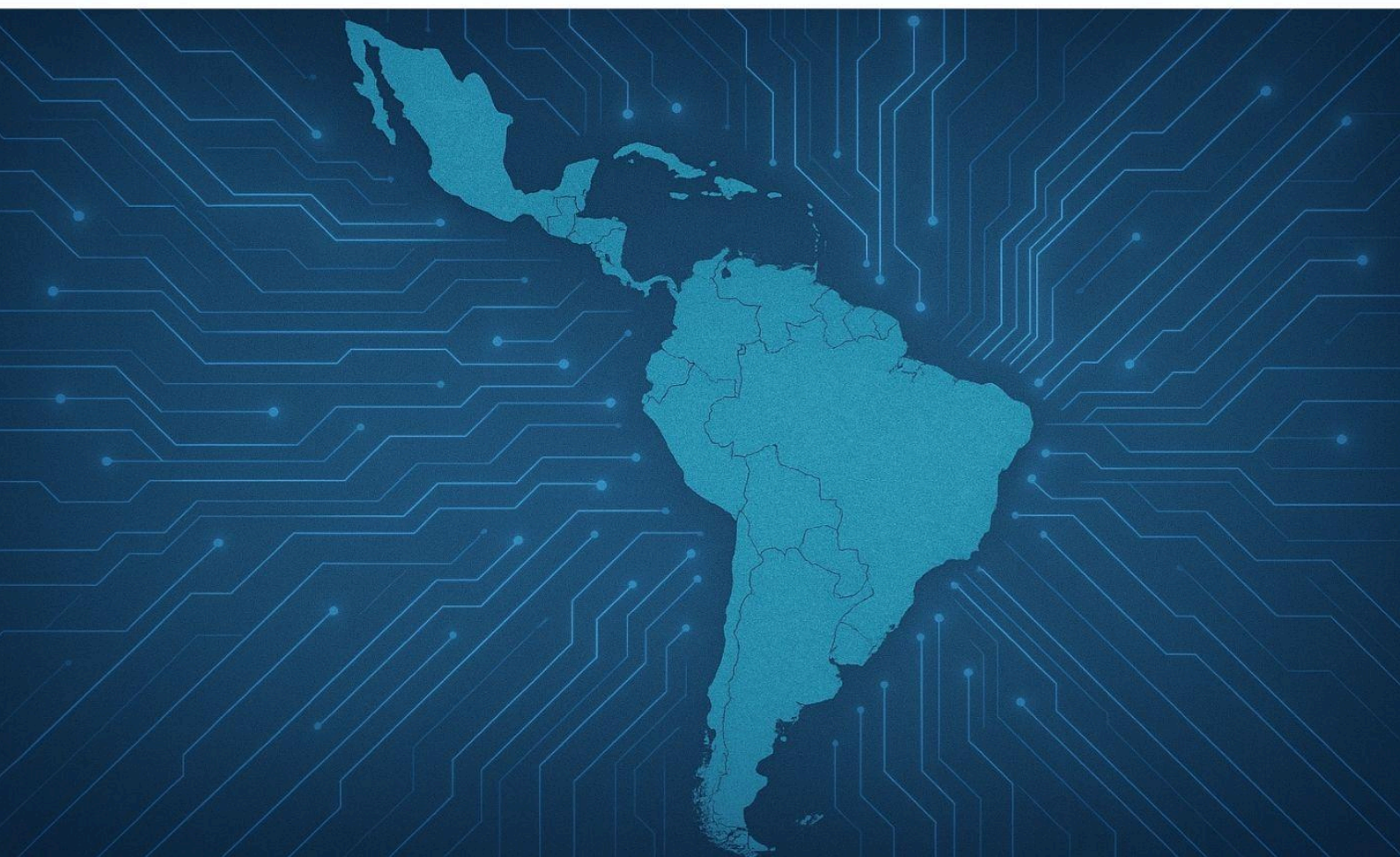


Regulación de la Compra Pública de Inteligencia Artificial

Estándares para la Contratación
Pública de Sistemas de IA Confiables,
Éticos y Centrados en las Personas



I. Resumen ejecutivo.....	3
II. Panorama actual de la contratación de sistemas de IA en América Latina y el Caribe.....	5
1. Aplicaciones de IA en LATAM y el Caribe.....	5
2. Estudio de casos relevantes: aplicaciones de alto impacto de IA en el Sector Público regional.....	12
3. Conclusiones preliminares acerca del panorama actual de la contratación de IA en la región.....	21
III. La relevancia de contar con instrumentos para la gobernanza ética de la contratación de sistemas de IA para el Sector Público.....	23
IV. Principios éticos aplicables a la compra pública de sistemas de IA.....	29
1. Proporcionalidad e inocuidad.....	29
2. Seguridad y robustez técnica.....	30
3. Equidad, no discriminación y diversidad.....	34
4. Sostenibilidad y bienestar.....	37
5. Privacidad y protección de datos personales.....	39
6. Supervisión y decisión humana.....	40
7. Transparencia y explicabilidad.....	42
8. Responsabilidad y rendición de cuentas.....	44
9. Concientización y alfabetización.....	48
10. Colaboración de múltiples partes interesadas.....	51
V. Gestión de tensiones entre principios.....	53
VI. Guía de estándares orientados a la Contratación Pública de sistemas de IA confiables, éticos y centrados en las personas.....	55
VII. Conclusiones y recomendaciones finales.....	59
Anexo I - Modelo de resolución de aprobación de estándares para la contratación de sistemas de IA confiables, éticos y centrados en las personas.....	66

I. Resumen ejecutivo

La Inteligencia Artificial (IA) se presenta como una herramienta estratégica con alto potencial para transformar el sector público. La posibilidad que brinda para automatizar diversas clases de procesos permite mejorar la eficiencia operativa, optimizar la toma de decisiones y ofrecer servicios más personalizados a la ciudadanía.

La IA puede aplicarse en múltiples áreas de la Administración Pública, desde la salud, la educación, la seguridad y la gestión de beneficios sociales, a la atención de la ciudadanía, entre muchos otros. Así, por ejemplo, permite automatizar la redacción de documentos administrativos, facilitar el acceso ciudadano a la información pública mediante asistentes conversacionales inteligentes, automatizar la creación de borradores de documentos, generar resúmenes de expedientes, normativa extensa o informes técnicos, y apoyar a las personas en la toma de decisiones, entre muchos otros casos de uso.

Actualmente, también es posible incorporar agentes de IA que operen de manera autónoma para supervisar procesos administrativos en tiempo real, detectando y notificando de forma automática vencimientos, errores o desvíos en la ejecución de tareas, entre otros posibles casos de uso, lo que contribuye a una gestión más eficiente, transparente y segura¹.

En cualquier caso, la incorporación de sistemas de IA en el Sector Público siempre debe realizarse de manera ética, responsable y en beneficio de la sociedad, garantizando la transparencia, el respeto de los derechos fundamentales y los principios democráticos.

Para lograrlo, resulta indispensable establecer marcos de gobernanza con estándares éticos claros, que orienten los procesos de adquisición de sistemas de IA en la Administración Pública, tanto para los organismos públicos que los contratan como para los proveedores que los desarrollan e implementan.

¹ Corvalán, J. G., & Sánchez Caparrós, M. (2025). Agentes de Inteligencia Artificial y Workflows Agénticos. La nueva frontera de la automatización. p. 40.

Este documento procura ofrecer un panorama actualizado sobre la contratación de sistemas de IA en América Latina y el Caribe, identificando tendencias, desafíos y oportunidades para una adopción responsable en el Sector Público.

A partir de este diagnóstico inicial, se aborda la necesidad de fortalecer la gobernanza ética en los procesos de adquisición de sistemas de IA, mediante la construcción de herramientas de contratación alineadas con marcos internacionales de referencia, como la Recomendación de la UNESCO sobre la Ética de la IA y los Principios de la OCDE².

En este punto, el desafío viene dado por la diversidad de sistemas de IA disponibles y la inexistencia de un método único para su contratación. La manera en que se contratan los sistemas de IA dependerá de la situación que enfrente el área de compras de la organización, e involucra diversos factores a considerar como, por ejemplo, si se contrata un sistema de IA a medida o si se contrata una licencia de uso de una aplicación de IA generativa bajo cláusulas generales predefinidas por el oferente; así como el sector de aplicación del sistema de IA a contratar y el impacto que tiene en los derechos de las personas.

Considerando estas particularidades, se proponen: **un modelo de pliego de bases y condiciones generales aplicable a la compra pública de sistemas de IA**, así como un **conjunto de estándares contractuales de base que pueden ser incorporados a los pliegos de bases y condiciones particulares, como a los contratos resultantes de los procedimientos de selección empleados por el organismo público.**

Estos estándares contractuales se estructuran bajo una lógica adaptativa y proporcional al nivel de impacto del sistema de IA a contratar, así como a la diversidad de modalidades de adquisición que pueden presentarse en la práctica. No es lo mismo contratar el desarrollo a medida de un sistema de IA, que adquirir licencias para utilizar modelos de lenguaje de propósito general o soluciones de software que incorporan

² UNESCO. (2021). Recomendación sobre la ética de la inteligencia artificial. Consultado el 22 de abril de 2025, en <https://www.unesco.org/es/articles/recomendacion-sobre-la-etica-de-la-inteligencia-artificial>
OCDE. (2024). Recommendation of the Council on Artificial Intelligence. Consultado el 22 de abril de 2025 en <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449>

funcionalidades potenciadas por IA. Tampoco es indiferente el sector de aplicación de la IA (ej. defensa, Justicia o atención ciudadana y turismo).

Atendiendo a esas diferencias, las herramientas aquí se proponen se han diseñado bajo un enfoque flexible que permite garantizar la ética, la transparencia y los derechos de las personas, sin obstaculizar la innovación ni complejizar innecesariamente los procedimientos de contratación ni la futura relación con el proveedor.

II. Panorama actual de la contratación de sistemas de IA en América Latina y el Caribe

A partir del relevamiento realizado y en base a investigaciones previas, hemos identificado un repositorio elaborado recientemente por la Universidad de los Andes³, titulado “Sistemas de IA en América Latina”, que sistematiza la incorporación de soluciones de inteligencia artificial (IA) en diversos ámbitos institucionales de los países de América Latina y el Caribe.

Este registro abarca aproximadamente 750 desarrollos distribuidos entre los poderes judiciales, legislativos, ejecutivos, y diversos órganos autónomos de países que conforman la región.

Dado que el presente trabajo se focaliza en las contrataciones de sistemas de IA en el ámbito de la Administración Pública, el análisis del repositorio se circunscribe a las implementaciones efectuadas por los poderes ejecutivos de los distintos países, siempre que las mismas se encuentren en etapa de ejecución, es decir, ya en funcionamiento y al servicio de las personas.

1. Aplicaciones de IA en LATAM y el Caribe

El repositorio de la Universidad de los Andes registra aproximadamente 350 casos de implementaciones de IA en el ámbito del Poder Ejecutivo. De allí, se identifican las

³ Universidad de los Andes. Sistemas de IA en América Latina. Consultado el 5 de mayo de 2025, de <https://sistemaspublicos.tech/>

aplicaciones más comunes en diversos países, clasificadas en las siguientes nueve categorías:

- **Chatbots**

Esta categoría abarca aproximadamente el 20.6 % de los casos analizados (73 en total)⁴. Las soluciones tecnológicas de este grupo comparten características comunes en cuanto a sus objetivos, ámbitos de aplicación y beneficios esperados.

Estas herramientas se diseñan para ofrecer asistencia automatizada a la ciudadanía, priorizando el acceso a información pública, la resolución de consultas frecuentes y la gestión remota de trámites administrativos.

El relevamiento realizado permite concluir que estas soluciones de IA constituyen una herramienta ampliamente adoptada y consolidada en el ecosistema de innovación pública regional y que su implementación obedece a criterios de eficiencia operativa, cercanía con la ciudadanía y mejora de la experiencia usuaria, con un enfoque marcadamente pragmático y centrado en la resolución de necesidades concretas, más que en el desarrollo de capacidades conversacionales sofisticadas.

- **Reconocimiento facial**

Esta categoría representa aproximadamente el 19.6 % del total de casos analizados (70 en concreto) y evidencia una clara orientación hacia su utilización en tareas de seguridad pública y vigilancia ciudadana.

Su funcionalidad principal consiste en la identificación automática de personas a partir de imágenes captadas en tiempo real o almacenadas, que son contrastadas con bases de datos biométricas. Estas soluciones permiten localizar individuos buscados por la

⁴ Un ejemplo representativo de esta clase de soluciones tecnológicas es “Boti”, el chatbot oficial del Gobierno de la Ciudad de Buenos Aires en Argentina. Esta solución fue implementada en el año 2019 como canal para “resolver dudas y consultas vinculadas a diversos temas de la ciudad”, tales como infracciones, educación, salud, actividades culturales, reclamos, licencias y DNI, además de otras funcionalidades. Ver en: <https://buenosaires.gob.ar/innovacionytransformaciondigital/boti>. Consultado el 3 de junio de 2025.

Justicia, detectar personas desaparecidas o monitorear comportamientos considerados sospechosos en espacios públicos.

Entre los beneficios genéricos informados se destacan la automatización de los procesos de identificación, la mejora en la eficiencia operativa de las fuerzas de seguridad, y la capacidad para generar alertas en tiempo real ante coincidencias relevantes. Estas características contribuyen a una respuesta más rápida y precisa por parte de los organismos responsables.

La implementación de estos sistemas se enmarca en una lógica de vigilancia expandida, con foco en la disuasión del delito, la trazabilidad de personas específicas y la modernización de los mecanismos estatales de control y monitoreo.

- **Detección de anomalías**

Esta categoría representa aproximadamente el 13.7 % del total de casos analizados (49 en concreto)⁵. Los sistemas de detección de anomalías se aplican en diversos ámbitos de la gestión pública, como educación, salud, finanzas, administración, seguridad informática y servicios sociales.

