marco se está utilizando en varias instituciones holandesas para orientar su desarrollo de nuevos algoritmos. En 2022, auditó nueve algoritmos importantes del sector público y descubrió que seis (67%) no cumplían los requisitos básicos, lo que expuso al gobierno a sesgos, filtraciones de datos y acceso no autorizado.

Suecia

La Oficina Nacional de Auditoría de Suecia llevó a cabo una auditoría de tres sistemas automatizados de adopción de decisiones utilizados por el Gobierno Sueco: la prestación parental en la Agencia de la Seguridad Social Sueca, el impuesto anual sobre la renta de los particulares en la Agencia Tributaria Sueca y los permisos de aprendiz de permiso de conducción en la Agencia Sueca de Transporte. La fiscalización tenía por objeto evaluar si estos sistemas funcionaban de manera efectiva y eficiente, salvaguardando al mismo tiempo la seguridad jurídica en la toma de decisiones. Evaluó el rendimiento de los sistemas con respecto a las normas legislativas en materia de eficiencia y seguridad jurídica, identificando deficiencias específicas.

Reino Unido

En 2024, la Oficina Nacional de Auditoría (NAO, por su sigla en inglés) del Reino Unido publicó un Informe sobre la IA en el gobierno, en el que se examina la efectividad con la que los organismos gubernamentales del Reino Unido están utilizando la IA para los servicios públicos. Concluyó que solo el 21 % de los 87 organismos analizados contaban con una estrategia de IA, como exige la política,

mientras que el 61 % tenía previsto desarrollar una. Entre las iniciativas destacadas se incluyen el establecimiento por parte del Departamento de Trabajo y Pensiones de una junta directiva de IA y un grupo independiente de asesoría y garantía, y el establecimiento por parte del Ministerio de Justicia de un grupo directivo de IA para revisar casos de uso individuales de IA, junto con la adopción de paneles de consulta sobre algoritmos, que incluyan a usuarios finales y especialistas en ética de datos.

Estados Unidos

La Oficina de Responsabilidad del Gobierno de Estados Unidos (GAO, por sus siglas en inglés) publicó un marco de rendición de cuentas del uso de IA para agencias federales y otras entidades, en el que se identifican prácticas clave de rendición de cuentas —centradas en torno a los principios de gobernanza, datos, desempeño y supervisión— para ayudar a las agencias federales y otras entidades a utilizar la IA de manera responsable.

Colaboración transfronteriza

Las Entidades Fiscalizadoras Superiores (EFS) de Alemania, Finlandia, Noruega, los Países Bajos y el Reino Unido emiten colectivamente *algoritmos de auditoría de aprendizaje automático: Un libro blanco para los auditores públicos*, que se actualiza a lo largo del tiempo.

Fuente: <https://www.ccomptes.fr/fr/publications/intelligence-artificielle-dans-les-politiques-publiques-l-exemple-du-ministere-de>,
<https://english.rekenkamer.nl/publications/reports/2021/01/26/understanding-algorithms>,
https://www.riksrevisionen.se/download/18.2008b69c18bd0f6ed3f25040/1608291082190/RiR_2020_22_en-GB.pdf, (Ada Lovelace Institute, 2021^[130]), <https://www.gao.gov/products/gao-21-519sp>, <https://www.auditingalgorithms.net>, (OECD, 2023^[131]),
<https://www.nao.org.uk/reports/use-of-artificial-intelligence-in-government>.

Los órganos de supervisión sirven de plataformas que reúnen diversos conocimientos especializados y perspectivas, lo que es importante para la efectividad de cualquier mecanismo de rendición de cuentas (Ada Lovelace Institute, 2021^[130]). Por ejemplo:

- En 2023, España creó la Agencia Española de Supervisión de la IA (AESIA), una entidad independiente (Pehlivan and Valín, 2023^[146]).
- En 2024, la Ley de IA de la UE estableció un Consejo Europeo Supervisor de Inteligencia Artificial (IA)⁸³.
- Los/as defensores/as del pueblo desempeñan un papel fundamental de supervisión del uso de la IA por parte de las organizaciones públicas. Por ejemplo, el Defensor del Pueblo Europeo ha investigado el uso de la IA por parte de la CE,⁸⁴ y el Defensor del Pueblo y la Autoridad de Protección de Datos de los Países Bajos han examinado activamente la toma de decisiones algorítmica y el impacto de la IA en los derechos de la ciudadanía en los Países Bajos⁸⁵.
- Los parlamentos y sus comités de supervisión están adoptando un papel más activo en diferentes países, como Australia y el Reino Unido⁸⁶, donde han realizado investigaciones sobre la IA y la transparencia algorítmica. Algunos órganos parlamentarios establecen comités ad hoc para investigar cuestiones específicas relacionadas con la IA, en particular cuando surgen riesgos o controversias emergentes.

Órganos consultivos

Los órganos consultivos también pueden ayudar a garantizar que los gobiernos estén utilizando la IA de forma confiable. Pueden ofrecer orientación, conocimientos y recomendaciones de expertos en respuesta a solicitudes específicas de gobiernos sobre cuestiones emergentes en materia de IA. Algunos pueden ser más prácticos, como el Consejo Asesor de IA de Irlanda, que, por ejemplo, desarrolla y ofrece su propio plan de trabajo de asesoría⁸⁷. Otros ejemplos son:

- El Grupo Consultivo sobre Ética de los Datos de Nueva Zelanda ofrece orientación no vinculante sobre el uso de sistemas algorítmicos por los organismos públicos. Sus recomendaciones abordan cuestiones como el cumplimiento de los derechos humanos, la validez científica, la privacidad y la ética (Ada Lovelace Institute, 2021^[130]) ⁸⁸.
- Grecia creó en noviembre de 2023 un Comité Consultivo de Alto Nivel sobre IA, que desempeña un papel fundamental en la configuración de la política nacional en materia de IA. Se centra en promover el crecimiento económico y social y, al mismo tiempo, abordar los riesgos asociados al uso irrestricto de la IA. Ha desarrollado "Un Plan General para la Transformación de la IA en Grecia" (2024^[47]). El país también ha creado un Comité Nacional de Tecnoética y Bioética para proporcionar conocimientos especializados independientes que proporcionen orientación estratégica y recomendaciones sobre las implicancias éticas de la IA, entre otras cosas ⁸⁹.
- España (2024^[147]) ha creado el Consejo Consultivo de Inteligencia Artificial como organismo independiente formal para proporcionar al gobierno análisis, asesoría y apoyo sobre el tema de la IA. Celebró su primera reunión en junio de 2024.
- En 2025, el gobierno de Australia Occidental creó una Junta Consultiva sobre IA para asesorar a los organismos gubernamentales de Australia Occidental en materia de mitigación de riesgos y para apoyar el uso seguro, responsable y ético de la IA en el sector público de Australia Occidental⁹⁰.

Institutos y unidades de seguridad y protección de IA

Los gobiernos de todo el mundo también se están centrando en este tema a través de esfuerzos que incluyen el establecimiento de institutos y unidades de seguridad y protección de la IA en varios países (OECD, 2024^[139]). Por ejemplo, Canadá, la UE, Francia, Japón, Corea, Singapur, el Reino Unido y Estados Unidos han puesto en marcha cada uno de estos institutos o unidades⁹¹, y estos tres países adicionales han decidido formar una red internacional de institutos (UK DSIT, 2024^[148]). El mandato de dichos institutos o unidades suele ser más amplio que el uso de la IA en el gobierno, pero algunos sí se centran en esto. Por ejemplo, el inminente Instituto de Seguridad de la IA de la India servirá como grupo de reflexión y acción para la innovación en gobernanza y ofrecerá orientación política, jurídica y técnica a las instituciones públicas que desplieguen la IA, entre otros objetivos⁹².

La participación para dar forma a una IA estratégica y responsable

Los sistemas de IA tienen el potencial de remodelar radicalmente la interacción entre la ciudadanía y sus gobiernos, y entre los propios ciudadanos. Las principales partes interesadas, incluidos la ciudadanía, deben poder opinar sobre la forma en que los gobiernos utilizan y gobiernan las tecnologías basadas en la IA. La participación de la ciudadanía y las partes interesadas puede generar una mayor confianza y legitimidad en los sistemas de IA utilizados por los gobiernos, así como en sistemas de IA que reflejen mejor las necesidades de todos (OECD, 2024^[149]). Estos esfuerzos pueden ayudar a promover la transparencia, la rendición de cuentas y la equidad en los sistemas de IA, evitando sesgos y posibles daños.

La participación de la ciudadanía y las partes interesadas —como científicos e ingenieros, comunidades afectadas, inversores, empresas o instituciones— puede enriquecer la comprensión de las cuestiones relacionadas con la tecnología de IA, ayudar a los responsables de la formulación de políticas a anticipar problemas de aceptación pública y promover una buena comunicación (OECD, 2024^[149]). Esta participación es crucial para alinear el uso y la gobernanza de la IA con los objetivos y las necesidades de la sociedad. Ayuda a las partes interesadas a comprender, cuestionar e influir en cómo se diseñan y funcionan los algoritmos y los mecanismos de gobernanza de la IA.

La participación temprana permite una evaluación exhaustiva de las posibles consecuencias y riesgos para diversos grupos, fomentando un desarrollo más inclusivo y ético de los sistemas de IA y la gobernanza. Este enfoque colaborativo ayuda a identificar y abordar preocupaciones desde diversas perspectivas, garantizando que los sistemas de IA se diseñen e implementen de forma responsable y beneficiosa para todas las partes interesadas (OECD, 2021^[113]). Los ciudadanos y las partes interesadas pueden participar en diferentes momentos de la formulación de políticas públicas:

- **Establecimiento de la agenda.** En el Reino Unido, el Centro de Ética de los Datos e Innovación ha estado involucrando a la ciudadanía a través de la *encuesta Public Attitudes to Data and AI tracker* (Postura Pública ante el Seguimiento de Datos e IA), que ahora se encuentra en su cuarta ola y que puede informar el enfoque del gobierno para el desarrollo de políticas futuras⁹³.
- **Diseño de tecnología.** En 2020, el Gobierno Francés puso en marcha PIAF, una iniciativa de colaboración con ciudadanos, académicos y la sociedad civil para crear bases de datos en francés con el fin de entrenar sistemas de IA⁹⁴. La fábrica griega de IA Pharos se está configurando para ser un centro de colaboración, intercambio de conocimientos, puesta en común de recursos y desarrollo de proyectos conjuntos entre el sector público, las instituciones académicas y la industria privada.
- **Evaluación tecnológica.** En los Estados Unidos, la Evaluación Participativa de Ciencia y Tecnología por parte de Expertos y Ciudadanos (ECAST, por su sigla en inglés) está aportando perspectivas públicas para influir en las decisiones críticas del gobierno en materia de ciencia y tecnología⁹⁵.
- **Regulación.** En 2024, la Unión Europea inició una consulta a múltiples partes interesadas sobre modelos de IA de uso general fiables en virtud de la Ley⁹⁶ de IA, así como una convocatoria de manifestaciones de interés para participar en la elaboración del primer código de buenas prácticas de IA de uso general⁹⁷.

Los gobiernos disponen de varias opciones para potenciar la participación pública a la hora de diseñar el desarrollo estratégico y operativo de la IA. Las secciones siguientes exploran procesos deliberativos, como **asambleas ciudadanas, participación con los servidores públicos y participación de los usuarios en el desarrollo de IA**. Además de la participación ciudadana en el uso y la gobernanza de la IA por parte del gobierno, los gobiernos también pueden utilizar la IA para fomentar la participación ciudadana, algo que se analiza en profundidad en el Capítulo 5 (sección sobre "La IA en la participación cívica y el gobierno abierto").

Asambleas ciudadanas

Las asambleas ciudadanas, también denominadas jurados o paneles de ciudadanos, generalmente se refieren a un grupo seleccionado al azar de personas que son ampliamente representativas de una comunidad que pasa mucho tiempo aprendiendo y colaborando a través de una deliberación facilitada para formar recomendaciones colectivas para los responsables de la formulación de políticas (OECD, 2020^[150]).

Un proceso deliberativo representativo es el más adecuado para abordar cuestiones como los dilemas basados en valores; problemas complejos que requieren concesiones y afectan a una variedad de grupos de diferentes maneras; o cuestiones a largo plazo que van más allá de los ciclos electorales (OECD, 2022^[151]). El desarrollo y la gobernanza de la IA son muy adecuados para este proceso. La IA implica debates éticos y sociales para decidir sobre sus usos en contextos específicos (por ejemplo, el reconocimiento facial) y puede considerarse una cuestión técnica con concesiones entre innovación y regulación. Lo que es más importante, la adopción de las tecnologías de IA sin duda dará forma a las interacciones sociales a largo plazo, con impactos que pueden extenderse a través de generaciones.

Como ejemplo de asamblea ciudadana sobre IA, en 2024, la Presidencia Belga del Consejo de la Unión Europea convocó a un grupo representativo de 60 belgas para recabar las opiniones de la ciudadanía sobre la IA con el bloque (BeEU, 2024^[152]). En otro ejemplo, en Estados Unidos, en 2023, un profesor de la Universidad de Siracusa se asoció con el Centro para Nuevos Procesos Democráticos (CNDP, por su sigla en inglés) para llevar a cabo el primer evento deliberativo nacional sobre IA en Estados Unidos (Atwood and Bozentko, 2023^[153]). También existen ejemplos internacionales y subnacionales de reuniones para dar forma a la gobernanza de la IA. En 2025 y 2026, la Coalición Global para la IA Inclusiva, una asociación entre el Laboratorio de Democracia Deliberativa de Stanford y la firma de consultoría en participación ciudadana Missions Publiques, llevará a cabo asambleas deliberativas que pretenden llegar a más de 10 000 ciudadanos en más de 100 países. Se han previsto actividades de seguimiento con los responsables de la toma de decisiones para garantizar el impacto de los procesos deliberativos (Vergne and Siu, 2025^[154]). La participación ciudadana también es un componente importante de otra tendencia emergente en la gobernanza de la IA: el localismo de la IA, en el que las comunidades, incluidos los gobiernos locales, actúan para debatir y regular el uso de las tecnologías de IA en función de sus necesidades (Marcucci, Kalkar and Verhulst, 2022^[155]). Por ejemplo, en 2018, el laboratorio de innovación del Distrito Federal, Laboratorio para la Ciudad de México (LabCDMX), llevó a cabo un ejercicio deliberativo para elaborar un Marco de Gobernanza Anticipatoria de la Ciudad de México sobre IA (Ramos, 2018^[156]).

Participación de funcionarios públicos e interlocutores sociales

La participación de los servidores públicos, que están en la primera línea de la prestación de servicios públicos, es fundamental para el uso de la IA en el gobierno. Sus funciones y responsabilidades se ven directamente afectadas por la introducción de las tecnologías de IA, y sus puntos de vista y experiencias son invaluables para dar forma a un uso responsable y eficaz de la IA. Como se comenta en el Capítulo 1, la automatización, creación o transformación de tareas debido a la introducción de la IA puede generar oportunidades para mejorar la eficiencia y la efectividad. Pero esos cambios también pueden generar preocupación sobre la seguridad laboral, las condiciones de trabajo y los derechos de los trabajadores para estas personas.

El diseño y la implementación de iniciativas de IA deberían llevarse a cabo de una manera que respete los derechos laborales, promueva el bienestar y utilice los conocimientos de los funcionarios públicos (OECD, 2023^[119]). Para lograrlo, es importante un diálogo transparente e inclusivo con los servidores públicos y los interlocutores sociales, como los sindicatos y las asociaciones de empleados. Se debe informar a los trabajadores sobre los objetivos de las iniciativas de IA, los posibles impactos en sus funciones y las medidas adoptadas para mitigar los impactos negativos. También deberían tener la oportunidad de expresar sus preocupaciones, ofrecer su opinión y contribuir con sus conocimientos a los enfoques de IA.

El diálogo social y la negociación colectiva son esenciales para que la adopción de la IA en el gobierno tenga éxito (OECD, 2023^[119]). Son clave para generar confianza y colaboración efectiva entre los servidores públicos y para garantizar el acceso a la formación con el fin de desarrollar las habilidades y capacidades necesarias para trabajar con IA. Los interlocutores sociales también deberían participar, ya que desempeñan un papel fundamental en la negociación de las condiciones de trabajo.

Integración de usuarios en el desarrollo de la IA

Integrar a los usuarios finales en el desarrollo de la IA en el gobierno ayuda a garantizar que las soluciones de IA estén centradas en el usuario y aborden de manera efectiva los problemas del mundo real. El Gráfico 4.5 muestra los pasos clave para entender a los usuarios y sus necesidades. Los gobiernos pueden utilizar métodos de investigación —como la revisión de pruebas existentes, la realización de entrevistas y la observación de los usuarios— para desarrollar un conocimiento profundo de estos

aspectos, mejorando así la relevancia y aceptación de las aplicaciones de la IA (OECD/UNESCO, 2024^[29]). En consonancia con los Principios de Buenas Prácticas de la OCDE para el Diseño y la Prestación de Servicios en la Era Digital, los usuarios pueden ayudar a identificar información útil para iterar el diseño de servicios, simplificar los procedimientos subyacentes y aumentar el acceso de todos los grupos de usuarios (2022^[157]). Por otra parte, convertir el diseño y la prestación de servicios basados en la IA en un proceso participativo e inclusivo empodera a los usuarios, dándoles un papel activo en la creación conjunta y el diseño conjunto de servicios públicos. Esto puede incluir la implementación de mecanismos para involucrar a los usuarios en las pruebas, la iteración y la mejora del servicio, así como la realización de experimentos rigurosos y diseñados éticamente con grupos de usuarios para ayudar a garantizar que el uso de la IA tenga los efectos deseados y que cualquier riesgo se identifique a pequeña escala antes de ampliarlo (véase más arriba el apartado "Creación de espacios para experimentar").

Gráfico 4.5. Pasos clave para entender a los usuarios y sus necesidades en los desarrollos de IA gubernamental



Fuente: (OECD/UNESCO, 2024^[29]),

Colaboración transfronteriza

Al igual que otras tecnologías digitales, la IA no conoce fronteras. Sus riesgos e impactos, así como sus posibles usos positivos, pueden ser transnacionales⁹⁸. El participación y la colaboración transfronterizos pueden ser fundamentales para reducir las brechas de conocimiento y desarrollo entre países, abordar desafíos comunes y complejos, gestionar los riesgos y aplicar políticas y servicios innovadores. La cooperación internacional puede ayudar a crear capacidad gubernamental en materia de IA en todo el mundo y en regiones específicas, como se ha visto en ALC (OECD/CAF, 2022^[30]). Esto puede incluir el intercambio de algoritmos abiertos, infraestructuras y conjuntos de datos intergubernamentales, así como la realización de esfuerzos conjuntos para el desarrollo responsable de tecnologías emergentes. Un informe de la OCDE (2021^[158]) ha identificado tres mecanismos que los gobiernos están utilizando para conectarse y colaborar con el fin de abordar cuestiones que cruzan fronteras entre entidades administrativas o áreas, incluso en áreas relacionadas con la innovación digital y la IA en el gobierno.

Los organismos de gobernanza transfronteriza pueden abordar cuestiones complejas o que abarcan el ámbito de competencias de múltiples jurisdicciones, como la integración y armonización de los enfoques de IA y la interoperabilidad. Los organismos de gobernanza permiten a los gobiernos coordinar y aprovechar los esfuerzos colectivos de agentes divididos por fronteras. Así lo demuestra la creación por parte de la Unión Europea de la Oficina de IA encargada de aplicar, supervisar y hacer cumplir la Ley de IA. Fuera del ámbito de la UE, el Grupo de trabajo de altos funcionarios de gobierno digital de la OCDE (E-Leaders) y su grupo temático sobre IA han sido plataformas para que los países colaboren en el desarrollo de orientaciones y productos analíticos sobre la IA en el gobierno, con la posibilidad de proponer instrumentos jurídicos no vinculantes pertinentes de la OCDE. Además, los países han estado colaborando en la elaboración de normas internacionales aplicadas por organismos internacionales, como la ISO/IEC 42001 sobre sistemas de gestión de la IA, pertinentes para organismos del sector público y empresas u organizaciones sin fines de lucro. Sin embargo, en la actualidad no hay pruebas de que se

hayan creado organismos de gobernanza transfronteriza formales para abordar de forma específica el desarrollo público y el uso de la IA.

En segundo lugar, los países también están utilizando **redes innovadoras** para abordar la colaboración transfronteriza. Las redes son estructuras horizontales, a menudo informales y de base que permiten la convergencia orgánica de ideas y conocimientos a través de las fronteras. Por ejemplo, la Red Europea de Administraciones Públicas (EUPAN, por su sigla en inglés) promueve el intercambio de conocimientos sobre IA⁹⁹.

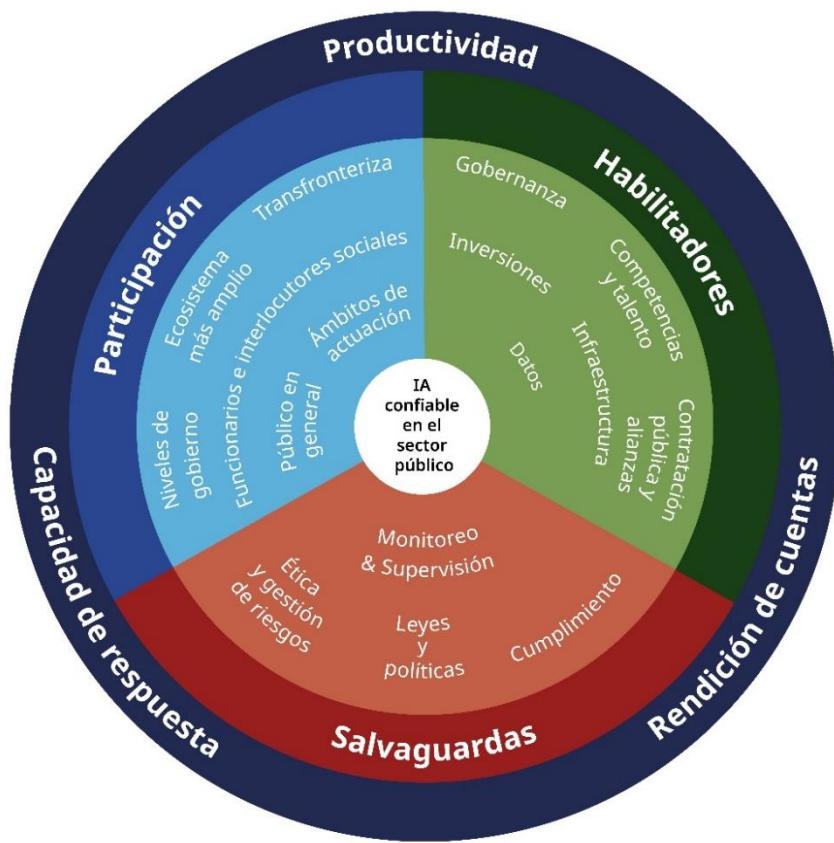
Y, en tercer lugar, algunos países están explorando **dinámicas emergentes de sistemas de gobernanza**, que son formas totalmente nuevas de trabajar juntos a través de las fronteras. Por ejemplo, los gobiernos han trabajado juntos para desarrollar infraestructuras digitales colectivas y enfoques de intercambio de datos con el fin de promover operaciones fluidas a nivel internacional. La Unión Europea es quizás la más avanzada en este ámbito, ya que su Ley Europea de Interoperabilidad, que entró en vigor en abril de 2024, establece un marco para mejorar la interoperabilidad dentro de las organizaciones del sector público, garantizando unos servicios transfronterizos sin fisuras. Los elementos clave incluyen la creación de una estructura de gobernanza de interoperabilidad, el fomento de la innovación y el intercambio de conocimientos, la implementación de entorno de prueba regulatorios para probar soluciones y la obligatoriedad de realizar evaluaciones de interoperabilidad para las administraciones públicas¹⁰⁰.

Un marco de trabajo para una IA confiable en el sector público

La adopción conjunta de las medidas de política pública analizadas en este capítulo constituye un *marco para una IA confiable en el sector público*, que facilita a las administraciones públicas la alineación de sus acciones relacionadas con el desarrollo y la implementación de la IA conforme a los Principios de IA de la OCDE (2024^[159]). El marco describe cómo los gobiernos pueden aprovechar la promesa de la IA en materia de productividad, capacidad de respuesta y rendición de cuentas mediante la combinación adecuada de habilitadores, salvaguardas y participación.

El Gráfico 4.6 presenta una representación visual del marco y la Tabla 4.1 detalla las preguntas y medidas de política en las que se basan sus elementos.

Gráfico 4.6. Marco de la OCDE para una inteligencia artificial confiable en el sector público



Fuente: (OECD, 2024^[3]), mejorado y finalizado a través del trabajo realizado para este informe.

Tabla 4.1. Preguntas y medidas de política pública que sustentan el marco para una IA confiable en el sector público

Pregunta de Política Pública	Medida de Política Pública	Descripción
¿Qué acciones y herramientas concretas pueden desarrollar los gobiernos para hacer frente a los desafíos existentes en aras de un uso confiable de la IA en la administración pública?	Habilitadores	Políticas e instrumentos que aborden las actuales limitaciones detectadas por los decisores políticos, para crear un entorno favorable y facilitar la adopción plena de la IA en el sector público. Entre ellos se incluyen la gobernanza, la infraestructura, los datos, las competencias y el talento, la gestión de las inversiones, la contratación pública y las alianzas.
	Salvaguardas	Herramientas de política pública que las administraciones pueden considerar desarrollar para un uso de la IA responsable, confiable y centrado en las personas. Estos pueden incluir instrumentos no vinculantes, leyes y reglamentos, instrumentos de transparencia y gestión de riesgos, o la supervisión (más allá del poder ejecutivo) y el monitoreo (dentro de los órganos ejecutivos).
¿A quién deberían involucrar los gobiernos a la hora de desarrollar e implementar los habilitadores y salvaguardas para el uso confiable de la IA en la administración pública?	Participación	Diferentes partes interesadas involucradas en la construcción de las bases para un uso responsable de la IA en el gobierno. Es necesario involucrar a diversos actores de todo el gobierno (por ejemplo, ministerios, servidores públicos o gobiernos subnacionales), del ecosistema en general y de fuera de las jurisdicciones nacionales, mediante acciones específicas para abordar de manera efectiva las oportunidades y los desafíos relacionados con el uso de la IA en el sector público.
¿Qué impacto debería lograr el gobierno al utilizar una IA confiable?	Impacto	La IA en el sector público puede ayudar a aumentar la productividad, la capacidad de respuesta y la rendición de cuentas.

Fuente: (OECD, 2024^[3]).

Futuro trabajo de la OCDE sobre estas problemáticas

Los habilitadores, las salvaguardas y los procesos de interacción que componen el *Marco de la OCDE para una IA confiable en el sector público* sirven como base sólida sobre la que los gobiernos pueden adoptar un enfoque estratégico y responsable con respecto a la IA. Sin embargo, un informe no puede reflejar plenamente la complejidad de este amplio espectro de actividades necesarias para adoptar una tecnología de IA en rápida evolución y, al mismo tiempo, gestionar tanto los riesgos críticos como los importantes desafíos de implementación. El trabajo futuro de la OCDE abordará elementos del marco más en profundidad, con ideas viables sobre cómo los gobiernos pueden poner en marcha estas bases. Por ejemplo, un informe sobre la experimentación con IA en el gobierno ya está en marcha y se publicará en los próximos meses.

Es fundamental que los gobiernos puedan identificar dónde priorizar las inversiones y los recursos en IA en función de diversas concesiones a la hora de considerar los posibles beneficios y riesgos de determinadas aplicaciones de IA. La OCDE ha recomendado (2024^[3]) y sigue alentando a los gobiernos a que den prioridad a las aplicaciones de IA de alto beneficio y bajo riesgo, especialmente cuando se está estableciendo un nivel inicial de madurez. Sin embargo, la mayoría no cuenta con procesos para la medición holística de los resultados potenciales o logrados —eficiencia del gasto, calidad de los servicios, daños potenciales— que les permitan hacer estas determinaciones. Esta debería ser una prioridad para los gobiernos como paso transversal que ayude a desbloquear el potencial de la IA, y será un tema central del trabajo futuro de la OCDE.

Bibliografía

- Ada Lovelace Institute (2025), *Learn fast and build things: Lessons from six years of studying AI in the public sector*, Ada Lovelace Institute, <https://www.adalovelaceinstitute.org/policy-briefing/public-sector-ai/>. [23]
- Ada Lovelace Institute (2021), *Algorithmic accountability for the public sector*, Ada Lovelace Institute (Ada), AI Now Institute (AI Now), and Open Government Partnerships (OGP), <https://www.adalovelaceinstitute.org/report/algorithmic-accountability-public-sector/>. [130]
- African Union (2024), *Continental Artificial Intelligence Strategy*, <https://au.int/en/documents/20240809/continental-artificial-intelligence-strategy>. [65]
- Agency for Digital Government (2024), *A Common Danish Language Resource*, <https://en.digst.dk/digital-governance/new-technologies/a-common-danish-language-resource/> (accessed on November 2024). [46]
- AI Hub (2024), “AI Data Finder”, *aihub.or.kr*, <https://www.aihub.or.kr/aihubdata/data/list.do?currMenu=115&topMenu=100> (accessed on November 2024). [53]
- AI Sweden (2024), *A shared digital assistant for the public sector*, <https://www.ai.se/en/project/shared-digital-assistant-public-sector> (accessed on October 2024). [51]
- Atwood, S. and K. Bozentko (2023), *U.S. Public Assembly on High Risk Artificial Intelligence 2023 Event Report*, <https://www.cndp.us/wp-content/uploads/2023/12/2023-U.S.-PUBLIC-ASSEMBLY-ON-HIGH-RISK-AI-EVENT-REPORT-final.pdf>. [153]

- Australia DTA (2024), *APS trials generative AI to explore safe and responsible use cases for government.* [102]
- Australian Government (2024), *Evaluation of the whole-of-government trial of Microsoft 365 Copilot*, <https://www.digital.gov.au/initiatives/copilot-trial>. [36]
- BeEU (2024), *A citizen's view of artificial intelligence within the EU*, <https://belgian-presidency.consilium.europa.eu/media/lzxauu4i/rapport-ia-en-web-v2.pdf>. [152]
- Berryhill, J. et al. (2019), "Hello, World: Artificial intelligence and its use in the public sector", *OECD Working Papers on Public Governance*, No. 36, OECD Publishing, Paris, <https://doi.org/10.1787/726fd39d-en>. [4]
- Brauneis, R. and E. Goodman (2017), "Algorithmic Transparency for the Smart City", *SSRN Electronic Journal*, <https://doi.org/10.2139/ssrn.3012499>. [134]
- Brizuela, A. et al. (2025), *Analysis of the generative AI landscape in the European public sector*, European Commission, <https://op.europa.eu/s/z4XY>. [68]
- Chilean Transparency Council (CPLT) (2024), *CPLT lanza recomendaciones de transparencia algorítmica en servicios públicos*, <https://www.consejotransparencia.cl/cplt-lanza-recomendaciones-de-transparencia-algoritmica-en-servicios-publicos/>. [127]
- CNIL (2023), « *Bac à sable* » données personnelles : la CNIL lance un appel à projets sur l'intelligence artificielle dans les services publics, <https://www.cnil.fr/fr/bac-sable-donnees-personnelles-la-cnil-lance-un-appel-projets-sur-lintelligence-artificielle-dans>. [41]
- CNIL (2023), « *Bac à sable* » intelligence artificielle et services publics : la CNIL accompagne 8 projets innovants, <https://www.cnil.fr/fr/bac-sable-intelligence-artificielle-et-services-publics-la-cnil-accompagne-8-projets-innovants>. [42]
- CONPES (2025), CONPES 4144, *Política Nacional de Inteligencia Artificial*, <https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/4144.pdf>. [31]
- Corrêa, N. et al. (2023), "Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance", *Patterns*, Vol. 4/10, p. 100857, <https://doi.org/10.1016/j.patter.2023.100857>. [115]
- Council of Europe (2024), *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>. [120]
- Cover-Kus, H. (2024), *UK Government doubles down efforts to deploy AI across the public sector*, <https://www.techuk.org/resource/uk-government-doubles-down-efforts-to-deploy-ai-across-the-public-sector.html>. [98]
- Deleware General Assembly (2024), *House Bill 333: An Act to Amend Title 29 of the Deleware Code Relating to the Artificial Intelligence Commission*, <https://legis.delaware.gov/BillDetail/140866>. [125]
- Digdir (2024), *The Norwegian Resource Centre for Sharing and Use of Data*, <https://www.digdir.no/datadeling/norwegian-resource-centre-sharing-and-use-data/2766> (accessed on November 2024). [60]

- Dombo, F. (2023), *Public Sector Considerations for Successful Cloud Adoption*, [73]
<https://news.sap.com/africa/2023/05/public-sector-considerations-for-successful-cloud-adoption/>.
- EC (2025), *A pioneering AI project awarded for opening Large Language Models to European languages*, [88]
<https://digital-strategy.ec.europa.eu/en/news/pioneering-ai-project-awarded-opening-large-language-models-european-languages> (accessed on 10 March 2025).
- EC (2025), *Common European Data Spaces*, European Commission, [55]
<https://digital-strategy.ec.europa.eu/en/policies/data-spaces>.
- EC (2025), *The AI Continent Action Plan*, European Commission, [67]
<https://digital-strategy.ec.europa.eu/en/library/ai-continent-action-plan>.
- EC (2025), *Updated EU AI model contractual clauses*, European Commission, [104]
<https://public-buyers-community.ec.europa.eu/communities/procurement-ai/resources/updated-eu-ai-model-contractual-clauses>.
- EC (2024), *Adopt AI study*, European Commission, [97]
<https://op.europa.eu/s/z2tg>.
- EC (2024), *GovTech: influencing factors, common requirements and recommendations*, [111]
European Commission, <https://interoperable-europe.ec.europa.eu/collection/public-sector-tech-watch/document/govtech-influencing-factors-common-requirements-and-recommendations>.
- EC (2024), *What factors influence perceived artificial intelligence adoption by public managers*, [8]
European Commission Joint Research Centre,
https://publications.jrc.ec.europa.eu/repository/bitstream/JRC138684/JRC138684_01.pdf.
- EC (2023), *EU model contractual AI clauses to pilot in procurements of AI*, [103]
<https://public-buyers-community.ec.europa.eu/communities/procurement-ai/resources/eu-model-contractual-ai-clauses-pilot-procurements-ai>.
- Fink, K. (2017), “Opening the government’s black boxes: freedom of information and algorithmic accountability”, *Information, Communication & Society*, Vol. 21/10, pp. 1453-1471, [133]
<https://doi.org/10.1080/1369118x.2017.1330418>.
- France Élysée (2025), *Make France and AI Powerhouse*, [64]
<https://www.elysee.fr/admin/upload/default/0001/17/d9c1462e7337d353f918aac7d654b896b77c5349.pdf>.
- Frazier, K. (2025), *The Dangers of AI Sovereignty*, [70]
<https://www.lawfaremedia.org/article/the-dangers-of-ai-sovereignty>.
- G7 (2023), *Hiroshima Process International Code of Conduct for Advanced AI Systems*, [138]
<https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-code-conduct-advanced-ai-systems>.
- G7 (2023), *Hiroshima Process International Guiding Principles for Advanced AI system*, [137]
<https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-guiding-principles-advanced-ai-system>.
- Gartner (2024), *Compare AI Software Spending in the Government Industry, 2023-2027*, [93]
<https://www.gartner.com/en/documents/5318363>.

- Gob.cl (2025), *Latam GPT: Learn about the Latin American AI model developed in Chile*, [87]
<https://www.gob.cl/en/news/latam-gpt-learn-about-the-latin-american-ai-model-developed-in-chile/> (accessed on 10 March 2025).
- Goodman, E. and J. Trehu (2022), *AI Audit-Washing and Accountability*, [144]
<https://www.gmfus.org/news/ai-audit-washing-and-accountability>.
- Government of Canada (2025), *AI Strategy for the Federal Public Service*, [13]
<https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/gc-ai-strategy-overview.html>.
- Government of Greece (2025), *The implementation of the new National Supercomputer "DAIDALOS" and the Data Center at the Lavrio TPP begins - The country acquires one of the most powerful computing systems in Europe*, [83]
<https://mindigital.gr/archives/7331>.
- Government of Greece (2024), *A Blueprint for Greece's AI Transformation*, [47]
https://foresight.gov.gr/wp-content/uploads/2024/11/Blueprint_GREECES_AI_TRANSFORMATION.pdf.
- Government of Ireland (2025), *Guidelines for the Responsible Use of AI in the Public Service*, [19]
<https://www.gov.ie/en/department-of-public-expenditure-ndp-delivery-and-reform/publications/guidelines-for-the-responsible-use-of-ai-in-the-public-service>.
- Government of Korea (2024), *A New Chapter in the Age of AI: Basic Act on AI Passed at the National Assembly's Plenary Session*, [123]
<https://www.msit.go.kr/eng/bbs/view.do?bbsSeqNo=42&mId=4&mPid=2&nttSeqNo=1071>.
- Government of New Zealand (2025), *Public Service AI Framework*, [18]
<https://www.digital.govt.nz/standards-and-guidance/technology-and-architecture/artificial-intelligence/public-service-artificial-intelligence-framework>.
- Government of Spain (2024), *Spain sets up the International Artificial Intelligence Advisory Council*, [147]
<https://www.lamoncloa.gob.es/lang/en/gobierno/news/Paginas/2024/20240621-ai-advisory-council-meeting.aspx>.
- Government of Spain (2024), *The Government approves the Artificial Intelligence Strategy 2024*, [110]
<https://portal.mineco.gob.es/en-us/comunicacion/Pages/20240514-Gobierno-aprueba-Estrategia-IA-2024.aspx>.
- Government of Switzerland (2025), *Strategy Use of AI systems in the Federal Administration*, [14]
<https://www.bk.admin.ch/bk/en/home/digitale-transformation-ikt-lenkung/ikt-vorgaben/strategien-teilstrategien/sb021-strategie-einsatz-von-ki-systemen-in-der-bundesverwaltung.html>.
- Government of the Dominican Republic (2024), *Estrategia Nacional de Inteligencia Artificial*, [12]
<https://innovacionrd.gob.do/enia/>.
- Government of Uruguay (2021), *AI Strategy for the Digital Government*, [15]
<https://www.gub.uy/agencia-gobierno-electronico-sociedad-informacion-conocimiento/comunicacion/publicaciones/ia-strategy-english-version/ia-strategy-english-version/ai-strategy-for>.

- GPAI (2024), *Algorithmic Transparency in the Public Sector: A state-of-the-art report of algorithmic transparency instruments*, Global Partnership on Artificial Intelligence, <https://gpai.ai/projects/responsible-ai/algorithmic-transparency-in-the-public-sector/algorithmic-transparency-in-the-public-sector.pdf>. [126]
- Gupta, A. et al. (2021), *The State of AI Ethics Report (Volume 5)*, <https://arxiv.org/abs/2108.03929>. [141]
- Hassani, A. et al. (2022), *Escaping the Big Data Paradigm with Compact Transformers*, <https://arxiv.org/abs/2104.05704>. [79]
- IEA (2025), *Energy and AI*, <https://iea.blob.core.windows.net/assets/34eac603-ecf1-464f-b813-2ecceb8f81c2/EnergyandAI.pdf>. [77]
- IEA (2023), *Data Centres and Data Transmission Networks*, <https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks>. [76]
- Jones, N. (2025), *Where AI Is Now: Smaller, Better, Cheaper Models*, <https://www.scientificamerican.com/article/ai-report-highlights-smaller-better-cheaper-models>. [80]
- Komaitis, K., E. Ponce de León and K. Thibaut (2024), *The sovereignty trap*, <https://www.atlanticcouncil.org/blogs/geotech-cues/the-sovereignty-trap>. [69]
- Lam, M. et al. (2023), “Sociotechnical Audits: Broadening the Algorithm Auditing Lens to Investigate Targeted Advertising”, *Proceedings of the ACM on Human-Computer Interaction*, Vol. 7/CSCW2, pp. 1-37, <https://doi.org/10.1145/3610209>. [143]
- Leslie, D. (2019), *Understanding artificial intelligence ethics and safety*, The Alan Turing Institute, <https://doi.org/10.5281/zenodo.3240529>. [117]
- Letzing, J. (2024), *What is ‘sovereign AI’ and why is the concept so appealing (and fraught)?*, <https://www.weforum.org/stories/2024/11/what-is-sovereign-ai-and-why-is-the-concept-so-appealing-and-fraught/>. [63]
- Long, D. and B. Magerko (2020), “What is AI Literacy? Competencies and Design Considerations”, *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1-16, <https://doi.org/10.1145/3313831.3376727>. [91]
- M Saiful Bari, Y. (2024), “ALLaM: Large Language Models for Arabic and English”. [48]
- Marcucci, S., U. Kalkar and S. Verhulst (2022), *AI Localism in Practice: Examining How Cities Govern AI*, <https://files.thegovlab.org/ailocalism-in-practice.pdf>. [155]
- McKinsey (2024), *AI power: Expanding data center capacity to meet growing demand*, <https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/ai-power-expanding-data-center-capacity-to-meet-growing-demand>. [75]
- Mehta, S., A. Rogers and T. Gilbert (2023), *Dynamic Documentation for AI Systems*, <https://arxiv.org/abs/2303.10854>. [142]
- METR (2023), *Responsible Scaling Policies (RSPs)*, <https://metr.org/blog/2023-09-26-rsp/>. [145]
- Metz, C. et al. (2025), *How A.I. is Changing the Way the World Builds Computers*, <https://www.nytimes.com/interactive/2025/03/16/technology/ai-data-centers.html>. [78]

- Monteiro, B., A. Hlacs and P. Boéchat (2024), "Public procurement for public sector innovation: Facilitating innovators' access to innovation procurement", *OECD Working Papers on Public Governance*, No. 80, OECD Publishing, Paris, <https://doi.org/10.1787/9aad76b7-en>. [101]
- Montgomery, C., F. Rossi and J. New (2023), *A Policymaker's Guide to Foundation Models*, <https://newsroom.ibm.com/Whitepaper-A-Policymakers-Guide-to-Foundation-Models>. [85]
- Mulgan, G. (2019), *Intelligence as an outcome not an input*, <http://www.nesta.org.uk/blog/intelligence-outcome-not-input>. [26]
- Netherlands Court of Audit (2024), *Focus on AI in central government*, <https://english.rekenkamer.nl/publications/reports/2024/10/16/focus-on-ai-in-central-government>. [128]
- NIST (2023), *AI Risk Management Framework*, <https://www.nist.gov/itl/ai-risk-management-framework>. [136]
- Noor, E. and B. Kanitroj (2025), *Speaking in Code: Contextualizing Large Language Models in Southeast Asia*, <https://carnegieendowment.org/research/2025/01/speaking-in-code-contextualizing-large-language-models-in-southeast-asia>. [89]
- OECD (2025), *Digital Government Review of Korea: Harnessing Digital and Data to Transform Government*, OECD Digital Government Studies, OECD Publishing, Paris, <https://doi.org/10.1787/9defc197-en>. [81]
- OECD (2025), "Effectively Managing Investments in Digital Government: An OECD Policy Framework", *OECD Public Governance Policy Papers*, No. 76, OECD Publishing, Paris, <https://doi.org/10.1787/5c324e91-en>. [94]
- OECD (2025), *Enhancing Access to and Sharing of Data in the Age of Artificial Intelligence*, OECD Publishing Paris, https://www.oecd.org/en/publications/enhancing-access-to-and-sharing-of-data-in-the-age-of-artificial-intelligence_23a70dca-en.html. [50]
- OECD (2025), "Mapping relevant data collection mechanisms for AI training", *OECD Artificial Intelligence Papers*, No. 48, OECD Publishing, Paris, <https://doi.org/10.1787/3264cd4c-en>. [38]
- OECD (2025), *OECD Regulatory Policy Outlook 2025*, OECD Publishing, Paris, <https://doi.org/10.1787/56b60e39-en>. [121]
- OECD (2024), "2023 OECD Digital Government Index: Results and key findings", *OECD Public Governance Policy Papers*, No. 44, OECD Publishing, Paris, <https://doi.org/10.1787/1a89ed5e-en>. [74]
- OECD (2024), "AI, data governance and privacy: Synergies and areas of international co-operation", *OECD Artificial Intelligence Papers*, No. 22, OECD Publishing, Paris, <https://doi.org/10.1787/2476b1a4-en>. [40]
- OECD (2024), "Assessing potential future artificial intelligence risks, benefits and policy imperatives", *OECD Artificial Intelligence Papers*, No. 27, OECD Publishing, Paris, <https://doi.org/10.1787/3f4e3dfb-en>. [139]
- OECD (2024), "Digital public infrastructure for digital governments", *OECD Public Governance Policy Papers*, No. 68, OECD Publishing, Paris, <https://doi.org/10.1787/ff525dc8-en>. [62]

- OECD (2024), *Enabling Digital Innovation in Government: The OECD GovTech Policy Framework*, OECD Digital Government Studies, OECD Publishing, Paris, [109]
<https://doi.org/10.1787/a51eb9b2-en>.
- OECD (2024), “Fixing frictions: ‘sludge audits’ around the world: How governments are using behavioural science to reduce psychological burdens in public services”, *OECD Public Governance Policy Papers*, No. 48, OECD Publishing, Paris, [2]
<https://doi.org/10.1787/5e9bb35c-en>.
- OECD (2024), *Framework for Anticipatory Governance of Emerging Technologies*, OECD Publishing, [149]
<https://doi.org/10.1787/0248ead5-en>.
- OECD (2024), *G20 Compendium on Data Access and Sharing Across the Public Sector and with the Private Sector for Public Interest*, OECD Publishing, Paris, [54]
https://www.oecd.org/en/publications/g20-compendium-on-data-access-and-sharing-across-the-public-sector-and-with-the-private-sector-for-public-interest_df1031a4-en.html.
- OECD (2024), “Governing with Artificial Intelligence: Are governments ready?”, *OECD Artificial Intelligence Papers*, No. 20, OECD Publishing, Paris, [3]
<https://doi.org/10.1787/26324bc2-en>.
- OECD (2024), *OECD Artificial Intelligence Review of Germany*, OECD Publishing, Paris, [95]
<https://doi.org/10.1787/609808d6-en>.
- OECD (2024), *OECD Digital Economy Outlook 2024 (Volume 1): Embracing the Technology Frontier*, OECD Publishing, Paris, [84]
<https://doi.org/10.1787/a1689dc5-en>.
- OECD (2024), *Public procurement*, <https://www.oecd.org/en/topics/public-procurement.html>. [107]
- OECD (2024), *Recommendation of the Council on Artificial Intelligence*, OECD Publishing, [159]
<https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.
- OECD (2024), *The Digital Transformation of Norway’s Public Sector*, OECD Digital Government Studies, OECD Publishing, Paris, [100]
<https://doi.org/10.1787/1620e542-en>.
- OECD (2023), “2023 OECD Open, Useful and Re-usable data (OURdata) Index: Results and key findings”, *OECD Public Governance Policy Papers*, No. 43, OECD Publishing, Paris, [52]
<https://doi.org/10.1787/a37f51c3-en>.
- OECD (2023), “Advancing accountability in AI: Governing and managing risks throughout the lifecycle for trustworthy AI”, *OECD Digital Economy Papers*, No. 349, OECD Publishing, Paris, [135]
<https://doi.org/10.1787/2448f04b-en>.
- OECD (2023), “AI language models: Technological, socio-economic and policy considerations”, *OECD Digital Economy Papers*, No. 352, OECD Publishing, Paris, [44]
<https://doi.org/10.1787/13d38f92-en>.
- OECD (2023), “Common guideposts to promote interoperability in AI risk management”, *OECD Artificial Intelligence Papers*, No. 5, OECD Publishing, Paris, [140]
<https://doi.org/10.1787/ba602d18-en>.
- OECD (2023), *Global Trends in Government Innovation 2023*, OECD Public Governance Reviews, OECD Publishing, Paris, [131]
<https://doi.org/10.1787/0655b570-en>.

- OECD (2023), *OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market*, OECD Publishing, Paris, <https://doi.org/10.1787/08785bba-en>. [119]
- OECD (2023), “Regulatory sandboxes in artificial intelligence”, *OECD Digital Economy Papers*, No. 356, OECD Publishing, Paris, <https://doi.org/10.1787/8f80a0e6-en>. [35]
- OECD (2022), *Going Digital to Advance Data Governance for Growth and Well-being*, OECD Publishing, Paris, <https://doi.org/10.1787/e3d783b0-en>. [56]
- OECD (2022), “Measuring the environmental impacts of artificial intelligence compute and applications: The AI footprint”, *OECD Digital Economy Papers*, No. 341, OECD Publishing, Paris, <https://doi.org/10.1787/7babf571-en>. [71]
- OECD (2022), “OECD Framework for the Classification of AI systems”, *OECD Digital Economy Papers*, No. 323, OECD Publishing, Paris, <https://doi.org/10.1787/cb6d9eca-en>. [58]
- OECD (2022), “OECD Good Practice Principles for Public Service Design and Delivery in the Digital Age”, *OECD Public Governance Policy Papers*, No. 23, OECD Publishing, Paris, <https://doi.org/10.1787/2ade500b-en>. [157]
- OECD (2022), *OECD Guidelines for Citizen Participation Processes*, OECD Publishing, <https://doi.org/10.1787/f765caf6-en>. [151]
- OECD (2021), “Achieving cross-border government innovation: Governing cross-border challenges”, *OECD Public Governance Policy Papers*, No. 10, OECD Publishing, Paris, <https://doi.org/10.1787/ddd07e3b-en>. [158]
- OECD (2021), *G20 survey on Agile approaches to the regulatory governance of innovation: Report for the G20 Digital Economy Task Force, Trieste, Italy, August 2021*, OECD Publishing, Paris, <https://doi.org/10.1787/f161916d-en>. [113]
- OECD (2021), *Good Practice Principles for Data Ethics in the Public Sector - OECD*, OECD Publishing, Paris, <https://www.oecd.org/gov/digital-government/good-practice-principles-for-data-ethics-in-the-public-sector.htm> (accessed on 14 April 2025). [43]
- OECD (2021), *OECD Report on Public Communication: The Global Context and the Way Forward*, OECD Publishing, <https://doi.org/10.1787/22f8031c-en>. [21]
- OECD (2021), *Public Sector Innovation Facets: Mission-oriented innovation*, OECD Publishing, <https://oecd-opsi.org/publications/facets-mission/>. [9]
- OECD (2021), *Recommendation of the Council for Agile Regulatory Governance to Harness Innovation*, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0464>. [122]
- OECD (2021), “The OECD Framework for digital talent and skills in the public sector”, *OECD Working Papers on Public Governance*, No. 45, OECD Publishing, Paris, <https://doi.org/10.1787/4e7c3f58-en>. [7]
- OECD (2020), *Innovative Citizen Participation and New Democratic Institutions: Catching the Deliberative Wave*, OECD Publishing, Paris, <https://doi.org/10.1787/339306da-en>. [150]
- OECD (2019), *Artificial Intelligence in Society*, OECD Publishing, Paris, <https://doi.org/10.1787/eedfee77-en>. [34]

- OECD (2019), "Data governance in the public sector", in *The Path to Becoming a Data-Driven Public Sector*, OECD Publishing, Paris, <https://doi.org/10.1787/9cada708-en>. [59]
- OECD (2019), *Recommendation of the Council on Public Service Leadership and Capability*, OECD Publishing, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0445>. [161]
- OECD (2019), *The Path to Becoming a Data-Driven Public Sector*, OECD Digital Government Studies, OECD Publishing, Paris, <https://doi.org/10.1787/059814a7-en>. [57]
- OECD (2017), *Systems Approaches to Public Sector Challenges: Working with Change*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264279865-en>. [1]
- OECD (forthcoming), *Generative AI Experimentation in Government: A review of current practices*, OECD Publishing. [37]
- OECD.AI (2025), *The OECD-African Union AI Dialogue 2.0: From strategy to implementation*, <https://oecd.ai/en/wonk/the-oecd-african-union-ai-dialogue-2-0-from-strategy-to-implementation>. [116]
- OECD/CAF (2022), *The Strategic and Responsible Use of Artificial Intelligence in the Public Sector of Latin America and the Caribbean*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/1f334543-en>. [30]
- OECD/UNESCO (2024), *G7 Toolkit for Artificial Intelligence in the Public Sector*, OECD Publishing, Paris, <https://doi.org/10.1787/421c1244-en>. [29]
- Pahlka, J. (2024), *AI meets the cascade of rigidity*, <https://www.niskanencenter.org/ai-meets-the-cascade-of-rigidity/>. [20]
- Parankusham, K., R. Rizk and K. Santosh (2025), *LakotaBERT: A Transformer-based Model for Low Resource Lakota Language*, <https://arxiv.org/abs/2503.18212>. [49]
- Pehlivan, C. and E. Valín (2023), *Spain establishes the EU's first AI supervisory agency*, <https://techinsights.linklaters.com/post/102intj/spain-establishes-the-eus-first-ai-supervisory-agency>. [146]
- Peixoto, T., O. Canuto and L. Jordan (2024), *AI and the Future of Government: Unexpected Effects and Critical Challenges*, <https://www.policycenter.ma/publications/ai-and-future-government-unexpected-effects-and-critical-challenges>. [45]
- Personal Information Protection Commission of Korea (2023), *Policy direction for safe use of personal information in the era of artificial intelligence [translated]*, <https://www.pipc.go.kr/np/cop/bbs/selectBoardArticle.do?bbsId=BS074&mCode=C020010000&nttId=9083>. [39]
- Policy Lab Digital, Work & Society within the German Federal Ministry of Labour and Social Affairs (2024), *Guidelines for the Use of AI in the Administrative Work of Employment and Social Protection Services*, https://www.denkfabrik-bmas.de/fileadmin/Downloads/Publikationen/Guidelines_for_the_use_of_ai_in_the_administrative_work_of_employment_and_social_protection_services.pdf. [118]
- Ramos, J. (2018), *Laboratorio Para La Ciudad (CDMX)*, <https://actionforesight.net/laboratorio-para-la-ciudad-cdmx/>. [156]

- Ray, T. (2025), *Sovereign remedies: Between AI autonomy and control*, [66]
<https://www.atlanticcouncil.org/in-depth-research-reports/issue-brief/sovereign-remedies-between-ai-autonomy-and-control/>.
- Redapt (2023), *On-Premises vs. Cloud for AI Workloads*, [72]
<https://www.redapt.com/blog/on-premises-vs-cloud-for-ai-workloads>.
- Rudra, S. (2024), *OSI Calls Out Meta for its Misleading 'Open Source' AI Models*, [160]
<https://news.itsfoss.com/osi-meta-ai/>.
- Ryseff, J., B. De Bruhl and S. Newberry (2024), *The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed*, RAND, [24]
https://www.rand.org/pubs/research_reports/RRA2680-1.html.
- Ryseff, J. and A. Narayanan (2025), *Why AI Projects Fail*, [25]
<https://www.rand.org/pubs/presentations/PTA2680-1.html>.
- SDAIA (2025), *Government Cloud (Deem)*, [82]
<https://sdaia.gov.sa/en/Services/Pages/Deem.aspx>.
- Seger, E. et al. (2024), *Open-Sourcing Highly Capable Foundation Models: An Evaluation of Risks, Benefits, and Alternative Methods for Pursuing Open-Source Objectives*, [86]
<https://www.governance.ai/research-paper/open-sourcing-highly-capable-foundation-models>.
- The Alan Turing Institute (2023), *AI Skills for Business Competency Framework*, [92]
<https://www.turing.ac.uk/skills/collaborate/ai-skills-business-framework>.
- U.S. General Services Administration (2024), *AI Capability Maturity - Operational maturity areas*, [61]
<https://coe.gsa.gov/coe/ai-guide-for-government/operational-maturity-areas/index.html#dataops>.
- UCL IIPP (2019), *A Mission-Oriented UK Industrial Strategy*, [10]
https://www.ucl.ac.uk/bartlett/public-purpose/sites/public-purpose/files/190515_iipp_report_moiis_final_artwork_digital_export.pdf.
- UK Committee of Public Accounts (2025), *Use of AI in Government*, [129]
<https://committees.parliament.uk/publications/47199/documents/244683/default/>.
- UK DSIT (2024), *Global leaders agree to launch first international network of AI Safety Institutes to boost cooperation of AI*, [148]
<https://www.gov.uk/government/news/global-leaders-agree-to-launch-first-international-network-of-ai-safety-institutes-to-boost-understanding-of-ai>.
- UK DSIT (2020), *Guidelines for AI procurement*, [27]
<https://www.gov.uk/government/publications/guidelines-for-ai-procurement/guidelines-for-ai-procurement>.
- UK Government (2025), *A blueprint for modern digital government*, [6]
<https://www.gov.uk/government/publications/a-blueprint-for-modern-digital-government/a-blueprint-for-modern-digital-government-html>.
- UK Government (2025), *Prime Minister: I will reshape the state to deliver security for working people*, [5]
<https://www.gov.uk/government/news/prime-minister-i-will-reshape-the-state-to-deliver-security-for-working-people>.

- UK Government Digital Service (2025), *Artificial Intelligence Playbook for the UK Government*, [22] <https://www.gov.uk/government/publications/ai-playbook-for-the-uk-government/artificial-intelligence-playbook-for-the-uk-government-html>.
- UK House of Commons (2024), *Governance of Artificial Intelligence (AI): Government Response*, [99] <https://committees.parliament.uk/publications/46145/documents/230927/default/>.
- UK NAO (2024), *Use of artificial intelligence in government*, [16] <https://www.nao.org.uk/wp-content/uploads/2024/03/use-of-artificial-intelligence-in-government.pdf>.
- UNESCO (2024), *Consultation paper on AI regulation: emerging approaches across the world*, [112] <https://unesdoc.unesco.org/ark:/48223/pf0000390979>.
- UNESCO (2023), *Ethical impact assessment. A tool of the Recommendation on the Ethics of Artificial Intelligence*, UNESCO, [114] <https://doi.org/10.54678/ytsa7796>.
- University of Rome Sapienza (2024), *AI made in Italy: here is Minerva, the first family of large language models trained “from scratch” for Italian*, [90] <https://www.uniroma1.it/en/notizia/ai-made-italy-here-minerva-first-family-large-language-models-trained-scratch-italian> (accessed on 10 March 2025).
- US IT Modernization Centers of Excellence (n.d.), *AI Guide for Government*, [28] <https://coe.gsa.gov/ai-guide-for-government>.
- US OMB (2025), *Accelerating Federal Use of AI through Innovation, Governance, and Public Trust*, [17] <https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-21-Accelerating-Federal-Use-of-AI-through-Innovation-Governance-and-Public-Trust.pdf>.
- US OMB (2025), *M-25-22 Driving Efficient Acquisition of Artificial Intelligence in Government*, [106] White House Office of Management and Budget, <https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-22-Driving-Efficient-Acquisition-of-Artificial-Intelligence-in-Government.pdf>.
- Valderrama, M., M. Hermosilla and R. Garrido (2023), *State of the Evidence: Algorithmic Transparency*, [132] <https://www.opengovpartnership.org/wp-content/uploads/2023/05/State-of-the-Evidence-Algorithmic-Transparency.pdf> (accessed on August 2024).
- van Noordt, C., R. Medaglia and L. Tangi (2023), “Policy initiatives for Artificial Intelligence-enabled government: An analysis of national strategies in Europe”, *Public Policy and Administration*, [96] <https://doi.org/10.1177/09520767231198411>.
- Vergne, A. and A. Siu (2025), *Global Coalition for an Inclusive AI*, [154] <https://global-ai-dialogue.org/>.
- Verhulst, S. and M. Sloane (2020), *Realizing the Potential of AI Localism*, [33] <https://www.project-syndicate.org/commentary/local-regulation-of-artificial-intelligence-uses-by-stefaan-g-verhulst-1-and-mona-sloane-2020-02?barrier=accesspaylog>.
- Vinnova (2022), *Designing missions*, [11] <https://www.vinnova.se/contentassets/1c94a5c2f72c41cb9e651827f29edc14/designing-missions.pdf?cb=20220311094952>.
- WAM (2024), *Hamdan bin Mohammed appoints 22 Chief AI Officers across government entities in Dubai*, [32] <https://www.wam.ae/en/article/b3kujwp-hamdan-bin-mohammed-appoints-chief-officers-across>.

- Werner, J. (2024), *New York Governor Signs AI Oversight Bill*, <https://babl.ai/new-york-governor-signs-ai-oversight-bill>. [124]
- World Bank (2025), *Global Trends in AI Governance: Evolving Country Approaches*, <https://openknowledge.worldbank.org/entities/publication/a570d81a-0b48-4cac-a3d9-73dff48a8f1a>. [108]
- World Economic Forum (2025), *AI Procurement Guideline*, <https://www.weforum.org/publications/ai-procurement-in-a-box/ai-government-procurement-guidelines/> (accessed on 10 March 2025). [105]

Notas

¹ Véase <https://oecd-opsi.org/work-areas/systems-approaches>.

² Véase <https://oecd-opsi.org/work-areas/anticipatory-innovation> y <https://www.oecd.org/en/about/programmes/strategic-foresight>.

³ El contexto y el uso de los "habilitadores" en este informe no son los mismos que los "habilitadores de IA" para sistemas de IA generalmente discutidos por las políticas de IA y las comunidades técnicas, que consisten en datos, algoritmos y poder computacional ("compute").

⁴ Véase <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023>, <https://www.gov.uk/government/publications/seoul-declaration-for-safe-innovative-and-inclusive-ai-ai-seoul-summit-2024>, y <https://www.elysee.fr/en/emmanuel-macron/2025/02/11/statement-on-inclusive-and-sustainable-artificial-intelligence-for-people-and-the-planet>, respectivamente.

⁵ Véase (OECD, 2021^[7]) para obtener material adicional relevante, incluidas habilidades y competencias para el liderazgo del gobierno digital. Véase también OCDE (2019^[161]) *Recomendación sobre el liderazgo y la capacidad de la función pública* para obtener información sobre cómo los países pueden inculcar una cultura y un liderazgo basados en valores y garantizar servidores públicos cualificados y eficaces, así como sistemas de empleo público receptivos y adaptables.

⁶ La OCDE apoya las innovaciones orientadas a misiones a través de su Laboratorio de Acción Misional, una iniciativa conjunta de la Dirección de Ciencia, Tecnología e Innovación (STI) de la OCDE, el Observatorio de Innovación del Sector Público (OPSI) de la Dirección de Gobernanza Pública (GOV) y la Dirección de Cooperación para el Desarrollo (DCD). Véase <https://oecd-missions.org>.

⁷ La ciencia del comportamiento es un enfoque interdisciplinario que abarca el estudio del comportamiento humano y el diseño de estrategias para cambiarlo. Véase <https://www.oecd.org/en/topics/behavioural-science>.

⁸ Véase <https://www.apsc.gov.au/initiatives-and-programs/workforce-information/research-analysis-and-publications/state-service/state-service-report-2023-24/fit-future/supporting-safe-and-responsible-use-artificial-intelligence>.

⁹ Véase <https://www.digmin.dk/digitalisering/mere-om-digitalisering/digital-taskforce-for-kunstig-intelligens->.

¹⁰ Véase <https://www.mitre.org/news-insights/fact-sheet/federal-ai-sandbox> y <https://www.mitre.org/news-insights/news-release/mitre-establish-new-ai-experimentation-and-prototyping-capability-us>.

¹¹ <https://ai.gov.uk>.

¹² <https://alliance.numerique.gouv.fr>.

¹³ Véase <https://govtech.justica.gov.pt/en/govtech-justica-english> y <https://www.ceei.es/legal&justicialechlab/?r=vpx6w6j70qr1jstvc6>, respectivamente.

¹⁴ <https://www.gov.uk/government/publications/the-magenta-book/guidance-on-the-impact-evaluation-of-ai-interventions-html>.

¹⁵ <https://www.gov.uk/government/publications/the-magenta-book/guidance-on-the-impact-evaluation-of-ai-interventions-html>.

¹⁶ Véase <https://www.gov.uk/guidance/repository-of-privacy-enhancing-technologies-pets-use-cases> para consultar un repositorio de casos de uso de diferentes países recopilados por el Reino Unido.

¹⁷ Otros países incluyen Brasil, Canadá, China, Egipto, Estonia, Finlandia, Alemania, Francia, Hungría, Islandia, India, Israel, Japón, Corea, Letonia, Noruega, Qatar, Singapur, Eslovenia, España, Sudáfrica, Tailandia, Turquía, el Reino Unido y Vietnam.

¹⁸ Véase también <https://www.ekt.gr/en/news/30774>.

¹⁹ Véase <https://www.athenarc.gr/el/news/meltemi-proto-anoihto-megalo-glossiko-montelo-gia-ta-ellinika> y <https://www.athenarc.gr/en/news/llama-krikri-new-greek-ai-language-model-featured-kathimerini>, respectivamente.

²⁰ Véase <https://static.pib.gov.in/WriteReadData/specificdocs/documents/2022/aug/doc202282696201.pdf> y <https://www.indiatoday.in/technology/news/story/bhashini-ceo-amitabh-naq-talks-about-how-their-ai-tool-is-bridging-indias-language-divide-2646101-2024-12-06>; información complementada por funcionarios del Gobierno de la India.

²¹ Véase <https://interoperable-europe.ec.europa.eu/collection/open-source-observatory-osor/news/spanish-authorities-release-alia-ai-models>.

²² Véase <https://www.gov.uk/guidance/national-data-strategy> y <https://www.dta.gov.au/digital-government-strategy>, respectivamente.

²³ Véase <https://www.gov.uk/guidance/national-data-strategy> y <https://www.dta.gov.au/digital-government-strategy>.

[government-strategy](#), respectivamente.

²⁴ Véase <https://www.gov.br/governodigital/pt-br/capacitacao/capacita-gov-br>.

²⁵ Véanse <https://sbsc.gov.ie/information-access/data-sharing-and-governance-act>, <https://eur-lex.europa.eu/eli/reg/2016/679/oj>, <https://eur-lex.europa.eu/eli/reg/2023/2854>, <https://eur-lex.europa.eu/eli/reg/2022/868/oj>, <https://eur-lex.europa.eu/eli/dir/2019/1024/oj>, y https://ec.europa.eu/isa2/eif_en, respectivamente.

²⁶ Las infraestructuras públicas digitales (IPD) se definen como sistemas digitales compartidos que son seguros e interoperables y que pueden apoyar la prestación y el acceso inclusivos a los servicios públicos y privados en toda la sociedad. Estos sistemas actúan como elementos básicos digitales comunes que sustentan los procesos y servicios públicos y permiten la transformación del gobierno digital a escala social (OECD, 2024^[62]).

²⁷ <https://www.gov.br/governodigital/pt-br/infraestrutura-nacional-de-dados>.

²⁸ Las ponderaciones del modelo son "las variables o valores numéricos utilizados para especificar cómo se transforma la entrada (por ejemplo, el texto que describe una imagen) en la salida (por ejemplo, la imagen misma). Estos se actualizan de forma iterativa durante la formación del modelo para mejorar su rendimiento en las tareas para las que ha sido formado" (Seger et al., 2024^[86]). El uso de modelos de "código abierto" para este informe no implica que dichos modelos se lancen bajo una licencia de código abierto aprobada por la Open Source Initiative (OSI), una organización sin fines de lucro administradora de la definición de código abierto (<https://opensource.org/osd>). OSI ha criticado a algunas compañías que llaman a sus modelos código abierto porque solo proporcionan los ponderadores para el modelo, y no otros elementos, como los datos de entrenamiento, código y prácticas de entrenamiento (Rudra, 2024^[160]). Algunos argumentan que estos modelos deberían llamarse "peso abierto" en lugar de "código abierto".

²⁹ <https://bigscience.huggingface.co/blog/bloom>.

³⁰ <https://www.tech.gov.sg/products-and-services/for-government-agencies/productivity-and-marketing/vica>.

³¹ Véase <https://oecd.ai/catalogue>.

³² <https://indiaai.s3.ap-south-1.amazonaws.com/docs/empowering-public-sector-leadership.pdf>.

Información complementada por funcionarios del Gobierno de la India.

³³ Véase <https://www.ipa.ie/ipa-overview/onelearning.2548.html> y <https://www.ypes.gr/ypourgeio-esoterikon-google-enarxi-epimorfotikis-drasis-gia-dimosious-ypallilous-me-thema-tin-techniti-noimosyni>, respectivamente.

³⁴ Véase <https://www.elementsofai.com>.

³⁵ Véase <https://www.canada.ca/en/government/system/digital-government/digital-talent-strategy.html>.

³⁶ Véase <https://chcoc.gov/content/skills-based-hiring-guidance-and-competency-model-artificial-intelligence-work>.

³⁷ Véase <https://innovadorespublicos.cl> y <https://innovadorespublicos.cl/events/652/>.

³⁸ https://catalogue.csps-efpc.gc.ca/product?catalog=DDN3-E45&cm_locale=en.

³⁹ <https://www.smart-city-dialog.de/en/about-us/international-smart-cities-network>.

⁴⁰ <https://alliance.numerique.gouv.fr>.

⁴¹ Véase: <https://cnai.swiss/en/products/community-of-practice/>

⁴² Véase <https://www.cio.bund.de/Webs/CIO/DE/digitale-loesungen/datenpolitik/daten-und-ki/daten-und-ki-node.html>. y <https://www.digitale-verwaltung.de/SharedDocs/downloads/Webs/DV/DE/Transformation/akteurssteckbrief-beki.pdf>.

⁴³ Véase <https://www.digital.gov.au/policy/ai/pilot-ai-assurance-framework>,
<https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-principles/australias-ai-ethics-principles> y <https://www.digital.gov.au/policy/ai>, respectivamente.

⁴⁴ Véase <https://www.dta.gov.au/advice/digital-and-ict-investments/digital-and-ict-investment-oversight-framework-iof>.

⁴⁵ Véase <https://www.gsa.gov/about-us/newsroom/news-releases/technology-modernization-fund-seeking-proposals-fo-02082024>.

⁴⁶ <https://www.numerique.gouv.fr/services/guichet-financement-exploitation-valorisation-des-donnees/>

⁴⁷ Véase <https://www.numerique.gouv.fr/offre-accompagnement/pilotage-panorama/>.

⁴⁸ <https://public-buyers-community.ec.europa.eu/communities/procurement-ai/resources/eu-model-contractual-ai-clauses-pilot-procurements-ai>

⁴⁹ https://www.buyict.gov.au/sys_attachment.do?sys_id=e535e2ca935caa10438b39cdfaba103d.

⁵⁰ Véase también <https://www.chilecompra.cl/2024/11/goblab-uai-presento-nueva-herramienta-para-una-ia-responsable-y-etica>.

⁵¹ https://ec.europa.eu/commission/presscorner/detail/en/ip_25_467.

⁵² Véase <https://www.elysee.fr/en/emmanuel-macron/2025/02/11/statement-on-inclusive-and-sustainable-artificial-intelligence-for-people-and-the-planet> y <https://www.elysee.fr/en/sommet-pour-l-action-sur-l-ia/public-interest-ai>.

⁵³ <https://www.dataobservatory.net>.

⁵⁴ Véase <https://www.rtu.lv/en/university/for-mass-media/news/open/latvia-establishes-artificial-intelligence-centre> y <https://digital-skills-jobs.europa.eu/en/latest/news/latvia-establishes-artificial-intelligence-centre>.

⁵⁵ Véanse <https://oecd.ai/ai-principles>, <https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence>, <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>, y

<https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-guiding-principles-advanced-ai-system>, respectivamente.

⁵⁶ <https://docs.un.org/en/A/78/L.49>.

⁵⁷ Véase <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023>, <https://www.gov.uk/government/publications/seoul-declaration-for-safe-innovative-and-inclusive-ai-ai-seoul-summit-2024> y <https://www.elysee.fr/en/emmanuel-macron/2025/02/11/statement-on-inclusive-and-sustainable-artificial-intelligence-for-people-and-the-planet>, respectivamente.

⁵⁸ <https://www.industry.gov.au/publications/voluntary-ai-safety-standard>.

⁵⁹ Véase también <https://oecd.ai/en/wonk/how-the-oecd-ai-policy-observatory-has-shaped-colombia-and-latin-americas-approach-to-ai-policy>.

⁶⁰ <https://aicm.ai.gov.eg/en/Resources/EgyptianCharterForResponsibleAIEnglish-v1.0.pdf>.

⁶¹ <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/guide-use-generative-ai.html>.

⁶² https://www.cio.bund.de/Webs/CIO/DE/digitale-loesungen/kuenstliche_intelligenz/kuenstliche_intelligenz-node.html.

⁶³ <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper>.

⁶⁴ Véase <https://hyscaler.com/insights/bahrain-pioneers-ai-regulation> y <https://www.ita.gov.om/itaportal/Data/SiteImgGallery/2024731125545486/National%20Artificial%20Intelligence%20Policy.pdf>, respectivamente.

⁶⁵ <https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-21-Accelerating-Federal-Use-of-AI-through-Innovation-Governance-and-Public-Trust.pdf>.

⁶⁶ <https://research-data.urosario.edu.co/file.xhtml?persistentId=doi:10.34848/YN1CRT/8OHRT0&version=1.0>.

⁶⁷ <https://www.gov.uk/algorithmic-transparency-records>.

⁶⁸ Véase <https://github.com/ombegov/2024-Federal-AI-Use-Case-Inventory> para consultar una consolidación centralizada de los inventarios de casos de uso de IA de todos los organismos del gobierno federal de los Estados Unidos.

⁶⁹ Véase <https://www.algoritmospublicos.cl/repositorio>, <https://odap.fr/inventaire>, y <https://algoritmes.overheid.nl>, respectivamente.

⁷⁰ Véase <https://algoritmeregister.amsterdam.nl/en/ai-register> y <https://ai.hel.fi/en/ai-register>, respectivamente.