Su finalidad común es identificar desviaciones respecto de patrones esperados en grandes volúmenes de datos, lo que permite generar alertas ante riesgos potenciales, comportamientos atípicos o irregularidades, tales como abandono escolar, fraudes, errores en contrataciones, omisiones clínicas o vulnerabilidades operativas.

⁵ Un ejemplo representativo de esta clase de soluciones tecnológicas es el Sistema de Alerta Temprana (SAT), diseñado en Entre Ríos, Argentina, para "anticipar el abandono escolar y cuidar las trayectorias de los estudiantes secundarios", para lo cual "toma la información disponible en el legajo de cada alumna y alumno dentro del sistema informático del CGE para generar un índice de riesgo de abandono escolar basado en datos de la trayectoria del estudiante y de su contexto familiar. Ver en: <https://cge.entrerios.gov.ar/2023/03/sistema-de-alertas-tempranas-para-prevenir-el-abandono-escolar/>. De manera similar, pero en lo relativo al área de la salud, en el Ministerio de Salud Nacional de Chile se encuentra funcionando "DART", una plataforma de tele-oftalmología que contiene una solución de inteligencia artificial y su objetivo es prevenir la causa más común de pérdida de visión y ceguera en la población en edad trabajadora, la retinopatía diabética. DART busca reducir las tasas de ceguera de la población al mejorar el acceso a salud preventiva para detectar signos de la enfermedad en sus etapas tempranas. Consultado el 3 de junio de 2025 en: <https://www.algoritmospublicos.cl/proyecto-dart-inteligencia-artificial-oftalmologia-minsal>.

Estos sistemas procesan datos históricos y en tiempo real provenientes de múltiples fuentes institucionales, aplicando modelos predictivos o de clasificación para emitir alertas automáticas cuando se detectan indicadores fuera de lo previsto.

Entre los beneficios generales de este tipo de soluciones se destacan la mejora en la supervisión institucional, la reducción de los tiempos de respuesta, el uso más eficiente de los recursos públicos y una mayor capacidad de intervención anticipada.

En síntesis, estas herramientas consolidan un enfoque preventivo en la gestión pública, basado en monitoreo continuo y evidencia empírica.

- **Reconocimiento de objetos, patrones y sonidos**

El análisis conjunto de las soluciones de IA basadas en reconocimiento de objetos (13% de los casos), patrones (9% de los casos) y sonidos (3% de los casos)⁶ implementadas en diversos países de América Latina evidencia una tendencia creciente hacia la automatización de funciones críticas del Sector Público.

Estas tecnologías se integran en procesos estatales con finalidades específicas: el reconocimiento de objetos se aplica principalmente al análisis de imágenes para tareas de fiscalización vehicular, monitoreo ambiental, supervisión de infraestructura y

⁶ Uno de los desarrollos tecnológicos vinculados al reconocimiento de patrones es el “Analista Virtual en gestión de licencias médicas de FONASA”, implementado en Chile. Esta herramienta permite clasificar automáticamente las licencias en aquellas que deben ser aceptadas y aquellas que requieren revisión de un médico contralor, además de calcular el Subsidio por Incapacidad Laboral que debe ser pagado, con lo que se agiliza la toma de decisiones. Consultado el 2 de junio de 2025 en: <https://algoritmospublicos.cl/proyecto-analista-virtual-licencias-medicas-fonasa>.

En lo atinente al desarrollo relativo al reconocimiento de sonido, encontramos que la Ciudad Autónoma de Buenos Aires, Argentina, junto con Microsoft, implementó un sistema con el propósito de mejorar la gestión y respuesta de los casos de violencia de género. Este sistema reconoce la voz de la llamada entrante a la línea 144, logra identificar el grado de riesgo del caso de manera autónoma, para poder tener un mayor control, hacer un mejor seguimiento y una gestión más ágil de los casos críticos. También permite reducir los tiempos operativos, y transcribe en tiempo real, almacenando el texto para su posterior análisis y resguardo. Consultado el 2 de junio de 2025 en: <https://news.microsoft.com/es-xl/la-ciudad-incorpora-inteligencia->

Por último, a modo ilustrativo del desarrollo tecnológico vinculado al reconocimiento de objetos, se encuentra “Delfos”, el cual es un sistema diseñado en Colombia para, a partir de imágenes aéreas satelitales, predecir y alertar sobre el daño ecológico causado por la minería ilegal en la flora, fauna y fuentes hídricas del país. Consultado el 2 de junio de 2025 en: <https://www.fac.mil.co/es/noticias/modelo-basado-en-inteligencia-artificial-nueva-herramienta-de-la-fuerza-aerea-para-la>

control del tránsito. El reconocimiento de patrones permite identificar regularidades o anomalías en grandes volúmenes de datos para clasificar documentos, predecir comportamientos o priorizar casos administrativos. Y el reconocimiento de sonido se utiliza para detectar eventos acústicos relevantes en contextos de emergencia, protección ambiental y evaluación educativa.

Estos sistemas de IA basados en “reconocimiento” constituyen herramientas transversales en la modernización de la gestión pública.

Su implementación no solo optimiza recursos y mejora la trazabilidad de la información, sino que también fortalece la capacidad del Estado para responder con mayor eficacia, previsión y equidad en ámbitos como seguridad, salud, educación, medio ambiente y administración tributaria.

- **Sistemas de predicción**

Esta solución representa el 10.8% del total de casos analizados (37 en total)⁷, y se utilizan principalmente para anticipar eventos o comportamientos relevantes en diversos sectores del ámbito público.

Entre los usos más frecuentes se destacan la predicción de brotes epidemiológicos, abandono escolar, riesgos de violencia de género, accidentes de tránsito, congestión hospitalaria, y fraudes financieros o administrativos. Estas implementaciones emplean técnicas de aprendizaje automático para analizar grandes volúmenes de datos históricos y en tiempo real, con el fin de identificar patrones que permitan generar alertas preventivas o apoyar la toma de decisiones anticipada.

⁷ Uno de los sistemas de predicción más relevantes es “Dara”, el cual se implementa actualmente en Brasil. Este sistema utiliza aprendizaje automático y aprendizaje profundo para ayudar a los profesionales a anticipar el pronóstico de enfermedades como la diabetes, apoyar la planificación de la infraestructura de la ciudad y la movilidad urbana, predecir la demanda de servicios públicos y anticipar sucesos como la deserción escolar, entre otras. Consultado el 2 de junio del 2025 en: <https://www.ipm.com.br/inteligencia-artificial-dara-ipm-governo-gestao-publica/>. Asimismo, en Chile, la Superintendencia de Educación, incorporó modelos predictivos basados en técnicas de machine learning, con el fin de identificar los establecimientos más propensos al incumplimiento de la normativa educacional y del uso de recursos, en base a dimensiones de información que indican un riesgo, ello con el propósito de hacer más eficiente el muestreo de establecimientos educacionales a fiscalizar. Consultado el 3 de junio de 2025 en: <https://algoritmospublicos.cl/proyecto-modelo-machine-learning-establecimientos-educacionales>.

Los beneficios genéricos de estos sistemas incluyen una mejora significativa en la capacidad de anticipación de las instituciones públicas, permitiendo respuestas más ágiles y focalizadas frente a situaciones críticas. Asimismo, contribuyen a una asignación más eficiente de los recursos estatales, al priorizar intervenciones según niveles de riesgo predichos o escenarios futuros simulados.

- **Sistemas de recomendación**

Esta categoría representa aproximadamente el 4.9% del total de casos analizados (17 en concreto)⁸ y su implementación tiene por objetivo el de optimizar la interacción entre el ciudadano y los servicios públicos.

En términos funcionales, los sistemas de recomendación identificados cumplen roles claves como: facilitar la toma de decisiones médicas (mediante la sugerencia de tratamientos o diagnósticos basados en evidencia clínica), mejorar la experiencia de búsqueda de información (con recomendaciones contextualizadas y relevantes), y orientar al ciudadano en portales institucionales mediante la personalización de servicios ofrecidos.

Algunos casos también apuntan a la zonificación técnica territorial o al diseño de políticas educativas focalizadas, empleando inferencias generadas por IA a partir de datos previos o inputs específicos del usuario.

- **Clasificadores**

⁸ El Ministerio de Economía de Brasil ha incorporado en el portal web del gobierno dos desarrollos basados en inteligencia artificial. En primer lugar, un sistema de recomendación que sugiere servicios a los usuarios en función de su historial de navegación. En segundo lugar, la posibilidad de crear perfiles personalizados para facilitar la identificación de servicios relevantes según los recorridos previos de cada ciudadano. El objetivo de estas iniciativas es mejorar la personalización de la experiencia y fortalecer la relación entre la ciudadanía y la Administración Pública. Ver en: <https://oecd-opsi.org/innovations/ai-and-user-profiles/>. Asimismo la Ciudad de Recife (Brasil) a través del programa EITA Recife, ha desarrollado un robot basado en IA que ayuda a los médicos con la derivación de pacientes. Este robot analiza los datos clínicos y recomienda el mejor procedimiento o unidad sanitaria para cada caso, optimizando el flujo de pacientes, reduciendo los tiempos de espera para la atención especializada. Consultado el 2 de junio del 2025 en: <https://portal.connectedsmartcities.com.br/2024/10/09/recife-como-protagonista-no-uso-de-inteligencia-artificial-no-brasil/>

Los sistemas de IA basados en clasificación representan casi el 4 % del total de implementaciones relevadas (12 casos en específico)⁹ y se emplean para organizar, categorizar o asignar automáticamente información en distintos procesos del Sector Público, tales como Justicia, educación, salud, propiedad intelectual y atención ciudadana.

Estas soluciones permiten procesar grandes volúmenes de datos con mayor eficiencia, estandarizar decisiones y reducir errores humanos, ofreciendo recomendaciones o clasificaciones preliminares que optimizan la labor administrativa.

- **Soporte humano-máquina**

Las soluciones de IA clasificadas como soporte humano-máquina, que representan el 1,15 % de los casos analizados (4 en total)¹⁰, se orientan a complementar tareas humanas mediante robots o sistemas inteligentes que interactúan directamente con personas.

Estas tecnologías se aplican en contextos sanitarios y de gestión pública para asistir en la atención, orientar usuarios o automatizar servicios, mejorando la eficiencia operativa, reduciendo la carga del personal y optimizando la experiencia ciudadana.

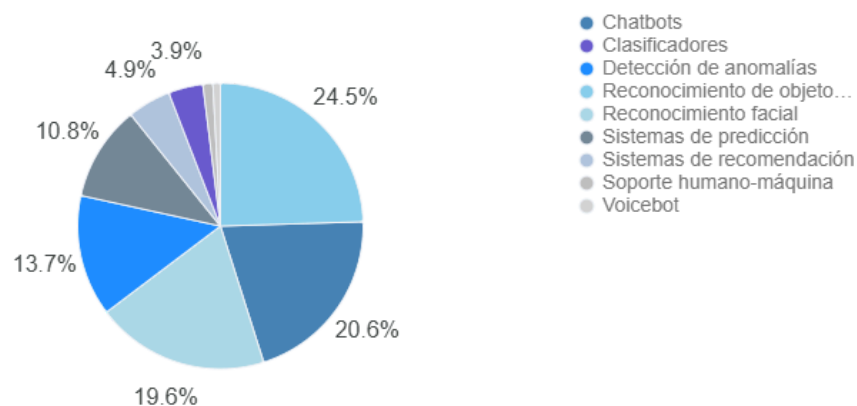
- **Voicebot**

⁹ En Uruguay, la Agencia de Gobierno Electrónico y Sociedad de la Información y del Conocimiento (AGESIC) ha diseñado un sistema para automatizar la clasificación de solicitudes de servicios de asistencia técnica mediante técnicas de IA y herramientas de código abierto para reducir los costos, optimizar los recursos y mejorar los tiempos de respuesta en beneficio de los ciudadanos y de las instituciones públicas como proveedores de servicios. Aquel sistema se encarga de clasificar los temas en diferentes categorías de servicios y vincular las tipologías de solicitudes con la información de los usuarios. Consultado el 2 de junio en: <https://oecd-opsi.org/innovations/ai-service-requests-classification/>.

¹⁰ Un ejemplo claro de este tipo de desarrollo tecnológico lo representa “Nao”, implementado en Colombia, el cual es un robot, que le cuenta de manera amena y agradable a los pacientes menores de edad del Hospital Militar detalles sobre cómo será su tratamiento, recomendaciones puntuales sobre los cuidados que deben tener, medicamentos y dietas que tendrán que llevar a cabo. Consultado el 2 de junio de 2025 en: <https://www.eltiempo.com/justicia/servicios/nao-el-robot-del-hospital-militar-que-ayuda-a-pacientes-con-cancer-834154>

Las soluciones de IA basadas en voicebots¹¹, que representan el 0,85 % de los casos analizados (3 en total)¹², se utilizan para automatizar la comunicación con usuarios en salud y educación. Permiten gestionar turnos, brindar información y mejorar la contactabilidad, reduciendo la carga operativa y agilizando la atención pública.

Aplicaciones de IA en LATAM y el Caribe



2. Estudio de casos relevantes: aplicaciones de alto impacto de IA en el Sector Público regional

Del relevamiento de implementaciones de sistemas de IA en el Sector Público de América Latina realizado por la Universidad de los Andes se seleccionaron, por su relevancia, cinco casos destacados para su estudio con más detalle. La elección se basó en criterios como el impacto tangible en la gestión pública, la transparencia en los procesos de contratación y la presencia de valoraciones positivas por parte de la opinión pública.

¹¹ Un voicebot es un software que utiliza inteligencia artificial y procesamiento de lenguaje natural para mantener conversaciones fluidas con personas a través de la voz. Estos sistemas pueden adaptarse al tono y las reacciones del interlocutor, brindando una experiencia más natural y eficiente. Ver más en: <https://blog.hubspot.es/service/que-es-voicebot>

¹² Con el objetivo de resolver una lista de espera de 28 mil cirugías y consultas de especialidad, acumulada durante la pandemia del Covid-19, el Servicio de Salud O'Higgins, Chile, implementó un sistema de contactabilidad automatizada, con inteligencia artificial, para llamar a los pacientes y confirmar si aún requerían la atención. Se trata de un voicebot, desarrollado por la empresa Elipse, entrenado con español latinoamericano para reconocimiento de lenguaje natural y que funciona con un equipo humano que valida y acompaña los llamados. Consultado el 2 de junio de 2025 en: <https://algoritmospublicos.cl/proyecto-Sistema-contactabilidad-automatizada-Hospital-Regional-Libertador-Bernardo-O-Higgins>

El análisis pormenorizado de estas experiencias busca evidenciar los alcances, desafíos y aprendizajes que surgen de iniciativas de alto valor estratégico para los gobiernos de la región.