⁷¹ <https://stip.oecd.org/stip/interactive-dashboards/policy-initiatives/2023%2Fdata%2FpolicyInitiatives%2F2329>.

⁷² <https://www.gov.uk/government/collections/algorithmic-transparency-recording-standard-hub>.

⁷³ <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592>.

⁷⁴ Para facilitar su análisis, los resultados de la calificación del derecho mundial a la información (RTI) evalúan el marco jurídico formal del derecho a la información en relación con cada país en función de una serie de categorías (por ejemplo, ámbito de aplicación, procedimiento de solicitud, procedimiento de recurso). Véase <https://www.rti-rating.org/country-data>.

⁷⁵ Véase también el trabajo del Grupo de expertos en IA de la OCDE (<https://oecd.ai/site/risk-accountability>).

⁷⁶ Véase <https://www.coe.int/en/web/portal/-/huderia-new-tool-to-assess-the-impact-of-ai-systems-on-human-rights>.

⁷⁷ Véase, por ejemplo, la cobertura del tema en (OECD/CAF, 2022^[30]).

⁷⁸ Véase <https://www.nist.gov/news-events/news/2024/05/nist-launches-aria-new-program-advance-sociotechnical-testing-and>.

⁷⁹ <https://www.forgov.qld.gov.au/information-and-communication-technology/qgea-directions-and-guidance/qgea-policies-standards-and-guidelines/foundational-artificial-intelligence-risk-assessment-guideline>.

⁸⁰ La OCDE está explorando en mayor medida el concepto de umbrales de riesgo para la IA, como demostró una consulta pública sobre el tema celebrada en septiembre de 2024 (<https://oecd.ai/wonk/seeking-your-views-public-consultation-on-risk-thresholds-for-advanced-ai-systems-deadline-10-september>).

⁸¹ Véase <https://www.gov.uk/government/publications/emerging-processes-for-frontier-ai-safety/emerging-processes-for-frontier-ai-safety#responsible-capability-scaling>.

⁸² Las Entidades Fiscalizadoras Superiores (EFS) son organismos públicos encargados de auditar los ingresos y los gastos públicos. Mediante el control de la gestión financiera pública y la presentación de informes, garantizan que los recursos se utilizan según lo prescrito. Véase <https://sirc.idi.no/about/what-are-sais> para más información.

⁸³ <https://digital-strategy.ec.europa.eu/en/policies/ai-board>.

⁸⁴ <https://www.ombudsman.europa.eu/en/doc/closing-note/en/196934>.

⁸⁵ <https://www.autoriteitpersoonsgegevens.nl/en/themes/algorithms-ai>.

⁸⁶ Véase https://www.aph.gov.au/Parliamentary_Business/Committees/Joint/Public_Accounts_and_Audit/Publicse

[ctoruseofAI/Report](#) y <https://committees.parliament.uk/work/6986/governance-of-artificial-intelligence-ai>, respectivamente.

⁸⁷ <https://www.gov.ie/en/publication/a3be3-membership-of-the-ai-advisory-council>.

⁸⁸ <https://data.govt.nz/leadership/advisory-governance/data-ethics-advisory-group>.

⁸⁹ <https://bioethics.gr>.

⁹⁰ <https://www.wa.gov.au/organisation/department-of-the-premier-and-cabinet/office-of-digital-government/western-australian-artificial-intelligence-advisory-board>.

⁹¹ Véanse <https://ised-isde.canada.ca/site/ised/en/canadian-artificial-intelligence-safety-institute>, <https://digital-strategy.ec.europa.eu/en/policies/ai-office>, <https://www.economie.gouv.fr/actualites/la-france-se-dote-d-un-institut-national-pour-levaluation-et-la-securite-de-l'intelligence>, <https://aisi.go.jp>, <https://www.aisi.re.kr>, <https://t.ly/vCtd1>, <https://www.gov.uk/government/publications/ai-safety-institute-overview>, y <https://www.nist.gov/aisi>, respectivamente.

⁹² <https://indiaai.gov.in/article/india-takes-the-lead-establishing-the-indiaai-safety-institute-for-responsible-ai-innovation>. Información complementada por funcionarios del Gobierno de la India.

⁹³ <https://www.gov.uk/government/publications/public-attitudes-to-data-and-ai-tracker-survey-wave-4/public-attitudes-to-data-and-ai-tracker-survey-wave-4-report>.

⁹⁴ <https://www.etalab.gouv.fr/ia-decouvrez-et-participez-au-projet-piaf-pour-des-ia-francophones>.

⁹⁵ <https://ecastnetwork.org> y <https://issues.org/thinking-like-citizen-participatory-technology-assessment-weller-govani-farooque>.

⁹⁶ <https://digital-strategy.ec.europa.eu/en/consultations/ai-act-have-your-say-trustworthy-general-purpose-ai>.

⁹⁷ <https://digital-strategy.ec.europa.eu/en/news/ai-act-participate-drawing-first-general-purpose-ai-code-practice>.

⁹⁸ La OCDE ha publicado una serie de informes sobre el tema "Lograr la innovación gubernamental transfronteriza", que abordan desafíos y casos de éxito en diversos ámbitos, incluida la IA. Véase <https://cross-border.oecd-opsi.org>.

⁹⁹ Véase, por ejemplo, https://www.eupan.eu/wp-content/uploads/2024/12/HU_eNews-on-Examples-of-recently-launched-digital-public-service-innovations.pdf.

¹⁰⁰ <https://www.consilium.europa.eu/en/press/press-releases/2024/03/04/interoperable-europe-act-council-adopts-new-law-for-more-efficient-digital-public-services-across-the-eu>.

5

Inmersión profunda: La situación actual y el futuro potencial de la IA en el sector público

Este capítulo ofrece un análisis en profundidad de la situación actual, el potencial sin explotar, la gestión de riesgos y desafíos y el camino a seguir para las funciones centrales del sector público. Consiste en funciones de política gubernamental (diseño y ejecución regulatoria, administración tributaria y gestión financiera pública), procesos gubernamentales clave (reforma de la función pública, contratación pública, lucha contra la corrupción y promoción de la integridad pública, evaluación de políticas, y participación cívica y gobierno abierto) y servicios gubernamentales y funciones de justicia (diseño y ejecución de servicios públicos, aplicación de la ley y gestión del riesgo de desastres, y administración de justicia y acceso a la justicia). Para ello, ofrece muchos ejemplos reales de casos prácticos de IA que las administraciones públicas de todo el mundo están empleando en aras de la productividad, la capacidad de respuesta y la rendición de cuentas. El contenido de este capítulo aportó mucha información a las conclusiones del Capítulo 2, que analiza las tendencias de la IA en el gobierno.

Las funciones de política pública

Las funciones de política pública, que incluyen la administración tributaria, gestión de finanzas públicas y el diseño e implementación de la regulación, son esenciales para una gobernanza efectiva. Perfilan los marcos normativos y generan ingresos públicos, además de ayudar a garantizar un uso transparente y responsable de los recursos, sentando las bases para unos resultados políticos estables y equitativos. La IA se está utilizando en estas funciones para fortalecer aún más la columna vertebral de unas administraciones públicas efectivas y eficientes.

IA en la administración tributaria

Durante muchos años, las administraciones tributarias han utilizado la IA para apoyar actividades en todo su modelo operativo y han estado explorando activamente su potencial para mejorar aún más sus operaciones y los servicios al contribuyente, aumentar el cumplimiento normativo de las obligaciones tributarias y prevenir el fraude fiscal. Están en condiciones de hacerlo; la recopilación y el análisis de datos son el eje de muchos procesos de la administración tributaria, lo que ha facilitado la adopción temprana de sistemas de IA basados en reglas. Estos sistemas han permitido a las administraciones tributarias analizar y extraer información de grandes volúmenes de datos, lo que ha agilizado la detección de incumplimientos y ha permitido destinar con mayor precisión los limitados recursos disponibles a los casos de alto riesgo.

Esta experiencia existente sustenta las investigaciones sobre cómo pueden desplegarse técnicas más avanzadas en un contexto de administración tributaria. Si bien los avances tecnológicos en el ámbito de la IA ofrecen un gran potencial para mejorar y transformar los servicios, también pueden amplificar los riesgos relacionados con el procesamiento de datos confidenciales e introducir nuevas preocupaciones. Por lo tanto, las administraciones tributarias deben hacer especial hincapié en la privacidad, la seguridad y la aplicación confiable de la IA.

Situación actual

Desde 2022, la encuesta del *Inventario de Iniciativas de Tecnológica Tributaria* (ITTI, por sus siglas en inglés) de la OCDE ha ofrecido una visión de cómo se está aplicando la IA en las administraciones tributarias de todo el mundo. La última encuesta muestra que las principales áreas de aplicación en los países de la OCDE son la detección de la evasión y el fraude fiscal, la asistencia en la toma de decisiones y la mejora de los servicios fiscales (Gráfico 5.1).

Gráfico 5.1. Despliegues de IA en los miembros de la OCDE que utilizan IA en la administración tributaria



Nota: 29 de los 38 miembros de la OCDE declaran utilizar la IA en la administración tributaria en el *Inventario de Iniciativas de Tecnológica Tributaria 2024*. Fuente: OECD Data Explorer - Inventory of Tax Technology Initiatives 2024 (<https://oe.cd/dx/ITTI2024>).

Mejoramiento del cumplimiento normativo y la detección de evasión y fraude

Para mejorar el cumplimiento normativo y detectar la evasión y el fraude, la IA se utiliza con frecuencia para descubrir patrones de comportamiento ocultos o nuevas conexiones entre transacciones, activos o contribuyentes dentro de las fuentes de datos que una administración tributaria ya podría tener. A medida que han aumentado las capacidades tecnológicas y los conocimientos internos, se están utilizando técnicas impulsadas por la IA para analizar conjuntos de datos no estructurados, como el texto de documentos escritos a mano y la información de publicaciones públicas en redes sociales, con el fin de profundizar el análisis y descubrir conexiones sobre actividades que podrían indicar evasión fiscal o incumplimiento. Estas técnicas también se utilizan para ayudar a detectar o prevenir ataques fraudulentos contra el sistema fiscal por parte de delincuentes que pueden utilizar la IA para crear solicitudes falsas de devolución a una escala potencialmente muy grande. Por ejemplo, la Autoridad Independiente de Ingresos Públicos (IAPR, por sus siglas en inglés) de Grecia está utilizando la IA para combatir la evasión fiscal detectando problemas de cumplimiento normativo, automatizando procedimientos complejos y permitiendo a los auditores responder en tiempo real¹.

Entre los usos más recientes cabe mencionar la aplicación de la capacidad de detección de patrones de los sistemas de IA de las administraciones tributarias a los datos de los sistemas de captación de imágenes para que las administraciones tributarias puedan analizar mapas e imágenes satelitales. Al comparar las imágenes satelitales a lo largo del tiempo, se está utilizando la IA, por ejemplo, para identificar alteraciones inmobiliarias o nuevos edificios o activos imponibles que pueden no haber sido declarados a la autoridad tributaria (Recuadro 5.1). Grecia también ha desarrollado un modelo de IA para geolocalizar piscinas que no han sido declaradas a efectos fiscales².

Recuadro 5.1. Detección de desarrollo no declarado de propiedades en Francia

Para mejorar el proceso de detección de construcciones o construcciones no declaradas, la Administración Tributaria francesa (DGFiP) utiliza inteligencia artificial y la mejora de datos a partir de fotografías aéreas tomadas por el Instituto Nacional de Información Geográfica y Forestal (*Institut national de l'information géographique et forestière*, IGN) en el marco del proyecto Tierra Innovadora (*Foncier Innovant*).

Los algoritmos permiten extraer los contornos de los edificios y las piscinas de las imágenes aéreas públicas del IGN, que pueden consultarse en el sitio web geoportail.gouv.fr.

A continuación, un proceso informático verifica, a partir de las declaraciones de los propietarios a la DGFiP, si los elementos detectados en las imágenes están correctamente gravados (en particular, en el caso de los impuestos sobre bienes inmuebles). A continuación, un funcionario de la DGFiP verifica sistemáticamente cada anomalía detectada antes de tomar ninguna medida para recordar o, en última instancia, gravar al propietario del inmueble.

Fuente: (OECD, 2022^[1]).

Asistencia en los procesos de toma de decisiones administrativas

Para que los procesos administrativos sean más eficientes, las administraciones tributarias están desplegando IA para automatizar sus procesos internos, por ejemplo, ayudando a clasificar y distribuir casos o tareas similares. Del mismo modo, la IA se utiliza habitualmente para diferenciar casos sencillos y estandarizados, que pueden resolverse mediante la automatización, de casos complejos, como los contenciosos fiscales, en los que los funcionarios de la administración tributaria son más necesarios tanto

por su experiencia como por su capacidad para emitir juicios. Esta aplicación permite un uso más eficiente de los recursos humanos y tecnológicos (Recuadro 5.2).

Recuadro 5.2. Mejoramiento en la toma de decisiones en Brasil

Brasil tenía alrededor de 140 000 millones USD en impuestos liquidados a la espera de resoluciones en recursos de los tribunales de lo contencioso-administrativo. La sentencia de apelación tarda unos seis años. En el marco del Proyecto de resolución de litigios en materia de IA, Brasil utilizó el ML supervisado (ML) para distribuir grupos de archivos similares entre los mismos funcionarios, con el fin de agilizar la administración del archivo y la adopción de decisiones. Los primeros ensayos, realizados con una muestra de 2 000 archivos etiquetados manualmente, mostraron que los algoritmos supervisados pueden alcanzar una sensibilidad y especificidad de más del 80 %.

Brasil también empleó algoritmos de agrupamiento para completar los archivos total o parcialmente. Además, se está desarrollando una herramienta de asistencia para la elaboración de informes basada en la web para apoyar el análisis de los funcionarios y ayudarles en su objetivo de reutilizar bloques de texto. Los recursos de la herramienta incluyen la presentación de grupos de archivos y párrafos sugeridos y el realce de oraciones que resultaron ser importantes para el proceso de agrupamiento. Los funcionarios pueden etiquetar archivos y párrafos, y las etiquetas se utilizan para mejorar sugerencias futuras.

Fuente: (OECD, 2022^[1]).

Mejoramiento de los procesos de evaluación de riesgos

Uno de los principales motivos por los que las administraciones tributarias adoptan la IA es su capacidad para analizar datos con el fin de puntuar y priorizar riesgos. Con la IA, las administraciones tributarias pueden procesar grandes volúmenes de datos de los contribuyentes —desde declaraciones históricas hasta registros de transacciones e información de pagos digitales— para elaborar sofisticados modelos de riesgo y, a continuación, asignar puntuaciones de riesgo a tipos específicos de comportamiento o transacciones. Estos modelos ayudan a las administraciones tributarias a identificar rápidamente los posibles incumplimientos y a centrar los recursos en ellos. Ahora que la capacidad predictiva de estos sistemas ha aumentado, las administraciones tributarias pueden tratar de identificar posibles problemas más rápidamente, evitar que se agraven y ayudar a los contribuyentes a presentar sus declaraciones correctamente. Para garantizar la confiabilidad de estos modelos de riesgo, es esencial que los datos de la formación sean exactos. Recuadro 5.3 en Austria se ofrece un ejemplo. Como otro ejemplo, en el mercado griego de combustibles, la IAPR ha desarrollado un sistema de análisis que utiliza datos de diversas fuentes para componer más de 100 criterios diferentes de análisis de riesgos³.

Recuadro 5.3. Mejoramiento de los controles de cumplimiento normativo con analítica predictiva en Austria

Desde 2014, la administración tributaria austriaca aplica algoritmos de ML a través del Centro de Competencia de Analítica Predictiva (PACC, por sus siglas en inglés), una unidad especializada del Ministerio Federal de Finanzas. Encargado de modernizar la gestión de riesgos, el PACC trabaja para mejorar la recaudación de impuestos, la auditoría y la detección de fraudes. Organizado en cuatro

áreas temáticas (Analítica Predictiva, Analítica Avanzada, Analítica Fiscal y Analítica Aduanera), el PACC aborda una amplia gama de desafíos en todo el sistema fiscal.

En 2023, los modelos de riesgo del PACC analizaron alrededor de 6,5 millones de casos en los sectores del impuesto sobre la renta, el impuesto de sociedades y el valor agregado, así como en las transacciones aduaneras. Los análisis detectaron casos de declaración falsa en las liquidaciones fiscales de los empleados e identificaron actividades fraudulentas, que dieron lugar a unos ingresos fiscales adicionales de aproximadamente 185 millones de euros. También se examinaron casi 27,5 millones de casos para comprobar su cumplimiento normativo, de los cuales 375 000 se consideraron para un examen más detallado debido a perfiles de riesgo inverosímiles. Las técnicas avanzadas, como los árboles de decisión, los modelos de regresión y la minería de texto, respaldan las auditorías retrospectivas y en tiempo real, mientras que los proyectos en curso tratan de ampliar las capacidades analíticas para incluir, por ejemplo, modelos generativos.

Fuente: <https://www.bmf.gv.at/en/press/press-releases/2024-New/August-2024/BMF-generated-around-EUR-185-million-in-tax-income-from-AI-in-2023-.html>.

Mejoramiento de servicios al contribuyente

Ayudar a los contribuyentes a cumplir sus obligaciones tributarias es una actividad de larga data de las administraciones tributarias que suelen utilizar métodos de implementación de servicios que requieren mucha mano de obra, como los centros de atención telefónica o las visitas a las oficinas físicas de la administración tributaria. Durante muchos años, las administraciones tributarias han adoptado la tecnología de asistentes virtuales basada en IA para ayudar a los contribuyentes de forma más rápida mediante consultas o procesos sencillos y liberar recursos. Los asistentes básicos emplean técnicas basadas en reglas que pueden orientar a los contribuyentes hacia la información pertinente, realizar pagos o comprobar su estado. Según datos recientes del ITTI, los asistentes virtuales que siguen un conjunto de reglas preprogramadas al interactuar con los contribuyentes son un caso de uso muy común para la IA⁴. Más recientemente, las administraciones tributarias han empezado a utilizar técnicas de IA más avanzadas, como los modelos de lenguaje amplio (ML), para mejorar y personalizar las interacciones con los contribuyentes. El ejemplo de Recuadro 5.4 ilustra la gama de servicios posibles y los niveles crecientes de sofisticación que permiten respuestas más personalizadas. En otro ejemplo, la IAPR de Grecia está desarrollando un asistente digital dirigido por un programa de aprendizaje permanente que presta asesoramiento a los contribuyentes respondiendo a preguntas durante la presentación de las declaraciones del impuesto sobre la renta⁵.

Recuadro 5.4. Mejoramiento de servicios al contribuyente en Singapur

La Agencia Tributaria de Singapur (IRAS, por sus siglas en inglés) ha empleado herramientas de automatización del usuario final (es decir, herramientas que simplifican las cosas para los usuarios del servicio), datos e IA para ofrecer servicios fluidos y personalizados a los contribuyentes. En 2021, la IRAS lanzó un chatbot para gestionar de manera más eficiente las consultas habituales sobre el impuesto sobre la renta de las personas físicas, el impuesto de sociedades, el impuesto sobre bienes y servicios, el impuesto sobre bienes inmuebles, los derechos de timbre, las retenciones a cuenta y los asuntos tributarios para empleadores. El chatbot funciona a través de la plataforma VICA (Virtual Intelligent Chat Assistant) del Gobierno de Singapur.

El chatbot IRAS VICA mejora la interacción del usuario mediante una experiencia de usuario humanizada, con elementos interactivos como carruseles e infografías. En 2023, se actualizó el chatbot

con la integración de un motor de Modelo Extenso de Lenguaje (LLM), lo que mejoró significativamente su capacidad para comprender la intención de las consultas de los contribuyentes.

Además de responder a las consultas, el chatbot también permite el acceso a servicios autenticados. Los contribuyentes pueden consultar cómodamente su saldo fiscal pendiente, consultar sus planes de pago, comprobar el estado de los pagos, cancelar o restablecer sus planes de pago y realizar pagos de impuestos a través de terminales de pago de autoservicio y canales móviles. Estas operaciones con bots ahorran a los contribuyentes unos 10 minutos por transacción, en comparación con los canales de servicio digitales tradicionales. En el ejercicio financiero de 2024, el chatbot VICA de la IRAS gestionó unas 70 000 consultas transaccionales, lo que podría suponer un ahorro estimado de 11.666 horas para los contribuyentes.

Fuente: (OECD, 2022^[1]), <https://www.iras.gov.sg/digital-services/others/iras-bot>, <https://www.tech.gov.sg/products-and-services/for-government-agencies/productivity-and-marketing/vica>, funcionarios del Gobierno de Singapur.

Asistencia a contribuyentes con su declaración

Además, la IA ha facilitado la cumplimentación previa de declaraciones de impuestos en gran medida; los algoritmos y otras herramientas examinan los extensos conjuntos de datos de las administraciones tributarias y, a continuación, los incluyen en la parte correcta de las declaraciones. Esto puede mejorar la exactitud de las declaraciones de impuestos y simplificar los procesos de presentación, reduciendo así el tiempo que los contribuyentes dedican a presentar sus declaraciones. En los casos en los que los contribuyentes presentan sus declaraciones completas, las administraciones tributarias están implementando técnicas de detección de patrones que pueden detectar anomalías y errores y pueden instar a los contribuyentes a comprobar sus declaraciones una y otra vez.

Recuadro 5.5. Prevención de errores en declaraciones de impuestos australianas

La Agencia Tributaria australiana (ATO, por sus siglas en inglés) está empleando tecnologías avanzadas para mejorar la exactitud y eficiencia de las declaraciones de impuestos. Entre ellos se incluyen la analítica en tiempo real, los formularios prellenados y los sistemas de detección de anomalías para ayudar a los contribuyentes a cumplir sus obligaciones y reducir los errores.

Como parte de sus esfuerzos, la ATO ofrece un servicio de prellenado que completa automáticamente las declaraciones de impuestos individuales con datos procedentes de empleadores, bancos, organismos públicos y otros terceros. Esto incluye información como el salario, los intereses bancarios, los dividendos y los detalles del seguro de salud privado. Los contribuyentes deben revisar y confirmar los datos prellenados antes de presentarlos. Este sistema no solo reduce las cargas administrativas, sino que también mejora la precisión al minimizar los errores de entrada manual.

Además, la ATO utiliza avisos en tiempo real durante el proceso de presentación para abordar posibles anomalías. Por ejemplo, si las cifras declaradas por un contribuyente son declaradas por un modelo como anormales, el sistema genera un mensaje en el que se le anima a comprobar sus aportes. Un mensaje podría indicar, por ejemplo: "Usted no ha declarado ningún ingreso por intereses. Por favor, revise y asegúrese de que se hayan declarado todos los intereses devengados." En 2023-24, se emitieron más de 636 000 avisos a individuos, lo que ayudó a proteger unos 78 millones de dólares australianos en ingresos.

Fuente: (OECD, 2022^[1]), <https://aws.amazon.com/institute/future-of-tax-technology-ai-ml-cloud-security>, <https://www.austlii.edu.au/au/journals/JCULawRw/2024/5.pdf>, <https://www.ato.gov.au/individuals-and-families/your-tax-return/how-to-lodge-your-tax-return/lodge-your-tax-return-online-with-mytax/pre-filling-your-online-tax-return>.

Gestión de riesgos y desafíos

Si bien todas las implementaciones de IA en las administraciones públicas suelen estar acompañadas de desafíos, la administración tributaria crea un entorno especialmente exigente para la integración de la IA debido a tres factores críticos:

1. Garantizar la protección de los datos y la privacidad: las administraciones tributarias tienen acceso a una gran cantidad de datos sensibles de los contribuyentes y tienen la obligación de proteger la privacidad y la confidencialidad de estos datos. Por lo tanto, es necesario adaptar los marcos jurídicos y de gobernanza para tener en cuenta el uso de la IA.
2. Garantizar los derechos de los contribuyentes: La capacidad de los contribuyentes para impugnar las decisiones fiscales adoptadas por las autoridades tributarias exige que la toma de decisiones basadas en los sistemas de IA y los procesos administrativos asociados sea explicable, transparente y responsable.
3. Mantener la confianza: las administraciones tributarias dependen en gran medida del cumplimiento voluntario de los contribuyentes, que requiere una confianza sostenida en la imparcialidad y la equidad de los procesos administrativos. Los contribuyentes deben poder confiar en el uso que hace el gobierno de los sistemas de IA, seguir creyendo que los sistemas tributarios son justos y seguir dispuestos a cumplir con sus obligaciones tributarias.

Riesgos asociados

- Datos inadecuados o sesgados en los sistemas de IA.
- Falta de transparencia y explicabilidad.

Si los sistemas de IA se basan en **datos inadecuados o sesgados**, podrían producirse resultados inexactos o adversos para algunas personas o grupos. En lo que respecta a la administración tributaria, esto podría dar lugar a evaluaciones de riesgos inexactas y, por tanto, a una orientación inadecuada de algunas personas o grupos para la adopción de medidas de control (por ejemplo, inspecciones). Véase Recuadro 5.6 por ejemplo.

Al adoptar la IA en la administración tributaria, la **inadecuada y deficiente calidad de los datos** plantea un riesgo significativo para la efectividad de los resultados, lo que puede socavar la confianza pública. Para que los sistemas de IA obtengan resultados confiables y confiables, los datos de entrada deben ser exactos, completos y bien estructurados. Las discrepancias en los formatos de los datos, los registros incompletos o mal escritos, los datos obsoletos de los contribuyentes o la falta de datos en el contexto necesario pueden dar lugar a predicciones inexactas, modelos de riesgo erróneos y resultados sesgados. Algunos sistemas de ML (ML) no requieren que los datos estén estructurados, pero la calidad y la precisión siguen siendo críticas. Los insumos incorrectos pueden hacer que los sistemas sean menos eficientes, socavar la confianza de los contribuyentes en el gobierno y tener consecuencias adversas para los contribuyentes.

Recuadro 5.6. Desafíos asociados a errores algorítmicos en la administración tributaria

El "Toeslagenaffaire" fue un escándalo de las prestaciones por hijos a cargo en los Países Bajos. Entre 2005 y 2019, la Administración Tributaria y Aduanera neerlandesa (Belastingdienst) acusó injustamente a unas 26 000 familias de solicitar fraudulentamente prestaciones de guardería, debido al uso de datos defectuosos y un algoritmo sesgado. Muchas de estas familias fueron marcadas debido a errores administrativos, como la falta de firmas, y familias con doble nacionalidad o origen migrante fueron objeto de medidas desproporcionadas. Estas familias se vieron obligadas a pagar decenas de

miles de euros en prestaciones, sumiéndolas en graves dificultades financieras, y algunas perdieron sus hogares, empleos y matrimonios. En casos extremos, los niños fueron separados de sus familias. El escándalo resultante resultó en el colapso del gobierno.

Después del Toeslagenaffaire, el Ministerio del Interior y Relaciones del Reino de los Países Bajos encargó a los investigadores de la Universidad de Utrecht que elaboraran algoritmos de evaluación de impacto sobre los derechos humanos (IAMA). En 2022, la Cámara de Representantes de los Países Bajos aprobó una moción para "hacer obligatoria la realización de esta evaluación de impacto antes de utilizar algoritmos cuando estos se utilicen para realizar evaluaciones o tomar decisiones sobre las personas".

Fuente: (OECD, 2023^[2]).

Las dificultades relativas a la **transparencia, la explicabilidad** (o el grado de comprensibilidad) e **interpretabilidad** de los sistemas complejos de IA pueden plantear riesgos para el Estado de Derecho y para el recurso jurídico de los contribuyentes. La capacidad de impugnar decisiones es un componente fundamental de los procesos de la administración tributaria. La naturaleza de "caja negra" de la IA en muchos sistemas de ML puede dar lugar a una falta de transparencia, explicabilidad e interpretabilidad que puede afectar a los derechos de los contribuyentes a entender cómo se toman las decisiones fiscales, lo que a su vez puede limitar sus derechos a impugnar una decisión.

Para mitigar estos riesgos, las autoridades tributarias deben invertir en marcos de gobernanza. Estos marcos deben garantizar que la IA se despliegue de forma que respete el equilibrio entre la recaudación tributaria de la administración y los derechos de los contribuyentes. Por lo tanto, los sistemas de IA y los procesos operativos asociados de las administraciones tributarias deben ser suficientemente transparentes y explicables y funcionar con mecanismos adecuados de supervisión y rendición de cuentas. Estos pueden incluir, por ejemplo, la gobernanza de datos (véase el Capítulo 4, sección "Crear una base de datos sólida"), en particular en términos de limpieza de datos, procesos de validación y actualizaciones periódicas para mantener la integridad y la confiabilidad de los datos utilizados para fundamentar la toma de decisiones impulsada por la IA. Solo con datos confiables y de alta calidad la IA puede mejorar realmente la administración tributaria al mejorar la precisión, el cumplimiento normativo y la eficiencia operativa para los contribuyentes.

Desafíos de implementación

- Brechas de habilidades.
- Elevados costos de adopción y escalabilidad de la IA.
- Entornos legales y regulatorios inflexibles u obsoletos.

Aprovechar toda la gama de oportunidades que la IA puede brindar a la administración tributaria pone de relieve tres desafíos clave.

En primer lugar, las administraciones tributarias informan de una **competencia feroz por empleados calificados** que pueden desarrollar e implementar estas tecnologías. Acompañando esto está el requisito de entrenar a un gran número de empleados existentes en el uso de esta nueva tecnología y la gestión de los cambios asociados en el proceso.

En segundo lugar, el despliegue de esta tecnología suele conllevar **importantes costos de inversión**. El proceso para garantizar los presupuestos puede resultar complicado, ya que el rendimiento de la inversión puede variar en función de la incertidumbre o a largo plazo. Además de contribuir a prestar nuevos servicios a los contribuyentes, esta inversión también es necesaria para ayudar a defender el sistema fiscal frente a los ataques fraudulentos.

Por último, la IA puede **alterar los marcos jurídicos existentes** que sustentan muchos procesos de las administraciones tributarias. La adaptación de estos marcos debe considerarse paralelamente a las directrices para garantizar el diseño y el despliegue efectivos de la IA. También cabe señalar que, si bien es posible que los países dispongan de directrices generales para el despliegue de la IA, a menudo no se adaptan a las necesidades específicas y las complejidades específicas de la administración tributaria, por lo que debe considerarse más a fondo.

Potencial sin explotar y camino a seguir

Las administraciones tributarias de todo el mundo ya han empezado a aprovechar el potencial de la IA y a desplegarla en una amplia gama de actividades para mejorar la eficiencia de sus operaciones. La IA aprovecha el potencial no solo para optimizar los modelos operativos existentes de las administraciones tributarias, sino también para transformarlos. El Marco de la Administración Tributaria 3.0 de la OCDE (2020^[3]) establece una visión de transformación digital en la que la tecnología digital se utiliza para crear ecosistemas fiscales fluidos y basados en eventos. Esto permite integrar la recaudación y la declaración de impuestos directamente en los sistemas naturales de los contribuyentes personas físicas y empresas.

Para materializar plenamente el potencial transformador de la IA, las administraciones tributarias deben adoptar un enfoque sistémico con respecto a su aplicación. La IA no es una herramienta aislada, sino que debe utilizarse como un componente integrado anexado en un ecosistema tributario más amplio, que abarque marcos regulatorios, estructuras organizativas, infraestructura técnica existente y personal experto. Es fundamental que las administraciones tributarias tengan en cuenta la necesidad de ajustar el proceso de elaboración de las normas. Por ejemplo, al aplicar los enfoques de reglas como código (RaC, por sus siglas en inglés), las administraciones tributarias pueden codificar la legislación y la normativa tributaria en formatos legibles por máquina, lo que permite a los sistemas de IA interpretar y aplicar normas con mayor exactitud y coherencia (Mohun and Roberts, 2020^[4])⁶.

Dado que la confianza de los contribuyentes es una parte esencial del buen funcionamiento de una administración tributaria, las administraciones tributarias también deben garantizar que la IA desplegada sea confiable, transparente y responsable. Para apoyar a las administraciones tributarias en este esfuerzo, el Foro de Administración Tributaria de la OCDE está actualmente probando un marco para apoyar a las administraciones tributarias en su despliegue de sistemas de IA. Sobre la base de los principios de la OCDE en materia de IA confiable, el marco describe consideraciones clave para cada fase del ciclo de vida de la IA y describe el aprendizaje y las experiencias de una serie de administraciones a partir de sus propias implementaciones. A través de esta cooperación en curso, las administraciones tributarias pueden adoptar un enfoque estructurado y basado en el riesgo para mejorar el uso de la IA en la administración tributaria, que también puede adaptarse a los cambios del contexto general. El Foro de Administración Tributaria tiene previsto publicar sus conclusiones sobre esta labor en 2026.

IA en la gestión de finanzas públicas

En las últimas décadas, la gestión financiera pública (GFP) ha integrado regularmente nuevas tecnologías, en particular mediante la adopción de sistemas de información de gestión financiera (FMIS, por sus siglas en inglés) cada vez más sofisticados. A medida que la tecnología ha avanzado y la calidad de los datos ha mejorado, las organizaciones de GFP han adoptado nuevos enfoques tecnológicos, como el análisis de datos, las herramientas de inteligencia empresarial y la automatización robótica de procesos (RPA, por sus siglas en inglés). Estos han sido los elementos básicos de la aplicación de la IA, que generalmente se ha utilizado para mejorar y profundizar en las capacidades existentes, al integrarse los sistemas y enfoques de IA en los productos y procesos existentes. En la GFP, la IA ha sido hasta ahora, en general, una continuación de la evolución tecnológica del sector, más que una revolución tecnológica.

Situación actual

Los usos actuales y previstos de la IA en la GFP están relacionados principalmente con la mejora de los procesos existentes, en lugar de con su revisión y rediseño completos (por ejemplo, eliminando cualquier intervención humana) o con la creación de procesos completamente nuevos. En la actualidad, la IA se utiliza como asistente que facilita la automatización de tareas pequeñas y a menudo rutinarias para los funcionarios, a gran velocidad y bajo costo de transacción, y como asesor que analiza datos históricos o en tiempo real para prever futuros eventos o comportamientos y que respalda los análisis propios de los funcionarios.

En este contexto, los enfoques de IA, especialmente el ML, tienen aplicaciones actuales para la GFP en las previsiones⁷ macroeconómicas y macrofiscales y en el apoyo a las decisiones de gasto; la planificación y el seguimiento presupuestarios; la gestión financiera, la presentación de informes y la supervisión; y la participación con partes interesadas externas.

Mejoramiento del pronóstico

La IA puede ayudar a afrontar los desafíos que plantean los métodos tradicionales de pronóstico financiera y económica, al mejorar la precisión y la puntualidad de las predicciones y permitir que los sistemas de IA superen a los modelos de pronóstico económica tradicionales (Jung, Patnam and Ter-Martirosyan, 2018^[5]).

Las capacidades de predicción de la IA también se consideran una oportunidad para desarrollar el “nowcasting”, identificando cambios en tiempo casi real y extrapolando posibles futuros a corto plazo. “Nowcasting” tiene en cuenta el pasado y el presente muy reciente para pronosticar el estado futuro muy cercano de los indicadores económicos que normalmente solo se determinan después de un retraso y están sujetos a revisión, como el producto interior bruto (PIB) o la inflación. Los bancos centrales de diferentes países están explorando la adopción de sistemas de IA para ofrecer predicciones más precisas y anticipadas que los modelos de series temporales tradicionales, incluso cuando se utilizan datos no estructurados, como el caso de Nueva Zelanda (Richardson, van Florenstein Mulder and Vehbi, 2019^[6]), Francia (Blanchet and Coueffe, 2020^[7]) y Perú (Tenorio and Perez, 2023^[8]).

Recuadro 5.7. Pronóstico del PIB con IA explicable en Suecia

La Autoridad Nacional de Gestión Financiera (ESV) de Suecia ha desarrollado una innovadora aplicación de pronóstico del PIB que aprovecha el ML explicable para mejorar tanto la precisión como la transparencia de las predicciones económicas. El modelo supera las previsiones oficiales de Suecia previas a la pandemia y aborda una limitación clave de las previsiones tradicionales basadas en la IA —su opacidad— visualizando el impacto de las variables a lo largo del tiempo. Esta herramienta permite a los responsables de la formulación de políticas, a los investigadores y al público en general generar proyecciones confiables, al tiempo que mantiene la interpretabilidad y respalda la toma de decisiones basada en datos. La iniciativa muestra cómo la IA puede integrarse en las previsiones macroeconómicas y, al mismo tiempo, garantizar la rendición de cuentas y la confianza.

Fuente: <https://oecd-opsi.org/innovations/forecasting-gdp>.

Facilitar las decisiones de gasto

Para facilitar la toma de decisiones en la GFP, la IA está respaldada por avances tecnológicos que se han generalizado. Entre ellos se incluyen los macrodatos para analizar grandes cantidades de información de múltiples fuentes, las herramientas de análisis de datos para desglosar en categorías financieras específicas y en función de los beneficiarios, y evaluar la efectividad de los gastos en función de tendencias y patrones, y las visualizaciones de datos para permitir una comunicación efectiva de información compleja.

La IA puede utilizarse para consolidar estos cimientos identificando tendencias y patrones y agrupando puntos de datos en función de la similitud o de características compartidas. En el caso específico de las decisiones de gasto, la IA puede analizar datos históricos de ejecución presupuestaria para identificar patrones de gasto insuficiente o excesivo, predecir necesidades de gasto futuras en función de parámetros clave (por ejemplo, cambios demográficos) y evaluar la efectividad del programa vinculando datos de gasto con métricas de resultados. Esto ofrece oportunidades para acelerar y mejorar los análisis de maneras que puedan automatizar o aumentar el trabajo de los seres humanos. Las técnicas de ML que aprovechan datos no estructurados ofrecen oportunidades para combinar conjuntos de datos que antes no se utilizaban para estos ejercicios (Recuadro 5.8).

Recuadro 5.8. IA para la gestión de las finanzas públicas en Corea

En 2022, Corea del Sur desarrolló e implementó dBrain+, un sistema avanzado de información de gestión financiera que aprovecha la IA para analizar datos económicos, fiscales y financieros en tiempo real, optimizando la evaluación de riesgos y la toma de decisiones en las finanzas públicas. Sus módulos clave, el Sistema de Información Fiscal de Corea (KFIS, por sus siglas en inglés) y la Evaluación de Riesgos y el Escaneo Horizontal de Corea (KORAHS, por sus siglas en inglés), utilizan la analítica impulsada por la IA para detectar riesgos financieros y apoyar las decisiones de política basadas en datos. Al centralizar todas las operaciones financieras nacionales, desde la elaboración de presupuestos y la gestión de fondos hasta la supervisión de la deuda y la evaluación del rendimiento, dBrain+ mejora la eficiencia, la transparencia y las capacidades de predicción en los gobiernos centrales y locales.

Un punto fuerte de dBrain+ es su integración con 63 sistemas de 46 instituciones, entre ellas el Servicio Nacional de Impuestos, el Servicio de Contratación Pública y el Banco de Corea, lo que permite una coordinación fluida de los contratos, la recaudación de impuestos y las transferencias de fondos. El análisis de estos datos en tiempo real impulsado por la IA mejora la ejecución presupuestaria, acelera

la presentación de informes financieros y apoya la identificación de riesgos, lo que permite tomar mejores decisiones sobre políticas fiscales y gasto público. Al proporcionar herramientas adaptadas a los diferentes usuarios —incluidos funcionarios públicos, investigadores y partes interesadas externas—, dBrain+ refuerza la rendición de cuentas y moderniza el enfoque de Corea del Sur para una gobernanza fiscal impulsada por la IA.

Fuente: (Korea Fiscal Information Service (KFIS), 2023^[9]),<https://www.adb.org/sites/default/files/publication/928976/governance-brief-052-digital-transformation-tax-administration-rok.pdf>.

Apoyo a la planificación y supervisión presupuestaria

La IA tiene la capacidad de respaldar los procesos de planificación y supervisión presupuestaria proporcionando productos que respaldan la formulación de bases de referencia de gastos precisas y la estimación de costos de las nuevas políticas. Por ejemplo, el Departamento de Asuntos de Veteranos de Australia desarrolló sistemas y herramientas de predicción para ayudar a simular los futuros impactos financieros de las decisiones de política. Estos incluyen el gasto fiscal anual de cada beneficiario, así como su promedio de años en beneficios, que se utilizan para la estimación de costos, las estimaciones presupuestarias y la evaluación de políticas (Australian Government, 2020^[10]).

Otro ámbito de aplicación prometedor de la IA es la identificación, el seguimiento y la mitigación de los riesgos fiscales mediante el análisis de grandes conjuntos de datos. Los riesgos fiscales para las administraciones públicas pueden deberse a diversas causas, como niveles de gasto o inversión insostenibles, que deben identificarse con antelación para poder adoptar medidas preventivas. La IA puede ayudar a identificar dichos riesgos, como se observa en el Recuadro 5.9. En otro ejemplo, Indonesia utiliza un sistema llamado Asesor Financiero AI (AIFA, por sus siglas en inglés) para procesar datos financieros y de rendimiento no estructurados con el fin de ofrecer análisis del rendimiento fiscal de los gobiernos subnacionales en tiempo real (Wisesa, 2023^[11]).

Recuadro 5.9. Francia utiliza IA en la supervisión presupuestaria

Desde hace varios años, la Agencia Tributaria francesa (DGFiP) viene implementando un "sistema de alerta" basado en la IA que tiene por objeto identificar municipios con dificultades financieras, proporcionarles asesoramiento financiero y apoyar proactivamente la aplicación de medidas correctoras.

Este sistema de alerta se basaba inicialmente en un algoritmo que utilizaba datos fiscales y financieros históricos para clasificar los municipios. Más recientemente, la DGFiP desarrolló un sistema de IA predictiva diseñado para identificar con antelación las dificultades financieras de los municipios. El sistema recibió formación sobre datos que abarcan cuatro años para predecir los resultados del quinto año. El sistema predictivo también se basa en técnicas de agrupamiento no supervisadas para clasificar los municipios con características financieras similares sin ejemplos de resultados predefinidos.

En 2022, un experimento con el sistema abarcó 2.500 municipios, de los cuales se identificó que alrededor del 40 % experimentaban dificultades financieras. De estos, alrededor del 17 % no habían sido detectados por el algoritmo anterior. Además, alrededor del 35 % de los municipios se identificaron con dificultades temporales no estructurales, lo que pone de relieve la capacidad del sistema para diferenciar entre problemas financieros permanentes y transitorios.

Fuente: (French Public Finances General Directorate (DGFiP), 2024^[12]).

Automatizar las actividades de gestión, presentación de informes y supervisión

La GFP y las actividades de reporte implican tareas importantes, pero a veces repetitivas, que son especialmente adecuadas para la automatización. Las técnicas de IA, como el procesamiento del lenguaje natural (NLP), pueden utilizarse para analizar imágenes digitales con el fin de extraer información de los documentos (por ejemplo, información sobre los proveedores), identificar y clasificar documentos (por ejemplo, facturas), realizar una comparación de documentos (por ejemplo, comparar la información sobre facturas y los proveedores) o identificar tendencias y patrones (por ejemplo, controles internos de las solicitudes de pago).

En Francia, por ejemplo, la Agencia Tributaria francesa (DGFiP) (2024^[13]) ha desarrollado una herramienta basada en IA como parte de los procesos de control interno habituales que "automatiza la selección de las solicitudes de pago que deben controlarse [y] optimiza la carga de trabajo y la calidad de los controles realizados". El Centro de Servicios Compartidos del Gobierno de Finlandia para Finanzas y Recursos Humanos (Palkeet) estableció un Centro de Excelencia para la Automatización Robótica de Procesos (RPA). Se centra en el desarrollo y despliegue de soluciones RPA en diversas actividades financieras y de recursos humanos, como la gestión de la información de proveedores, el equilibrio de los datos contables y el procesamiento de transacciones financieras, y la integración de la IA en procesos de automatización en los que es necesario tomar decisiones complejas o procesar datos (Palkeet, 2024^[14]).

Recuadro 5.10. Transparencia fiscal impulsada por IA en Brasil

La Tesorería Nacional de Brasil (STN) está utilizando IA para mejorar la transparencia fiscal al clasificar el gasto público subnacional de acuerdo con la norma internacional COFOG. La adopción de la IA, que antes era una tarea manual que requería muchos recursos y que utilizaba modelos de ML con redes neuronales convolucionales y recurrentes, ha reducido el tiempo de clasificación de 1000 horas de trabajo humano a solo 8 horas, al tiempo que ha logrado una precisión superior al 97 %. Este avance llevó a la publicación del informe Gasto por Función del Gobierno General en 2024, un hito en las estadísticas fiscales de Brasil.

Sobre la base de este éxito, STN está ampliando ahora las aplicaciones de IA a nuevas áreas, incluida la clasificación de los gastos relacionados con el clima. En colaboración con el Banco Interamericano de Desarrollo (BID), Brasil está fortaleciendo su capacidad para evaluar las implicaciones fiscales del cambio climático. Al ser pionero en la gestión de las finanzas públicas impulsada por la IA, Brasil sienta un ejemplo para otros países que buscan modernizar las estadísticas fiscales y mejorar la transparencia en un panorama económico cada vez más complejo.

Fuente: <https://blog-pfm.imf.org/en/pfmblog/2024/12/ai-is-enhancing-fiscal-transparency-in-brazil>.

Además de reforzar la presentación de informes, el desarrollo de verificaciones específicas para identificar errores (por ejemplo, pagos inadecuados) y fraude (por ejemplo, robo de identidad) se han convertido en un objetivo importante para muchos gobiernos. Esto es especialmente cierto tras la pandemia de COVID-19, que puso de manifiesto las vulnerabilidades de los sistemas de pago de varios países. La capacidad de la IA para identificar tendencias y patrones puede ayudar en este sentido. Por ejemplo, durante la pandemia, la Autoridad Empresarial Danesa desarrolló controles basados en la IA para las solicitudes de ayuda de empresas para diversos regímenes de ayuda (van Noordt and Tangi, 2023^[15]).

Facilitar la interacción con las partes interesadas y usuarios

Como se analiza a lo largo de este capítulo y se sintetiza en el Capítulo 2, los chatbots impulsados por NLP y modelos de lenguaje se emplean cada vez más en el gobierno para proporcionar servicios directamente. En cuanto a la gestión de las finanzas públicas, los Emiratos Árabes Unidos (EAU) han

desarrollado U-Ask, un chatbot unificado basado en IA para servicios públicos que también puede utilizarse para responder a preguntas⁸ sencillas sobre la declaración de datos fiscales. En México, el gobierno ha introducido una herramienta de asistencia virtual basada en la IA como parte de su Plataforma de Apoyo Inteligente, diseñada para guiar a los usuarios a través de los programas y apoyos del gobierno (2023^[16]). La herramienta proporciona información sobre los beneficios, la elegibilidad y los procesos de solicitud para individuos, empresas y gobiernos locales, utilizando una simple búsqueda de palabras clave o cuestionarios personalizados para adaptar la información al perfil del usuario.

Gestión de riesgos y desafíos

Riesgos asociados

- Falta de transparencia y explicabilidad.
- Datos inadecuados o sesgados en los sistemas de IA.

Debido a su naturaleza de caja negra, los sistemas basados en IA con los mejores resultados de predicción representan tanto un paso adelante en precisión como un paso atrás en transparencia fiscal (Jung, Patnam and Ter-Martirosyan, 2018^[5]). Esta **falta de transparencia** dificulta a los gobiernos la verificación de los procesos de toma de decisiones de estos modelos, lo que es crucial para la rendición de cuentas y el cumplimiento normativo. En consecuencia, esta limitación ha llevado a los gobiernos a dar prioridad al uso de sistemas de IA más simples para mejorar la modelización humana y el análisis de sensibilidad.

Si bien los gobiernos y las organizaciones de gestión de la productividad se han vuelto cada vez más expertos en la identificación de riesgos y desafíos relacionados con la IA, muchos siguen desarrollando marcos integrales y enfoques prácticos para gestionar estos riesgos. Sus esfuerzos se centran en establecer estructuras de gobernanza, desarrollar capacidad técnica y crear protocolos claros para el despliegue de la IA en los sistemas financieros públicos.

Los gobiernos también están trabajando para desarrollar métodos para "desempaquetar" los sistemas de IA y hacer que su razonamiento sea más transparente, explicable e interpretable, todas condiciones importantes para usar estos sistemas en la GFP. Por ejemplo, como se ha comentado anteriormente (Recuadro 5.7), la Autoridad Nacional de Gestión Financiera (ESV) sueca ha desarrollado una aplicación para analizar el impacto que cada variable de datos tiene en la predicción de los "modelos de caja negra", como parte de un programa de trabajo más amplio para integrar la IA en la gestión financiera del Gobierno Sueco (Boström et al., 2020^[17]).

Los casos de uso de la IA muestran que **riesgos éticos como datos incompletos o insuficientes y algoritmos sesgados** pueden ser significativos en el campo de la GFP. La IA puede amplificar los patrones de desigualdad incorporados en los datos financieros y conducir a la exclusión financiera de las personas percibidas como de alto riesgo (Crisanto et al., 2024^[18]). Este fenómeno se observa en el sector bancario, donde una distribución sesgada del crédito puede perpetuar las prácticas discriminatorias de préstamo (Bailey, 2023^[19]; Klein, 2020^[20]). En el ámbito de la gestión de las finanzas públicas, los algoritmos sesgados pueden afectar de manera similar a la distribución de los fondos públicos, las prestaciones sociales y el acceso a los programas gubernamentales, perpetuando las desigualdades existentes y obstaculizando un trato justo. Aunque el sistema australiano de Robodebt es más un sistema automatizado de toma de decisiones que una verdadera IA, ilustra los problemas que pueden surgir como resultado de errores algorítmicos si no son detectados y abordados por los seres humanos (Recuadro 5.11).

Recuadro 5.11. El esquema Robodebt: desafíos para cobrar pagos indebidos

El esquema Robodebt de Australia, introducido en 2016, fue un programa automatizado de recuperación de deudas diseñado para identificar y recuperar los pagos excesivos de asistencia social. Sustituyó un proceso manual por un algoritmo de coincidencia de datos que comparaba los datos quincenales sobre ingresos declarados a Centrelink, el organismo responsable de los pagos de la seguridad social, con las cifras medias de ingresos anuales de la Agencia Tributaria Australiana (ATO). Las discrepancias se señalaron como pagos en exceso y las notificaciones de deuda se emitieron automáticamente sin verificación humana. Este método de "promediado de ingresos" ignoraba las fluctuaciones en las ganancias reales, generando a menudo falsas deudas. El sistema también invirtió la carga de la prueba, al exigir que los receptores proporcionaran registros de pago históricos para impugnar las deudas, una tarea exigente para muchos. A lo largo de su funcionamiento, el régimen emitió 470 000 notificaciones de deuda incorrectas por un total de 775 millones de euros, lo que causó dificultades económicas y económicas generalizadas.

Los cálculos del régimen fueron declarados ilegales en 2019. En 2022 se creó una Comisión Real para investigar el establecimiento, el diseño y la implementación del esquema Robodebt; el uso de cobradores de deudas externos en el marco del esquema Robodebt; las preocupaciones planteadas tras la implementación del esquema Robodebt; y los resultados previstos o reales del esquema Robodebt. En 2023, la Comisión Real publicó un informe en el que se analizaban los efectos del régimen sobre los beneficiarios, incluidos los relacionados con la retención y el embargo de ingresos, los efectos emocionales y psicológicos y la pérdida de fe en el gobierno. Robodebt ejemplifica los riesgos de automatizar sistemas sociales complejos sin una supervisión humana adecuada o pruebas rigurosas. Las consecuencias incluyeron acuerdos legales significativos y pedidos de regulaciones más estrictas sobre el uso de algoritmos en las políticas públicas. Si bien las conclusiones de la Comisión Real no eran necesariamente representativas de las opiniones del Gobierno Australiano, el gobierno estuvo de acuerdo, o estuvo de acuerdo en principio, con 56 de las 57 recomendaciones de la Comisión.

El esquema Robodebt, que utilizaba cotejo automatizado de datos, promedio de ingresos y cálculo de pagos en exceso, puede describirse como un sistema automatizado de toma de decisiones. Si bien el esquema no apalancó la IA, ayuda a ilustrar problemas de gobernanza, supervisión humana y diseño algorítmico.

Fuente: (OECD, 2023^[2]), <https://robodebt.royalcommission.gov.au>, <https://ministers.dss.gov.au/media-releases/13091>.

Desafíos de implementación

- Hacer coincidir problemas con soluciones de IA.
- Entornos legales y regulatorios inflexibles u obsoletos.
- Sistemas obsoletos de tecnología de la información.
- Falta de datos de alta calidad y capacidad para compartirlos.
- Brechas de habilidades.
- Falta de marcos viables y directrices sobre el uso de la IA.

Uno de los desafíos es **coincidir las necesidades de GFP y tecnologías de IA**. La mayoría de los Ministerios de Hacienda que ya han implementado proyectos de IA en la GFP hacen hincapié en la importancia de cartografiar procesos y actividades para identificar áreas de ineficiencia y posibles aumentos de eficiencia como requisito previo para desplegar la IA⁹. Una vez completada esta fase inicial,

el siguiente paso es evaluar la idoneidad de la tecnología de IA u otras tecnologías para la integración con el fin de ayudar a abordarlas.

A pesar de la capacidad de la IA para resumir o redactar textos utilizando modelos lingüísticos, los Ministerios de Hacienda han actuado con cautela a la hora de implantar nuevas tecnologías en los informes fiscales (por ejemplo, la elaboración automática de informes fiscales). Esto puede deberse a la **preocupación por la precisión y por si la IA cumple la normativa vigente**. También puede haber dudas sobre quién es el responsable del uso de la IA, y sobre lo que significa su uso para las personas que ocupan puestos de responsabilidad, como los auditores externos, y las personas que utilizan los informes, como los legisladores.

Los sistemas de tecnología de la información (TI) son cruciales para que los Ministerios de Hacienda puedan aprovechar las oportunidades que ofrece la IA. Sin embargo, muchos países de la OCDE afirman que están **atrapados en tecnologías heredadas** que están significativamente fragmentadas, a menudo obsoletas y carecen de la infraestructura y la compatibilidad necesarias para integrar las funcionalidades avanzadas de IA. Por ejemplo, en la mayoría de los países de la OCDE, los sistemas FMIS gestionados de forma centralizada tienen más de 10 años (OECD, 2024^[21]). Estas tecnologías no son específicas de los países de la OCDE y están frenando el uso de la IA en la GFP en todo el mundo (Rivero del Paso et al., 2023^[22]).

Como ocurre con cualquier sistema informático, la calidad de los resultados obtenidos con un sistema de IA depende de la calidad de las entradas. Los Ministerios de Hacienda también indican que la **fragmentación de los datos, junto con las restricciones al intercambio de datos**, suele impedir el inicio de proyectos de IA (véase el Capítulo 4, sección "Creación de una base de datos sólida"). Estos desafíos ponen de relieve la necesidad de mejorar las prácticas y políticas de gestión de datos que facilitan una accesibilidad e intercambio de datos más efectivos. En consecuencia, varios países de la OCDE planean importantes actualizaciones de su FMIS, reconociendo también la necesidad de fundamentos más sólidos para los datos (Gráfico 5.2).

Gráfico 5.2. Objetivos para las actualizaciones del FMIS en países de la OCDE, 2022



Nota: Se refiere solo a los países que actualmente están llevando a cabo un desarrollo importante o sustituciones de su sistema central de información financiera y de gobernanza (18 países). Las calificaciones presentan el nivel promedio de importancia asignado a cada objetivo en una escala de 0 a 4 por todos los encuestados. No se dispone de datos de Chile, Colombia, Eslovenia, Estados Unidos, Israel y México.

Fuente: (OECD, 2022^[23]), pregunta 24.

Dado que la GFP es un ámbito de actuación muy técnico, la implementación de sistemas de IA requiere una formación sustancial que implique la intervención y supervisión humanas (y la retroalimentación) de los expertos en GFP. Un desafío importante es que, si bien las organizaciones de gestión de las finanzas públicas pueden identificar eficazmente los riesgos relacionados con la IA, muchas carecen de las **competencias del personal especializado y de las capacidades institucionales** necesarias para desarrollar y aplicar los marcos de gestión de riesgos necesarios. Además, resulta tentador sobreestimar

las capacidades de la IA. Los sistemas de IA producen resultados que se basan en probabilidades, lo que significa que, por naturaleza, son inciertos. Pueden producir salidas de datos que son incorrectas o textos que suenan altamente autoritarios pero que son incorrectos ("alucinaciones"). Por lo tanto, no importa cuánto puedan recibir los sistemas de capacitación, los especialistas en gestión de finanzas públicas deben ser capaces de ejercer un juicio crítico al utilizar los resultados que generan. Esta supervisión requiere especialistas que combinen habilidades técnicas en GFP y un conocimiento básico del funcionamiento de la IA.

Garantizar la transparencia y la explicabilidad en las decisiones basadas en IA **requiere marcos y normas sólidas que gobiernen y supervisen los procesos de IA**, incluida la trazabilidad y la accesibilidad de los conjuntos de datos de formación (por ejemplo, la procedencia de los datos y los registros de datos) y, cuando sea posible, el código fuente de los algoritmos de IA. Sin embargo, hasta hace poco, en muchos países de la OCDE dichos marcos y normas no existían. Los órganos de supervisión tradicionales, como las entidades fiscalizadoras superiores (EFS) y las instituciones fiscales independientes, están empezando a adaptar sus metodologías y a desarrollar nuevas competencias para supervisar eficazmente los procesos fiscales impulsados por la IA (Recuadro 4.8). Además, es fundamental identificar e implementar salvaguardias contra el posible uso indebido o la dependencia excesiva de la IA en la gobernanza fiscal, a medida que los sistemas de IA adquieran un mayor protagonismo en los procesos de toma de decisiones. Los especialistas en GFP deben abordar estos problemas para diseñar un nuevo marco de gobernanza fiscal que aproveche el potencial de la IA y, al mismo tiempo, mantenga la transparencia, la responsabilidad y la integridad.

Potencial sin explotar y camino a seguir

Los Ministerios de Hacienda están adoptando actualmente un enfoque prudente con respecto a la IA, dando prioridad a la automatización de tareas y a las aplicaciones predictivas frente a una IA más prescriptiva. Mientras que la IA predictiva se centra en pronosticar los resultados, la IA prescriptiva va más allá al sugerir cursos de acción para lograr los objetivos deseados o mitigar los riesgos. Sin embargo, una aplicación sistemática de la IA prescriptiva podría tener un profundo impacto en las funciones y responsabilidades de los sistemas de gestión de las finanzas públicas, lo que podría alterar las actividades de las partes interesadas en materia fiscal y los órganos de supervisión.

Cualquier cambio de los juicios humanos a los resultados basados en el sistema en las decisiones fiscales requiere una reevaluación de la rendición de cuentas basada en el riesgo, y la asignación de funciones y responsabilidades en un entorno cada vez más automatizado. Esto también requiere considerar cómo la IA podría alterar la letra y espíritu de las instituciones y procesos de GFP, y cómo podría remodelar las funciones de las partes interesadas en el ámbito fiscal, incluidos los órganos de supervisión externa como las EFS y las instituciones fiscales independientes. En este contexto, deben abordarse las cuestiones relativas al futuro de la transparencia y la rendición de cuentas.

- ¿Qué mecanismos serán necesarios para garantizar la transparencia de las decisiones automatizadas de GFP? Dado que la IA podría llevar a cabo cada vez más el pronóstico y la planificación y seguimiento presupuestarios, debería existir un marco sólido que garantice la transparencia y la explicabilidad en lo que respecta al uso que hacen los gobiernos de los sistemas de IA y sus datos de base. Esto implica crear e implementar estándares que gobiernen el uso y la supervisión de la IA.
- ¿Cómo evolucionarán las funciones de los órganos de supervisión tradicionales a medida que se adopte la IA? Las EFS y las instituciones fiscales independientes tendrán que adaptar sus metodologías para supervisar y auditar eficazmente los procesos fiscales impulsados por la IA. Esto podría incluir el desarrollo de la ciencia de datos y la formación en IA para los miembros del

personal o la actualización de los procesos de auditoría para incorporar consideraciones específicas de la IA, como la integridad y la exhaustividad de los datos.

- ¿Qué nuevas salvaguardias serán necesarias para proteger contra el uso indebido de la IA en la gobernanza fiscal? Dado que los sistemas de IA desempeñan un papel más destacado en la toma de decisiones fiscales, resulta fundamental identificar posibles tipos de usos indebidos intencionados o no intencionados y protegerse frente a ellos.

Al abordar estas cuestiones, los especialistas en GFP pueden ayudar a dar forma a un nuevo marco de gobernanza fiscal que dé cabida al pleno potencial innovador de la IA, salvaguardando al mismo tiempo los principios clave de la GFP.

Los esfuerzos de IA de los Ministerios de Finanzas necesitan un enfoque coordinado y deben integrar las lecciones de otros proyectos gubernamentales. Esto podría minimizar en gran medida los riesgos de fracaso y los desafíos de implementación, al tiempo que mejoraría la calidad de los resultados. Por ejemplo, la comunidad de práctica en proyectos de IA puede ayudar a prevenir errores comunes y evitables y a mitigar los riesgos de los proyectos de IA intrínsecamente complejos. Estos riesgos se derivan, por ejemplo, de la participación de múltiples partes interesadas con intereses diversos o con una falta de comprensión de las tecnologías y los sistemas.

Si bien muchas agencias de gestión financiera identifican un potencial de aumento de la productividad a gran escala gracias a la IA, los estudios públicos sobre viabilidad y costos son poco frecuentes y las evaluaciones de resultados e impactos siguen siendo anecdóticas. No se recogen pruebas de los costos y los impactos o no se ponen a disposición del público debido al carácter provisional de los proyectos que todavía se encuentran en fases piloto o iniciales de desarrollo.

Idealmente, los resultados y el impacto del uso de la IA en la GFP deberían evaluarse utilizando marcos de evaluación para realizar un seguimiento de los costos de los proyectos y los indicadores clave de rendimiento de los proyectos finalizados, incluidos el ahorro de costos, el aumento de la efectividad y la eficiencia, la reducción de errores y la mejora del cumplimiento normativo. Esto implicaría, entre otras cosas, monitorear todos los costos de los proyectos, comparar las métricas del antes y el después de la implementación de la IA, realizar encuestas de satisfacción a las partes interesadas y analizar los datos para ver cómo los resultados de la IA se ajustan a los objetivos de política fiscal.

Debido a la falta de datos de impacto, órganos de supervisión externa pidieron recientemente una mayor transparencia en los proyectos de IA de todo el gobierno; dijeron que se requiere un mayor escrutinio y recopilación de pruebas, además de una inversión sustancial (véase el Capítulo 4, secciones sobre "Invertir a propósito" y "Capacitar a los órganos de supervisión y asesoramiento para orientar la IA responsable").

Es fundamental establecer marcos sólidos de selección y evaluación de proyectos en los Ministerios de Hacienda. Estos deben alinearse con los marcos de todo el gobierno. Deberían realizar un seguimiento no solo de los indicadores de costos y rendimiento, sino también de los impactos cualitativos de la IA, como la reducción de errores y la mejora del cumplimiento normativo. Esto es necesario, ya que es probable que en el futuro varios proyectos compitan por recursos de inversión limitados.

Para ampliar sus ambiciones de forma progresiva y segura, los Ministerios de Hacienda podrían adoptar un enfoque secuencial para la introducción de las tecnologías de IA. Esto queda bien ilustrado en el caso del Centro de Servicios Compartidos del Gobierno Finlandés para Finanzas y Recursos Humanos (Palkeet), como se ha comentado anteriormente, que comenzó con "pequeños" usos de RPA y ahora está avanzando hacia la hiperautomatización con ML. Según Palkeet, empezar desde pequeña con su avanzada trayectoria de digitalización también ayudó a aumentar la aceptación de las tecnologías de IA más sofisticadas por parte de los funcionarios públicos (Palkeet, 2024^[14]).

IA en el diseño e implementación de la regulación

La IA puede contribuir a mejoras significativas en el diseño e implementación de la regulación. Puede permitir a los gobiernos adaptar las estrategias regulatorias para obtener mejores resultados económicos, sociales y ambientales, al tiempo que abordan de manera efectiva los desafíos y preocupaciones relevantes, como la rendición de cuentas, la transparencia y las cargas regulatorias. Para lograrlo, los gobiernos deben alejarse del enfoque tradicional de "regular y olvidar" de la formulación de políticas regulatorias y adoptar un enfoque de "adaptar y aprender" (OECD, 2024^[24])¹⁰. Las tecnologías digitales, incluida la IA, desempeñan un papel fundamental en este cambio, no solo al ampliar el conjunto de herramientas regulatorias, sino también al permitir una elaboración de normas más innovadora, efectiva y eficiente mediante el diseño, la toma de decisiones y la aplicación de la ley basados en datos (OECD, 2021^[25]).

Situación actual

En lo que respecta a la mejora del *diseño* normativo, la IA ofrece a los responsables de la formulación de políticas una serie de oportunidades significativas. Los sistemas regulatorios pueden ser complejos, con miles de leyes en numerosos ministerios y agencias competentes. La IA puede ayudar a navegar por un conjunto existente de regulaciones y puede analizar conjuntos de datos vastos y complejos para identificar brechas, solapamientos y patrones en los marcos regulatorios. Esto puede permitir tomar decisiones de diseño más informadas y específicas para mejorar los resultados. La IA también puede automatizar tareas rutinarias, lo que puede generar eficiencias al agilizar procesos como el análisis de políticas, Análisis de Impacto Regulatorio (RIA, por sus siglas en inglés) y redacción de textos legales. Puede fortalecer el involucramiento de las partes interesadas al facilitar entendimientos regulatorios simplificados y analizar las consultas públicas. La IA también puede permitir el análisis anticipatorio y la experimentación que pueden informar la gestión de riesgos, mejorando el diseño de marcos regulatorios listos para el futuro.

En cuanto a la *implementación* regulatoria, la IA puede ayudar a mejorar el rendimiento de los organismos de prestación de servicios. Por ejemplo, pueden utilizar algoritmos de IA para optimizar los recursos de inspección, mejorar la seguridad y reducir las cargas regulatorias para las empresas. Los reguladores han estado utilizando la IA para mejorar las actividades y los procesos de estas instituciones con el fin de proteger mejor los intereses públicos y la eficiencia de los recursos (OECD, 2021^[25]). Los enfoques específicos incluyen:

- Mejorar los enfoques basados en el riesgo con análisis de datos que ayuden a aumentar la precisión y los parámetros de las evaluaciones de riesgo para mejorar la forma en que se dirigen las inspecciones.
- Mejorar los esfuerzos para monitorear y detectar el incumplimiento, como en las plataformas de redes sociales, lo que mejora la eficiencia de la supervisión de las entidades regulatorias.
- Potenciar las diversas actividades de los reguladores económicos para entender y supervisar mejor sus sectores y mercados.

Redacción de regulaciones y documentos conexos

Los sistemas de IA que utilizan NLP, como los LLM, ofrecen grandes oportunidades para simplificar y mejorar los aspectos escritos de la gobernanza regulatoria. Los gobiernos pueden utilizar estos sistemas para elaborar borradores a partir de plantillas y normativas existentes, garantizando así el cumplimiento de las normas y ahorrando tiempo y recursos. En Estados Unidos (EE.UU.), por ejemplo, la legislatura estatal de California se convirtió en la primera del país en utilizar la IA para redactar una resolución en 2023 (Tribune News Service, 2023^[26]), Costa Rica siguió de cerca este tema y utilizó ChatGPT para redactar los proyectos de ley para regular la IA (Guio and Müller-Daubermann, 2024^[27]). La IA también puede utilizarse para cruzar nuevos proyectos de ley con la legislación vigente, identificando conflictos y reduciendo los errores humanos, lo que da lugar a una normativa más sencilla, coherente y accesible (Recuadro 5.12).

Recuadro 5.12. Uso de la IA en la redacción de leyes y respuestas en el Reino Unido

La incubadora de inteligencia artificial del Gobierno del Reino Unido (i.AI) ha desarrollado dos sistemas basados en IA, Lex y Parlex, para mejorar la redacción de leyes y la formulación de políticas.

Lex: este sistema tiene como objetivo mejorar el proceso de redacción legislativa proporcionando herramientas avanzadas de IA para navegar, explicar e interrogar la legislación del Reino Unido. Entre las características clave se incluyen las capacidades de búsqueda semántica que permiten a los usuarios localizar materiales legislativos relevantes para el contexto de manera eficiente, y las herramientas de redacción asistidas por IA que generan notas explicativas para los proyectos de ley del gobierno, lo que reduce el esfuerzo manual y aumenta la precisión en el lenguaje jurídico. El sistema también se diseñó con un profundo conocimiento de los términos jurídicos específicos del Reino Unido, lo que le permitió capturar con precisión los matices de la terminología jurídica, promoviendo así la innovación y la colaboración en el sector jurídico.

Parlex: este sistema de IA está diseñado para ayudar a los responsables de la formulación de políticas mediante el pronóstico de las reacciones parlamentarias a las políticas propuestas. Mediante el análisis de los registros parlamentarios históricos, Parlex proporciona información sobre cómo los parlamentarios podrían responder a las nuevas políticas, lo que permite a los funcionarios desarrollar estrategias efectivas para la implementación de políticas. Por ejemplo, puede realizar una "verificación de vibración parlamentaria" de posibles leyes, prediciendo el apoyo o la oposición entre los parlamentarios basándose en debates pasados. Esta herramienta ayuda a los responsables de la formulación de políticas a comprender el clima político y a anticipar los desafíos o el apoyo a las políticas antes de formular propuestas formales.

Tanto Lex como Parlex ejemplifican el compromiso del Reino Unido de integrar la IA en los servicios públicos, con el objetivo de mejorar la eficiencia, la precisión y la planificación estratégica en el desarrollo legislativo y político.

Fuente: <https://ai.gov.uk/projects/parlex-and-lex>, <https://www.thetimes.com/uk/politics/article/parlex-ai-to-advise-ministers-on-how-policies-will-be-received-99txwlwph>.

Agilización de los análisis regulatorios

El Análisis de Impacto Regulatorio (RIA) y la evaluación ex post son elementos fundamentales de una buena gobernanza regulatoria. No deben considerarse como requisitos discretos que han de aplicarse sucesivamente, sino como herramientas mutuamente complementarias integradas en el ciclo de políticas

para fundamentar la adaptación adecuada de los enfoques regulatorios (o alternativos) (OECD, 2024^[24]; 2021^[28]).

Los reguladores pueden utilizar la IA para mejorar cada uno de estos elementos, al tiempo que utilizan la IA para mejorar la velocidad y la precisión de la evaluación y minimizar la carga que suponen dichas actividades. Esto puede permitir una evaluación más frecuente de la regulación para ayudar a crear regulaciones adaptables y adaptables al futuro que sigan siendo relevantes y efectivas en un entorno rápidamente cambiante. Este uso de la IA puede ser especialmente necesario para garantizar una gobernanza adecuada de la IA. Los expertos en IA señalan como riesgo principal que los mecanismos e instituciones de gobernanza no puedan mantenerse al día con el rápido ritmo de los avances tecnológicos (OECD, 2024^[29]).

Como paso inicial, la IA puede ayudar a estimar la carga que representan las posibles nuevas regulaciones y otras políticas, lo que puede influir en la facilidad con que las empresas pueden adaptarse a los cambios regulatorios y en la probabilidad de que cumplan. Naturalmente, esto afecta al éxito de cualquier medida regulatoria. También es importante que los reguladores consideren la posibilidad de comparar los diferentes enfoques de gobernanza y sus correspondientes compensaciones costo-beneficio para una formulación de políticas más eficiente y económicamente racional (Recuadro 5.13).

Recuadro 5.13. Desarrollo de IA para medir los costos de cumplimiento en Alemania

En Alemania, el Centro de Servicios para la Mejoramiento de la Legislación de la Oficina Federal de Estadística está desarrollando un sistema de ML para ayudar a estimar los costos de cumplimiento con el fin de respaldar el impacto regulatorio. Este enfoque implica identificar partes de borradores legales que influyen en los costos de cumplimiento, utilizando la IA para la extracción de datos web de textos legales. Después utiliza la IA para predecir como el nuevo texto jurídico modifica los costos de cumplimiento y estimar si estos costos son bajos o altos. Si los costos son probablemente bajos, la oficina utilizará la IA para calcular los costos de cumplimiento. Pero si los costos probablemente sean altos, la estimación será hecha manualmente por humanos para ayudar a garantizar la exactitud. Este proceso debería permitir a la Oficina centrarse en estimaciones de gran esfuerzo y liberar recursos para otros proyectos. Sin embargo, sigue habiendo desafíos relacionados con la estructura de los datos extraídos, la comprensibilidad de los textos jurídicos alemanes, la calidad de los datos, la explicabilidad de las variables utilizadas en el sistema y la correspondencia entre las fuentes de datos.

Fuente: (Walprecht and Lewerenz, 2024^[30]).

La IA también puede mejorar la evaluación de políticas y apoyar una toma de decisiones más informada e iterativa al permitir una sofisticada experimentación y evaluación de políticas. Mediante la simulación de escenarios regulatorios del futuro, la IA permite a los responsables de la formulación de políticas modelar y predecir los posibles impactos de diferentes opciones regulatorias (OECD, 2025^[31]). Esto ayuda a realizar evaluaciones de impacto normativo, ya que ayuda a los responsables de la toma de decisiones a comprender las consecuencias de las regulaciones propuestas para los diferentes sectores y partes interesadas. Puede ayudar a prever de forma preventiva la efectividad de posibles políticas, al descubrir patrones que pueden no resultar evidentes de inmediato a través de los métodos de análisis tradicionales.

Por ejemplo, la Universidad de Dublín lideró el desarrollo de la Simulación de Políticas de Innovación para la Economía Inteligente (IPSE, por sus siglas en inglés), una herramienta que simula el efecto de instrumentos de política basados en perfiles regionales e información sectorial. El objetivo es comprender mejor los posibles factores impulsores de la innovación y su impacto antes de que se pongan en marcha los instrumentos de política (Nesta, 2024^[32]). Del mismo modo, PolicyEngine es una herramienta que

utiliza legislación lista para la digitalización con el fin de fundamentar los cambios en las políticas y los modelos para ver cómo afectarían los cambios a los gobiernos y la ciudadanía. Los usuarios pueden seleccionar su país (Canadá, Reino Unido, Estados Unidos y Nigeria están actualmente disponibles), el ámbito de actuación pertinente y los parámetros de actuación que van a modificarse, así como calcular el impacto económico y presupuestario para ver cómo afectan los cambios a los ingresos públicos, por ejemplo. PolicyEngine ha integrado ChatGPT en su sistema para ofrecer un análisis más detallado y explicaciones a los usuarios (Martin, 2023^[33]).

Promoción de la participación de las partes interesadas en el diseño regulatorio

Si bien los gobiernos están mejorando lentamente sus actividades de participación de las partes interesadas en materia de gobernanza regulatoria, la mayoría de los países de la OCDE tienen un amplio margen para mejorar estos esfuerzos (OECD, 2023^[34]). La IA puede mejorar la eficiencia y la efectividad de los procesos al involucrar a las partes interesadas en el diseño de las políticas. Para ello, puede utilizar análisis avanzados e interfaces de usuario inteligentes, contribuir a un proceso de diseño regulatorio más inclusivo, mejorar la capacidad de respuesta de los gobiernos y aumentar la transparencia y la confianza en el gobierno. Un número cada vez mayor de chatbots de IA están facilitando las consultas públicas sobre regulaciones nuevas o revisadas al interactuar con muchas partes interesadas de forma simultánea, y sintetizando datos para que los gobiernos ajusten aún más sus propuestas regulatorias. Los chatbots también pueden responder instantáneamente a las consultas de las partes interesadas, guiarlas a través del proceso de consulta y simplificación de textos legislativos complejos para mejorar la comprensión del público. Esto hace que el proceso de consulta regulatoria sea más accesible para las partes interesadas y reduce las cargas para los reguladores en lo que respecta al diseño de políticas participativas. La sección "Participación cívica y gobierno abierto" de este capítulo analiza en mayor medida cómo se está utilizando la IA para la participación de las partes interesadas y la participación pública.

Mejoramiento de las funciones de los reguladores económicos

Los reguladores económicos se encargan de diversas funciones en las que puede aplicarse la IA, como la fijación de tarifas, la autorización de actividades, la tramitación de reclamaciones, la mediación en conflictos, la supervisión de los mercados y la realización de inspecciones y actividades de aplicación de la ley. Más allá de las funciones regulatorias esenciales, los reguladores económicos llevan a cabo muchas actividades para ser efectivos en su trabajo, como la investigación para comprender mejor los sectores y mercados que supervisan.

Si bien se están realizando algunos esfuerzos, los reguladores se encuentran en las primeras etapas del uso de la IA, y pocos exploran o prueban su uso¹¹. Según una encuesta del Organismo de Reguladores Europeos de Comunicaciones Electrónicas (BEREC, por sus siglas en inglés), "la adopción de IA por parte de las agencias reguladoras nacionales está todavía en sus inicios" y muy pocos reguladores han "realizado o tienen previsto realizar estudios para explorar formas de adoptar la IA" (BEREC, 2023^[35]). No obstante, los usos actuales y potenciales de la IA en los enfoques regulatorios reflejan una tendencia más amplia hacia un mayor uso de los datos y las tecnologías digitales. Un análisis llevado a cabo por la OCDE (OECD, 2025^[31]; 2020^[36]) muestra que las herramientas digitales ayudan a los reguladores al permitir un mejor uso de los datos y una toma de decisiones informada. Ante la presión de hacer más con menos, los reguladores recurren cada vez más a la IA para hacer que los procesos sean más eficientes y efectivos.

Los recientes debates entre la Red de Reguladores Económicos (NER)¹² de la OCDE y los ejemplos de la literatura reciente revelan que la IA ya se utiliza en las inspecciones, el monitoreo del mercado y en áreas de cara al consumidor, como la tramitación de quejas o la respuesta a consultas de los consumidores. Por ejemplo¹³:

- Los reguladores de las comunicaciones electrónicas de la UE utilizan la IA para modelar las cadenas de radio y optimizar el uso compartido del espectro, detectar contenidos ilegales o prohibidos en línea, gestionar las relaciones con los clientes, por ejemplo, clasificando las reclamaciones, y medir a gran escala las experiencias en línea de la ciudadanía y los comportamientos de las plataformas. Esto puede ir acompañado de una supervisión del mercado para garantizar que los productos que se venden en línea cumplen las normas de seguridad de los productos (Recuadro 5.14) (BEREC, 2023^[35]; Faculty, 2021^[37]).
- El Regulador de Energía de Austria (E-control) está desarrollando una aplicación de IA para ayudar a los consumidores a entender sus facturas de energía y un chatbot para responder a las consultas de los consumidores.
- El Regulador del Agua del Perú (Sunass) (2024^[38]) está aplicando IA en la elaboración de informes de inspección. La aplicación automatiza la generación de informes en función de las variables registradas por los inspectores en las tablas, lo que simplifica significativamente el proceso y reduce el tiempo dedicado a la redacción de informes. Los informes son validados por los especialistas para garantizar su exactitud. Sunass también ha desarrollado una herramienta que utiliza análisis geoespaciales y un algoritmo de clasificación para calcular las necesidades de inversión y las brechas en el sector hídrico del Perú.
- La Autoridad del Mercado de Capitales Israelí ha comenzado a desarrollar una herramienta impulsada por IA que agrega información de sitios web sobre seguros y ahorros. El proyecto integra modelos financieros avanzados, ML y NLP para permitir una identificación más temprana y precisa de los riesgos, anomalías y casos sospechosos de fraude en los mercados de capitales. Además, proporciona herramientas para aumentar la transparencia para el público en general y los inversores. El sistema, que se encuentra actualmente en la fase de prueba de concepto, ofrece herramientas visuales para los modelos de formación, la supervisión del rendimiento, la toma de decisiones y la generación de conocimientos prácticos para los empleados de la Autoridad¹⁴.
- La Agencia Nacional de Transporte Terrestre (ANTT) de Brasil utiliza la IA en su supervisión de las infraestructuras de transporte para respaldar el cumplimiento efectivo de su mandato como regulador económico en el sector. El Sistema de Información Vial combina datos sobre aspectos como accidentes, asistencia en carretera, posibles infractores, peajes, cámaras de velocidad y sensores de tráfico en 26 concesionarios. El sistema registra 15 000 entradas por segundo con datos en tiempo real y combina herramientas de IA con una interfaz humana con un equipo las 24 horas del día, los 7 días de la semana. La información permite a la ANTT supervisar las actividades del sector del transporte por carretera de una forma más efectiva y respalda la toma de decisiones regulatorias basadas en datos por parte de la ANTT y otros agentes públicos.

Recuadro 5.14. La herramienta SAFE AI de Dinamarca

Dinamarca desarrolló en 2021 una herramienta de IA para rastrear productos peligrosos en Internet en 16 países europeos, automatizando así un proceso que anteriormente requería búsquedas manuales. La herramienta, denominada SAFE, fue desarrollada por la Autoridad Danesa de Tecnología de la Seguridad en cooperación con una empresa privada de TI para automatizar el trabajo que exige muchos recursos y mejorar la seguridad de los consumidores.

La herramienta SAFE utiliza el reconocimiento de imágenes y textos para buscar en la web productos peligrosos o deficientes, utilizando información del Safety Gate, el sistema europeo de alerta rápida para productos peligrosos (RAPEX) y el sistema de información y comunicación sobre vigilancia del mercado. La herramienta recibe formación continua a través de los comentarios de los usuarios para

mejorar la precisión de sus conclusiones. Las conclusiones de SAFE pueden utilizarse para advertir a las autoridades de cualquier producto no conforme que se comercialice en sus mercados.

SAFE es la continuación de una herramienta de IA anterior de la autoridad, llamada AIME, desarrollada para el mercado danés en 2020. La autoridad recibió fondos de la Comisión Europea para desarrollar una herramienta similar para su uso en toda la UE, lo que condujo a la creación de la herramienta SAFE.

Fuente: <https://www.sik.dk/en/business/nyheder/new-digital-fine-toothed-comb-denmark-about-make-more-300-million-european-consumers-safer>.

Refinación de validación de criterios de riesgo

Una implementación efectiva de la regulación depende de una comprensión precisa de los riesgos. Las autoridades públicas deben basar sus actividades de inspección y aplicación de la ley en criterios de riesgo. Por lo tanto, es fundamental supervisar la evolución de los riesgos en la vida real, permitir respuestas adaptativas a los posibles cambios y actualizar periódicamente los criterios de riesgo.

Las técnicas de ML pueden calcular la precisión de las evaluaciones de riesgos y perfeccionar los parámetros (Recuadro 5.15). Las aplicaciones recientes de ML son prometedoras a la hora de identificar los principales predictores de riesgo, mejorando considerablemente la efectividad de la focalización basada en el riesgo (OECD, 2021^[28]). A medida que los algoritmos de ML evolucionan con nuevos datos, deben actualizarse periódicamente para seguir siendo confiables (Cary Coglianese, 2024^[39]). Si bien estas herramientas tienen un potencial considerable, los reguladores deben usar su discreción al utilizar aplicaciones de IA para predecir el cumplimiento normativo y garantizar así la exactitud y evitar sesgos.

Recuadro 5.15. Criterios de riesgo para las contribuciones agrícolas en la Provincia Autónoma de Trento

En Italia, la Provincia Autónoma de Trento trabajó con la OCDE para analizar los parámetros de riesgo utilizados por la Agencia de Pagos Agrícolas (APPAG, Agenzia Provinciale per i Pagamenti), proponer revisiones e introducir prácticas estándar sólidas de ML. La APPAG paga contribuciones agrícolas relacionadas con la gestión de tierras agrícolas, como la siega manual o el uso limitado de pesticidas. Las contribuciones son proporcionales a la superficie de la tierra de acuerdo con la solicitud presentada por el agricultor. Debido a que los recursos de supervisión son limitados, la selección de sitios para inspeccionar con la mayor precisión posible es particularmente importante. Los criterios de riesgo de la APPAG para planificar las inspecciones pertinentes se validaron aplicando técnicas de ML. Un algoritmo capaz de predecir las solicitudes de contribución agrícola de mayor riesgo permitió la revisión de los parámetros existentes, haciéndolos significativamente más efectivos en la focalización de riesgos (solicitudes de alto riesgo) y la identificación de situaciones de incumplimiento.

Fuente: (OECD/EU, 2024^[40]).

Mejoramiento de la modelización de riesgos para orientar mejor las inspecciones

Los reguladores pueden utilizar la IA para mejorar la modelización de riesgos y orientar mejor las inspecciones. Mediante un enfoque basado en el riesgo, los reguladores destinan sus recursos a actividades que suponen una amenaza para los bienes públicos y, en términos más generales, para la consecución de los objetivos deseados. La mejora de la eficiencia de las inspecciones y la aplicación de

la normativa depende de que se comprendan con precisión los riesgos para el interés público (OECD, 2014^[41]). Es necesario establecer y evaluar criterios de riesgo para medir el nivel de riesgo planteado por los operadores privados y orientar los esfuerzos de aplicación de la normativa. Esto protege mejor los bienes públicos, garantiza el uso eficiente de los recursos y fomenta una relación basada en la confianza con las empresas (OECD, 2018^[42]; Blanc, 2018^[43]). Véase Recuadro 5.16 como ejemplo de este enfoque.

Recuadro 5.16. La Toscana y la OCDE elaboran criterios de riesgo para las solicitudes de subvención

La aplicación de técnicas de IA por parte de los reguladores ayuda a mejorar el conocimiento con respecto de qué características de las empresas pueden ser predictores efectivos de riesgo. Esto podría mejorar considerablemente la focalización basada en el riesgo. Mediante el uso de herramientas de IA y análisis de datos avanzados, las autoridades regulatorias pueden adaptar sus estrategias de supervisión, centrando los recursos en las áreas más críticas y minimizando los riesgos, garantizando al mismo tiempo que sus acciones sean eficientes y tengan impacto.

La región italiana de la Toscana aplicó una metodología basada en el riesgo a los controles documentales de las solicitudes de financiamiento de actividades económicas que operan en la región, en particular de subvenciones europeas para incentivar la innovación. La OCDE ayudó a desarrollar parámetros de riesgo que orientan el control de estas solicitudes de incentivos. Mediante el análisis de la relación entre las características de las pequeñas y medianas empresas (PYMES) que solicitan financiamiento público y una adecuada estimación del riesgo potencial, se construyeron sistemas predictivos de ML. Dependiendo del tipo de solicitud, las administraciones públicas podrían utilizar estos múltiples sistemas para orientar, mejorar y acelerar la toma de decisiones sobre subvenciones y solicitudes de financiamiento.

Estas herramientas y algoritmos podrían servir de base para clasificar futuras aplicaciones, una vez actualizados los sistemas con los datos pertinentes y a medida que los conjuntos de datos crezcan con el tiempo. El uso continuo de estas herramientas podría poner de relieve la recaída asociada al incumplimiento malicioso o la mejora del comportamiento de los establecimientos a medida que mejoran en la preparación de sus aplicaciones.

Fuente: (OECD/EU, 2024^[40]).

La modelización de riesgos basada en la IA también puede mejorarse a través de los datos recopilados en las redes sociales. Los resultados de las inspecciones históricas y las bases de datos ya sirven para orientar mejor las futuras inspecciones y las estrategias preventivas. Las autoridades de cumplimiento de la regulación también han empezado a adquirir y utilizar los datos disponibles en las plataformas de redes sociales para detectar posibles incumplimientos. Mientras que las fuentes tradicionales de información dependen de las autoridades y los inspectores, las redes sociales permiten el acceso directo a la ciudadanía expuesta a riesgos. Esto contribuye a fundamentar aún más los enfoques basados en el riesgo y a garantizar que las decisiones estén orientadas a los resultados.

Las inspecciones deben guiarse por la evaluación de riesgos, y las quejas públicas a menudo ponen de manifiesto los riesgos emergentes. Las mejores inspecciones comienzan con decisiones informadas y coherentes. Las denuncias deben utilizarse para mejorar la planificación de las inspecciones basada en el riesgo y, sólo en unos pocos casos, deben dar lugar a inspecciones sin demora. Por lo tanto, un sistema de gestión de reclamaciones basado en el riesgo es esencial para garantizar el equilibrio adecuado entre las inspecciones proactivas —que se producen después de una cuidadosa planificación basada en el

riesgo— y las inspecciones reactivas —que no se planifican y se producen ad hoc en respuesta a las reclamaciones consideradas graves—.

Las denuncias oficiales se suelen presentar a través de aplicaciones específicas. Sin embargo, el público puede mostrarse reacio a utilizar estas iniciativas o simplemente ignorar que esto puede ayudar a hacer un seguimiento de los problemas de salud y seguridad. Por lo tanto, las autoridades reguladoras podrían utilizar plataformas de redes sociales y ML para apoyar el análisis de riesgos y obtener información más amplia y oportuna que con los métodos tradicionales (OECD, 2021^[28]). La OCDE apoyó una iniciativa en la región italiana del Lacio para poner a prueba las quejas en las redes sociales como fuente de los CMS inspirados en iniciativas internacionales (Recuadro 5.17).

Recuadro 5.17. Mejoramiento de los criterios de riesgo a través de reclamos de clientes en la región de Lacio

Las autoridades regulatorias deben utilizar la evaluación de riesgos para orientar sus estrategias de aplicación de la ley con el fin de evitar o abordar los impactos negativos potenciales más probables o más graves del incumplimiento. Las reclamaciones de la ciudadanía disponibles en Internet podrían ser una fuente esencial de información y desempeñar un papel crucial a la hora de hacer cumplir la normativa vigente. Al mejorar sus herramientas para recopilar y analizar las opiniones del público, las autoridades pueden mejorar la identificación de riesgos y la planificación de inspecciones.

Cuando la ciudadanía no utilizan los canales oficiales para dar su opinión, incluidas las quejas por riesgos, es probable que lo hagan a través de las redes sociales y otros sitios web. Se recogieron dos millones de opiniones sobre restaurantes de TripAdvisor y se seleccionó una muestra de 5 000 comentarios. Cada revisión de esta muestra se categorizó manualmente en función de la presencia de problemas de higiene o intoxicación alimentaria. Este conjunto de datos preclasificados se utilizó para entrenar un algoritmo de ML, un enfoque basado en una memoria bidireccional de largo plazo (LSTM). Este algoritmo LSTM se aplicó para analizar nuevas revisiones en el sitio web y clasificarlas, identificando revisiones negativas (con problemas de higiene). Incluso con los limitados datos preclasificados utilizados para entrenar el algoritmo, el rendimiento del sistema osciló entre el 81 % y el 83 %, lo que demuestra su eficiencia.

Fuente: (OECD/EU, 2024^[40]).

Mejoramiento de la identificación de incumplimientos

Un registro de datos preciso debería sustentar una supervisión del cumplimiento normativo más inteligente y la focalización de las situaciones de incumplimiento. Cuando los inspectores se enfrentan al desafío de aumentar el número de datos potencialmente sospechosos, las soluciones de ML pueden ayudar a identificar los datos no confiables y detectar el incumplimiento.

Las herramientas innovadoras agilizan la presentación de datos, detectan anomalías y analizan patrones de cumplimiento normativo (Recuadro 5.18), mejorando la eficiencia de la supervisión por parte de las autoridades reguladoras.

Recuadro 5.18. Buena reputación de aprobaciones y cumplimiento normativo en Israel

En 2024, la Unidad de Fideicomiso del Registro de Organizaciones Benéficas del Ministerio de Justicia de la Autoridad de Corporaciones de Israel introdujo un proceso automatizado basado en IA en su sistema central de aprobación regulatoria para emitir aprobaciones de "Good Standing" (gestión adecuada). Esta aprobación es vital para más de 23 000 organizaciones sin fines de lucro al año, ya que sirve como requisito previo para los reembolsos de impuestos de los donantes y los criterios de elegibilidad para los beneficios públicos y las oportunidades de contratación.

El nuevo sistema automatizado basado en IA superó importantes desafíos en el reconocimiento óptico de caracteres (OCR, por sus siglas en inglés) y la detección de objetos para identificar el cumplimiento normativo de las normas de gestión adecuadas y detectar signos sospechosos de corrupción o gestión incorrecta. Ahora los supervisores pueden centrarse en los casos detectados, mientras que el tiempo medio de respuesta se ha reducido de 45 días a una hora, lo que supone un nuevo nivel de servicio público.

El proyecto fue una de las nueve iniciativas ganadoras de una convocatoria abierta de implementación de la IA anunciada por el Ministerio de Innovación, Ciencia y Tecnología, en colaboración con la Agencia Nacional Digital de Israel.

Fuente: Funcionarios del Gobierno de Israel, https://www.gov.il/en/pages/most_ai_government_agencies_open_call_winners.

Gestión de riesgos y desafíos

Riesgos asociados

- Datos inadecuados o sesgados en los sistemas de IA.
- Falta de transparencia y explicabilidad.

Si los sistemas de IA se basan en **datos inadecuados o sesgados**, podrían producirse resultados inexactos o adversos para algunas personas o grupos. Con respecto a la GFP, esto podría dar lugar, por ejemplo, a resultados regulatorios negativos cuando algunas personas o grupos se seleccionan indebidamente para acciones de aplicación de la ley.

Los sistemas avanzados de IA suelen tomar decisiones dentro de una caja negra, a menudo sin que ni siquiera los operadores del sistema comprendan cómo llegaron a la decisión (Valderrama, Hermosilla and Garrido, 2023^[44]; OECD, 2024^[29]). Este **riesgo de explicabilidad limitada** requiere una supervisión y evaluación humanas del sistema de IA y sus resultados para garantizar una toma de decisiones regulatorias transparente y responsable. Si la IA se utiliza para el diseño o la entrega de la regulación, como la redacción de textos, la realización de evaluaciones y la participación de las partes interesadas, se debe prestar especial atención a: las entradas y salidas de dicha aplicación; la calidad de los datos utilizados; la precisión y confiabilidad de las salidas de la IA; la explicabilidad de tales salidas; la transparencia del uso de IA en la toma de decisiones; y la rendición de cuentas por los impactos asociados en el diseño e implementación de la regulación.

Este riesgo puede dificultar la **promoción de la rendición de cuentas y la creación de confianza** en la capacidad de los gobiernos para utilizar la IA con el fin de mejorar la elaboración de normas. En general, los sistemas de IA de los gobiernos deben ser responsables y auditables, lo que ayuda a reforzar el principio de rendición de cuentas de la OCDE en materia de IA. Cuando sea factible y apropiado, los gobiernos deben dar prioridad a que los sistemas de IA sean abiertos y transparentes para fomentar la

confianza pública y permitir el escrutinio y la validación externos lo que puede incluir hacer públicos los datos, desarrollar algoritmos de código abierto y garantizar procesos transparentes de toma de decisiones para impulsar la confianza en las decisiones respaldadas por la IA. Además, deben existir estructuras claras para garantizar unos mecanismos de rendición de cuentas y supervisión adecuados, teniendo en cuenta quién es responsable de cada elemento de los resultados del sistema de IA y quién es responsable de la calidad o la revisión de los resultados de toda la iniciativa de IA.

Desafíos de implementación

- Entornos legales y regulatorios inflexibles u obsoletos.
- Falta de datos de alta calidad y capacidad para compartirlos.
- Brechas de habilidades.

Si bien la IA ofrece muchas oportunidades de gobernanza regulatoria adaptable, los cambios frecuentes en el diseño regulatorio pueden perturbar tanto a las empresas como al público. **Los ajustes basados en el análisis continuo de datos pueden dar lugar a un entorno normativo volátil**, lo que dificulta a las empresas la planificación de estrategias a largo plazo y a que el público se mantenga informado sobre la legislación vigente. Además, los responsables de la formulación de políticas y los organismos reguladores pueden enfrentarse a **desafíos en cuanto al acceso, la recopilación y el procesamiento de datos**, lo que limita el alcance en el que los sistemas de IA llevarán a cabo análisis confiables y recomendaciones válidas. Por ejemplo, Australia cuenta con una Ley de Disponibilidad y Transparencia de los Datos que proporciona una base legal para compartir los datos del Gobierno australiano. Sin embargo, en muchos casos el intercambio de datos se ve obstaculizado por la falta de protocolos o de sistemas incompatibles (Productivity Committee, 2024^[45]). La mejora de las capacidades de gestión de datos es fundamental para cualquier uso de la IA que esté evaluando e informando la base empírica para el diseño o la entrega regulatoria.

Dado que la IA en la gobernanza regulatoria sigue requiriendo la intervención humana, la falta de conocimientos técnicos puede dar lugar a malos resultados y a un uso indebido de la IA. Los datos recopilados por la OCDE durante el debate con sus miembros muestran que los reguladores tienen dificultades para igualar la experiencia de las empresas tecnológicas debido al alto costo y la escasez de competencias digitales. Para cerrar esta brecha y mejorar la efectividad regulatoria, **los gobiernos deben invertir en el desarrollo de competencias digitales** y promover la colaboración con el sector tecnológico, por ejemplo, a través de alianzas y contrataciones públicas para adquirir competencias y capacidades si aún no están presentes.

Potencial sin explotar y camino a seguir

La IA en el diseño e implementación de la gobernanza regulatoria todavía se encuentra en sus etapas iniciales en comparación con las aplicaciones comerciales. Sin embargo, existe un mayor margen para que la IA se utilice en la política regulatoria —con una serie de temas clave derivados de estudios de casos que los gobiernos pueden considerar— para avanzar en la madurez de la IA en el diseño y la entrega de la regulación. En cuanto al diseño normativo:

- La IA puede **aprovecharse aún más para diseñar regulaciones**, ya que actualmente se invierte poco en el alcance de su aplicación potencial. Aunque los gobiernos ya utilizan la IA, muchas de estas aplicaciones se centran en la toma de decisiones operativas, las medidas de cumplimiento, los procesos internos generales y la implementación de servicios o productos públicos. Sin embargo, estas aplicaciones de IA se pueden adaptar fácilmente al diseño normativo. Por ejemplo, los chatbots utilizados en la prestación de servicios públicos o en las revisiones de solicitudes y subvenciones pueden reutilizarse para recabar la opinión de las partes interesadas y analizar la

legislación para su revisión reglamentaria. La IA utilizada para la prestación de servicios y el seguimiento de las actividades del mercado también puede utilizarse para monitorear de forma continua los impactos de las regulaciones existentes, proporcionando información en tiempo real y permitiendo ajustes oportunos por parte de los reguladores.

- **Una legislación más digital permite una aplicación más amplia de IA** a lo largo de todo el ciclo de formulación de políticas. La IA puede utilizarse para simplificar la legislación, lograr que el lenguaje sea neutro desde el punto de vista tecnológico y mejorar el diseño para el procesamiento automatizado de casos. Si la legislación puede administrarse parcial o totalmente de forma digital, la IA puede utilizarse para respaldar la aplicación, el cumplimiento y la revisión ex post, así como para generar datos que ayuden a fundamentar mejor el diseño y la experimentación de la política regulatoria.
- **El uso de la IA para anticipar escenarios y riesgos futuros con miras a un diseño normativo más informado.** La IA puede ofrecer información valiosa sobre previsiones, ofreciendo a los gobiernos la posibilidad de ver las tendencias emergentes y los cambios en diversos sectores para planificar proactivamente las respuestas normativas. Por ejemplo, la IA puede prever el crecimiento de las tecnologías o las aplicaciones tecnológicas en el sector sanitario, permitiendo a los reguladores probar los marcos existentes o desarrollar otros nuevos para garantizar la seguridad y la efectividad antes de su adopción generalizada. Esto puede ayudar a aumentar la confianza en el gobierno, ya que los gobiernos asumen un papel proactivo, en lugar de reactivo, en la protección de la ciudadanía (OECD, 2024^[46]).

En cuanto a otras áreas de política pública para avanzar en el uso de la IA en el diseño regulatorio, los gobiernos deben abordar las incertidumbres legales y de gobernanza. Es posible que muchos gobiernos no dispongan de las estructuras y marcos jurídicos adecuados para desplegar con confianza la IA en el ámbito de la gobernanza regulatoria. Esto puede deberse, por ejemplo, a ambigüedades en el texto jurídico sobre el cumplimiento y la rendición de cuentas, que requieren un juicio jurídico. Aunque los países están tomando medidas para aclarar el uso de la IA en la administración, se necesitan directrices más maduras sobre el uso de la IA en el diseño normativo. Una gobernanza de la IA más sólida es crucial, no solo para los sistemas regulatorios, sino para todas las aplicaciones gubernamentales, con el fin de garantizar un despliegue responsable y confiable de la IA.

Las autoridades públicas también deben adaptar los mecanismos de cumplimiento de la regulación para brindar una buena protección a los bienes públicos y a la ciudadanía de manera efectiva en el contexto de la globalización y la tecnología, aprovechando estos cambios para lograr mejores resultados. Los pasos adelante incluyen garantizar que los reguladores tengan los mandatos, poderes, funciones y mecanismos de rendición de cuentas necesarios para utilizar la IA, y proporcionar acceso a datos de alta calidad, precisos y seguros para protegerse contra las ciberamenazas. A la hora de implementar la IA para la prestación de servicios regulatorios, nos enfrentamos a tres áreas principales de atención:

- **Otorgando a los reguladores las facultades adecuadas de recopilación de datos.** Para recopilar e interpretar de manera efectiva a los datos procedentes de diversas fuentes, los reguladores necesitan una legislación sólida. La regulación de los mercados basados en datos requiere un conjunto de herramientas adecuadas, de modo que los reguladores puedan solicitar y recibir suficiente información en tiempo real y en la forma pertinente (OECD, 2020^[36]). Dado que las aplicaciones de la IA suelen basarse en la recopilación masiva de datos, las autoridades reguladoras deben estar legalmente facultadas para recopilarlos, procesarlos y publicarlos cuando proceda, respetando al mismo tiempo los principios de privacidad y protección de datos y garantizando que las condiciones jurídicas e institucionales favorezcan el acceso y el intercambio de datos en aras del interés público (OECD, 2021^[47]). Los reguladores deben colaborar con los gobiernos y parlamentos nacionales para garantizar un marco jurídico adecuado que defina sus competencias y facultades.

- **Mejoramiento de las competencias y conocimientos de los reguladores en materia de IA.** Para aprovechar el potencial de las técnicas de ML y el análisis de macrodatos, los reguladores deben aumentar su experiencia en IA contratando a científicos de datos y expertos en ciber riesgos. Atraer y retener talento en materia de datos en medio de la competencia del sector privado es un desafío para los reguladores y la administración pública en general. Además, todos los responsables de la formulación de políticas y organismos reguladores deberían invertir en el desarrollo de capacidades para reforzar la confianza de los responsables de la formulación de políticas en las aplicaciones de la IA. Estos esfuerzos se realizan tanto a nivel nacional como internacional. Por ejemplo, se pueden desarrollar nuevos foros de intercambio y cooperación internacional para que los funcionarios compartan sus conocimientos y experiencias, como la Red Internacional de Cooperación en materia de Regulación Digital (INDRC) establecida para fomentar el debate entre los reguladores (DRCF, 2023^[48]).
- **Garantizar una sólida gobernanza de datos y estrategias que sustenten el uso de IA.** La transición digital ofrece beneficios, pero también plantea nuevos desafíos para la gobernanza de los datos (véase el Capítulo 4, sección "Creación de una base de datos sólida", en el que se ofrece un análisis detallado sobre la gestión, la recopilación, el suministro y el uso de datos para la IA). Es esencial contar con estrategias sólidas para mitigar los riesgos potenciales. Desarrollar una estrategia de datos es una forma global de abordar la gobernanza de los datos. Una encuesta realizada entre los miembros de la *Red de Reguladores Económicos* (NER) de la OCDE en 2024 reveló que el 55 % de los encuestados estaban en proceso de desarrollar una estrategia de datos, mientras que el 29 % ya tenía una en funcionamiento¹⁵. Estos hallazgos ponen de relieve la necesidad de contar con estrategias sólidas de gobernanza y datos para apoyar un uso efectiva de los datos y la IA.

Procesos clave del gobierno

Los procesos clave del gobierno —como la reforma de la función pública, la contratación pública, los esfuerzos anticorrupción, la evaluación de políticas y la participación cívica— son vitales para construir instituciones eficientes, transparentes y responsables. Estos procesos fortalecen la confianza pública, mejoran la prestación de servicios y fomentan una gobernanza basada en evidencia. Los gobiernos están utilizando la IA para garantizar la integridad de la gestión y el gasto de los fondos de los contribuyentes, así como para impulsar el talento del sector público de cara al futuro.

IA en la reforma de la función pública

La IA puede transformar significativamente la forma en que se organizan y gestionan los funcionarios públicos. Al igual que en otras áreas de políticas, las herramientas de IA podrían utilizarse para orientar y personalizar mejor los servicios de recursos humanos (RR. HH.), acelerar y mejorar los procesos de RR. HH. con el fin de reducir las cargas administrativas e impulsar la productividad de los empleados. Algunos gobiernos ya están experimentando con la aplicación de la IA a funciones discretas de RR. HH., como los procesos de contratación, aprendizaje y desarrollo de los empleados (por ejemplo, de formación). Este tipo de aplicación suele centrarse en la automatización de procesos con el objetivo de garantizar procesos de RR. HH. más rápidos y precisos. El potencial de la IA en este campo podría ser mucho mayor. Si se cuenta con los datos adecuados, los sistemas de IA podrían eventualmente ayudar a emparejar mejor a las personas con los puestos de trabajo, y predecir el desempeño en función de los requisitos laborales y las características y antecedentes de las personas. Sin embargo, para que esto funcione es necesario una serie de factores. Por ejemplo, las administraciones públicas tendrían que disponer de datos sólidos en al menos tres ámbitos: las características de su organización, las demandas de puestos de trabajo específicos y los indicadores de rendimiento individual. Lamentablemente, la mayoría de las administraciones públicas carecen de datos confiables en las tres áreas. Para aprovechar este potencial, los gobiernos tendrán que invertir en más y mejores datos de RR. HH., así como en los conjuntos de competencias asociadas a la función de gestión de recursos humanos.

Situación actual

La IA podría aplicarse a una amplia gama de procesos y actividades estratégicos de gestión de recursos humanos, identificando patrones en los datos de la administración pública para orientar una planificación más estratégica de la organización, orientar mejor las políticas de recursos humanos o incluso identificar las competencias necesarias durante emergencias y crisis. Los datos disponibles y los casos prácticos indican dos ámbitos de reforma de la función pública en los que algunos países estaban experimentando activamente con aplicaciones de la IA: la contratación; y el aprendizaje y el desarrollo. En esta sección se analizan y analizan más a fondo.

Hasta la fecha, las iniciativas han sido a menudo proyectos piloto fragmentados, sin un enfoque estratégico para su adopción sistemática. Sin embargo, gobiernos como el francés están adoptando un enfoque coherente y estratégico para explorar los beneficios de la IA en la reforma de la función pública, al tiempo que reconocen los posibles efectos negativos y establecen medidas de seguridad para hacerles frente (Recuadro 5.19).

Recuadro 5.19. Estrategia de Francia para el uso de IA en la gestión de recursos humanos de la administración pública

Francia ha desarrollado una estrategia estructurada para integrar la IA en la gestión de recursos humanos en toda su administración pública. Esta estrategia se centra en tres áreas clave: la integración de la IA, la planificación de la organización y la formación de los funcionarios. Su objetivo es garantizar que la IA se utilice de manera responsable, ética y efectiva, mejore la productividad y respalde procesos complejos de toma de decisiones.

La estrategia se centra en primer lugar en identificar las partes interesadas, las tareas y las herramientas adecuadas para la adopción de la IA en las actividades de RR. HH. Incluye definir objetivos claros para el uso de la IA, seleccionar herramientas de IA confiables y establecer metodologías para garantizar la transparencia y la rendición de cuentas en los procesos impulsados por la IA. El plan también incorpora mapas de riesgo y auditorías internas para monitorear el impacto de la IA y ayudar a garantizar su uso ético.

La estrategia destaca la necesidad de un enfoque estratégico de planificación de la organización específicamente adaptado para integrar la IA en los procesos de RR. HH. La adopción de herramientas de IA para realizar tareas que tradicionalmente desempeñan los servidores públicos tiene implicaciones significativas y polifacéticas para el MRH. Para garantizar una transición fluida y una planificación sostenible de la fuerza laboral, es esencial anticipar y evaluar eficazmente estos impactos. Este enfoque implica integrar consideraciones relacionadas con la IA en los marcos de planificación de recursos humanos existentes, centrándose en:

- Definir las tareas delegadas a la IA, garantizando la alineación con los objetivos operativos y estratégicos;
- Evaluar los requisitos de habilidades actuales y futuros, incluidas las competencias específicas de IA y la experiencia complementaria;
- Evaluar el impacto de la IA en las profesiones de RR. HH. y las funciones laborales, identificando áreas de mejora de las competencias y de transformación de funciones; y
- Atraer y reclutar profesionales con conocimientos sobre IA para que apoyen, orienten y supervisen el uso responsable de la IA en la gestión de recursos humanos.

Para apoyar esta integración, el Gobierno Francés está haciendo hincapié en la formación en IA de los servidores públicos, en particular los directivos y los profesionales de recursos humanos. Ha creado programas de formación integrales para mejorar la calificación de los empleados y ha desarrollado directrices éticas que fusionan los recursos humanos y la ética digital. Este enfoque ayuda a garantizar que la IA se aplique de forma que se equilibre la tecnología con la supervisión humana, al tiempo que se prepara a la organización para el futuro de la IA en la administración pública.

Fuente: (French Public Finances General Directorate (DGFIP), 2024^[12]).

Mejoramiento de los procesos de contratación

La IA puede respaldar la reforma de la función pública durante todo el proceso de contratación, haciéndolo más rápido y eficiente. Las herramientas de IA pueden automatizar tareas transaccionales como redactar descripciones de puestos de trabajo, diseñar metodologías de evaluación adaptadas, comprobar los documentos de antecedentes de los candidatos (por ejemplo, títulos universitarios) y responder a sus consultas. En Singapur, por ejemplo, diez organismos públicos crearon un servicio de contratación con IA

para automatizar tareas repetitivas en el proceso de preselección, como la revisión y el screening de solicitudes. Un chatbot diseñado a medida también realiza una prueba escrita, revisando y puntuando el componente escrito de los candidatos. El servicio redujo significativamente la carga de trabajo de las agencias, lo que hizo que el proceso fuera más eficiente y efectiva¹⁶. El Reino Unido ha realizado esfuerzos especialmente sólidos en este ámbito (Recuadro 5.20).

Recuadro 5.20. Automatización de la contratación mediante IA en el Reino Unido

La Agencia Tributaria y Aduanera del Reino Unido (HMRC) utiliza una plataforma basada en la IA llamada Outmatch para automatizar el proceso de contratación de principio a fin en el caso de algunos puestos subalternos. Se pide a los candidatos que registren sus respuestas a seis preguntas vinculadas a un marco de competencias. A continuación, la herramienta de IA analiza sus respuestas y les asigna un puntaje. La herramienta está diseñada para gestionar grandes volúmenes de candidatos mediante la automatización de la etapa de evaluación y entrevista.

Un ámbito de gran interés es el uso de herramientas de IA para evaluar, simplificar y redefinir las descripciones de los puestos de trabajo con el fin de atraer a los mejores candidatos. La HMRC también está explorando cómo ayudar a los responsables de contratación con una herramienta de IA capaz de generar descripciones de puestos de trabajo, preguntas para entrevistas y publicaciones en redes sociales, mientras que otro prototipo permite el análisis de los mercados laborales regionales para respaldar campañas de contratación adaptadas. La Oficina del Gabinete ha elaborado una prueba de concepto, el Optimizador de Anuncios de Empleo (JAO, por sus siglas en inglés), que traduce en asesoramiento los resultados de descripciones de puestos de trabajo que anteriormente habían tenido éxito y que tenían objetivos similares. De este modo, los responsables de contratación podrán adaptar sus descripciones de empleo para dirigirse a candidatos de alta calidad con las competencias y la experiencia adecuadas. También se está estudiando la posibilidad de integrar la IA en otros aspectos de la contratación, como el apoyo a los candidatos —desde que identifican una función hasta que aceptan un puesto—, el traslado de un gran número de solicitudes, la programación y planificación de entrevistas, y la adecuación de los candidatos de las listas de reserva a otras funciones disponibles.

Fuente: <https://www.thetimes.com/uk/society/article/inside-hmrc-job-without-speaking-human-ai-tmrljvv2r>.

Además, la IA aplicada a la contratación en la función pública tiene un gran potencial para ampliar los grupos de candidatos y mejorar la selección de candidatos. A menudo, las administraciones públicas tienen dificultades para ser más proactivas a la hora de contratar personal, para solventar las carencias de competencias de sus administraciones y para impulsar activamente las oportunidades de empleo en el mercado (OECD, 2021^[49]). En este sentido, cabe citar, por ejemplo, uno procedente de Canadá en el Recuadro 5.21. Además, en Suecia, el gobierno municipal de Upplands-Bro desarrolló un robot de entrevista con IA adaptado a las necesidades de una empresa privada de consultoría en IA para garantizar un proceso de contratación más preciso y mejorar la eficiencia. Aunque se excluyen los datos relativos a la edad, el sexo, la ropa y el aspecto físico, el bot de IA realiza entrevistas a ciegas y evalúa a los candidatos de la primera ronda¹⁷.

Recuadro 5.21. Mayor representación de minorías visibles en el liderazgo de la defensa canadiense

En septiembre de 2020, el Departamento de Defensa Nacional de Canadá (DND, por sus siglas en inglés) lanzó una campaña piloto externa de reclutamiento EX-01 destinada a aumentar la representación de las minorías visibles en el nivel superior. Esta iniciativa puso a prueba nuevos enfoques al tiempo que adoptaba la inclusión, los métodos innovadores y la tecnología con el fin de lograr un cambio fundamental y duradero que ayudara a aumentar la representación y la diversidad e inclusión (D&I) en la DND. Los principales objetivos del programa piloto eran:

- Identificar oportunidades para miembros de grupos minoritarios visibles;
- Introducir nuevas herramientas y tecnología para apoyar las evaluaciones sin barreras ni sesgos; y
- Evaluar el proceso piloto frente a los enfoques tradicionales para identificar barreras y sesgos sistémicos e identificar recomendaciones para futuros procesos de contratación.

En colaboración con diversas partes interesadas, se establecieron barreras de protección adecuadas para garantizar la privacidad y la garantía de calidad. El proceso logró resultados clave al facilitar una evaluación objetiva y justa de los candidatos, al tiempo que eliminó sesgos y barreras en las diversas etapas. Como resultado, este programa piloto creó oportunidades de promoción profesional y mejoró los enfoques y procesos tradicionales para lograr mejores resultados.

Fuente: <https://www.canada.ca/en/department-national-defence/corporate/reports-publications/ex-01-report-visible-minorities-recruitment-campaign>.

Facilitar el aprendizaje y el desarrollo

La reforma de la función pública puede usar la IA generativa para crear contenido didáctico, como módulos de aprendizaje y material didáctico basado en documentos fuente e información. Los esfuerzos ya se están arraigando en todo el mundo. La IA también puede utilizarse para hacer recomendaciones personalizadas y orientar el aprendizaje y el desarrollo profesional a través de una compleja y gran cantidad de información y datos (Johnson, Coggburn and Llorens, 2022^[50]). A continuación, algunos ejemplos en estas áreas:

- La Comisión Australiana de Servicios Públicos (APSC, por sus siglas en inglés) colaboró recientemente con IBM en 2023 para diseñar, estructurar e implementar un sistema que generara contenido de cursos basado en documentos insertados por los usuarios (Recuadro 5.22).
- El Instituto Nacional de Administraciones Públicas (INAP) de España está incorporando la IA en las funciones de organización, catalogación y búsqueda de sus plataformas digitales, oferta de aprendizaje y biblioteca en general. El "gráfico de conocimientos" resultante hace que una gran parte de los recursos sean fácilmente localizables, susceptibles de búsqueda y compartibles con los funcionarios españoles, los países socios y el público en general.
- La plataforma de aprendizaje y desarrollo del Ministerio de Gestión del Personal de Corea incorpora funcionalidad basada en IA para ayudar a clasificar, organizar y recomendar contenidos. El sistema de IA que se está implementando pretende analizar el papel del usuario y su historial de aprendizaje para recomendar formación y material para desarrollar determinadas habilidades a partir del vasto catálogo de plataformas de 1,4 millones de piezas de contenido (OECD, 2023^[51]).

Recuadro 5.22. Generación de contenido didáctico en Australia

La Comisión Australiana de Administración Pública (APSC) llevó a cabo un proyecto piloto de seis semanas de duración para utilizar la IA con el fin de diseñar, estructurar e implantar un curso de aprendizaje en línea sobre competencias digitales para el liderazgo. El sistema piloto permitió a los profesionales "alimentar" con IA una variedad de materiales de información como artículos, libros y transcripciones de voz en un sistema cerrado de información que se utilizó para crear el contenido del curso. El sistema creó un esquema del curso, objetivos, módulos y contenido, seguido de un cuestionario. Los resultados de la prueba piloto mostraron que alrededor del 60-70 % de lo que el sistema producía era utilizable, relevante, bien estructurado y preciso. En una encuesta realizada a usuarios/as, el 91 % consideró que el resultado del programa piloto era valioso. Para aumentar estas cifras, el profesional que crea el contenido podría dar su opinión al sistema de IA para ajustarlo.

Esta iniciativa reportó varios beneficios. En primer lugar, la redacción fue muy, muy rápida. En segundo lugar, la información se extraía de un sistema cerrado, lo que eliminaba la incertidumbre en torno a dónde venía la información y si era correcta. La tecnología puso de relieve las esferas de los módulos y señaló dónde se había obtenido la información. En tercer lugar, el sistema también incorporaba "comprobadores" programáticos para examinar el material en relación con cuestiones preocupantes, como el lenguaje discriminatorio.

Si bien el sistema piloto podía redactar el contenido del curso muy rápidamente, no creaba contenido listo para la producción e implementación. Los problemas identificados en el programa piloto fueron:

- La IA puede sintetizar bien el contenido existente, pero no puede crear contenido que no existe. Por ejemplo, no pudo escribir un curso sobre el uso de la IA en la función pública en esa fase, ya que no existía contenido.
- La confiabilidad del contenido todavía requería la revisión de expertos, y es aquí donde normalmente ya existen cuellos de botella en la producción de contenido.
- El sistema cerrado proporcionó resultados más precisos; sin embargo, es una opción costosa si solo se utiliza para la producción del curso.

Tras la fase piloto, el APSC reconoce que los sistemas de IA cerrados y locales son útiles para ayudar en la creación de contenidos. Sin embargo, ha remitido una nueva investigación sobre la IA in situ a otro organismo público australiano, Services Australia. Se pueden encontrar más observaciones sobre el uso de la IA por parte de las agencias gubernamentales australianas en sus declaraciones sobre la IA.

Fuente: Información proporcionada a la OCDE por la Comisión de Administración Pública de Australia, <https://www.apsacademy.gov.au/news/piloting-generative-ai-address-aps-skills-gap>. APSC's AI statement (<https://www.apsc.gov.au/initiatives-and-programs/workforce-information/research-analysis-and-publications/state-service/state-service-report-2023-24/fit-future/supporting-safe-and-responsible-use-artificial-intelligence>), Services Australia's AI statement (<https://www.servicesaustralia.gov.au/automation-and-artificial-intelligence-ai-use>).

Evidencia de impacto

Dada la naturaleza incipiente de las aplicaciones descritas anteriormente, todavía es demasiado pronto para proporcionar evidencia empírica de su impacto. Los primeros experimentos como los descritos anteriormente muestran un potencial significativo para reducir el tiempo y el esfuerzo necesarios para manejar grandes volúmenes (por ejemplo, número de solicitantes, cantidad de material de aprendizaje existente), ampliar los grupos de candidatos y reducir el error humano en la toma de decisiones. Sin

embargo, en la mayoría de los casos, existe una notable falta de pruebas empíricas rigurosas que demuestren su efectividad e impacto. Muchas implementaciones de IA se basan en el potencial teórico o en historias de éxito anecdóticas, en lugar de evaluaciones científicas sólidas. Abordar esta limitación requiere un esfuerzo concertado para diseñar y llevar a cabo evaluaciones rigurosas de las aplicaciones de la IA en la gestión de los RR.HH. Las futuras implementaciones deberían basarse en evidencia empírica y adaptarse a las necesidades y limitaciones específicas de las organizaciones gubernamentales.

Además, hay muchas preguntas sin respuesta sobre los posibles impactos negativos. Por ejemplo, la contratación de herramientas de IA puede ayudar a reducir los errores humanos, dependiendo de cómo se utilicen. Sin embargo, también pueden limitar la autonomía de los gerentes de contratación de maneras que impacten la calidad de la contratación relacionada con la cultura, la idoneidad del equipo u otras decisiones de contratación. También pueden limitar aún más la capacidad de los gerentes para crear sus propios equipos para lograr los resultados que necesitan. La calidad real de las decisiones tomadas a través de los sistemas de IA es difícil de evaluar, ya que requiere una visión a largo plazo del rendimiento y la aptitud para el empleo. Además, hay bases de referencia limitadas para medir. Los sistemas y procesos tradicionales (es decir, anteriores a la IA) se enfrentan a desafíos a la hora de medirlos y evaluarlos, y no hay indicadores estándar acordados, especialmente en sistemas sin medidas objetivas de rendimiento, como suele ocurrir en las administraciones públicas.

Como tal, puede llevar mucho tiempo antes de que los gobiernos puedan medir adecuadamente el impacto real a largo plazo de la calidad de las decisiones impulsada por la IA en comparación con las decisiones tomadas por sistemas tradicionales o seres humanos. También será muy difícil cuantificar el impacto de la IA en la productividad de los sistemas de recursos humanos, ya que existen muy pocas medidas y puntos de referencia comparables estándar. La OCDE está trabajando actualmente con un conjunto básico de países miembros para tratar de establecer este tipo de indicadores, que pueden ayudar a realizar un seguimiento de las mejoras impulsadas por la IA en el futuro.

Gestión de riesgos y desafíos

Riesgos asociados

- "Sesgo de automatización".
- Datos inadecuados o sesgados en los sistemas de IA.
- Uso indebido o cuestionable de la IA, que genera problemas de vigilancia y privacidad.
- Falta de transparencia y explicabilidad.

En el ámbito de la reforma de la función pública existe **un riesgo documentado del llamado sesgo de automatización**, por el que los seres humanos prefieren no cuestionar los resultados de los asistentes para la toma de decisiones automatizadas, incluso cuando tienen la responsabilidad y la responsabilidad últimas de tomar la decisión final. Para optimizar el uso de la IA en la reforma de la función pública será necesario mejorar las capacidades analíticas estratégicas en muchas actividades de RR. HH. y entre los responsables de contratación (Broecke, 2023^[52]).

Otro desafío bien documentado se refiere a evitar, detectar y abordar la **parcialidad en los propios sistemas de IA**, en particular si se está utilizando la IA para fundamentar la toma de decisiones relacionadas con la selección de empleo y el avance profesional. El problema es que los datos históricos de cualquier organización se basan en decisiones pasadas tomadas por humanos, y existe un riesgo significativo de que los sistemas de IA codifiquen estas perspectivas en sus algoritmos. Si a esto se añade la falta general de buenos datos de los empleados y de indicadores de rendimiento, resulta difícil imaginar cómo podría ofrecerse un asesoramiento personalizado de calidad a través de estos sistemas de IA. La

introducción de cualquier sistema en este ámbito requerirá mecanismos de supervisión y evaluación cuidadosos para detectar y corregir los sesgos (Johnson, Cogburn and Llorens, 2022^[50]).

Los derechos a la privacidad y los datos de los servidores públicos requieren una atención especial en un contexto de servicio público, en el que valores como el mérito y la equidad orientan la contratación, limitando de forma inherente los tipos de datos que pueden utilizarse en los sistemas de IA. Por ejemplo, algunas herramientas de contratación del sector privado comprueban periódicamente las cuentas en redes sociales de los solicitantes y utilizan estos datos para evaluar a los candidatos. Hay poca evidencia empírica de que las publicaciones en medios sociales o las características físicas que pueden evaluarse en las entrevistas en video tengan algún impacto real en el desempeño laboral. Esto plantea interrogantes éticos y de efectividad sobre muchas herramientas de evaluación basadas en la IA actualmente disponibles en el mercado (Broecke, 2023^[52]).

El uso de herramientas de "gestión algorítmica"—software para automatizar aspectos de la gestión, por ejemplo, en la asignación de horarios de trabajo, el seguimiento de las actividades laborales o la fijación de objetivos para los trabajadores—está aumentando significativamente, alcanzando una tasa de adopción del 90 % en las empresas estadounidenses y del 79 % en la Unión Europea (Milanez, Lemmens and Ruggiu, 2025^[53]). Hasta la fecha, no se han realizado estudios específicos de los gobiernos. Si bien algunas de estas herramientas pueden ayudar a aumentar la productividad cuando se aplican de manera efectiva, se han planteado preocupaciones tangibles sobre los impactos negativos existentes de la IA y las herramientas algorítmicas en la calidad del empleo, incluida la intensificación del trabajo, el aumento del estrés y la reducción percibida de la equidad (OECD, 2023^[54]). La IA podría hacer que los trabajos sean menos gratificantes al incentivar **nuevos tipos de vigilancia en el lugar de trabajo** que podrían perjudicar la salud mental (APA, 2023^[55]), o nuevas formas de "taylorismo digital" hipereficiente pero agotador en las que el trabajo está sujeto a una mayor vigilancia y regulación (UC Berkeley, 2021^[56]). La gestión de tareas mediante IA también puede tener el potencial de erosionar la autonomía y la voz de los trabajadores, reduciendo así la percepción humana sobre cómo se gestiona el trabajo (Gmyrek, Berg and Bescond, 2023^[57]). Muchas de las herramientas de IA de RR. HH. existentes están diseñadas para "optimizar" la gestión de la organización (por ejemplo, supervisar, controlar y reducir la autonomía en la toma de decisiones y la resolución de problemas), en contra de décadas de ciencia que demuestra cómo el empoderamiento de los empleados fomenta el compromiso, el rendimiento y la confianza. Algunas investigaciones sugieren que encuadrar la IA como una herramienta utilizada para apoyar a los empleados, en lugar de sustituirlos o limitar su autonomía, es fundamental para fomentar percepciones positivas de la IA en el lugar de trabajo (Brougham and Haar, 2017^[58]).

Cuanto más complejos se vuelven los sistemas y las predicciones de IA, menos se pueden entender y explicar. Esto reduce la rendición de cuentas si los empleadores no pueden explicar sus decisiones y obstaculiza la capacidad de los empleados para comprender cómo desarrollarse para avanzar en sus carreras. Los sistemas de contratación basados en el mérito son la base del buen funcionamiento de los sistemas públicos de empleo, y para que funcionen correctamente se requieren transparencia y rendición de cuentas. Los empleados y sus empleadores necesitan entender claramente por qué se toman las decisiones de nombramiento, y cómo se analizan las habilidades y el desempeño de las personas (Cappelli and Rogovsky, 2023^[59]).

Desafíos de implementación

- Falta de datos de alta calidad y capacidad para compartirlos.
- Explicabilidad.
- Brechas de habilidades.

Disponer de datos de calidad es esencial para implementar la IA en la reforma de la función pública. Lamentablemente, los **países de la OCDE carecen de grandes cantidades de datos en la mayoría de**

las áreas relevantes, y a menudo no están estandarizados sistemáticamente en todas las organizaciones para permitir un análisis más riguroso y predictivo. Los datos descriptivos suelen limitarse a la edad, el sexo, el nivel educativo y la trayectoria profesional. Es muy difícil evaluar el rendimiento de forma objetiva y coherente en los distintos equipos y organizaciones. Los puestos de trabajo también suelen clasificarse de forma amplia. Si la IA se basa en datos erróneos o incompletos, hará malas predicciones.

La implementación de IA en los sistemas de gestión de la función pública requiere profesionales de recursos humanos con las habilidades y la mentalidad adecuadas. Aunque puede que no se requieran conocimientos técnicos sobre IA, los responsables de RR. HH. tendrían que entender la posible aplicación de la IA a sus sistemas y contar con las competencias adecuadas para ser compradores inteligentes de herramientas en el mercado. Los profesionales de recursos humanos que trabajan con herramientas de IA suelen necesitar capacidades analíticas para comprender los principios de estas herramientas y el uso que hacen de la analítica de datos, con el fin de interpretar y cuestionar los resultados. Aunque abundan en las administraciones públicas, los **departamentos de RR. HH. carecen a menudo de las competencias en IA necesarias**.

Potencial sin explotar y camino a seguir

Analítica fortalecida para el presente y el futuro

Las organizaciones gubernamentales podrían recopilar una gran cantidad de datos e información sobre sus empleados que podrían analizarse para mejorar el rendimiento y la experiencia de los empleados; sin embargo, la IA suele estar infravalorizada en el ámbito de los recursos humanos debido a varios desafíos. Las capacidades de IA podrían permitir a los profesionales y a la dirección de RR. HH. examinar las tendencias actuales de la organización (envejecimiento, competencias y rendimiento, remuneración) para ofrecer una visión clave de los principales desafíos y cuestiones del día a día: atractivo o competitividad del gobierno como empleador; necesidades de recalificación y mejora de las competencias; oportunidades de aprendizaje y desarrollo mejor orientadas; o los factores que impulsan el rendimiento y la satisfacción de los empleados y del equipo.

Los algoritmos de IA pueden utilizar datos de series temporales para el análisis predictivo con el fin de identificar tendencias y hacer predicciones sobre la función pública en relación con el futuro. Si bien se puede lograr un análisis de regresión simple sin IA, se podrían desarrollar operaciones más sofisticadas con modelos y creación de escenarios más complejos. Por ejemplo, las organizaciones podrían reducir potencialmente la rotación de los empleados al predecir los empleados de alto riesgo para la salida en función de su permanencia en un puesto, sus niveles de compromiso del equipo y otros factores. Mediante el análisis de grandes conjuntos de datos, la IA puede identificar factores y patrones que generan una rotación excesiva, lo que resulta costoso para las organizaciones y va en detrimento del rendimiento, y permite a los profesionales adoptar medidas preventivas para mejorar. Además de la rotación de personal, el análisis predictivo puede ayudar en la planificación de la fuerza laboral al anticipar la escasez de habilidades o personal, predecir los mejores desempeños para tipos específicos de roles, apoyar la diversidad y la inclusión, o impulsar el compromiso y el bienestar.

Actualmente, estas aplicaciones son extremadamente incipientes en la fuerza laboral gubernamental; la OCDE todavía no ha podido identificar muchos casos de uso concretos en esta área. Es probable que esto se deba a una variedad de desafíos, como los enumerados anteriormente, así como a preocupaciones éticas y de privacidad válidas, que se detallan más adelante.

Camino a seguir

Las herramientas de IA pueden brindar a los líderes una nueva oportunidad de desarrollar una visión y una dirección estratégicas para su servicio público y las actividades de RR. HH. necesarias para lograrlo. La IA tiene el potencial de remodelar la organización y aumentar sus competencias en muchas áreas. En el ámbito de los recursos humanos, la IA puede acelerar los procesos de RR. HH., mejorar la focalización de los servicios, el conocimiento y las campañas de contratación/marca, y generar información valiosa para los directores de RR. HH. y los altos directivos sobre una serie de cuestiones, desde las futuras carencias en materia de competencias hasta la efectividad en la contratación. La IA puede ser una herramienta transformadora en el aprendizaje y el desarrollo, al aportar conocimientos a los servidores públicos y, al mismo tiempo, desarrollar sus habilidades y capacidades esenciales. Todo esto depende de una visión clara de cómo debería ser la futura función pública, respaldada por recursos y equipos de recursos humanos capaces. Esto implica una estrategia conjunta para el desarrollo de la fuerza laboral en la que la IA desempeñe un papel clave.

Aumentar la transparencia y dar a los empleados una explicación y capacidad de impugnar las decisiones automatizadas. Los estudios sugieren que muchos candidatos y empleados pueden percibir los procesos de evaluación automatizada como más justos que los realizados por humanos si entienden cómo funcionan y por qué se los está utilizando, siempre y cuando confíen en que un humano será responsable de la decisión final (Broecke, 2023^[52]). Esta transparencia debe incluir los insumos para la decisión, asegurando que se tenga cuidado de obtener el consentimiento informado y gestionar los problemas de privacidad de los empleados. Los gobiernos deben disponer de una base empírica clara para los datos de partida que decidan utilizar.

Incluir a los profesionales de RR. HH. y otros empleados en el diseño, la implementación y la evaluación de las herramientas de RR. HH. e IA. Los gobiernos deben ser especialmente cautelosos a la hora de introducir herramientas de IA que puedan reducir la autonomía de los empleados. En algunos casos, la automatización impulsada por la IA podría reducir la autonomía de los trabajadores y devaluar su experiencia, lo que daría lugar a una menor motivación, participación y compromiso. Esto es especialmente cierto cuando las herramientas de IA se aplican directamente para optimizar su productividad, orientándolas sobre cómo usar su tiempo y monitoreando sus actividades laborales. Esto puede tener consecuencias adversas no deseadas; puede aumentar el estrés y la ansiedad, lo que podría aumentar la ausencia del trabajo y la rotación de personal, reduciendo así la productividad a largo plazo. Estas consecuencias pueden evitarse incluyendo a profesionales de RR. HH. y a otros empleados en el diseño, la implementación y la evaluación de las herramientas de RR. HH.

Mejorar la calificación de los profesionales de RR. HH. y recalificarlos para la era de la IA. La automatización tiene muchas ventajas y reduce la carga administrativa para los profesionales de RR. HH. y los empleados a los que prestan apoyo. Pueden encargarse de muchas de las tareas repetitivas y aburridas, dejando que los profesionales de RR. HH. se centren en tareas más complejas y de mayor valor añadido. Entre ellas se incluyen la resolución de casos complejos, la contratación de trabajadores más especializados y el desarrollo de estrategias (OECD, 2024^[60]). Sin embargo, en muchos casos, esto requerirá mejorar y recalificar a la profesión de RR. HH. En el caso de algunos tipos de empleados, los gobiernos tendrán que invertir en la mejora de las competencias en áreas más técnicas, como la ciencia de datos y la programación, para ayudar a los profesionales de RR. HH. a comprender y utilizar los sistemas de IA de forma efectiva.

Si los sistemas de IA se utilizan como entrada en las evaluaciones de los candidatos, los gobiernos deberían disponer de personas bien calificadas que interpreten los resultados y tomen las decisiones finales. Los gobiernos deben tener cuidado de diseñar un "sesgo de automatización". Esto puede incluir realizar primero evaluaciones tradicionales y, a continuación, añadir información sobre la IA después, para ofrecer información adicional sobre una breve lista de candidatos. De este modo, la IA puede ayudar a auditar las prácticas de contratación y mejorar la toma de decisiones humanas sin sustituirla. Esto es

esencial para garantizar el nivel adecuado de responsabilidad necesario para la toma de decisiones basada en el mérito en los servicios públicos.

Al tiempo que incorporan la IA en el ámbito de la analítica de personas, los gobiernos deben perseguir los beneficios de la IA teniendo en cuenta los costos y los riesgos. También deben tener establecidos determinados requisitos previos, como datos rigurosos y confiables, capacidades estadísticas para comprender y verificar los resultados y análisis generados por la IA y sus debilidades, y medidas para proteger la privacidad de los empleados y evitar sesgos.

IA en la contratación pública

Teniendo en cuenta que la contratación pública representa alrededor del 13 % del PIB en los países de la OCDE (OECD, 2024^[61]), la adopción de la IA en la contratación pública suele estar impulsada por la necesidad de mejorar la eficiencia y la toma de decisiones operativas y reducir costos (Friton et al., 2024^[62]; Hickok, 2022^[63]). También se utiliza para ayudar a abordar desafíos como las limitaciones de la fuerza laboral. Con la IA, la contratación pública puede volverse más dinámica y reactiva, capaz de satisfacer las demandas de un entorno rápidamente cambiante durante todo el ciclo de vida de la contratación (Hickok, 2022^[63]; Gastaldi et al., 2024^[64]) (Gráfico 5.3). Esta transformación digital también brinda la oportunidad de repensar de manera fundamental la contratación y administración pública, mejorar las conexiones entre las entidades públicas y los proveedores, y permitir una colaboración más dinámica (Glas and Kleeman, 2016^[65]). Sin embargo, para materializar todo el potencial de la IA se requiere una implementación efectiva, una gobernanza de datos sólida y un enfoque centrado en el usuario.

Gráfico 5.3. Posible uso de IA y análisis de datos a lo largo del ciclo de la contratación pública

Fase previa a la licitación				Fase de licitación				Fase posterior a la adjudicación		
Evaluación de necesidades y análisis del mercado	Planificación y elaboración de presupuestos	Desarrollo de especificaciones y requisitos	Elección del procedimiento adecuado	Solicitud de propuesta/oferta	Presentación de la oferta	Evaluación de pujas	Adjudicación del contrato	Gestión de contratos	Pedido y pago	Informes / evaluación del rendimiento
IA	Pronóstico, análisis automatizado de datos, identificación de riesgos y oportunidades.	Optimización de presupuestos, predicción de costos, análisis de escenarios presupuestarios.	Automatización del desarrollo de especificaciones mediante el análisis de documentos y la extracción de info.	Optimización del proceso de selección, reducción del tiempo de toma de decisiones.	Automatización de redacción del RFPs mediante el análisis de documentos anteriores y requisitos actuales.	Automatización del proceso de evaluación		Reducción de la supervisión manual, supervisión automatizada y comprobación del cumplimiento.	Procesamiento automatizado de pedidos y verificación de pagos.	Informes precisos sobre el desempeño de los proveedores y los resultados de las adquisiciones.
Macro-datos y análisis	Análisis de grandes conjuntos de datos para conocer las tendencias del mercado, evaluación de la demanda y mapeo de las capacidades de los proveedores.	Decisiones presupuestarias basadas en datos, identificación de oportunidades de ahorro de costes.	Determinación de especificaciones más efectivas sobre la base de contrataciones anteriores y análisis de mercado.	Análisis de los resultados de las adquisiciones anteriores para identificar el procedimiento más efectivo.		Toma de decisiones basada en datos, identificación de patrones y anomalías.				Generación de información sobre el rendimiento de proveedores, cumplimiento de contratos y eficiencia de adquisiciones.

Licitación

Firma del contrato

Fin del contrato

Nota: RPF = Solicitud de propuesta.

Fuente: Ilustración de la OCDE basada en (EC, 2020^[66]).

Situación actual

Las entidades públicas están integrando la IA y la toma de decisiones algorítmica en sus procesos, utilizando estas tecnologías para mejorar servicios, agilizar operaciones y mejorar la toma de decisiones, así como para fortalecer la gestión de riesgos, la supervisión y la rendición de cuentas. Un estudio reciente identificó las funcionalidades basadas en IA ofrecidas por las plataformas de contratación pública, revelando que el 54 % de las soluciones admiten la prelicitación y la planificación, el 31 % se centra en las actividades de licitación y solo el 4 % aborda las actividades más operativas de la fase de suministro

(Guida et al., 2023^[67]). Además, el 11 % de las soluciones, incluyendo los asistentes digitales y la automatización de actividades sin valor añadido, abarcan todo el ciclo de vida de la contratación. Entre las funcionalidades habituales se incluyen el análisis de gastos, la gestión de riesgos, el financiamiento de la cadena de suministro, la identificación de proveedores y la optimización de negociaciones.

Agilización de tareas operativas

La IA puede ayudar a clasificar el gasto para estandarizar los informes. Clasificar los datos sobre gasto en taxonomías estándar en la contratación pública significa organizar el gasto en categorías predefinidas, lo que ayuda a hacer seguimiento y analizar el gasto de manera más efectiva. Su importancia radica en la capacidad de tomar mejores decisiones, aumentar la transparencia e identificar oportunidades de ahorro de costos. Por ejemplo, el sistema ucraniano de contratación electrónica ProZorro utiliza una solución ML para predecir el código correcto del Vocabulario Común de Contratos Públicos (CPV, por sus siglas en inglés) para productos y servicios (Recuadro 5.23). Estos códigos establecen un sistema único de clasificación de la contratación pública, destinado a estandarizar las referencias utilizadas para describir los contratos públicos con el fin de aumentar la transparencia y facilitar la identificación de oportunidades por parte de los proveedores potenciales (EC, 2020^[66]).

Recuadro 5.23. El sistema ucraniano de contratación electrónica ProZorro

Reconociendo los desafíos planteados por la falta de competencia leal y los riesgos de corrupción en las prácticas de contratación, las organizaciones de la sociedad civil ucraniana, las asociaciones de expertos y el gobierno colaboraron en 2014 para desarrollar ProZorro. Este sistema de contratación pública de código abierto utiliza IA y analítica avanzada para mejorar la eficiencia, la transparencia y la rendición de cuentas. Una década después de su creación, ProZorro ha evolucionado hasta convertirse en un ecosistema basado en datos que integra herramientas de inteligencia de negocios (BI, por sus siglas en inglés) de vanguardia para apoyar la toma de decisiones basada en evidencia en el gobierno, los órganos de supervisión, las empresas y la sociedad civil.

Los módulos de análisis de BI ProZorro y ProBI del sistema permiten el monitoreo en tiempo real y el análisis profundo de los datos de compras. El módulo público BI Prozorro ofrece 49 paneles de control que cubren todas las etapas de la contratación, lo que permite a los usuarios evaluar las tendencias de precios, la contratación médica, el rendimiento del comprador y los indicadores de riesgo. El módulo ProBI proporciona una herramienta de creación de informes para que los usuarios avanzados creen análisis personalizados, facilitando la supervisión regulatoria y la planificación estratégica de las contrataciones.

El impacto de la analítica de ProZorro es sustancial. Estas herramientas se han adoptado ampliamente, y más de 30 000 usuarios han analizado anualmente el mercado de contratación pública de Ucrania, abarcando transacciones por valor de 21 000 millones de euros en 2023. Más de 80 estudios de contratación pública han servido de base para mejoras regulatorias, incluyendo un ahorro de 250 millones de euros desde 2021 gracias a cambios en las políticas relacionados con correcciones en licitaciones. Además, las evaluaciones de riesgos basadas en la IA ayudan a los órganos de supervisión a detectar irregularidades, mientras que las entidades de contratación utilizan información basada en datos para optimizar las estrategias de compra. El modelo colaborativo de gobernanza del sistema, que integra a las partes interesadas del gobierno, la sociedad civil y el sector privado, promueve la innovación continua y la capacidad de respuesta a los desafíos emergentes.

Fuente: <https://prozorro.gov.ua>.

La IA puede ayudar a los gobiernos a mejorar los procesos simplificando y optimizando rápidamente flujos de trabajo altamente reglamentados de principio a fin. Simplificar y racionalizar los flujos de trabajo integrales y altamente regulados en la contratación pública implica reducir la complejidad y automatizar las tareas, por ejemplo, con la introducción de RPA o LLM para mejorar la eficiencia. Esto puede lograrse integrando herramientas y tecnologías digitales para minimizar los errores y acelerar el proceso de contratación. Por ejemplo, en Chile, ChileCompra ha transformado la forma en que se realiza la contratación pública y ha evolucionado con el tiempo para utilizar tecnologías novedosas en apoyo de los equipos gubernamentales (Recuadro 5.24).

Recuadro 5.24. Chile avanza en contratación pública con ChileCompra

En Chile, la contratación pública se ha transformado a través de ChileCompra, el organismo central de compras del país creado en 2000. ChileCompra opera Mercado Público, una plataforma electrónica que centraliza y agiliza la contratación de bienes y servicios para entidades públicas. La plataforma soporta los acuerdos marco, que permiten a múltiples proveedores ofrecer productos en condiciones estandarizadas, fomentando la inclusión y la competencia. Con el tiempo, ChileCompra se ha convertido en la tienda virtual más grande de Chile, promoviendo la transparencia y la eficiencia en la contratación pública. Además, ha introducido herramientas innovadoras como ChileCompra Express, un mercado en línea que permite las compras directas a proveedores previamente aprobados sin necesidad de procedimientos de licitación adicionales.

ChileCompra ha evolucionado continuamente para hacer frente a desafíos, como la participación desigual de los proveedores y las ineficiencias de acuerdos marco. En 2014, más de 850 entidades públicas utilizaban la plataforma, generando aproximadamente 810 000 órdenes de compra por valor de 1.800 millones de USD al año. Entre 2010 y 2015, el número de proveedores que realizaron operaciones en el sistema aumentó casi un 180 %, creando oportunidades para las pequeñas y medianas empresas (PYMES). Sin embargo, el sistema se enfrentaba a desafíos operativos, como una elevada concentración de ingresos entre un pequeño número de proveedores y una competencia limitada en la segunda fase de acuerdos marco. Para abordar estos problemas, ChileCompra rediseñó sus acuerdos marco estandarizando las categorías de productos e introduciendo mecanismos competitivos que redujeron los precios hasta en un 28 % en comparación con los precios de mercado.

En los últimos años, ChileCompra ha integrado IA para seguir modernizando las prácticas de contratación. Introdujo plantillas de licitación estandarizadas para proyectos de IA y ciencia de datos como parte de su iniciativa Algoritmos Éticos en colaboración con la Universidad Adolfo Ibáñez y BID Lab. Estas plantillas incluyen requisitos de transparencia, privacidad, no discriminación y explicabilidad para garantizar un uso responsable de la IA en los contratos públicos. Además, el Observatorio de Contratación Pública de ChileCompra utiliza herramientas de IA, como los LLM, para analizar los datos de las contrataciones en busca de irregularidades y mejorar la supervisión del cumplimiento. Estos avances han permitido una supervisión más eficiente y, al mismo tiempo, han promovido normas éticas en la contratación pública.

Fuente: <https://www.chilecompra.cl>.

La IA puede ayudar a agilizar los esfuerzos por gestionar los marcos jurídicos y reglamentarios de la contratación pública. Por ejemplo, en Rumanía, la Agencia Nacional de Contratación Pública (ANAP) desarrolló una herramienta para mejorar su capacidad de examinar la nueva legislación, que incluye la recuperación de documentos en tiempo real de los sitios web de las instituciones públicas y la conversión de los documentos escaneados en texto que permite realizar búsquedas (The World Bank, 2023^[68]).

Mejoramiento de las relaciones compradores -proveedores y de las capacidades de los servidores públicos

La IA proporciona comunicación y asistencia en tiempo real a través de chatbots, que pueden responder a consultas, proporcionar actualizaciones y facilitar interacciones más fluidas. Esta tecnología ayuda a agilizar los procesos, reducir los tiempos de respuesta y mejorar la satisfacción general de ambas partes. Por ejemplo, en los Estados Unidos, el departamento de compras y aprovisionamiento (PSS, por sus siglas en inglés) del Concejo de El Paso (PSS) integró una solución de chatbot, llamada Ask Laura, en su página web. Ask Laura utiliza algoritmos de código abierto para interpretar y tratar las preguntas y obtener información sobre los proveedores potenciales en función de los perfiles de negocio y las preguntas formuladas (Collins, 2020^[69]).

Las herramientas de IA pueden facilitar el apoyo y orientación en tiempo real a los profesionales de la contratación y los proveedores, agilizando la comunicación y agilizando la resolución de consultas. Esto puede ayudar a mejorar la colaboración al ofrecer asesoramiento oportuno, llenar automáticamente los formularios y crear paneles personalizados. Por ejemplo, en Estados Unidos, el Departamento de Tecnología de la Información de Carolina del Norte (NCDIT, por sus siglas en inglés) ha introducido un chatbot impulsado por IA para ayudar al personal de las agencias estatales con los procesos de compra de TI. Disponible las 24 horas del día, los 7 días de la semana, el chatbot ofrece respuestas instantáneas a las consultas más habituales, como el acceso a los formularios de contratación, la presentación de solicitudes de excepción y la comprensión de los plazos de contratación, lo que reduce los tiempos de espera y mejora la eficiencia (NCDIT, 2024^[70]).