Caso N°1: Argentina –Sistema Federal de Identificación Biométrica para la Seguridad (SIBIOS)

Objetivo: instrumentar un servicio informático para permitir la comprobación idónea y oportuna de identificación de personas y rastros para fines de seguridad pública y de investigación judicial, contribuyendo al desarrollo de políticas eficientes de prevención y conjuración de ilícitos en el ámbito de la seguridad ciudadana y al mejoramiento de las diligencias investigativas requeridas por autoridades judiciales¹³.

SIBIOS es un sistema nacional, por lo que las veinticuatro provincias del país han firmado convenios de cooperación con el Ministerio de Seguridad de Argentina, incluidas sus cuatro fuerzas de seguridad federales, y el Ministerio del Interior de Argentina, que abarca el RENAPER y la Dirección Nacional de Migraciones. Estos convenios garantizan que las policías locales puedan acceder y actualizar la base de datos. En 2017, mediante el Decreto 243/17, el Gobierno extendió el acceso a SIBIOS a cualquier organismo público de los Poderes Ejecutivo y Judicial, tanto a nivel nacional como provincial.¹⁴

Descripción técnica e institucional:

Mecanismo de contratación: La implementación de SIBIOS se realizó mediante contratación directa con la empresa proveedora de tecnología biométrica (la empresa francesa Morpho Safran, que, tras una fusión, se convirtió en IDEMIA)¹⁵. No se

¹³ Poder Ejecutivo Nacional. (2011). Decreto 1766/2011. República Argentina. Consultado el 2 de junio de 2025 en <http://www.infoleg.gob.ar/infolegInternet/anexos/185000-189999/189382/norma.htm>

¹⁴ Asociación por los Derechos Civiles (ADC). (2022). Surveillance Technology in Argentina. Consultado el 2 de junio de 2025 en: <https://adc.org.ar/wp-content/uploads/2022/03/ADC-Surveillance-Technology-in-Arentina.pdf>

¹⁵ Asociación por los Derechos Civiles (ADC). (2022). Surveillance Technology in Argentina. Consultado el 2 de junio de 2025 en: <https://adc.org.ar/wp-content/uploads/2022/03/ADC-Surveillance-Technology-in-Arentina.pdf>

encontraron registros públicos de licitaciones abiertas o pliegos de contratación específicos disponibles en línea relacionados con este sistema.

Críticas públicas: La principal crítica dirigida al sistema SIBIOS radica en su posible contradicción con las normas de protección de datos del país y en las implicancias que conlleva para la seguridad ciudadana, debido a la ausencia de mecanismos de control claramente definidos.

En esta línea, la organización local Fundación Vía Libre, junto con la Electronic Frontier Foundation (EFF), han alertado sobre los riesgos que representa la implementación de SIBIOS para la privacidad de las personas. Desde hace tiempo, la EFF advierte sobre los efectos perjudiciales que puede generar “el seguimiento perfecto” en el marco de una sociedad libre y democrática. En sintonía con estas preocupaciones, Julián Assange, fundador de WikiLeaks, afirmó que Argentina —si bien no alcanza la escala de vigilancia de China o Estados Unidos— posee uno de los regímenes de monitoreo más agresivos de América Latina¹⁶.

Caso N° 2: Brasil – Sistema de Inteligencia Artificial “Dara” para la gestión educativa en Rio do Sul

Objetivo: Optimizar la gestión de la educación pública municipal mediante IA, con foco en la predicción del abandono escolar, la planificación de infraestructura educativa y la mejora de políticas públicas. El sistema identifica a estudiantes con alto riesgo de deserción escolar, anticipa la demanda de plazas en guarderías y centros de educación infantil, y propone acciones concretas para racionalizar recursos y mejorar la calidad educativa. A su vez, permite generar un diagnóstico detallado y predictivo a nivel de alumno, escuela y municipio, con el fin de orientar el desarrollo del Plan Municipal de Educación¹⁷.

¹⁶ Giswatch / Asociación por los Derechos Civiles (ADC). (2014). Vigilancia de las comunicaciones en Argentina. En GISWatch 2014: Vigilancia de las comunicaciones en la internet. Recuperado de <https://www.giswatch.org/es/country-report/communications-surveillance/argentina>

¹⁷ IPM Sistemas. (2023, 13 de octubre). Rio do Sul es pionera nacional en el uso de Inteligencia Artificial en la educación pública. Consultado el 2 de junio de 2025 en <https://www.ipm.com.br/rio-do-sul-e-pioneira-no-uso-de-inteligencia-artificial-na-educacao-publica/>

Descripción técnica e institucional: Dara es una herramienta de IA desarrollada por la empresa IPM Sistemas. Fue implementada por el municipio de Rio do Sul, en el estado de Santa Catarina (Brasil), posicionándose como el primer municipio del país en utilizar una IA de este tipo en el sistema educativo público.

Mediante algoritmos de aprendizaje automático, la IA actúa como un asistente estratégico para los gestores públicos, permitiendo definir la asignación óptima de guarderías y centros preescolares. La herramienta proporciona datos clave como ubicación recomendada, capacidad requerida, proyección de demanda y distribución de alumnos. Esto se traduce en una mejora directa en la eficiencia del uso de los recursos públicos, ya que permite orientar las inversiones con precisión, evitando gastos innecesarios y aumentando la efectividad en la toma de decisiones. Al mismo tiempo, contribuye a mejorar la calidad de vida de la población, dado que cada acción de gestión se fundamenta en un análisis riguroso de datos respaldado por IA¹⁸.

Mecanismo de contratación: La adquisición de Dara se realizó mediante contratación directa del municipio de Rio do Sul con IPM Sistemas, la empresa desarrolladora de la IA. Aunque no se encontraron pliegos públicos en línea, el vínculo con IPM forma parte del proceso de digitalización institucional liderado por la gestión municipal, y se enmarca en el uso de soluciones tecnológicas integradas a sistemas de gestión administrativa ya contratados. No se identificaron registros de licitación pública abierta para esta implementación en particular.

Críticas públicas : Hasta el momento, no se han identificado críticas públicas formales o cuestionamientos institucionales a la implementación de Dara. Por el contrario, ha recibido valoraciones positivas por parte de autoridades locales, como la Secretaria Municipal de Educación de Rio do Sul, Janara Mafra, quien destacó que el gran apoyo que representa la IA¹⁹.

¹⁸ Universidad de los Andes – Escuela de Gobierno Alberto Lleras Camargo. (s.f.). Sistemas de Algoritmos Públicos. Recuperado el 3 de junio de 2025, de <https://sistemaspublicos.tech/>

¹⁹ Universidad de los Andes – Escuela de Gobierno Alberto Lleras Camargo. (s.f.). Sistemas de Algoritmos Públicos. Recuperado el 3 de junio de 2025, de <https://sistemaspublicos.tech/>

Sin embargo, la aplicación de IA para la gestión de datos sensibles como los educativos podría eventualmente abrir debates sobre privacidad, uso de algoritmos en políticas públicas y el control de las decisiones automatizadas, especialmente si su uso se expande a otras jurisdicciones sin evaluación ética previa.

Caso N°3: Brasil – Sistema de IA para la Identificación de Acciones de Zeladoria (Municipio de Piracicaba)

Objetivo: Implementar una solución basada en IA para detectar de forma automatizada y proactiva situaciones que requieran tareas de mantenimiento urbano, optimizando la respuesta estatal y liberando recursos humanos previamente afectados a tareas de patrullaje y reporte manual.

El objetivo final es mejorar la calidad de vida de la población mediante la modernización de los servicios municipales de zeladoria. Entre las ciudades que utilizan el servicio están Hortolândia (SP), Palmas (TO), Santa Rosa (RS), Luís Eduardo Magalhães (BA), Miguel Pereira (RJ) y Uberaba (MG)²⁰.

Descripción técnica e institucional: El sistema consiste en un vehículo equipado con una cámara de alta resolución y sensores, integrados a un software de IA desarrollado por la empresa Mapzer Inteligencia Artificial Ltda. Este vehículo recorre periódicamente las calles de Piracicaba, captando imágenes y procesándolas en tiempo real para identificar más de 30 tipos de incidencias urbanas.

Los eventos detectados son georreferenciados y derivados automáticamente a las secretarías municipales competentes para su atención. El proyecto es coordinado por la Secretaría de Movilidad, Tránsito y Transportes (Semuttran).

Mecanismo de contratación: El sistema fue contratado por el Municipio de Piracicaba mediante contratación directa, invocando la exclusividad tecnológica de la empresa

²⁰ Da Redação. (2024, 1 de febrero). Prefeitura vai usar inteligência artificial para identificar buracos e mato alto. Sampi Piracicaba. Recuperado el 3 de junio de 2025, de <https://sampi.net.br/piracicaba/noticias/2813556/cidade/2024/02/prefeitura-vai-usar-inteligencia-artificial-para-indenticar-buracos-e-mato-alto>

proveedora, Mapzer Inteligencia Artificial Ltda. El contrato, firmado sin licitación pública, tiene una duración de 12 meses y un valor de R\$ 1,6 millones²¹.

Críticas públicas: algunos habitantes expresaron su preocupación en redes sociales, considerando el gasto innecesario y afirmando que esos recursos podrían aplicarse en necesidades más urgentes²².

Valoración institucional: Desde la Municipalidad se valora positivamente la implementación. La secretaria de Semuttran, Jane Franco de Oliveira, destacó que la herramienta permite tomar decisiones con base en datos actualizados, mejora la planificación de intervenciones y optimiza la asignación de recursos públicos²³.

Se la considera una innovación útil para aumentar la eficacia de la gestión urbana y mejorar la calidad de vida de la población mediante una respuesta más ágil y precisa a los problemas cotidianos de mantenimiento.

Caso N°4: Uruguay – Sistema de Inteligencia Artificial para Videoanalítica Proactiva y Atención a Personas en Situación de Calle

Objetivo: Incorporar una solución de IA al sistema nacional de videovigilancia para identificar de manera proactiva tanto conductas asociadas a delitos como situaciones de vulnerabilidad extrema, especialmente personas durmiendo en la vía pública. La finalidad es mejorar la eficacia de la seguridad ciudadana y, a su vez, fortalecer los

²¹ Prefeitura de Piracicaba. (2024, 2 de julio). Prefeitura inicia uso de inteligência artificial na identificação de ações de zeladoria. Recuperado el 3 de junio de 2025, de <https://piracicaba.sp.gov.br/noticias/prefeitura-inicia-uso-de-inteligencia-artificial-na-identificacao-de-acoes-de-zeladoria/>

²² Da Redação. (2024, 1 de febrero). Prefeitura vai usar inteligência artificial para identificar buracos e mato alto. Sampi Piracicaba. Recuperado el 3 de junio de 2025, de <https://sampi.net.br/piracicaba/noticias/2813556/cidade/2024/02/prefeitura-vai-usar-inteligencia-artificial-para-identificar-buracos-e-mato-alto>

²³ Da Redação. (2024, 1 de febrero). Prefeitura vai usar inteligência artificial para identificar buracos e mato alto. Sampi Piracicaba. Recuperado el 3 de junio de 2025, de <https://sampi.net.br/piracicaba/noticias/2813556/cidade/2024/02/prefeitura-vai-usar-inteligencia-artificial-para-identificar-buracos-e-mato-alto>

mecanismos de respuesta humanitaria, garantizando una intervención estatal más rápida y oportuna²⁴.

Descripción técnica e institucional: El sistema se integra al Centro de Comando Unificado del Ministerio del Interior de Uruguay y utiliza algoritmos de visión artificial aplicados sobre una red de más de 11.900 cámaras de videovigilancia instaladas en todo el país.

La herramienta permite identificar automáticamente una cara que coincide con una base de datos de personas requeridos o cuando localiza a un señor que se vista de determinada manera²⁵, cuando las detecte, el sistema emite una alerta al funcionario que en ese momento esté visualizando las cámaras de videovigilancia, el cual tendrá que verificar la información y enviar el aviso correspondiente.

La iniciativa tenía previsto ser implementada de forma conjunta entre el Ministerio del Interior y el MIDES a fines de 2023, adaptando para fines sociales una herramienta originalmente diseñada para tareas de seguridad.

Mecanismo de contratación: La adquisición del sistema se realizó mediante contrataciones públicas del Ministerio del Interior, que incluyeron un piloto con participación de cuatro empresas tecnológicas locales. Posteriormente, el sistema fue compartido con el MIDES para ampliar su finalidad social. El desarrollo se enmarca dentro de una estrategia de cooperación interinstitucional con apoyo técnico del sector privado y bajo supervisión estatal.

²⁴ Infobae. (2023, 21 de agosto). El gobierno uruguayo usará inteligencia artificial para ofrecer refugio a personas que viven en la calle. Recuperado el 3 de junio de 2025, de <https://www.infobae.com/america/america-latina/2023/08/21/el-gobierno-uruguayo-usara-inteligencia-artificial-para-detectar-personas-que-viven-en-la-calle/>

²⁵ Infobae. (2023, 21 de agosto). El gobierno uruguayo usará inteligencia artificial para ofrecer refugio a personas que viven en la calle. Recuperado el 3 de junio de 2025, de <https://www.infobae.com/america/america-latina/2023/08/21/el-gobierno-uruguayo-usara-inteligencia-artificial-para-detectar-personas-que-viven-en-la-calle/>

Críticas públicas: Si bien la iniciativa fue bien recibida en su dimensión social, generó cuestionamientos vinculados a la privacidad y el conflicto que entraña la incorporación de este sistema con la Ley de Protección de datos vigente en Uruguay²⁶.

Se señaló, además, la importancia de implementar protocolos éticos que regulen el tratamiento de datos sensibles y las consecuencias de una vigilancia automatizada en el espacio público.

Valoración institucional: Las autoridades valoran positivamente la implementación como un salto cualitativo en materia de gestión pública y seguridad con enfoque social. El sistema de IA propuesto permite a la policía actuar preventivamente lo que mejora significativamente la capacidad de respuesta frente a personas que requieren asistencia urgente, especialmente durante el invierno, como así también la detección de situaciones de inseguridad. Se lo considera un ejemplo de uso dual de la IA orientado tanto a la disuasión del delito como al fortalecimiento de políticas públicas inclusivas.

Caso N°5: Chile – Sistema de IA para Automatización del Pago de Licencias Médicas (Ministerio de Salud)

Objetivo: Implementar un sistema de IA destinado a automatizar la evaluación y resolución de licencias médicas presentadas por trabajadores, con el propósito de reducir los tiempos de espera y garantizar el pago oportuno de subsidios por incapacidad laboral. El objetivo final es optimizar el funcionamiento de la Comisión de Medicina Preventiva e Invalidez (COMPIN), fortaleciendo así el acceso a derechos laborales en el ámbito de la salud.

Descripción técnica e institucional: El sistema, desarrollado por el Ministerio de Salud, se basa en IA entrenada con los mismos insumos utilizados por los médicos contralores que revisan manualmente las licencias: guías clínicas, listado de especialistas

²⁶ Telenoche. (2023, 21 de agosto). Mides e Interior prevén usar Inteligencia Artificial. Recuperado el 3 de junio de 2025, de <https://www.telenoche.com.uy/informes-especiales/mides-e-interior-preven-usar-inteligencia-artificial-n5353957>

habilitados, y promedios estadísticos de días de permiso según diagnóstico. La herramienta cruza y analiza estos datos de forma automatizada para determinar si una licencia puede ser aprobada directamente o requiere revisión por parte de un profesional humano²⁷.

Mecanismo de contratación: El sistema fue implementado por el Ministerio de Salud en el marco de un plan de modernización del sistema COMPIN. Si bien la fuente consultada no menciona el proveedor exacto del software, el desarrollo fue conducido por el Estado con apoyo técnico y recursos públicos. No consta que haya sido mediante licitación abierta, por lo que se presume que se trató de una contratación directa o desarrollo interno con colaboración especializada.

Críticas públicas: No se identificaron críticas públicas directas al sistema, pero algunos especialistas señalaron que la automatización no resuelve por sí sola los problemas estructurales del sistema de licencias médicas, como la fragmentación de competencias entre instituciones o la deficiente calidad de algunos certificados médicos. También se advirtió que el uso de IA debe ser acompañado por una evaluación constante de sus decisiones y criterios, para evitar errores o inequidades en la aprobación o rechazo de solicitudes.

Valoración institucional: La implementación fue ampliamente valorada por las autoridades sanitarias. La Subsecretaria de Salud Pública, Paula Daza, destacó: “Las licencias médicas electrónicas que llegan a estas regiones hoy se están incorporando al sistema de red neuronal, un sistema que nos permite ser más eficientes”; “el objetivo más importante son las personas y que reciban su pago de forma oportuna”; y “el sistema funciona durante la noche y en promedio está ejecutando alrededor de 3.500 licencias al día, lo que equivale a la carga de trabajo de 14 médicos contralores”²⁸.

²⁷ La Tercera. (2019, 5 de febrero). Inteligencia artificial en Compín redujo en 60,5% el plazo de pago de licencias médicas. Recuperado el 3 de junio de 2025, de <https://www.latercera.com/nacional/noticia/inteligencia-artificial-compin-redujo-605-plazo-pago-licencias-medicas/515629/>

²⁸ La Tercera. (2019, 5 de febrero). Inteligencia artificial en Compín redujo en 60,5% el plazo de pago de licencias médicas. Recuperado el 3 de junio de 2025, de <https://www.latercera.com/nacional/noticia/inteligencia-artificial-compin-redujo-605-plazo-pago-licencias-medicas/515629/>

Según datos oficiales, el sistema logró reducir el tiempo promedio de pago de licencias de 106 días a 42 días en los primeros meses de funcionamiento, y se proyectaba llegar a menos de 30 días.

Un resultado adicional fue la detección de patrones de uso indebido del sistema. Según informó el Ministerio, el análisis de datos permitió mapear el comportamiento de ciertos médicos emisores, identificando 28 profesionales con presuntas irregularidades²⁹.

Además, el uso de IA permitió agilizar el cálculo de los pagos, mejorar la trazabilidad de los trámites y descongestionar las oficinas de COMPIN, donde históricamente se acumulaban reclamos y expedientes. El Ministerio de Salud calificó la experiencia como un caso exitoso de digitalización pública centrada en el usuario, al priorizar el pago oportuno a personas enfermas que dependen de ese ingreso.

3. Conclusiones preliminares acerca del panorama actual de la contratación de IA en la región

El análisis del estado actual de la contratación de sistemas de IA en América Latina y el Caribe revela un proceso de expansión sostenida, impulsado por la necesidad de modernizar los servicios públicos mediante soluciones tecnológicas capaces de mejorar la eficiencia operativa, optimizar recursos y ampliar el acceso ciudadano.

Esta dinámica ha dado lugar a una implementación creciente de sistemas de IA en áreas sensibles como la salud, la seguridad, la educación y la gestión administrativa.

Sin embargo, esta adopción no siempre ha ido acompañada de procedimientos contractuales transparentes, exigencias normativas estandarizadas o esquemas de gobernanza adecuados para abordar los riesgos específicos que implican estas tecnologías.

²⁹ La Tercera. (2019, 5 de febrero). Inteligencia artificial en Compín redujo en 60,5% plazo de pago de licencias médicas. Recuperado el 3 de junio de 2025, de <https://www.latercera.com/nacional/noticia/inteligencia-artificial-compin-redujo-605-plazo-pago-licencias-medicas/515629/>

Una de las principales limitaciones identificadas durante la investigación radica en la falta de acceso público a pliegos contractuales, cláusulas específicas o documentación técnica que permita conocer con precisión las condiciones bajo las cuales fueron adquiridas e implementadas muchas de estas soluciones.

En la mayoría de los casos, el análisis debió sustentarse exclusivamente en bibliografía secundaria, reportes institucionales, notas de prensa o menciones indirectas en publicaciones académicas. Esta opacidad estructural en la información pública disponible representa un obstáculo significativo para la evaluación independiente del cumplimiento de estándares éticos, principios de transparencia y criterios de supervisión humana en la contratación pública de sistemas basados en IA.

A ello se suma la prevalencia de mecanismos de contratación directa que, si bien responden a criterios de celeridad operativa, limitan las posibilidades de garantizar procesos competitivos, inclusivos y verificables.

En este contexto, la evaluación técnica y jurídica de dichas contrataciones se torna compleja, y se dificulta la exigibilidad de principios como la explicabilidad algorítmica, la seguridad de los datos tratados, o la responsabilidad ante fallas o sesgos de los sistemas.

III. La relevancia de contar con instrumentos para la gobernanza ética de la contratación de sistemas de IA para el Sector Público

El último quinquenio da cuenta de distintas situaciones en las cuales el despliegue de sistemas de IA ha provocado un impacto negativo en áreas sensibles como la justicia³⁰ o la seguridad social³¹.

Estas fallas suelen estar marcadas por situaciones que afectan directamente los derechos de las personas³², muchas veces con repercusiones políticas de relevancia,

³⁰ Un ejemplo claro de ello versa sobre un estudio llevado a cabo por Sino Esthappan, estudiante de posgrado de la Universidad Northwestern, quien investigó sobre el uso de algoritmos de evaluación de riesgo en audiencias previas al juicio en Estados Unidos. Aunque estas herramientas pretenden reducir sesgos humanos al predecir la probabilidad de que un acusado no comparezca en el tribunal o cometa otro delito, el estudio revela que los jueces las utilizan selectivamente para respaldar decisiones que ya desean tomar. El factor decisivo para los jueces a menudo no era si el algoritmo parecía confiable, sino si les ayudaría a justificar la decisión que querían tomar. Asimismo, el estudio revela que en los casos de delitos menores los jueces, mayormente, confían en los algoritmos para agilizar procesos, pero en delitos graves, como violencia doméstica o agresión sexual, prefieren basarse en su criterio personal, preocupados por su reputación y la precisión de las herramientas. El estudio de Esthappan cuestiona la idea de que si los jueces deciden cuándo confiar en las puntuaciones basándose en factores como el riesgo reputacional, es posible que no estén reduciendo el sesgo humano, sino que podrían estar *legitimándolo* y dificultando su detección. Ver en The Verge “Judges are using algorithms to justify doing what they already want”. 2024. Consultado el 3 de junio de 2025 en: <https://www.theverge.com/2024/10/30/24281924/pretrial-risk-assessment-algorithms-research>

³¹ Ferrer, I. “El Gobierno holandés dimite en bloque por el escándalo en las ayudas al cuidado de niños” Diario El País España. 2021. Consultado el 3 de junio de 2025 en: <https://elpais.com/internacional/2021-01-15/el-gobierno-holandes-estudia-la-dimision-en-bloque-por-el-escandalo-en-las-ayudas-al-cuidado-de-ninos.html>.

³² Ejemplo claro de ello lo encontramos en cómo el uso de tecnologías de reconocimiento facial por parte de las fuerzas de seguridad en Estados Unidos han llevado a arrestos erróneos de al menos siete personas, seis de ellas afroamericanas. Esto ha generado preocupación sobre la menor fiabilidad de estos sistemas para personas de color y el riesgo de que, sin una validación rigurosa, se profundicen las desigualdades raciales en la actuación policial y en el sistema judicial penal. Ver en Innocence Project “Artificial Intelligence Is Putting Innocent People at Risk of Being Incarcerated”. 2024. Consultado el 3 de junio de 2025 en <https://innocenceproject.org/news/artificial-intelligence-is-putting-innocent-people-at-risk-of-being-incarcerated/>

Otros ejemplos dan cuenta del impacto negativo de la incorporación de Inteligencia Artificial dentro de la órbita de la seguridad social. Podemos aquí referir al uso de “Meu INSS”, un sistema utilizado por el Instituto Nacional del Seguro Social de Brasil, que emplea inteligencia artificial para procesar solicitudes de beneficios sociales. Si bien dicho sistema agiliza trámites simples, también generó rechazos automáticos en casos complejos, afectando especialmente a trabajadores rurales con baja alfabetización digital. Errores menores, como datos personales incorrectos, han llevado a denegaciones sin opciones claras de apelación, evidenciando cómo la automatización puede perjudicar a los más vulnerables si no se implementa con salvaguardas adecuadas. Ver en Rest of World “La aplicación de seguridad social de Brasil, impulsada por inteligencia artificial, rechaza solicitudes erróneamente”. 2025. Consultado el 3 de junio de 2025 en: <https://restofworld.org/2025/brazil-ai-social-security-app-rejected/>.

Otro ejemplo lo encontramos en el sistema utilizado por Reino Unido para detectar fraudes en beneficios sociales el cual mostró sesgos estadísticamente significativos en función de la edad, el estado

que incrementan la desconfianza de la ciudadanía hacia sus gobiernos, lo que a su vez frena el avance de la transformación digital en los países.

Por tal motivo, la contratación de **sistemas de IA**³³ para el Sector Público debería regirse por un **sólido marco de gobernanza ética**, basado en el cumplimiento de estándares contractuales que apliquen principios éticos ampliamente consensuados en materia de IA, así como la normativa vigente aplicable en cada jurisdicción. Esto incluye, en particular, las leyes de protección de datos personales, de acceso a la información pública y de protección contra toda forma de discriminación.

Dicho marco debe asegurar que los sistemas de IA contratados presenten un **diseño, desarrollo y despliegue ético, adecuado, responsable y centrado en las personas**, en pleno respeto de sus derechos fundamentales, las libertades civiles y los valores democráticos.

A ese fin, el enfoque de gobernanza que se propone en este informe busca alinear cada fase del ciclo de vida de los sistemas de IA con los principios internacionales establecidos en la Recomendación sobre la Ética de la IA de la UNESCO y en los Principios de la OCDE sobre Inteligencia Artificial, promoviendo la transparencia, la rendición de cuentas, la equidad, la sostenibilidad y la legalidad de las soluciones adoptadas.

civil, la nacionalidad y la discapacidad de los solicitantes. Un análisis interno reveló que el sistema seleccionaba incorrectamente a personas de algunos grupos más que de otros al recomendar a quién investigar por posible fraude. Los activistas respondieron acusando al gobierno de aplicar una política de “primero dañar, luego arreglar” y pidieron a los ministros que fueran más abiertos acerca de qué grupos probablemente serían erróneamente sospechosos por el algoritmo de intentar engañar al sistema. Respecto de ello el DWP sostiene que la herramienta no sustituye el criterio humano y que un asistente social siempre analizará toda la información para tomar la decisión. Ver en The Guardian “Revealed: bias found in AI system used to detect UK benefits fraud”. 2024. Consultado el 3 de junio de 2025 en: <https://www.theguardian.com/society/2024/dec/06/revealed-bias-found-in-ai-system-used-to-detect-uk-benefits>

³³ A los fines de este documento se toma un concepto amplio de sistema de IA, entendiendo por tal a “...un sistema basado en máquinas que, para objetivos explícitos o implícitos, infiere, a partir de la información que recibe, cómo generar resultados como predicciones, contenido, recomendaciones o decisiones que pueden influir en entornos físicos o virtuales. Los diferentes sistemas de IA varían en sus niveles de autonomía y adaptabilidad tras su implementación”, conforme la definición revisada de OCDE de 2023. Ver en Legal Instruments “Recommendation of the Council on Artificial Intelligence” OECD. 2024. <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449>.