Mejoramiento de la gestión de riesgos, supervisión y rendición de cuentas

Los controles de cumplimiento automatizados, los algoritmos de detección de fraude y los mecanismos de detección de anomalías refuerzan la rendición de cuentas al señalar irregularidades y desviaciones de los protocolos de contratación establecidos. Las funcionalidades de IA utilizan técnicas de ML para identificar automáticamente errores y fraude, y para gestionar el riesgo de manera eficiente y efectiva (Guida et al., 2023^[67]).

La IA también puede utilizarse para detectar vulneraciones de la integridad en las licitaciones. Tanto las administraciones como la ciudadanía están investigando el potencial de la IA en este ámbito. Como ejemplo relevante, utilizando datos de licitaciones públicas disponibles, los investigadores han podido usar algoritmos de IA para detectar colusión con una tasa de precisión del 81-95 %. Una vez entrenados los algoritmos, pueden actualizarse automáticamente con las últimas subastas y con poco esfuerzo por parte del usuario para supervisar sus resultados (Garcia Rodriguez et al., 2022^[71]). Además, en Hungría, un estudio analizó 119 000 licitaciones públicas entre 2011 y 2020 para identificar estrategias sutiles basadas en texto utilizadas por agentes corruptos para favorecer a licitadores específicos. Utilizando enfoques de ML como los bosques aleatorios, el estudio descubrió que los datos de texto mejoraron la precisión del sistema del 77 % al 82 %, lo que demuestra el potencial de la minería de texto para descubrir comportamientos corruptos y mejorar las políticas anticorrupción (Katona and Fazekas, 2024^[72]). Otro ejemplo relevante es el de España, donde los investigadores desarrollaron un sistema de IA para proporcionar un "sistema de alerta temprana" que predice la corrupción pública. La herramienta utiliza datos sobre factores económicos y políticos —como el crecimiento económico y la duración en el poder de un partido político— junto con datos sobre casos de corrupción para predecir el riesgo de corrupción en las provincias españolas (López-Iturriaga and Sanz, 2017^[73]). La lucha contra la corrupción, más allá de la contratación pública, se analiza más adelante en la sección titulada "La IA en la lucha contra la corrupción y la promoción de la integridad pública".

Más allá de identificar anomalías en los datos existentes, las capacidades predictivas de la IA pueden identificar posibles riesgos e irregularidades y optimizar los procesos de contratación pública. En Brasil,

por ejemplo, la Contraloría General desarrolló Alice, una herramienta que utiliza IA para detectar posibles casos de fraude, lo que permite la gestión y supervisión de riesgos en tiempo real (Recuadro 5.25).

Recuadro 5.25. Supervisión de la contratación impulsada por IA en Brasil con Alice

En Brasil, la contratación pública representa una parte significativa del gasto público, lo que la convierte en un área crítica para garantizar la eficiencia y la transparencia. Para abordar vulnerabilidades como el fraude, las ineficiencias y los errores en el proceso de contratación, la Contraloría General (CGU) desarrolló Alice, un sistema impulsado por IA diseñado para analizar licitaciones, contratos y avisos públicos. Alice utiliza la inteligencia artificial y la RPA para supervisar de forma continua las actividades de contratación en todos los organismos federales, identificando los riesgos y las irregularidades en tiempo real. Al automatizar estos procesos, Alice permite la auditoría a gran escala y ayuda a los funcionarios públicos a tomar decisiones informadas para mejorar la supervisión.

Desde su implementación, Alice ha dado resultados notables. Solo en 2023, analizó casi 191 000 adquisiciones y activó 203 auditorías relativas a contratos por un valor equivalente de 4 150 millones de euros. Entre 2019 y 2022, sus alertas dieron lugar a la suspensión o cancelación de ofertas por un total equivalente a aproximadamente 1 500 millones de euros. Además, el sistema ha acelerado considerablemente los procesos de auditoría, reduciendo el tiempo medio necesario de 400 días a solo ocho. Alice aprovecha algoritmos de NLP adaptados capaces de gestionar las complejidades únicas de los datos de contratación pública brasileños, mejorando aún más su efectividad en la identificación de riesgos a través de aproximadamente 40 tipologías predefinidas.

Al combinar tecnología avanzada con marcos institucionales sólidos, Brasil mejoró la supervisión, logró ahorros financieros sustanciales y mejoró la rendición de cuentas en el gasto público.

Fuente: <https://oecd-opsi.org/innovations/robot-alice-bid-contract-and-notice-analyser>.

Los órganos de supervisión, incluidas las instituciones superiores de auditoría, son agentes clave que garantizan la legalidad, eficiencia e integridad de los procesos de contratación (OECD, 2015^[74]). Cada año, el Tribunal de Cuentas de Portugal (Tribunal de Contas, TdC) lleva a cabo un número significativo de revisiones y auditorías relacionadas con los procesos de contratación pública (antes, durante y después), que requieren amplios recursos humanos y financieros. TdC está trabajando con la OCDE para desarrollar capacidades de control más sólidas y una asignación de recursos más eficiente. Esto incluye el desarrollo y las pruebas de los flujos de datos y los sistemas de ML para identificar señales de alarma, centrándose en los riesgos de irregularidades en los procesos de contratación pública (OECD, 2024^[75]).

Empoderamiento de agentes externos y fortalecimiento de la confianza en el gobierno

La IA también puede empoderar a agentes externos, como la ciudadanía y organizaciones de la sociedad civil, para que realicen evaluaciones por terceros de los programas de contratación pública y el gasto público (Santiso, 2022^[76]). Los gobiernos han desarrollado plataformas que permiten a las partes interesadas acceder a datos abiertos sobre contratación pública, facilitando así un intercambio de información transparente (Attard et al., 2015^[77]). Estos portales publican datos tanto estructurados como no estructurados que abarcan varias fases de los procesos de contratación. Para mejorar la presentación de estos datos, algunos organismos gubernamentales han integrado paneles de control impulsados por la IA en estas plataformas que muestran estadísticas e indicadores relevantes para los procesos de contratación (Ansari, Barati and Martin, 2022^[78]). Países como Colombia (2024^[79]) Chile (2024^[80]) y México (2024^[81]) han adoptado tales iniciativas.

Las tecnologías de IA pueden permitir a las partes interesadas supervisar las actividades de contratación pública y gestionar los riesgos con una granularidad sin precedentes, al proporcionar acceso en tiempo real a los datos de las contrataciones, las pistas de auditoría y las métricas de rendimiento. Además, las iniciativas de transparencia impulsadas por la IA pueden promover una mayor confianza pública en las prácticas de contratación pública, fomentando una cultura de integridad y rendición de cuentas. Las herramientas de IA pueden ayudar a las administraciones a anticipar la demanda, identificar posibles riesgos y optimizar sus procesos de contratación. En Brasil, por ejemplo, el proyecto Labcontas (GLOBO, 2018^[82]) que reunió 96 bases de datos con información relevante para el trabajo del Tribunal Federal de Cuentas brasileño (TCU), ha permitido realizar controles automatizados de las licitaciones públicas que plantean un riesgo potencial de corrupción (EC, 2020^[66]).

Gestión de riesgos y desafíos

Varios riesgos y desafíos potenciales son motivo de preocupación en cuanto el uso de la IA en la administración pública (Andersson, Arbin and Rosenqvist, 2025^[83]; Shark, 2024^[84]). A continuación, se desarrollan las cuestiones que más se abordan en el trabajo de la OCDE y la evaluación de los casos de uso analizados en esta sección.

Riesgos asociados

- Datos inadecuados o sesgados en los sistemas de IA.
- Falta de transparencia y explicabilidad.

Los sistemas de IA con **datos de entrenamiento inadecuados o sesgados** pueden presentar problemas, ya que los sistemas utilizados para evaluar las ofertas podrían favorecer a determinados licitadores, dando lugar a decisiones de contratación injustas. Además, la capacidad de los equipos de contratación pública para comprender el funcionamiento preciso de los sistemas algorítmicos está limitada por sesgos humanos en la percepción, complejidades múltiples y otros factores (Hickok, 2022^[63]). Estos tipos de sistemas de IA pueden tener resultados perjudiciales que podrían afectar a una parte mayor de la sociedad. En las licitaciones, un empleado público humano revisa las ofertas y decide una por una, mientras que un sistema de IA puede revisar muchas ofertas y tomar decisiones en cuestión de segundos y minutos. Por lo tanto, si existe un sesgo en el sistema de IA, la velocidad y la magnitud del daño también superarán al de una revisión humana (Hickok, 2022^[63]). Para evitar estos resultados, los contratantes deben asegurarse de que la IA ha sido entrenada con datos representativos. Además, los sistemas de IA deben diseñarse teniendo en cuenta criterios de equidad, que vayan más allá de los costos y la eficiencia. Por último, el rendimiento de la IA debe supervisarse en escenarios reales para detectar y hacer frente a los daños emergentes.

También surgen preocupaciones respecto a la **falta de transparencia algorítmica** de la IA. Cuando se implementa un sistema de IA en la administración pública, los gobiernos deben cumplir sus compromisos con los principios de equidad, rendición de cuentas y transparencia. El sistema debe ser suficientemente explicable; la autoridad contratante debe obtener suficiente información sobre el funcionamiento del sistema y los datos sobre los que ha sido entrenado para extraer sus conclusiones. De lo contrario, los agentes públicos dispondrán de sistemas integrados sin una capacidad independiente para mantenerlos y supervisarlos. Sin estas capacidades, tampoco se dispondrá de mecanismos alternativos de supervisión y rendición de cuentas debido a la falta inicial de transparencia o acuerdos de subcontratación (Hickok, 2022^[63]). En el Reino Unido, la Oficina de Inteligencia Artificial (OAI) y el Servicio de Gobierno Digital (GDS) elaboraron una guía en colaboración con el Alan Turing Institute para salvaguardar la confianza pública en el uso de la IA en la contratación mediante el uso de los principios FAST Track (por sus siglas en inglés): equidad, rendición de cuentas, sostenibilidad y transparencia (GOV.UK, 2019^[85]).

Desafíos de implementación

- Entornos legales y regulatorios inflexibles u obsoletos.
- Falta de datos de alta calidad y capacidad para compartirlos.
- Brechas de capacidades.
- Aversión al riesgo.
- Dependencia en datos y proveedores.

Muchas jurisdicciones **carecen de normativas y directrices formales sobre el uso de la IA**, lo que genera ambigüedades jurídicas y posibles dificultades para los licitadores infructuosos que cuestionan la equidad del proceso. Dadas estas brechas regulatorias, existe una creciente necesidad de marcos regulatorios y directrices que garanticen una orientación clara para la IA en la contratación pública, promoviendo la transparencia y la equidad, reduciendo las ambigüedades legales y minimizando los desafíos de los licitadores infructuosos. El Recuadro 5.26 muestra cómo algunos gobiernos se están uniendo para superar este desafío.

Recuadro 5.26. Coalición GovAI para la contratación responsable de IA y despliegue en los Estados Unidos

La Coalición GovAI es una iniciativa interinstitucional dedicada a promover el uso responsable y ético de la IA en el Gobierno. Fundada en 2023 por la ciudad de San José, la coalición se ha ampliado desde entonces para incluir agencias locales, estatales y federales en todo Estados Unidos. Sirve como plataforma para la colaboración interinstitucional, el intercambio de conocimientos y las mejores prácticas de gobernanza de la IA, ayudando a los gobiernos simultáneamente a promover la innovación y garantizar la rendición de cuentas.

Los miembros de la Coalición GovAI han colaborado para crear un conjunto de plantillas de contratación pública y herramientas de intercambio de conocimientos que cualquier organismo público puede utilizar para poner en marcha su propio programa de gobernanza de la IA. Entre estos recursos, la Coalición GovAI desarrolló el Centro de Contratos de IA, lanzado en febrero de 2025 en colaboración con Pavilion. Esta plataforma agiliza la contratación de IA al ofrecer un repositorio compartido de modelos de contrato, acuerdos cooperativos y mejores prácticas. El objetivo del centro es reducir los costos y los plazos de la contratación, mejorar la transparencia de los contratos y ampliar el acceso a los proveedores de IA. A medida que crece la adopción de la IA en el gobierno, hasta alcanzar los 3 300 millones de USD en contratos federales relacionados con la IA en 2022, la Coalición GovAI ayuda a garantizar que los organismos públicos dispongan de las herramientas necesarias para adquirir soluciones de IA de manera eficiente, al tiempo que defiende los valores públicos.

Fuente: <https://www.sanjoseca.gov/your-government/departments-offices/information-technology/ai-reviews-algorithm-register/govai-coalition>.

La IA puede mejorar significativamente todo el ciclo de contratación pública. Sin embargo, su impacto es limitado sin **datos estandarizados y accesibles**, lo que requiere una estrategia de gobernanza de datos coherente para todo el Gobierno. Además, los organismos gubernamentales suelen implementar diversos sistemas de IA sin estándares unificados, lo que da lugar a sistemas incompatibles, datos fragmentados e inefficiencias en la agregación de la información de contratación. Las restricciones al intercambio de datos contribuyen a estos desafíos y dificultan la obtención del máximo provecho de los sistemas de IA al ampliar el alcance del entrenamiento y el análisis de los sistemas (Andersson, Arbin and Rosenqvist, 2025^[83]). Aunque los países de la OCDE han realizado considerables progresos, los datos que contienen

toda la información pertinente sobre contratación pública siguen sin estar disponibles en gran medida y existen muy pocos datos abiertos (definidos como datos reutilizables en un formato accesible) en la mayoría de los países evaluados (da Rosa, 2023^[86]).

Las carencias en materia de competencias digitales, así como la falta de comprensión del potencial de la IA, son obstáculos importantes para el despliegue satisfactorio de la IA en los procesos de contratación pública (Guida et al., 2023^[67]). El poder de la IA por sí solo no es suficiente para que pueda adoptar con éxito plataformas avanzadas de contratación pública. La gestión de datos, el cambio cultural y el desarrollo de competencias son fundamentales (Handfield, Jeong and Choi, 2019^[87]). Si persisten estas deficiencias y el conocimiento que se tiene es limitado, las entidades públicas podrían tener dificultades para implementar y gestionar eficazmente los sistemas de IA o mitigar sus riesgos, dando lugar a procesos de contratación ineficientes y potencialmente inadecuados.

Otro hallazgo interesante es que **los gerentes de contratación son en su mayoría escépticos de la IA**, creyendo que las habilidades típicas del comprador humano están estrechamente relacionadas con la negociación y que este conocimiento, a menudo tácito y no formalizado, no puede transferirse a agentes o sistemas autónomos (Guida et al., 2023^[67]). En consecuencia, consideran que este conocimiento no puede transferirse de manera efectiva a agentes o sistemas autónomos. Esta aversión al riesgo pone de relieve la importancia de la preparación digital y el nivel de competencias digitales de los funcionarios públicos, ya que estos factores son cruciales para integrar con éxito la IA en los procesos de contratación.

Los acuerdos de licencia de datos restrictivos o con un diseño deficiente pueden **crear dependencia de los datos**, impidiendo que la autoridad contratante comparta los datos necesarios con el desarrollador de IA, limitando así la efectividad del sistema de IA. Además, existe el riesgo de **bloqueo de proveedores**, lo que hace que la autoridad contratante dependa en gran medida de la tecnología y los formatos de datos propios del proveedor de IA.

Potencial sin explotar y camino a seguir

Los resultados de la OCDE y las investigaciones externas identifican poca investigación específica sobre la IA para la contratación pública, y un nivel bastante bajo de madurez en IA en las entidades de contratación pública (Andersson, Arbin and Rosenqvist, 2025^[83]). Para ayudar a las organizaciones de contratación pública a evaluar su propia madurez e identificar los factores necesarios para el crecimiento, el Centro para el Negocio del Gobierno de IBM ha desarrollado un modelo de madurez de IA para la contratación pública que puede ser una referencia útil (2023^[88]).

Si se adopta con éxito en el terreno, el potencial de la IA en la contratación pública incluye la evaluación automatizada de los proveedores, sistemas predictivos para anticipar acontecimientos —como la escasez de productos y el momento óptimo para garantizar los mejores precios—, la detección de posibles influencias de cuestiones económicas y geopolíticas más amplias y la creación de plataformas de licitación inteligentes que satisfagan automáticamente las necesidades de la contratación con los licitadores más adecuados (Shark, 2024^[84]). Los LLM podrían utilizarse para promover la integridad en el gasto público. Por ejemplo, modelos como los que impulsan ChatGPT pueden ayudar a los funcionarios de contratación pública a analizar grandes cantidades de datos sobre una empresa y un posible contratista para detectar riesgos de fraude o corrupción (Ugale and Hall, 2024^[89]).

La IA también podría incidir en el establecimiento de los requisitos y especificaciones que necesita un funcionario de compras, la evaluación de las ofertas y la fijación de precios justos y razonables, la optimización de la selección de proveedores y el cumplimiento de la regulación (IBM, 2023^[88]). La IA podría ayudar a los funcionarios de contratación pública a determinar las especificaciones de una compra presentando información sobre ejercicios de contratación anteriores o identificando y presentando dinámicamente productos y servicios relevantes en el mercado (IBM, 2023^[88]). La mejora del

conocimiento del mercado a través de contenidos recopilados y sintetizados por IA también podría ayudar a los funcionarios de contratación a identificar precios razonables y justos para diversos productos y servicios que ayuden a servir de referencia para la compra directa o para considerar ofertas competitivas. Además, la IA podría ayudar a sintetizar la información de los proveedores para la toma de decisiones en materia de contratación pública. La información sobre los proveedores puede analizarse utilizando métodos de NLP de diversas fuentes, como perfiles de empresa, datos financieros y opiniones en Internet (Burger, Nitsche and Arlinghaus, 2023^[90]).

Otras aplicaciones potenciales se pueden encontrar en el uso de herramientas de IA por parte del sector privado. La contratación privada ha sido objeto de mucha más investigación, aunque las empresas tardan en adoptar la IA en comparación con otras funciones empresariales (Andersson, Arbin and Rosenqvist, 2025^[83]). Estos enfoques pueden servir de inspiración a las oficinas de contratación pública, que no pueden replicar directamente las soluciones del sector privado debido a los marcos jurídicos y reglamentarios. Algunos ejemplos son:

- Walmart ha implementado Pactum AI, un chatbot de negociación para proveedores que proporcionan "bienes no destinados para reventa", cuyo objetivo era mejorar las condiciones de pago, asegurar descuentos y ofrecer avisos flexibles de terminación de contratos¹⁸.
- AutogenAI ha desarrollado una herramienta de IA diseñada para agilizar el proceso de redacción de ofertas en contrataciones. Esta herramienta ayuda a las empresas a elaborar propuestas de manera más eficiente, reduciendo el tiempo y el esfuerzo necesarios para responder a las oportunidades de contratación¹⁹.
- Sievo, una empresa de analítica de contratación, ofrece una plataforma para mejorar los procesos de contratación que utiliza la IA para analizar datos de gastos, prever la demanda y optimizar la selección de proveedores, mejorando así la toma de decisiones y la eficiencia operativa²⁰.
- DocuSign desarrolló una herramienta de gestión de contratos basada en IA que utiliza NLP para escanear e interpretar documentos legales, identificando oportunidades de ahorro de costos y garantizando el cumplimiento²¹.

Los poderes adjudicadores de los gobiernos deben tomar medidas para asegurarse de que están adoptando acciones y decisiones informadas al incorporar la IA en sus procesos de contratación. Deberían centrarse en reducir tanto los riesgos como la aversión al riesgo, mejorar las capacidades, alentar a los funcionarios de contratación pública a dialogar con los proveedores y mejorar la recopilación de datos y el seguimiento de los resultados. La Recomendación de la OCDE sobre Contratación Pública (2015^[74]) y los Principios de IA de la OCDE (2024^[91]) — entre otros estándares internacionales y de la OCDE — ayudan a las entidades adjudicadoras en sus esfuerzos por implementar una IA confiable en sus respectivos procesos. Sin un enfoque informado y confiable, los agentes públicos podrían no aprovechar los beneficios de los sistemas de IA. También podrían terminar con sistemas altamente integrados pero con fallas, sin una capacidad independiente para mantener estos sistemas, o capacidades para monitorear su desempeño (Hickok, 2022^[63]).

El éxito de la IA en los procesos de contratación pública depende tanto de la aplicación efectiva por parte de la entidad compradora como del compromiso de las principales partes interesadas. Es esencial adoptar un enfoque centrado en el usuario para la transformación digital de la contratación pública. El éxito también depende de una sólida gobernanza e infraestructura de datos, que incluya la estandarización, el intercambio y el uso de los datos de las licitaciones, así como de los modernos sistemas informáticos necesarios para albergar y transferir datos de manera eficiente y gestionar los sistemas de IA.

Los mecanismos de gobernanza y las estructuras de rendición de cuentas siguen siendo esenciales. Los responsables de las políticas y las organizaciones intergubernamentales están estableciendo regulaciones para regular el uso de la IA en la contratación pública. Para garantizar una gobernanza y

una rendición de cuentas sólidas, las obligaciones y la documentación de la contratación pública específicas de la IA deben aplicarse por igual al desarrollo tanto externo como interno (Heikkila, 2022^[62]).

Los gobiernos también deberían invertir en programas de desarrollo de capacidades, iniciativas de formación y plataformas de intercambio de conocimientos para apoyar la adopción de la IA en la contratación pública. Al invertir en programas de capacitación y desarrollo de capacidades, los profesionales de la contratación pueden estar equipados con las capacidades y los conocimientos necesarios para utilizar eficazmente las tecnologías de IA.

Además, la colaboración entre los sectores público y privado, el mundo académico, la sociedad civil y, en algunos casos, el público es esencial para fomentar la innovación y difundir las mejores prácticas en los sistemas de contratación pública basados en la IA. Al promover el intercambio de conocimientos y la colaboración, los gobiernos pueden acelerar la adopción de las tecnologías de IA y maximizar los posibles beneficios para la sociedad. Esta colaboración también ayuda a determinar si la IA es la mejor solución para un desafío determinado en relación con otros enfoques o tecnologías, un paso importante pero a menudo ignorado (Hickok, 2022^[63]).

Los organismos públicos de adquisición deben dar prioridad a la evaluación continua y la iteración de los sistemas de IA utilizados en la contratación pública, por ejemplo, el seguimiento del rendimiento de los algoritmos de IA. Deberían evaluar su impacto en los resultados de la contratación pública y en las personas y solicitar la opinión de las partes interesadas.

IA en la lucha contra la corrupción y promoción de la integridad pública

Los actores de la integridad, como las agencias anticorrupción (ACA), las entidades fiscalizadoras superiores (EFS), la auditoría interna y otros órganos de supervisión, están cosechando los beneficios de adoptar la IA para mejorar sus operaciones, auditorías e investigaciones. A medida que se intensifica la disponibilidad de datos y aumenta la complejidad digital, estas instituciones se ven obligadas a adaptarse para seguir siendo relevantes. Ante el creciente uso de la tecnología por parte de todos los actores, incluidos aquellos que intentan evadir el escrutinio y la supervisión, es una cuestión de legitimidad que las instituciones de integridad sigan el ritmo de la evolución digital.

En los últimos años, la IA ha sido un tema especialmente importante para los actores relacionados con la integridad. Esto se debe a las rápidas mejoras técnicas que hacen que estas tecnologías estén más disponibles y sean más asequibles, así como a las posibilidades que ofrece la IA para procesar todos los tipos de datos de manera eficiente, descubrir patrones y relaciones complejos, mejorar la precisión y la precisión analíticas y, en general, aumentar las actividades de integridad actuales (OECD, 2024^[93]). Al igual que en otras áreas de política pública, la IA es principalmente valiosa para el análisis de conjuntos de datos vastos y complejos, donde las computadoras pueden ser más precisas o más rápidas que los humanos.

Situación actual

Entre los actores de la integridad, la IA tiene numerosas aplicaciones potenciales, que actualmente se están utilizando, probando o que se espera que surjan en los próximos años. Las actividades principales se han clasificado en tres categorías en función de su posible impacto: detección de actividades fraudulentas; aumento de la eficiencia en la gestión y síntesis de conocimientos; y mejora de la analítica predictiva y pronósticos.

Detección de actividades fraudulentas

Una de las primeras y todavía más profundas aplicaciones de la IA por parte de los actores de la integridad radica en detectar actividades fraudulentas, transformando los métodos analíticos tradicionales en enfoques sofisticados basados en datos. Los algoritmos de IA se destacan por aplicar técnicas estadísticas para identificar valores atípicos, patrones, transacciones y comportamientos que se desvían de las normas establecidas que justifican una mayor investigación humana (OECD, 2024^[93]; Taşdöken, 2024^[94]). Estas herramientas son capaces de analizar datos numéricos de poblaciones enteras, lo que supone un cambio significativo respecto al análisis tradicional basado en muestras, que permite detectar asignaciones erróneas o informes erróneos de fondos, entre otras irregularidades (Köbis, Starke and Rahwan, 2021^[95]) (Recuadro 5.27). Como ejemplos cabe citar:

- Recientemente, la OCDE colaboró con la Contraloría General de la Administración del Estado (IGAE) española en una iniciativa de prueba de concepto que empleaba analítica avanzada y ML para detectar anomalías y patrones indicativos de posibles riesgos de corrupción o fraude (OECD, 2021^[96]).

- El Tribunal de Cuentas Europeo (TCE), en el marco de una auditoría sobre las actividades de defensa de intereses de la UE, ha utilizado algoritmos ML para comprobar el registro de transparencia de la Comisión Europea e identificar valores atípicos con el objetivo de aislar una muestra más interesante basada en el riesgo para su análisis por los auditores (ECA, 2024^[97]).
- En Brasil, la herramienta Alice analiza los procesos diarios de compra y adquisición para detectar áreas de riesgo e incoherencias. Los patrones inusuales activan una alerta que suspende la compra, que se marca para más consultas (Recuadro 5.25).
- En el Reino Unido, el Departamento de Trabajo y Pensiones está utilizando la IA para identificar patrones en las reclamaciones que podrían sugerir fraude o error, de modo que estas reclamaciones puedan ser revisadas por los equipos pertinentes del departamento (Recuadro 5.28).

Recuadro 5.27. La herramienta europea DATAACROS para detectar anomalías en la propiedad de empresas

En la Unión Europea, el Proyecto DATAACROS ha desarrollado una herramienta innovadora impulsada por IA para detectar anomalías en las estructuras de propiedad de las empresas que puedan indicar riesgos de corrupción, lavado de dinero y otros delitos financieros. Financiado por el Fondo de Seguridad Interior - Policía de la UE y coordinado por Transcrime en la Università Cattolica del Sacro Cuore, el proyecto aborda la necesidad crítica de herramientas analíticas avanzadas en los esfuerzos de supervisión pública e integridad.

El prototipo de la herramienta DATAACROS consta de dos componentes principales: el Área Restringida para usuarios autorizados como agencias de aplicación de la ley y anticorrupción; y el Área Pública, accesible a todos la ciudadanía. El Área Restringida emplea algoritmos de IA para analizar las estructuras de propiedad de más de 70 millones de empresas en 44 países europeos, identificando señales de alerta y patrones ocultos que pueden indicar actividades ilícitas. Incorpora datos transfronterizos y aprovecha el amplio conocimiento de las redes delictivas, al tiempo que se adhiere a la normativa sobre privacidad y protección de datos.

La actividad de prueba de la herramienta DATAACROS en 2021 ha mostrado resultados prometedores. Los sistemas predictivos identificaron correctamente al 83 % de las empresas sancionadas y al 88 % de las empresas con propietarios sancionados. Las agencias socias informaron de una alta satisfacción (4.3/5) y probabilidad de adopción futura (4.3/5). Al proporcionar tanto capacidades de investigación a las autoridades como datos agregados transparentes para la supervisión pública, DATAACROS representa un ejemplo notable de cómo el uso de la IA puede ayudar a mejorar la integridad pública y combatir los delitos financieros.

Fuente: https://www.transcrime.it/wp-content/uploads/2021/10/Datacros_report.pdf.

Recuadro 5.28. Uso de la IA para hacer frente al fraude de beneficios en el Reino Unido

El Departamento de Trabajo y Pensiones del Reino Unido (DWP, por sus siglas en inglés) lleva utilizando sistemas de IA desde 2021-2022 para analizar datos históricos e identificar patrones relacionados con reclamaciones fraudulentas de crédito universal. Está invirtiendo 70 millones de GBP en analítica avanzada entre 2022 y 2025 para reducir el fraude y el error en las solicitudes de beneficios, con un ahorro proyectado de 1.600 millones de GBP para 2030. Los sistemas adoptados evalúan el riesgo en función de las incoherencias en los ingresos declarados, las frecuencias inusuales de las reclamaciones y los desajustes con los registros externos. Las reclamaciones de alto riesgo se marcan para su revisión manual, en la que los pagos pueden interrumpirse hasta que se complete una verificación adicional. Las principales áreas de interés incluyen las solicitudes de empleo por cuenta propia, el fraude en las prestaciones por vivienda y el capital no declarado, ámbitos en los que la detección del fraude ha sido históricamente más difícil.

Si bien estos sistemas mejoran la eficiencia, el DWP reconoce que garantizar la equidad y la transparencia sigue siendo un desafío. Para mitigar este sesgo, realiza un análisis de equidad antes del despliegue y supervisa continuamente las reclamaciones detectadas. Sin embargo, la falta de datos demográficos dificulta la evaluación completa de si determinados grupos se ven afectados de manera desproporcionada. Los trabajadores sociales conservan plena autoridad para tomar decisiones y no se les informa de por qué se ha presentado una reclamación para evitar un sesgo de confirmación. La selección aleatoria de reclamaciones también se utiliza para probar la precisión del sistema y evitar la dependencia excesiva de la automatización.

Para reforzar la confianza pública, el DWP se ha comprometido a publicar informes anuales sobre el impacto de la IA en la detección de fraudes, en particular en lo que respecta a los grupos vulnerables. Se están realizando esfuerzos para perfeccionar el análisis de justicia mejorando la recopilación de datos y ajustando los métodos de clasificación de riesgos cuando sea necesario.

Fuente: (UK National Audit Office, 2023^[98]).

El análisis de redes es otra técnica valiosa para detectar posibles fraudes o escenarios de riesgo similar. Implica examinar las interacciones entre las entidades dentro de un sistema, como los funcionarios públicos, los contratistas y los proveedores, para descubrir y representar visualmente relaciones potencialmente ocultas. Este método es especialmente útil para identificar conflictos de intereses, así como para detectar actividades de corrupción que antes eran difíciles o imposibles de detectar (debido al volumen y la complejidad de los datos), y proporciona información crucial que ayuda a mantener la integridad de las operaciones públicas. Por ejemplo, en una licitación, el reconocimiento de patrones puede ayudar a identificar secuencias de transacciones que pueden ser indicativas de colusión o sobornos²². La lucha contra la corrupción en este ámbito se analiza con más detalle en la sección "La IA en la contratación pública".

Mejoramiento de la gestión y síntesis del conocimiento

Para los actores de la integridad, que dedican gran parte de su jornada laboral a leer, procesar y extraer significado de la documentación basada en texto, la IA puede ayudar en la gestión del conocimiento y el análisis de texto. Esto puede incluir, por ejemplo, extraer significado y automatizar la recuperación de hechos de documentos.

Dentro del ámbito de la NLP, el reconocimiento de entidades con nombre permite la detección y clasificación rápidas de información esencial en textos y otros formatos de datos no estructurados, como los nombres de personas, lugares, fechas y organizaciones, a una velocidad y escala que habría sido imposible recuperar manualmente. El reconocimiento de entidades con nombre se utiliza en chatbots,

análisis de sentimientos, extracción de información y detección de fraudes. Brasil, por ejemplo, lo integró en el Tribunal Federal de Cuentas (TCU) (Odilla, 2023^[99]). La traducción automática también está facilitando el acceso a una mejor información textual con mayor rapidez, mejorando así la efectividad de las auditorías e investigaciones.

El rápido desarrollo de GenAI también se está utilizando para mejorar la eficiencia. Los LLM pueden ser de gran ayuda al permitir la lectura, interpretación, resumen y categorización de grandes cantidades de información textual con un alto grado de precisión y en una fracción del tiempo de un ser humano. Los actores de la integridad están ampliando su uso de los programas de aprendizaje permanente para aplicaciones internas y de cara al público, haciendo que no solo su trabajo sea más eficiente, sino también que el público esté más involucrado y sea más consciente de sus esfuerzos. La creación de ChatTCU en Brasil es un ejemplo (Recuadro 5.29). Otros ejemplos son:

- Los actores de la integridad en Finlandia, Francia, Grecia, Lituania, el Reino Unido y en las instituciones de la UE están utilizando los grandes modelos de aprendizaje para ayudar en la redacción de documentos, el análisis de hojas de cálculo, la síntesis de textos, el tratamiento de informes de corrupción, el análisis de la jurisprudencia y para su descubrimiento en la auditoría de rendimiento (Ugale and Hall, 2024^[89]).
- El Tribunal de Cuentas de los Países Bajos está llevando a cabo una prueba piloto de un sistema GenAI para que la ciudadanía y otras partes interesadas puedan desplazarse por los informes públicos y encontrar respuestas y fuentes a sus preguntas sobre el trabajo de auditoría pública²³.
- El ECA está llevando a cabo un proyecto piloto con otras tres EFS (Alemania, Países Bajos y Suecia) para mejorar las relaciones interinstitucionales y el intercambio de conocimientos utilizando los grandes modelos de aprendizaje para analizar los informes de auditoría nacionales existentes (en todas las lenguas europeas), automatizar la traducción y crear fichas informativas en inglés y resúmenes orientados²⁴.

Recuadro 5.29. Brasil mejora sus procesos de auditoría mediante IA generativa con ChatTCU

El Tribunal Federal de Cuentas brasileño (TCU) ha sido pionero en el uso de la IA generativa en actividades de auditoría y supervisión a través de ChatTCU, un asistente impulsado por la IA que se puso en marcha en febrero de 2023. Desarrollado como una iniciativa institucional, ChatTCU se integra con los sistemas internos de TCU, proporcionando a los auditores acceso en tiempo real a resúmenes de casos, orientación regulatoria y apoyo administrativo. A diciembre de 2023, la herramienta contaba con más de 1.400 usuarios, lo que demuestra su rápida adopción en toda la institución.

ChatTCU, que actualmente funciona con el GPT-4 de 32k, mejora la eficiencia al permitir a los auditores recuperar y sintetizar información rápidamente, reduciendo la carga de trabajo manual y manteniendo la precisión y seguridad. Alojada en Microsoft Azure, la plataforma garantiza la protección de los datos y evita el intercambio no autorizado de información confidencial con proveedores externos de IA. Entre los desarrollos futuros se incluyen una mayor integración de los sistemas y la automatización de los flujos de trabajo, así como una mayor integración de la IA en los procesos de auditoría.

Más allá de su aplicación interna, ChatTCU ahora forma parte de un esfuerzo internacional de transferencia de tecnología. En octubre de 2024, la TCU firmó un acuerdo con la Entidad Fiscalizadora Superior (EFS) de Honduras, lo que la convierte en el primer receptor internacional del código fuente de la ChatTCU. Esta alianza, establecida a través de la OLACEFS, refleja una estrategia más amplia para fortalecer la adopción de la IA en las instituciones de auditoría de ALC. Desde su lanzamiento, 33 instituciones brasileñas también han comenzado a adaptar ChatTCU para sus necesidades específicas, destacando su escalabilidad y adaptabilidad.

Fuente: (Ugale and Hall, 2024^[89]).

Con la entrada de GenAI en el discurso popular a finales de 2022, los actores de la integridad han tenido poco tiempo para comprender la oportunidad que presenta la IA generativa para su trabajo, por no hablar de integrarla completamente en sus actividades. Cuando la OCDE (2024^[89]), encuestó el año pasado a los actores implicados en la lucha contra la corrupción, los encuestados calificaron en general a sus organizaciones como «en fase inicial de madurez» en lo que respecta al uso de la IA generativa y los LLM, en el caso de que se utilizaran, a pesar de que se esté convirtiendo en un área de interés prioritario.

Analítica preventiva y predictiva, pronósticos, prospectiva

Los sistemas de IA predictiva pueden ayudar a anticipar posibles riesgos de corrupción y fraude, entre otros, permitiendo priorizar casos para un examen humano más detallado. Si bien muchos órganos de integridad carecen del mandato para llevar a cabo acciones ex ante, identificar los riesgos crecientes de manera temprana puede ayudar a mitigarlos de manera más efectiva. Aunque las acciones ex ante a menudo están fuera del mandato de muchos órganos de integridad, anticipar un aumento de los riesgos asociados con la corrupción y el fraude antes de que se manifiesten puede ayudar a mitigar estos riesgos. Por ejemplo, en Colombia, la Universidad del Rosario desarrolló VigIA para analizar los contratos de la Alcaldía de Bogotá y señalar a quienes corrían un alto riesgo de corrupción, anticipando posibles conductas ilegales. La herramienta ha sido adoptada por la Ciudad de Bogotá (Salazar, Pérez and Gallego, 2024^[100]).

Otros investigadores demostraron cómo se pueden adoptar técnicas de IA y ciencia de datos para predecir la corrupción. Por ejemplo, en Brasil, los investigadores demostraron que, en comparación con las inspecciones aleatorias, un sistema de IA que utilizara datos presupuestarios como predictores podría detectar casi el doble de municipios corruptos con la misma tasa de inspección (Ash, Galletta and Giommoni, 2020^[101]). En Italia, los investigadores utilizaron archivos policiales y aplicaron algoritmos de IA para predecir delitos de cuello blanco, demostrando cómo los pronósticos algorítmicos podrían apoyar la política anticorrupción dirigida a objetivos (de Blasio, D'Ignazio and Letta, 2022^[102]). En España, un equipo de investigación también desarrolló un sistema de IA que sirve como sistema de alerta temprana para predecir la corrupción pública en función de factores económicos y políticos en las provincias españolas (López-Iturriaga and Sanz, 2017^[73]). La transferencia de estos enfoques de la investigación a los gobiernos sigue siendo limitada, pero va en aumento. Por ejemplo, en Lituania, el Servicio Especial de Investigación (STT) está desarrollando una herramienta basada en IA para ayudar a los funcionarios de prevención de la corrupción con sus análisis legislativos (Recuadro 5.30).

Recuadro 5.30. Prevención de la corrupción mediante análisis de proyectos de ley en Lituania

El Servicio Especial de Investigación de la República de Lituania (STT), el principal organismo anticorrupción del país está desarrollando y probando actualmente una herramienta para ayudar a los funcionarios encargados de la prevención de la corrupción a evaluar las leyes promulgadas y a redactar proyectos de ley con fines anticorrupción.

La herramienta utiliza modelos de aprendizaje permanente específicamente formados en datos sobre riesgos de corrupción y metodologías de evaluación de riesgos. Mediante la identificación de posibles factores de riesgo de corrupción legal (por ejemplo, lagunas, procedimientos y medidas insuficientes o débiles), ayuda a los funcionarios a analizar rápida y eficazmente grandes volúmenes de legislación con el fin de identificar posibles deficiencias o vulnerabilidades de los actos jurídicos.

Al utilizar esta herramienta, el STT trata de reforzar las salvaguardias legislativas y garantizar que, cuando se adopte un proyecto de acto jurídico, se consideren cuidadosamente las posibles consecuencias de su aplicación.

Fuente: Trabajo en curso con la OCDE y el STT (informe de próxima publicación).

Evidencia de impacto

La IA es muy prometedora para los actores de la integridad, pero su impacto empírico sigue siendo difícil de medir en esta etapa. La evidencia anecdótica muestra que la IA mejora la eficiencia operativa al ahorrar tiempo, permitir automatizaciones inteligentes y aumentar la precisión de los análisis. Los casos de uso en los que la NLP, incluidos los grandes programas de aprendizaje, resume, traduce o extrae información de fuentes basadas en texto se encuentran entre aquellos en los que el valor añadido de la IA para las actividades relacionadas con la integridad es más fácil de ver y evaluar. Sin embargo, cuantificar el impacto general de la IA en este ámbito es complejo. En primer lugar, la madurez en este ámbito suele ser baja; muchas iniciativas de IA de organismos de integridad siguen en fase exploratoria (prueba de conceptos, proyectos piloto) y todavía no han alcanzado la escala deseada.

Además, la evaluación del estado actual de la IA en este ámbito se ve obstaculizada por la falta de transparencia y por la falta de intercambio de información y de lecciones aprendidas entre instituciones y a través de las fronteras. La información sigue estando muy fragmentada, y el intercambio de casos prácticos, mejores prácticas y conocimientos sobre las herramientas de IA en materia de integridad sigue planteando dificultades. Los desafíos surgen de una combinación de factores, como las cuestiones de confianza en torno a la confidencialidad y la sensibilidad de las actividades de supervisión e investigaciones, y las diferencias culturales, así como los distintos niveles de madurez institucional para adoptar la IA. Además, el rápido ritmo de la evolución tecnológica, unido a los limitados recursos en términos de tiempo y experiencia, complica aún más los esfuerzos por compartir información y aprender de diversos casos prácticos. Esta falta de apertura dificulta la capacidad de evaluar y comprender la efectividad de las aplicaciones de la IA y aprender de diferentes instituciones y contextos.

Gestión de riesgos y desafíos

Riesgos asociados

- Datos inadecuados o sesgados en los sistemas de IA.
- Falta de transparencia y explicabilidad.
- Uso indebido o cuestionable de la IA.

Al desplegar la IA en los esfuerzos de integridad pública y anticorrupción, es fundamental abordar los **posibles errores que pueden estar integrados en los sistemas de IA, como los causados por datos inadecuados o sesgados** (Köbis, Starke and Rahwan, 2022^[103]). Esto es especialmente importante en ámbitos como la detección de fraudes y el análisis de conflictos de intereses, en los que la equidad y la precisión son fundamentales. Continúan los esfuerzos para identificar indicadores objetivos que puedan servir como indicadores de la corrupción, un fenómeno que, por definición, está oculto. Si bien es lo más práctico, la dependencia de datos históricos (casos o decisiones pasadas) para entrenar algoritmos puede perpetuar errores. Esto podría dar lugar a percepciones sesgadas y limitar la posibilidad de descubrir dinámicas nuevas y en evolución de comportamiento corrupto, lo que podría dar lugar a prácticas desleales de selección o supervisión, socavando la credibilidad de los esfuerzos en pro de la integridad. Para contrarrestar este sesgo, los sistemas de IA deben desarrollarse con controles y equilibrios sólidos, que incluyan pruebas rigurosas, un seguimiento continuo y mecanismos de retroalimentación que permitan identificar y rectificar errores involuntarios y los daños resultantes.

La opacidad de muchos sistemas avanzados de IA también plantea importantes desafíos a los actores que se ocupan de la integridad. Si el público percibe que las acciones y decisiones del gobierno son opacas y percibe que se basan en **algoritmos de "caja negra" —donde los resultados generados por la IA carecen de explicabilidad y transparencia**— es probable que la ciudadanía se resistan al uso de la IA o desconfíen de él (OECD, 2024^[93]). En la profesión de auditor, en particular, su uso para generar

información que se utilizará como prueba de auditoría sigue siendo un tema controvertido (Mpfu, 2023^[104]). Plantea dudas sobre su razonabilidad, confiabilidad y cumplimiento de las normas profesionales y reglamentarias (como las Normas Internacionales de Auditoría o ISAs) que aún no se han actualizado para tener en cuenta los recientes cambios tecnológicos.

La integración de las tecnologías de IA en los organismos de integridad plantea importantes preocupaciones en materia de responsabilidad, ética y protección de datos, especialmente en lo que respecta a los datos personales. El **posible uso indebido de los datos**, incluida la fusión inadecuada de datos o la divulgación pública de información confidencial, plantea graves desafíos éticos y legales. Es fundamental garantizar la protección de la información sensible, como las declaraciones de intereses. Por lo tanto, las instituciones deberían guiarse por la protección de datos y las normas y regulaciones éticas al implementar la IA. Además, los organismos de integridad deben adoptar marcos sólidos de gobernanza de datos e implementar medidas integrales de protección de datos para mitigar los riesgos de privacidad, incluida la realización de auditorías periódicas, el empleo de técnicas avanzadas de cifrado y el fomento de una cultura de uso ético de los datos en sus organizaciones (véase el Capítulo 4, sección "Crear una base de datos sólida").

Desafíos de implementación

- Brechas de habilidades.
- Falta de datos de alta calidad y capacidad para compartirlos.
- Falta de directrices exhaustivas y marcos prácticos.
- Aversión al riesgo.
- Elevados costos de adopción y escalabilidad de la IA.

El informe de la OCDE (2024^[89]) ha puesto de manifiesto que **la escasez de habilidades y experiencia relevantes** son uno de los mayores desafíos en la adopción de la IA en la integridad y la supervisión públicas. Esto incluye competencias técnicas (tanto avanzadas en el ámbito digital como calificación del personal en general), así como habilidades en materia de gestión del cambio, gobernanza de datos y liderazgo.

El acceso a **datos exhaustivos y de alta calidad es un requisito previo** para el uso efectiva de la IA por parte de cualquier institución, pero los actores de la integridad a menudo carecen de autonomía en los datos. Si bien algunos actores de la integridad "son dueños" de los datos que utilizan (por ejemplo, las bases de datos de declaraciones de intereses), muchos de ellos, incluidos los ACA y las EFS, suelen depender de datos generados externamente sobre los que tienen poco o ningún control. Dependen de las normas y regulaciones gubernamentales, las prácticas históricas y los memorandos de entendimiento (MoU) con otras instituciones para adquirir los datos necesarios para sus actividades. Esto significa que los organismos de integridad a menudo se enfrentan a retrasos en el acceso a los datos necesarios, lo que dificulta el análisis en tiempo real y la toma de decisiones oportuna.

La incoherencia y la baja calidad de los datos también pueden obstaculizar gravemente el desarrollo y la efectividad de las herramientas de IA, lo que dificulta la creación de sistemas confiables y sólidos. Los organismos de integridad suelen depender de datos de fuentes externas, por lo que la disponibilidad y la precisión de estos datos son fundamentales para lograr una implementación exitosa. La calidad de los datos obtenidos de fuentes externas suele ser desconocida o no estar controlada hasta que se reciben, lo que puede dar lugar a inexactitudes e incoherencias que pueden limitar o complicar el desarrollo de la IA y la formación de los sistemas. Dado que la calidad de los datos es fundamental para el desarrollo y uso efectivos de la IA, es necesario que los marcos regulatorios faciliten el intercambio de datos y la colaboración entre organismos gubernamentales y entre sectores, según proceda y de conformidad con las normas de protección de datos y privacidad, y que se establezcan estrategias y sistemas para garantizar la calidad y la integridad de los datos.

La interoperabilidad de los datos también supone una barrera a la hora de compartirlos o combinarlos. La capacidad para cruzar referencias y analizar las relaciones dentro de los datos es fundamental para descubrir cuestiones como los conflictos de intereses. Sin embargo, en la actualidad, la madurez de la mayoría de las plataformas de datos de los gobiernos significa que la comprobación cruzada de la información es extremadamente difícil, si no imposible. Las limitaciones regulatorias —por motivos de seguridad o protección— pueden obstaculizar aún más la capacidad de integrar diferentes fuentes de datos, limitando así la posibilidad de una supervisión efectiva e integral.

Tal vez debido a la naturaleza del trabajo, las instituciones de integridad suelen ser reacias al riesgo. Muchos actores de la integridad desean integrar tecnologías avanzadas, pero no están seguros de por dónde empezar o cómo proceder. Esta aprensión se ve reforzada por el **miedo a cometer errores en el proceso de adopción**. Las directrices gubernamentales suelen centrarse en los consejos de precaución, haciendo hincapié en lo que no se debe hacer, en lugar de proporcionar medidas concretas y viables para la implementación efectiva de la IA. Esta falta de asesoramiento práctico y viable deja a muchos órganos de integridad sin una vía clara hacia adelante, lo que estanca el progreso y la innovación. Una **orientación exhaustiva y marcos prácticos** sobre el uso confiable de la IA para los actores relacionados con la integridad, de los que a menudo carecen, podrían ayudar.

Por último, el **costo de implementación y mantenimiento de las herramientas de IA plantea un reto considerable**, especialmente para las instituciones con presupuestos y conocimientos técnicos limitados. La complejidad de garantizar medidas sólidas de mitigación de los riesgos ciberneticos complica aún más este panorama. En respuesta, la colaboración entre instituciones se vuelve esencial. Al utilizar recursos centralizados, como entornos de prueba de IA gestionados por el gobierno o herramientas compartidas, los organismos de integridad pueden reducir costos y, al mismo tiempo, acceder a tecnologías avanzadas.

Potencial sin explotar y camino a seguir

De cara al futuro, los sistemas de aprendizaje adaptativo podrían perfeccionar continuamente sus predicciones en función de nuevos datos, ya que pueden desempeñar un papel importante en la actualización dinámica de las evaluaciones de riesgos, lo que permitirá a los organismos de integridad anticiparse a las posibles amenazas en evolución. La IA, en particular su capacidad para analizar y reconocer patrones en datos no numéricos, también puede ser la fuente de futuras mejoras en la lucha contra la corrupción. Al combinarlo con la tecnología de sensores e imágenes (satélites, drones, aeronaves que captan imágenes u otros datos como lecturas térmicas o señales de radar), la IA crea nuevas oportunidades para monitorear y analizar patrones a gran escala y comprender las actividades y dinámicas relacionadas con la corrupción (Zinnbauer, 2025^[105]).

Más adelante, el modelado basado en agentes (ABM, por sus siglas en inglés) realizado por entidades modeladoras (funcionarios, empresas, etc.) y sus interacciones, podría simular cómo evolucionan la corrupción y la influencia en diferentes condiciones (U4, 2025^[106]). Esto podría ayudar a evaluar el impacto de las políticas, leyes y estrategias anticorrupción o de integridad antes de su implementación.

El análisis de redes impulsado por la IA también podría ayudar a mapear de forma más eficiente y exhaustiva las relaciones entre los grupos de presión, los políticos, las empresas y otros agentes, revelando patrones de influencia, conflictos de intereses y dinámicas de puertas giratorias. Si bien la integración de las tecnologías de IA en los esfuerzos de integridad pública y anticorrupción ofrece un gran potencial, garantizar su uso efectivo y ético requiere una cuidadosa atención a consideraciones políticas clave que promuevan una adopción confiable. Un reciente estudio de la OCDE (2024^[89]) proporcionó tres recomendaciones a los agentes responsables de la integridad a la hora de explorar el uso de la IA:

- Empezar por incorporar IA generativa en áreas y procesos de bajo riesgo. Este enfoque puede ayudar a crear capacidad en áreas en las que los errores son menos costosos, ya sea

financieramente o desde la perspectiva del cumplimiento, antes de que se amplíen a tareas más arriesgadas y más analíticas, incluidas las que requieren más recursos financieros.

- Considerar los requisitos de TI tanto para el pilotaje como para la ampliación. Considerar qué recursos informáticos y de almacenamiento, capacidades de almacenamiento y gestión de datos pueden ser necesarios y asegurarse de que las decisiones tempranas no complican excesivamente los planes futuros.
- Recurrir a los datos generados internamente o abiertos para demostrar su valor y lograr resultados rápidos. Este enfoque es de bajo costo y ayuda a demostrar los casos de uso con mayor rapidez.

A medida que los gobiernos continúan explorando la IA, deben establecer mecanismos sólidos de transparencia y rendición de cuentas para garantizar que las decisiones impulsadas por la IA no solo sean precisas, sino que también se interpreten y documenten adecuadamente. Por ejemplo, si bien los sistemas de IA pueden proporcionar información valiosa al procesar grandes cantidades de datos, y los sistemas futuros podrían ser cada vez más autónomos, la autoridad última sobre las decisiones sensibles debe permanecer en manos de los operadores humanos para preservar la equidad, el juicio profesional (o el escepticismo) y, en última instancia, mantener la confianza pública. Los organismos de integridad pública deben tratar de garantizar que: el razonamiento detrás de las decisiones sea comprensible tanto para los expertos como para el público general; existan mecanismos de recurso y de impugnación; y, en la medida de lo posible, los procesos de IA sean explicables. Las pruebas generadas por la IA deben cumplir los requisitos de documentación para proporcionar una pista de auditoría adecuada, incluido el cumplimiento de las normas actuales para demostrar cómo se obtuvieron y evaluaron las pruebas.

El intercambio de conocimientos también es crucial para la correcta implementación y ampliación de las tecnologías de IA en funciones de integridad pública. Las instituciones deben participar y crear plataformas que permitan el intercambio de mejores prácticas, estrategias, lecciones aprendidas e incluso modelos, códigos y datos. El intercambio ayudará a construir una comprensión más integral de la efectividad de la IA en los roles de integridad e impulsará el desarrollo de mejores prácticas que puedan adoptarse ampliamente. Ejemplos de ello son iniciativas como la Comunidad de Práctica en Tecnología y Análisis de la División Anticorrupción del Gobierno de la OCDE o la iniciativa Conexión Tecnológica para la Integridad²⁵. Estos foros reúnen a profesionales de los sectores público y privado para debatir desafíos, explorar soluciones innovadoras y desarrollar estrategias colectivas para mejorar la adopción de la IA y otras tecnologías. Esto puede permitir un mejor flujo de información, la puesta en común de recursos, un aprendizaje rápido y un desarrollo y aplicación más rápidos de metodologías y soluciones.

Sobre la base de las directrices existentes sobre el uso ético y responsable de la IA, los gobiernos deben ir más allá de las directrices generales y proporcionar orientaciones y marcos más detallados y viables para implementar las tecnologías de IA. Esto puede incluir instrucciones detalladas paso a paso, estudios de casos y plantillas que las instituciones puedan adaptar a sus necesidades específicas. Estas orientaciones también podrían servir de orientación para identificar y mitigar los sesgos en los datos, utilizar técnicas explicables de IA, dotar de transparencia a los procesos de toma de decisiones y cómo incorporar mecanismos de reparación.

Una base de datos unificada e interoperable es esencial para el uso efectivo de la IA. Los gobiernos deben racionalizar sus operaciones de datos para garantizar que las bases de datos y los sistemas puedan comunicarse e integrarse sin problemas. Esto implica crear formatos de datos estandarizados, establecer marcos de gobernanza de datos sólidos e invertir en infraestructura que admita la interoperabilidad de los datos. Al promover la estandarización y la integración de los datos, así como políticas para el intercambio y el cruce seguro de datos, los gobiernos pueden mejorar la capacidad de los sistemas de IA para analizar los datos de forma exhaustiva, identificar patrones y proporcionar información procesable en apoyo de la integridad pública.

IA en la evaluación de políticas públicas

La evaluación de las políticas públicas aporta datos empíricos cruciales para ayudar a los responsables de formular políticas a comprender qué funciona, para quién y en qué circunstancias (OECD, 2025^[107]; 2020^[108]). La IA puede desempeñar un papel importante en el apoyo a la evaluación de políticas públicas. El aumento exponencial de los datos generados a nivel mundial, junto con el rápido desarrollo de nuevas tecnologías que reducen los costos de almacenamiento e informática, está impulsando la innovación en técnicas que pueden capturar, analizar y visualizar instantáneamente estos vastos repositorios de datos, enriqueciendo las evaluaciones (Petersson et al., 2017^[109]) (Rinaldi, Giuffrida and Negrete, 2017^[110]). Esto puede tener consecuencias importantes sobre qué y cómo evalúan los evaluadores.

La IA tiene el potencial de acelerar y automatizar tareas esenciales, como la recopilación y el análisis de datos, y de apoyar a los evaluadores en diferentes tareas de gestión, accediendo a un amplio conjunto de datos internos y externos o sintetizando resultados. La IA también puede respaldar las evaluaciones ex ante mediante la creación de sistemas predictivos y simulaciones que ayuden a los responsables de la formulación de políticas a anticipar los posibles impactos antes de su implementación (Bénassy-Quéré, 2022^[111]). Sin embargo, su uso dentro del gobierno para la evaluación de políticas ha sido limitado y ha avanzado más lentamente que en otras funciones analizadas en este capítulo.

Situación actual

Apoyo al diseño e implementación de la evaluación

La IA puede ayudar a los evaluadores de políticas a procesar una gran cantidad de contenido que puede ser útil para diseñar e implementar evaluaciones, ya sea ex ante o ex post. Esto abarca desde enfoques más sofisticados y estructurados, como el apoyo a la síntesis de la evidencia existente, hasta funciones más básicas, como proporcionar resúmenes sencillos de evaluaciones anteriores.

La síntesis de pruebas es especialmente útil para fundamentar la elaboración de evaluaciones ex ante o ex post. Estos métodos suelen implicar la combinación de los resultados de diferentes estudios que investigan el mismo tema para tener una comprensión global del tema. En el campo de la síntesis de evidencias, la minería de textos y otros tipos de herramientas de IA se han aplicado a la búsqueda de literatura, el screening de texto completo, la extracción de datos y el análisis durante más de 10 años. Los investigadores en el campo han cartografiado una variedad de herramientas que pueden apoyar las diferentes etapas de la síntesis de evidencias, particularmente para revisiones sistemáticas y hacerlas más efectivas. Ejemplos de herramientas son Rayyan, para examinar títulos y resúmenes, y Robot Reviewer, para evaluar determinados riesgos (Khalil, Ameen and Zarnegar, 2022^[112]). Se elaboraron orientaciones y recomendaciones para garantizar el uso responsable de la IA en la síntesis de pruebas (RAISE), lo que demuestra el creciente interés por aplicar estas herramientas en este ámbito (Thomas et al., 2024^[113]). Los beneficios de utilizar la IA para la síntesis de pruebas se basan en la capacidad de acceder a una gran cantidad de bibliografía y procesar la información con mayor rapidez que otros enfoques alternativos. De hecho, el tiempo promedio para realizar una síntesis de pruebas es de 15 meses, pero el uso de la IA puede reducir radicalmente algunos de los pasos (Blaizot et al., 2022^[114]). Por ejemplo, un riesgo de sesgo en la evaluación de datos de 30 artículos de ensayos controlados aleatorizados (ECA) puede realizarse con precisión utilizando un LLM en un promedio de 53 segundos.

Esto habría requerido mucho más tiempo a un ser humano; un estudio reciente proporcionó una estimación de alrededor de 28 minutos para cada estudio en la revisión sistemática utilizando una herramienta para ensayos aleatorizados (RoB 2), que ahora se considera el estándar de referencia (Minozzi et al., 2020^[115]; Lai et al., 2024^[116]; Odell, 2024^[117]). Del mismo modo, los evaluadores del Banco Mundial, que utilizan minería de texto e IA, han podido duplicar el tamaño de la base empírica que utilizan para tomar determinadas decisiones sobre programas. Probablemente eso habría sido imposible utilizando los enfoques tradicionales de identificación de carteras, sentando las bases para las evaluaciones de pertinencia y efectividad en el análisis de carteras que llevaron a cabo (Bohni Nielsen, Mazzeo Rinaldi and Petersson, 2024^[118]).

Apoyo a la analítica

La IA puede ayudar a procesar grandes cantidades de datos y textos largos, como informes de documentación y entrevistas, para identificar patrones, lo que puede ser especialmente útil para las evaluaciones. Al utilizar la minería de texto habilitada para la NLP, la IA puede ayudar a los evaluadores a comprender las conclusiones y proporcionar evaluaciones detalladas que, de otro modo, requerirían mucho tiempo (Næss et al., 2025^[119]).

Las aplicaciones de análisis cuantitativo de texto para la evaluación de programas son prometedoras, ya que la IA permite un análisis más rápido de un gran número de documentos en comparación con las herramientas clásicas de análisis de texto (Gatto and Bundi, 2025^[120]). Por ejemplo, analizar las respuestas a preguntas abiertas necesarias para reconstruir la teoría de un programa podría implicar diferentes métodos que requieren mucho tiempo, como identificar afirmaciones que tienen una forma específica y reformularlas en proposiciones condicionales "si-entonces" (Leeuw, 2003^[121]), que pueden facilitarse mediante técnicas novedosas, como los modelos temáticos (Gatto and Bundi, 2025^[120]). Si bien el análisis de texto convencional puede contribuir a la evaluación de varias maneras empíricas y conceptuales, como medir las preferencias de las partes interesadas o identificar teorías de programas subyacentes, el análisis cuantitativo de texto basado en IA puede crear aún más oportunidades. Por ejemplo, el Banco Mundial utilizó el ML no supervisado para analizar 392 informes de proyectos en 64 países receptores de ayuda, identificando con éxito factores novedosos y perspicaces que influyen en el éxito y el fracaso de los proyectos (Franzen et al., 2022^[122]). Además, el análisis de textos con IA resulta especialmente útil para identificar temas subyacentes en informes o programas de políticas existentes, ya que permite cartografiar conceptos clave incorporados en la documentación pertinente, una técnica también conocida como modelización temática (Cintron and Montrosse-Moorhead, 2021^[123]). Algunos de estos métodos pueden utilizarse para comprender las posiciones de diferentes partes interesadas y proporcionar información interesante para la evaluación de políticas. En algunos países, agentes relevantes, como las EFS, están adoptando técnicas de análisis cuantitativo de textos para realizar auditorías de desempeño que comparten varias características con las evaluaciones de políticas. En Noruega, por ejemplo, se utilizaron aplicaciones de minería de textos y ML para realizar una auditoría del rendimiento del tratamiento policial de la ciberdelincuencia (Recuadro 5.31). En el ámbito de la política ambiental, la OCDE se asoció con importantes instituciones de investigación que utilizaron la IA para llevar a cabo la primera evaluación global integral de las medidas de política ambiental, analizando más de 1 500 intervenciones de política en cuatro sectores de 1998 a 2022 en 41 países de seis continentes (Recuadro 5.32).

Más allá de estas técnicas de análisis de textos, la IA también tiene potencial para mejorar la inferencia causal en la evaluación de políticas, para apoyar diseños quasi experimentales que se basan en modelos probabilísticos, ayudando a generar algunos de los datos que faltan que se necesitan para sofisticados métodos no paramétricos y variables instrumentales para simular diversos escenarios de impacto. Esto puede ayudar a mover las fronteras de la investigación y permitir evaluaciones mucho más potentes sin la necesidad de costosos enfoques de ensayos controlados aleatorios (Miller, 2020^[124]).

Recuadro 5.31. Noruega utiliza minería de textos y ML en inspecciones policiales

En 2018, la Oficina del Auditor General de Noruega creó un Laboratorio de Innovación destinado a integrar la ciencia de datos en la auditoría. El laboratorio incluía a científicos de datos con experiencia en áreas como la evaluación/auditoría del rendimiento, codificación y ML, para ayudar a los auditores automatizando tareas, ayudando en la adquisición de datos (como el web scraping) y el análisis (por ejemplo, la minería de texto), así como proporcionando formación interna en herramientas de codificación y análisis.

En 2021, la Oficina del Auditor General de Noruega llevó a cabo una auditoría del desempeño de las actividades de la policía nacional noruega en la lucha contra la ciberdelincuencia. El objetivo de la auditoría era examinar si la policía noruega había mejorado su capacidad para combatir la delincuencia cibernética. Durante la auditoría, la colaboración entre los científicos de datos y los auditores fue clave para garantizar un uso hábil de los nuevos métodos en dichas auditorías, aportando contribuciones en sus respectivas competencias. Si bien los auditores abordaron varias tareas, como el diseño de investigaciones, la recopilación y el análisis de datos, así como la redacción, los científicos de datos desempeñaron un papel clave en la aplicación de métodos avanzados de minería de texto y ML para clasificar los casos penales. Más concretamente, durante la auditoría, se utilizaron 1.000 casos codificados para formar a un sistema de ML que aplicaba técnicas de minería de texto para extraer el texto de 334 544 casos de diferentes tipos de delitos (por ejemplo, lucro/robo ilícito, tráfico, violencia, delitos financieros). La tarea de clasificar los casos penales como delitos cibernéticos o delitos no cibernéticos es crucial para evaluar la comprensión y la gestión de esos delitos por parte de la policía. Los datos que poseía anteriormente la policía noruega eran poco confiables y carecían de información sobre el alcance de las investigaciones y las resoluciones de casos.

Fuente: (Næss et al., 2025^[119]),<https://www.adb.org/sites/default/files/publication/928976/governance-brief-052-digital-transformation-tax-administration-rok.pdf>.

Recuadro 5.32. Evaluación de la efectividad de las políticas ambientales con IA

Entender qué políticas reducen efectivamente las emisiones es un desafío apremiante para los gobiernos de todo el mundo. En colaboración con el Instituto de Potsdam para la Investigación del Impacto Climático (PIK), la Universidad de Oxford y la Universidad de Victoria, la OCDE ha contribuido a un estudio que utiliza la IA para evaluar la efectividad de este tipo de políticas en 41 países. Esta investigación, publicada en *Science*, hace un balance de más de 1.500 intervenciones políticas que abarcan el período comprendido entre 1998 y 2022, ofreciendo una visión sin precedentes de lo que funciona en la lucha contra el cambio climático.

El núcleo del estudio es el Marco de Medición de Acciones y Políticas Climáticas de la OCDE. Utilizando IA, el estudio identificó 63 políticas que han llevado a reducciones significativas de las emisiones. Este enfoque impulsado por la IA no solo mejora la capacidad de detectar medidas de política acertadas, sino que también revela patrones y combinaciones de políticas que, de otro modo, podrían pasar desapercibidos. Una de las principales conclusiones destaca que las combinaciones de políticas —en lugar de las medidas aisladas— son más efectivas para reducir las emisiones, lo que refuerza la importancia del diseño estratégico de las políticas.

Fuente: <https://www.oecd.org/en/blogs/2025/01/what-works-groundbreaking-evaluation-on-the-effectiveness-of-climate-policies.html>.

Apoyo a la gestión y comunicación

Los gestores de las evaluaciones pueden beneficiarse de una variedad de herramientas basadas en la IA desarrolladas para facilitar las actividades de apoyo, como los procesos administrativos, la redacción, la traducción y las herramientas de búsqueda. Al examinar la planificación y la gestión, las evaluaciones gubernamentales suelen subcontratarse a evaluadores externos, lo que aumenta su complejidad de gestión. Las funciones basadas en IA de las herramientas de gestión de proyectos, como monday.com o Asana, ayudan a los evaluadores a optimizar recursos y plazos, por ejemplo, recomendando una dotación de personal y una programación óptimas en función de evaluaciones anteriores, o garantizando que los recursos se asignan donde más se necesitan. Ambas plataformas utilizan la IA para ofrecer automatización y análisis, como la automatización de los flujos de trabajo, las tareas predictivas o la asistencia a través de asistentes de IA. Además, la IA generativa también puede facilitar la redacción de mandatos o tareas de gestión similares (Jacob, 2025^[125]).

Las herramientas que utilizan los programas de aprendizaje permanente pueden ayudar a mejorar la comunicación de los resultados de las evaluaciones; resumen rápidamente los informes largos en productos más breves que pueden compartirse con los responsables de la toma de decisiones o con el público. La Comisión Europea, por ejemplo, ha desarrollado una herramienta basada en el aprendizaje permanente que puede elaborar resúmenes e informes de políticas a partir de documentos elaborados en diferentes idiomas (Recuadro 5.33).

Por último, las herramientas de IA pueden ser útiles para desarrollar repositorios de evaluación digital y herramientas de gestión del conocimiento. Varios países de la OCDE han desarrollado repositorios de evaluación que garantizan que todas las evaluaciones realizadas sean fáciles de encontrar, o plataformas para describir y alinear los esfuerzos de evaluación en todo el gobierno. Un ejemplo es la plataforma Kudos de Noruega²⁶. Aunque el desarrollo manual de estos repositorios puede requerir mucho tiempo, los LLM pueden automatizar gran parte del trabajo y permitir búsquedas más precisas en un gran volumen de informes. En Francia, la Inspección General de Finanzas (IGF) está utilizando actualmente modelos de gestión de la liquidez para desarrollar una generación aumentada de recuperación interna (Recuadro 5.34) llamada Fragments, que recopila informes del IGF y del Tribunal de Cuentas francés desde 2006 y permite una búsqueda precisa en todos estos documentos.

Recuadro 5.33. Herramientas de resumen electrónico de la Comisión Europea para apoyar administraciones públicas

La Dirección General de Traducción de la Comisión Europea ha desarrollado una serie de servicios basados en IA que pueden apoyar el análisis de las políticas, incluida su evaluación. Por ejemplo, eSummary es un servicio basado en IA que puede realizar una rápida descripción general de un texto enviado y enviar una versión abreviada; utiliza algoritmos de IA para elegir dónde se encuentra el énfasis en el documento y proporcionar una síntesis relevante. eSummary está conectado a otra herramienta de traducción basada en IA (eTranslation), lo que le permite crear texto en todas las lenguas de la UE. La herramienta es accesible para una variedad de actores en toda la Unión Europea, incluidas las administraciones públicas de los Estados miembros.

Fuente: <https://language-tools.ec.europa.eu>.

Recuadro 5.34. Generación de recuperación aumentada para LLMs

La Generación Aumentada de Recuperación (RAG, por sus siglas en inglés) es una técnica desarrollada para mejorar la forma en que los LLM, como los que están detrás de los chatbots avanzados y los asistentes virtuales, manejan la información. Por diferentes razones, incluida la dependencia de datos antiguos, los grandes modelos pueden proporcionar respuestas incorrectas y puede resultar difícil comprender cómo se derivaron de una respuesta concreta. La RAG puede ayudar a abordar estos desafíos al permitir que los LLM accedan a fuentes de datos adicionales que puedan mantener actualizada la información, lo que es particularmente útil cuando se aplica a dominios o áreas de conocimiento especializadas. Para los actores gubernamentales, las RAG pueden ser un medio efectiva para delimitar sus fuentes internas de datos, mejorando al mismo tiempo la precisión, la pertinencia y la confiabilidad de los resultados de un modelo.

La RAG comienza identificando la documentación pertinente y extrayendo el texto vital de la misma. A continuación, divide este texto en partes más pequeñas y las transforma en un formato (es decir, incrustaciones) que el modelo puede entender y almacenar de forma eficiente. Estas piezas de información se mantienen en una base de datos especial (es decir, bases de datos vectoriales). Cuando alguien le hace una pregunta al modelo, éste puede buscar en esta base de datos información actualizada y precisa para agregar a lo que ya sabe antes de dar una respuesta.

En situaciones en las que es fundamental que un modelo proporcione datos actualizados y precisos, como cuando se trata de información confidencial o se necesita mantener un registro claro de las fuentes de datos, el manual de IA del Reino Unido para el Gobierno del Reino Unido (Recuadro 4.2) recomienda utilizar las RAG. Este enfoque puede ayudar a garantizar que las respuestas del modelo se basen en datos confiables, lo que lo convierte en un instrumento especialmente valioso para las organizaciones centradas en mantener altos niveles de precisión y responsabilidad.

Fuente: (Ugale and Hall, 2024^[89]).

Evidencia de impacto

Dado que el uso de la IA en la evaluación de políticas se encuentra todavía en una fase incipiente, su impacto en la práctica de evaluación de políticas sigue siendo modesto y difícil de medir. Un estudio reciente realizado a 758 consultores del Boston Consulting Group (BCG) sobre tareas complejas y realistas que exigen un alto nivel de conocimiento (como la evaluación de políticas), demostró que, de un conjunto de 18 tareas de consultoría realistas dentro de las capacidades conocidas de la IA, los consultores que utilizan IA eran mucho más productivos. En comparación con un grupo de control, en promedio, completaron un 12 % más de tareas, un 25 % más rápido y un 40 % más de calidad (Dell'Acqua et al., 2023^[120]). Sin embargo, se observó un efecto diferente en las tareas "fuera de la capacidad actual de la IA", al mostrar menos errores del grupo que no utilizaba IA. Otros estudios sobre el impacto de la IA en la síntesis de evidencias muestran resultados prometedores de la utilización de modelos de aprendizaje permanente para llevar a cabo algunos elementos de revisiones sistemáticas, como las evaluaciones del riesgo de sesgo, con un acuerdo en el juicio entre los seres humanos y los sistemas de IA que oscila entre el 41 % para el juicio global y el 71 % para la medición de resultados (Eisele et al., 2024^[127]). Al mismo tiempo, el estudio también pone de relieve que el juicio sobre la IA todavía no puede sustituir a la evaluación humana.

Gestión de riesgos y desafíos

Si bien hay muy poca investigación sobre los riesgos y desafíos del uso de la IA en la evaluación de políticas públicas (Jacob, 2025^[125]), esa investigación, junto con el trabajo de la OCDE con los gobiernos y el análisis de casos de uso individuales, han identificado varios riesgos asociados y desafíos de implementación para el uso de la IA en este campo.

Riesgos asociados

- Datos inadecuados o sesgados en los sistemas de IA.
- Sesgo de automatización.
- Falta de transparencia y explicabilidad.

El uso de la IA durante los procesos de evaluación puede reforzar algunos errores que pueden surgir en múltiples puntos de la cadena de suministro de IA. El primer riesgo de utilizar la IA en la evaluación de políticas procede de **algoritmos de formación con datos sesgados o incompletos**. Esto puede dar lugar a sistemas que generen predicciones erróneas o que refuerzen o agraven perspectivas preexistentes (Jacob, 2025^[125]). Si bien este riesgo no solo afecta a la evaluación de las políticas —y puede aplicarse a otros usos de la IA en la formulación de políticas—, su posible riesgo sobre el terreno es significativo si se tiene en cuenta su impacto en el mantenimiento o la interrupción de intervenciones potencialmente positivas o perjudiciales (Marcucci and Verhulst, 2025^[128]). Por esta razón, es esencial mitigar este riesgo en la medida de lo posible asegurándose de que los datos utilizados para entrenar estos sistemas sean de buena calidad y representativos.