Los estándares propuestos a partir de este documento resultan aplicables, no sólo a la contratación de **sistemas de IA desarrollados a medida para una organización**, sino también a la adquisición de **otras soluciones de software**, incluyendo **sistemas de IA de propósito general**, **sistemas que incorporen IA** para potenciar algunas de sus funcionalidades, **flujos agénticos y agentes de IA**.

En algunos casos, como sucede con los flujos agénticos y agentes de IA, así como con los sistemas de IA de propósito general, esa aplicación debe realizarse **de forma diferenciada y proporcional**, considerando el grado de autonomía, el nivel de intervención humana requerido y el impacto potencial que los sistemas puedan provocar en los derechos de las personas, los procesos u objetivos institucionales, y atendiendo a que se trata de soluciones que presentan particularidades en cuanto a su modo de funcionamiento y a las posibilidades reales del proveedor de acceder o conocer información específica sobre los modelos subyacentes —incluido su entrenamiento—, lo que exige adaptar las exigencias contractuales de manera razonable a las limitaciones estructurales inherentes a estas soluciones.

Los estándares contractuales que se proponen junto con este documento son instrumentos que colaboran a la **gobernanza³⁴ de la contratación de sistemas de IA, sistemas de IA propósito general, sistemas que incorporen IA para potenciar sus funcionalidades básicas, flujos agénticos y agentes de IA**.

A tal fin, entendemos por gobernanza de la compra pública de IA, al conjunto de normas, políticas, prácticas y lineamientos que componen el marco de gestión destinado a orientar la compra pública responsable de sistemas de IA, con el objetivo de maximizar los beneficios para las organizaciones y las personas, garantizando al mismo tiempo el respeto por principios éticos fundamentales como la transparencia y

³⁴ Existen distintas definiciones para describir el alcance del concepto de gobernanza de la IA. A los fines de este documento, tomaremos un concepto que integra los aspectos éticos, organizativos y tecnológicos del ciclo de vida de la IA, que resulta aplicable a nivel de organización. Ello así, entendemos por gobernanza de la IA al: sistema de reglas (soft y hard law), prácticas, procesos y herramientas tecnológicas que se emplean para garantizar que el uso de tecnologías de IA por parte de una organización se alinee con las estrategias, objetivos y valores de la organización; cumpla con los requisitos legales; y cumpla con los principios de IA éticos aplicables. En línea con: Matti Mäntymäki, Matti Minkkinen, Teemu Birkstedt & Mika Viljanen, “Defining organizational AI governance”, Volume 2, pages 603–609, (2022), <https://link.springer.com/article/10.1007/s43681-022-00143-x#Sec2>.

explicabilidad, la rendición de cuentas, la equidad y el respeto por los derechos humanos.

En términos más simples, la gobernanza de la contratación de IA por parte del Sector Público, representa su **capacidad para ejercer un control estratégico sobre las compras de sistemas inteligentes en general**, de modo que pueda canalizar su potencial de manera segura y alineada con las normas jurídicas y principios éticos aplicables³⁵.

En consecuencia, una **adecuada gobernanza** requiere de la aplicación de estándares internacionales consolidados en materia de ética de la IA, como aquellos enunciados en el documento de Principios Éticos de IA de la OCDE y en la Recomendación sobre Ética de la IA de la UNESCO³⁶.

No obstante, como se señaló, dichos estándares éticos generales deben ser trabajados y adaptados para su aplicación concreta al contexto de la compra pública, de modo tal que orienten eficazmente los procesos de adquisición, contratación y supervisión de sistemas de IA desde una perspectiva centrada en los derechos y en el interés público.

Ambos documentos, el de OCDE y el de UNESCO, se toman como referencia no solo por su contenido normativo y técnico, sino también por el amplio consenso internacional que han logrado.

La Recomendación de UNESCO —aplicable a los 194 Estados miembros de la UNESCO— establece un marco ético integral para el diseño, desarrollo y despliegue de sistemas de IA³⁷ e incorpora principios como la centralidad de la persona, la equidad, la transparencia y la rendición de cuentas. Por su parte, los Principios de la OCDE,

³⁵ En la misma línea se expresan las Recomendaciones de UNESCO, cuando señalan que la gobernanza de la IA busca lograr que los beneficios que trae asociados su empleo puedan ser compartidos por todos y contribuir al desarrollo sostenible. Recomendación sobre la Ética de la Inteligencia Artificial. Pág 23. UNESCO. 2021. <https://www.unesco.org/es/articles/recomendacion-sobre-la-etica-de-la-inteligencia-artificial> consultado el 22/05/2025.

³⁶ UNESCO. (2021). Recomendaciones sobre la Ética de la Inteligencia Artificial (p. 20). Recuperado el 3 de junio de 2025, de https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

³⁷ UNESCO. (2021). Recomendaciones sobre la Ética de la Inteligencia Artificial (p. 20). Recuperado el 3 de junio de 2025, de https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

adoptados en 2019 y actualizados en 2023 y 2024, constituyen la primera norma intergubernamental sobre IA respaldada por economías avanzadas y emergentes, y ofrecen un enfoque pragmático basado en valores democráticos, desarrollo sostenible y enfoque de ciclo de vida.

Ambos marcos se complementan mutuamente, ofrecen una base sólida y legitimada para guiar la contratación ética y responsable de sistemas de IA en el Sector Público, y presentan la suficiente flexibilidad para adaptar sus lineamientos al contexto de la compra pública, al tipo de sistema de IA que se contrata y su nivel de impacto.

Así, las pautas básicas que orienten la contratación estatal deberían contemplar, como mínimo y de manera adaptada conforme los factores ya señalados, los siguientes principios³⁸:

- Proporcionalidad e inocuidad;
- Seguridad y robustez técnica;
- Equidad, no discriminación y diversidad;
- Sostenibilidad y bienestar;
- Privacidad y protección de datos;
- Supervisión y decisión humana;
- Transparencia y explicabilidad;
- Responsabilidad y rendición de cuentas;
- Concientización y alfabetización;
- Colaboración de múltiples partes interesadas.

³⁸ UNESCO. (2021). Recomendaciones sobre la Ética de la Inteligencia Artificial (p. 20). Recuperado el 3 de junio de 2025, de https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa.

En dicho marco, los instrumentos que se proponen en este documento constituyen una aplicación concreta de los referidos principios, diseñada para colaborar a promover la **innovación responsable guiada por principios éticos**, mediante la definición de reglas claras que aseguren la confiabilidad de los sistemas de IA adquiridos y el respeto a los derechos de las personas. Al mismo tiempo, estos instrumentos se han diseñado para colaborar a que la gobernanza de la IA **no sume barreras que obstaculicen o entorpezcan la capacidad del Estado de avanzar con decisión en la adquisición de tecnologías de IA** que permitan mejorar sustancialmente la prestación de servicios públicos.

Existe hoy una demanda ciudadana legítima por respuestas más ágiles, efectivas y transparentes, y **los sistemas de IA, cuando son bien diseñados y adecuadamente gobernados, pueden ser una herramienta estratégica** para satisfacer esa expectativa social creciente.

Estas herramientas pretenden funcionar como un **marco de referencia** que puede ser adoptado y adaptado por cada organismo público en función de sus necesidades específicas y del tipo de sistema de IA que se pretenda contratar.

Se alienta a que su utilización como insumo para el diseño o adecuación de normativa interna en materia de compras públicas de sistemas de IA siga una **lógica de aplicación proporcional**, diferenciando entre sistemas según su impacto en derechos y áreas sensibles, como se explicará más adelante, y también considerando la particularidad de algunas soluciones, como sucede con los sistemas de IA de propósito general, el software que incorpora IA en algunas de sus funciones, los flujos agénticos y agentes de IA.

Ello pues se trata de sistemas de IA cuyas características, en ciertos casos, requiere de enfoques específicos que aseguren estándares mínimos sin obstaculizar su integración en entornos públicos. O bien que se contemplen adecuaciones, como sucede con las contrataciones de software que, si bien no tiene a la IA como funcionalidad principal, integra capacidades de IA para potenciar determinados procesos o tareas. En estos

supuestos, los **requisitos contractuales deberán aplicarse de manera adaptada**, evitando cargas excesivas sin renunciar a los principios éticos fundamentales.

Es importante no perder de vista que, de manera adicional a los estándares, se deben incorporar al contrato todas aquellas exigencias que surgen de la normativa vigente en cada jurisdicción, ya que los contratos públicos de sistemas de IA, en tanto especie de un género más amplio (el contrato administrativo) deben siempre cumplir con la regulación aplicable.

En suma, para una gobernanza ética de la contratación de sistemas de IA, se necesita consolidar al contrato como una herramienta integral, que contemple los aspectos técnicos, pero también los éticos y jurídicos relacionados con los sistemas de IA, facilitando que el Sector Público pueda adoptar soluciones innovadoras que le permitan mejorar los servicios que brinda a la ciudadanía, sin poner en riesgo los derechos de las personas ni la rendición de cuentas institucional.

IV. Principios éticos aplicables a la compra pública de sistemas de IA

1. Proporcionalidad e inocuidad

Los sistemas de IA, así como los procesos relacionados con su ciclo de vida, deben orientarse exclusivamente al logro de propósitos u objetivos legítimos, garantizando siempre su adecuación al contexto y evitando intervenciones que excedan lo necesario³⁹.

Este enfoque busca asegurar que el diseño, desarrollo y uso de la IA se realicen de manera ética, responsable, proporcional y contextualizada. Al limitar su aplicación a fines legítimos y necesarios, se promueve la protección de derechos fundamentales, se previenen abusos o usos desmedidos de la tecnología, y se fortalece la confianza pública en los sistemas de IA.

³⁹ UNESCO. (2021). Recomendaciones sobre la Ética de la Inteligencia Artificial (p. 22). Recuperado el 3 de junio de 2025, de https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

En razón de ello, la adopción de sistemas de IA, así como la selección del enfoque técnico correspondiente, deben estar debidamente fundamentados, para lo cual se debería verificar que el enfoque elegido⁴⁰:

- Es pertinente y proporcional respecto del objetivo legítimo que se busca alcanzar;
- Respetar principios éticos fundamentales, como los enunciados en este documento, y su empleo no afecta derechos humanos o incurre en usos abusivos;
- Es adecuado para el contexto de aplicación elegido y se sustenta en evidencia científica sólida.

Cuando se trate de sistemas de alto impacto, la responsabilidad última de la decisión debe recaer en una persona humana.

La integración de este principio a la compra pública de sistemas de IA busca promover la confianza pública en el sistema contratado mediante un diseño y uso ético, responsable y proporcional, protegiendo los derechos fundamentales, previniendo abusos y asegurando que la tecnología se aplique únicamente para fines legítimos y necesarios, en alineación con estándares éticos internacionales.

2. Seguridad y robustez técnica

Los sistemas de IA deben ser robustos, seguros y confiables durante todo su ciclo de vida para que, en condiciones de uso normal, uso previsible, mal uso u otras condiciones adversas, funcionen adecuadamente y no planteen riesgos irrazonables de seguridad.

El proveedor debe seguir el principio de “seguridad desde el diseño y por defecto”, asegurando que la solución cuenta con controles incorporados desde su concepción. El

⁴⁰ UNESCO. (2021). Recomendaciones sobre la Ética de la Inteligencia Artificial (p. 22). Recuperado el 3 de junio de 2025, de https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

contrato exigirá que el sistema alcance niveles apropiados de precisión (performance), solidez (robustez ante distintas condiciones) y resiliencia frente a ataques o interferencias, manteniendo un comportamiento consistente y seguro a lo largo de todo su ciclo de vida.

Antes de la entrega, el proveedor implementará medidas técnicas y organizativas concretas, descritas posiblemente en un anexo específico (por ejemplo, cifrado de datos en tránsito y en reposo, autenticación robusta de usuarios, protección contra inyecciones de datos maliciosos, testeos de penetración, redundancias para alta disponibilidad, entre otros).

Igualmente, se requerirá un plan de gestión de vulnerabilidades que prevea que el proveedor notifique prontamente cualquier vulnerabilidad o brecha de seguridad descubierta y aplique parches o soluciones en tiempo oportuno. Durante la vigencia contractual, el proveedor deberá garantizar la actualización periódica del sistema para afrontar nuevas amenazas de ciberseguridad.

El contrato también puede prever derechos del organismo público a realizar pruebas de seguridad independientes o requerir certificaciones de seguridad (p. ej. cumplimiento de ISO/IEC 27001, 27701, o evaluaciones de conformidad con esquemas de ciberseguridad nacionales).

En suma, este principio vela porque la IA opere de forma segura y confiable, protegiendo la integridad de los datos y evitando consecuencias perjudiciales por fallos técnicos o ataques externos.

La robustez, la seguridad y la confiabilidad, como conceptos técnicos, están estrechamente vinculados al principio de prevención de daños⁴¹. Es fundamental evitar y gestionar los efectos adversos no intencionales (riesgos de seguridad) y las debilidades que facilitan ataques (riesgos de protección), considerándolos,

⁴¹ Grupo de Alto Nivel sobre Inteligencia Artificial de la Comisión Europea. (2019). Directrices éticas para una IA fiable. Recuperado el 3 de junio de 2025, de <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>

previniéndolos y mitigándolos durante todas las etapas del ciclo de vida de los sistemas de IA. Esto es crucial para asegurar la tutela de los derechos de las personas⁴².

Para los sistemas de IA, la **robustez** describe su capacidad de mantener su nivel de rendimiento, según lo previsto por sus desarrolladores, bajo cualquier circunstancia⁴³. Y la **confiabilidad** se refiere a la capacidad que le permite proporcionar la predicción, recomendación y decisión requeridas de manera consistente y correcta durante su etapa de operación⁴⁴. Ambos se vinculan con la noción de **resiliencia**, que es la capacidad del sistema para recuperar la condición operativa rápidamente después de un incidente⁴⁵.

En sectores críticos, como el transporte, la salud, la educación, la seguridad y las finanzas públicas, un fallo o intrusión puede tener consecuencias muy graves para las comunidades. Por ello, las normas internacionales subrayan la necesidad de que la IA sea **segura por diseño**.