Muchas personas perciben que los sistemas de IA y las decisiones que toman son neutrales e imparciales, lo que les lleva a aceptar resultados sin escrutinio, a pesar de la posibilidad de que existan imprecisiones. Esta tendencia de los operadores humanos a depender excesivamente de la automatización se conoce como "**sesgo de automatización**" (Horowitz, 2023^[129]; Alon-Barkat and Busuioc, 2022^[130]). El exceso de automatización podría reducir el papel del juicio humano y, potencialmente, simplificar en exceso supuestos sociales y económicos complejos. Esto puede hacer que los evaluadores acepten las recomendaciones propuestas por la IA sin examinar completamente los supuestos o datos subyacentes.

Además, la **falta de transparencia de determinadas herramientas de IA** puede complicar aún más las tareas de los responsables de la formulación de políticas para entender y justificar los conocimientos basados en la IA. Esto puede ser especialmente problemático en la formulación de políticas, en la que se necesita un enfoque objetivo para abordar de manera proporcionada las diferentes poblaciones y justificar las decisiones tomadas.

Desafíos de implementación

- Falta de datos de alta calidad y capacidad para compartirlos.
- Brechas de habilidades.

Así como los gobiernos se enfrentan a una serie de desafíos a la hora de promover la evaluación de políticas, también se enfrentan a desafíos específicos al intentar utilizar la IA en este ámbito (OECD/UNESCO, 2024^[131]). La **gobernanza de datos inadecuada** ha sido durante mucho tiempo un desafío para los evaluadores, lo que ha limitado la capacidad de los gobiernos para generar los datos necesarios para producir pruebas y evaluaciones (OECD, 2020^[108]). Este problema persiste a pesar de que, al adherirse a la Recomendación de la OCDE sobre la Evaluación de Políticas Públicas (2023^[132]), todos los países miembros se comprometieron a garantizar la disponibilidad de resultados, datos administrativos y de desempeño de alta calidad, oportunos, accesibles, desglosados y reutilizables para la evaluación de políticas.

La implementación de la IA requiere **competencias digitales y de cálculo**. Si bien no todos los miembros de un equipo de evaluación deben ser expertos, es importante asegurarse de que todos los evaluadores comprendan cómo la IA puede respaldar la evaluación de políticas. Por lo tanto, los evaluadores deben contar con el apoyo adecuado de equipos digitales y de ciencia de datos, que actualmente se están creando, como el de la Inspección General de Finanzas (IGF) de Francia. Los datos muestran que los evaluadores de las políticas han tardado más en adaptarse a los nuevos avances a nivel general, aunque algunos países cuentan con prácticas avanzadas. A menudo, la formación de los evaluadores en materia de desarrollo en el ámbito del análisis de macrodatos y la IA es limitada. Este es también el caso de las evaluaciones dentro del gobierno, que a menudo adolecen de capacidades analíticas internas y habilidades técnicas limitadas, lo que dificulta el desarrollo de evaluaciones de calidad (OECD, 2020^[108]).

El conjunto de herramientas para la aplicación de la *Recomendación de la OCDE sobre evaluación de políticas públicas* (2025^[107]) proporciona información sobre cómo los gobiernos pueden evaluar sus capacidades actuales de evaluación de políticas, identifica las mejores prácticas y los casos prácticos en todo el mundo e ilustra las posibles soluciones prácticas y herramientas para los responsables de la formulación de políticas y los evaluadores.

Potencial sin explotar y camino a seguir

En el ámbito de la evaluación de políticas, la IA puede realizar algunas tareas, lo que permite a los analistas gubernamentales utilizar una gama más amplia de evidencias y procesarlas más rápidamente. Si bien se identificaron algunas aplicaciones iniciales de la IA en el diseño, el análisis y la evaluación, la comunicación y la gestión de las evaluaciones, el uso de la IA en la evaluación de políticas sigue siendo limitado. Por este motivo, existen ámbitos en los que la IA puede tener un impacto significativo en la evaluación de políticas en el futuro.

En primer lugar, para apoyar el diseño de la evaluación, los chatbots podrían ayudar a los evaluadores a desarrollar su conocimiento en campos específicos. De hecho, si se les solicita, los chatbots pueden realizar varias actividades que pueden apoyar el aprendizaje. Como muestran algunos ejemplos iniciales de diseño de evaluación, también pueden apoyar el pensamiento creativo y utilizarse como herramientas útiles para el intercambio de ideas (Ferretti, 2023^[133]). Aunque estas herramientas no generen nuevas evidencias, sí pueden proporcionar nuevos conocimientos útiles para las etapas iniciales de un proceso de evaluación. Recientemente, por ejemplo, la Investigación Profunda de ChatGPT intenta automatizar una gran parte del proceso de revisión y síntesis de evidencias. Usando el razonamiento de la cadena de pensamiento (CoT, por sus siglas en inglés), herramientas como la Investigación Profunda dividen preguntas de investigación complejas en subpreguntas más pequeñas y comprensibles que responde en secuencia. Este enfoque permite al sistema elaborar un informe detallado basado en el examen de las pruebas disponibles. Estas técnicas de CoT tienen el potencial de automatizar una gran parte del proceso de revisión y síntesis de evidencias. Esto podría permitir a los investigadores que anteriormente desarrollaban algunas revisiones desde cero automatizar, asegurar la calidad y basarse en docenas de revisiones de investigación generadas por la IA.

En segundo lugar, desde una perspectiva analítica, existe un gran potencial para que la IA se utilice en mayor medida para realizar evaluaciones ex ante y ex post más ambiciosas, utilizando una gama más amplia de datos y evaluando el impacto a través de métodos quasi experimentales. Por ejemplo, la previsión del comportamiento impulsada por la IA puede analizar grandes cantidades de datos históricos y comportamientos observados para identificar patrones, anticipar decisiones y optimizar las experiencias de los usuarios mediante la integración de variables contextuales y estímulos externos. Las herramientas de ML pueden utilizarse para la predicción contrafactual en los casos en los que falta un grupo de control. Esto puede utilizarse, por ejemplo, en el caso de las evaluaciones de los precios del carbono, en las que los evaluadores de las políticas carecen de una perspectiva ex post. Un estudio propone un enfoque de evaluación de políticas que utiliza herramientas de ML y la teoría económica para la predicción contra

fáctica con el fin de analizar los costos y los impactos de las emisiones del CPS del Reino Unido, "un impuesto al carbono aplicado a todas las centrales eléctricas de combustibles fósiles" (Abrell, Kosch and Rausch, 2022^[134]).

Por último, a más largo plazo, la IA puede cambiar el enfoque de la formulación de políticas desde un ciclo de formulación de políticas y permitir que las evaluaciones se incorporen a la toma de decisiones en múltiples etapas. Dado que la IA agiliza las evaluaciones y, en cierta medida, las hace menos costosas, los académicos sugieren la posibilidad de pasar de un sistema en el que las evaluaciones suelen llegar demasiado tarde para la toma de decisiones a un enfoque en el que las pruebas de evaluación están disponibles para dar forma, ajustar y rediseñar las políticas casi en tiempo real. Esto se conoce como ciclo dinámico de políticas públicas (Jacob, 2025^[125]). Dado que los países de todo el mundo se han enfrentado a una serie de crisis en los últimos años, es esencial que los gobiernos tengan acceso a pruebas de evaluación en las etapas clave de la toma de decisiones. Las evaluaciones rápidas se elaboran para fundamentar la adopción de decisiones urgentes y se han utilizado eficazmente con este fin, por ejemplo, en Australia (Better Evaluation Knowledge, 2022^[135]). Aunque estas evaluaciones rápidas ahora se basan principalmente en datos cualitativos, la IA podría desempeñar un papel importante para que estas evaluaciones sean más sólidas y comunes en el futuro.

Sin embargo, para que la IA apoye eficazmente la evaluación, los gobiernos deben invertir en las competencias de los funcionarios públicos y en el desarrollo de una sólida infraestructura de datos. Una mayor colaboración internacional también puede potenciar el potencial de la IA en la evaluación de políticas. Los evaluadores necesitan conocer bien los posibles beneficios, riesgos y limitaciones de la IA para tomar decisiones informadas sobre cuándo y cómo utilizarla. Por este motivo, los gobiernos deben invertir en cursos de formación para evaluadores con el fin de asegurarse de que comprenden las diferentes herramientas de que disponen. Se han impartido formaciones en todos los gobiernos de la OCDE (véase el Capítulo 4, sección "Fomento de las competencias y el talento"). Sin embargo, estos se centran principalmente en el uso de la IA en la administración pública y no se adaptan al ámbito de la evaluación. Además de las formaciones, es importante apoyar la experimentación con IA y el aprendizaje práctico. El desarrollo de una red interministerial para intercambiar aplicaciones pertinentes de IA puede ser una buena forma de apoyar la adopción de la IA en diferentes tareas de evaluación. En la actualidad se están desarrollando algunas incubadoras, pero es necesario prestar más atención a la evaluación.

Como en el caso de otras áreas de política, los gobiernos deberían invertir en infraestructuras de datos pertinentes e intercambio de datos seguro y protegido (para un análisis detallado, véanse el Capítulo 4, secciones sobre "Creación de una base de datos sólida" y "Construcción de una infraestructura digital"). Algunas organizaciones gubernamentales, como el Centro Australiano de Evaluación, han elaborado guías para facilitar el descubrimiento de datos y el acceso a las actividades de apoyo a la evaluación (ACE, 2025^[136]). Algunos países de la OCDE han desarrollado formas en las que se pueden vincular y acceder a diferentes conjuntos de datos en un entorno seguro para garantizar el análisis de políticas. En Dinamarca, por ejemplo, Statistics Denmark (2025^[137]) facilita el uso de estas bases de datos a nivel micro con fines de investigación para analistas aprobados, universidades, organizaciones de investigación o ministerios. En los Países Bajos, el gobierno puso en marcha la Agenda de Datos, en la que se esbozan planes para mejorar la gestión de los datos personales, los datos abiertos y los macrodatos, aprovechando el análisis y la integración para la formulación de políticas informadas y abordando los desafíos sociales (Netherlands Ministry of the Interior and Kingdom Relations, 2019^[138]).

Por último, la IA tiene potencial en la síntesis de evidencias. Existe un llamamiento más amplio en favor de una mayor colaboración en la generación de pruebas empíricas entre países, tras las iniciativas estratégicas apoyadas por países como el Reino Unido y Australia (Halpern and Maru, 2024^[139]). Esta agenda reconoce la necesidad de una síntesis más rápida y confiable a nivel internacional, dado que la IA ya está ayudando a reducir los plazos de producción de pruebas. Esto podría ayudar a subsanar algunas de las deficiencias existentes con mayor rapidez.

IA en la participación cívica y gobierno abierto

Participación de las partes interesadas se refiere a "todas las formas en que la ciudadanía y las partes interesadas pueden participar en el ciclo de políticas públicas y en el diseño e implementación de servicios públicos" (OECD, 2017^[140]). La participación cívica, que se centra en la participación de la ciudadanía, puede reportar beneficios tanto a las administraciones públicas como a la ciudadanía. No solo genera confianza en el sector público, sino que también refuerza la democracia al ofrecer oportunidades significativas para que la ciudadanía tengan voz en el diseño e implementación de las políticas y servicios del gobierno (Paunov and Planes-Satorra, 2023^[141]; 2024^[142]; 2022^[143]). En lo que respecta a las políticas y los productos de IA, la participación temprana en el ciclo de desarrollo tecnológico enriquece la comprensión de las cuestiones, ayuda a alinear la innovación tecnológica con las necesidades de la sociedad e influye en la forma en que se mitigan los posibles daños (OECD, 2024^[144]; UK Government Office for Science, 2023^[145]).

Los gobiernos están utilizando las tecnologías digitales para ampliar las oportunidades de participación cívica al permitir nuevas formas y canales de interacción entre la ciudadanía y las instituciones públicas. Esto contribuye a la creciente tendencia de Civic Tech: el uso de tecnologías digitales para reforzar la democracia al permitir que el público esté informado, participe en la toma de decisiones y la formulación de políticas, y aumente la capacidad de respuesta y la rendición de cuentas de los gobiernos (OECD, 2024^[146]). En particular, las herramientas de IA tienen aplicaciones relevantes en la participación cívica y, más ampliamente, en las políticas y prácticas de gobierno abierto, incluido el acceso a la información, la comunicación gubernamental y la protección del espacio cívico.

Situación actual

Mejoramiento de la transparencia y el acceso a la información

La capacidad de la IA para clasificar, filtrar y resumir grandes cantidades de información podría reducir las barreras a la divulgación de información, facilitando así que los gobiernos sean transparentes para el público (OECD, 2024^[29]). Por ejemplo, el Gobierno de Estados Unidos está probando una herramienta impulsada por IA para revisar documentos clasificados para su futura divulgación en virtud de la Ley de Libertad de Información (US FOIA, 2023^[147]; MITRE, 2023^[148]). La IA también puede ayudar a la ciudadanía a comprender los complejos procesos gubernamentales y crear oportunidades para una mayor participación y escrutinio por parte de la ciudadanía y las organizaciones de la sociedad civil (OSC) (Araszkiewicz and Rodríguez-Doncel, 2019^[149]).

Los gobiernos pueden ofrecer herramientas basadas en la IA directamente al público para ayudar a las personas a navegar e interpretar grandes cantidades de datos e información, proporcionando así a la ciudadanía un acceso personalizado y accesible a la información pública. Por ejemplo, el Distrito de Columbia (EE. UU.) ha lanzado una versión beta de la plataforma DC Compass²⁷ que utiliza IA para responder a las consultas de los usuarios y generar mapas basados en datos abiertos de la ciudad. En Grecia, la plataforma opencouncil.gr utiliza la IA para transcribir automáticamente las reuniones de los

consejos locales y generar resúmenes, contenido de redes sociales y actualizaciones personalizadas de los barrios a través de aplicaciones de mensajería, lo que hace que la gobernanza local sea más accesible y comprensible para la ciudadanía²⁸. Otro ejemplo de Grecia es la plataforma DidaktorikaAI, puesta en marcha por el Centro Nacional de Documentación (EKT), que mejora la accesibilidad del conocimiento académico y científico para los responsables de la formulación de políticas y para la sociedad en general a través de una biblioteca en línea impulsada por la IA que reúne más de 50 000 publicaciones (National Documentation Center (EKT), 2025^[150]). Además, el Parlamento Europeo proporciona una herramienta de IA para ayudar a la ciudadanía a acceder a su archivo formulando preguntas en lenguaje natural (Recuadro 5.35).

Recuadro 5.35. Democratizar el acceso a los archivos del Parlamento Europeo

Los archivos del Parlamento Europeo han adoptado la IA generativa para transformar el acceso a sus extensos registros históricos. Desarrollada por la Unidad de Archivos, la innovadora plataforma Archibot integra los grandes modelos de lenguaje con técnicas RAG para permitir a los usuarios interactuar directamente con documentos digitalizados. Los ciudadanos, investigadores y responsables políticos pueden ahora realizar consultas en lenguaje natural para obtener información sobre acontecimientos históricos, cifras clave y avances institucionales a partir de un índice de más de 100 000 documentos que se remontan a 1952. La herramienta emplea métodos RAG para identificar documentos relevantes a partir de las consultas de los usuarios y sintetiza la información extraída en respuestas concisas y fácilmente comprensibles. Accesible a través del sitio web de archivos del Parlamento Europeo, la plataforma admite la traducción automática, ampliando así su alcance a los usuarios en 55 idiomas, a pesar del contenido predominantemente francés.

El proyecto, que reduce el tiempo de búsqueda y análisis de documentos en aproximadamente un 80 %, mejora significativamente la eficiencia de la investigación histórica y el análisis de datos. Además de facilitar un acceso rápido a la información legislativa y normativa, Archibot apoya la creación de informes exhaustivos y productos analíticos, que sirven como un valioso recurso para educadores, investigadores y responsables de la formulación de políticas. Esta transformación digital del acceso a los archivos no solo aborda desafíos de larga data en la navegación por vastas colecciones de registros, sino que también mejora la transparencia y el compromiso público dentro de la Unión Europea.

Fuente: <https://historicalarchives.europarl.europa.eu/en/sites/historicalarchive/home/cultural-heritage-collections/news/ai-dashboard.html>.

Gracias en parte a la mejora de la transparencia y el acceso a la información, las herramientas de IA ofrecen oportunidades para promover el espacio cívico (OECD, 2023^[151]). El espacio cívico es la piedra angular del funcionamiento de las democracias. Consiste en el conjunto de condiciones jurídicas, políticas, institucionales y prácticas que los agentes no gubernamentales necesitan para acceder a la información, expresarse, asociarse, organizarse y participar en la vida pública (OECD, 2022^[152]). Por ejemplo, en Israel, la plataforma Kol Zchut (Todos los derechos) —un repositorio exhaustivo de derechos de los residentes desarrollado por la OSC homónima junto con el Ministerio de Justicia y la Agencia Digital Nacional del país— ha integrado recientemente un chatbot de IA para permitir a los usuarios formular sus consultas en hebreo, árabe y ruso (OECD OPSI, 2019^[153]). Los expertos en IA han identificado el empoderamiento de la ciudadanía, las OSC y los interlocutores sociales (por ejemplo, los sindicatos) como uno de los beneficios potenciales más críticos del uso de la IA, y han pedido a los gobiernos que tomen más medidas para empoderar a estos actores (OECD, 2024^[29]).

La IA también puede ampliar el acceso a la información al ayudar a las personas con discapacidad que pueden enfrentarse a obstáculos para acceder a la información. Por ejemplo, la ciudadanía afectados por

deficiencias visuales o cognitivas pueden beneficiarse de la conversión de voz a texto o de imagen a voz aplicada a sitios web y portales gubernamentales, así como de interfaces accesibles impulsadas por IA con textos adaptables, fuentes, colores y contrastes (Welker, 2023^[154]). El Reino Unido creó su sitio web gov.uk para que fuera totalmente compatible con las tecnologías que utilizan la IA con el fin de mejorar su accesibilidad. La IA también puede mejorar el acceso a la información superando la brecha lingüística, en particular en los sitios web gubernamentales, mediante la traducción automática de contenidos, que se ha generalizado. Esto puede mejorar la capacidad de participar en procesos participativos al fomentar el diálogo entre los participantes que hablan diferentes idiomas (European Parliament, 2022^[155]). La plataforma Bhashini de la India, analizada en el Capítulo 4, es otro ejemplo de cómo se intenta superar las brechas lingüísticas, por ejemplo, mediante el uso de técnicas participativas.

Hacer que los procesos participativos sean efectivos e inclusivos

Los gobiernos pueden utilizar la IA para ayudar a llevar a cabo y facilitar los debates y diálogos públicos. A menudo, estos diálogos se enfrentan a una disyuntiva entre la amplitud (número y variedad de participantes) y la profundidad del debate y la creación de consenso (Landemore, 2022^[156]). Si bien Internet ha permitido espacios de comunicación de muchos a muchos, estos espacios no siempre están diseñados para fomentar la deliberación y la creación de consenso (Weyl, Tang and Plurality, 2022^[157]). La IA puede proporcionar un espacio deliberativo acogedor y productivo (Landemore, 2022^[156]; Ovadya, 2023^[158]) al:

- Extraer información significativa de la contribución de muchos participantes, que podría proporcionarse en diferentes formatos, como texto y voz.
- Detectar contenidos nocivos y apoyar a los moderadores humanos en la lucha contra la incitación al odio.
- Facilitar la verificación de datos proporcionando datos e información procedente de fuentes en tiempo real.
- Ayudar a realizar una facilitación en vivo para favorecer múltiples opiniones y crear puentes hacia el consenso y apoyar la deliberación.

La IA también puede mejorar la calidad del diálogo. Por ejemplo, puede ayudar a los organizadores a redactar materiales de aprendizaje accesibles y basados en hechos, así como **simplificar documentos burocráticos complejos en un lenguaje fácil de leer para permitir un debate informado**. Por ejemplo:

- En Estados Unidos, la herramienta impulsada por IA MAPLE (Plataforma de Massachusetts para la Participación Legislativa) permite a la ciudadanía comprender mejor el contexto y los objetivos de los borradores de textos legales a través de resúmenes generados por la IA que envían sus aportes y comentarios²⁹.
- En España, la plataforma participativa Decide Madrid, basada en el software de código abierto Consul, experimentó con un sistema de NLP para ayudar a la ciudadanía a agregar y desarrollar propuestas (Arana-Catania et al., 2021^[159]).

Las investigaciones indican que el uso de **herramientas de deliberación basadas en IA podría tener un impacto significativo en la disminución de la polarización** (Tsai et al., 2024^[160]; OECD, 2020^[161]). Los ejemplos en el Recuadro 5.36 ponen de manifiesto que estos instrumentos no solo son beneficiosos para obtener información comprensible sobre un tema específico, sino que también pueden ayudar a reunir los resultados de las múltiples deliberaciones. Pueden garantizar la coherencia preservando al mismo tiempo la riqueza de los debates (Landemore, 2022^[156]), así como fortalecer las conexiones entre los participantes y la sociedad en general haciendo que el contenido y los resultados de las discusiones sean más accesibles para el público.

Recuadro 5.36. La IA como facilitadora de procesos deliberativos

Deliberación masiva en línea con Polis

Polis es un software de código abierto impulsado por IA diseñado para permitir a grandes grupos de personas debatir, colaborar y alcanzar consensos sobre cuestiones complejas. Los participantes comparten sus opiniones sobre una pregunta determinada y expresan su acuerdo o desacuerdo sobre las declaraciones de otros participantes. El algoritmo mapea las respuestas de los participantes y agrupa las opiniones, facilitando la identificación de posiciones que son apoyadas por la mayoría de los participantes. El polaco se ha utilizado para facilitar el debate y crear consenso en Austria sobre cuestiones ambientales, en Uruguay sobre un referéndum nacional, en Nueva Zelanda para facilitar el desarrollo de la política gubernamental, en Filipinas para generar política municipal, en Estados Unidos para contrarrestar la polarización a nivel municipal y en Alemania para desarrollar una plataforma de partido político.

En Chile, la IA ayuda a moderar los procesos deliberativos

En 2020, el Senado de Chile se asoció con el Centro de Stanford para la Democracia Deliberativa para involucrar a la ciudadanía en el proceso deliberativo en línea "LXS 400 – Chile Delibera" realizado a través de la [Plataforma de Deliberación en Línea de Stanford](#). La plataforma utilizó una herramienta de IA para mantener el tiempo y distribuir equitativamente los derechos de expresión oral, lo que permitió a los participantes formar colas de expresión y debatir en pequeños grupos con agendas programadas. Más de 500 ciudadanos representativos de la sociedad chilena debatieron en pequeños grupos y en sesiones plenarias sobre las reformas sanitarias y de las pensiones.

Fuente: (The Computational Democracy Project, 2024^[162]; Fishkin and al, 2021^[163]).

Además de facilitar la participación, la IA puede ayudar a los gobiernos y la ciudadanía a analizar y dar sentido a los aportes de los procesos participativos y deliberativos. Las funciones de IA, como la automatización y el procesamiento de grandes cantidades de datos y texto, pueden facilitar la deliberación y el diálogo en línea masivos y multilingües. Por lo general, los gobiernos se enfrentan al reto de analizar grandes cantidades de insumos cualitativos recibidos en las consultas públicas, que pueden remediarse con la ayuda de la IA. Por ejemplo, la Incubadora de Inteligencia Artificial (i.AI) del Reino Unido ha desarrollado el i.AI Consultation Analyser, cuyo objetivo es ampliar la capacidad del gobierno para analizar los aportes de la ciudadanía en las consultas públicas legales. Se espera que la herramienta reduzca el costo de esta actividad en 80 millones de GBP (AI.GOV.UK, 2024^[164])³⁰.

Las herramientas de interpretación con IA permiten la identificación y evolución de temas recurrentes, la agrupación automática de opiniones e ideas, la detección de valores atípicos y el análisis de sentimientos de la percepción de un tema determinado (Arana-Catania et al., 2021^[159]; Berditchevskaia and Baeck, 2020^[165]; Schneider and Sanders, 2023^[166]). Las herramientas de sensibilización pueden ayudar a los servidores públicos a comprender y visualizar las prioridades y opiniones de la ciudadanía sobre cuestiones concretas y ayudarles a traducir grandes volúmenes de aportes en recomendaciones viables que sean representativas de las opiniones de los participantes. Los ciudadanos también pueden utilizar estas herramientas para comprender las perspectivas de los demás, informar y elaborar sus propias aportes, y comprender mejor los resultados de los procesos participativos y deliberativos. Ejemplos de ello son el *Gran Debate* nacional de Francia, o *Gran Debate* (Recuadro 5.37), así como:

- El Ayuntamiento de Cambridge, Reino Unido, utilizó la herramienta de interpretación impulsada por IA de la plataforma Go Vocal (anteriormente CitizenLab), que agrupa y prioriza

automáticamente las opiniones de la ciudadanía para analizar sus contribuciones a su Código de Diseño, ahorrando un 50 % del tiempo estimado para el procesamiento manual (Fillet, 2024^[167])³¹.

- El Comité Económico, Social y Medioambiental francés (CESE) organizó un proceso deliberativo (Convención Citoyenne) para informar a la legislación sobre la delicada cuestión política del final de la vida. La herramienta Panorámica, desarrollada por Make.org y desarrollada por IA, permite a la ciudadanía, periodistas y OSC acceder y navegar por los debates y los resultados del proceso deliberativo a través de una interfaz interactiva³².

Recuadro 5.37. Mejoramiento de la participación en el Gran Debate Nacional de Francia

En Francia, el Gran Debate Nacional, lanzado a principios de 2019, es un ejemplo de un esfuerzo pionero para democratizar el debate público a través de la innovación digital. En respuesta al creciente descontento social, el Gobierno Francés estableció esta consulta nacional para reunir la opinión de la ciudadanía sobre cuestiones clave, como el empleo, la fiscalidad y la reforma democrática, utilizando una plataforma en línea dedicada. La inteligencia artificial desempeñó un papel central en la gestión e interpretación de la amplia gama de contribuciones.

Se emplearon técnicas impulsadas por la IA, como la NLP y los algoritmos de clustering, para analizar sistemáticamente millones de propuestas. Al identificar temas recurrentes y sintetizar diversos puntos de vista, estas herramientas permitieron a los responsables de la formulación de políticas poner de relieve puntos de consenso y tendencias emergentes. Este análisis automatizado transformó la información bruta aportada por la ciudadanía en información accesible, lo que propició una toma de decisiones más informada y transparente.

Fuente: <https://granddebat.fr>.

Por último, la IA puede facilitar la innovación en actividades de participación cívica al **generar simulaciones y visualizaciones sobre diversos escenarios y alternativas**, incluso en combinación con otras tecnologías inmersivas emergentes como la realidad virtual o aumentada³³. Las herramientas generativas de IA pueden ayudar a los gobiernos a interactuar con el público en proyectos complejos relacionados con un espacio físico (por ejemplo, movilidad urbana y rediseño). Estas herramientas pueden simular escenarios futuros y alternativos, permitiendo a la ciudadanía visualizar los resultados finales y proporcionar comentarios para su consideración por parte del gobierno (OECD, 2023^[2]). La IA generativa puede reducir significativamente los costos de creación de simulaciones visuales al automatizar el proceso de diseño y generar representaciones realistas. Por ejemplo, la ciudad de Hamburgo, en Alemania, utilizó tecnologías 3D e IA en la plataforma CityScope, desarrollada en el MIT Media Lab, para involucrar a las partes interesadas relevantes en el proceso de toma de decisiones para 161 lugares viables para acoger a los refugiados (Campagnucci et al., 2025^[168]). Del mismo modo, Finlandia utilizó la herramienta UrbanistAI para fomentar el debate sobre la peatonalización de las calles (Recuadro 5.38).

Recuadro 5.38. Finlandia utiliza IA generativa para la planificación urbana participativa

En Finlandia, la ciudad de Helsinki utilizó la IA generativa para mejorar la planificación urbana participativa a través de UrbanistAI. Esta plataforma integra herramientas de visualización impulsadas por IA con funciones de diseño conjunto que permiten a residentes y empresas dar forma de forma colaborativa a espacios urbanos. Durante la iniciativa Calles de Verano 2023, Helsinki utilizó la plataforma en talleres comunitarios, donde los participantes tradujeron sus ideas en diseños de calles generados por IA, fomentando procesos de planificación urbana más inclusivos y atractivos.

Al permitir la visualización en tiempo real de las propuestas de la comunidad, UrbanistAI ayudó a reducir la brecha entre la aportación de la ciudadanía y las decisiones finales de diseño. La capacidad de la herramienta para alinear las propuestas con las políticas locales también racionalizó el proceso de planificación, garantizando el cumplimiento de la regulación al tiempo que se mantiene la flexibilidad creativa. El éxito del caso de Helsinki ha inspirado una adopción más amplia de la plataforma, con ciudades como Tallin, Pristina y Berlín integrándola en sus iniciativas participativas de diseño urbano.

Fuente: <https://interoperable-europe.ec.europa.eu/collection/public-sector-tech-watch/reimagining-helsinki-participatory-urban-planning-generative-ai>.

Apoyo a una comunicación pública efectiva y adaptada

Un software de análisis basado en IA puede ayudar a los comunicadores públicos a monitorear y analizar el debate en línea, las percepciones de la audiencia y las actitudes sobre un tema determinado³⁴. La IA puede ayudarles a mejorar la calidad, cantidad, accesibilidad y pertinencia de la información que los gobiernos facilitan a diferentes segmentos de la población. Además, las herramientas de IA pueden mejorar la eficiencia y efectividad de los comunicadores públicos al ayudar en la lluvia de ideas, automatizar tareas rutinarias como la planificación de medios y garantizar la coherencia del contenido en todos los canales de comunicación. Estas capacidades permiten a los comunicadores interactuar mejor con sus audiencias, adaptar los mensajes a diversos grupos y responder rápidamente a las preocupaciones públicas, fomentando así una ciudadanía más informada y comprometida. Véase el ejemplo en el Recuadro 5.39.

Recuadro 5.39. Innovación en las comunicaciones públicas con IA en el Reino Unido

En el Reino Unido, el Servicio de Comunicación Gubernamental (GCS) ha estado desarrollando y probando GCS Assist, una herramienta conversacional impulsada por IA diseñada para ayudar a los comunicadores gubernamentales a mejorar la eficiencia y efectividad de su trabajo. La herramienta integra directrices profesionales, marcos de comunicación e información sobre la audiencia privada, lo que garantiza que los resultados generados por la IA se ajusten a las normas gubernamentales. Dado que la IA es más efectiva cuando se utiliza como asistente colaborativo, GCS Assist está diseñado para apoyar a profesionales con y sin experiencia racionalizando las tareas rutinarias, mejorando la creatividad y permitiendo más tiempo para el trabajo estratégico.

GCS Assist ha experimentado un proceso de desarrollo iterativo, que comenzó con un prototipo inicial pilotado dentro de la Oficina del Gabinete a finales de 2023. La herramienta fue diseñada para apoyar a los comunicadores gubernamentales generando borradores de textos, planes e ideas estratégicas, garantizando al mismo tiempo la seguridad de los datos dentro de los sistemas gubernamentales. En la evaluación alfa de marzo de 2024 se calificó el progreso de la herramienta como ámbar, reconociendo su potencial y destacando al mismo tiempo las áreas susceptibles de mejora. Actualmente se está llevando a cabo una implementación gradual, que comienza con un grupo limitado de usuarios antes de ampliar el acceso de forma más amplia.

Fuente: <https://gcs.civilservice.gov.uk/wp-content/uploads/2024/04/2024-05-GCS-Innovating-with-Impact-Strategy-interactive.pdf>.

La producción y difusión de materiales informativos, así como la asistencia y respuesta a las consultas de la ciudadanía, es una función fundamental pero que requiere muchos recursos para permitir una participación efectiva. La IA también puede potenciar las iniciativas de comunicación pública mediante la

selección y adaptación de materiales de comunicación. La IA generativa puede, por ejemplo, resumir y sintetizar información y crear contenido para diferentes públicos en diferentes formatos (vídeo, imagen, texto, etc.). Por otra parte, GenAI puede ayudar a los gobiernos a proporcionar a la ciudadanía información accesible y basada en pruebas, permitiéndoles comprender mejor las cuestiones políticas en juego, así como el papel del gobierno en su regulación. Los ejemplos de los Países Bajos y Brasil que figuran en el Recuadro 5.40 ilustran cómo se puede hacer esto en el ámbito gubernamental.

Recuadro 5.40. Prácticas de IA para la conservación de la información en la participación ciudadana

Ayudar a informar sobre las deliberaciones verbales en los Países Bajos

La Cámara de Representantes de los Países Bajos implementó el sistema Speech2Write, que permite a su Oficina de Informes Parlamentarios (PRO) convertir voz en texto, así como traducir voz en informes escritos y producir un informe "casi listo" para su finalización humana. El sistema se basa en un modelo de lenguaje parlamentario que se desarrolló desde cero a partir de más de 2 000 horas de audio de actas y archivos de texto.

Hacer más accesible la información compleja en Brasil

La Cámara de Diputados brasileña está utilizando la IA para mejorar el apoyo prestado a los parlamentarios y construir una institución legislativa más amigable con el público. Su Ulysses AI Suite utiliza un sistema de NLP para clasificar y ayudar a la ciudadanía a navegar por la gran cantidad de información producida por la Cámara, incluidas leyes, discursos, datos y documentos de investigación. Ulises también ayuda a los parlamentarios a redactar textos legislativos y a encontrar proyectos de ley similares ya adoptados o rechazados.

Fuente: <https://www.ipu.org/innovation-tracker/story/artificial-intelligence-innovation-in-parliaments>, (Arana-Catania et al., 2021^[159]), <https://www.camara.leg.br/noticias/548730-camara-lanca-ulysses-robo-digital-que-articula-dados-legislativos>.

Los asistentes virtuales, como los chatbots, pueden respaldar las iniciativas de comunicación ayudando a la ciudadanía a navegar por la información gubernamental, respondiendo a sus consultas en un lenguaje sencillo y enviando notificaciones y contenido personalizados en función de sus intereses (Androutsopoulou et al., 2019^[169]; van Noordt and Misuraca, 2019^[170]; Cortés-Cediel, 2023^[171]). En algunos casos, se han utilizado chatbots como canal de participación cívica (véanse algunos ejemplos en el Recuadro 5.41).

Recuadro 5.41. Chatbots apoyando la participación en España y Colombia

En 2021, la ciudad de Madrid (España) integró Clara, una asistente virtual basada en IA, en su plataforma participativa existente, Decide Madrid, basada en el software de código abierto Consul. Clara responde a las preguntas de los usuarios sobre el funcionamiento de la plataforma y de los procesos participativos de la ciudad.

En 2022, la ciudad de Bogotá (Colombia) desplegó Chatico. Este asistente virtual apoya la interacción entre la ciudadanía y la administración local mejorando la experiencia de los procedimientos y servicios y permitiendo la participación ciudadana. Chatico está disponible tanto en el sitio web como a través de WhatsApp para mejorar su accesibilidad.

Fuente: <https://www.madrid.es/portales/munimadrid/es/Inicio/Actualidad/Noticias/Nace-Clara-el-asistente-virtual-de-Decide-Madrid>, <https://oecd-opsi.org/innovations/chatico-virtual-agent>, <https://gobiernoabiertobogota.gov.co>.

Fortalecimiento de los procesos electorales

Los debates en materia de políticas sobre el potencial de la IA en el proceso electoral se han centrado en proteger los procesos electorales contra los riesgos de la IA, en concreto, empoderando a la sociedad y creando resiliencia, protegiendo los principios democráticos y promoviendo la integridad de la información. Sin embargo, también existen oportunidades relevantes para que la IA refuerce los procesos electorales evitando el fraude y la privación de derechos, y reduciendo las barreras de entrada para los candidatos con fondos insuficientes (Eisen et al., 2023^[172]; OECD, 2024^[29]).

Los órganos de gestión electoral (EBM) pueden utilizar las herramientas de IA durante todo el período electoral (Juneja, 2024^[173]). Antes de las elecciones, se pueden adoptar herramientas de IA para apoyar la gestión de listas de votantes, el registro de votantes, la planificación de la asignación de recursos, la previsión de costos electorales o la publicidad dirigida. Por ejemplo, en Estados Unidos, el Centro de Información de Registro Electrónico (ERIC) (2024^[174]) utiliza ML para analizar el registro de votantes y los registros de vehículos de motor. Esto ayuda a los estados a identificar registros duplicados, votantes fallecidos e individuos elegibles pero no registrados, lo que mejora la precisión del censo electoral. Durante las elecciones, las herramientas de IA pueden contribuir al seguimiento de las campañas y los medios de comunicación, a la verificación de votantes, al seguimiento de los centros de votación y al recuento de votos. Por último, la IA puede ayudar a las autoridades en las auditorías postelectorales detectando incidentes o fraudes (Juneja, 2024^[173]).

Gestión de riesgos y desafíos

La IA ofrece importantes oportunidades para empoderar a la ciudadanía y a la sociedad civil y reconfigurar las relaciones entre los gobiernos y su ciudadanía para mejor. Sin embargo, persisten riesgos potenciales, que pueden tener implicaciones directas en los espacios democráticos. Si bien los avances tecnológicos pueden abordar progresivamente estos problemas, los responsables de la formulación de políticas deben tenerlos en cuenta al adoptar herramientas de IA en este ámbito.

Riesgos asociados

- Datos inadecuados o sesgados en los sistemas de IA.
- Dependencia excesiva de la IA.
- Falta de transparencia y explicabilidad.
- Resistencia: falta de comprensión pública sobre el uso de la IA por parte del gobierno.
- Uso indebido o cuestionable de la IA, que genera problemas de vigilancia y privacidad.
- Exacerbación de la exclusión social y las brechas digitales.

Si los sistemas de IA se basan en **datos inadecuados o sesgados**, podrían producirse resultados inexactos o adversos para algunas personas o grupos. En lo que respecta a la participación cívica y el gobierno abierto, esto podría dar lugar a una consideración inexacta o imprecisa de la aportación de la ciudadanía a los procesos de participación habilitados por la IA.

Además, los procesos deliberativos no solo pretenden aprovechar la inteligencia colectiva y permitir el diálogo social sobre cuestiones políticas específicas, sino también mejorar el conocimiento de los participantes, así como su comprensión mutua y empatía (OECD, 2021^[175]). **Una dependencia excesiva de las herramientas de IA** para mejorar la eficiencia de los procesos deliberativos podría pasar por alto otros aspectos fundamentales de la deliberación.

La falta de transparencia y explicabilidad de la IA, especialmente en sistemas complejos, puede reducir la confianza en las herramientas de IA y, cuando se utilizan en espacios democráticos, puede, a su vez, afectar a la confianza en el proceso participativo y sus resultados. Para hacer frente a este desafío,

algunos gobiernos han puesto en marcha registros de algoritmos que divultan información detallada y técnica sobre los algoritmos, como su finalidad, diseño, entradas de datos, procesos de toma de decisiones y posibles sesgos, entre otros (para más detalles y ejemplos, véase el Capítulo 4, sección "Establecimiento de barreras para orientar la IA estratégica y responsable").

Además, todavía hay muchas incógnitas sobre las percepciones de los usuarios con respecto a la IA en los procesos participativos y la influencia que puede tener en sus contribuciones y su **disposición a participar**. Se están realizando experimentos para evaluar la reacción de los participantes a la introducción de herramientas de IA como intermediarios en procesos participativos y deliberativos (Hadfi et al., 2021^[176]; Kim et al., 2021^[177]). El uso de herramientas de IA para ayudar a los participantes a formular sus insumos y aportes debe manejarse con cautela para evitar que comprometan la creatividad en el lenguaje y la reflexión de los participantes, centrándose más bien en la eficiencia y el acuerdo.

Algunos gobiernos han utilizado indebidamente las tecnologías digitales para la vigilancia o incluso para silenciar a la población y a la oposición digital, dañando así el espacio cívico (OECD, 2022^[152]). El uso de sistemas de IA con fines de vigilancia, la censura de contenidos y las formas inadecuadas de actuación policial predictiva amenazan el libre ejercicio de las libertades civiles, como el derecho de reunión pacífica y la libertad de expresión. Estos usos inadecuados corren el riesgo de generar desconfianza general entre la ciudadanía con respecto a la adopción y el despliegue de sistemas de IA en las funciones gubernamentales y los servicios públicos. La OCDE (2022^[152]) constató que la mayoría de las estrategias nacionales de IA no abordan el impacto de la IA en la capacidad de ejercer libremente sus derechos, aunque alrededor de la mitad propone mecanismos concretos de supervisión y reparación (OECD, 2022^[152]).

Por último, **muchos idiomas no están suficientemente representados en los sistemas de IA**, que se forman principalmente en inglés, español y mandarín (Peixoto, Canuto and Jordan, 2024^[178]) (véase el Capítulo 1, sección "Exacerbación de las brechas digitales", para un análisis detallado de la cuestión). En el contexto de la participación ciudadana, esto significa que los insumos presentados en otros idiomas podrían no procesarse y valorarse de la misma manera, **creando nuevos desequilibrios democráticos** (Romberg and Escher, 2024^[179]). Por ejemplo, para abordar la brecha lingüística y preservar el idioma islandés, el Gobierno de Islandia se asoció con OpenAI para formar al LLM GPT-4 en islandés (Government of Iceland, 2023^[180]). Del mismo modo, la Universidad de Turku (Finlandia) se asoció en 2023 con la empresa SiloAI para construir el conjunto de sistemas Poro, una familia de LLM multilingües de código abierto para todos los idiomas oficiales europeos (University of Turku, 2023^[181]). Iniciativas similares existen en China para otros idiomas, incluidos los indígenas y los amenazados (OECD, 2023^[182]).

Desafíos de implementación

- Brechas de habilidades.
- Elevados costos de adopción y escalabilidad de la IA.

La adopción de herramientas de IA por parte de los funcionarios públicos encargados de los procesos participativos y deliberativos, así como por parte de los profesionales de la participación ciudadana, requiere unas competencias adecuadas. Este conjunto de competencias abarca desde la identificación de los usos pertinentes de las herramientas de IA para apoyar las prácticas participativas y deliberativas (OECD, 2025^[183]), a la selección y compra o desarrollo de soluciones de IA que cumplan los objetivos definidos, al uso de herramientas de IA en el contexto de procesos participativos y deliberativos, a la evaluación de la calidad de sus resultados.

La adopción, ya sea a través del desarrollo interno o de la contratación pública, y el uso de herramientas de IA para apoyar la participación ciudadana requieren un enfoque flexible que incluya iniciativas piloto, pruebas y ajustes iterativos. Este enfoque conlleva costos considerables y requiere una elevada tolerancia

al riesgo. Los desafíos relacionados con los costos y la escala de las herramientas de IA son particularmente relevantes para la participación ciudadana, ya que los gobiernos locales suelen organizar e implementar procesos participativos (OECD, 2022^[184]).

Potencial sin explotar y camino a seguir

Dada la novedad del uso de la IA en la participación civil y la habilitación, es difícil evaluar las pruebas de impacto y las lecciones generales aprendidas más allá de los estudios de casos individuales. Los casos prácticos presentados en este capítulo sugieren que la IA tiene el potencial de mejorar la eficiencia, reducir los costos y ampliar el alcance de la participación, reduciendo en última instancia las barreras para que la ciudadanía y las OSC opinen y opinen en las distintas etapas de la formulación de políticas. El uso de herramientas de IA para procesar los aportes de la ciudadanía, automatizar tareas, apoyar la creación de materiales de información y comunicación y proporcionar a la ciudadanía asistencia virtual las 24 horas del día, los 7 días de la semana, podría reducir significativamente algunos de los costos asociados al diseño y la implementación de los procesos pertinentes. Estos ahorros, junto con la creciente disponibilidad de herramientas de código abierto, podrían permitir oportunidades de participación ciudadana en contextos con recursos limitados, como en ciudades más pequeñas y regiones menos desarrolladas.

Las tecnologías digitales han permitido una participación asincrónica y han reducido significativamente la carga de las barreras geográficas, permitiendo una colaboración más amplia (OECD, 2022^[184]). La IA tiene el potencial de reducir otras barreras, como la brecha lingüística, permitiendo niveles más altos de participación ciudadana. Además, los asistentes virtuales y los chatbots pueden mejorar aún más la accesibilidad de los procesos participativos y deliberativos al ayudar a los participantes a navegar por información compleja. Por último, las herramientas de IA para la moderación, la facilitación y la interpretación podrían resultar esenciales para abordar el desafío de la escala en los procesos deliberativos (Landemore, 2022^[156]).

Los primeros esfuerzos podrían conducir a concepciones a más largo plazo de transparencia y capacidad de respuesta. Las tecnologías digitales han afectado durante mucho tiempo la forma en que las personas se relacionan con nociones como su llamado "derecho a saber" y expresan opiniones sobre operaciones complejas y sus resultados (Margetts, 2011^[185]; Song and Lee, 2015^[186]). Las tecnologías de IA también ofrecen formas novedosas de mejorar el diseño, la implementación y la transparencia de los servicios. Las personas pueden empezar a considerar que los sistemas operativos actuales son anticuados y difíciles de usar, por lo que pueden expresar sus opiniones sobre cómo los servicios deben integrar tecnologías que puedan mejorar el acceso a la información (Carrasco et al., 2019^[187]). Esto podría ayudar a acelerar las reformas institucionales destinadas a mejorar la transparencia y la accesibilidad. Estos cambios pueden estar impulsados no solo por la adopción de tecnologías de IA, sino también por su capacidad para cambiar la percepción pública y las expectativas de capacidad institucional.

De cara al futuro, los gobiernos deberían explorar el potencial de la IA para mejorar los procesos de participación cívica existentes ampliando las capacidades gubernamentales de análisis de las contribuciones de la ciudadanía, permitiendo la deliberación masiva y reduciendo las barreras a la participación. Los gobiernos deben considerar cuidadosamente los desafíos asociados con el uso de la IA en la participación cívica, en particular los relacionados con una mayor intermediación tecnológica en las interacciones entre la ciudadanía y el gobierno, y su impacto en el discurso público y la confianza. La adopción y el despliegue de herramientas de IA para la participación cívica deben seguir estrategias claras que incluyan la transparencia y la rendición de cuentas como principios fundamentales. Los ciudadanos y la sociedad civil deben participar activamente en la elaboración de dichas estrategias para garantizar su alineación con los principios democráticos y aumentar la confianza y la aceptabilidad del proceso (véase el Capítulo 4, sección "Participación para dar forma a una IA estratégica y responsable").

Los servicios públicos y las funciones judiciales

Los servicios públicos y las funciones judiciales son esenciales para satisfacer las necesidades de la ciudadanía y mantener el Estado de derecho. Esto incluye el diseño y la implementación de servicios públicos que sean accesibles, eficientes y receptivos, así como una aplicación efectiva de la ley y una gestión del riesgo de desastres que garantice seguridad y resiliencia. La administración de justicia y el acceso a la justicia son igualmente fundamentales, ya que proporcionan procesos legales justos y oportunos que protegen los derechos y fomentan la confianza pública. Estas áreas son las que tienen mayor impacto en la ciudadanía. Los gobiernos están utilizando IA de diversas maneras para garantizar que sus operaciones sean eficientes y efectivas, y que la ciudadanía y los residentes estén seguros y sean tratados con equidad.

IA en el diseño e implementación de servicios públicos

El desarrollo y uso de la IA ha permeado el diseño e implementación de servicios públicos en los últimos años. Por ejemplo, el 67 % de los países de la OCDE han utilizado la IA para mejorar esta función (OECD, 2024^[188]). Cada vez más, los funcionarios públicos se encargan de experimentar con la IA en sus funciones laborales (Bright et al., 2025^[189]). En general, el objetivo es que sean más eficientes, efectivos y respondan mejor a las necesidades de la ciudadanía. La IA puede ayudar a conseguirlo mediante la automatización de los procesos, una asignación de recursos más eficiente (por ejemplo, optimizando los niveles de personal), el apoyo a la toma de decisiones, la mejora de la participación ciudadana y la adaptación y personalización de los contenidos y las vías de servicio para la ciudadanía (2024^[60]; 2023^[190]; 2019^[191]; OECD, 2025^[192]). La IA también puede aumentar la comprensión de los gobiernos sobre los desafíos socioeconómicos a los que se enfrentan los usuarios de servicios —al detectar tendencias y relaciones causales que no son detectables por los seres humanos—, mejorando así la capacidad de respuesta y la personalización de los servicios públicos (Anshari et al., 2024^[193]). Esto puede permitir a los gobiernos diseñar nuevos servicios públicos que sean más proactivos a la hora de anticipar y responder a las necesidades de los usuarios (Madan and Ashok, 2023^[194]).

En determinadas circunstancias, la toma de decisiones automatizada en los servicios públicos también puede percibirse como más justa y menos propensa a errores humanos o corrupción (Miller and Keiser, 2020^[195]; Leyer and Schneider, 2021^[196]). Sin embargo, diversos factores de comportamiento humano —la forma en que las personas perciben, confían e interactúan con las decisiones impulsadas por la IA— determinarán su efectividad y legitimidad. Una mejor implementación de servicios y su mayor efectividad pueden impulsar la satisfacción de la ciudadanía y su confianza en el gobierno (Sun and Medaglia, 2019^[197]; OECD, 2024^[46]). Algunos sostienen que este momento requiere un replanteamiento amplio de cómo los efectos de la IA en las estructuras de las administraciones públicas subyacentes a la implementación de servicios públicos, más allá de centrarse en los aumentos de eficiencia (Babšek et al., 2025^[198]).

Situación actual

Automatización y optimización de tareas

Las tareas más comunes en los servicios públicos son rutinarias: registrar, preparar, clasificar, clasificar, archivar y verificar la exactitud de la información, y tienen un alto potencial de ser automatizadas (Straub et al., 2024^[199]). La IA puede agilizar las tareas burocráticas, liberando tiempo a los servidores públicos para que se centren en aquellas tareas que requieren juicio humano, creatividad, discreción y toma de decisiones. Los casos prácticos reales ponen de relieve la importancia de la rentabilidad a la hora de integrar aplicaciones de IA en la prestación de servicios. Por ejemplo, Brasil emplea una solución de RPA para detectar y desestimar demandas legales predatorias y repetitivas —demandas diseñadas por entidades más ricas para cargar a los acusados más pequeños con solicitudes de datos y costos exorbitantes y extraños—, agilizando así los procedimientos judiciales y ahorrando un tiempo valioso³⁵.

En España, la tecnología RPA automatiza o acelera las tareas repetitivas de bajo valor en los servicios públicos, garantizando una distribución eficiente de las pensiones de invalidez y jubilación no contributivas³⁶. Se pueden consultar ejemplos adicionales de Austria y España en el Recuadro 5.42.

Recuadro 5.42. Soluciones de IA en procesos de planificación

Revisión de permisos digitales de construcción de Viena (Austria)

El Sistema de Información sobre Normas de Construcción para la Presentación de Solicitudes (BRISE-Viena) de Austria introduce una innovación digital en las revisiones de permisos de construcción, transformando los procesos tradicionales basados en papel en un flujo de trabajo digital completo. En el proyecto BRISE, las rutinas de verificación basadas en IA comparan automáticamente los modelos digitales de información sobre edificios en 3D presentados por los planificadores con un modelo de referencia digital municipal que engloba los requisitos regulatorios. Este proceso automatizado no solo identifica rápidamente cualquier desviación de las normas establecidas, sino que también proporciona una retroalimentación inmediata, reduciendo significativamente el período de revisión tradicionalmente largo.

Automatización el informe de pobreza energética en Cataluña (España)

El Informe sobre pobreza energética automatizada de Cataluña (España) es una iniciativa del Consorcio de Administración Abierta de Cataluña (AOC) para desarrollar una plataforma impulsada por la IA para automatizar la generación de informes sobre pobreza energética. Esta plataforma se centra en el proceso manual y burocráticamente costoso de determinar la vulnerabilidad social y elaborar informes sobre pobreza energética para los municipios. La innovación radica en la automatización del proceso de generación de informes de pobreza energética, que tradicionalmente implicaba métodos manuales, lentos y propensos a errores. La iniciativa aborda el problema de la determinación de la vulnerabilidad social mediante el uso de la automatización robótica de procesos (RPA), el modelado de procesos comerciales (BPM) y las tecnologías de IA para agilizar el proceso. Anteriormente, los municipios gestionaban esta tarea de forma independiente, lo que daba lugar a resultados incoherentes y a una asignación desigual de los recursos. La automatización de los informes sobre pobreza energética conlleva varias ventajas, como un soporte centralizado basado en IA, una integración fluida con diversos sistemas de información y una mayor interoperabilidad entre empresas, municipios y otras administraciones. En general, la plataforma permite determinar la vulnerabilidad social de manera más eficiente y precisa, agilizar los procesos administrativos y asignar recursos de forma equitativa para mitigar la pobreza energética.

Automatización de registros de propiedad en Grecia

El Catastro Helénico, anteriormente obstaculizado por el procesamiento manual de papel que retrasaba las transacciones inmobiliarias durante meses o incluso años, ha introducido un sistema impulsado por la IA para automatizar la lectura, la categorización y la evaluación jurídica de los contratos inmobiliarios. Esta innovación ha reducido los tiempos de evaluación de varias horas a menos de 10 minutos y ha reducido los costos de 15 euros a tan solo 0,11 euros por caso. También ha mejorado la seguridad jurídica de los propietarios e impulsado la actividad económica al acelerar la finalización de las transacciones.

Fuente: <https://oecd-opsi.org/innovations/brise-vienna-building-regulations-information-for-submission-envolvement>, <https://uia-initiative.eu/en/news/piloting-brisevienna-results-journal-ndeg-3>, <https://www.aoc.cat/blog/2023/pilot-pobresa-energetica>, <https://www.ekathimerini.com/economy/1245436/smart-solutions-for-land-register>, <https://www.microsoft.com/en/customers/story/20308-hellenic-cadastre-azure-open-ai-service>, funcionarios del Gobierno de Grecia.

La automatización es relativamente fácil con la IA cuando se trata de tareas repetitivas y poco discrecionales si la calidad de los datos está garantizada. La automatización resulta más difícil cuando se trata de tareas discretionales de mayor nivel, pero se utilizan para descubrir nuevas opciones, detectar anomalías y riesgos y planificar mejor los servicios. Sin embargo, deben tenerse en cuenta los posibles desafíos para garantizar la confiabilidad de los sistemas, como se observa en el Recuadro 5.43,

Recuadro 5.43. Automatización para una implementación de servicios más fluida en Suecia

El Servicio Público de Empleo de Suecia puso en marcha el programa Prepare and Match en todo el país en 2021 para mejorar la coherencia y la precisión de las evaluaciones relacionadas con el mercado laboral y, por tanto, mejorar la eficiencia de la asignación de recursos. Los solicitantes de empleo inscritos reciben apoyo en forma de capacitación u orientación de un proveedor elegido. Las decisiones sobre si los solicitantes de empleo deben ser aprobados para participar en Prepare and Match están asistidos por una herramienta de apoyo a la decisión llamada BÄR.

BÄR es un sistema ML basado en datos históricos que consta de 1,1 millones de perfiles recopilados durante un período de 10 años. Ofrece diferentes recomendaciones de acción basadas en la probabilidad del solicitante de empleo de encontrar un trabajo dentro de los seis meses. Los factores considerados por el sistema incluyen la edad, el sexo, la educación, la ubicación residencial (y la información demográfica asociada) y las actividades previas de (des)empleo.

Las decisiones relativas a la preparación y el emparejamiento las toman formalmente los trabajadores sociales, quienes en las directrices reciben instrucciones de adherirse principalmente a la recomendación automatizada. Anular una recomendación negativa del sistema es difícil y requiere ponerse en contacto con un grupo de trabajo especial dentro del organismo. El papel asignado a la IA en el proceso de toma de decisiones pone de relieve el posible impacto de la IA para los demandantes de empleo, al tiempo que subraya los desafíos que supone apartarlos parcial o totalmente de las decisiones que les afectan. El sistema de apoyo a la toma de decisiones no tiene en cuenta el propio juicio de los solicitantes de empleo sobre su necesidad de apoyo

Fuente: (Berman, de Fine Licht and Carlsson, 2024_[200]; OECD, 2024_[201]).

Nuevas fuentes de información para la obtención de conocimientos avanzados

La IA puede ayudar a maximizar la calidad y la utilidad de los datos, así como la capacidad de los seres humanos y las máquinas para procesarlos y analizarlos (Jarrahi et al., 2023_[202]). Esto puede contribuir a que la IA se utilice para un análisis de impacto de los servicios públicos más avanzado y en tiempo real. Capaces de reconocer patrones que pueden ser indetectables por los humanos, las aplicaciones de IA pueden identificar rápidamente problemas que pueden haber pasado desapercibidos (Gesk and Leyer, 2022_[203]). También pueden procesar datos de diversas fuentes estructuradas y no estructuradas (por ejemplo, imágenes, texto o metadatos), facilitando enormemente la gestión de la información (Cockx, Lechner and Bollens, 2023_[204]). En Finlandia, la ciudad de Tampere recopila datos de las cámaras de tráfico para alimentar un algoritmo predictivo que prediga los flujos de visitantes en el centro de la ciudad, ayudando a las empresas y a los servidores públicos a tomar decisiones informadas, lo que representa un servicio público completamente nuevo (Recuadro 5.44).

Recuadro 5.44. Pronóstico de flujo de visitantes impulsado por IA con Tampere Pulse en Finlandia

Tampere Pulse es un servicio de análisis predictivo basado en IA que prevé los flujos de visitantes en el centro de la ciudad de Tampere, Finlandia. Desarrollado en respuesta a los desafíos que enfrentan las empresas locales para predecir el tráfico de clientes, el servicio utiliza datos de la plataforma de Internet de las cosas (IoT, por sus siglas en inglés) de la ciudad, incluidas cámaras de tráfico equipadas con visión artificial, condiciones climáticas y horarios de eventos. El resultado es una previsión del flujo de visitantes de un mes de duración, accesible tanto en dispositivos móviles como de escritorio a través de una interfaz intuitiva basada en mapas.

El servicio ha demostrado un impacto significativo, con una precisión de predicción del 79,33 % para la categoría exacta del flujo de visitantes y del 97,55 % para la categoría correcta o adyacente hasta con tres semanas de antelación. Esta precisión permite a las empresas optimizar la planificación de la fuerza laboral, la gestión de inventarios y las horas de funcionamiento, lo que conduce a ahorros de costos y a un mejor servicio al cliente. Los comentarios de los usuarios han sido muy positivos, con una calificación de facilidad de uso de 4,47 sobre 5 y una probabilidad de recomendación de 7,16 sobre 10.

Más allá de los beneficios comerciales, Tampere Pulse ayuda a los residentes y visitantes de la ciudad a planificar sus movimientos y ayuda a los equipos de mantenimiento y seguridad en la asignación de recursos. Las mejoras futuras tienen como objetivo ampliar la cobertura, integrar la información sobre el comportamiento de compra y vincularla con los sistemas de planificación de recursos empresariales (ERP). La iniciativa muestra el potencial de la IA y el Internet de las cosas para fomentar un entorno urbano inteligente y basado en datos, al tiempo que sirve de modelo para la replicación en otras ciudades de todo el mundo.

Fuente: <https://oecd-opsi.org/innovations/tampere-pulse>.

Estos nuevos conocimientos avanzados suelen ser el resultado de una combinación de información procedente de diversas fuentes y tecnologías. En Portugal, el Sistema de Vigilancia de la Cubierta Terrestre utiliza la IA para integrar las necesidades de las partes interesadas con la tecnología satelital con el fin de crear una cartografía abierta, promoviendo una visión eficiente, accesible y transparente de la cubierta terrestre (Recuadro 5.45). La Autoridad Saudita de Datos e Inteligencia Artificial (SDAIA) desarrolló Smart C para apoyar la planificación urbana impulsada por datos y la toma de decisiones en tiempo real, lo que permite herramientas como la simulación de tráfico, el monitoreo ambiental y la detección de contaminación visual basada en IA (SDAIA, 2024^[205]).

Recuadro 5.45. Sistema de vigilancia de la cubierta terrestre (SMOS) de Portugal

El sistema portugués de vigilancia de la cubierta terrestre (SMOS) es un servicio público que proporciona mapas de uso de la tierra y cubierta terrestre de acceso abierto basados en IA para apoyar la toma de decisiones basada en pruebas. Desarrollado por la Dirección General de Territorio (DGT), SMOS garantiza que los responsables de la formulación de políticas, los investigadores y el público tengan datos confiables sobre los cambios en la tierra, mejorando la transparencia y la gestión sostenible de la tierra.

SMOS analiza imágenes satelitales Sentinel-2, ortofotos de alta resolución y datos geográficos oficiales utilizando IA para producir productos cartográficos precisos, como mapas de uso de la tierra y cubierta terrestre (COS), mapas dinámicos de cubierta terrestre coyuntural (COSc) y productos de monitoreo especializados como el mapa anual de cultivos agrícolas temporales (MACAT). Estos productos se basan en la clasificación automatizada de imágenes satelitales en categorías temáticas de cobertura terrestre, utilizando la IA para distinguir características como zonas urbanas, bosques, masas de agua y campos agrícolas.

SMOS ha progresado a través de tres fases: investigación y desarrollo iniciales; participación con los usuarios para alinear los productos con las necesidades públicas; y ahora pleno despliegue operativo. Desde 2022, SMOS ha publicado mapas anuales de uso del suelo y cubierta vegetal (COS), mapas dinámicos de cubierta vegetal coyuntural (COSc) y productos de seguimiento especializados, todos ellos accesibles a través de la Infraestructura Nacional de Datos Espaciales (SNIG) de Portugal.

Fuente: <https://smos.dgterritorio.gov.pt>.

Entrega constante de información y servicios accesibles

Los asistentes virtuales y chatbots se encuentran entre los usos más comunes de la IA en los servicios públicos (Ramires Hernández, Valle-Cruz and Mendoza Méndez, 2022^[206]; Chen, Gascó-Hernandez and Esteve, 2023^[207]). Estas herramientas ofrecen a los usuarios de la función pública acceso rápido a la información las 24 horas del día. En la actualidad, los chatbots van desde los más básicos (algoritmos basados en la recuperación de información) hasta los más avanzados que utilizan sistemas de IA generativa, como los LLM, y utilizan datos relacionados con los servicios en tiempo real. Este último puede responder a las consultas con mensajes personalizados, que van desde conversaciones programadas hasta información y recomendaciones adaptadas a las necesidades específicas de los usuarios, incluido el lenguaje sencillo o el lenguaje de señas basado en la entrada de texto o voz. Además de ser conveniente para los usuarios, los chatbots pueden hacer que los servicios públicos sean más accesibles, aumentar la aceptación de los servicios e impulsar la participación.

Los esfuerzos realizados en Portugal son un ejemplo de cómo los gobiernos pueden integrar los últimos modelos de cimientos en los servicios públicos (Recuadro 5.46). En Estados Unidos, el chatbot Emma del Departamento de Seguridad Nacional proporciona a los usuarios información sobre servicios de inmigración, tarjetas verdes y pasaportes (Chen, Gascó-Hernandez and Esteve, 2023^[207]). En Singapur, el chatbot Gov.sg permite a los usuarios presentar y realizar un seguimiento del estado de las quejas sobre los servicios públicos. En Brasil, el municipio de Recife permite que más de 700 servicios públicos de la plataforma "Conecta Recife" se distribuyan y entreguen automáticamente a sus ciudadanos, con un chatbot que aprovecha los sistemas de IA para proporcionar información de inmediato a los usuarios. En Chile, se puso a prueba un chatbot en ChileAtiende para responder a las consultas de la ciudadanía sobre los servicios públicos, lo que demuestra el potencial de prestar asistencia personalizada y reducir los tiempos de respuesta a través de millones de solicitudes³⁷.

Además de las interacciones directas con la ciudadanía, las administraciones públicas también están utilizando chatbots basados en IA para ayudar al personal a prestar servicios de manera más eficiente. Por ejemplo, Caddy, desarrollado por i.AI del Reino Unido, actúa como copiloto para los agentes de servicio al cliente, ayudándoles a recuperar rápidamente información confiable de recursos gubernamentales y, al mismo tiempo, garantizar la exactitud a través de un sistema de validación basado en la participación de personas³⁸. Del mismo modo, la plataforma AIJobs de Singapur permite a los equipos gubernamentales personalizar asistentes de IA para procesos internos, agilizando los flujos de trabajo y mejorando la toma de decisiones³⁹. En Francia, se desarrolló un chatbot para ayudar a los servidores públicos a garantizar una mejor implementación de servicios en diferentes ámbitos

(Recuadro 5.46). El Gobierno de los Emiratos Árabes Unidos ha tratado de innovar el concepto de asistente virtual mediante la representación de asistentes virtuales y humanos en forma de holograma (Recuadro 5.47). Al centrarse en un ámbito específico, en mayo de 2025 la Autoridad de Planificación de Israel —el organismo nacional responsable de supervisar la planificación territorial, la infraestructura y los desarrollos habitacionales— comenzó a probar el concepto de un chatbot diseñado para ayudar a los profesionales de la planificación a revisar los planes de acuerdo con los requisitos nacionales, agilizando los largos procesos de revisión, que a menudo tardan años, y garantizando el cumplimiento de las normas jurídicas y profesionales.

Recuadro 5.46. Chatbots que agilizan el acceso a la información

El chatbot de Portugal proporciona información a la ciudadanía

En mayo de 2023, la Agencia de Modernización Administrativa de Portugal (AMA) lanzó un chatbot con un avatar realista para ayudar a la ciudadanía con preguntas sobre la clave digital móvil (MDK). Desarrollado en colaboración con Microsoft, DareData Engineering y Defined.ai, el chatbot se basa en el modelo turbo GPT-3.5 de Azure OpenAI y puede reconocer y reproducir texto y voz en portugués. Entre mayo y diciembre de 2023, facilitó 23 780 conversaciones (un promedio de 101 al día) y contribuyó a un aumento del 10 % en las activaciones de MDK.

En enero de 2025, AMA amplió el chatbot para convertirlo en un asistente virtual completo en el portal gov.pt, reforzando el compromiso de Portugal con la inclusión digital y la transformación de los servicios públicos. Ahora impulsado por ChatGPT con una base de conocimientos especializados, proporciona información sobre más de 2 300 servicios públicos y ofrece apoyo en 12 idiomas, eliminando las barreras lingüísticas para la ciudadanía de diversas nacionalidades. Integrado en más de 5 000 páginas de gov.pt, el asistente virtual introducirá nuevas características transaccionales en 2025, como el seguimiento de procesos y la programación de citas, lo que reforzará aún más su papel como herramienta clave de servicio público.

Albert, de Francia, apoya a los empleados de los servicios públicos

En abril de 2024, el Gobierno Francés introdujo Albert, una IA generativa soberana desarrollada por el DataLab de la Direction Interministérielle du Numérique (DINUM) para ayudar a los agentes de las administraciones públicas a responder a las solicitudes de la ciudadanía de manera más eficiente. Formado en modelos de código abierto y enriquecido con guías prácticas de service-public.fr, Albert ayuda a los empleados gubernamentales a buscar regulaciones, redactar resúmenes y proporcionar información precisa a la ciudadanía. Albert se basa en modelos de código abierto de Meta (Llama) y Mistral. Como norma digital común, la interfaz de programación de aplicaciones (API) de Albert proporciona modelos de IA generativa abiertos y reutilizables, reduciendo el umbral de adopción para los organismos públicos y contribuyendo a la creación de una infraestructura pública digital nacional para la IA.

A principios de 2024, Albert había sido sometido a pruebas por 60 agentes de 30 centros de servicios de Francia, que ofrecían respuestas sugeridas, fuentes verificadas y referencias útiles, mientras mantenían a los agentes humanos en control de las interacciones finales. Albert, que se puso en marcha oficialmente en todo el país en abril de 2024, se prevé que se amplíe a toda la administración pública, ayudando a los servicios tributarios, a la revisión de proyectos ambientales, a la transcripción judicial y a otros ámbitos.

Sofía, Francia, mejora el acceso público a la información

En 2025, el Ministerio de Transición Ecológica de Francia liberó a Sofía, una agente conversacional que facilita el acceso a los conocimientos científicos y técnicos de la organización. Sofía permite a los

usuarios consultar una base de datos documental especializada y agrupada, lo que agiliza y hace más eficiente la búsqueda de información. El sistema se basa en un LLM desarrollado por Mistral y puede proporcionar respuestas relevantes basadas en las necesidades de los usuarios. Sofía se ha desarrollado con atención a los impactos ecológicos. Por esta razón, se decidió mostrar a los usuarios la huella de carbono de cada búsqueda.

Los asistentes de IA de Grecia ayudan a los usuarios a navegar y encontrar información

Maigov de Grecia es un asistente impulsado por IA del portal gov.gr que ayuda a los usuarios a navegar por más de 1 300 servicios digitales y más de 3 200 procedimientos administrativos respondiendo a preguntas en griego natural mediante texto o voz. Maigov puede comprender y responder a las consultas enviadas en lenguaje natural, proporcionando una forma intuitiva para que los usuarios localicen los servicios que necesitan. El chatbot incorpora capacidades de razonamiento y mantiene el contexto de las conversaciones para proporcionar una experiencia óptima al usuario. Accesible tanto a través de teléfonos móviles como de ordenadores, maigov se lanzó inicialmente en modo piloto en diciembre de 2023 y puede gestionar hasta 240 conversaciones simultáneas en griego. Desde entonces, su funcionalidad se ha ampliado para dar soporte a 25 idiomas, atendiendo a una gama más amplia de usuarios, incluidos la ciudadanía griegos en el extranjero. Con más de 1.6 millones de consultas respondidas, Maigov ha demostrado ser una herramienta valiosa para mejorar la interacción ciudadana con el estado. Los futuros desarrollos para Maigov incluyen la capacidad de emitir certificados y programar citas en nombre de la ciudadanía, racionalizando aún más la implementación de servicios públicos. Además, el Ministerio de Turismo heleno, en colaboración con el Ministerio de Gobernanza Digital, ha desarrollado MaiGreece, un asistente de IA digital para mejorar la experiencia de viaje en Grecia. La aplicación proporciona información administrativa y de salud y seguridad, y recomendaciones para explorar Grecia.

Fuente: <https://oecd-opsi.org/innovations/virtual-assistant-for-public-services>; <https://digital.gov.pt/noticias/assistente-virtual-do-gov-pt-facilita-a-interacao-em-12-idiomas>, <https://www.info.gouv.fr/actualite/ia-connaissez-vous-albert>, <https://www.lefigaro.fr/secteur/high-tech/albert-l-ia-generative-de-l-etat-desormais-deploye-a-grande-echelle-20250212>, <https://digital-skills-jobs.europa.eu/en/latest/news/maigov-new-ia-digital-assistant-govgr-service-citizens>, funcionarios del Gobierno de Grecia.

Recuadro 5.47. Misión Inteligente de los Emiratos Árabes Unidos para servicios consulares

Implementada de acuerdo con la Estrategia de Servicios Gubernamentales de los Emiratos Árabes Unidos y la Estrategia Nacional de IA 2031, la Misión Inteligente está diseñada para agilizar y digitalizar los servicios consulares, reduciendo significativamente los tiempos de procesamiento, mejorando la implementación de servicios y mejorando la experiencia del usuario. El núcleo de esta transformación es el Servicio de Certificación de Documentos impulsado por IA, que utiliza un sistema de verificación de certificación impulsado por IA para autenticar documentos oficiales mediante el reconocimiento de sellos oficiales de documentos personales y comerciales. Esto garantiza una mayor precisión en los procesos de certificación y mejora las medidas de seguridad para proteger contra el fraude y el uso indebido.

Como complemento se encuentra la asistencia holográfica impulsada por IA, un asistente virtual interactivo de IA que puede interactuar con los usuarios directamente a través de un holograma. Estos agentes de IA prestan apoyo a los usuarios durante todo el trayecto, respondiendo a consultas al instante en varios idiomas y prestando asistencia en los servicios. Si se necesita más asistencia, el asistente de IA puede transferir la conversación a un funcionario consular humano en forma

holográfica. Los sistemas de IA predictiva analizan el comportamiento de los usuarios, los datos históricos y los patrones de interacción para proporcionar apoyo consular proactivo. Esto permite al sistema anticipar solicitudes frecuentes, sugerir servicios relevantes y notificar a los usuarios las acciones requeridas con anticipación.

Por último, también se proporciona un reconocimiento facial basado en IA como servicio opcional para un acceso sin problemas y servicios proactivos, donde los usuarios pueden autenticar su identidad utilizando el reconocimiento facial. Esto elimina la necesidad de documentos físicos y permite una asistencia personalizada sin necesidad de una introducción manual, como la expedición de documentos de retorno para ciudadanos de los Emiratos Árabes Unidos con pasaportes perdidos o dañados, sin necesidad de documentación alguna.

Fuente: Funcionarios del Gobierno de los Emiratos Árabes Unidos, <https://u.ae/en/about-the-uae/digital-uae/digital-transformation/platforms-and-apps/smart-mission>.

Investigaciones recientes también exploraron el papel de los chatbots en intervenciones conductuales (por ejemplo, durante la salud pública y las comunicaciones de riesgo durante la crisis del COVID-19), como aumentar la voluntad de distanciarse socialmente. Por ejemplo, un estudio aleatorizado controlado descubrió que un chatbot de IA, diseñado con ciencias del comportamiento, mejoró la comprensión de las pautas de COVID-19 y aumentó las intenciones de buscar servicios de pruebas al reducir la incertidumbre en torno a los comportamientos protectores. Mediante el uso de animaciones y mensajes atractivos, el chatbot abordó eficazmente las necesidades informativas y emocionales de los usuarios, demostrando cómo la IA puede influir en la toma de decisiones y promover comportamientos de salud pública (van Baal, 2024^[208]).

Los chatbots a menudo están diseñados principalmente para agilizar las operaciones, en lugar de optimizar la experiencia del usuario para la ciudadanía. Sin embargo, los estudios muestran que los usuarios interactúan más positivamente con los servicios basados en IA cuando los chatbots mantienen la continuidad en la identidad y retienen la información de los usuarios a través de las interacciones. Se ha descubierto que los sistemas de IA migrables, que preservan tanto la identidad de un agente como el historial de usuario en diferentes interacciones, aumentan la confianza, la competencia y la presencia social (Tejwani et al., 2020^[209]). Esto sugiere que el diseño de chatbots con un mayor énfasis en la familiaridad y la continuidad del usuario, no solo en la eficiencia, puede mejorar significativamente la participación. La aplicación de nociones sobre el comportamiento puede mejorar aún más su efectividad al garantizar que los chatbots detecten cuando los usuarios necesitan apoyo humano, en lugar de frustrarlos con interacciones automatizadas prolongadas. Al reconocer señales de confusión o urgencia, los chatbots de IA pueden transmitir sin dificultades problemas complejos a los agentes humanos, reduciendo la fricción y mejorando la accesibilidad del servicio. Priorizar la continuidad y el diseño centrado en el usuario en los chatbots no solo mejora la eficiencia, sino que también fomenta la confianza, la accesibilidad y la participación con los servicios públicos.

Mejoramiento de la personalización y adaptación de servicios

El uso de la IA para la implementación de servicios públicos también permite pasar de la implementación descendente de los servicios públicos a un enfoque de diseño e implementación basado en las necesidades de los usuarios. Al permitir la personalización de los servicios, la IA aumenta la relevancia y la capacidad de respuesta de las interacciones con cada ciudadano, haciendo que los servicios sean más accesibles y fomentando significativamente la participación pública (OECD/UNESCO, 2024^[131]). Por ejemplo, Francia está utilizando la IA para mejorar la implementación de servicios públicos, garantizando una relación más fluida entre los usuarios, sus necesidades individuales y los servidores públicos (Recuadro 5.48).

Recuadro 5.48. La asistencia de Francia en materia de IA para la implementación de servicios públicos personalizados

En Francia, la Direction Interministérielle de la Transformation Publique (Dirección Interministerial de Transformación Pública, DITP) ha puesto en marcha una innovadora plataforma impulsada por IA llamada Servicios Públicos+ para mejorar la calidad y la eficiencia de los servicios públicos en todo el país. La plataforma, en la que participan 45 millones de ciudadanos y 2,5 millones de servidores públicos, incorpora tres componentes principales de IA: funcionalidad de voz a texto para mejorar la accesibilidad; asistencia automatizada a los funcionarios en la redacción de respuestas; y análisis en tiempo real de las opiniones de los usuarios para identificar tendencias y puntos débiles.

La implementación del sistema mantiene un cuidadoso equilibrio entre la automatización y la supervisión humana. Los servidores públicos conservan el control total sobre las respuestas generadas por la IA, con la capacidad de modificar o rechazar sugerencias antes de enviarlas a la ciudadanía. De conformidad con los principios de la Ley de IA de la UE, se notifica a la ciudadanía cuando las respuestas han sido asistidas por IA, y todas las respuestas generadas por IA se publican como datos abiertos para garantizar la transparencia. El sistema se ha desplegado en 12 ministerios y 40 servicios públicos, y más de 4 000 funcionarios lo utilizan con regularidad.

El impacto de esta innovación ha sido significativo, con tiempos de respuesta promedio que han caído de 19 días a tres días. El sistema ha logrado altos índices de satisfacción: el 78 % de los servidores públicos consideran que las sugerencias de IA son útiles y el 70 % de los elementos sugeridos por la IA están validados por los servidores públicos. Lo que es más importante, la satisfacción de la ciudadanía con las respuestas ha aumentado: el 68 % de los usuarios encuentra útiles las respuestas asistidas por IA, frente al 57 % de las respuestas tradicionales. Esto demuestra cómo la IA puede integrarse de manera efectiva en los servicios públicos al tiempo que se mantienen los conocimientos humanos y se mejora tanto la eficiencia como la satisfacción del usuario.

Fuente: <https://www.plus.transformation.gouv.fr>, presentación de funcionarios franceses en la Cumbre de Acción sobre la IA (febrero de 2025).

Además de su uso en chatbots, los sistemas de IA generativa (como generadores de texto, imagen, vídeo, código o audio) pueden personalizar aún más los servicios para la ciudadanía. Con capacidad para crear perfiles de usuario y microclasificar a la ciudadanía en tipos específicos de grupos de usuarios, la IA puede adaptar la información y los propios servicios a las complejas necesidades de los usuarios y grupos de usuarios, incluidos los que se encuentran en condiciones vulnerables y desfavorecidas. Esto puede utilizarse, por ejemplo, para garantizar que las personas que reúnen los requisitos tengan conocimiento de los programas sociales y las prestaciones disponibles, en caso de que puedan proporcionarlas de forma preventiva (OECD, 2024_[210]) (Recuadro 5.49). También puede adaptar los servicios en torno a eventos vitales anticipando necesidades de servicios dependientes del contexto en una variedad de ámbitos, como la educación, el transporte y el empleo (Kopponen et al., 2024_[211]). También puede ayudar en los servicios públicos de empleo, por ejemplo, emparejando empleadores y demandantes de empleo (Recuadro 5.50).

Recuadro 5.49. Cómo los sistemas de protección social están adoptando IA

Los avances en los datos y la tecnología están ayudando a mejorar la accesibilidad, la cobertura y la implementación de prestaciones y servicios sociales, aunque la no aceptación sigue siendo un problema importante. Los avances recientes se centran en vincular datos administrativos de diferentes fuentes para mejorar la inscripción en programas sociales. Al vincular conjuntos de datos, los gobiernos pueden agilizar los procesos de solicitud y renovación, informar a los posibles beneficiarios de su elegibilidad para los beneficios y (en casos limitados) inscribir automáticamente a los usuarios elegibles.

A pesar de su potencial, los ministerios de asuntos sociales han utilizado poco los usos avanzados de la IA, incluso para la adopción de programas. Los usos más comunes de la IA en los programas sociales de los países de la OCDE han sido la atención automatizada al cliente (por ejemplo, chatbots y asistentes digitales), la automatización de los procesos internos (por ejemplo, entrada de datos y procesamiento de documentos), la detección de fraudes y la detección de anomalías en las solicitudes de prestaciones.

A pesar de este uso limitado, la IA ofrece una promesa significativa en materia de protección social, incluso a través de la identificación y la inscripción de usuarios de servicios, la mejora de la asistencia al usuario y la mejora de la eficiencia y la puntualidad de la implementación de beneficios y servicios en respuesta a las necesidades cambiantes. Entre los usos emergentes se incluyen:

- Avances en los usos actuales de chatbots, automatización de tareas rutinarias y detección de fraudes.
- Mejorar el servicio al cliente acompañando a los proveedores de servicios humanos en situaciones delicadas.
- Una mejor automatización de las decisiones sobre prestaciones basada en las normas legales de elegibilidad.
- Nuevos tipos de análisis predictivo, como la identificación de riesgos y vulnerabilidades recurrentes para orientar las intervenciones de los programas sociales hacia personas y grupos de riesgo.

Quizás el posible caso de uso más gratificante de la IA en la protección social sea identificar a los beneficiarios de alto riesgo y llegar a ellos de forma proactiva, o inscribirlos automáticamente, para mejorar la cobertura de la protección social. Se está empezando a emplear en las esferas de la prevención y el alivio de las personas sin hogar y de las evaluaciones de riesgos para las sobrevivientes de la violencia doméstica. En este último caso, las sobrevivientes se conectan con trabajadores sociales y otros proveedores de servicios pertinentes con el objetivo de prevenir la repetición de la violencia.

Estas oportunidades conllevan riesgos, y los gobiernos (por ejemplo, Alemania) están implementando una serie de medidas para garantizar que se mitigan, incluidos los marcos legales, regulatorios y de responsabilidad adecuados para proteger la privacidad de las personas y regular el uso de sistemas automatizados.

Fuente: (OECD, 2024^[210]), Frey et al (de próxima publicación), "Inteligencia artificial y el futuro de la protección social en los países de la OCDE"; respuestas al cuestionario *Harnessing Technology and Data to Improve Social Protection Coverage and Social Service Delivery*, administrado por la División de Política Social de la OCDE; (OECD/UNESCO, 2024^[131]); (Policy Lab Digital, Work & Society within the German Federal Ministry of Labour and Social Affairs, 2024^[212]).

Recuadro 5.50. Los PES de los países de la OCDE recurren cada vez más a la IA para la implementación de servicios

La mitad de los Servicios Públicos de Empleo (PES) de los países de la OCDE han desplegado soluciones basadas en la IA, incluidas soluciones generativas de IA que utilizan modelos de aprendizaje permanente, para ayudar en sus actividades y servicios a solicitantes de empleo, personas en riesgo de perder el empleo y empleadores. En la actualidad, están más acostumbrados a:

- Perfillar a los demandantes de empleo y predecir la probabilidad de que un solicitante de empleo permanezca en el desempleo.
- Emparejar demandantes de empleo con vacantes adecuadas.
- Contribuir al diseño de los anuncios de vacantes, incluida la clasificación profesional.
- Proporcionar información a los clientes a través de chatbots o asistentes virtuales.
- contribuir a la gestión de las carreras y a la orientación en la búsqueda de empleo de los demandantes de empleo.

Por ejemplo, desde 2024, el Servicio de Empleo de Israel ha integrado motores de IA en sus sistemas centrales para agilizar los procesos de búsqueda de empleo de los solicitantes de empleo. Las soluciones de IA ofrecen numerosas oportunidades para los SPE, como la mejora de la orientación de las medidas y los servicios, la optimización del uso de los datos, la reducción de las cargas administrativas y la habilitación de sistemas digitales más dinámicos. Sin embargo, los PES también deben tener en cuenta varios desafíos y riesgos, como la necesidad de definir una responsabilidad clara para los sistemas de IA, abordar las preocupaciones sobre la transparencia y la explicabilidad, garantizar la calidad y la privacidad de los datos y gestionar la posible resistencia o la falta de competencias tanto entre el personal como entre los clientes. Por ejemplo, el PES austriaco implementó un chatbot habilitado para ChatGPT a principios de 2024 para proporcionar información sobre formación y orientación profesional a los demandantes de empleo. Sin embargo, esta innovación fue criticada poco después de su lanzamiento, sobre todo debido a los informes de que las recomendaciones del chatbot perpetuaban las diferencias entre hombres y mujeres.

Para hacer frente a estos desafíos, algunos países de la OCDE están liderando medidas innovadoras para promover una mejor gobernanza de la IA en sus SPE. Esto incluye el establecimiento de comités de ética multidisciplinarios por parte de los SPE en Francia y Bélgica, y el desarrollo de una estrategia específica por parte de los SPE de Noruega para establecer las ambiciones, habilitadores y principios para el uso de la IA.

Fuente: (Brioscú et al., 2024^[213]), <https://www.maariv.co.il/business/career/article-1122739>.

Los servicios de salud representan un ámbito especialmente prometedor, ya que los expertos en IA lo consideran una de las áreas con mayor potencial (OECD, 2024^[29]). Los sistemas de IA pueden sugerir diagnósticos y predecir factores de riesgo para ayudar a introducir medidas preventivas (Khalifa and Albadawy, 2024^[214]; Alowais et al., 2023^[215]). Esto fue particularmente útil durante la pandemia de COVID-19. Por ejemplo, Grecia desarrolló y utilizó ML para optimizar las pruebas en frontera, centrándose en los viajeros asintomáticos en función de los datos demográficos y los resultados de pruebas anteriores (Bastani et al., 2021^[216]). En algunos contextos, la IA también puede utilizarse para predecir la probabilidad de enfermedad o lesión. Por ejemplo, en Portugal, la aplicación PrevOcupAI se aplica para predecir enfermedades relacionadas con el trabajo y factores de riesgo relacionados en la administración pública (EC JRC, 2021^[217]). En Arabia Saudí, la IA se está empleando para predecir y reducir la ausencia de pacientes en los consultorios ambulatorios, mejorando significativamente la utilización de los recursos⁴⁰.

En Corea, la ciudad de Incheon está utilizando tecnología de IA para mejorar la accesibilidad sanitaria de residentes y visitantes extranjeros (Recuadro 5.51).

Combinada con el conocimiento de los médicos y otros expertos médicos, la IA puede conducir a una mejor precisión, una mayor eficiencia y resultados más positivos para las personas. Los beneficios pueden ir más allá del bienestar de la ciudadanía; en el Reino Unido, se estima que los ahorros para el presupuesto público a través de la expansión de la atención médica preventiva habilitada por la IA valdrán 6 000 millones de libras esterlinas al año para 2040 (TBI, 2024^[218]).

Recuadro 5.51. Comunicación de salud multilingüe impulsada por IA en Corea

En Corea, la ciudad de Incheon ha implementado SIMTOMI, un servicio de apoyo médico basado en IA que cierra la brecha lingüística entre los pacientes extranjeros y los proveedores de atención médica. Este servicio combina un verificador de síntomas de IA multilingüe con capacidades avanzadas de traducción de idiomas, lo que permite a los extranjeros y turistas comunicar con precisión sus preocupaciones de salud a los profesionales médicos y farmacéuticos locales. El sistema genera informes resumidos de síntomas anonimizados que pueden compartirse a través de códigos QR durante las visitas en persona al hospital o a la farmacia. El impacto de esta innovación ha sido significativo, ya que el servicio ha logrado una alta puntuación de satisfacción de 84,4 puntos y una tasa de intención de reutilización del 95,0 %, superando los objetivos iniciales.

Fuente: <https://oecd-opsi.org/innovations/simtomi-ai-powered-multilingual-medical-support>.

Previsión de futuras necesidades de servicio público

La IA también tiene un amplio potencial de uso en la realización de análisis predictivos para prever futuras necesidades de servicios, así como para preparar y ajustar la capacidad de respuesta en consecuencia. Esto puede complementar la elaboración de perfiles de los usuarios de servicios y sus necesidades, ayudando a dirigirse y proponer mejor los servicios públicos a los usuarios. El potencial del análisis predictivo es bien conocido en el campo de la policía y la justicia (Macnish, Wright and Jiya, 2021^[219]), como se analiza en las secciones "La IA en la aplicación de la ley y la gestión del riesgo de desastres" y "La IA en la administración de justicia y el acceso a la justicia" de este capítulo. Para optimizar los sistemas a gran escala, los sistemas predictivos de IA han demostrado mejoras concretas en la calidad del servicio, como la optimización del tráfico, la disminución de los accidentes de tráfico y el aumento de la eficiencia del combustible en los sistemas de transporte público (Jevinger et al., 2023^[220]). En Bélgica, la Agencia de Carreteras y Tráfico utilizó la IA para predecir las condiciones resbaladizas de las carreteras y ayudar a asignar recursos para mantenerlas libres de hielo. En los Países Bajos, la IA se utiliza para experimentar con las predicciones de tráfico y la previsión de accidentes de tráfico. En Portugal, la aplicación PrevOcupAI se utiliza para predecir enfermedades relacionadas con el trabajo y factores de riesgo relacionados en la administración pública (EC JRC, 2021^[217]).

Un paso más allá es la tendencia emergente de implementación de servicios proactivos, donde la IA puede ayudar a anticipar las necesidades de las personas más allá del contexto de servicio público actual y responder rápidamente a ellas (Pawlowski and Scholte, 2023^[221])⁴¹. Por ejemplo, esto puede implicar sistemas ciudadanos digitales gemelos (CDT), que recopilan datos sobre la situación general de una persona y luego ayudan a estructurar respuestas de servicio de andamiaje a sus eventos vitales (Kopponen et al., 2024^[211]). De este modo, se pueden crear combinaciones de servicios adaptadas de forma proactiva para satisfacer las necesidades específicas de las personas en función de su situación actual o de los acontecimientos vitales previstos en el futuro (por ejemplo, terminar la escuela e

incorporarse al mercado laboral). En Singapur, la aplicación " Moments of Life " utiliza algoritmos de IA para analizar la información de los usuarios y enviar notificaciones personalizadas (por ejemplo, para próximas vacunas o plazos de inscripción escolar) (Aldemir and Uçma Uysal, 2025^[222]).

Recuadro 5.52. Cómo están utilizando las ciudades las herramientas de IA para prestar servicios

La IA está empezando a desempeñar un papel fundamental en la implementación de servicios públicos y en la resolución de los desafíos de desarrollo urbano. Por ejemplo:

- Las ciudades están utilizando la IA para optimizar el tráfico y cartografiar las brechas y necesidades de transporte. Los Ángeles (Estados Unidos) utiliza la IA para detectar y responder a los incidentes de tráfico, como accidentes o cierres de carreteras, minimizando su impacto en el flujo de tráfico. Los sistemas basados en IA analizan los datos de las cámaras de tráfico, los sensores y las redes sociales para ofrecer alertas en tiempo real y sugerir rutas óptimas para los conductores, reduciendo así la congestión general. Sydney (Australia) está utilizando IA para optimizar el flujo de tráfico y probar autobuses sin conductor, sentando las bases para una adopción más amplia.
- Las ciudades utilizan la IA para ofrecer servicios públicos e información. Buenos Aires (Argentina) creó un chatbot llamado Boti, primero como canal oficial del gobierno para la vacunación contra la COVID-19, y ahora como canal para proporcionar información sobre temas como el uso compartido de bicicletas y el cuidado social. Yokosuka (Japón) está desarrollando un servicio de prevención de la demencia que utiliza IA generativa conversacional para reforzar la salud mental de la ciudadanía de edad avanzada.
- El Proyecto Ellinikon de Grecia tiene como objetivo transformar el antiguo aeropuerto internacional de Atenas en un distrito inteligente de vanguardia. Las tecnologías avanzadas, como el Internet de las cosas, el análisis de inteligencia artificial, la conectividad 5G y la fibra óptica de ultra alta velocidad, se integrarán perfectamente en el tejido de la ciudad. Se desplegarán sensores inteligentes y sofisticados algoritmos de IA para gestionar la infraestructura de manera eficiente, mejorando la asignación de recursos y la implementación de servicios. Las soluciones de movilidad inteligentes, como la asistencia inteligente para estacionarse y las opciones de transporte compartido, agilizarán aún más el movimiento urbano. Además, la IA desempeñará un papel fundamental en la optimización de la gestión del agua y los residuos, garantizando la sostenibilidad ambiental. Se creará un Gemelo Digital de la ciudad para supervisar y optimizar el consumo de recursos y las condiciones ambientales en tiempo real.

La IA generativa, en particular, tiene el potencial de transformar la forma en que funcionan las ciudades; puede predecir la demanda y racionalizar los servicios públicos, mejorar la capacidad de innovación de los gobiernos locales, mejorar la toma de decisiones basada en datos, optimizar la asignación de recursos y agilizar los procesos administrativos. Esto es posible, ya que algunas ciudades recopilan datos de forma continua de vehículos, edificios, infraestructuras e individuos a través de sensores, drones y dispositivos móviles para presentar una imagen de hasta el segundo minuto. Por ejemplo

- Ámsterdam (Países Bajos), Eslöv (Suecia), Helsinki (Finlandia) y Múnich (Alemania) utilizan gemelos digitales para mejorar la planificación y la gestión urbanas y construir ciudades más resilientes ante amenazas futuras. La réplica digital, actualizada constantemente con datos de su gemelo físico, permite a las ciudades analizar datos, monitorear sistemas y ejecutar simulaciones y construir escenarios, cada uno evaluado con arreglo a criterios clave como la

accesibilidad social, la sostenibilidad ambiental y el rendimiento financiero, facilitando la toma de decisiones sobre vivienda, uso del suelo, densidad, infraestructura verde y transporte.

- Shanghái (China) utiliza gemelos digitales e Internet de las cosas impulsados por GenAI para gestionar su compleja infraestructura urbana combinando datos en tiempo real de dispositivos de Internet de las cosas con sofisticados análisis. Esto ayuda a la ciudad a reducir el tráfico, aumentar la efectividad del transporte público, optimizar el uso de la energía, mejorar la preparación ante desastres y planificar infraestructuras.

A pesar de su potencial, las ciudades se encuentran en una fase temprana del uso de GenAI. Según una encuesta realizada por Bloomberg Philanthropies en 2023 a 100 ciudades de todo el mundo, solo el 2 % implementaba activamente la GenAI en sus actividades, mientras que el 69 % exploraba o probaba la tecnología, pero todavía no había desarrollado capacidades ni políticas. Entre los obstáculos se incluyen la falta de concienciación, las limitaciones presupuestarias, la falta de directrices y de gobernanza de datos pertinentes, la falta de competencias, las preocupaciones en materia de privacidad y seguridad de los datos, y los desafíos administrativos y técnicos a la hora de integrar tecnologías de IA como la IA geoespacial (GeoAI) en los procesos de planificación y diseño de servicios.

De cara al futuro, para aprovechar el potencial y la adopción de la IA, las ciudades deberían:

- Cerrar las brechas digitales, en coordinación con el gobierno nacional, para garantizar el acceso a los servicios de Internet para toda la población.
- considerar y abordar proactivamente los futuros desafíos, riesgos y posibles inconvenientes de la IA (por ejemplo, problemas de privacidad de los datos y de calidad y disponibilidad de los datos que limitan la toma de decisiones).
- Desarrollar, con el apoyo de los gobiernos nacionales, los sistemas regulatorios para garantizar que los sistemas de IA utilizados sean seguros, transparentes, rastreables y no discriminatorios para preservar la privacidad, los valores éticos y los derechos individuales.
- Explorar entornos de prueba regulatorios para fomentar alianzas público-privadas en un ambiente controlado e innovador y probar ideas de manera segura.
- invertir en alfabetización en IA para empleados públicos locales, desarrollar experiencia especializada en IA dentro de sus departamentos jurídicos.
- Establecer alianzas con instituciones académicas y empresas tecnológicas para compartir conocimientos.

Fuente: (OECD, 2024^[223]; 2023^[224]; Bloomberg Philanthropies, 2023^[225]),<https://theellinikon.com.gr/en/smart-and-sustainable/smart-city>.

Evidencia de impacto

En un estudio reciente sobre la automatización potencial de los servicios públicos del Reino Unido, el Instituto Alan Turing descubrió que la IA podría ayudar a automatizar el 84 % de las transacciones relacionadas con los servicios y ahorrar el equivalente a aproximadamente 1.200 personas/año trabajado cada año (Straub et al., 2024^[199]). Además, la reducción de la burocracia y las mejoras en la regulación en torno al diseño y la implementación de los servicios públicos podrían beneficiar enormemente a los usuarios en términos de ahorro de tiempo, esfuerzo y recursos, o en la aceleración del tiempo de respuesta del gobierno.

Si bien la promesa es grande, hasta el momento hay poca evidencia empírica de que la IA tenga o vaya a tener estos beneficios. El conocimiento actual sobre la adopción de la IA en la función pública se basa en estudios de casos que sintetizan las primeras evidencias extraídas de las prácticas implementadas⁴².

Los casos prácticos se están multiplicando rápidamente y en las siguientes secciones se ofrece una visión general de las tendencias de innovación que están surgiendo.

La mayor parte de la evidencia sobre el impacto de las últimas innovaciones en materia de IA en los servicios públicos se basa en proyectos y casos individuales y, con mayor frecuencia, se denuncia por sí misma. La revisión de casos en la literatura indica una indicación general positiva de los beneficios de la IA. Sin embargo, no es muy frecuente que los proyectos de IA divulguen por adelantado sus posibles impactos negativos (Birhane et al., 2022^[226]). Esto es de esperar, ya que las aplicaciones son todavía relativamente nuevas, siguen surgiendo problemas y a menudo no se llevan a cabo evaluaciones.

A medida que la IA sigue remodelando los servicios públicos, es esencial realizar pruebas rigurosas de las intervenciones conductuales para garantizar que estas tecnologías no solo sean eficientes, sino también efectivas para influir en la participación ciudadana y la toma de decisiones (Soman, 2024^[227]). Algunos equipos gubernamentales ya están creando prototipos y probando rápidamente soluciones basadas en la IA. Sin embargo, los métodos de evaluación más estructurados, como los RCTs, el análisis de diferencias en las diferencias y las pruebas A/B, pueden ayudar a determinar qué aplicaciones de IA funcionan mejor, para quién y en qué contextos. La colaboración con el mundo académico para que estos enfoques experimentales formen parte de la aplicación de la IA puede aportar pruebas empíricas más sólidas sobre el impacto de la IA en la vida real, garantizando que los servicios públicos se diseñan y adaptan en función de análisis conductuales, en lugar de supuestos.

Gestión de riesgos y desafíos

Si bien la IA tiene un enorme potencial para mejorar la eficiencia, la efectividad y la capacidad de respuesta, existen riesgos y desafíos que pueden limitar su adopción y provocar daños en el mundo real. Un estudio de 2022 sobre las lecciones aprendidas de 61 iniciativas de automatización de la función pública canceladas en varios países revela que muchas han fracasado debido a la falta de efectividad, la protección de la sociedad civil, la investigación crítica de los medios de comunicación y las acciones legales, entre otras cuestiones (Redden et al., 2022^[228]). Los principales riesgos y desafíos identificados se analizan a continuación.

Riesgos asociados

- Uso indebido o cuestionable de la IA, que genera problemas de vigilancia y privacidad.
- Datos inadecuados o sesgados en los sistemas de IA.
- Falta de transparencia y explicabilidad.
- Exacerbación de la exclusión social y las brechas digitales.

El uso de IA en las innovaciones de servicio público requiere un mayor acceso a la información personal de los usuarios, especialmente cuando se conecta a una implementación de servicios proactiva y personalizada. A menudo esto también requiere tener acceso a una variedad de bases de datos de organizaciones públicas e información de servicio para identificar el mejor enfoque de servicio. La implementación proactiva de servicios, por ejemplo, suele requerir un procesamiento algorítmico de datos dispersos entre múltiples fuentes de datos públicas y, en ocasiones, privadas (Nikiforova et al., 2023^[229]). El uso de la IA plantea preocupaciones sobre **la privacidad y el uso intencionado y responsable de los datos** y puede hacer que estos servicios sean más contestables (véase el Capítulo 1, sección "Para aprovechar los beneficios de la IA, es necesario gestionar sus riesgos"). Una suposición comúnmente aceptada es que las herramientas inteligentes de IA adaptadas y entrenadas en datos enfrentan concesiones ineludibles entre la capacidad (predictiva) y la privacidad. Sin embargo, ha habido un gran interés reciente en la investigación para eludir esta hipótesis general gracias a las tecnologías de

preservación de la privacidad (PPT) o las tecnologías de mejora de la privacidad (PETs), que utilizan datos anonimizados o cifrados (OECD, 2023^[230])⁴³.

Si los sistemas de IA se basan en **datos inadecuados o sesgados**, podrían producirse resultados inexactos o adversos para algunas personas o grupos. Con respecto al diseño y la implementación de servicios públicos, esto podría dar lugar, por ejemplo, a un trato preferencial indebido para algunas personas o grupos o a un refuerzo de estereotipos inexactos. En parte debido a que los datos son inadecuados o sesgados, las brechas de datos digitales —donde algunos grupos están más representados que otros en los datos— **limitan el potencial de los servicios de IA personalizados**, lo que solo los hace útiles y precisos para poblaciones ricas en datos y puede dejar a algunos grupos menos representados (UNESCO, 2019^[231]; Perry and Lee, 2019^[232]; Dieterle, Dede and Walker, 2024^[233]).

Dada la naturaleza de los servicios públicos, **es crucial garantizar la rendición de cuentas, transparencia y explicabilidad de la toma de decisiones y resultados de los sistemas y herramientas de IA**. Para ello, es necesario crear e implementar marcos de servicio público que definan líneas claras de responsabilidad para los sistemas de IA a lo largo de todo su ciclo de vida dentro de los servicios públicos. Cuanto mayor sea la transparencia en el uso de la IA en los servicios públicos, más fácil será hacer un seguimiento de los efectos adversos y generar una mayor sensación de explicabilidad y confianza en los resultados de la IA (Sun and Medaglia, 2019^[197]; Thiebes, Lins and Sunyaev, 2020^[234]).

Desafíos de implementación

- Elevados costos de adopción y escalabilidad de la IA.
- Falta de marcos viables y directrices sobre el uso de la IA.
- Dificultad para establecer un seguimiento continuo del progreso de la vía y de los riesgos superficiales.
- Falta de datos de alta calidad y capacidad para compartirlos.
- Brechas de habilidades.

No se puede ignorar el **costo asociado a la creación y el mantenimiento de sistemas de IA para los servicios públicos** (desde la infraestructura, el seguimiento, la formación del personal y la educación). Por ejemplo, el costo de las opciones de suscripción a la IA comercial podría aumentar cuando las empresas traten de volverse rentables y cubrir grandes gastos generales y de explotación para satisfacer las expectativas de los inversores (Shark, 2025^[235]).

La adopción confiable de sistemas de IA en el diseño y la implementación de servicios públicos requiere un **seguimiento continuo para hacer un seguimiento de los avances y ayudar a identificar los riesgos emergentes**, pero los gobiernos carecen en general de esta función. Los sistemas de IA totalmente autónomos que tienen un impacto en la vida de la ciudadanía pueden plantear problemas jurídicos y contribuir aún más a la desconfianza y la resistencia en la adopción. En consecuencia, en las políticas públicas a menudo se requiere o se da prioridad a la determinación y la supervisión humanas. Por ejemplo, la Ley de IA de la UE impone requisitos de supervisión humana en casos de uso de "alto riesgo", que se consideran muchos usos gubernamentales. Sin embargo, confiar solo en la intervención humana puede crear una falsa sensación de seguridad al adoptar algoritmos (Green, 2022^[236]). Contar con seres humanos en el bucle no anula la necesidad de llevar a cabo análisis de costo-beneficio y beneficio-riesgo para sopesar las compensaciones ni de realizar una evaluación a lo largo del ciclo de vida de la IA. En consecuencia, la adopción de la IA en la implementación de servicios públicos requiere una evaluación técnica y detallada desde la perspectiva del valor público en un contexto en el que pueden manifestarse efectos adversos durante la implementación del sistema de IA (Medaglia, Gil-Garcia and Pardo, 2021^[237]; Madan and Ashok, 2023^[194]).

Algunos de los primeros fallos en el uso de la IA en los servicios públicos y casos de errores se han relacionado con **problemas de calidad de los datos**. Un estudio de la OCDE (2019^[191]) ha identificado la necesidad de verificar continuamente la exactitud, confiabilidad y adecuación de las entradas de datos a los sistemas de IA, lo que ayuda a los proveedores de servicios a identificar y abordar los errores presentes en los datos.

Por último, **los distintos niveles de competencias en materia de IA y las competencias asociadas** en el gobierno son problemáticos (Medaglia, Mikalef and Tangi, 2025^[238]). Si bien las herramientas generativas de IA han pasado a ser ampliamente accesibles para la adaptación ascendente, no ha habido suficiente capacitación en IA en el gobierno ni directrices claras para su uso en organizaciones públicas para desarrollar e informar los servicios públicos. Además, dado el rápido desarrollo de los sistemas de IA, el aprendizaje continuo es una necesidad para mantenerse al día. Las competencias necesarias para adaptar los sistemas de IA no son solo tecnológicas, sino también éticas, jurídicas, de gestión y sociotécnicas; los administradores públicos deben comprender cómo interactúan los sistemas de IA tanto con la organización como con el público en general (Trajkovski, 2024^[239]). Por ejemplo, la creatividad y las habilidades de pensamiento de diseño son necesarias para implementar servicios centrados en el ser humano adaptados a las necesidades públicas. Esto también requiere altos niveles de habilidades interpersonales, ya que los administradores necesitan colaborar de manera efectiva con científicos de datos, abogados, especialistas en ética y socios externos. La falta de estas habilidades puede provocar retrasos en la adopción de la IA o peores impactos negativos, como sistemas algorítmicos sesgados u opacos. En la actualidad, faltan en gran medida enfoques sistémicos para desarrollar dichos conjuntos de habilidades en los distintos gobiernos. Algunos han propuesto una nueva profesión de profesionales funcionales de la IA que actúen como intermediarios entre el lado técnico y el de las políticas públicas y ayuden a formar a los funcionarios públicos en consecuencia (Labanava et al., 2022^[240]).

Potencial sin explotar y camino a seguir

Existe un gran potencial para utilizar la IA en otras áreas de servicio y en nuevos casos prácticos. Por ejemplo, en el ámbito de la salud, la IA podría permitir intervenciones preventivas y adaptadas, así como "estímulos" conductuales informados, que conduzcan a mejores resultados y permitan a los profesionales sanitarios disponer de más tiempo para la atención. También podría generar nuevas técnicas para liberar valor de los vastos activos de datos de salud, el 97 % de los cuales siguen sin explotarse en los países de la OCDE (Sumner et al., 2023^[241]; Bennett Institute, 2024^[242]; OECD, 2024^[243]). En este ámbito, la norma de la OCDE (2017^[244]) para promover el uso de datos de salud en el interés público guía a los gobiernos para que estos datos estén más disponibles para la innovación en la atención médica. El análisis de los desafíos para la implementación de los principios de esta recomendación se puede encontrar en (OECD, 2022^[245]) y en un próximo informe de la OCDE: "Facilitación del uso secundario de datos sanitarios con fines de interés público a través de las fronteras". Los avances de la IA también podrían ayudar a reducir la escasez de personal sanitario prevista en 3,5 millones para 2030 en los países de la OCDE (OECD, 2024^[243]).

El uso de la IA en el diseño y la implementación de servicios públicos todavía se encuentra en una fase experimental. Si bien hay ejemplos de automatización, racionalización e incluso nuevos tipos de servicios emergentes, las organizaciones gubernamentales todavía están lejos de una amplia adopción y uso operativo. El bombo publicitario de la IA y las demandas de eficiencia de costos pueden empujar a los gobiernos hacia una rápida automatización. Sin embargo, las evaluaciones y críticas de las prácticas tempranas, como se analiza en esta sección, sugieren que es más probable que un enfoque más gradual adoptado junto con los servidores públicos tenga éxito. Los gobiernos necesitan estrategias sobre cómo utilizar modelos híbridos en los que la IA aumente la toma de decisiones humana, y sobre cómo hacer frente al papel cambiante de los funcionarios públicos en este nuevo formato de burocracia algorítmica (Vogl et al., 2019^[246]). Si bien la promesa de la IA puede consistir en liberar tiempo de los servidores

públicos para tareas de mayor valor —como colaborar con la ciudadanía para abordar cuestiones sociales complejas—, también puede tener el efecto contrario, ya que el distanciamiento de los servidores públicos de la sociedad disminuye a medida que su papel en la recopilación directa de datos y la interacción con los usuarios de los servicios disminuye (Bullock, Young and Wang, 2020^[247]; Madan and Ashok, 2023^[194]) (véase también el debate sobre la "IA en la reforma de la función pública" en este capítulo).

Otro enfoque emergente para la automatización en los servicios públicos son las reglas como código (RaC), que implica codificar las reglas gubernamentales en formatos consumibles por máquina junto con el texto legal tradicional (Mohun and Roberts, 2020^[4]). Dado que los servicios dependen cada vez más de la IA para apoyar la toma de decisiones, la RaC proporciona una base para aplicaciones de IA más precisas, transparentes y escalables en el gobierno. Al incorporar normas en formatos digitales, la RaC podría mejorar la coherencia de la implementación de servicios impulsada por la IA, garantizando que los sistemas automatizados interpretan y aplican la normativa correctamente en diferentes contextos. Varios gobiernos han comenzado a utilizar la RaC para agilizar el cumplimiento normativo y mejorar la implementación de servicios impulsada por la IA. Por ejemplo, los Emiratos Árabes Unidos han puesto en marcha una plataforma nacional de Normas como Código cuyo objetivo es desarrollar leyes y reglamentos basados en la IA, transformando el ecosistema financiero⁴⁴.

A diferencia de los enfoques convencionales basados en tareas, en los que la implementación de servicios se segmenta en distintas operaciones secuenciales, los enfoques basados en IA pueden integrar funciones y procesos de toma de decisiones de forma simultánea. Esto representa un cambio potencialmente transformador en la forma en que se organizan los servicios públicos.

En lo que respecta a una mayor aceptación de los servicios públicos basados en la IA por parte de la ciudadanía y los residentes, de las investigaciones existentes se desprenden algunas ideas claras. En general, se acepta que la IA es mejor para procesar grandes cantidades de datos e identificar patrones (Rane, Choudhary and Rane, 2024^[248]). Sin embargo, una combinación de factores, como la percepción de la utilidad de los sistemas de IA, la expectativa de rendimiento, las actitudes, la confianza y la expectativa de esfuerzo tienden a influir en la voluntad de uso de la IA en todos los sectores (Kelly, Kaye and Oviedo-Trespalacios, 2023^[249]). La percepción del riesgo y la confianza en los sistemas de IA difieren cuando se trata de servicios públicos generales (es decir, prestados por el gobierno sin una solicitud específica y que conciernen a todos o a la mayoría de la ciudadanía, como la educación básica y la seguridad pública) frente a servicios públicos específicos (es decir, solicitados explícitamente por la ciudadanía y que solo afectan a uno o pocos ciudadanos, como los programas de atención a las personas mayores o la asistencia para la vivienda para poblaciones de bajos ingresos). Los servicios generales son más abstractos y los usuarios tienen un menor nivel de conocimiento de la situación, lo que significa que también pueden aceptar la IA con mayor facilidad⁴⁵. En servicios específicos, la oportunidad de decidir adquiere relevancia; los usuarios en general desean tener la posibilidad de elegir si se ha utilizado una aplicación de IA en su decisión sobre el servicio (Gesk and Leyer, 2022^[203]).

Las primeras investigaciones también indican que la confianza en las aplicaciones de IA es mayor si se integran en servicios ya existentes en lugar de en servicios completamente nuevos (Aoki, 2020^[250]). Esto significa que las innovaciones de servicios más proactivas pueden tener que demostrarse ante los usuarios y pueden enfrentar un mayor escrutinio. En todos los casos, la participación de los usuarios en la creación conjunta de servicios públicos está correlacionada con una percepción positiva de las decisiones en materia de IA y podría impulsar una mayor adopción (Gesk and Leyer, 2022^[203]). Esto es muy importante, ya que la experiencia de las personas con los servicios puede influir en la forma en que perciben a sus gobiernos en general, y la satisfacción pública con los servicios administrativos y sociales es un importante motor de confianza (OECD, 2024^[46]). Las organizaciones públicas pueden poner en riesgo su legitimidad democrática si la ciudadanía no confían en los servicios que el gobierno pretende prestar con IA (Aoki, Tay and Yarime, 2024^[251]; Aoki, 2020^[250]).

Asimismo, los gobiernos deben comprender y tener más en cuenta las necesidades de los usuarios a la hora de diseñar servicios basados en IA que probablemente deseen utilizar. Esto implicará garantizar que las políticas y los servicios estén diseñados para la forma en que las personas se comportan realmente, en lugar de para la forma en que se supone que se comportan. Incluso las políticas bien intencionadas pueden fallar cuando introducen fricciones innecesarias (“de fricción”), lo que dificulta el acceso de las personas a los servicios, la cumplimentación de las solicitudes o la toma de decisiones informadas. Un mayor uso de la ciencia del comportamiento puede contribuir a ello, al ofrecer instrumentos prácticos para reducir la fricción, mejorar la accesibilidad y garantizar que los servicios estén centrados en el usuario y sean confiables para la ciudadanía. Al abordar las barreras cognitivas, como la complejidad, la fatiga de la decisión y la inercia, los gobiernos pueden diseñar servicios más simples, intuitivos y fáciles de navegar. Las herramientas impulsadas por IA pueden identificar servicios que requieren auditorías de fricción, evaluar la efectividad de las intervenciones y mejorar el diseño de los servicios. La NLP y el análisis de opiniones pueden analizar las opiniones del público en tiempo real, indicando los puntos débiles de las plataformas en línea y los canales de atención al cliente. La IA también puede cuantificar las interacciones de los usuarios y adaptar las intervenciones a los distintos grupos demográficos, garantizando así que las políticas estén mejor orientadas y sean más inclusivas. A medida que los gobiernos avancen hacia la gobernanza digital por diseño e impulsada por los datos, la integración de nociones sobre el comportamiento con la IA será clave para que los servicios públicos sean más fáciles de usar, eficientes y receptivos (OECD, 2024^[252]).

Por último, existe una brecha considerable entre la velocidad a la que se está introduciendo la IA en los servicios públicos y la medida en que se llevan a cabo evaluaciones sólidas. Los análisis más exhaustivos ya apuntan a muchos efectos imprevistos y, en ocasiones, adversos que pueden aparecer junto a los aumentos generales de eficiencia. Esta área necesita más inversión.

En todos estos ámbitos, la cooperación internacional en el desarrollo de la capacidad del sector público en materia de IA es crucial. Cuanto más compartan los gobiernos las prácticas de IA en el desarrollo de los servicios públicos (por ejemplo, compartiendo algoritmos abiertos, infraestructuras, conjuntos de datos intergubernamentales y esfuerzos conjuntos para el desarrollo responsable de tecnologías emergentes), más probabilidades habrá de garantizar la calidad en general. Iniciativas multilaterales como la Global Partnership on Artificial Intelligence (GPAI) pueden desempeñar un papel crucial en este sentido.

IA en la aplicación de la ley y la gestión del riesgo de desastres

Las herramientas de IA se utilizan en diversos organismos públicos con un papel en la gobernanza de riesgos críticos, como la gestión de catástrofes, la aplicación de la ley y ámbitos interrelacionados como la lucha contra el terrorismo, las aduanas y la gestión de fronteras. Los organismos pertinentes utilizan tecnologías de IA para mejorar la escala, velocidad y precisión de las capacidades de análisis anticipatorio, vigilancia y supervisión en tiempo real y procesos administrativos, al tiempo que reducen los costos operativos. Esto puede mejorar las capacidades de investigación y gestión de crisis, optimizar la asignación de recursos y mejorar los tiempos de respuesta.

Dado que el uso de la IA tiene un gran potencial de impacto en este ámbito, especialmente en lo que respecta a la aplicación de la ley, se considera que estas entidades son usuarios finales de IA de alto riesgo (OECD, 2020^[253]) y, por lo tanto, deben tener consideraciones adicionales a la hora de recurrir a la IA en aras del bien público y la seguridad. Por encima de todos los demás tipos de instituciones públicas, la ciudadanía deposita el mayor grado de confianza en la policía (OECD, 2024^[46]). Los organismos pertinentes deben mantener esta confianza garantizando que la IA se adopte de una manera confiable que mitigue los desafíos éticos y los riesgos para la protección de los derechos personales, en particular, mediante la vigilancia de la forma en que se obtienen, mantienen y aplican los datos subyacentes (OECD, 2020^[254]; 2019^[191]; 2022^[255]). Se pueden extraer lecciones de los casos de uso existentes y, como se analiza más adelante, algunos gobiernos han establecido mayores expectativas de rendición de cuentas por el uso de la IA por parte de estos agentes.

Situación actual

Identificación de presuntos delincuentes y personas desaparecidas

La tecnología de reconocimiento facial basada en IA mejora la capacidad de emparejar imágenes grabadas o en tiempo real de material de vigilancia, redes sociales y otras fuentes con un conjunto de bases de datos de imágenes. Esto puede utilizarse para ayudar a identificar a sospechosos de delitos, personas de interés público y personas desaparecidas. Esto también ha demostrado tener cierto efecto disuasorio en las personas que de otro modo podrían intentar entrar en zonas en las que se les prohíbe el acceso. Por ejemplo, en muchos países, las grabaciones de circuito cerrado de televisión (CCTV) se combinan con la tecnología de reconocimiento facial para permitir la comprobación en tiempo real de los espectadores que entran o salen de los principales recintos deportivos sin detener las actividades dentro de las instalaciones (Hutchins and Andrejevic, 2021^[256]). En Estados Unidos, el sistema de identificación de nueva generación, gestionado por la Oficina Federal de Investigaciones (FBI, por sus siglas en inglés), ha implementado sistemas basados en IA para ofrecer funciones biométricas más precisas, como la coincidencia de nombres, huellas dactilares, reconocimiento facial, reconocimiento del iris y coincidencia de impresiones latentes. El sistema genera listas clasificadas de coincidencias potenciales, que luego son revisadas manualmente por analistas biométricos para garantizar la precisión. Este proceso ha

demonstrado su efectividad a la hora de facilitar las investigaciones, proporcionando pistas en aquellos casos en los que otros identificadores están degradados o faltantes (US GAO, 2024^[257]).

También en lo que respecta a la biometría, el uso de tecnologías de reconocimiento facial ha suscitado preocupación por la violación de la privacidad en muchas jurisdicciones (OECD, 2024^[258]) La necesidad de un marco jurídico y normativo sólido que limite las intervenciones basadas en dicho análisis se analiza más adelante en el apartado "Gestión de riesgos y desafíos".

Además de la identificación biométrica, los algoritmos de IA ayudan a los analistas a examinar grandes cantidades de datos para identificar patrones y anomalías que pueden indicar una probabilidad mayor o menor de la participación de una persona concreta, o para descubrir actividades delictivas que, de otro modo, habrían pasado desapercibidas. Esto incluye analizar publicaciones en redes sociales, transacciones financieras e inteligencia de señales en busca de pruebas de conductas ilícitas que ya se hayan producido.

Anticipación a conductas criminales e identificación de factores de riesgo

Mediante el análisis de los perfiles psicológicos, el comportamiento y las acciones pasadas de las personas, los sistemas de IA pueden anticipar posibles conductas delictivas y delitos futuros (Hung and Yen, 2020^[259]). Entre las fuentes de información que permiten realizar estas predicciones se incluyen las redes sociales, las transacciones en línea, los registros públicos, la documentación existente sobre las fuerzas del orden y las grabaciones de CCTV, entre otros. Estas aplicaciones se han utilizado para vigilar a delincuentes conocidos o identificar a personas previamente desconocidas que pueden estar planeando ataques terroristas u otros delitos violentos (McKendrick, 2019^[260]). También se utilizan para delitos no violentos de alta frecuencia y para identificar comportamientos sospechosos que destacan como anomalía estadística (Wastupranata, Kong and Wang, 2024^[261]).

Los sistemas basados en IA también se han utilizado en la gestión de fronteras. Por ejemplo, el Servicio de Aduanas y Protección Fronteriza de los Estados Unidos (CBP, por sus siglas en inglés) ha implementado un Modelo de Evaluación de la Seguridad de los Pasajeros para identificar posibles riesgos en los cruces de pasajeros, que se centra en varios indicadores preocupantes. Este modelo está diseñado para ayudar al personal de la CBP a reconocer rápidamente los cruces que pueden justificar un escrutinio adicional. (US DHS, 2024^[262]). Otro ejemplo es el uso por parte de la CBP de un sistema de ML para identificar patrones potencialmente sospechosos en los historiales de vehículos que cruzan la frontera, que ha dado lugar a 240 incautaciones de miles de kilogramos de drogas ilícitas (US DHS, 2024^[263]).

Las herramientas de IA pueden utilizarse para analizar las redes sociales y las relaciones con el fin de identificar a personas que pueden correr un mayor riesgo de participar en actividades delictivas, como autores o víctimas. Esto puede ayudar a focalizar las intervenciones y los esfuerzos de la policía comunitaria. Las tecnologías de IA están diseñadas para proporcionar asistencia en las investigaciones, por ejemplo, alertando a la policía sobre a quién observar antes de que se cometan delitos.

Hay muchos ejemplos relevantes en todo el mundo. Esto se puede ver en usos a pequeña escala. Por ejemplo, en Japón, un software de detección de robos con IA detecta un lenguaje corporal potencialmente sospechoso en las grabaciones de las cámaras de seguridad y alerta a las autoridades sobre posibles ladrones a través de una aplicación para teléfonos inteligentes (Belova, 2020^[264]). También pueden adoptar la forma de operaciones a mayor escala, como la Herramienta de Evaluación del Riesgo de Daño (HART, por sus siglas en inglés) de la Policía de Durham en el Reino Unido. La vigilancia predictiva ha suscitado controversia, como se analiza más adelante.

Recuadro 5.53. La herramienta de evaluación de riesgos de Alemania para el terrorismo islamista

La Evaluación de Riesgos para el Terrorismo Islamista (RADAR-iTE) es una herramienta algorítmica estandarizada de evaluación de riesgos introducida por la policía alemana en todo el país en 2017. La herramienta está específicamente diseñada para su uso en contextos de protección del Estado (Staatsschutz) para evaluar el riesgo de actos terroristas graves de motivación religiosa y violencia por motivos políticos cometidos por personas ya conocidas por las fuerzas del orden.

RADAR-iTE implica un proceso estandarizado de preparación de casos en el que se evalúan los factores de riesgo y de protección de una persona, asignándolos a una categoría de riesgo de dos niveles (moderado o alto). Esto permite priorizar a las personas y contribuye a una asignación eficiente de los recursos policiales. La evaluación se basa en comportamientos observables e información que ya están a disposición de los agentes de policía o que estos pueden obtener legalmente, en lugar de en atributos como la ideología o la religiosidad. La herramienta utiliza un formulario de evaluación de riesgos con preguntas y categorías de respuesta estandarizadas, pero no está automatizada; la evaluación de las características siempre requiere la evaluación profesional de especialistas capacitados.

RADAR-iTE facilita las decisiones de priorización al identificar la presencia de factores de riesgo y de protección. Las personas consideradas de alto riesgo se debaten a continuación en conferencias sobre casos gestionadas por la Oficina Federal de Policía Criminal (Bundeskriminalamt, BKA) en el seno del Grupo de Trabajo de Gestión de Riesgos del Centro Conjunto de Lucha contra el Terrorismo (GTAZ) (AGRIMA). La herramienta está diseñada para mejorar la estructuración y documentación de historias biográficas de individuos conocidos dentro del espectro militante-salafista, mejorando la transparencia de la evaluación de riesgos y la comparabilidad en toda Alemania. Sirve como herramienta para asistir, pero no para sustituir, a las evaluaciones de casos individuales, complementando las medidas estándar aplicadas a las posibles amenazas y a las personas pertinentes en la protección policial del Estado.

Fuente: https://www.bka.de/DE/UnsereAufgaben/Deliktsbereiche/PMK/Radar/radar_node.html, <https://penal.org/sites/default/files/files/A-02-23.pdf>.

Predicción de horas y lugares en riesgo de actividad criminal

Además de predecir quién puede estar en riesgo de cometer o ser víctima de un delito, los algoritmos de policía predictiva analizan datos históricos de delitos para ayudar a predecir cuándo y dónde puede ser más probable que ocurra un delito. Estos algoritmos tienen en cuenta diversos factores, como la hora del día, las condiciones meteorológicas y los datos socioeconómicos. Algunos departamentos de policía utilizan estos análisis para centrar las patrullas en los puntos críticos identificados, lo que ayuda a los organismos a determinar las rutas de patrulla y desplegar los recursos de manera más estratégica.

Algunos sistemas de policía anticipatoria utilizan datos en tiempo real para ajustar dinámicamente las predicciones y recomendaciones tanto en contextos de aplicación de la ley como de gestión de emergencias. Varios países utilizan la vigilancia CCTV, los informes de incidentes en vivo y los datos de los sensores para perfeccionar las predicciones en tiempo real. En Corea, por ejemplo, la Agencia de Internet y Seguridad (KISA) y la Agencia Nacional de Policía (NPA) colaboraron para desarrollar tecnología destinada a ayudar a las posibles víctimas de acoso a reconocer y evitar el peligro. Este sistema detecta precursores de delitos de acecho, emitiendo advertencias a posibles víctimas y alertando a la policía. Su objetivo es detener a los infractores antes de que se les haga daño. Esta aplicación utiliza

tecnología de IA para analizar información de cámaras de vigilancia en tiempo real (que antes solo estaban disponibles para la detención de los delincuentes después del delito), lo que permite identificar con antelación las señales de delitos de acecho e informar a las víctimas, evitando así cualquier daño (OECD, 2025^[265]).

Mejoramiento de la identificación y anticipación del riesgo de desastres

La IA está transformando la forma en que los países identifican y anticipan los riesgos en diferentes sectores. Los algoritmos de ML se están aplicando a vastos conjuntos de datos para detectar señales de alerta temprana para una amplia gama de riesgos críticos, desde peligros naturales hasta amenazas de salud pública. El Recuadro 5.54 destaca cómo se utiliza la IA en una provincia canadiense para anticiparse a los riesgos de incendios forestales y posicionar los activos de antemano. Grecia también está utilizando IA para la detección temprana de incendios en bosques y arboledas, así como para la prevención de inundaciones⁴⁶.

Recuadro 5.54. Anticipación de riesgos de incendios forestales en Canadá

La provincia de Alberta, en Canadá, utiliza la inteligencia artificial para respaldar sus decisiones de asignación de recursos y posicionamiento previo de activos para combatir los incendios forestales. El innovador sistema predice la probabilidad de que al día siguiente se produzcan incendios forestales en bosques protegidos, que cubren más de la mitad de la provincia.

El sistema de IA, entrenado en una década de datos sobre incendios, meteorología y ecología, incorpora pronósticos meteorológicos diarios y factores como el aumento de la actividad humana durante los fines de semana largos. El sistema tiene una tasa de éxito del 50 % en la predicción de incendios forestales y es preciso el 80 % del tiempo cuando indica que es probable un incendio. Un aspecto crucial es que la herramienta no sustituye a la toma de decisiones en el ámbito humano, sino que presta apoyo al personal.

Desarrollado mediante la colaboración entre los sectores público y privado, el sistema de IA sigue mejorando cada año a medida que se integran nuevos datos sobre incendios forestales.

Fuente: <https://www.oafc.on.ca/about/announcements/fire-news-headlines/how-alberta-uses-artificial-intelligence-wildland>.

En Guatemala, se han utilizado técnicas analíticas de IA aplicadas a la observación de la tierra e imágenes a nivel de calle para identificar estructuras susceptibles de colapso en caso de un terremoto. El enfoque fue capaz de identificar aproximadamente el 85 % de los edificios identificados como vulnerables en base a las evaluaciones tradicionales de ingeniería civil utilizando IA para identificar las áreas dañadas (GFDRR, 2018^[266]; Tilon et al., 2020^[267]).

La observación de la Tierra, los dispositivos conectados y las herramientas de IA también se están utilizando para mejorar la previsión del riesgo de incendios forestales (Nelson, He and Moore, 2024^[268]). Las imágenes de observación de la Tierra proporcionan información regular sobre el crecimiento de la vegetación para determinar las cargas de combustible que impulsan el potencial de incendios forestales. En Australia, se está desarrollando un satélite con tecnologías de infrarrojos calibrado específicamente para monitorear la vegetación local dominante (eucalipto y arbusto) para proporcionar una mayor precisión en el pronóstico de incendios forestales (Amos, 2020^[269]).

Brasil creó el Sistema de Gestión de la Transmisión Geoespacial (GGT), que utiliza IA y procesamiento de imágenes satelitales para inspeccionar las líneas de transmisión y, por lo tanto, evitar incendios forestales (Recuadro 5.55).

Recuadro 5.55. Procesamiento de imágenes satelitales e IA para la protección contra incendios forestales en Brasil

Los incendios forestales son una de las principales causas de apagones en Brasil. Para monitorear las miles de líneas de transmisión brasileñas (TL), la Agencia Brasileña de Regulación de la Electricidad (ANEEL) creó el Sistema de Gestión de Transmisión Geoespacial (GGT), que consiste en una herramienta que utiliza IA y procesamiento de imágenes satelitales, para inspeccionar preventivamente el mantenimiento de los TL contra incendios forestales. El sistema GGT utiliza IA para procesar imágenes satelitales e identificar todos los tramos de las líneas de transmisión afectadas por incendios. Detecta automáticamente las variaciones en las cantidades de vegetación y puede detectar las actividades de supresión de vegetación llevadas a cabo sobre el terreno por las empresas de transmisión sin necesidad de desplazamiento por parte de los equipos de inspección de la ANEEL.

El sistema también genera automáticamente informes de monitoreo que contienen diagnósticos sobre los ajustes de mantenimiento realizados por las compañías de transmisión en los carriles de seguridad de las líneas de transmisión. Esto permite a los equipos de inspección de la ANEEL actuar de manera preventiva y con la suficiente firmeza para mitigar los riesgos de apagones provocados por incendios forestales en todo Brasil. Las empresas de transmisión utilizan la información generada por la GGT y las alertas emitidas por la ANEEL para mejorar sus procesos de mantenimiento de las líneas de transmisión y reducir los riesgos de apagones causados por incendios forestales. Desde su lanzamiento en 2017 hasta 2022, se produjo una disminución del 89 % en los confinamientos causados por incendios forestales en las líneas de transmisión monitoreadas, aunque el número de incendios forestales anuales ha aumentado en todo el país.

Fuente: <https://oecd-opsi.org/innovations/ai-and-geospatial-data-to-protect-against-wildfires>.

En Estados Unidos, por ejemplo, un importante proveedor de servicios energéticos de los estados de Carolina del Norte y Carolina del Sur utilizó técnicas de IA antes del huracán Florence en 2018 para determinar dónde ubicar previamente a los técnicos necesarios para restablecer rápidamente la energía si las líneas de transmisión resultaban dañadas (Seto, 2018^[270]).

La IA puede procesar grandes cantidades de datos de diversas fuentes, como historiales médicos, redes sociales y dispositivos de vigilancia ambiental, para detectar patrones y anomalías que puedan indicar la aparición de una enfermedad, lo que permite a las autoridades tomar medidas proactivas. (El Morr et al., 2024^[271]). La IA también se está implementando en aplicaciones que abordan riesgos inducidos por el hombre, no solo los derivados de peligros naturales. El Recuadro 5.56 muestra cómo puede utilizarse la IA para avanzar en la anticipación de riesgos y la aplicación de la ley.

Recuadro 5.56. CCTV coreano basado en IA para proteger contra delincuencia y emergencias

En Corea, a medida que las zonas urbanas enfrentan crecientes demandas de vigilancia, las ciudades están desplegando sistemas de vigilancia selectiva basados en IA para mejorar la seguridad, la respuesta a emergencias y la prevención del delito. Estos sistemas utilizan la infraestructura CCTV existente y algoritmos avanzados de aprendizaje profundo para detectar y clasificar automáticamente objetos, personas y eventos, incluidos accidentes, incendios y comportamientos peatonales inusuales (por ejemplo, merodear, luchar contra el acceso no autorizado).

Al dar prioridad al material de archivo relevante, el monitoreo CCTV basado en IA reduce la dependencia de la vigilancia manual, lo que permite la detección de incidentes en tiempo real y una

respuesta rápida. Entre las funciones principales se incluyen el análisis de trayectorias de varias palabras clave para el seguimiento de objetos, la búsqueda de vídeo de alta velocidad y el reconocimiento inteligente de eventos para apoyar a las fuerzas del orden y la gestión de desastres. Para garantizar la privacidad, el sistema anonimiza la información confidencial, como rostros y matrículas, y restringe la exportación de vídeo mediante cifrado.

Fuente: <https://smartcity.go.kr/wp-content/uploads/2024/09/EN02.AI-Based-Selective-Monitoring-System.pdf>.

Mejoramiento de la evaluación de riesgos y modelización predictiva

Los sistemas avanzados de IA permiten realizar evaluaciones de riesgos más precisas al integrar conjuntos de datos variados y complejos. El modelado predictivo impulsado por IA puede simular varios escenarios de riesgo, lo que permite a los responsables de la toma de decisiones evaluar los impactos potenciales y planificar estrategias de mitigación. Estos sistemas utilizan datos de múltiples fuentes, como registros climáticos, indicadores socioeconómicos y redes de infraestructura, para generar perfiles de riesgo más completos. La capacidad de la IA para procesar grandes volúmenes de datos mejora la precisión de los modelos de riesgo, en particular en entornos muy dinámicos como las catástrofes. Al identificar las correlaciones y las relaciones causales entre las diferentes variables, los sistemas de IA pueden ofrecer información más detallada sobre la probabilidad y la gravedad de los riesgos futuros. Esta mejor comprensión respalda la toma de decisiones basada en evidencia y la asignación de recursos. La IA también puede ayudar a los responsables de la modelización a ejecutar modelos más rápidos, con menos uso de recursos y más precisos. Por ejemplo, el Centro Europeo de Previsiones Meteorológicas a Medio Plazo (ECMWF) utiliza el Sistema de Predicción de Inteligencia Artificial (AIFS, por sus siglas en inglés) para mejorar la predicción meteorológica numérica. El AIFS ofrece hasta un 20 % más de precisión que los principales modelos basados en la física en varias áreas, incluido el seguimiento de ciclones tropicales (ECMWF, 2025^[272]). Los sistemas de IA pueden perfeccionarse continuamente a medida que se dispone de nuevos datos, lo que garantiza que las evaluaciones de riesgos se mantengan actualizadas. Esta adaptabilidad es especialmente valiosa para hacer frente a los riesgos emergentes, como los asociados a los ciberataques.

Recuadro 5.57. El gemelo virtual de Singapur basado en datos

Singapur desarrolló Virtual Singapore, el primer gemelo digital de un país, para apoyar la planificación urbana basada en datos y mejorar la gestión de desastres. Liderada por la Autoridad de Tierras de Singapur (SLA, por sus siglas en inglés), esta iniciativa utiliza modelado 3D avanzado e integración de datos en tiempo real para crear una réplica digital muy detallada de la ciudad-estado. La tecnología de escaneo láser, desplegada a través de aeronaves y vehículos, captura información precisa del terreno y la superficie, mapeando estructuras sobre el terreno, servicios subterráneos, condiciones ambientales y movimientos de población. Como parte clave de la estrategia de nación inteligente de Singapur, Singapur Virtual proporciona a los responsables de la formulación de políticas y a los investigadores una plataforma interactiva para probar escenarios, anticipar riesgos, optimizar servicios y mejorar la resiliencia ante futuras emergencias.

La plataforma permite a los usuarios simular el impacto de nuevos proyectos, evaluar los flujos de transporte y modelar los factores ambientales (como las islas de calor urbanas) y los riesgos (como las inundaciones). Al facilitar la colaboración intersectorial, Virtual Singapore apoya una mejor asignación de recursos, mejora la toma de decisiones y fomenta la innovación tanto en el sector público como en el privado. Los organismos gubernamentales pueden utilizar la plataforma para mejorar la resiliencia

de las infraestructuras, mientras que las empresas y los investigadores utilizan sus datos para desarrollar soluciones de ciudades inteligentes. Además, el modelo desempeña un papel crucial en la gestión de emergencias y desastres, ayudando a las autoridades a planificar las respuestas a los fenómenos meteorológicos extremos y optimizar la coordinación de crisis. Con actualizaciones en curso que incorporan nuevos enfoques de IA, Internet de las cosas y nuevas fuentes de datos, Singapur Virtual sigue evolucionando como una herramienta dinámica para dar forma al futuro de la vida urbana.

Fuente: <https://oecd-opsi.org/innovations/virtual-twin-singapore>.

Mejoramiento de la comunicación de riesgos

Las herramientas de IA están mejorando la forma en que los responsables de la toma de decisiones y el público entienden la información sobre riesgos. Al traducir datos complejos a formatos visuales y un lenguaje accesible, la IA hace que las evaluaciones de riesgos sean más comprensibles. Los paneles interactivos, chatbots y asistentes virtuales proporcionan información adaptada a los distintos públicos, fomentando una mejor concienciación y preparación frente a los riesgos (UNU Institute for Environment and Human Security, 2024^[273]).

La IA también puede personalizar la comunicación de riesgos y enviar mensajes específicos a las personas en función de su ubicación, vulnerabilidad y preferencias. Este enfoque ayuda a garantizar que la información llegue a quienes más la necesitan, especialmente durante emergencias (Zhao et al., 2025^[274]). Además, las herramientas de análisis de sentimientos impulsadas por la IA pueden medir la percepción pública de los riesgos y permitir a las autoridades ajustar sus estrategias de comunicación en consecuencia. Las plataformas de comunicación de riesgos en tiempo real mejoran la gestión de crisis al proporcionar actualizaciones sobre la evolución de las situaciones y orientación sobre medidas de protección. Este flujo dinámico de información refuerza la confianza pública y fomenta un comportamiento más proactivo de mitigación de riesgos (Recuadro 5.58).

Recuadro 5.58. Asistente de planificación de mitigación de riesgos de Estados Unidos

En Estados Unidos, la Agencia Federal para el Manejo de Emergencias (FEMA, por sus siglas en inglés) está explorando actualmente el uso de una solución generativa de IA, llamada el Asistente de Planificación para Comunidades Resilientes (PARC, por sus siglas en inglés). Este nuevo sistema genera eficiencias para el proceso de planificación de mitigación de riesgos para los gobiernos locales, incluidas las comunidades desatendidas. Los planes de mitigación de riesgos son un paso fundamental que las comunidades pueden tomar para desarrollar su resiliencia; sin embargo, pueden ser largos de producir y desafiantes para las comunidades que carecen de recursos para hacerlo. PARC apoyará específicamente la comprensión de los gobiernos estatales, locales, tribales y territoriales (SLTT, por sus siglas en inglés) sobre cómo diseñar un plan que identifique los riesgos y las estrategias de mitigación, así como generar elementos de borrador del plan, de fuentes públicas e investigadas, que los gobiernos podrían personalizar para satisfacer sus necesidades. Esto podría dar lugar a que más comunidades puedan presentar solicitudes de subvención para obtener financiamiento, lo que las haría más resilientes y reduciría los riesgos de desastre.

Fuente: <https://www.dhs.gov/ai/use-case-inventory/fema>.

Mejoramiento de la respuesta ante catástrofes y crisis

La IA está mejorando la conciencia situacional durante los desastres mediante el procesamiento de imágenes satelitales, datos de sensores y actualizaciones de redes sociales en tiempo real. Esta información ayuda a los servicios de emergencia a mapear las áreas afectadas, priorizar la asignación de recursos y coordinar los esfuerzos de respuesta. Los sistemas de apoyo a la toma de decisiones impulsados por IA pueden recomendar rutas de evacuación óptimas, identificar infraestructuras críticas en riesgo y predecir la propagación de peligros. La IA también puede monitorear los cambios en un entorno operativo de respuesta ante desastres. Mediante el análisis de fuentes de datos en directo, las herramientas de IA pueden proporcionar actualizaciones sobre cierres de carreteras, apagones e interrupciones del transporte. Estos conocimientos ayudan a las autoridades y las comunidades a adaptarse a las condiciones cambiantes y a mantener la continuidad de los servicios esenciales.

La IA puede ayudar en la recopilación de datos complejos directamente de las poblaciones afectadas. En Indonesia, los residentes de Yakarta han usado los medios sociales para compartir actualizaciones sobre las inundaciones, promovidas por Bencana Bot, un chatbot de la Fundación Indonesia de Mapeo de Desastres (PetaBencana.id). El chatbot invita a los usuarios a informar sobre los detalles de la inundación, que luego se mapean en tiempo real y se puede acceder en línea de forma gratuita. La Agencia de Mitigación de Desastres de Yakarta también actualiza el mapa con datos oficiales. Inmediatamente después de las inundaciones a gran escala en la ciudad, la plataforma ha visto un aumento del 2 000 % en la actividad dentro de las 12 horas, ayudando a los residentes a evitar áreas peligrosas y tomar decisiones de seguridad informadas (PetaBencana.id, 2021^[275]).

Las agencias de respuesta también están utilizando algoritmos de detección de cambios. Estos pueden aplicarse a las imágenes de observación de la Tierra para identificar rápidamente las regiones geográficas afectadas, así como los daños a edificios e infraestructuras individuales.

Mejoramiento de la evaluación de necesidades después de los desastres y la recuperación/reconstrucción

La IA está mejorando la evaluación de las necesidades después de un desastre al integrar datos de imágenes satelitales, redes sociales y registros administrativos para producir evaluaciones de daños y necesidades más rápidas y exhaustivas. Estos sistemas pueden identificar automáticamente la infraestructura dañada, evaluar el número de personas desplazadas y estimar los costos de recuperación. Esta automatización acelera el proceso de evaluación, permitiendo que los esfuerzos de recuperación comiencen antes.

Los sistemas de apoyo a la toma de decisiones impulsados por IA ayudan a las autoridades a priorizar proyectos de recuperación identificando las necesidades más urgentes y optimizando la asignación de recursos. Estos sistemas pueden modelar diferentes escenarios de recuperación, lo que permite a los responsables de la toma de decisiones evaluar las compensaciones entre la reconstrucción de la infraestructura, la restauración de los medios de vida y la mejora de la resiliencia.

La IA apoya una reconstrucción transparente y responsable mediante el seguimiento del progreso de los proyectos de recuperación y la identificación de retrasos o irregularidades. Esta supervisión ayuda a garantizar que los recursos se utilicen de manera eficiente y que las comunidades vulnerables reciban asistencia oportuna. La IA también puede ayudar a optimizar el despliegue de recursos, por ejemplo, asignando funcionarios a áreas o tareas específicas en función de las necesidades previstas. Estas tecnologías también pueden utilizarse con fines de seguridad pública que escapan a los contextos predictivos, por ejemplo, para mejorar los tiempos de respuesta y el uso eficiente general de recursos limitados. Un ejemplo destacado de la Agencia Federal para la Gestión de Emergencias (FEMA) de EE.UU. puede encontrarse en el Recuadro 5.59.

Recuadro 5.59. Aceleramiento con IA de la recuperación post huracanes en Estados Unidos

Tras el paso del huracán Ian, que azotó el estado de Florida en 2022, la Agencia Federal de Estados Unidos para el Manejo de Emergencias (FEMA, por sus siglas en inglés) implementó una herramienta de IA conocida como modelo de Evaluación de Daños Geoespaciales (GDA, por sus siglas en inglés). Tradicionalmente, las evaluaciones de los daños posteriores a los desastres se basaban en inspecciones manuales, un proceso que podía tardar semanas y a menudo retrasaba la entrega de ayuda financiera a las víctimas. El modelo GDA transformó este enfoque al automatizar la clasificación de daños.

Esta innovación utilizó imágenes aéreas, datos satelitales y ML para evaluar rápidamente el daño estructural en las áreas afectadas, mejorando significativamente la velocidad y precisión de los esfuerzos de respuesta ante desastres. Identificó y clasificó los niveles de daños estructurales en más de un millón de estructuras afectadas. Al integrar datos adicionales, como las velocidades del viento, las características de los edificios y la proximidad a la trayectoria del huracán, el modelo proporcionó una comprensión más completa de la gravedad de los daños. Redujo el número de estructuras que requerían revisión humana de más de 1 millón a solo 77 000, reduciendo los tiempos de evaluación de semanas a días. A las 72 horas de la llegada del huracán Ian, FEMA tuvo conocimiento de la magnitud de los daños en las regiones afectadas, lo que permitió una asignación más rápida de recursos y una planificación de la recuperación.

Fuente: <https://gis-fema.hub.arcgis.com/pages/geospatial-da-training>.

Gestión de riesgos y desafíos

Las herramientas de IA pueden mejorar las capacidades gubernamentales relacionadas con la aplicación de la ley, el riesgo de desastres y la gestión de crisis presenta beneficios claros; sin embargo, su despliegue debe tener en cuenta varios riesgos y desafíos. Un análisis de 61 iniciativas de automatización de la función pública canceladas en varios países reveló que la mitad de los sistemas cancelados estaban relacionados con la policía y la aplicación de la ley (en parte debido a la interrupción del reconocimiento facial), a menudo como resultado de revisiones críticas por parte de organismos gubernamentales, medios de comunicación o la sociedad civil (Redden et al.[228]). A continuación, se describen los principales riesgos y desafíos identificados, que los gobiernos deben tener en cuenta en su desarrollo e implementación de las herramientas de IA.

Riesgos asociados

- Aunque el uso de la IA para determinados aumentos de productividad, como la realización automatizada de tareas rutinarias y el papeleo, suele ser de bajo riesgo y no controvertido, el uso de la IA en la vigilancia predictiva plantea varias consideraciones importantes de política pública. Deben tenerse en cuenta y abordarse para garantizar que dichos usos proporcionen un análisis creíble e incluyan salvaguardias para garantizar un uso responsable de sus capacidades.

Si los sistemas de IA se basan en **datos inadecuados o sesgados**, podrían producirse resultados inexactos o adversos para algunas personas o grupos. En lo que respecta la aplicación de la ley y la gestión del riesgo de desastres, esto podría, por ejemplo, dar lugar a una discriminación ilegal (en la medida en que un país tenga tales leyes). Los responsables de la formulación de políticas deben tratar de garantizar que estos sistemas se prueben y validen rigurosamente antes de su implementación.

Deberían exigir que se registren y pongan a disposición del escrutinio los índices de falsos positivos y verdaderos positivos. En los últimos años han proliferado las investigaciones que utilizan diversos métodos para la validación de sistemas (Stettinger, Weissensteiner and Khastgir, 2024^[276]).

Los expertos en IA han destacado que la vigilancia invasiva y las infracciones de la privacidad que socavan la capacidad de las personas para ejercer libremente sus derechos humanos y libertades son uno de los riesgos más acuciantes de la IA (OECD, 2024^[29]). La amplia recopilación de datos necesaria para la actuación policial predictiva amenaza con vulnerar involuntariamente los derechos a la privacidad personal y a la protección de los datos. El uso de la IA con fines de vigilancia puede dar lugar a una **vigilancia intrusiva de las personas, lo que puede suponer una violación de los derechos de privacidad**. Los responsables de las políticas deben tratar de promover la seguridad pública y, al mismo tiempo, garantizar la protección de las personas a través de regulaciones sobre privacidad y protección de datos para evitar su uso indebido y extralimitación.

Al agrupar los riesgos de sesgo y las infracciones de la privacidad, la identificación biométrica remota (es decir, el reconocimiento facial) ha resultado particularmente controvertida porque los sistemas involucrados pueden conllevar un sesgo tecnológico inherente (por ejemplo, la discriminación por origen racial o étnico) (OECD, 2020^[277]). El uso de la IA en espacios públicos también puede socavar la capacidad de las personas de ejercer libremente los derechos aplicables relacionados con la privacidad, la protección de datos y la libertad de expresión y reunión, según algunas organizaciones de la sociedad civil (Access Now, 2022^[278]; CSOs, 2021^[279]), organismos de aplicación de la UE (EDPB-EDPS, 2021^[280]), y algunos Relatores Especiales de la ONU (Khan et al., 2022^[281]). En el Recuadro 5.60, se examinan varias cuestiones relacionadas con el uso del reconocimiento facial en este contexto.