En tal contexto, la seguridad y la robustez de un sistema de IA puede evaluarse mediante cuatro criterios clave⁴⁶:

- **Resiliencia ante ataques y seguridad:** los sistemas de IA, al igual que todos los sistemas de software, deben estar protegidos contra vulnerabilidades que puedan permitir su explotación por adversarios, como el envenenamiento de datos, la fuga de modelos o la infraestructura, tanto de software como de hardware (robustez adversarial).

⁴² UNESCO. (2021). Recomendaciones sobre la Ética de la Inteligencia Artificial (p. 20). Recuperado el 3 de junio de 2025, de https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

⁴³ ISO/IEC. (2022). Information technology – Artificial intelligence – Artificial intelligence concepts and terminology (ISO/IEC Standard No. 22989).

⁴⁴ ISO/IEC. (2022). Information technology – Artificial intelligence – Artificial intelligence concepts and terminology (ISO/IEC Standard No. 22989).

⁴⁵ ISO/IEC. (2022). Information technology – Artificial intelligence – Artificial intelligence concepts and terminology (ISO/IEC Standard No. 22989).

⁴⁶ Koshiyama, A., Kazim, E., Treleaven, P., Rai, P., Szpruch, L., Pavey, G., Ahamat, G., Leutner, F., Goebel, R., Knight, A., Adams, J., Hitrova, C., Barnett, J., Nachev, P., Barber, D., Chamorro-Premuzic, T., Klemmer, K., Gregorovic, M., Khan, S., Lomas, E., Hilliard, A., & Chatterjee, S. (2024, 15 de mayo). Towards algorithm auditing: managing legal, ethical and technological risks of AI, ML and associated algorithms. Royal Society Open Science. Recuperado el 3 de junio de 2025, de <https://royalsocietypublishing.org/doi/10.1098/rsos.230859#d1e3410>

- **Plan de contingencia y seguridad general:** los sistemas de IA (y la infraestructura asociada) deben contar con salvaguardas que permitan un plan de contingencia en caso de problemas. El nivel de seguridad requerido dependerá de la magnitud del riesgo presente en el sistema.
- **Precisión:** refiere a la capacidad de un sistema de IA para emitir juicios correctos, por ejemplo, clasificar correctamente la información en las categorías adecuadas. Responde a la pregunta: ¿qué tan bien funciona el sistema? Por ejemplo, en 8 de 10 casos clasifica correctamente. Se estima cuantificando el rendimiento de generalización esperado para el sistema.
- **Fiabilidad:** un sistema de IA fiable es aquel que funciona correctamente con diversas entradas y en diversas situaciones.

Incorporar estos conceptos en la contratación colabora a que el proveedor adopte prácticas de desarrollo seguras (como pruebas exhaustivas, uso de frameworks robustos, validación en escenarios extremos) y que implemente salvaguardas contra amenazas. En la práctica, esto protege tanto al organismo público como a los ciudadanos: asegura la continuidad del servicio (evitando caídas del sistema) y la integridad de los datos y decisiones (evitando fugas o manipulaciones).

La resiliencia contra ataques busca prevenirlos, pero también tener capacidad de detectar anomalías y recuperarse de ellas. Al requerir al proveedor planes de contingencia y actualización, el contrato promueve esta resiliencia dinámica.

En esencia, una IA no es confiable si no es segura. Ningún beneficio probable o comprobado justifica exponer sistemas públicos a riesgos de ciberataques o errores que se salgan de control.

Por eso, este principio es crucial para garantizar la confiabilidad técnica, complementando otras que garantizan confiabilidad ética. Para el Sector Público, asegurar la robustez de la solución significa salvaguardar la inversión del proyecto y su sostenibilidad a largo plazo en entornos muchas veces desafiantes, con infraestructuras de IT variables, posibles intentos de fraude, entre otros.

3. Equidad, no discriminación y diversidad

Los sistemas de IA deben concebirse bajo un enfoque inclusivo que asegure la disponibilidad y accesibilidad universal de sus beneficios, atendiendo a las necesidades específicas de diversos grupos: etarios, culturales, lingüísticos, personas con discapacidad, niñas, mujeres y poblaciones desfavorecidas, marginadas o en situación de vulnerabilidad⁴⁷.

Para garantizar este objetivo, se debería exigir al proveedor que realice esfuerzos razonables para prevenir y minimizar la generación o el refuerzo de resultados discriminatorios o sesgados durante las fases de diseño, desarrollo y despliegue de la IA.

Además, se deberían establecer mecanismos efectivos para impugnar situaciones de discriminación inducidas por la IA y para identificar decisiones que impliquen un trato, impacto o representación desigual, incluso si no constituyen discriminación en términos legalmente definidos.

Antes de la entrega, el proveedor deberá evaluar exhaustivamente el sistema en busca de potenciales sesgos⁴⁸ que puedan derivar en daños de asignación o representación perjudiciales para las personas. Y si durante la operación el organismo público detecta indicios de resultados sesgados que provoquen daños de asignación o representación, debe poder requerir al proveedor la realización inmediata de las correcciones o ajustes necesarios para subsanar la situación.

⁴⁷ UNESCO. (2021). Recomendaciones sobre la Ética de la Inteligencia Artificial (p. 20). Recuperado el 3 de junio de 2025, de https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

⁴⁸ Es importante tener presente que el concepto de sesgo varía dependiendo del contexto en el que se lo utilice. En efecto, en el contexto del machine learning, el sesgo es un término que refiere a la idea de que las situaciones distintas merecen un trato diferente y es lo que permite que los sistemas de IA diferencien entre situaciones y se comporten en consecuencia (cfr. Standard ISO/IEC 22989:2022). En cambio, en el contexto social, se refiere a aquellas situaciones en las que la diferencia de trato se visualiza como injusta, lo que no equivale a sostener que es necesariamente discriminatoria.

Es fundamental reconocer que los sesgos se incorporan a los sistemas de IA a través de los datos de entrenamiento y de las múltiples decisiones de diseño e ingeniería que se adoptan a lo largo de su ciclo de vida, incluyendo la etapa de implementación⁴⁹.

En otras palabras, los equipos involucrados en el diseño, desarrollo y despliegue de sistemas de IA incorporan sus propios sesgos, tanto individuales como grupales⁵⁰. El sesgo no sólo viene dado por los datos con que se entrena un modelo, también influye en las suposiciones de los equipos acerca de qué datos utilizar, cuáles serán los objetivos del sistema, qué modelos desarrollar, dónde implementar la IA e, incluso, sobre la necesidad misma de su uso⁵¹.

La equidad en los resultados de la IA se ha erigido como una preocupación central para su adopción en el Sector Público. Un sistema que reproduzca sesgos y genere un trato desigual podría vulnerar el derecho a la igualdad y no discriminación, o conducir a resultados igualmente injustos desde una perspectiva ética⁵².

⁴⁹ Sánchez Caparrós, M. (2023). Inteligencia artificial, sesgos y categorías sospechosas. Prevenir y mitigar la discriminación algorítmica. En J. G. Corvalán (Ed.), Tratado de inteligencia artificial y derecho (Tomo I, pp. 749–774).

Sánchez Caparrós, M. (2022, julio). Los riesgos de la inteligencia artificial para el principio de igualdad y no discriminación. Planteo de la problemática y algunas aclaraciones conceptuales necesarias bajo el prisma del Sistema Interamericano de Derechos Humanos. El Dial. Recuperado el 16 de abril de 2025, de https://www.researchgate.net/publication/368202273_Los_riesgos_de_la_inteligencia_artificial_para_el_principio_de_igualdad_y_no_discriminacion_Planteo_de_la_problematika_y_algunas_aclaraciones_conceptuales_necesarias_bajo_el_prisma_del_Sistema_Interam

⁵⁰ National Institute of Standards and Technology (NIST). (2022, March). AI Risk Management Framework: Identifying, measuring, managing, and reducing harmful bias in AI (p. 5). U.S. Department of Commerce. Recuperado el 22 de abril de 2025, de <https://www.nist.gov/artificial-intelligence/ai-research-identifying-managing-harmful-bias-ai>

⁵¹ National Institute of Standards and Technology (NIST). (2022, March). AI Risk Management Framework: Identifying, measuring, managing, and reducing harmful bias in AI (p. 5). U.S. Department of Commerce. Recuperado el 22 de abril de 2025, de <https://www.nist.gov/artificial-intelligence/ai-research-identifying-managing-harmful-bias-ai>.

⁵² Sánchez Caparrós, M. (2023). Principios éticos para una inteligencia artificial antropocéntrica: consensos actuales desde una perspectiva global y regional. En J. G. Corvalán (Ed.), Tratado de inteligencia artificial y derecho (Tomo I, pp. 459–510).

Sánchez Caparrós, M. (2022, julio). Los riesgos de la inteligencia artificial para el principio de igualdad y no discriminación. Planteo de la problemática y algunas aclaraciones conceptuales necesarias bajo el prisma del Sistema Interamericano de Derechos Humanos. El Dial. Recuperado el 16 de abril de 2025, de https://www.researchgate.net/publication/368202273_Los_riesgos_de_la_inteligencia_artificial_para_el_principio_de_igualdad_y_no_discriminacion_Planteo_de_la_problematika_y_algunas_aclaraciones_conceptuales_necesarias_bajo_el_prisma_del_Sistema_Interam

Organismos internacionales como la UNESCO, en su Recomendación sobre la Ética de la IA (2021), y los Principios de la OCDE, enfatizan la no discriminación y la equidad como principios rectores fundamentales para el ciclo de vida de la IA, instando a una gestión proactiva de los riesgos de sesgo y a la promoción de la inclusión.

La inclusión de pautas específicas en la documentación contractual traduce estos valores en acciones concretas, definiendo formal y claramente las funciones, obligaciones y metodologías aplicables en cada fase del ciclo de vida de la IA.

A modo de ejemplo, según el caso, el contrato debería especificar: criterios para el mapeo y medición de sesgos; protocolos de validación técnica; mecanismos de supervisión legal; cronograma de auditorías periódicas; políticas de actualización de modelos; y planes de contingencia para riesgos operativos.

Además, debería exigir que la documentación del modelo incluya descripciones claras de su funcionamiento, en la medida de lo razonable según las técnicas de IA empleadas, facilitando al organismo público la evaluación adecuada del riesgo de sesgo. Esta documentación actuará como una fuente centralizada de información para la toma de decisiones y la resolución de conflictos.

Este tipo de cláusulas también promueven la transparencia algorítmica al obligar al proveedor a documentar los sesgos conocidos y las medidas de mitigación implementadas, generando visibilidad sobre las limitaciones del sistema. Aunque algunos proveedores pueden argumentar la confidencialidad de sus modelos, las condiciones contractuales orientadas a la equidad les exigirán proporcionar evidencia si el organismo público lo requiere.

Para el Sector Público, es crucial que los sistemas de IA que financia o desarrolla no perpetúen ni agraven brechas de género, étnicas o socioeconómicas, ni generen un trato desigual e ilegítimo entre las personas.

4. Sostenibilidad y bienestar

Para que una sociedad sea sostenible, se necesita avanzar objetivos relacionados con distintas dimensiones: humanas, sociales, culturales, económicas y ambientales. En este sentido, la OCDE ha reconocido que la IA tiene el potencial de mejorar el bienestar de las personas, contribuir a una actividad económica mundial positiva y sostenible, aumentar la innovación y la productividad y ayudar a responder a desafíos mundiales clave⁵³.

Por tal motivo, destaca al desarrollo sostenible y bienestar como un principio ético para una IA fiable. Y resalta el deber de las partes interesadas de “...poner en marcha activamente la administración responsable en aras de una IA fiable en busca de resultados beneficiosos para las personas y el planeta, como aumentar las capacidades del ser humano y estimular la creatividad, avanzar en la inclusión de poblaciones infrarrepresentadas, reducir las desigualdades económicas, sociales, de género y de otra índole, y proteger los entornos naturales, reforzando de este modo el crecimiento inclusivo, el bienestar, el desarrollo sostenible y la sostenibilidad medioambiental...”⁵⁴.

Los sistemas de IA, sin dudas, pueden ayudar a lograr comunidades más sostenibles, pero también pueden ser un obstáculo, dependiendo de cómo se usen y del contexto de estado de desarrollo de cada país⁵⁵.

Por eso, es fundamental evaluar de forma constante cómo impacta la IA en todas estas dimensiones. Esa evaluación debe hacerse con una mirada amplia, entendiendo que la sostenibilidad no es un objetivo fijo, sino un conjunto de metas que evolucionan con el tiempo, como ocurre con los Objetivos de Desarrollo Sostenible (ODS) que propone la ONU⁵⁶.

⁵³ OECD. (2024). Recommendation of the Council on Artificial Intelligence. Recuperado el 22 de abril de 2025, de <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449>

⁵⁴ OECD. (2024). Recommendation of the Council on Artificial Intelligence. Recuperado el 22 de abril de 2025, de <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449>

⁵⁵ UNESCO. (2021). Recomendaciones sobre la Ética de la Inteligencia Artificial (p. 20). Recuperado el 3 de junio de 2025, de https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

⁵⁶ UNESCO. (2021). Recomendaciones sobre la Ética de la Inteligencia Artificial (p. 20). Recuperado el 3 de junio de 2025, de https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

La relevancia de este principio se observa de cara a la contratación de sistemas de IA por parte del Sector Público. La Administración Estatal, al adquirir o desarrollar sistemas de IA, tiene la responsabilidad de asegurar que la inversión no solo sea eficiente y cumpla objetivos específicos, sino que también contribuya al bienestar general y al progreso hacia una sociedad más justa y equitativa.

En este contexto, el principio de sostenibilidad y bienestar se convierte en un faro que ha de guiar la decisión en materia de compra pública de sistemas de IA; e implica que la contratación de sistemas de IA debe ir más allá de la mera optimización de procesos internos o la reducción de costos. Debe enfocarse también en la capacidad de la IA para generar valor público genuino, abordando desafíos sociales, económicos y ambientales de manera integral.

La contratación pública de IA debe priorizar aquellos sistemas que demuestren un potencial claro para beneficiar a las personas, especialmente a los grupos marginados o en riesgo de exclusión. Esto significa buscar soluciones de IA que puedan mejorar el acceso a servicios esenciales como salud, educación y justicia, que faciliten la inclusión laboral y económica, y que contribuyan a la construcción de comunidades más resilientes y sostenibles.

De esta manera, el Sector Público puede asegurarse de que la IA se convierta en una herramienta con potencial para reducir desigualdades, proteger el medio ambiente y fomentar un crecimiento inclusivo, que marque una diferencia tangible en la vida de los ciudadanos y permita avanzar hacia un futuro más próspero y sostenible para todos.

Incorporar cláusulas que aborden la integración de este principio en el contrato colaborará para que los sistemas de IA que se contraten contribuyan activamente al bienestar de las personas y al desarrollo sostenible.

5. Privacidad y protección de datos personales

La privacidad es un derecho fundamental para la salvaguarda de la dignidad, la autonomía y la libertad de acción de las personas, por lo que debe garantizarse durante todo el ciclo de vida de los sistemas de IA.

En ese contexto, el manejo de los datos de los sistemas de IA —desde su recopilación y procesamiento hasta su almacenamiento y eliminación— debería alinearse con las normas del derecho internacional y respetar los valores y principios establecidos en documentos éticos como la Recomendación de UNESCO y de la OCDE.

Dado que los sistemas de IA pueden implicar el tratamiento de datos personales, el contrato debería incluir una cláusula que asegure el estricto cumplimiento de la normativa de protección de datos vigente en la jurisdicción del organismo contratante y de los estándares internacionales que se estimen relevantes.

El proveedor debe garantizar que todo tratamiento de datos personales (recolección, uso, divulgación, archivo y supresión) se realice con un objetivo legítimo y una base jurídica válida.

Adicionalmente, se exigirá la implementación de medidas técnicas y organizativas de protección de la privacidad, tales como la anonimización o seudonimización cuando sea pertinente, el control de accesos y la eliminación segura de los datos una vez cumplida su finalidad.

Los sistemas de IA podrán requerir evaluaciones adecuadas del impacto en la privacidad, así como la integración de la privacidad desde el diseño.

El proveedor y el organismo público asumirán la responsabilidad conjunta de concebir y aplicar los sistemas de IA de manera que se garantice la protección de la información personal durante todo su ciclo de vida.

Para ello, el contrato debería estipular que el proveedor ofrezca apoyo en la realización de evaluaciones de impacto en la privacidad, si el proyecto lo requiere, y que colabore

con el organismo público en el ejercicio de los derechos ARCO (Acceso, Rectificación, Cancelación y Oposición) y otras obligaciones derivadas de la legislación sobre datos personales.

La relevancia de cláusulas que integren este principio en la contratación de sistemas de IA radica en la necesidad legal de garantizar el cumplimiento normativo y, desde una perspectiva ética, en la importancia de fomentar la confianza en el uso de los datos personales.

En la práctica, impulsa al proveedor a integrar la privacidad desde el diseño (privacy by design), incluyendo la anonimización o la eliminación segura de la información personal cuando ya no sea necesaria. También proporciona al organismo público herramientas contractuales para exigir transparencia sobre el uso de los datos e intervenir ante incidentes como brechas de seguridad.

Adicionalmente, cuando se estime pertinente, el contrato podría exigir que tanto el proveedor como el organismo público cuenten con personal especializado en protección de datos (como un Delegado de Protección de Datos - DPD) y establezcan programas continuos de formación y concientización para todo el personal involucrado en el manejo y supervisión de los sistemas de IA.

6. Supervisión y decisión humana

Al contratar sistemas de IA, el Sector Público debe garantizar que en cualquier etapa de su ciclo de vida sea posible asignar responsabilidades éticas y legales a personas humanas o jurídicas⁵⁷. Para ello, es crucial establecer contractualmente los mecanismos concretos de supervisión humana a lo largo del ciclo de vida del sistema.

Esto implica definir claramente los roles y responsabilidades de las personas designadas para la supervisión, especificando los puntos de intervención obligatorios y los protocolos de actuación ante posibles desviaciones o resultados inesperados.

⁵⁷ UNESCO. (2021). Recomendaciones sobre la Ética de la Inteligencia Artificial. Recuperado el 3 de junio de 2025, de https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

El contrato debería detallar los procesos para la revisión y validación de las salidas del sistema de IA, y prever mecanismos de auditoría y trazabilidad que permitan verificar la efectividad de la supervisión humana y la capacidad de rendición de cuentas, especialmente en aquellos casos con potencial impacto significativo en los derechos de los ciudadanos o en sectores sensibles del Estado (ej. salud o educación).

La supervisión humana no se limita a los individuos, sino que también puede involucrar mecanismos de supervisión pública colectiva, cuando sea pertinente.

Adicionalmente, y esto resulta crítico, si bien en ciertos casos los humanos pueden optar por delegar funciones en sistemas de IA para ganar eficiencia, la autoridad última sobre dicha delegación siempre debería recaer en personas humanas⁵⁸.

La supervisión humana, además, "...incluye abordar riesgos derivados de usos distintos del fin perseguido, usos indebidos intencionados o usos indebidos no intencionados, de forma apropiada al contexto y coherente con el estado actual de la tecnología..."⁵⁹. El principio de supervisión humana procura que los sistemas de IA operen siempre bajo la vigilancia y dirección de las personas.

La implementación efectiva de la supervisión humana requiere considerar la capacitación y las herramientas proporcionadas a los supervisores. Estos deben poseer la comprensión necesaria del funcionamiento del sistema de IA, sus limitaciones y los posibles riesgos éticos y legales asociados.

Así el contrato debería exigir al proveedor la entrega de documentación clara y accesible sobre el diseño, la lógica y los datos utilizados por el sistema, adaptado a las particularidades del sistema que se contrata, así como la formación adecuada para el personal del Sector Público encargado de la supervisión.

⁵⁸ UNESCO. (2021). Recomendaciones sobre la Ética de la Inteligencia Artificial. Recuperado el 3 de junio de 2025, de https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

⁵⁹ UNESCO. (2021). Recomendaciones sobre la Ética de la Inteligencia Artificial. Recuperado el 3 de junio de 2025, de https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

7. Transparencia y explicabilidad

Se exigirá que la solución de IA cumpla con los principios de transparencia y explicabilidad, conforme a las definiciones establecidas en la Recomendación sobre Ética de la IA de la UNESCO.

La **transparencia** procura que los usuarios dispongan de información adecuada para comprender el funcionamiento general de la solución, mientras que la **explicabilidad** persigue la inteligibilidad de las entradas, salidas y el funcionamiento de cada componente algorítmico, así como su contribución a los resultados específicos. El proveedor deberá diseñar y desarrollar la solución de IA de modo que la entidad pública pueda entender, dentro de un marco razonable y atendiendo al estado actual de la tecnología, cómo opera el sistema y cómo se generan sus decisiones.

Esto puede incluir, a modo de ejemplo y cuando corresponda de acuerdo con las características del sistema contratado, que antes de la entrega el proveedor documente detalladamente el comportamiento del modelo, los datos empleados y las variables más influyentes en los resultados.

Asimismo, comprende la proporción de documentación técnica adecuada y guías de uso. Esto permitirá al organismo público informar a ciudadanos o partes interesadas, de manera clara, sobre el funcionamiento y las decisiones del sistema de IA.

Es importante también incluir una cláusula de “explicabilidad bajo demanda”, que obligue al proveedor a asistir al organismo público —durante la vigencia del contrato— en la interpretación de cualquier resultado específico, identificando los factores determinantes y los posibles cambios en las entradas que puedan haber modificado la salida.

En definitiva, el contrato debe asegurar los derechos de la entidad contratante a recibir información clara sobre el funcionamiento general (transparencia) y a comprender las decisiones particulares (explicabilidad) del sistema de IA, siempre dentro de un marco razonable y en la medida de las posibilidades técnicas del sistema, para mitigar el

efecto “caja negra”. Estos principios fortalecen la confianza pública y la rendición de cuentas, y alinean el contrato con estándares éticos internacionales.

Como señala el documento de Recomendaciones de la UNESCO, la transparencia y la explicabilidad son condiciones previas esenciales para garantizar el respeto, la protección y la promoción de los derechos humanos, las libertades fundamentales y los principios éticos⁶⁰.

Desde una perspectiva sociotécnica, un mayor grado de transparencia fomenta la construcción de sociedades más pacíficas, equitativas, democráticas e inclusivas.

La transparencia colabora para que las personas comprendan cómo se implementa cada fase de un sistema de IA, según el contexto y la sensibilidad de la solución en particular.

En ese marco, también puede brindar información acerca de los factores que influyen en las decisiones del sistema de IA, así como sobre la existencia o no de medidas de seguridad y equidad que pudiera haber tomado el proveedor. Si un organismo público no comprende cómo opera una IA que utiliza, difícilmente podrá responder por sus decisiones o corregir errores, socavando la confianza ciudadana.

La explicabilidad, por su parte y con el alcance señalado, se vincula estrechamente con la transparencia. y apunta a que los resultados y subprocesos que han conducido a esos resultados sean, en la medida de lo posible, comprensibles y trazables, de una forma apropiada al contexto de aplicación. Por tal motivo, los proveedores deberían garantizar la proporción de una explicación satisfactoria de las decisiones que hayan dado lugar a las acciones tomadas, a fin de transparentar el resultado. Sobre todo cuando se trata de decisiones no temporales o difícilmente reversibles.

Adicionalmente a las Recomendaciones de UNESCO, otros estándares internacionales relevantes, como de la OCDE (2019, revisados en 2024), resaltan la relevancia de ambos principios en cuanto instan a que los actores de la IA brinden información

⁶⁰ Ibidem; págs 22-23.

significativa, adecuada para la situación específica y ajustada al estado actual de la tecnología, de cara a fomentar la comprensión general de los sistemas de IA; el conocimiento por las partes interesadas acerca de sus interacciones con los sistemas de IA; en la medida que sea viable y útil, la proporción de información sobre fuentes de datos, factores, procesos y/o razonamiento subyacente a las predicciones, contenidos, recomendaciones y decisiones del sistema de IA, para que el afectado pueda comprender los resultados; el suministro de información que permita la impugnación de resultado del sistema de IA por parte de quien se ha visto afectado negativamente. los sistemas de IA sean transparentes y explicables en la medida necesaria para que tanto usuarios como supervisores entiendan sus resultados.

El nivel de transparencia y explicabilidad debe ajustarse siempre al tipo de sistema de IA, contexto de uso e impacto del sistema en los derechos de las personas, dado que puede requerirse un balance entre estos aspectos y otros principios, como la privacidad, la seguridad y la protección.

Incluir en el contrato cláusulas de transparencia y explicabilidad adaptadas traduce esos principios en obligaciones concretas hacia el proveedor. Su importancia práctica radica en posibilitar que funcionarios públicos expliquen decisiones automatizadas a los ciudadanos afectados, satisfaciendo expectativas de transparencia e incluso disposiciones legales emergentes sobre el “derecho a explicación”.

En síntesis, la transparencia y la explicabilidad se vinculan directamente con la rendición de cuentas y la responsabilidad, a la vez que permiten el escrutinio público y técnico, indispensables también para el debido proceso cuando las decisiones automatizadas inciden en derechos. Sin estas cláusulas, la solución elegida un sistema opaco podría generar arbitrariedad o indefensión, mientras que con ellas se avanza hacia una IA ética y confiable, que refuerza la legitimidad de su uso en lo público.

8. Responsabilidad y rendición de cuentas

Los proveedores y organismos públicos deben asumir la “...responsabilidad ética y jurídica respectiva, de conformidad con el derecho nacional e internacional, en

particular las obligaciones de los Estados Miembros en materia de derechos humanos, y con las directrices éticas establecidas durante todo el ciclo de vida de los sistemas de IA, incluso con respecto a los actores de la IA dentro de su territorio y bajo su control efectivos...”⁶¹.

En ese marco, la definición clara de responsabilidades contractuales constituye un pilar fundamental de la rendición de cuentas en proyectos de IA. Dado que los sistemas de IA pueden provocar daños y afectar los derechos de las personas, es esencial establecer de antemano quién asume qué tareas y qué obligaciones. Como señala UNESCO, la “...responsabilidad ética y la obligación de rendir cuentas de las decisiones y las acciones basadas de alguna manera en un sistema de IA siempre deberían ser atribuibles, en última instancia, a los actores de la IA conforme a la función que tengan en el ciclo de vida del sistema de IA...”⁶².

Por ello, los organismos públicos y los proveedores deben detallar sus respectivas responsabilidades en el contrato, ya que las normas sobre asignación de responsabilidad pueden variar según la jurisdicción, pero también de cara a la contratación de los distintos sistemas de IA.

El contrato debe definir con claridad las obligaciones del proveedor del sistema de IA y del organismo público contratante. Debe incluir disposiciones que establezcan las respectivas responsabilidades y un marco general de rendición de cuentas, en el cual el proveedor se compromete a entregar un sistema de IA que cumpla con todos los requisitos estipulados (legales, técnicos y éticos), y a cumplir las obligaciones a su cargo durante toda la vigencia del contrato.

Por su parte, el organismo público debe comprometerse a usar el sistema conforme a la finalidad prevista y a facilitar, cuando sea necesario, información o recursos al proveedor para asegurar el correcto funcionamiento (por ejemplo, datos de entrenamiento pertinentes o feedback sobre el desempeño). Se especificará qué

⁶¹ UNESCO. (2021). Recomendaciones sobre la Ética de la Inteligencia Artificial. Recuperado el 3 de junio de 2025, de https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

⁶² Ibidem.

personal o unidades de cada parte asumirán responsabilidades específicas (ej. un responsable técnico y uno legal por parte del proveedor), de modo que exista una asignación clara y transparente de roles en la gestión del sistema de IA.

Para garantizar un adecuado seguimiento del comportamiento de la IA, se podrán incluir cláusulas que habiliten la auditoría y supervisión continua del sistema, sobre todo frente a sistemas de alto impacto.