Recuadro 5.60. Reconocimiento facial y preocupaciones sobre privacidad y prejuicios

A medida que la tecnología de reconocimiento facial ha madurado, se ha vuelto cada vez más capaz de identificar rostros en una multitud, emparejar imágenes de fuentes que van desde grabaciones de CCTV hasta bases de datos policiales para proporcionar vigilancia en tiempo real, e identificar a sospechosos criminales o personas desaparecidas, entre otras aplicaciones. La preocupación por la privacidad es que la tecnología de vigilancia pública permite a los gobiernos recopilar grandes cantidades de información sobre la ciudadanía con un solo propósito (sin su consentimiento) que podría usarse para diferentes propósitos, no autorizados e ilegales.

Las herramientas de IA para mejorar la tecnología de reconocimiento facial, cuando se forman en conjuntos de datos inadecuados o sesgados, pueden reducir la precisión de la identificación en algunos grupos.

También han surgido casos de procedimientos inapropiados por parte de las fuerzas policiales que han llevado al uso de datos de entrada de mala calidad, lo que ha debilitado sustancialmente la precisión del software de reconocimiento facial. Por ejemplo, las fuerzas policiales de Estados Unidos han intentado emparejar dibujos de sospechosos, imágenes de CCTV de baja calidad, imágenes mejoradas por computadora e incluso una foto del famoso colaborador de un sospechoso con bases de datos de imágenes. Estos ejemplos ponen de relieve la importancia de disponer de normas claras sobre cómo utilizar exactamente el software y aclarar si una coincidencia es motivo suficiente para una detención.

Aunque se ha prestado mucha atención a los casos en los que la tecnología funciona mal (por ejemplo, falsos positivos) o se utiliza indebidamente, algunos defensores de la privacidad subrayan que el uso general de la tecnología de vigilancia o monitoreo subyacente socava la capacidad de ejercer libremente los derechos al disminuir la expectativa de anonimato en los espacios públicos, que es una condición para el ejercicio de derechos como la libertad de reunión. Por ejemplo, las autoridades de

protección de datos de la UE han emitido dictámenes en los que declaran que el uso de estos sistemas en espacios públicos es incompatible con el derecho a la protección de datos, en virtud de la legislación de la UE, ya que el escaneo de todas las personas que pasan por estos espacios interfiere con sus derechos en virtud de la legislación de la UE. En consecuencia, en la UE, la Ley de IA de la UE solo permite el uso de sistemas de identificación biométrica a distancia "en tiempo real" en espacios de acceso público a efectos de aplicación de la ley si dicho uso está autorizado en virtud de la legislación nacional de un Estado miembro y sujeto a las salvaguardas adecuadas. En particular, cada uso debe recibir una autorización previa de una autoridad judicial o una autoridad administrativa independiente, y debe notificarse a la autoridad nacional de protección de datos pertinente. Otras economías logran un equilibrio diferente a la hora de garantizar los beneficios de las tecnologías y, al mismo tiempo, hacer frente a los riesgos relacionados con ellas.

Fuente: (OECD, 2020^[254]), <https://towardsdatascience.com/how-ethical-is-facial-recognition-technology-8104db2cb81b>, www.bbc.co.uk/news/technology-47117299, <https://medium.com/@AINowInstitute/after-year-of-tech-scandals-our-10-recommendations-for-ai-95b3b2c5e5>, www.flawedfacedata.com, www.americaunderwatch.com, www.bbc.co.uk/news/technology-48222017, <https://edri.org/our-work/remote-biometric-identification-a-technical-legal-guide>, <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>.

Muchos sistemas de IA predictiva son privados o se basan en modelos de "caja negra", lo que dificulta que el público y los responsables de la formulación de políticas comprendan plenamente por qué se toman algunas decisiones operativas en este campo. **Esta falta de transparencia puede socavar la confianza en las fuerzas del orden.** Las herramientas de IA extrapolan los resultados de las correlaciones estadísticas entre las diferentes fuentes de datos que agregan. Al diseñar y utilizar estos sistemas de apoyo a la toma de decisiones, los desarrolladores e investigadores deben reflexionar sobre la relevancia de utilizar variables sociodemográficas específicas en los modelos predictivos. En el caso de una popular herramienta de IA utilizada por la Gendarmería francesa, el modelo predictivo utiliza 15 variables sociodemográficas que, según los desarrolladores, están fuertemente correlacionadas con la delincuencia. Sin embargo, no existe transparencia sobre la naturaleza de estas variables ni se intenta demostrar una relación causa-efecto genuina (Lecorps and Tissandier, 2022^[282]).

Por último, existe el riesgo de que los organismos encargados de hacer cumplir la ley **dependan excesivamente de la IA**, descuidando otros aspectos críticos de la actuación policial, como la participación de la comunidad y los métodos de investigación tradicionales. Deberían realizarse estudios evaluativos de las herramientas alternativas de IA utilizadas en la actuación policial predictiva para identificar los puntos fuertes comparativos en cuanto a la delincuencia que se produce en las zonas previstas y el número de intervenciones policiales efectivas (Birks, Townsley and Hart, 2023^[283]). En relación con esto, el uso de IA en la vigilancia policial y la gestión del riesgo de desastres plantea **preguntas éticas sobre la delegación de la toma de decisiones morales a las máquinas**. Los responsables de las políticas deberían considerar las implicaciones morales del uso de la IA para tomar decisiones que podrían alterar la vida en el contexto de la aplicación de la ley y la gestión del riesgo de desastres. Es crucial garantizar que se mantenga la supervisión humana y las determinaciones finales, y que la IA se utilice para aumentar el juicio humano en lugar de sustituirlo.

Desafíos de implementación

- Elevados costos de adopción y escalabilidad de la IA y de la demostración de resultados y rendimiento de la inversión.
- Brechas de habilidades.
- Entornos legales y regulatorios inflexibles u obsoletos.

La implementación de tecnologías de IA en la actuación policial requiere una inversión significativa. La mayoría de los departamentos de policía son organismos locales que cubren una pequeña zona

geográfica con poblaciones pequeñas. Las funcionalidades incorporadas en el software estándar del lugar de trabajo y los sistemas de información geográfica pueden admitir muchos métodos predictivos. Sin embargo, las herramientas, paquetes de software o ordenadores de IA más caros pueden superar la asequibilidad (y tal vez las necesidades) de los programas locales de vigilancia predictiva. Las autoridades deben analizar los **costos relativos de los beneficios y procurar que los recursos se asignen de manera efectiva**. El personal encargado de hacer cumplir la ley necesita una formación adecuada para utilizar y entender eficazmente las herramientas de IA. Los responsables de la formulación de políticas deben tratar de garantizar que se implementen programas de capacitación adecuados (EUCPN, 2022^[284]) lo que a menudo no es el caso.

Es posible que las leyes actuales no aborden adecuadamente los desafíos únicos que plantea la IA en la actuación policial. Los responsables de las políticas deben desarrollar y actualizar los marcos jurídicos para regular el uso de la IA en la aplicación de la ley. Es esencial establecer normas y mejores prácticas para el desarrollo y el despliegue de tecnologías de policía predictiva a fin de garantizar la coherencia y la calidad.

Potencial sin explotar y camino a seguir

Las herramientas de IA siguen ofreciendo cada vez más oportunidades para mejorar las capacidades de los cuerpos policiales y los organismos de gestión de emergencias al mejorar la seguridad pública y la capacidad de respuesta. Se espera que el uso de estas herramientas en una gama más amplia de organismos públicos con un papel en la gobernanza de riesgos críticos continúe aumentando a medida que avancen las tecnologías subyacentes y se vuelva más frecuente la capacitación del personal. Proliferan anécdotas sobre los beneficios de las herramientas de IA en contextos relevantes, al igual que narrativas sobre usos irresponsables. **Es necesario continuar con la investigación empírica para fortalecer la base de evidencias tanto sobre los beneficios medibles como sobre los inconvenientes del uso de la IA en estos contextos**. Esto ayudará a calibrar su adopción allí donde se pueda demostrar una mayor efectividad y eficiencia. O, por el contrario, ayudará a respaldar una supervisión más estricta debido a los impactos negativos que los resultados erróneos pueden tener en las vidas humanas.

Las implicaciones éticas, jurídicas y sociales de la adopción de la IA exigen una cuidadosa consideración y medidas proactivas para garantizar que estas tecnologías se despliegan de una manera confiable y respetuosa con los derechos y libertades individuales. **Los responsables de las políticas deben colaborar con las comunidades para generar confianza y garantizar que se tengan en cuenta sus preocupaciones sobre el despliegue de la IA**. Comunicar claramente cómo la IA en la actuación policial y en la gestión del riesgo de desastres, incluidos sus objetivos, beneficios y limitaciones, es vital para mantener la confianza pública (Evans, 2022^[285]). Abordar los desafíos descritos anteriormente requiere una colaboración integral y de múltiples partes interesadas entre las entidades usuarias finales, los responsables de la formulación de políticas, los tecnólogos y las comunidades afectadas. Las investigaciones basadas en este tipo de consultas pueden ampliar la evidencia de los beneficios medibles del uso de la IA, así como profundizar en la comprensión de las consecuencias no deseadas.

Los gobiernos y los organismos encargados de hacer cumplir la ley deberían reforzar su comprensión de las limitaciones de la IA en la actuación policial predictiva para reducir la delincuencia. Los modelos estadísticos avanzados y el poder computacional no son ni necesarios ni suficientes para reducir la delincuencia. Los algoritmos pueden ayudar a predecir el riesgo de eventos futuros, pero no los eventos en sí. Aunque la IA puede simplificar la búsqueda de patrones, las herramientas extrapolan datos de acontecimientos pasados y solo son tan precisas como los datos subyacentes utilizados para elaborarlos. Los seres humanos son los elementos más importantes en el proceso de vigilancia predictiva; necesitaban encontrar y reunir datos pertinentes, prepararlos para el

análisis, revisar e interpretar los resultados de estos análisis para eliminar las conclusiones incorrectas y ofrecer recomendaciones sobre cómo proceder basadas en los hallazgos analíticos (Perry et al., 2013^[286]).

El uso de la IA en la aplicación de la ley es una de las áreas de aplicación de la tecnología que presenta más desafíos. Proporciona grandes beneficios potenciales para la seguridad nacional y la seguridad pública, pero también posibles impactos negativos. Por ejemplo, los algoritmos sesgados pueden dar lugar a falsos positivos que den lugar a detenciones y encarcelamiento erróneos de personas inocentes, o a falsos negativos que den lugar a una omisión de la actividad ilícita (Petra et al., 2024^[287]). Estas acciones refuerzan **la necesidad de desarrollar mecanismos consultivos destinados a obtener el apoyo público para el despliegue de herramientas de IA dentro de las condiciones y modalidades de las barreras de protección aceptadas**. Estas acciones introducen una incertidumbre significativa en cuanto al camino a seguir por la IA en este ámbito.

IA en la administración y acceso a la justicia

Las tecnologías digitales y los datos tienen un gran potencial para apoyar el acceso a la justicia, así como para fortalecer su resiliencia, eficiencia, efectividad y equidad, en consonancia con la *Recomendación de la OCDE sobre el acceso a la justicia y sistemas de justicia centrados en las personas* (2023^[288]). De entre todas las tecnologías que impulsan esta transformación, la IA destaca por ser tanto notable como novedosa.

Hasta ahora, la IA está impactando de manera desigual en los sistemas de justicia. Si bien algunos países han adoptado la IA para aplicaciones prácticas dentro de sus sectores judiciales, otros todavía se encuentran en las primeras etapas de integración de avances tecnológicos más básicos. Las aplicaciones de IA en este ámbito van desde la automatización de tareas administrativas rutinarias relacionadas con la gestión de casos, hasta la oferta de análisis predictivo y el apoyo a la investigación jurídica. A menudo enfrentados a retrasos, largos requisitos procesales y una accesibilidad limitada, los sistemas de justicia tienen mucho que ganar con la integración de la IA. La IA posee un gran potencial para mejorar la eficiencia y la capacidad de respuesta, mejorar la implementación de servicios adaptados a las necesidades de las personas, apoyar las tareas rutinarias y ampliar capacidades en diversos ámbitos de la justicia (Redden and O'Donovan Dix, 2020^[289]; CEPEJ, 2018^[290]).

Un desarrollo o uso inadecuado de la IA en el sector judicial puede dar lugar a un trato desigual entre personas o grupos. Los marcos de gobernanza y salvaguardas para evaluar los riesgos e incidentes, así como las soluciones para los afectados por los sistemas de IA, pueden ayudar a garantizar un uso confiable de la IA en la justicia.

Situación actual

Mayor eficiencia de las operaciones internas

El impacto de la IA en la productividad y la capacidad de respuesta del sistema judicial puede ser transformador. Los algoritmos avanzados de IA pueden analizar los casos entrantes, clasificarlos en función de su complejidad o urgencia y asignarlos a los departamentos o jueces pertinentes (Reiling, 2020^[291]). Esto puede reducir significativamente los cuellos de botella administrativos, ayudar a garantizar un uso más eficiente de los recursos y mejorar la efectividad en la administración de justicia. Por ejemplo:

- El Tribunal de Casación francés (2024^[292]) desarrolló un sistema de IA que clasifica y redirige las apelaciones y peticiones entrantes a las salas adecuadas basándose los datos de casos anteriores. Esto hace que la gestión de casos sea más eficiente y garantiza que los casos tengan el nivel de escrutinio adecuado.
- En Brasil, el sistema VICTOR AI automatiza el examen de los recursos ante la Corte Suprema de Justicia, identificando casos con "repercusiones generales", un requisito para la tramitación de recursos (Supreme Court of Brazil, 2024^[293]). Mientras que un secretario judicial tarda 44 minutos en evaluar si un recurso cumple las condiciones para avanzar, VICTOR AI lo hace en cuestión de segundos. Además, el Consejo Superior de Justicia Laboral de Brasil lanzó Chat-JT, una herramienta de IA generativa que ayuda a los jueces, al personal judicial y a los becarios a automatizar tareas (por ejemplo, investigación jurídica, análisis de documentos o redacción de

resúmenes estandarizados). Lanzada en febrero de 2025, la plataforma también permite a los usuarios crear asistentes personalizados adaptados a las necesidades específicas del flujo de trabajo judicial. Al integrarse con otros sistemas internos, Chat-JT proporciona capacidades avanzadas de búsqueda en bases de datos internas, con el objetivo de agilizar las operaciones diarias y mejorar los procesos de toma de decisiones en el ámbito de la Justicia Laboral (CNJ, 2025^[294]).

- La transcripción automatizada puede ayudar a mejorar el acceso a los procedimientos judiciales. En Eslovenia, por ejemplo, está facilitando el acceso a los procedimientos judiciales (Slovenia Ministry of Justice, 2024^[295]). Esta tecnología convierte rápidamente las palabras habladas en texto escrito, lo que garantiza una mayor transparencia y permite que un mayor número de personas revisen y comprendan los procedimientos judiciales (CMS, 2023^[296]). En España, los servicios de transcripción, denominados "textualización", llevan funcionando desde 2019, integrados en el visor de archivos judiciales electrónicos HORUS. Esta combinación de sistemas no sólo proporciona transcripciones, sino que también permite la navegación rápida de grabaciones de las audiencias. Esto se realiza al: 1) asociar cada segmento de la textualización al marco de tiempo correspondiente de la grabación y 2) identificar las diferentes voces que aparecen en la grabación, permitiendo a los usuarios rastrear qué partes intervienen en un momento dado. Esto resulta especialmente importante en España, donde la propia grabación constituye el registro oficial de las audiencias.
- El sistema PretorIA de Colombia asiste a la Corte Constitucional en la gestión del elevado volumen de casos de tutela o acciones legales presentadas para proteger los derechos humanos. Con más de 600 000 tutelas recibidas anualmente, PretorIA ayuda a preseleccionar casos para revisión mediante la detección de criterios legales predefinidos. Garantizar la supervisión humana y la explicabilidad ha sido una prioridad clave, y PretorIA se fue rediseñado en 2020 para utilizar modelados temáticos interpretables en lugar de redes neuronales opacas. Los jueces siguen siendo los máximos responsables de la toma de decisiones, manteniendo la autoridad humana en la selección de casos (OECD/CAF, 2022^[297]).
- En Grecia, la implementación de la IA facilita los procesos judiciales a través del procesamiento automatizado de documentos alineado con los marcos jurídicos nacionales y la estrategia europea de IA. Las tecnologías de NLP analizan documentos jurídicos, lo que permite realizar búsquedas de jurisprudencia y gestionar los registros judiciales de forma eficiente⁴⁷. Otras iniciativas actuales integran la IA para la conversión de voz a texto, la traducción de documentos y la interpretación de declaraciones de testigos. Aplicaciones adicionales incluyen el apoyo a los servicios forenses y la optimización de la administración judicial, todo ello bajo supervisión humana para garantizar la protección de los derechos humanos y, al mismo tiempo, mejorar la eficiencia⁴⁸.

Los sistemas generativos de IA también están ayudando a mejorar la eficiencia de los sistemas judiciales. La transcripción automatizada —una de las formas más básicas de IA generativa— puede mejorar el acceso a los procedimientos judiciales, como se está haciendo en Eslovenia y España. Esta tecnología convierte rápidamente las palabras habladas en texto escrito, lo que garantiza una mayor transparencia y permite que un mayor número de personas revisen y comprendan los procedimientos judiciales.

La NLP y la IA generativa también han contribuido a mejorar la gestión de casos y el procesamiento de documentos judiciales. En el caso de España, al automatizar las tareas repetitivas y mejorar la extracción de metadatos, los servicios basados en IA han agilizado los flujos de trabajo diarios del personal del sistema judicial, lo que permite una asignación más efectiva de los recursos humanos y posibilita que el personal se centre en actividades de mayor valor, como los análisis jurídicos complejos (Recuadro 5.61)⁴⁹.

Recuadro 5.61. Mejoramiento de la gestión documental en España

El Ministerio de la Presidencia, Justicia y Relaciones con las Cortes de España ha desarrollado internamente un conjunto de herramientas de NLP y de IA generativa para promover una gestión y un tratamiento eficientes de los documentos judiciales. Estas herramientas ayudan a los profesionales del derecho a clasificar, analizar, resumir y anonimizar los textos relacionados con los tribunales, garantizando al mismo tiempo el cumplimiento de las normas de protección de datos.

La funcionalidad de clasificación permite ordenar los documentos por jurisdicción, tipo de registro y categoría de documento (por ejemplo, presentaciones legales o notificaciones), mejorando así la recopilación y sistematización de la información en diversas ramas del sistema judicial. El paquete también incluye funciones analíticas avanzadas que extraen y visualizan entidades con nombre, fechas clave, valores monetarios y relaciones entre entidades. Además, el sistema realiza detecta similitudes semánticas, ayudando a armonizar inconsistencias en nombres entre documentos.

Las capacidades de resumen se adaptan a diversos tipos de documentos jurídicos, como sentencias y órdenes, y ofrecen cuatro modos distintos, que van desde extractivos hasta generativos, incluida una opción de "lenguaje jurídico a lenguaje sencillo" diseñada para mejorar la accesibilidad.

Estas herramientas están disponibles a través de tres canales:

1. Un portal seguro en línea para uso del personal de justicia.
2. La Carpeta de Justicia, la ventanilla única para la ciudadanía, donde también pueden acceder a las herramientas de anonimización y resumen.
3. Integración con otros productos de justicia digital.

Un ejemplo destacado de esta integración es LexNET, la plataforma española de intercambio electrónico de información que conecta a los órganos judiciales con profesionales del Derecho, cuerpos y fuerzas de seguridad, hospitales y otras partes interesadas. Dentro de LexNET, NLP escanea documentos para identificar y extraer las entidades nombradas en ellos, para luego autocompletar campos de formulario obligatorios que antes tenían que rellenarse manualmente. Esta automatización reduce significativamente el tiempo de procesamiento de varios minutos a cuestión de segundos.

Las mejoras futuras del paquete de NLP e IA generativa podrían incluir capacidades de traducción para garantizar el acceso bilingüe a todas las lenguas oficiales de España (español, catalán, gallego, valenciano y euskera).

Fuente: Información facilitada a la OCDE por funcionarios del Ministerio de Presidencia, Justicia y Relaciones con las Cortes.

Sistemas más sofisticados de IA generativa están contribuyendo a crear contenido y a realizar análisis de información. MARIA (acrónimo de "Módulo de Apoyo a la Redacción de IA") es una herramienta basada en LLM que ayuda a la Corte Suprema de Brasil a resumir documentos, elaborar primeros borradores de informes y realizar evaluaciones preliminares para identificar casos entrantes con precedentes judiciales (STF, 2024^[298]). La implementación de MARÍA aporta importantes mejoras de eficiencia; la automatización de tareas repetitivas libera tiempo a los empleados judiciales para actividades más complejas. Del mismo modo, Argentina está introduciendo un programa de aprendizaje permanente para complementar su actual batería de herramientas de IA con el fin de prestar apoyo al sector judicial (Recuadro 5.62).

Recuadro 5.62. Prometea y ChatGPT para hacer más eficiente el sector judicial de Argentina

Prometea es un sistema de IA desarrollado en 2017 por el Ministerio Público Fiscal de la Ciudad Autónoma de Buenos Aires y el Laboratorio de Innovación e Inteligencia Artificial (IALAB) de la Universidad de Buenos Aires. Diseñado para automatizar tareas judiciales repetitivas y agilizar los procedimientos, Prometea actuaba como asistente virtual, prediciendo soluciones de caso a partir de sentencias previas y ayudando a reunir archivos de caso. Al reducir significativamente los tiempos de tramitación, permitía a los funcionarios judiciales centrarse en casos más complejos que requerían un juicio humano. Entre 2017 y 2020, Prometea fue fundamental para resolver 658 casos relacionados con el derecho a la vivienda, conflictos laborales y derechos de las personas con discapacidad. El sistema alcanzó una tasa de precisión del 90 % en sus predicciones, y las decisiones de los fiscales se ajustaron a las recomendaciones de la IA en la mayoría de los casos. Al mejorar la eficiencia del flujo de trabajo, Prometea aumentó la productividad en casi un 300 %, lo que permitió la tramitación de alrededor de 490 casos al mes, en comparación con los 130 anteriores a su implementación.

Con el tiempo, la Corte Interamericana de Derechos Humanos y otros órganos jurídicos adoptaron Prometea. Sin embargo, su expansión a nuevas categorías de casos requirió un extenso reentrenamiento y ajustes algorítmicos, lo que llevó a preocupaciones sobre la posibilidad de aplicarlo a mayor escala. En 2024, los desarrolladores comenzaron a eliminar Prometea en favor de los sistemas de IA generativa, incluido ChatGPT, para determinadas tareas legales. Desde mayo de 2024, el Ministerio Público Fiscal utiliza ChatGPT para predecir fallos en algunos casos de empleo público relacionados con demandas salariales. Los empleados de justicia han subido documentos de casos a ChatGPT, que analiza patrones, ofrece una clasificación preliminar a partir de un catálogo de plantillas y redacta borradores de decisiones. Hasta ahora, ChatGPT se ha utilizado en 20 resoluciones judiciales, reduciendo el tiempo necesario de redacción de una hora a unos 10 minutos.

Fuente: (OECD/CAF, 2022^[297]),<https://restofworld.org/2024/buenos-aires-courts-adopt-chatgpt-draft-rulings>.

El tiempo que los tribunales tardan en proporcionar o rechazar medidas de protección para personas en situaciones vulnerables, como la violencia doméstica, puede tener graves consecuencias para las víctimas. Las respuestas judiciales rápidas son esenciales para garantizar la seguridad y el bienestar de las personas que se enfrentan a amenazas inmediatas. Los períodos prolongados de espera para recibir órdenes de protección pueden exponer a las víctimas a nuevos abusos. Algunos países han implementado soluciones utilizando una combinación de métodos, entre ellos la IA generativa, para ayudar a ofrecer respuestas judiciales oportunas y, cuando proceda, adoptar medidas⁵⁰ de protección para defender los derechos de las víctimas (Recuadro 5.63).

Recuadro 5.63. Rediseño de los servicios de justicia con IA para medidas de protección efectivas en Perú

El sistema *Amauta Pro* de Perú, desarrollado internamente por la Corte Superior de Justicia de Lima Norte, se ha probado con casos de violencia doméstica. Utiliza un sistema de IA y modelos normalizados de resoluciones judiciales para llevar a cabo diversas tareas esenciales para la determinación de la resolución que otorgue o deniegue medidas de protección. Las tareas automatizadas son:

- Conversión y categorización: el sistema convierte los archivos legibles por máquina en texto, clasifica cada página en función del contenido y verifica la validez de los códigos de archivo.

Esta categorización ayuda a organizar y gestionar los expedientes de casos de manera más eficiente.

- Correspondencia de casos: *Amauta Pro* busca procedimientos abiertos que involucren a las mismas partes o hechos relacionados. Si se encuentran coincidencias, el sistema prepara proyectos de resolución, racionalizando el proceso de consolidación de casos.
- Extracción y limpieza de datos: en los casos en los que no hay coincidencias, el sistema extrae datos de los informes policiales y otros documentos pertinentes, limpia el texto e identifica información clave.
- Redacción de resoluciones: el sistema redacta resoluciones a partir de datos extraídos y analizados, que los jueces pueden revisar y ultimar. Esto reduce el tiempo necesario para elaborar resoluciones.

El proyecto tiene por objeto mejorar el acceso a la justicia para las víctimas de la violencia proporcionando medidas de protección oportunas y efectivas. El sistema reduce el tiempo necesario para redactar una resolución de 3 horas a 40 segundos. El proyecto ha recibido el apoyo de los más altos niveles del poder judicial peruano, incluido el Presidente del Poder Judicial, lo que pone de relieve la importancia del proyecto y la confianza depositada en su capacidad para transformar la implementación de servicios de justicia en Perú. Con más de 800 propuestas de sentencia generadas durante su fase piloto, *Amauta Pro* permite a los jueces dedicar más tiempo y recursos a los aspectos críticos de la toma de decisiones que a tareas administrativas que consumen mucho tiempo.

Fuente: (Government of Peru, 2023^[299]).

Los algoritmos de IA también pueden entrenarse para identificar y clasificar datos, incluido texto que contenga información personal (por ejemplo, nombre, dirección) de documentos judiciales y transcripciones. Esto puede utilizarse para proteger información sensible y mantener la privacidad, aspectos críticos de la tramitación de casos judiciales. Varias prácticas nacionales demuestran la aplicación de la IA en este ámbito (Recuadro 5.64).

Recuadro 5.64. Anonimización de las decisiones judiciales impulsada por IA

España: herramienta automatizada de anonimización de documentos jurídicos

El Ministerio español de la Presidencia, Justicia y Relaciones con las Cortes ofrece una sólida herramienta de anonimización como parte de su conjunto de instrumentos de NLP e IA generativa. La herramienta elimina la información de identificación personal de los documentos a través de tres modos: automático, basado en iniciales y totalmente personalizable, lo que permite la creación de documentos y conjuntos de datos compatibles y compatibles con el Reglamento General de Protección de Datos (RGPD). Además de elegir el modo, los usuarios pueden especificar los tipos de información que desean anonimizar. Esto incluye datos relacionados con personas (por ejemplo, nombres), personas jurídicas, direcciones, documentos de identidad (por ejemplo, números de DNI o pasaporte), así como direcciones de correo electrónico y perfiles de redes sociales.

Croacia: anonimización automatizada de las decisiones judiciales

Desde enero de 2025, Croacia ha puesto en funcionamiento ANON, una herramienta de IA diseñada para automatizar la anonimización y la publicación de las decisiones judiciales. Integrada con el sistema nacional de gestión de casos eSpis, ANON anonimiza automáticamente las sentencias de primera instancia y propone su indexación para su publicación en el portal público, Buscador de Decisiones de los Tribunales de la República de Croacia (Tražilica odluka sudova Republike Hrvatske).

En el caso de los tribunales de instancias superiores, el sistema facilita la anonimización y la verificación final corre a cargo de personal autorizado de los tribunales. Esta iniciativa, respaldada por el Plan Nacional de Recuperación y Resiliencia 2021-2026, tiene como objetivo mejorar la transparencia judicial, mejorar el acceso a la información jurídica y garantizar el cumplimiento de las normas de protección de datos personales.

Austria: una combinación de enfoques para la anonimización de decisiones judiciales

El Ministerio Federal de Justicia (BMJ) de Austria está llevando a cabo un proyecto integral para automatizar la anonimización de las decisiones judiciales. Los objetivos principales del proyecto son mejorar la eficiencia, reducir el esfuerzo manual en el proceso de anonimización y permitir una publicación más amplia y el acceso a las decisiones judiciales anonimizadas. Esta iniciativa utiliza una combinación de enfoques de IA, incluidos métodos de búsqueda que utilizan datos de registros y expresiones regulares, NLP para el reconocimiento de entidades nombradas y algoritmos de ML personalizados para perfeccionar continuamente la calidad de los resultados en función de los comentarios.

Francia: anonimización impulsada por IA para la publicación de datos abiertos

El motor de anonimización impulsado por IA del Tribunal de Casación francés es un componente clave en sus esfuerzos por aumentar la transparencia y la accesibilidad de la información jurídica. Este sistema está diseñado para identificar y proteger automáticamente los datos personales en las sentencias, preparándolas para su publicación como parte de una iniciativa de datos abiertos.

Grecia: resumen y anonimización gracias a la IA

El Ministerio de Justicia de Grecia está utilizando la IA para traducir, resumir y anonimizar los documentos. Además, el Ministerio está introduciendo un asistente de IA para impartir instrucciones a la ciudadanía y abogados sobre los servicios electrónicos y no electrónicos del sistema judicial griego.

Nota: Los datos del registro son bases de datos estructuradas o registros oficiales que contienen nombres, números de caso u otra información identificativa. Las expresiones regulares son reglas de búsqueda de coincidencia de patrones utilizadas en el procesamiento de texto que ayudan a detectar y eliminar patrones específicos, como nombres, fechas o direcciones, en función de formatos predefinidos. Fuente: (Spain Ministry of Justice, 2022^[300]; Austria Federal Ministry of Justice, 2021^[301]; Sommer, 2021^[302]), Funcionarios del Ministerio de Justicia croata, <https://ministryofjustice.gr/?p=13929>, <https://ministryofjustice.gr/?p=15317>.

Mejoramiento de la implementación de servicios de justicia

En los países de la OCDE, la IA se está implementando para apoyar la implementación de servicios legales y judiciales y el acceso a ellos, facilitando las consultas iniciales a través de chatbots y asistentes virtuales. Estos sistemas de IA pueden ofrecer asesoramiento jurídico preliminar en un lenguaje accesible, clasificar los casos en función de la urgencia y la relevancia, y guiar a los usuarios en los procedimientos judiciales, como se ha hecho en Portugal (Recuadro 5.65).

Recuadro 5.65. El chatbot de Portugal para ofrecer información práctica

El Ministerio de Justicia de Portugal ha implementado una *Guía Práctica de la Justicia* (GPJ, por sus siglas en portugués), un chatbot impulsado por IA que proporciona al público un fácil acceso a la información relacionada con la justicia. La GPJ está capacitada para responder a diversas formulaciones de preguntas de manera accesible. Proporciona respuestas basadas en información de la Plataforma de Justicia Digital, ofreciendo a las personas orientación práctica sobre cuestiones

relacionadas con la justicia. El chatbot está diseñado para acelerar las interacciones entre las personas y el sistema judicial, ayudándolas a encontrar la información que necesitan de manera más eficiente. La GPJ ofrece respuestas en un lenguaje sencillo a las consultas de la gente. Contribuye a reducir la brecha informativa entre el complejo sistema jurídico y el público en general, haciendo que la información relacionada con la justicia sea más accesible para todos.

Fuente: (Government of Portugal, 2024^[303]).

Los servicios de apoyo a las víctimas también pueden beneficiarse de las innovaciones impulsadas por la IA. Estos servicios, que a menudo implican proporcionar información oportuna, apoyo jurídico y recursos a las víctimas de delitos, pueden ampliarse mediante la integración de tecnologías de IA. Los chatbots basados en IA, por ejemplo, pueden ofrecer apoyo personalizado, proporcionando a las víctimas acceso inmediato a información sobre sus derechos, los recursos disponibles y las medidas a adoptar tras un delito (véase el ejemplo en Recuadro 5.66).

Recuadro 5.66. Chatbot brinda apoyo a menores víctimas con información jurídica personalizada

El proyecto *i-ACCESS My Rights* ejecutado en Grecia, Bulgaria y Rumanía proporciona información jurídica adaptada a la edad y adaptada a los niños a partir de los 13-18 años, especialmente a los que han sido víctimas de abusos o delitos. Al ofrecer información jurídica general y orientación sobre los servicios de apoyo disponibles, el chatbot *i-ACCESS* está diseñado para empoderar a las víctimas jóvenes informándolos de sus derechos en los procedimientos judiciales y dirigiéndolos a las redes y autoridades de apoyo pertinentes. Lo que distingue a este proyecto es su fuerte énfasis en la participación de los niños en el proceso de desarrollo a través de sesiones de diseño conjunto y Juntas Consultivas Infantiles en cada país. Para garantizar que la herramienta de IA responda realmente a las necesidades y preferencias de sus jóvenes usuarios, se contó con la participación activa de niños y niñas con experiencia directa en el sistema judicial en el diseño y la funcionalidad del chatbot.

Fuente: <https://justicewithchildren.org/en/i-access-my-rights>.

Facilitar las predicciones para un sistema de justicia más proactivo

La IA puede mejorar la proactividad del sistema judicial mediante análisis predictivos. Este enfoque permite a los sistemas judiciales pasar de un enfoque reactivo a un enfoque proactivo, centrándose en la prevención y la toma de decisiones informadas. Dos de las aplicaciones más habituales de la IA en la administración y el acceso a la justicia son la vigilancia policial predictiva —en la que la IA analiza datos históricos para identificar patrones y prever posibles acontecimientos (véase la sección "La IA en la aplicación de la ley y la gestión del riesgo de catástrofes")— y las herramientas de evaluación de riesgos —en la que la IA se utiliza para evaluar la probabilidad de que las personas se conviertan en víctimas o reincidan en conductas delictivas similares en el futuro, respaldando las decisiones sobre fianza, sentencia y libertad condicional.

Algunos países han venido avanzando en la aplicación de la vigilancia predictiva y los instrumentos de evaluación de riesgos, especialmente en el contexto de la violencia doméstica y la prevención de la violencia comunitaria, con algunas implementaciones prometedoras. Tanto España como los Estados Unidos cuentan con un amplio sistema de evaluación del riesgo de las víctimas en el ámbito de la violencia

doméstica (Recuadro 5.67). Otros ejemplos relevantes sobre la actuación policial predictiva se incluyen en la sección de este capítulo sobre "La IA en la aplicación de la ley y la gestión del riesgo de desastres".

Recuadro 5.67. Sistema de evaluación de riesgos de violencia doméstica en España

El Sistema de Seguimiento Integral en los casos de Violencia de Género (VioGén) de España es una herramienta de evaluación de riesgos diseñada para mejorar la protección de las víctimas y prevenir la escalada de la violencia doméstica. Su objetivo es fortalecer la capacidad del país para responder a la violencia contra las mujeres y las niñas utilizando evaluaciones basadas en evidencia. Desde su introducción en 2007, VioGén ha tramitado más de 670 000 casos, facilitando una respuesta coordinada entre el poder judicial y las fuerzas policiales nacionales y locales. El sistema funciona mediante un proceso estructurado de evaluación de riesgos, que comienza cuando una víctima denuncia un abuso o cuando la policía interviene en un incidente doméstico. Los agentes completan un cuestionario detallado, evaluando factores como el historial del agresor, la gravedad de la violencia previa y las vulnerabilidades de la víctima. A continuación, el algoritmo asigna un nivel de riesgo — bajo, medio, alto o muy alto — que determina las medidas de protección necesarias, como las visitas policiales de control o alarmas de emergencia. Los casos considerados "de especial relevancia", en función de los niveles de riesgo, pueden ser señalados para su revisión por fiscales y jueces. Evaluaciones periódicas de seguimiento ajustan las clasificaciones de riesgos, garantizando que las respuestas sigan siendo adecuadas a lo largo del tiempo.

Fuente: (Government of Spain, 2007^[304]; Algorithm Watch, 2020^[305]).

Estas herramientas ayudan a identificar patrones de victimización y a anticipar y abordar posibles problemas antes de que se intensifiquen. Sin embargo, también requieren una implementación cuidadosa para mitigar los sesgos y mantener los estándares éticos dentro del sistema de justicia. En España, si bien los datos sugieren que la herramienta de evaluación de riesgos de VioGén (Recuadro 5.67) refuerza la protección de las víctimas, las preocupaciones relativas a la "caja negra" algorítmica, los riesgos de una posible clasificación errónea y la limitada alfabetización sobre el uso de herramientas digitales y enfoques sensibles para las mujeres y las niñas subrayan la importancia de una formación adecuada para los usuarios, así como de mecanismos de transparencia y supervisión. El algoritmo de VioGén parece seguir siendo en gran medida opaco; su fórmula exacta no se ha revelado, lo que plantea interrogantes sobre la rendición de cuentas y la equidad. Los informes también indican que, en el 95 % de los casos, los agentes se adhieren a la clasificación de riesgos de VioGén sin modificaciones, a pesar de tener la facultad discrecional de invalidarla (Algorithm Watch, 2020^[305]).

Otro ejemplo, en Estados Unidos, la Gestión de perfiles penitenciarios para la aplicación de sanciones alternativas (COMPAS por sus siglas en inglés), es una herramienta patentada de evaluación de riesgos que evalúa la probabilidad de que un acusado reincida. COMPAS, utilizado en varias jurisdicciones de Estados Unidos, fue diseñado para ayudar a los tribunales a tomar decisiones informadas con respecto a la libertad bajo fianza, la sentencia y la libertad condicional al predecir riesgos de reincidencia en general y de reincidencia violenta (Equivant, 2025^[306]). El sistema analiza los datos a partir de un cuestionario exhaustivo que abarca factores como los antecedentes penales, el consumo de sustancias, las conexiones sociales y los datos demográficos personales. Esta información se procesa a través de un algoritmo patentado para clasificar a los individuos con un riesgo bajo, medio o alto de reincidencia. Estas puntuaciones de riesgo ayudan a los profesionales jurídicos a adaptar las intervenciones y los niveles de supervisión al riesgo evaluado de cada infractor.

Si bien COMPAS ha supuesto un aumento eficiente del tiempo de procesamiento y la estandarización en todas las jurisdicciones, el sistema ha sido el centro de atención debido a su nivel de precisión del 68 %

(De Miguel Beriain, 2018^[307]; Grgić-Hlača et al., 2018^[308]). Al igual que en el caso de VioGén, la limitada transparencia algorítmica de COMPAS también planteó dudas éticas sobre si era un sistema confiable para tomar decisiones tan fundamentales como la libertad de los acusados. Desde entonces, algunas jurisdicciones han buscado alternativas más transparentes, por ejemplo, el sistema de Evaluación de la Seguridad Pública (PSA, por sus siglas en inglés), un método no patentado de evaluación del riesgo estadístico (APPR, 2025^[309]) — mientras que otros siguen utilizando herramientas como COMPAS bajo un escrutinio más estricto.

La IA tiene un enorme potencial para mejorar la proactividad de los sistemas judiciales (véase más adelante la sección "Potencial por explotar y camino a seguir"). Sin embargo, todavía queda mucho camino por recorrer antes de cosechar los beneficios de los sistemas predictivos de IA. Esto incluye invertir en la formación de los profesionales de la justicia sobre el uso de herramientas de IA, una mayor transparencia algorítmica y la supervisión humana (véase la sección "Gestión de riesgos y desafíos" más adelante).

Monitoreo del desempeño judicial y mejoramiento de la rendición de cuentas

La IA ha demostrado su efectividad en diversos escenarios para mejorar la transparencia y la rendición de cuentas de las actividades gubernamentales (OECD, 2022^[152]). Las aplicaciones de la IA para la supervisión externa e interna del sector judicial se centran en la supervisión del rendimiento de los sistemas judiciales y en la mejora de la rendición de cuentas. Los sistemas de IA pueden ofrecer información sobre el rendimiento de los tribunales, identificar posibles incoherencias en la toma de decisiones y apoyar la evaluación de las políticas. Estos sistemas deben ir acompañados de mecanismos de gobernanza preexistentes en materia de controles y equilibrios (por ejemplo, independencia administrativa y control judicial autónomo) para garantizar la autonomía del poder judicial. La supervisión del desempeño judicial basada en la IA podría conllevar riesgos de interferencias indebidas si no existen protecciones adecuadas. Los controles y contrapesos (como los órganos de supervisión independientes, la independencia administrativa y el control judicial autónomo) ayudan a evitar que las herramientas de IA se utilicen para influir en los resultados de los casos o para evaluar injustamente a los jueces. La IA puede poner de relieve patrones en las decisiones judiciales, pero, sin mecanismos de gobernanza independientes, estos conocimientos podrían utilizarse indebidamente (por ejemplo, penalizando a los jueces por resoluciones impopulares en lugar de centrarse en mejoras sistémicas).

Aunque son menos comunes que otras aplicaciones, estas actividades de supervisión representan un ámbito importante con potencial de desarrollo de la IA en los sistemas judiciales. El proyecto RAFA 2030 implementado por el Tribunal Supremo de Brasil (STF, por sus siglas en portugués) destaca como un ejemplo innovador de uso de la IA para la supervisión (Recuadro 5.68).

Recuadro 5.68. Aprovechando la IA para la supervisión en el Tribunal Supremo de Brasil

El proyecto RAFA 2030 implementado por el Tribunal Supremo de Brasil (STF) utiliza el ML y el aprendizaje profundo (DL) para clasificar los datos por categoría, procesos que figuran en la agenda plenaria del STF, origen del proceso, clase, procesos completados o en curso, y Objetivos de Desarrollo Sostenible (SDGs) 2030. RAFA 2030 no solo apoya la evaluación interna del desempeño, sino que también facilita la supervisión externa al proporcionar un marco estandarizado para evaluar la implementación de la agenda de los SDGs en el contexto de las decisiones del STF. La clasificación automatizada de casos con arreglo a los SDGs ofrece una visión clara de cómo la resolución del tribunal se ajusta a la Agenda 2030.

Además, al alinear los casos judiciales con los objetivos mundiales de sostenibilidad, la RAFA 2030 permite comparar el desempeño del STF con otros tribunales superiores internacionales. El sistema

ya ha respaldado la clasificación de 3 315 asuntos judiciales, lo que demuestra su potencial para mejorar la transparencia y alinear las actividades judiciales con objetivos sociales más amplios.

Fuente: <https://portal.stf.jus.br/hotsites/agenda-2030>.

En Israel, el Ministerio de Justicia ha puesto en marcha una iniciativa impulsada por IA dentro de la Defensoría Pública para mejorar la calidad de la representación legal y reforzar los mecanismos de supervisión. El proyecto integra tecnologías de ML, NLP y reconocimiento óptico de caracteres (OCR) para agilizar la revisión y la gestión de los expedientes de casos penales. Esta iniciativa apoya la supervisión de la calidad de la defensa al permitir el análisis automatizado de los documentos jurídicos, mejorar los controles de coherencia e identificar posibles problemas en la tramitación de los casos⁵¹.

Evidencia de impacto

Iniciativas pioneras como las discutidas anteriormente han contribuido a hacer que la información legal sea más accesible, rompiendo barreras que durante mucho tiempo han impedido que muchas personas pudieran participar plenamente en el sistema de justicia. Estas herramientas impulsadas por la IA ayudan a abordar las necesidades de muchas personas, como la identificación de los servicios judiciales más adecuados para cada situación individual, promueven una ciudadanía más informada y allanan el camino para sistemas judiciales más efectivos y eficientes.

Uno de los impactos más significativos de la IA ha sido su capacidad para acelerar diversos procesos judiciales. Al reducir el tiempo necesario para la revisión y categorización de documentos, la IA permite tratar más casos en menos tiempo. Por ejemplo, la solución anteriormente mencionada del Tribunal de Casación francés para la preclasificación de recursos consiguió un 87 % de aciertos en sus priorizaciones (Court of Cassation of France, 2024^[292]). Otro caso emblemático de mejora de la capacidad de respuesta judicial lo representa el sistema *Amauta Pro AI* del Perú (Recuadro 5.63), que ha transformado la velocidad a la que los tribunales pueden responder a las víctimas de la violencia reduciendo el tiempo necesario para redactar proyectos de resolución sobre medidas de protección de tres horas a 40 segundos. Además, está mejorando de manera fundamental la capacidad general del sistema judicial para abordar los casos de violencia contra las mujeres y los miembros de la familia, contribuyendo así a un acceso efectivo a la justicia.

Si bien existen datos de resultados e impacto para algunos casos de uso específicos, como se ha comentado en los ejemplos anteriores, los gobiernos aún no han recopilado ni documentado de manera sistemática los resultados y el impacto en muchos otros. **A pesar del creciente interés por las aplicaciones de IA en los sistemas judiciales, los países no han establecido mecanismos sistemáticos para recopilar y analizar datos sobre su impacto.** Los datos empíricos existentes siguen estando fragmentados, y a menudo se limitan a estudios piloto o evaluaciones sectoriales, con escasas evaluaciones exhaustivas a nivel organizativo o nacional (Queen Mary University of London; Centre for European Policy Studies, 2021^[310]). Esta falta de evidencia a nivel organizativo y nacional impide que los reformadores aprendan de la experiencia. Esto podría llevar a que innovaciones potencialmente efectivas en materia de IA pasen desapercibidas, mientras que herramientas no probadas se apliquen a gran escala de forma inadecuada.

Aunque algunas jurisdicciones informan de ganancias en eficiencia gracias a la gestión de casos asistida por IA o a la investigación jurídica automatizada, hay pocos datos empíricos sobre si estas herramientas mejoran la administración de justicia centrada en las personas o la precisión de la toma de decisiones. Sin una supervisión estructurada, los gobiernos se arriesgan a desplegar soluciones de IA sin evidencias claras de sus beneficios o consecuencias no deseadas, lo que limita su capacidad para tomar decisiones informadas de política pública. La falta de pruebas del impacto refleja un problema sistémico más amplio,

que es que los sistemas judiciales a menudo se han enfrentado a dificultades a la hora de recopilar y utilizar los datos de forma efectiva (World Justice Project, 2021^[311]). Históricamente, los sistemas judiciales se han basado en gran medida en procesos documentales (Byrom, Piccinin-Barbieri and Wells, 2024^[312]) (véase la sección siguiente sobre "Gestión de riesgos y desafíos").

Gestión de riesgos y desafíos

Riesgos asociados

- Datos inadecuados o sesgados en los sistemas de IA.
- Falta de transparencia y explicabilidad.
- Inexactitud, falsedades y falta de fiabilidad.
- Dependencia excesiva de la IA y resistencia a su uso por parte del gobierno.

Si los sistemas de IA se basan en **datos inadecuados o sesgados**, podrían producirse resultados inexactos o adversos para determinadas personas o grupos. En lo que respecta a la administración y el acceso a la justicia, esto podría dar lugar, por ejemplo, a evaluaciones de riesgos inexactas (por ejemplo, de la probabilidad de reincidencia) que podrían generar impactos negativos para algunas personas.

La falta de transparencia de sistemas de IA utilizados para la administración de justicia se identificó como otra preocupación en la revisión de los casos de uso. Tomando los ejemplos mencionados en la sección "Facilitar las predicciones para un sistema de justicia más proactivo", aunque se diseñaron para mejorar la eficiencia y la toma de decisiones, sistemas como VioGén y COMPAS a menudo operan como cajas negras, con sus metodologías subyacentes, la ponderación de factores y los posibles sesgos protegidos del escrutinio público. La naturaleza propietaria de determinados sistemas de IA o la falta de auditorías externas impide que los demandados comprendan o cuestionen las puntuaciones de riesgo que influyen en las decisiones judiciales, lo que suscita preocupación por el derecho a un proceso con todas las garantías y la equidad en las decisiones judiciales. Si bien la falta de transparencia algorítmica no es una cuestión poco frecuente, tiene una dimensión muy sensible en la administración y el acceso a la justicia, ya que puede ser perjudicial para los derechos humanos y las libertades fundamentales y reforzar las desigualdades sistémicas en las sociedades. **La falta de transparencia algorítmica tiene importantes repercusiones en la explicabilidad de las resoluciones judiciales.** Cuando los sistemas de IA no son transparentes, resulta complicado entender la justificación de las decisiones. Esta opacidad puede dar lugar a dificultades a la hora de cuestionar y garantizar la equidad de las decisiones, lo que genera desconfianza en el uso de las herramientas y, en última instancia, en los sistemas judiciales.

La dependencia de la IA para la rendición de cuentas, y las posibles preocupaciones por la **excesiva dependencia** de la misma en este sentido, introducen nuevos niveles de complejidad que requieren mecanismos de supervisión sólidos. **No todos los jueces están convencidos de los beneficios de utilizar sistemas de IA** en los sistemas judiciales, y tienen dudas sobre sus aplicaciones. Por ejemplo, el 27 % de los encuestados por la UNESCO entre los operadores judiciales expresaron su preocupación por la calidad de los resultados del chatbot de IA, destacando cuestiones como la inexactitud, las falsedades y la falta de fiabilidad. Los gobiernos pueden mitigar estos riesgos implementando marcos regulatorios y de políticas sólidos. España, por ejemplo, ha establecido un marco para el uso responsable y ético de la IA en la justicia, reconociendo el potencial de la misma para apoyar, pero no sustituir, la toma de decisiones jurisdiccionales (Recuadro 5.69). El Comité Científico Permanente del Ministerio de Justicia de Grecia, con el objetivo de examinar las repercusiones de la introducción de la IA en el sistema judicial, ha elaborado unas directrices que deben seguir los órganos competentes de los tribunales y las fiscalías⁵².

Recuadro 5.69. España: política nacional sobre el uso de IA en la administración de justicia

En junio de 2024, el Comité Técnico Estatal de la Administración Judicial Electrónica (CTEAJE) de España aprobó la Política de Uso de la Inteligencia Artificial en la Administración de Justicia. Este documento establece un marco fundacional para el uso responsable, legítimo y ético de la IA en contextos judiciales. Fundamentada en las disposiciones de la Ley de Justicia y Eficiencia Digital de España (Real Decreto-Ley 6/2023), que regula el uso de la IA en el sector judicial, esta política pretende equilibrar la innovación tecnológica con la integridad judicial y reconoce el potencial de la IA para apoyar, pero no sustituir, la toma de decisiones jurisdiccionales.

La Política se aplica a cualquier sistema de IA que gestione datos judiciales, desde la presentación inicial de un caso hasta su resolución final. Establece una línea clara entre los sistemas que podrían afectar a la independencia judicial y los destinados exclusivamente a fines administrativos. Las directrices son vinculantes para todo el personal que opera dentro del sistema judicial, incluidos los colaboradores externos tanto del sector público como del privado que tienen acceso a los datos judiciales. La política fue acordada por los principales actores institucionales del sistema de justicia español: el Consejo General del Poder Judicial (CGPJ), la Fiscalía General, todas las administraciones regionales con competencia sobre la Administración de Justicia y el Ministerio de Presidencia, Justicia y Relaciones con las Cortes.

Además de estos principios, la política proporciona orientación operativa, en la que se describen los usos aceptables de la IA, como la síntesis interna de documentos y la automatización de tareas administrativas no sensibles, al tiempo que prohíbe explícitamente otros. Los usos prohibidos incluyen la toma de decisiones automatizada sin supervisión humana y la generación de contenido legal vinculante basado en datos protegidos. La política también requiere un etiquetado claro de los productos generados por la IA y afirma la soberanía nacional sobre los datos y algoritmos implementados.

Las responsabilidades de supervisión están claramente definidas en el documento. Cuando las herramientas de IA tienen un impacto en las funciones judiciales, el CGPJ es responsable de realizar auditorías algorítmicas para salvaguardar la independencia judicial y garantizar la transparencia. En el caso de las herramientas utilizadas en contextos no judiciales, la supervisión se asigna a CTEAJE o a la autoridad administrativa pertinente. Otros mecanismos son el mantenimiento de registros FAT (Fairness, Accuracy, and Transparency) (Justicia, Precisión y Transparencia) para los sistemas de IA y la colaboración interinstitucional sostenida para promover la innovación ética en todo el ecosistema judicial español.

Fuente: (Spain Ministry of Justice, 2024^[313]); Información proporcionada por funcionarios del Ministerio de Presidencia, Justicia y Relaciones con las Cortes a la OCDE.

Además de la transparencia algorítmica, los países deberían considerar la posibilidad de implementar una supervisión humana efectiva del uso de las herramientas de IA en sus sistemas judiciales, incluidos sistemas de IA con procesos claros de toma de decisiones que permitan a las partes interesadas comprender las capacidades y limitaciones e interpretar los resultados, realizar un seguimiento continuo del rendimiento de la IA y permitir la intervención y el control humanos (OECD, 2024^[91]). Las directrices éticas, como la Carta Ética Europea sobre el uso de la IA en los Sistemas Judiciales y su entorno, de 2018, fueron un primer paso en este sentido, al esbozar los principios para garantizar que el desarrollo de la IA en los sistemas judiciales se ajuste a los derechos fundamentales conceptualizados en el Derecho de la UE (CEPEJ, 2018^[290]). El seguimiento y la mejora continuos deben integrarse en un marco operativo que permita auditar periódicamente el rendimiento y los resultados de la IA. Además, los planes operativos

deben tener en cuenta la naturaleza evolutiva de la legislación y ayudar a garantizar que los sistemas de IA puedan actualizarse para reflejar los cambios en la legislación o los precedentes.

Desafíos de implementación

- Falta de datos de calidad y capacidad para compartirlos.
- Brechas de habilidades.
- Falta de marcos de actuación y directrices sobre el uso de la IA.

Históricamente, los sistemas de justicia se basaban principalmente en el papel y, hasta hace poco, los **esfuerzos por establecer una gobernanza de datos sólida** en los sistemas de justicia eran limitados (Byrom, Piccinin-Barbieri and Wells, 2024^[312]). Este legado sigue estando muy presente hoy en día. Por ejemplo, según el índice OURdata de la OCDE, los países obtienen resultados relativamente mejores en la publicación de estadísticas y datos geoespaciales, mientras que los datos relativos a la justicia siguen hallándose entre aquellos con un desempeño más débil (OECD, 2023^[314]). Además, cuando se recopilan datos, a menudo se almacenan en formatos a los que no se puede acceder fácilmente (por ejemplo, en formato narrativo, PDF o archivos en papel), o se registran de forma desigual en diferentes organismos (Byrom, Piccinin-Barbieri and Wells, 2024^[312]). El sistema de justicia se ocupa de cantidades ingentes de datos complejos y matizados, como la jurisprudencia, los estatutos y los detalles de casos individuales. Es fundamental garantizar que estos datos sobre justicia sean exactos, exhaustivos y libres de sesgos históricos. Esto implica mejorar la **calidad y representatividad de los datos utilizados para formar sistemas de IA, la solidez de los sistemas que albergan dichos datos y la capacidad de gestionar de forma segura la información confidencial** (Byrom, Piccinin-Barbieri and Wells, 2024^[312]). Además, la infraestructura técnica debe ser capaz de procesar estos datos de manera eficiente, manteniendo al mismo tiempo los más altos estándares de seguridad y privacidad.

La idoneidad operativa se centra en **integrar IA sin fisuras en los procesos y flujos de trabajo existentes, mientras se mejoran las competencias digitales**. Esto incluye impartir formación a los profesionales jurídicos, incluidos los jueces, para que utilicen eficazmente las herramientas de IA e interpreten los conocimientos asistidos por la IA, así como adaptar los procesos que deben optimizarse para incorporar soluciones basadas en la IA. Algunos países han tomado medidas en este sentido conforme ponen en marcha sus herramientas de IA (CNJ, 2025^[294]).

El uso de la IA plantea interrogantes de cumplimiento normativo, que son particularmente importantes en el sistema judicial debido a la gran importancia que revisten las decisiones judiciales y los derechos fundamentales o constitucionales en juego. La Ley de IA de la UE ejemplifica la evolución del panorama normativo, al clasificar como de alto riesgo los sistemas de IA en la aplicación de la ley y la administración de justicia. Esto significa que, en la UE, tales sistemas deben cumplir obligaciones estrictas antes de poder comercializarse en el mercado europeo, incluido el requisito de someterse a una evaluación por terceros. Si bien cada vez se desarrollan más marcos generales, a menudo **faltan directrices operativas concretas sobre cómo utilizar la IA de forma confiable en un ámbito tan sensible como la administración de justicia**. El Reino Unido y Colombia han intentado establecer dichas directrices (Recuadro 5.70 y Recuadro 5.71).

Recuadro 5.70. Directrices judiciales sobre IA en el Reino Unido

La Guía Judicial sobre IA del Reino Unido proporciona directrices para los titulares de cargos judiciales sobre el uso de IA en cortes y tribunales, en las que se describen los riesgos asociados y las medidas para gestionarlos. La guía garantiza que cualquier uso de IA se ajuste a las obligaciones judiciales, en particular, a la integridad de la justicia. La orientación se aplica a todos los titulares de cargos judiciales bajo la autoridad de la Presidenta del Tribunal Supremo y Presidenta Superior de los Tribunales, así como a sus secretarios y personal de apoyo.

La guía incluye consideraciones clave, como la confidencialidad, la rendición de cuentas, los sesgos y la seguridad. Establece principios para utilizar las herramientas de IA y especifica las tareas en las que puede ayudar, como el resumen de textos y las actividades administrativas, mientras recomienda precaución en ámbitos como la investigación y el análisis jurídico. También pone de relieve los riesgos de desinformación y el uso de contenidos generados por IA en los procedimientos judiciales. Además de las directrices clave, el documento también incluye algunos ejemplos para ilustrar los posibles usos y riesgos de la IA generativa en cortes y tribunales.

Elaborada por un grupo judicial interjurisdiccional tras consultar a todos los titulares del poder judicial, la guía representa un paso inicial en el apoyo a la participación de la judicatura con la IA. Entre las iniciativas futuras se incluye la publicación de un documento de preguntas frecuentes para responder a las preguntas planteadas por el poder judicial, con revisiones en curso para adaptarse a los avances tecnológicos.

Fuente: (United Kingdom, 2023^[315]).

Recuadro 5.71. Directiva colombiana sobre normas y directrices en materia de IA para el poder judicial

En diciembre de 2024, el Consejo Superior de la Judicatura de Colombia aprobó el Acuerdo PCSJA24-12243, estableciendo directrices para el uso respetuoso, responsable, seguro y ético de la IA en el Poder Judicial. El objetivo de esta Directiva es mejorar el acceso, la transparencia y la eficiencia en la administración judicial mediante la integración de la IA en la gestión de documentos, las tareas administrativas y las actividades de apoyo judicial. El documento esboza tres componentes para su aplicación: la necesidad de programas de formación integrales; la aplicación de iniciativas y el intercambio de casos prácticos; y la elaboración de directrices sobre el uso de los sistemas de IA en el poder judicial.

La directiva fue precedida por el uso de ChatGPT en una resolución sobre la exención de los gastos médicos de un niño autista. En enero de 2023, un juez de Cartagena incorporó respuestas generadas por IA a su decisión de agilizar el proceso de redacción. Esta acción desató un importante debate sobre la confiabilidad y ética de la IA en las resoluciones judiciales. En respuesta, el Consejo Superior de la Judicatura emitió el Acuerdo PCSJA24-12243, haciendo hincapié en la necesidad de directrices claras para garantizar que las herramientas de IA se utilicen adecuadamente. El incidente subraya la necesidad urgente de desarrollar la alfabetización digital dentro del poder judicial sobre las capacidades y limitaciones de la IA para evitar la dependencia excesiva de la tecnología y mantener la supervisión humana en la toma de decisiones judiciales. Desde el Acuerdo PCSJA24-12243, se han puesto en marcha varias acciones e iniciativas, como encuestas para evaluar el uso de la IA por parte del personal judicial, programas de formación con profesionales del derecho para introducir herramientas asistidas por IA en las operaciones judiciales, y asociaciones con universidades y organizaciones internacionales para talleres y programas de formación.

Fuente: (Colombia, 2024^[316]; Colombia, 2024^[316]; UNESCO, 2025^[317]).

Potencial por explotar y camino a seguir

Las personas y los profesionales del derecho pueden utilizar aplicaciones de IA para servicios como el asesoramiento jurídico, el acceso a la información o la preparación de documentos. Los sistemas de justicia bien diseñados -eficientes, transparentes y fáciles de usar- tienen más probabilidades de ser accesibles y responder a las necesidades del público. Los algoritmos avanzados y las técnicas de ML de la IA pueden contribuir a garantizar que los procesos judiciales sean más coherentes, transparentes y libres de sesgos humanos, y pueden ayudar a los tribunales a supervisar su eficiencia y recursos. La IA también puede mejorar la rendición de cuentas al señalar incoherencias y posibles sesgos en tiempo real, programando los sistemas para monitorear y alertar a las partes interesadas sobre patrones inusuales que podrían indicar sesgo o mala conducta. Por ejemplo, si los datos revelan que se está actuando de manera desproporcionada contra determinados grupos demográficos o se les están imponiendo penas más severas, estas discrepancias pueden identificarse y abordarse sin demora. Estas alertas y controles en tiempo real pueden dar lugar a revisiones inmediatas y medidas correctivas, lo que mejora la equidad y la integridad generales del sistema judicial.

Los sistemas de IA deben ir acompañados de mecanismos de gobernanza preexistentes en materia de controles y contrapesos (por ejemplo, independencia administrativa y control judicial autónomo) para garantizar la autonomía del poder judicial. Estos mecanismos de gobernanza se ajustan a la visión establecida en el *Convenio Marco del Consejo de Europa sobre Inteligencia Artificial, Derechos Humanos, Democracia y Estado de Derecho* (2024^[318]), que subraya la importancia de preservar la independencia judicial y el acceso a la justicia a la hora de implantar sistemas de IA.

La confianza es quizás el aspecto más crucial del sistema judicial, puesto que las decisiones pueden tener un profundo impacto en la vida de las personas y en la confianza de la sociedad. Este aspecto abarca cuestiones de transparencia, explicabilidad, equidad y rendición de cuentas, igualdad de trato, preocupaciones éticas y brechas en materia de privacidad y justicia que deben ser atendidas con especial cuidado. Del mismo modo, el potencial de los sistemas de IA para perpetuar o amplificar las disparidades sociales existentes y las brechas de justicia también debe ser abordado con similar atención. Los sistemas de IA deben diseñarse de manera que no se perpetúen o agraven los sesgos existentes en el sistema judicial, en particular por lo que respecta a la raza, el sexo o la situación socioeconómica.

La transparencia algorítmica es esencial no solo por razones éticas, sino también para cumplir con los requisitos legales del proceso con todas las garantías. La explicabilidad de las decisiones basadas en la IA es crucial para mantener la confianza pública y permitir apelaciones significativas. Las consideraciones éticas deberían extenderse a todas las partes interesadas, incluidos los acusados, las víctimas, los profesionales del Derecho y la sociedad en general.

El potencial transformador de la IA en el ámbito de la justicia exige un enfoque centrado en las personas que promueva la transformación digital y proteja los derechos humanos. Una gobernanza efectiva implica la colaboración entre responsables políticos, profesionales del derecho y técnicos para alinear las aplicaciones de la IA con los principios de justicia, equidad e integridad procesal. Las evaluaciones de impacto continuas, la participación de las partes interesadas y los mecanismos de reparación mejoran tanto la legitimidad como la efectividad. Esto pone de relieve la importancia de diseñar sistemas de IA como herramientas que apoyen la toma de decisiones judiciales y la administración de justicia y que refuerzen la capacidad de impartir justicia centrada en las personas.

En este contexto, sigue siendo fundamental establecer marcos de gobernanza sólidos y mecanismos de supervisión continua para garantizar un uso confiable y equitativo de la IA en los sistemas judiciales. Para apoyar la gobernanza de la IA en la justicia y garantizar su uso responsable, los *Principios de la OCDE sobre IA* (Tabla 2.2), el *Marco de la OCDE para una IA confiable en el gobierno* (Gráfico 4.6) y el *Marco Común de la OCDE de Notificación de Incidentes de IA* (2025^[319]) pueden servir como buenos puntos de partida. Los *Principios de IA de la OCDE* abogan por el crecimiento inclusivo, el desarrollo sostenible y el

bienestar, haciendo hincapié en la importancia de los valores y la equidad centrados en el ser humano. Se exige que los sistemas de IA sean transparentes y que sus decisiones sean comprensibles para quienes se ven afectados por ellas. El *Marco para una IA confiable en el Gobierno* puede ser una herramienta útil para apoyar el uso responsable de la IA en el sector judicial. Por último, el *Marco Común de la OCDE para la Notificación de Incidentes de IA* puede ayudar al sector judicial a supervisar los incidentes de IA, garantizando que estos sistemas rindan cuentas, sean confiables y beneficiosos, al tiempo que mitigan los posibles riesgos asociados a su despliegue.

Bibliografía

- Abrell, J., M. Kosch and S. Rausch (2022), "How effective is carbon pricing?—A machine learning approach to policy evaluation", *Journal of Environmental Economics and Management*, Vol. 112, p. 102589, <https://doi.org/10.1016/j.jeem.2021.102589>. [134]
- Access Now (2022), *Ban Biometric Surveillance*, <https://www.accessnow.org/campaign/ban-biometric-surveillance/>. [278]
- ACE (2025), *Government administrative data sources for evaluation in Australia*, <https://evaluation.treasury.gov.au/sites/evaluation.treasury.gov.au/files/2025-01/government-administrative-data-sources-for-evaluation-in-australia.pdf>. [136]
- Agencia Nacional de Contratación Pública –Colombia Compra Eficiente (2024), *VISUALIZACIONES COMPROA PÚBLICA*, <https://www.colombiacompra.gov.co/content/visualizaciones-compra-publica>. [79]
- AI.GOV.UK (2024), *i.AI Consultation Analyser*, <https://ai.gov.uk/consultations/>. [164]
- Aldemir, C. and T. Uçma Uysal (2025), "Artificial Intelligence for Financial Accountability and Governance in the Public Sector: Strategic Opportunities and Challenges", *Administrative Sciences*, Vol. 15/2, p. 58, <https://doi.org/10.3390/admsci15020058>. [222]
- Algorithm Watch (2020), *Automating Society Report*, <https://automatingsociety.algorithmwatch.org/report2020/spain/spain-story/>. [305]
- Alon-Barkat, S. and M. Busuioc (2022), "Human–AI Interactions in Public Sector Decision Making: “Automation Bias” and “Selective Adherence” to Algorithmic Advice", *Journal of Public Administration Research and Theory*, Vol. 33/1, pp. 153-169, <https://doi.org/10.1093/jopart/muac007>. [130]
- Alowais, S. et al. (2023), "Revolutionizing healthcare: the role of artificial intelligence in clinical practice", *BMC Medical Education*, Vol. 23/1, <https://doi.org/10.1186/s12909-023-04698-z>. [215]
- Amos, J. (2020), "Bushfires: Australian satellite would be 'tuned' to eucalypt vegetation", <https://www.bbc.com/news/science-environment-51727231> (accessed on 4 March 2025). [269]
- Andersson, P., K. Arbin and C. Rosenqvist (2025), "Assessing the value of artificial intelligence (AI) in governmental public procurement", *Journal of Public Procurement*, Vol. 25/1, pp. 120-139, <https://doi.org/10.1108/jopp-05-2024-0057>. [83]

- Androutsopoulou, A. et al. (2019), "Transforming the communication between citizens and government through AI-guided chatbots.", *Government information quarterly*, Vol. 36/2, pp. 358-367, <https://doi.org/10.1016/j.giq.2018.10.001>. [169]
- Ansari, B., M. Barati and E. Martin (2022), "Enhancing the usability and usefulness of open government data: A comprehensive review of the state of open government data visualization research", *Government Information Quarterly*, Vol. 39/1, p. 101657, <https://doi.org/10.1016/j.giq.2021.101657>. [78]
- Anshari, M. et al. (2024), "Public service delivery, artificial intelligence and the sustainable development goals: trends, evidence and complexities", *Journal of Science and Technology Policy Management*, Vol. 16/1, pp. 163-181, <https://doi.org/10.1108/jstpm-07-2023-0123>. [193]
- Aoki, N. (2020), "An experimental study of public trust in AI chatbots in the public sector", *Government Information Quarterly*, Vol. 37/4, p. 101490, <https://doi.org/10.1016/j.giq.2020.101490>. [250]
- Aoki, N., M. Tay and M. Yarime (2024), "Trustworthy public sector AI: research progress and future agendas", in *Research Handbook on Public Management and Artificial Intelligence*, Edward Elgar Publishing, <https://doi.org/10.4337/9781802207347.00026>. [251]
- APA (2023), *Worries about artificial intelligence, surveillance at work may be connected to poor mental health*, <https://www.apa.org/news/press/releases/2023/09/artificial-intelligence-poor-mental-health>. [55]
- APPR (2025), *About the Public Safety Assessment*, <https://advancingpretrial.org/psa/about/>. [309]
- Arana-Catania, M. et al. (2021), "Citizen participation and machine learning for a better democracy", *Digital Government: Research and Practice*, Vol. 2/3, pp. 1-22, <https://doi.org/10.1145/3452118>. [159]
- Araszkiewicz, M. and V. Rodríguez-Doncel (2019), *Legal Knowledge and Information Systems: JURIX 2019: The Thirty-second Annual Conference*, <https://ebooks.iospress.nl/volume/legal-knowledge-and-information-systems-jurix-2019-the-thirty-second-annual-conference>. [149]
- Ash, E., S. Galletta and T. Giommoni (2020), "A Machine Learning Approach to Analyzing Corruption in Local Public Finances", *SSRN Electronic Journal*, <https://doi.org/10.2139/ssrn.3589545>. [101]
- Attard, J. et al. (2015), "A systematic review of open government data initiatives", *Government Information Quarterly*, Vol. 32/4, pp. 399-418, <https://doi.org/10.1016/j.giq.2015.07.006>. [77]
- Australian Government (2020), *Department of Veterans' Affairs Annual Report 2019-20*, <https://www.transparency.gov.au/publications/veterans-s-affairs/department-of-veterans-affairs/department-of-veterans-affairs-annual-report-2019-20>. [10]
- Austria Federal Ministry of Justice (2021), *Anonymization of court decisions in Austria: Webinar on the use of AI in the justice field*, https://commission.europa.eu/system/files/2021-04/anonymisation_webinar_29032021_austria.pdf (accessed on 19 February 2025). [301]
- Babšek, M. et al. (2025), "Artificial Intelligence Adoption in Public Administration: An Overview of Top-Cited Articles and Practical Applications", *AI*, Vol. 6/3, p. 44, <https://doi.org/10.3390/ai6030044>. [198]

- Bailey, K. (2023), *The Ethics of AI in Finance: How to Detect and Prevent Bias*, [19]
<https://corporatefinanceinstitute.com/resources/data-science/ai-ethics-in-finance-detect-prevent-bias/> (accessed on 10 April 2025).
- Bastani, H. et al. (2021), “Efficient and targeted COVID-19 border testing via reinforcement learning”, *Nature*, Vol. 599/7883, pp. 108-113, <https://doi.org/10.1038/s41586-021-04014-z>. [216]
- Bénassy-Quéré, A. (2022), *Ex ante, ex post: tuning the two pillars of policy evaluation*, [111]
<https://www.tresor.economie.gouv.fr/Articles/2022/02/25/ex-ante-ex-post-tuning-the-two-pillars-of-policy-evaluation>.
- Bennett Institute (2024), *Public Health gets personal: The case for an AI-driven personalised prevention platform*, <https://www.bennettinstitute.cam.ac.uk/wp-content/uploads/2024/03/The-case-for-an-AI-driven-personalised-prevention-platform.pdf>. [242]
- Berditchevskaia, A. and P. Baeck (2020), *The Future of Minds and Machines: How artificial intelligence can enhance collective intelligence*. [165]
- BEREC (2023), *Report on the impact of Artificial Intelligence (AI) solutions in the telecommunications sector on regulation*, <https://www.berec.europa.eu/en/document-categories/berec/reports/berec-report-on-the-impact-of-artificial-intelligence-ai-solutions-in-the-telecommunications-sector-on-regulation>. [35]
- Berman, A., K. de Fine Licht and V. Carlsson (2024), “Trustworthy AI in the public sector: An empirical analysis of a Swedish labor market decision-support system”, *Technology in Society*, Vol. 76, p. 102471, <https://doi.org/10.1016/j.techsoc.2024.102471>. [200]
- Berryhill, J. et al. (2019), “Hello, World: Artificial intelligence and its use in the public sector”, *OECD Working Papers on Public Governance*, No. 36, OECD Publishing, Paris, <https://doi.org/10.1787/726fd39d-en>. [191]
- Better Evaluation Knowledge (2022), *Rapid Evaluation*, [135]
<https://www.betterevaluation.org/methods-approaches/approaches/rapid-evaluation>.
- Birhane, A. et al. (2022), “The Values Encoded in Machine Learning Research”, *2022 ACM Conference on Fairness, Accountability, and Transparency*, <https://doi.org/10.1145/3531146.3533083>. [226]
- Birks, D., M. Townsley and T. Hart (2023), *Predictive policing in an Australian context: assessing viability and utility*, Australian Institute of Criminology, <https://doi.org/10.52922/ti78870>. [283]
- Blaizot, A. et al. (2022), “Using artificial intelligence methods for systematic review in health sciences: A systematic review”, *Research Synthesis Methods*, Vol. 13/3, pp. 353-362, <https://doi.org/10.1002/jrsm.1553>. [114]
- Blanc, F. (2018), *From Chasing Violations to Managing Risks Origins, challenges and evolutions in regulatory inspections*, [43]
<https://www.elgaronline.com/monobook/9781788112482/9781788112482.xml> (accessed on 28 May 2024).
- Blanchet, M. and M. Coueffe (2020), *Improved GDP nowcasting using large datasets*, Ministère de l’Économie et des Finances, Direction générale du Trésor, [7]
<https://www.tresor.economie.gouv.fr/Articles/2020/04/09/tresor-economics-no-254-improved-gdp-nowcasting-using-large-datasets>.

- Bloomberg Philanthropies (2023), *State of Cities: Generative AI in Local Governments*, [225]
https://cityaiconnect.jhu.edu/pdfs/Final-Gen-AI-In-Cities-Report_10.18.2023.pdf.
- Bohni Nielsen, S., F. Mazzeo Rinaldi and G. Petersson (2024), *Artificial Intelligence and Evaluation*, Routledge, New York, [118]
<https://doi.org/10.4324/9781003512493>.
- Boström, H. et al. (2020), *Explaining Multivariate Time Series Forecasts: An Application to Predicting the Swedish GDP*, [17]
https://ceur-ws.org/Vol-2796/xi-ml-2020_bostrom.pdf.
- Bright, J. et al. (2025), “Generative AI is already widespread in the public sector: evidence from a survey of UK public sector professionals”, *Digital Government: Research and Practice*, [189] Vol. 6/1, pp. 1-13, <https://doi.org/10.1145/3700140>.
- Bright, J. et al. (2019), “Data Science for Local Government”, *SSRN Electronic Journal*, [327]
<https://doi.org/10.2139/ssrn.3370217>.
- Brioscú, A. et al. (2024), “A new dawn for public employment services: Service delivery in the age of artificial intelligence”, *OECD Artificial Intelligence Papers*, No. 19, OECD Publishing, Paris, [213]
<https://doi.org/10.1787/5dc3eb8e-en>.
- Broecke, S. (2023), “Artificial intelligence and labour market matching”, *OECD Social, Employment and Migration Working Papers*, No. 284, OECD Publishing, Paris, [52]
<https://doi.org/10.1787/2b440821-en>.
- Brougham, D. and J. Haar (2017), “Smart Technology, Artificial Intelligence, Robotics, and Algorithms (STARA): Employees’ perceptions of our future workplace”, *Journal of Management & Organization*, Vol. 24/2, pp. 239-257, [58]
<https://doi.org/10.1017/jmo.2016.55>.
- Burger, M., A. Nitsche and J. Arlinghaus (2023), “Hybrid intelligence in procurement: Disillusionment with AI’s superiority?”, *Computers in Industry*, Vol. 150, [90]
<https://doi.org/10.1016/j.compind.2023.103946>.
- Byrom, N., M. Piccinin-Barbieri and P. Wells (2024), “Towards effective governance of justice data”, *OECD Working Papers on Public Governance*, No. 74, OECD Publishing, Paris, [312]
<https://doi.org/10.1787/d2950e02-en>.
- Campagnucci, F. et al. (2025), *Artificial Intelligence for Participation*, [168]
https://scope.uni-muenster.de/wp-content/uploads/2025/01/ENG-Policy-Brief-AI_Jan2025.pdf.
- Cappelli, P. and N. Rogovsky (2023), “Artificial intelligence in human resource management”, [59]
ILO working paper, <https://doi.org/10.54394/ohvv4382>.
- Carrasco, M. et al. (2019), *The Citizen’s Perspective on the Use of AI in Government*, [187]
<https://www.bcg.com/en-ca/publications/2019/citizen-perspective-use-artificial-intelligence-government-digital-benchmarking>.
- Cary Coglianese (2024), “How to Regulate Artificial Intelligence”, *The Regulatory Review*, [39]
<https://www.theregreview.org/2024/01/15/coglianese-how-to-regulate-artificial-intelligence/>
 (accessed on 28 May 2024).
- CEPEJ (2018), *European ethical charter on the use of artificial intelligence in judicial systems and their environment*, [290]
<https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>.

- Chen, T., M. Gascó-Hernandez and M. Esteve (2023), "The Adoption and Implementation of Artificial Intelligence Chatbots in Public Organizations: Evidence from U.S. State Governments", *The American Review of Public Administration*, Vol. 54/3, pp. 255-270, <https://doi.org/10.1177/02750740231200522>. [207]
- ChileCompra (2024), *Datos Abiertos de las compras públicas de Chile*, <https://datos-abiertos.chilecompra.cl/>. [80]
- Cintron, D. and B. Montrosse-Moorhead (2021), "Integrating Big Data Into Evaluation: R Code for Topic Identification and Modeling", *American Journal of Evaluation*, Vol. 43/3, pp. 412-436, <https://doi.org/10.1177/10982140211031640>. [123]
- CMS (2023), *CMS Expert Guide to Digital Litigation in Slovenia*, <https://cms.law/en/int/expert-guides/cms-expert-guide-to-digital-litigation/slovenia> (accessed on June 2024). [296]
- CNJ (2025), *Labor Court launches artificial intelligence to assist professionals at the institution*, <https://www.cnj.jus.br/justica-do-trabalho-lanca-inteligencia-artificial-para-auxiliar-profissionais-da-instituicao/>. [294]
- Cockx, B., M. Lechner and J. Bollens (2023), "Priority to unemployed immigrants? A causal machine learning evaluation of training in Belgium", *Labour Economics*, Vol. 80, p. 102306, <https://doi.org/10.1016/j.labeco.2022.102306>. [204]
- Collins, B. (2020), *Transformational Results - Purchasing and Strategic Sourcing*, <https://www2.elpasotexas.gov/municipal-clerk/agenda/04-27-20/1.1.2.pdf>. [69]
- Colombia (2024), *Information System of the Presidential Rapporteurship*, <https://actosadministrativos.ramajudicial.gov.co/web/Acto%20Administrativo/Default.aspx?ID=19280>. [316]
- Cortés-Cediel, M. (2023), "Trends and challenges of e-government chatbots: Advances in exploring open government data and citizen participation content.", *Government Information Quarterly*, Vol. 40/4, p. 101877, <https://doi.org/10.1016/j.giq.2023.101877>. [171]
- Council of Europe (2024), *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>. [318]
- Court of Cassation of France (2024), *The Court of Cassation facing digital technology and artificial intelligence*, <https://www.vie-publique.fr/parole-dexpert/278415-la-cour-de-cassation-face-au-numerique-et-lintelligence-artificielle>. [292]
- Crisanto, J. et al. (2024), *Regulating AI in the financial sector: recent developments and main challenges*, <https://www.bis.org/fsi/publ/insights63.pdf>. [18]
- CSOs (2021), *An EU Artificial Intelligence Act for Fundamental Rights: A Civil Society Statement*, <https://www.accessnow.org/wp-content/uploads/2021/11/joint-statement-EU-AIA.pdf>. [279]
- da Rosa, I. (2023), *Repository of existing projects on transparency in public procurement*, Publications Office of the European Union, https://ted.europa.eu/documents/d/ted/final_analysis-report-on-the-inventory-of-contract-registers_isabel-da-rosa. [86]

- de Blasio, G., A. D'Ignazio and M. Letta (2022), "Gotham city. Predicting 'corrupted' municipalities with machine learning", *Technological Forecasting and Social Change*, Vol. 184, p. 122016, <https://doi.org/10.1016/j.techfore.2022.122016>. [102]
- De Miguel Beriain, I. (2018), "Does the use of risk assessments in sentences respect the right to due process? A critical analysis of the Wisconsin v. Loomis ruling", *Law, Probability and Risk*, Vol. 17/1, pp. 45 - 53, <https://doi.org/10.1093/lpr/mgy001>. [307]
- Dell'Acqua, F. et al. (2023), "Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality", *SSRN Electronic Journal*, <https://doi.org/10.2139/ssrn.4573321>. [126]
- Dieterle, E., C. Dede and M. Walker (2024), "The cyclical ethical effects of using artificial intelligence in education", *AI & SOCIETY*, Vol. 39/2, pp. 633-643, <https://doi.org/10.1007/s00146-022-01497-w>. [233]
- DRCF (2023), *Launch of the International Network for Digital Regulation Cooperation (INDRC)*, <https://www.drcf.org.uk/news-and-events/news/launch-of-international-network-for-digital-regulation-cooperation-indrc>. [48]
- EC (2020), *Study on up-take of emerging technologies*, European Commission, https://joinup.ec.europa.eu/sites/default/files/news/2020-06/D.01.06_Final_report_v3.00.pdf. [66]
- EC JRC (2021), *Selected AI cases in the public sector (JRC129301)*, <http://data.europa.eu/89h/7342ea15-fd4f-4184-9603-98bd87d8239a>. [217]
- ECA (2024), *Special report 05/2024: EU Transparency Register – provides useful but limited information on lobbying activities*, <https://www.eca.europa.eu/en/publications/SR-2024-05>. [97]
- ECMWF (2025), *ECMWF's AI forecasts become operational*, <https://www.ecmwf.int/en/about/media-centre/news/2025/ecmwfs-ai-forecasts-become-operational> (accessed on 4 March 2025). [272]
- EDPB-EDPS (2021), *EDPB-EDPS Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*, https://www.edpb.europa.eu/our-work-tools/our-documents/edpbedps-joint-opinion/edpb-edps-joint-opinion-52021-proposal_en. [280]
- Eisele, A. et al. (2024), "Exploring the potential of Claude 2 for risk of bias assessment: Using a large language", p. 10.1101/2024.07.16.24310483. [127]
- Eisen, N. et al. (2023), *AI can strengthen U.S. democracy—and weaken it*, <https://www.brookings.edu/articles/ai-can-strengthen-u-s-democracy-and-weaken-it/>. [172]
- El Morr, C. et al. (2024), "AI-based epidemic and pandemic early warning systems: A systematic scoping review", *Health Informatics Journal*, Vol. 30/3, <https://doi.org/10.1177/14604582241275844>. [271]
- Equivant (2025), *Solutions: Risk Needs Assessments*, <https://equivant-supervision.com/solutions/risk-needs-assessments/#COMPAS-R-P>. [306]
- ERIC (2024), *Electronic Registration Information Center (ERIC): Technology and*, Electronic Registration Information Center, <https://ericstates.org/wp-content/uploads/documents/ERIC-Tech-Security-Brief.pdf>. [174]

- EUCPN (2022), *Artificial intelligence and predictive policing: risks and challenges*, [284]
<https://eucpn.org/document/recommendation-paper-artificial-intelligence-and-predictive-policing-risks-and-challenges>.
- European Parliament (2022), *Report on Artificial Intelligence in a Digital Age*, [155]
https://www.europarl.europa.eu/doceo/document/A-9-2022-0088_EN.html.
- Evans, N. (2022), “Policing the pandemic in Australia and New Zealand: lessons for trust and legitimacy”, *Journal of Criminological Research, Policy and Practice*, Vol. 9/2, pp. 106-122, [285]
<https://doi.org/10.1108/jcrpp-10-2022-0050>.
- Faculty (2021), *Automated approaches to measuring online experiences*, [37]
https://www.ofcom.org.uk/_data/assets/pdf_file/0015/220425/automated-tooling-report.pdf.
- Ferretti, S. (2023), “Hacking by the prompt: Innovative ways to utilize ChatGPT for evaluators”, [133]
New Directions for Evaluation, Vol. 2023/178-179, pp. 73-84,
<https://doi.org/10.1002/ev.20557>.
- Fillet, S. (2024), *Cambridge City Council analyzes input 50% faster with Go Vocal's AI assistant*, [167]
<https://www.govocal.com/case-studies/cambridge-city-council-analyzes-input-50-faster-with-go-vocals-ai-assistant>.
- Fishkin, J. and A. al (2021), *Final Report - LXS 400 - Chile Delibera*, [163]
https://drive.google.com/file/d/1Vt5XaTk2MW9M_VBN2oztK2nxYZdckI89/view.
- Franzen, S. et al. (2022), *Advanced Content Analysis: Can Artificial Intelligence Accelerate Theory-Driven Complex Program Evaluation?*, World Bank Group, [122]
<https://documents1.worldbank.org/curated/en/400031645128516191/pdf/Advanced-Content-Analysis-Can-Artificial-Intelligence-Accelerate-Theory-Driven-Complex-Program-Evaluation.pdf>.
- French Public Finances General Directorate (DGFiP) (2024), *Digital Transformation of the Public Finances General Directorate (DGFiP)*. [13]
- French Public Finances General Directorate (DGFiP) (2024), *The predictive model*. [12]
- Friton, P. et al. (2024), *Public Procurement of AI: Current and Future Challenges*, [62]
<https://www.lexology.com/library/detail.aspx?g=2c131357-03ed-4965-aa19-e28a65dd9ad6>
 (accessed on 30 May 2024).
- Garcia Rodriguez, M. et al. (2022), “Collusion detection in public procurement auctions with machine learning algorithms”, *Automation in Construction*, Vol. 133, p. Article 104047, [71]
<https://doi.org/10.1016/j.autcon.2021.104047>.
- Gastaldi, L. et al. (2024), *AI in Public Settings: Status and Next Steps*, [64]
https://www.dt.mef.gov.it/export/sites/sitdt/modules/documenti_it/analisi_progammazione/note_tematiche/Nota-Tematica-n-1-February-2024.pdf.
- Gatto, L. and P. Bundi (2025), *The Use of Quantitative Text Analysis in Evaluations*, Routledge, [120]
<https://library.oapen.org/bitstream/handle/20.500.12657/94001/1/9781040128510.pdf>.
- Gesk, T. and M. Leyer (2022), “Artificial intelligence in public services: When and why citizens accept its usage”, *Government Information Quarterly*, Vol. 39/3, p. 101704, [203]
<https://doi.org/10.1016/j.giq.2022.101704>.

- GFDRR (2018), *Machine Learning for Disaster Risk Management*, [266]
<http://creativecommons.org/licenses/by/3.0/igo/www.worldbank.org>.
- Giest, S. and S. Grimmelikhuijsen (eds.) (2020), “Artificial intelligence, bureaucratic form, and discretion in public service”, *Information Polity*, Vol. 25/4, pp. 491-506, [247]
<https://doi.org/10.3233/IP-200223>.
- Glas, A. and F. Kleeman (2016), “The Impact of Industry 4.0 on Procurement and Supply Management: A Conceptual and Qualitative Analysis”, *International Journal of Business and Management Invention*, Vol. 5/6, pp. 55-66, [65]
<https://doi.org/10.1590/0103-6513.20180104>.
- GLOBO (2018), *Como as robôs Alice, Sofia e Monica ajudam o TCU a caçar irregularidades em licitações*, [82]
<https://g1.globo.com/economia/tecnologia/noticia/como-as-robos-alice-sofia-e-monica-ajudam-o-tcu-a-acabar-irregularidades-em-licitacoes.ghtml>.
- Gmyrek, P., J. Berg and D. Bescond (2023), *Generative AI and Jobs: A global analysis of potential effects on job quantity and quality*, ILO, [57]
<https://doi.org/10.54394/FHEM8239>.
- GOV.UK (2019), *Guidance - Understanding artificial intelligence ethics and safety*, [85]
<https://www.gov.uk/guidance/understanding-artificial-intelligence-ethics-and-safety>.
- Government of Iceland (2023), *Head Start for Icelandic*, [180]
<https://www.government.is/diplomatic-missions/embassy-article/2023/03/14/Head-start-for-Icelandic/> (accessed on 23 May 2024).
- Government of Mexico (2024), *Programa Anual de Adquisiciones, Arrendamientos, Servicios y Obra Pública*, [81]
<https://upcp-compranet.hacienda.gob.mx/paaasopashboard>.
- Government of Mexico (2023), *Plataforma inteligente de apoyos del Gobierno Federal*, [16]
<https://www.transparenciapresupuestaria.gob.mx/es/apoyosdelgobierno> (accessed on 20 April 2024).
- Government of Peru (2023), *Sistema con inteligencia artificial permitirá a jueces resolver solo casos de violencia severo o muy severo*, [299]
<https://www.gob.pe/institucion/pj/noticias/774687-sistema-con-inteligencia-artificial-permitira-a-jueces-resolver-solo-casos-de-violencia-severo-o-muy-severo>.
- Government of Portugal (2024), *Guia Prático da Justiça*, [303]
<https://justica.gov.pt/Servicos/Guia-pratico-da-Justica>.
- Government of Spain (2007), *VioGén*, [304]
<https://www.interior.gob.es/opencms/es/servicios-al-ciudadano/violencia-contra-la-mujer/sistema-viogen/>.
- Green, B. (2022), “The flaws of policies requiring human oversight of government algorithms”, *Computer Law & Security Review*, Vol. 45, p. 105681, [236]
<https://doi.org/10.1016/j.clsr.2022.105681>.
- Grgić-Hlača, N. et al. (2018), “Beyond Distributive Fairness in Algorithmic Decision Making: Feature Selection for Procedurally Fair Learning”, *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 32/1, [308]
<https://doi.org/10.1609/aaai.v32i1.11296>.
- Guida, M. et al. (2023), “The role of artificial intelligence in the procurement process: State of the art and research agenda”, *Journal of Purchasing and Supply Management*, Vol. 29/2, p. Article 100823, [67]
<https://doi.org/10.1016/j.pursup.2023.100823>.

- Guio, A. and K. Müller-Daubermann (2024), *Insights and Recommendations for Policy Makers in Latin America and the Caribbean*, Global Network of Internet and Society Research Centers, <https://networkofcenters.net/sites/networkofcenters.net/files/Summary%20Report%20AI%20Governance%20Costa%20Rica%20NoC.pdf>. [27]
- Hadfi, R. et al. (2021), "Argumentative Conversational Agents for Online Discussions", *Journal of Systems Science and Systems Engineering*, <https://doi.org/10.1007/s11518-021-5497-1>. [176]
- Halpern, D. and D. Maru (2024), *Global Evidence Report: A blueprint for better international collaboration on evidence*, <https://www.bi.team/wp-content/uploads/2024/08/ESRC-Global-Evidence-Report-September-2024.pdf>. [139]
- Handfield, R., S. Jeong and T. Choi (2019), "Emerging procurement technology: data analytics and cognitive analytics", *International Journal of Physical Distribution & Logistics Management*, Vol. 49/10, <https://doi.org/10.1108/IJPDLM-11-2017-0348>. [87]
- Heikkila, M. (2022), *AI: Decoded: A Dutch algorithm scandal serves a warning to Europe — The AI Act won't save us*, <https://www.politico.eu/newsletter/ai-decoded/a-dutch-algorithm-scandal-serves-a-warning-to-europe-the-ai-act-wont-save-us-2/>. [92]
- Hickok, M. (2022), *Public procurement of artificial intelligence systems: new risks and future proofing.*, <https://doi.org/10.1007/s00146-022-01572-2>. [63]
- Horowitz, M. (2023), *Bending the Automation Bias Curve: A Study of Human and AI-based Decision Making in National Security Contexts*, <https://arxiv.org/abs/2306.16507>. [129]
- Hung, T. and C. Yen (2020), "On the person-based predictive policing of AI", *Ethics and Information Technology*, Vol. 23/3, pp. 165-176, <https://doi.org/10.1007/s10676-020-09539-x>. [259]
- Hutchins, B. and M. Andrejevic (2021), *Olympian Surveillance: Sports Stadiums and the Normalization of Biometric Monitoring*, <https://ijoc.org/index.php/ijoc/article/view/16377>. [256]
- IBM (2023), *Government Procurement and Acquisition: Opportunities and challenges presented by artificial intelligence and machine learning*, IBM Center for The Business of Government, https://www.businessofgovernment.org/sites/default/files/Government%20Procurement%20and%20Acquisition_0.pdf. [88]
- Jacob, S. (2025), "Artificial Intelligence and the Future of Evaluation: From Augmented to Automated Evaluation", *Digital Government: Research and Practice*, Vol. 6/1, pp. 1-10, <https://doi.org/10.1145/3696009>. [125]
- Jarrahi, M. et al. (2023), "Artificial intelligence and knowledge management: A partnership between human and AI", *Business Horizons*, Vol. 66/1, pp. 87-99, <https://doi.org/10.1016/j.bushor.2022.03.002>. [202]
- Jevinger, Å. et al. (2023), "Artificial intelligence for improving public transport: a mapping study", *Public Transport*, Vol. 16/1, pp. 99-158, <https://doi.org/10.1007/s12469-023-00334-7>. [220]
- Johnson, B., J. Coggburn and J. Llorens (2022), "Artificial Intelligence and Public Human Resource Management: Questions for Research and Practice", *Public Personnel Management*, Vol. 51/4, pp. 538-562, <https://doi.org/10.1177/00910260221126498>. [50]
- Juneja, P. (2024), *Artificial Intelligence for Electoral Management*, International Institute for Democracy and Electoral Assistance, <https://doi.org/10.31752/idea.2024.31>. [173]

- Jung, J., M. Patnam and A. Ter-Martirosyan (2018), *An Algorithmic Crystal Ball: Forecasts-based on Machine Learning*, International Monetary Fund, [5]
<https://www.imf.org/en/Publications/WP/Issues/2018/11/01/An-Algorithmic-Crystal-Ball-Forecasts-based-on-Machine-Learning-46288>.
- Katona, E. and M. Fazekas (2024), *Hidden barriers to open competition: Using text mining to uncover corrupt restrictions to competition in public procurement*, [72]
https://www.govtransparency.eu/wp-content/uploads/2024/05/Katona-Fazekas_HU_text-miningPP-corruption240417_GTIpublish240430final.pdf.
- Kelly, S., S. Kaye and O. Oviedo-Trespalacios (2023), “What factors contribute to the acceptance of artificial intelligence? A systematic review”, *Telematics and Informatics*, Vol. 77, p. 101925, <https://doi.org/10.1016/j.tele.2022.101925>. [249]
- Khalifa, M. and M. Albadawy (2024), “Artificial Intelligence for Clinical Prediction: Exploring Key Domains and Essential Functions”, *Computer Methods and Programs in Biomedicine Update*, Vol. 5, p. 100148, <https://doi.org/10.1016/j.cmpbup.2024.100148>. [214]
- Khalil, H., D. Ameen and A. Zarnegar (2022), *Tools to support the automation of systematic reviews: a scoping review*, *J Clin Epidemiology*, <https://doi.org/10.1016/j.jclinepi.2021.12.005>. [112]
- Khan, I. et al. (2022), *Mandates of the Special Rapporteur* (Ref.: OL IRL 3/2022), [281]
<https://spcommreports.ohchr.org/TMResultsBase/DownLoadPublicCommunicationFile?gId=27594>.
- Kim, S. et al. (2021), “Moderator Chatbot for Deliberative Discussion”, *Proceedings of the ACM on Human-Computer Interaction*, Vol. 5/CSCW1, pp. 1-26, <https://doi.org/10.1145/3449161>. [177]
- Klein, A. (2020), *Reducing bias in AI-based financial services*, [20]
<https://www.brookings.edu/articles/reducing-bias-in-ai-based-financial-services>.
- Kleinberg, J. et al. (2017), “Human Decisions and Machine Predictions*”, *The Quarterly Journal of Economics*, <https://doi.org/10.1093/qje/qjx032>. [326]
- Köbis, N., C. Starke and I. Rahwan (2022), “The promise and perils of using artificial intelligence to fight corruption”, *Nature Machine Intelligence*, Vol. 4/5, pp. 418-424, <https://doi.org/10.1038/s42256-022-00489-1>. [103]
- Köbis, N., C. Starke and I. Rahwan (2021), *Artificial Intelligence as an Anti-Corruption Tool (AI- ACT) -- Potentials and Pitfalls for Top-down and Bottom-up Approaches*, [95]
<https://arxiv.org/abs/2102.11567>.
- Kopponen, A. et al. (2024), “Personalised public services powered by AI: the citizen digital twin approach”, in *Research Handbook on Public Management and Artificial Intelligence*, Edward Elgar Publishing, <https://doi.org/10.4337/9781802207347.00020>. [211]
- Korea Fiscal Information Service (KFIS) (2023), *Digital Transformation in Public Finance: Korean FMIS (dBrain+)*. [9]
- Labanova, A. et al. (2022), “Capacity Building in Government: Towards Developing a Standard for a Functional Specialist in AI for Public Services”, in *Communications in Computer and Information Science, Future Data and Security Engineering. Big Data, Security and Privacy, Smart City and Industry 4.0 Applications*, Springer Nature Singapore, Singapore, [240]
https://doi.org/10.1007/978-981-19-8069-5_34.

- Lai, H. et al. (2024), "Assessing the Risk of Bias in Randomized Clinical Trials With Large Language Models", *JAMA Network Open*, Vol. 7/5, p. e2412687, [116]
<https://doi.org/10.1001/jamanetworkopen.2024.12687>.
- Landemore, H. (2022), "Can AI bring deliberative democracy to the masses?", [156]
<https://www.law.nyu.edu/sites/default/files/Helen%20Landemore%20Can%20AI%20bring%20deliberative%20democracy%20to%20the%20masses.pdf>.
- Lecorps, Y. and G. Tissandier (2022), "PAVED With Good Intentions? An Evaluation of a French Police Predictive Policing System", *SSRN Electronic Journal*, [282]
<https://doi.org/10.2139/ssrn.4314831>.
- Leeuw, F. (2003), "Reconstructing Program Theories: Methods Available and Problems to be Solved", *The American Journal of Evaluation*, Vol. 24/1, pp. 5-20, [121]
[https://doi.org/10.1016/s1098-2140\(02\)00271-0](https://doi.org/10.1016/s1098-2140(02)00271-0).
- Leyer, M. and S. Schneider (2021), "Decision augmentation and automation with artificial intelligence: Threat or opportunity for managers?", *Business Horizons*, Vol. 64/5, pp. 711-724, [196]
<https://doi.org/10.1016/j.bushor.2021.02.026>.
- Linders, D., C. Liao and C. Wang (2018), "Proactive e-Governance: Flipping the service delivery model from pull to push in Taiwan", *Government Information Quarterly*, Vol. 35/4, pp. S68-S76, [325]
<https://doi.org/10.1016/j.giq.2015.08.004>.
- López-Iturriaga, F. and I. Sanz (2017), "Predicting Public Corruption with Neural Networks: An Analysis of Spanish Provinces", *Social Indicators Research*, Vol. 140/3, pp. 975-998, [73]
<https://doi.org/10.1007/s11205-017-1802-2>.
- Lorenz, P., K. Perset and J. Berryhill (2023), "Initial policy considerations for generative artificial intelligence", *OECD Artificial Intelligence Papers*, No. 1, OECD Publishing, Paris, [190]
<https://doi.org/10.1787/fae2d1e6-en>.
- Macnish, K., D. Wright and T. Jiya (2021), "Correction to: Predictive Policing in 2025: A Scenario", in *Advanced Sciences and Technologies for Security Applications, Policing in the Era of AI and Smart Societies*, Springer International Publishing, Cham, [219]
https://doi.org/10.1007/978-3-030-50613-1_13.
- Madan, R. and M. Ashok (2023), "AI adoption and diffusion in public administration: A systematic literature review and future research agenda", *Government Information Quarterly*, Vol. 40/1, p. 101774, [194]
<https://doi.org/10.1016/j.giq.2022.101774>.
- Marcucci, S. and S. Verhulst (2025), *Reimagining the Policy Cycle in the Age of Artificial Intelligence*, Elsevier BV, [128]
<https://doi.org/10.2139/ssrn.5137557>.
- Margetts, H. (2011), "The Internet and Transparency", *The Political Quarterly*, Vol. 82/4, pp. 518-521, [185]
<https://doi.org/10.1111/j.1467-923x.2011.02253.x>.
- Margetts, H. and C. Dorobantu (2019), "Rethink government with AI", *Nature*, Vol. 568/7751, pp. 163-165, [324]
<https://doi.org/10.1038/d41586-019-01099-5>.
- Martin, P. (2023), *Policy modelling with Rules as Code (RaC)*, <https://salsa.digital/insights/policy-modelling-with-rules-as-code-rac>. [33]

- McKendrick, K. (2019), *Artificial Intelligence Prediction and Counterterrorism*, [260]
<https://www.chathamhouse.org/sites/default/files/2019-08-07-AICounterterrorism.pdf>.
- Medaglia, R., J. Gil-Garcia and T. Pardo (2021), "Artificial Intelligence in Government: Taking Stock and Moving Forward", *Social Science Computer Review*, Vol. 41/1, pp. 123-140, [237]
<https://doi.org/10.1177/08944393211034087>.
- Medaglia, R., P. Mikalef and L. Tangi (2025), *Competences and governance practices for artificial intelligence in the public sector*, European Commission: Joint Research Centre, [238]
<https://doi.org/10.2760/7895569>.
- Milanez, A., A. Lemmens and C. Ruggiu (2025), "Algorithmic management in the workplace: New evidence from an OECD employer survey", *OECD Artificial Intelligence Papers*, No. 31, OECD Publishing, Paris, [53]
<https://doi.org/10.1787/287c13c4-en>.
- Miller, S. (2020), "Causal Forest estimation of heterogeneous and time-varying environmental policy effects", *Journal of Environmental Economics and Management*, Vol. 103, [124]
<https://doi.org/10.1016/j.jeem.2020.102337>.
- Miller, S. and L. Keiser (2020), "Representative Bureaucracy and Attitudes Toward Automated Decision Making", *Journal of Public Administration Research and Theory*, Vol. 31/1, pp. 150-165, [195]
<https://doi.org/10.1093/jopart/muaa019>.
- Minozzi, S. et al. (2020), *The revised Cochrane risk of bias tool for randomized trials (RoB 2) showed low interrater reliability and challenges in its application*, Journal of Clinical Epidemiology, [115]
<https://doi.org/10.1016/j.jclinepi.2020.06.015>.
- MITRE (2023), *Searching for Solutions: MITRE Tool Simplifies Freedom of Information Act Requests*, [148]
<https://www.mitre.org/news-insights/impact-story/mitre-tool-simplifies-freedom-information-act-requests>.
- Mohun, J. and A. Roberts (2020), "Cracking the code: Rulemaking for humans and machines", [4]
OECD Working Papers on Public Governance, No. 42, OECD Publishing, Paris,
<https://doi.org/10.1787/3afe6ba5-en>.
- Mpofu, F. (2023), "The application of Artificial Intelligence in external auditing and its implications on audit quality? A review of the ongoing debates", *International Journal of Research in Business and Social Science* (2147- 4478), Vol. 12/9, pp. 496-512, [104]
<https://doi.org/10.20525/ijrbs.v12i9.2737>.
- Næss, T. et al. (2025), *Text Mining and Machine Learning in a Performance Audit of Police Handling of Cybercrime in Norway*, Routledge, [119]
<https://library.oapen.org/bitstream/handle/20.500.12657/94001/1/9781040128510.pdf>.
- National Documentation Center (EKT) (2025), , [150]
<https://www.didaktorika.gr/eadd/>.
- NCIT (2024), *New Chatbot Helps Answer IT Procurement Questions*, [70]
<https://it.nc.gov/blog/2024/04/04/new-chatbot-helps-answer-it-procurement-questions>
 (accessed on 19 February 2024).
- Nelson, D., Y. He and G. Moore (2024), "Trends and applications in wildfire burned area mapping: Remote sensing data, cloud geoprocessing platforms, and emerging algorithms", [268]
Geomatica, Vol. 76/1, p. 100008, <https://doi.org/10.1016/j.geomat.2024.100008>.

- Nesta (2024), *Innovation Policy Simulation for the Smart Economy*, [32]
<https://www.nesta.org.uk/feature/smarter-policy-through-simulation/innovation-policy-simulation-for-the-smart-economy/>.
- Netherlands Ministry of the Interior and Kingdom Relations (2019), *NL Digital Data Agenda Government*, [138]
<https://www.nldigitalgovernment.nl/wp-content/uploads/sites/11/2019/04/data-agenda-government.pdf>.
- Nikiforova, A. et al. (2023), *PPPS'2023-Proactive and Personalised Public Services: Searching for Meaningful Human Control in Algorithmic Government*, [229]
<https://ceur-ws.org/Vol-3449/paper27.pdf>.
- Odell, C. (2024), *Why we need to explore the use of AI in evidence synthesis: Reflections from the Global Evidence Summit 2024*, [117]
<https://researchforevidence.fhi360.org/need-explore-use-ai-evidence-synthesis-reflections-global-evidence-summit-2024>.
- Odilla, F. (2023), “Bots against corruption: Exploring the benefits and limitations of AI-based anti-corruption technology”, *Crime, Law and Social Change*, Vol. 80/4, pp. 353-396, [99]
<https://doi.org/10.1007/s10611-023-10091-0>.
- OECD (2025), *How Innovation Ecosystems Foster Citizen Participation Using Emerging Technologies in Portugal, Spain and the Netherlands*, OECD Public Governance Reviews, [192]
OECD Publishing, Paris, <https://doi.org/10.1787/2cb37a30-en>.
- OECD (2025), *Implementation Toolkit for the OECD Recommendation on Public Policy Evaluation*, OECD Public Governance Reviews, OECD Publishing, Paris, [107]
<https://doi.org/10.1787/77faa4fe-en>.
- OECD (2025), *OECD Regulatory Policy Outlook 2025*, OECD Publishing, Paris, [31]
<https://doi.org/10.1787/56b60e39-en>.
- OECD (2025), *Prevention of Stalking Crimes Powered by Smart CCTV Video Analysis Technology*, [265]
<https://oecd-opsi.org/innovations/prevention-of-stalking-crimes-powered-by-smart-cctv-video-analysis-technology/> (accessed on 12 March 2025).
- OECD (2025), “Tackling civic participation challenges with emerging technologies: Beyond the hype”, *OECD Public Governance Policy Papers*, No. 72, OECD Publishing, Paris, [183]
<https://doi.org/10.1787/ec2ca9a2-en>.
- OECD (2025), “Towards a common reporting framework for AI incidents”, *OECD Artificial Intelligence Papers*, No. 34, OECD Publishing, Paris, [319]
<https://doi.org/10.1787/f326d4ac-en>.
- OECD (2024), “2023 OECD Digital Government Index: Results and key findings”, *OECD Public Governance Policy Papers*, No. 44, OECD Publishing, Paris, [188]
<https://doi.org/10.1787/1a89ed5e-en>.
- OECD (2024), *AI in Health: Huge potential, huge risks*, [243]
<https://www.oecd.org/health/AI-in-health-huge-potential-huge-risks.pdf>.
- OECD (2024), *Anti-Corruption and Integrity Outlook 2024*, OECD Publishing, Paris, [93]
<https://doi.org/10.1787/968587cd-en>.

- OECD (2024), "Assessing potential future artificial intelligence risks, benefits and policy imperatives", *OECD Artificial Intelligence Papers*, No. 27, OECD Publishing, Paris, <https://doi.org/10.1787/3f4e3dfb-en>. [29]
- OECD (2024), *Enabling Digital Innovation in Government: The OECD GovTech Policy Framework*, OECD Digital Government Studies, OECD Publishing, Paris, <https://doi.org/10.1787/a51eb9b2-en>. [146]
- OECD (2024), *Facts not Fakes: Tackling Disinformation, Strengthening Information Integrity*, OECD Publishing, Paris, <https://doi.org/10.1787/d909ff7a-en>. [24]
- OECD (2024), "Financial Management Information Systems in OECD countries", *OECD Papers on Budgeting*, No. 2024/02, OECD Publishing, Paris, <https://doi.org/10.1787/ce8367cd-en>. [21]
- OECD (2024), "Fixing frictions: 'sludge audits' around the world: How governments are using behavioural science to reduce psychological burdens in public services", *OECD Public Governance Policy Papers*, No. 48, OECD Publishing, Paris, <https://doi.org/10.1787/5e9bb35c-en>. [252]
- OECD (2024), "Framework for Anticipatory Governance of Emerging Technologies", *OECD Science, Technology and Industry Policy Papers*, No. 165, OECD Publishing, Paris, <https://doi.org/10.1787/0248ead5-en>. [144]
- OECD (2024), *Global Trends in Government Innovation 2024: Fostering Human-Centred Public Services*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/c1bc19c3-en>. [201]
- OECD (2024), "Governing with Artificial Intelligence: Are governments ready?", *OECD Artificial Intelligence Papers*, No. 20, OECD Publishing, Paris, <https://doi.org/10.1787/26324bc2-en>. [60]
- OECD (2024), *Modernising Access to Social Protection: Strategies, Technologies and Data Advances in OECD Countries*, OECD Publishing, Paris, <https://doi.org/10.1787/af31746d-en>. [210]
- OECD (2024), *OECD Digital Economy Outlook 2024 (Volume 1): Embracing the Technology Frontier*, OECD Publishing, Paris, <https://doi.org/10.1787/a1689dc5-en>. [258]
- OECD (2024), *OECD Survey on Drivers of Trust in Public Institutions – 2024 Results: Building Trust in a Complex Policy Environment*, OECD Publishing, Paris, <https://doi.org/10.1787/9a20554b-en>. [46]
- OECD (2024), *Public procurement*, <https://www.oecd.org/en/topics/public-procurement.html>. [61]
- OECD (2024), *Recommendation of the Council on Artificial Intelligence*, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449> (accessed on 2 June 2024). [91]
- OECD (2024), *Reinforcing democracy initiative*, <https://www.oecd.org/en/about/programmes/reinforcing-democracy-initiative>. [142]
- OECD (2024), *Shaping smart cities of all sizes*, OECD Publishing, <https://oe.cd/cities-2024-roundtable>. [223]

- OECD (2024), *Strengthening Oversight of the Court of Auditors for Effective Public Procurement in Portugal: Digital Transformation and Data-driven Risk Assessments*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/35aeab1e-en>. [75]
- OECD (2023), “2023 OECD Open, Useful and Re-usable data (OURdata) Index: Results and key findings”, *OECD Public Governance Policy Papers*, No. 43, OECD Publishing, Paris, <https://doi.org/10.1787/a37f51c3-en>. [314]
- OECD (2023), “AI language models: Technological, socio-economic and policy considerations”, *OECD Digital Economy Papers*, No. 352, OECD Publishing, Paris, <https://doi.org/10.1787/13d38f92-en>. [182]
- OECD (2023), “Emerging privacy-enhancing technologies: Current regulatory and policy approaches”, *OECD Digital Economy Papers*, No. 351, OECD Publishing, Paris, <https://doi.org/10.1787/bf121be4-en>. [230]
- OECD (2023), *Global Trends in Government Innovation 2023*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/0655b570-en>. [2]
- OECD (2023), *Government at a Glance 2023*, OECD Publishing, Paris, <https://doi.org/10.1787/3d5c5d31-en>. [34]
- OECD (2023), *OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market*, OECD Publishing, Paris, <https://doi.org/10.1787/08785bba-en>. [54]
- OECD (2023), *Open Government for Stronger Democracies: A Global Assessment*, OECD Publishing, <https://doi.org/10.1787/5478db5b-en>. [151]
- OECD (2023), *Public Employment and Management 2023: Towards a More Flexible Public Service*, OECD Publishing, Paris, <https://doi.org/10.1787/5b378e11-en>. [51]
- OECD (2023), *Recommendation of the Council on Access to Justice and People-Centred Justice Systems*, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0498>. [288]
- OECD (2023), *Smart City Data Governance: Challenges and the Way Forward*, OECD Urban Studies, OECD Publishing, Paris, <https://doi.org/10.1787/e57ce301-en>. [224]
- OECD (2022), “2022 OECD Financial Management and Reporting Survey”, <https://data-explorer.oecd.org> (accessed on 22 April 2024). [23]
- OECD (2022), *Building Trust to Reinforce Democracy: Main Findings from the 2021 OECD Survey on Drivers of Trust in Public Institutions*, Building Trust in Public Institutions, OECD Publishing, Paris, <https://doi.org/10.1787/b407f99c-en>. [143]
- OECD (2022), *Exploring Innovation in Law Enforcement: Opportunities and critical considerations*, OECD Publishing, <https://oecd-opsi.org/blog/exploring-innovation-in-law-enforcement/>. [255]
- OECD (2022), *OECD Guidelines for Citizen Participation Processes*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/f765caf6-en>. [184]
- OECD (2022), *Supporting Health Innovation with Fair Information Practice Principles*, <https://www.oecd.org/content/dam/oecd/en/topics/policy-sub-issues/digital-health-systems/oecd-israel-health-data-governance-workshop-report.pdf>. [245]

- OECD (2022), *Tax Administration 2022: Comparative Information on OECD and other Advanced and Emerging Economies*, OECD Publishing, Paris, <https://doi.org/10.1787/1e797131-en>. [1]
- OECD (2022), *The Protection and Promotion of Civic Space: Strengthening Alignment with International Standards and Guidance*, OECD Publishing, Paris, <https://doi.org/10.1787/d234e975-en>. [152]
- OECD (2021), *Countering Public Grant Fraud in Spain: Machine Learning for Assessing Risks and Targeting Control Activities*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/0ea22484-en>. [96]
- OECD (2021), *Data-Driven, Information-Enabled Regulatory Delivery*, OECD Publishing, Paris, <https://doi.org/10.1787/8f99ec8c-en>. [25]
- OECD (2021), *Evaluation Guidelines for Representative Deliberative Processes*, OECD Publishing, Paris, <https://doi.org/10.1787/10ccbfcb-en>. [175]
- OECD (2021), *OECD Regulatory Policy Outlook 2021*, OECD Publishing, Paris, <https://doi.org/10.1787/38b0fdb1-en>. [28]
- OECD (2021), *Public Employment and Management 2021: The Future of the Public Service*, OECD Publishing, Paris, <https://doi.org/10.1787/938f0d65-en>. [49]
- OECD (2021), *Recommendation of the Council for Agile Regulatory Governance to Harness Innovation*, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0464>. [328]
- OECD (2021), *Recommendation of the Council on Enhancing Access to and Sharing of Data*, OECD Publishing, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0463>. [47]
- OECD (2020), *Artificial Intelligence & Responsible Business Conduct*, <https://mneguidelines.oecd.org/RBC-and-artificial-intelligence.pdf>. [253]
- OECD (2020), *Improving Governance with Policy Evaluation: Lessons From Country Experiences*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/89b1577d-en>. [108]
- OECD (2020), *Innovative Citizen Participation and New Democratic Institutions: Catching the Deliberative Wave*, OECD Publishing, <https://doi.org/10.1787/339306da-en>. [161]
- OECD (2020), *Public Provider versus Big Brother*, OECD Publishing, <https://trends.oecd-opsi.org/trend-reports/public-provider-versus-big-brother/>. [254]
- OECD (2020), *Shaping the Future of Regulators: The Impact of Emerging Technologies on Economic Regulators*, The Governance of Regulators, OECD Publishing, Paris, <https://doi.org/10.1787/db481aa3-en>. [36]
- OECD (2020), *Tax Administration 3.0: The Digital Transformation of Tax Administration*, OECD Publishing, Paris, <https://doi.org/10.1787/ca274cc5-en>. [3]
- OECD (2020), “Tracking and tracing COVID: Protecting privacy and data while using apps and biometrics”, *OECD Policy Responses to Coronavirus (COVID-19)*, OECD Publishing, Paris, <https://doi.org/10.1787/8f394636-en>. [277]
- OECD (2018), *OECD Regulatory Enforcement and Inspections Toolkit*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264303959-en>. [42]

- OECD (2017), *Recommendation of the Council on Health Data Governance*, [244]
<https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0433>.
- OECD (2017), *Recommendation of the Council on Open Government*, [140]
<https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0438>.
- OECD (2015), *Recommendation of the Council on Public Procurement*, [74]
<https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0411> (accessed on 2 June 2024).
- OECD (2014), *Regulatory Enforcement and Inspections*, OECD Best Practice Principles for Regulatory Policy, OECD Publishing, Paris, <https://doi.org/10.1787/9789264208117-en>. [41]
- OECD OPSI (2019), *Kol Zchut*, <https://oecd-opsi.org/innovations/kol-zchut/>. [153]
- OECD/CAF (2022), *The Strategic and Responsible Use of Artificial Intelligence in the Public Sector of Latin America and the Caribbean*, OECD Public Governance Reviews, OECD Publishing, Paris, <https://doi.org/10.1787/1f334543-en>. [297]
- OECD/EU (2024), *Reforming Regulatory Inspections in Italy at Regional and National Level 21IT12 (Final Report)*, <https://reform-support.ec.europa.eu/document/download/e2ad74a6-9bf9-4eef-8ffc-a8f2aaee3ccdc?filename=05nov2024%20Final%20Report%2021IT12%20Reforming%20inspections%20in%20Italy.pdf>. [40]
- OECD/UNESCO (2024), *G7 Toolkit for Artificial Intelligence in the Public Sector*, OECD Publishing, Paris, <https://doi.org/10.1787/421c1244-en>. [131]
- Ojo, A., S. Mellouli and F. Ahmadi Zeleti (2019), “A Realist Perspective on AI-era Public Management*”, *Proceedings of the 20th Annual International Conference on Digital Government Research*, <https://doi.org/10.1145/3325112.3325261>. [323]
- Ovadya, A. (2023), “Bridging Systems: Open Problems for Countering Destructive Divisiveness across Ranking, Recommenders, and Governance”, *Arxiv*, <https://arxiv.org/abs/2301.09976>. [158]
- Palkeet (2024), *RPA/AI in Financial Management*. [14]
- Paunov, C. and S. Planes-Satorra (2023), “Engaging citizens in innovation policy: Why, when and how?”, *OECD Science, Technology and Industry Policy Papers*, No. 149, OECD Publishing, Paris, <https://doi.org/10.1787/ba068fa6-en>. [141]
- Pawlowski, C. and H. Scholte (2023), “A taxonomy for proactive public services”, *Government Information Quarterly*, Vol. 40/1, p. 101780, <https://doi.org/10.1016/j.giq.2022.101780>. [221]
- Peixoto, T., O. Canuto and L. Jordan (2024), *AI and the Future of Government: Unexpected Effects and Critical Challenges*, <https://www.policycenter.ma/publications/ai-and-future-government-unexpected-effects-and-critical-challenges>. [178]
- Perry, A. and N. Lee (2019), *AI is coming to schools, and if we're not careful, so will its biases*, <https://www.brookings.edu/articles/ai-is-coming-to-schools-and-if-were-not-careful-so-will-its-biases/>. [232]
- Perry, W. et al. (2013), *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*, RAND Corporation, Santa Monica, CA, <https://doi.org/10.7249/RR233>. [286]

- PetaBencana.id (2021), *As Jakarta floods again, humanitarian chatbots on social media support community-led disaster response*, <https://info.petabencana.id/2021/02/22/as-jakarta-floods-again-humanitarian-chatbots-on-social-media-support-community-led-disaster-response/> (accessed on 4 March 2025). [275]
- Petersson, G. et al. (2017), *Cyber Society, Big Data, and Evaluation*, Routledge, https://www.routledge.com/Cyber-Society-Big-Data-and-Evaluation/Petersson-Breul/p/book/9781138483033?srsltid=AfmBOooqz5iNGr7Tpm5k4aiNFcJc0RAx_yZwMW42X6AiuvNZIs0Q3bi_. [109]
- Petra, B. et al. (2024), *Strategies to Counter Artificial Intelligence in Law Enforcement: Cross-Country Comparison of Citizens in Greece, Italy and Spain*, <https://arxiv.org/abs/2405.19970>. [287]
- Policy Lab Digital, Work & Society within the German Federal Ministry of Labour and Social Affairs (2024), *Guidelines for the Use of AI in the Administrative Work of Employment and Social Protection Services*, https://www.denkfabrik-bmas.de/fileadmin/Downloads/Publikationen/Guidelines_for_the_use_of_ai_in_the_administrative_work_of_employment_and_social_protection_services.pdf. [212]
- Productivity Committee (2024), *Making the Most of the AI Opportunity: Data Availability and Use*, <https://www.pc.gov.au/research/completed/making-the-most-of-the-ai-opportunity/ai-paper3-data.pdf>. [45]
- Queen Mary University of London; Centre for European Policy Studies (2021), *Criminal Justice, Fundamental Rights and the Rule of Law in the Digital Age*, <https://cdn.ceps.eu/wp-content/uploads/2021/05/Criminal-Justice-Fundamental-Rights-and-the-Rule-of-law-in-the-Digital-Age.pdf>. [310]
- Ramires Hernández, P., D. Valle-Cruz and R. Mendoza Méndez (2022), “Review on the Application of Artificial Intelligence-Based Chatbots in Public Administration”, in *Handbook of Research on Applied Artificial Intelligence and Robotics for Government Processes, Advances in Computational Intelligence and Robotics*, IGI Global, <https://doi.org/10.4018/978-1-6684-5624-8.ch007>. [206]
- Rane, N., S. Choudhary and J. Rane (2024), “Artificial intelligence for enhancing resilience”, *Journal of Applied Artificial Intelligence*, Vol. 5/2, pp. 1-33, <https://doi.org/10.48185/jaai.v5i2.1053>. [248]
- Redden, J. et al. (2022), *Automating Public Services: Learning from Cancelled Systems*, Carnegie UK, <https://carnegieuk.org/publication/automating-public-services-learning-from-cancelled-systems/>. [228]
- Redden, J. and M. O'Donovan Dix (2020), *Artificial intelligence in the criminal justice system: Demystifying artificial intelligence, its applications, and potential risks*, <https://www.ojp.gov/ncjrs/virtual-library/abstracts/artificial-intelligence-criminal-justice-system-demystifying>. [289]
- Reiling, D. (2020), *Courts and artificial intelligence*, <https://doi.org/10.36745/ijca.343>. [291]
- Richardson, A., T. van Florenstein Mulder and T. Vehbi (2019), *Nowcasting GDP using machine learning algorithms: A real-time assessment*, <https://www.rbnz.govt.nz-/media/c0bb0238351244c986e33550b36bd2d4.ashx>. [6]

- Rinaldi, F., G. Giuffrida and T. Negrete (2017), *Real-Time Monitoring and Evaluation—Emerging News as Predictive Process Using Big Data-Based Approach*, Routledge. [110]
- Rivero del Paso, L. et al. (2023), *Digital Solutions Guidelines for Public Financial Management*, <https://www.imf.org/en/Publications/TNM/Issues/2023/10/06/Digital-Solutions-Guidelines-for-Public-Financial-Management-537781>. [22]
- Romberg, J. and T. Escher (2024), “Making Sense of Citizens’ Input through Artificial Intelligence: A Review of Methods for Computational Text Analysis to Support the Evaluation of Contributions in Public Participation”, *Digital Government: Research and Practice*, pp. 1-30, <https://doi.org/10.1145/3603254>. [179]
- Salazar, A., J. Pérez and J. Gallego (2024), “VigIA: prioritizing public procurement oversight with machine learning models and risk indices”, *Data & Policy*, Vol. 6, <https://doi.org/10.1017/dap.2024.83>. [100]
- Santiso, C. (2022), “Govtech against corruption: What are the integrity dividends of government digitalization?”, *Data & Policy*, Vol. 4, <https://doi.org/10.1017/dap.2022.31>. [76]
- Schneider, B. and N. Sanders (2023), *AI could shore up democracy - here’s one way*, <https://theconversation.com/ai-could-shore-up-democracy-heres-one-way-207278> (accessed on 23 May 2024). [166]
- SDAIA (2024), *SDAIA Showcases National Smart City Platform and National Algorithms for Smart Cities at Barcelona Expo*, <https://sdaia.gov.sa/en/MediaCenter/News/Pages/NewsDetails.aspx?NewsID=228>. [205]
- Seto, J. (2018), *Using AI and IoT for disaster management*, <https://azure.microsoft.com/en-us/blog/using-ai-and-iot-for-disaster-management/> (accessed on 3 March 2025). [270]
- Shark, A. (2025), *What the Rising Costs of AI Means for Government*, <https://statetechmagazine.com/article/2025/01/what-rising-costs-ai-means-government>. [235]
- Shark, A. (2024), *Artificial Intelligence: The Impact on Public Procurement - Opportunity & Risk*, <https://www.nigp.org/blog/ai-and-public-procurement>. [84]
- Slovenia Ministry of Justice (2024), *Information request to Maja Velic, advisor of the Ministry of Justice of Slovenia*. [295]
- Soman, D. (2024), *What Works, What Doesn’t (and When): Case Studies in Applied Behavioral Science*, <https://utppublishing.com/doi/book/10.3138/9781487548735>. [227]
- Sommer, J. (2021), *La Cour de cassation à l’épreuve du numérique et de l’intelligence artificielle*, <https://www.vie-publique.fr/parole-dexpert/278415-la-cour-de-cassation-face-au-numerique-et-l-intelligence-artificielle> (accessed on 19 February 2025). [302]
- Song, C. and J. Lee (2015), “Citizens’ Use of Social Media in Government, Perceived Transparency, and Trust in Government”, *Public Performance & Management Review*, Vol. 39/2, pp. 430-453, <https://doi.org/10.1080/15309576.2015.1108798>. [186]
- Spain Ministry of Justice (2024), *Política de uso de la IA en la Administración de Justicia*, <https://www.administraciondejusticia.gob.es/w/approbadala-politica-de-uso-de-la-ia-en-la-administracion-de-justicia>. [313]

- Spain Ministry of Justice (2022), *Digital transformation within the administration of justice*, [300] [https://www.mjusticia.gob.es/es/JusticiaEspana/ProyectosTransformacionJusticia/Documents/202209%20Digital%20Transformation%20\[ENG\].pdf](https://www.mjusticia.gob.es/es/JusticiaEspana/ProyectosTransformacionJusticia/Documents/202209%20Digital%20Transformation%20[ENG].pdf) (accessed on 19 February 2025).
- Statistics Denmark (2025), *Microdata schemes in Statistics Denmark*, [137] <https://www.dst.dk/en/TilSalg/Forskningservice/Dataadgang/Projektoprettelse>.
- Stettinger, G., P. Weissensteiner and S. Khastgir (2024), “Trustworthiness Assurance Assessment for High-Risk AI-Based Systems”, *IEEE Access*, Vol. 12, pp. 22718-22745, [276] <https://doi.org/10.1109/access.2024.3364387>.
- STF (2024), *STF lança MARIA, ferramenta de inteligência artificial que dará mais agilidade aos serviços do Tribunal*, [298] <https://noticias.stf.jus.br/postsnoticias/stf-lanca-maria-ferramenta-de-inteligencia-artificial-que-dara-mais-agilidade-aos-servicos-do-tribunal/>.
- Straub, V. et al. (2024), *AI for bureaucratic productivity: Measuring the potential of AI to help automate 143 million UK government transactions*, [199] https://www.turing.ac.uk/sites/default/files/2024-03/ai_for_bureaucratic_productivity.pdf.
- Sumner, J. et al. (2023), “Developing an Artificial Intelligence-Driven Nudge Intervention to Improve Medication Adherence: A Human-Centred Design Approach”, *Journal of Medical Systems*, Vol. 48/1, [241] <https://doi.org/10.1007/s10916-023-02024-0>.
- Sunass (2024), *Use of big data and AI in water regulation: Experience from Sunass, Peru*, [38] <https://cdn.www.gob.pe/uploads/document/file/6872996/5943561-presentacion-en-la-red-de-reguladores-economicos-o-network-economic-regulators-ner.pdf>.
- Sun, T. and R. Medaglia (2019), “Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare”, *Government Information Quarterly*, Vol. 36/2, [197] pp. 368-383, <https://doi.org/10.1016/j.giq.2018.09.008>.
- Supreme Court of Brazil (2024), *Information request to Victor Durigan, advisor to the Supreme Court*. [293]
- Taşdöken, Ö. (2024), “Use of Artificial Intelligence and Audit Analytics in Internal Audit Processes in the Public Sector”, *EDPACS*, Vol. 69/9, pp. 1-15, [94] <https://doi.org/10.1080/07366981.2024.2376790>.
- TBI (2024), *The Economic Case for Reimagining the State*, Tony Blair Institute for Global Change, [218] <https://institute.global/insights/economic-prosperity/the-economic-case-for-reimagining-the-state>.
- Tejwani, R. et al. (2020), *Migratable AI: Effect of identity and information migration on users perception of conversational AI agents*, [209] <https://ieeexplore.ieee.org/document/9223436>.
- Tenorio, J. and W. Perez (2023), *GDP nowcasting with Machine Learning and Unstructured Data to Peru*, [8] <https://perueconomics.org/wp-content/uploads/2023/11/WP-197.pdf>.
- The Alan Turing Institute (2023), *AI Skills for Business Competency Framework*, [132] <https://www.turing.ac.uk/skills/collaborate/ai-skills-business-framework>.
- The Computational Democracy Project (2024), *The Computational Democracy Project*, [162] <https://compdemocracy.org/>.

- The World Bank (2023), *OUTPUT 14 - Report on the proposed tool for screening of legislation*, [68] <https://documents1.worldbank.org/curated/en/099011924055735846/pdf/P169141129e0ad0f618892137a09a4eb9f9.pdf>.
- Thiebes, S., S. Lins and A. Sunyaev (2020), "Trustworthy artificial intelligence", *Electronic Markets*, Vol. 31/2, pp. 447-464, [234] <https://doi.org/10.1007/s12525-020-00441-4>.
- Thomas, J. et al. (2024), *Responsible Alin Evidence SynthEsis (RAISE): guidance and recommendations*, [113] <https://osf.io/fwaud/>.
- Tilon, S. et al. (2020), "Post-Disaster Building Damage Detection from Earth Observation Imagery Using Unsupervised and Transferable Anomaly Detecting Generative Adversarial Networks", *Remote Sensing*, Vol. 12/24, p. 4193, [267] <https://doi.org/10.3390/rs12244193>.
- Trajkovski, G. (2024), "Bridging the public administration-AI divide: A skills perspective", *Public Administration and Development*, Vol. 44/5, pp. 412-426, [239] <https://doi.org/10.1002/pad.2061>.
- Tribune News Service (2023), *State Legislature Adopts Resolution on AI ... Drafted by AI*, [26] <https://insider.govtech.com/california/news/state-legislature-adopts-resolution-on-ai-drafted-by-ai>.
- Tsai, L. et al. (2024), "Generative AI for Pro-Democracy Platforms", *An MIT Exploration of Generative AI*, [160] <https://doi.org/10.21428/e4baedd9.5aaaf489a>.
- U4 (2025), *Unlocking AI's potential in anti-corruption: Hype vs. reality*, [106] <https://www.u4.no/blog/unlocking-ai-s-potential-in-anti-corruption-hype-vs-reality>.
- UC Berkeley (2021), *Positive AI Economic Futures*, World Economic Forum, [56] <https://www.weforum.org/reports/positive-ai-economic-futures>.
- Ugale, G. and C. Hall (2024), "Generative AI for anti-corruption and integrity in government: Taking stock of promise, perils and practice", *OECD Artificial Intelligence Papers*, No. 12, OECD Publishing, Paris, [89] <https://doi.org/10.1787/657a185a-en>.
- UK Government Office for Science (2023), *Future Risks of Frontier AI*, [145] <https://assets.publishing.service.gov.uk/media/653bc393d10f3500139a6ac5/future-risks-of-frontier-ai-annex-a.pdf>.
- UK National Audit Office (2023), *Report on Accounts - Department for Work & Pensions*, [98] <https://www.nao.org.uk/wp-content/uploads/2023/07/dwp-report-on-accounts-2022-23.pdf>.
- UNESCO (2025), *Artificial Intelligence and the Rule of Law*, [317] <https://www.unesco.org/en/artificial-intelligence/rule-law>.
- UNESCO (2019), *Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development*, [231] <https://repositorio.minedu.gob.pe/bitstream/handle/20.500.12799/6533/Artificial%20intelligence%20in%20education%20challenges%20and%20opportunities%20for%20sustainable%20development.pdf>.
- United Kingdom (2023), *Artificial Intelligence (AI) – Judicial Guidance*, [315] <https://www.judiciary.uk/guidance-and-resources/artificial-intelligence-ai-judicial-guidance/>.
- University of Turku (2023), *TurkuNLP*, [181] <https://turkunlp.org/> (accessed on 23 May 2024).

- UNU Institute for Environment and Human Security (2024), *5 Ways AI Can Strengthen Early Warning Systems*, <https://unu.edu/ehs/series/5-ways-ai-can-strengthen-early-warning-systems> (accessed on 4 March 2025). [273]
- US DHS (2024), *DHS Launches First-of-its-Kind Initiative to Hire 50 Artificial Intelligence Experts in 2024*, <https://www.dhs.gov/archive/news/2024/02/06/dhs-launches-first-its-kind-initiative-hire-50-artificial-intelligence-experts-2024>. [263]
- US DHS (2024), *United States Customs and Border Protection – AI Use Cases*, <https://www.dhs.gov/ai/use-case-inventory/cbp#deployed> (accessed on 3 March 2025). [262]
- US FOIA (2023), *FOIA Reference Model White Paper*, FOIA.gov, <https://www.foia.gov/chief-foia-officers-council/foia-reference-model-white-paper-april-19-2023>. [147]
- US GAO (2024), *Facial Recognition Technology: Federal Law Enforcement Agency Efforts Related to Civil Rights and Training*, United States Government Accountability Office, <https://www.gao.gov/products/gao-24-107372>. [257]
- Valderrama, M., M. Hermosilla and R. Garrido (2023), *State of the Evidence: Algorithmic Transparency*, <https://www.opengovpartnership.org/documents/state-of-the-evidence-algorithmic-transparency/>. [44]
- van Baal, S. (2024), “Testing behaviour change with an artificial intelligence chatbot in a randomized controlled study.”, *Journal Public Health*, Vol. 45, pp. 506-522, <https://doi.org/10.1057/s41271-024-00500-6>. [208]
- van Noordt, C. and G. Misuraca (2019), “New Wine in Old Bottles: Chatbots in Government”, in *Lecture Notes in Computer Science, Electronic Participation*, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-030-27397-2_5. [170]
- van Noordt, C. and L. Tangi (2023), “The dynamics of AI capability and its influence on public value creation of AI within public administration”, *Government Information Quarterly*, Vol. 40/4, p. 101860, <https://doi.org/10.1016/j.giq.2023.101860>. [15]
- Vogl, T. et al. (2019), “Algorithmic Bureaucracy”, *Proceedings of the 20th Annual International Conference on Digital Government Research*, <https://doi.org/10.1145/3325112.3325240>. [246]
- Walprecht, S. and C. Lewerenz (2024), *Facilitating Regulatory Impact Assessments: The Benefits of Machine Learning in Legislation*, https://www.destatis.de/EN/About-Us/Events/Machine-Learning/Slides/s2_walprecht.pdf. [30]
- Wastupranata, L., S. Kong and L. Wang (2024), “Deep Learning for Abnormal Human Behavior Detection in Surveillance Videos—A Survey”, *Electronics*, Vol. 13/13, p. 2579, <https://doi.org/10.3390/electronics13132579>. [261]
- Welker, Y. (2023), *Generative AI holds great potential for those with disabilities - but it needs policy to shape it*, <https://www.weforum.org/agenda/2023/11/generative-ai-holds-potential-disabilities/> (accessed on 24 May 2024). [154]
- Weyl, E., A. Tang and C. Plurality (2022), *Augmented Deliberation*, <https://www.plurality.net/v/chapters/5-4/eng>. [157]

- Wisesa, A. (2023), "Inovasi artificial intelligence sebagai financial advisor kementerian keuangan", *Swatantra*, Vol. XX, pp. 1-8, <https://jurnal.umj.ac.id/index.php/SWATANTRA/article/viewFile/15611/8230>. [11]
- World Justice Project (2021), *Grasping the Justice Gap: Opportunities and Challenges for People-Centered Justice Data*, <https://worldjusticeproject.org/news/grasping-justice-gap-opportunities-and-challenges-people-centered-justice-data>. [311]
- Xavier, O. et al. (2022), "Tax evasion identification using open data and artificial intelligence", *Revista de Administração Pública*, Vol. 56/3, pp. 426-440, <https://doi.org/10.1590/0034-761220210256x>. [322]
- Yao, M. (ed.) (2025), "Tailoring generative AI chatbots for multiethnic communities in disaster preparedness communication: extending the CASA paradigm", *Journal of Computer-Mediated Communication*, Vol. 30/1, <https://doi.org/10.1093/jcmc/zmae022>. [274]
- Zhang, G., H. Atasoy and M. Vasarhelyi (2022), "Continuous monitoring with machine learning and interactive data visualization: An application to a healthcare payroll process", *International Journal of Accounting Information Systems*, Vol. 46, p. 100570, <https://doi.org/10.1016/j.accinf.2022.100570>. [321]
- Zilka, M., H. Sargeant and A. Weller (2022), "Transparency, Governance and Regulation of Algorithmic Tools Deployed in the Criminal Justice System: a UK Case Study", *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, <https://doi.org/10.1145/3514094.3534200>. [320]
- Zinnbauer, D. (2025), *Un-Plateauing Corruption Research? Perhaps less necessary, but more exciting than one might think*, Elsevier BV, <https://doi.org/10.2139/ssrn.5094531>. [105]
- Ziyadin, S. et al. (eds.) (2020), "Experience of Artificial Intelligence Implementation in Japan", *E3S Web of Conferences*, Vol. 159, p. 04035, <https://doi.org/10.1051/e3sconf/202015904035>. [264]

Notas

¹ https://www.aade.gr/sites/default/files/2024-12/dt_11.12.24.pdf.

² <https://www.tovima.gr/2023/05/29/finance/aade-pos-i-texniti-noimosyni-apokalyptei-tis-adilotes-pisines>.

³ <https://www.aade.gr/sites/default/files/2024-12>.

⁴ <https://www.oecd.org/en/data/datasets/inventory-of-tax-technology-initiatives.html>

⁵ <https://www.mononews.gr/oikonomia/aade-sta-skaria-chatbox-gia-tis-forologikes-diloseis>.

⁶ Reglas como código (RaC, por sus siglas en inglés) es la práctica de traducir las normas y reglamentos jurídicos a formatos legibles por máquina (Mohun and Roberts, 2020^[4]). Este enfoque permite la

verificación y aplicación automatizadas del cumplimiento y apoya la toma de decisiones basada en marcos jurídicos predefinidos. Al aplicar la IA en la administración tributaria, la RaC puede ayudar a garantizar que la toma de decisiones apoyada por la IA sea transparente, coherente en todos los casos y que se ajuste a la ley. La RaC puede contribuir así a aumentar la confianza pública y la rendición de cuentas en la aplicación de la IA.

⁷ El pronóstico macrofiscal es el proceso de predicción de indicadores económicos y financieros gubernamentales clave, que incluyen el crecimiento del PIB, la inflación, los ingresos tributarios, el gasto público y los niveles de deuda. Estos pronósticos ayudan a los gobiernos a planificar los presupuestos, establecer la política fiscal y administrar las finanzas públicas durante períodos de varios años.

⁸ <https://ask.u.ae/en>.

⁹ Basado en debates de la OCDE y las reuniones de los grupos de trabajo con los funcionarios de la GFP.

¹⁰ La Recomendación de la OCDE para una Gobernanza Regulatoria Ágil (2021^[328]) tiene como objetivo ayudar a los gobiernos a desarrollar e implementar enfoques regulatorios ágiles y resilientes y facilitar la cooperación institucional en respuesta y para estimular la innovación. Se destaca la necesidad de ajustar las herramientas de gestión regulatoria y permitir una mayor experimentación para fomentar el aprendizaje continuo y la adaptación.

¹¹ La 22^a reunión de la Red de Reguladores Económicos de la OCDE se celebró en París en abril de 2024.

¹² <https://www.oecd.org/en/networks/network-of-economic-regulators.html>.

¹³ A menos que se indique lo contrario, los detalles se han obtenido a partir de la colaboración de la OCDE y del trabajo con los reguladores nacionales.

¹⁴ Información proporcionada por el Gobierno de Israel a la OCDE.

¹⁵ Reunión interna de la Red de Reguladores Económicos (2024).

¹⁶ <https://www.singaporetech.edu.sg/news/impressai-revolutionises-talent-acquisition-government-agencies>.

¹⁷ Véase <https://www.mynewsdesk.com/tengai-interview-robot/pressreleases/swedish-municipality-first-in-the-world-to-use-social-interview-robot-tengai-in-recruitment-2878357>.

¹⁸ <https://hbr.org/2022/11/how-walmart-automated-supplier-negotiations>.

¹⁹ <https://autogenai.com>.

²⁰ <https://sievo.com>.

²¹ <https://www.docusign.com>.

²² Véase también la sección de este capítulo sobre el uso de la "IA en la contratación pública".

²³ Entrevista con el Tribunal de Cuentas neerlandés.

²⁴ Información facilitada por funcionarios del TCE.

²⁵ Véase <https://www.oecd.org/en/about/projects/tech-connect-for-integrity.html>.

²⁶ <https://kudos.dfo.no>.

²⁷ <https://opendata.dc.gov/pages/compass>.

²⁸ <https://opencouncil.gr>.

²⁹ <https://www.mapletestimony.org>.

³⁰ <https://ai.gov.uk/consultations>.

³¹ <https://www.govocal.com>.

³² <https://panoramic.make.org/customer/cece/event/convention-citoyenne-sur-la-fin-de-vie>.

³³ Véase <https://oe.cd/immersive-brief> para consultar el documento informativo de la OCDE sobre tecnologías inmersivas.

³⁴ Las oportunidades clave para la comunicación pública identificadas en este párrafo son los resultados de las entrevistas realizadas por la OCDE en 2023-2024 con miembros de la Red de Comunicación Pública de la OCDE.

³⁵ <https://oecd-opsi.org/innovations/use-of-robotic-process-automation-in-forensic-routines/>

³⁶ <https://oecd-opsi.org/innovations/streamlined-benefits-for-a-more-equal-society/>

³⁷ Véase https://cms.hacienda.cl/cmod_cl/assets/documento/descargar/3a29739655166/1712674476.

³⁸ Véase <https://ai.gov.uk/projects/caddy>.

³⁹ <https://www.tech.gov.sg/gtpc-product-1>.

⁴⁰ <https://oecd-opsi.org/innovations/no-show-prediction-at-outpatient-clinic>.

⁴¹ Los servicios públicos proactivos son servicios que una organización pública "impulsa" hacia sus homólogos en función de sus "necesidades, circunstancias, preferencias personales, eventos vitales y ubicación". (Linders, Liao and Wang, 2018^[325]).

⁴² Los casos prácticos de IA se han centrado en áreas de alto riesgo y ricas en datos, como la seguridad (por ejemplo, detección de riesgos, predicción y simulación) (Margetts and Dorobantu, 2019^[324]; Xavier et al., 2022^[322]; Zhang, Atasoy and Vasarhelyi, 2022^[321]), reconocimiento facial en la policía (Zilka, Sargeant and Weller, 2022^[320]), reincidencia, predicción en justicia penal (Kleinberg et al., 2017^[326]), automatización de procesos (Ojo, Mellouli and Ahmadi Zeleti, 2019^[323]) y la previsión de las necesidades futuras de servicios (Bright et al., 2019^[327]).

⁴³ Véase <https://www.oecd.org/en/topics/privacy-enhancing-technologies.html>.

⁴⁴ <https://u.ae/en/about-the-uae/digital-uae/digital-technology/artificial-intelligence/rules-as-code>.

⁴⁵ En primer lugar, los servicios públicos pueden dividirse en generales (por ejemplo, semáforos) y específicos (solicitados individualmente y dirigidos específicamente).

⁴⁶ Véase <https://www.youtube.com/watch?v=JzLhcQijsz4>,
https://epiteliki.civilprotection.gov.gr/sites/default/files/2024-01/%CE%94%CE%B9%CE%B1%CE%BA%CE%AE%CF%81%CF%85%CE%BE%CE%B7%CE%A0%CF%85%CF%81%CE%B1%CE%BD%CE%AF%CF%87%CE%BD%CE%B5%CF%85%CF%83%CE%B7_24PROC014095623_03.01.2024.pdf y <https://civilprotection.gov.gr/en/deltia-tupou/b-kikiliias-i-tehniti-noimosyni-stin-ypiresia-tis-asfaleias-ton-politon> (en griego).

⁴⁷ <https://www.ktpae.gr/erga/anavathmisi-kai-epektasi-upodomon-tpe-tomea-dikaiosinis>.

⁴⁸ Información facilitada por funcionarios del Gobierno de Grecia.

⁴⁹ Información proporcionada a la OCDE por funcionarios del Ministerio de Presidencia, Justicia y Relaciones con las Cortes.

⁵⁰ Las medidas de protección se refieren a las acciones o mecanismos jurídicos establecidos para salvaguardar los derechos, la seguridad o los intereses de las personas o los grupos, a menudo en el contexto de partes vulnerables o durante los procedimientos judiciales.

⁵¹ Información proporcionada por funcionarios del Gobierno de Israel a la OCDE.

⁵² https://ministryofjustice.gr/?page_id=7483.

Gobernar con la inteligencia artificial

Panorama actual y hoja de ruta en las funciones centrales de gobierno

La inteligencia artificial (IA) es una de las fuerzas más transformadoras del siglo XXI y se está convirtiendo en una parte integral de la transformación digital de los gobiernos y sus administraciones públicas. El uso de la IA por parte de los gobiernos puede generar impactos tangibles al agilizar procesos internos y ofrecer servicios públicos automatizados y personalizados; fomentar una mejor toma de decisiones y capacidad de pronóstico; fortalecer la detección de riesgos de fraude; y elevar la calidad del empleo público y el aprendizaje de los servidores públicos. Este informe explora 200 casos de uso concretos sobre cómo los gobiernos están utilizando la IA en 11 funciones gubernamentales centrales, desde la prestación de servicios públicos y la administración de justicia hasta la lucha contra la corrupción, la gestión de las finanzas públicas y la reforma de la función pública. Así mismo, destaca las oportunidades y riesgos únicos que conlleva la IA en las administraciones públicas, indaga en los desafíos que enfrentan los gobiernos al adoptar estas tecnologías y ofrece información sobre los habilitadores, salvaguardas y estrategias de participación necesarios para garantizar que la IA se utilice de manera confiable y efectiva.



IMPRESA ISBN 978-92-64-55690-4
PDF ISBN 978-92-64-72718-2



9 789264 556904