La auditoría es el proceso sistemático e independiente para obtener evidencia y evaluarla objetivamente para determinar en qué medida se cumplen los criterios de auditoría. Puede ser interna, externa o combinada, y conducida por la organización o terceras partes⁶³.

También se podrá prever que el proveedor conserve registros detallados de la operación del sistema (datos de entrada, salidas, decisiones, fechas, etc.) y ante la solicitud del organismo público, deba facilitar dichos registros sin demora.

El contrato otorgará al organismo público el derecho a auditar el sistema y verificar su conformidad con los requisitos acordados en cualquier momento razonable durante la vigencia contractual.

Esto implica que el proveedor debe cooperar plenamente con auditorías, ya sean realizadas por el personal público o por auditores independientes designados, proporcionando toda la información necesaria y acceso a sus instalaciones o sistemas relacionados. Igualmente, se podrá pactar la realización de auditorías periódicas (por ejemplo, anuales) o tras cualquier modificación sustancial del sistema.

De detectarse incumplimientos o riesgos durante una auditoría, el proveedor estará obligado a implementar medidas correctivas en plazos definidos en el contrato, bajo apercibimiento de incurrir en incumplimiento contractual.

⁶³ ISO/IEC. (2022). Information technology – Artificial intelligence – Artificial intelligence concepts and terminology (ISO/IEC Standard No. 22989). p 3.

Cláusulas que incorporen las pautas señaladas aseguran un monitoreo activo de la IA y la posibilidad de corregir rumbos ante desviaciones.

Al especificar obligaciones como la entrega de resultados, el mantenimiento, el soporte o el uso adecuado, se establece un marco bilateral de rendición de cuentas que motiva a cada parte a cumplir su rol. Ello se alinea con el principio de responsabilidad de la OCDE, que promueve una distribución clara de la rendición de cuentas entre desarrolladores, proveedores y usuarios de sistemas de IA⁶⁴.

En la práctica, esta claridad reduce ambigüedades: el proveedor comprende qué se espera de él más allá de entregar un sistema de IA (p. ej., colaborar en auditorías, corregir fallos, asesorar en la implementación), mientras que el organismo público reconoce sus compromisos (p. ej., proporcionar datos necesarios, designar personal capacitado para la supervisión, emplear la solución conforme a las instrucciones).

Adicionalmente, responde a la necesidad de confianza mutua: un contrato con responsabilidades bien definidas fomenta la confianza entre las partes y hacia terceros —como la ciudadanía o los entes de control—, al evidenciar un esquema sólido de gobernanza del proyecto de IA.

Por otro lado, la capacidad de auditar y monitorear un sistema de IA durante su uso es esencial desde la perspectiva de la gestión de riesgos continua y se sustenta en la noción de “control posterior”. Incluso después de implementado, un sistema debe estar sujeto a verificación y mejora. Esto responde a la realidad de que los sistemas de IA pueden degradar su rendimiento con el tiempo (por cambios en los datos o el entorno) o presentar comportamientos inesperados no detectados en las pruebas iniciales.

En línea con ello, la OCDE recomienda garantizar la trazabilidad de los conjuntos de datos, los procesos y las decisiones que se toman durante el ciclo de vida del sistema de

⁶⁴ OECD. (2024). Recommendation of the Council on Artificial Intelligence. Recuperado el 22 de abril de 2025, de <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449> (p.5)

IA, para permitir el análisis de los resultados del sistema y garantizar que las respuestas sean adecuadas al contexto específico y coherentes con el estado de la tecnología⁶⁵.

A ese fin, los mecanismos de auditoría y monitoreo permiten al organismo público inspeccionar su funcionamiento, verificar el cumplimiento de las métricas acordadas (ej. tasa de acierto, tiempos de respuesta, etc.) y detectar posibles fallos emergentes. También facilitan evaluaciones externas e independientes, que aportan objetividad en el escrutinio (ej. auditorías algorítmicas realizadas por universidades o entes reguladores, cuando proceda).

Por otro lado, el monitoreo activo permite cumplir con el principio de no ocasionar daño. Si se advierte a tiempo que el sistema está causando efectos adversos, el organismo puede suspender su uso o retirarlo.

En suma, este tipo de disposiciones tienen como finalidad asegurar un monitoreo activo y continuo del sistema de IA, que permita al organismo público supervisar su comportamiento, evaluar su rendimiento, verificar el cumplimiento de los estándares éticos, técnicos, legales y de protección de datos acordados, y corregir o mitigar desviaciones y riesgos de manera oportuna. La supervisión debería incluir la evaluación de la efectividad de las medidas de seguridad, la transparencia del funcionamiento, la explicabilidad de las decisiones (en la medida de lo posible), y el impacto del sistema en los derechos y libertades de las persona

El organismo público podrá utilizar la información obtenida a través de la auditoría y la supervisión para tomar decisiones informadas sobre la continuidad del contrato y la implementación de mejoras en el sistema.

9. Concientización y alfabetización

En el marco de la contratación de sistemas de inteligencia artificial por parte del Sector Público, es esencial garantizar que los operadores estatales cuenten con una formación adecuada para utilizar de manera responsable, ética, efectiva y segura los sistemas

⁶⁵ OECD. (2024). Recommendation of the Council on Artificial Intelligence. Recuperado el 22 de abril de 2025, de <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449> (p.5)

adquiridos, y que los funcionarios responsables de su contratación comprendan los principios, riesgos y potencialidades de los sistemas que incorporan al Estado.

Del mismo modo, se necesita que la ciudadanía tenga acceso a información clara y accesible sobre el funcionamiento y los impactos de estos sistemas en su vida cotidiana. Esta formación no debe ser uniforme, sino que debe contemplar distintos niveles, técnico, no técnico e interdisciplinario, para responder a la diversidad de roles, formación y responsabilidades de los actores involucrados, contemplando la brecha digital existente hacia el interior de cada organización, así como de las comunidades de cada jurisdicción.

En dicho contexto, es necesario promover espacios de educación abierta, participación ciudadana y capacitación continua, orientados a desarrollar competencias digitales, alfabetización en la técnica y en la ética de la IA.

Estos procesos formativos deben ser impulsados por el Estado en articulación con universidades, organizaciones de la sociedad civil, organismos internacionales, medios de comunicación y actores del sector privado, y considerar la pluralidad lingüística, social y cultural de las comunidades, para garantizar una participación democrática informada y una apropiación crítica de la tecnología.

La capacitación debe también incorporar enfoques de diseño y desarrollo de sistemas de IA centrados en las personas, reconociendo el carácter sociotécnico de estas tecnologías⁶⁶.

Esto implica no solo comprender sus componentes técnicos de los sistemas de IA, sino también comprender su impacto en los derechos humanos, la equidad, el acceso a servicios públicos y el entorno social y ambiental.

⁶⁶ Sánchez Caparrós, M. (2023). Principios éticos para una inteligencia artificial antropocéntrica: consensos actuales desde una perspectiva global y regional. En J. G. Corvalán (Ed.), Tratado de inteligencia artificial y derecho (T. I, pp. 459–510).

El aprendizaje sobre IA en el ámbito estatal debe, por tanto, estar orientado a garantizar que las decisiones tecnológicas se alineen con los principios de justicia, inclusión y sostenibilidad.

En definitiva, formar a los distintos actores del ecosistema público en estas dimensiones es clave para que las políticas de contratación e implementación de sistemas de IA sean responsables, legítimas y alineadas con el interés público, evitando usos opacos o discriminatorios y fortaleciendo la confianza de la ciudadanía en la administración digital.

Una estrategia integral de formación en el contexto de la contratación pública de IA debe concebirse como un eje transversal que permee todas las etapas del ciclo de vida de los sistemas, desde la identificación de necesidades hasta la evaluación de su impacto y su eventual desincorporación. Esto implica no solo la capacitación inicial de los operadores y funcionarios, sino también el desarrollo de programas de formación continua que permitan adaptarse a la rápida evolución de la tecnología y a la aparición de nuevos riesgos y oportunidades.

Además, la estrategia debe considerar la creación de comunidades de práctica y el intercambio de conocimientos entre diferentes organismos públicos, así como la promoción de la investigación aplicada y la colaboración con centros de excelencia en IA para mantener actualizadas las competencias y los enfoques pedagógicos.

Finalmente, la sostenibilidad de la formación en IA para el Sector Público requiere la institucionalización de mecanismos que garanticen su continuidad y su adaptación a largo plazo. Esto podría incluir la designación de responsables de formación en IA dentro de los organismos, la asignación de recursos presupuestarios específicos y la integración de la formación en los planes estratégicos de transformación digital.

Asimismo, es fundamental fomentar una cultura de aprendizaje continuo y de actualización proactiva de conocimientos, tanto a nivel individual como organizacional, para asegurar que el Sector Público esté siempre a la vanguardia en la comprensión y la gestión de las tecnologías de IA.

La integración de este principio en el contexto de la compra pública de sistemas de IA busca asegurar que el proveedor capacite al organismo público para utilizar el sistema de manera adecuada, promoviendo su alineación con principios éticos y el interés público, mientras el organismo público asume la responsabilidad de profundizar y sostener dichas competencias.

10. Colaboración de múltiples partes interesadas

La contratación de sistemas de IA de alto impacto por parte del Sector Público debería prever mecanismos de participación significativa y continua de múltiples partes interesadas a lo largo del ciclo de vida del sistema, desde el diseño de los términos de referencia hasta su implementación, monitoreo y eventual retiro.

Esta participación es relevante para asegurar una gobernanza inclusiva, contextualizada y ética, que contribuya a maximizar los beneficios sociales del sistema y mitigar sus riesgos, en línea con los objetivos de desarrollo sostenible y la protección de los derechos fundamentales.

Se considerarán como partes interesadas relevantes a los organismos públicos convocantes, organismos de control, organizaciones de la sociedad civil, entidades académicas y técnicas, defensorías, organismos de protección de derechos, pueblos indígenas, representantes de comunidades vulnerables y otros actores pertinentes según el contexto de aplicación del sistema.

En los procesos de contratación de sistemas de IA de alto impacto, debería procurarse:

- **La identificación temprana de grupos afectados:** desde la fase inicial, se deben identificar los individuos, comunidades o sectores que puedan ser impactados directa o indirectamente por el sistema, con especial atención a grupos vulnerables o históricamente marginados.
- **La consulta abierta y documentada:** se establecerán procesos de consulta transparentes en etapas críticas, incluyendo el diseño de pliegos, la evaluación de impacto ético, la definición de criterios de supervisión y el monitoreo

post-implementación. Estas consultas deben ser accesibles, documentadas y sus resultados deben integrarse en la toma de decisiones.

- **La revisión interdisciplinaria:** antes de la adjudicación, se implementarán mecanismos formales de evaluación liderados por comités interdisciplinarios que analicen riesgos éticos, sociales y técnicos, proponiendo salvaguardas específicas.
- **La adaptación al contexto sociocultural y tecnológico:** los procesos de participación deben diseñarse considerando las particularidades culturales, lingüísticas y tecnológicas del entorno, asegurando accesibilidad (por ejemplo, traducción a lenguas indígenas, formatos accesibles para personas con discapacidad) y estrategias inclusivas para comunidades marginadas.
- **La actualización y trazabilidad:** los mecanismos de participación deben revisarse y actualizarse periódicamente para adaptarse a la evolución tecnológica, los impactos emergentes del sistema y las necesidades cambiantes de las partes interesadas. Todo el proceso debe ser transparente, con registros públicos que garanticen su trazabilidad.
- **La capacitación y sensibilización:** se promoverán actividades de formación dirigidas a las partes interesadas, especialmente a comunidades menos familiarizadas con la IA, para garantizar una participación informada y efectiva.

Los documentos contractuales, cuando corresponda, deberán incluir cláusulas específicas que exijan a los proveedores de IA demostrar cómo integrarán la participación de las partes interesadas en el desarrollo e implementación del sistema.

Además, se recomienda establecer indicadores de cumplimiento para evaluar la efectividad de estos procesos participativos, como la cantidad de consultas realizadas, la diversidad de actores involucrados y la incorporación de sus aportes en el diseño final.

En sistemas clasificados como de impacto medio o bajo, el organismo público podrá disponer, justificadamente, la implementación matizada y proporcional de mecanismos de participación cuando lo considere pertinente en función del contexto, los grupos potencialmente afectados o los objetivos de la contratación.

A modo de ejemplo, en un asistente virtual de atención ciudadana se podría implementar una mesa de prueba con representantes de adultos mayores, personas con discapacidad o migrantes para validar la accesibilidad de las respuestas.

En conclusión, la participación inclusiva y continua no solo fortalece la legitimidad y confianza en los sistemas de IA desplegados por el Sector Público, sino que también asegura que estos respondan a las necesidades reales de la sociedad, respetando los principios éticos y los derechos fundamentales.

V. Gestión de tensiones entre principios

Es importante reconocer que la aplicación de los principios éticos en la contratación de sistemas de inteligencia artificial (IA) puede dar lugar a situaciones en las que, por ejemplo, la búsqueda de transparencia colisione con la protección de la privacidad o la demanda de eficiencia entre en tensión con la obligación de garantizar equidad.

Es decir, a situaciones de tensión entre principios éticos que será necesario dirimir de una manera que preserve el enfoque de derechos y la rendición de cuentas, incluso en contextos de incertidumbre.

Esta tarea interpretativa debería intentar lograr una coherencia sistémica, evitando que la aplicación de un principio ético desplace por completo o anule otro igualmente relevante para el caso de uso y contexto de aplicación.

En ese marco, una buena técnica interpretativa consiste en evitar un excesivo rigor que desnaturalice el espíritu que inspira cada principio ético o que conduzca a la pérdida de derechos de las personas.

Ello así, se debe procurar armonizar los principios entre sí dentro del espíritu general que les dio origen, asegurando que su aplicación contribuya al fortalecimiento de un sistema de contratación pública justo, ético, transparente, inclusivo y respetuoso de los derechos humanos.

A la luz de estas pautas, cada organismo público debería prever cómo dirimir estas tensiones atendiendo a sus particularidades organizativas; al nivel de impacto del sistema de IA sobre los distintos sectores de gobierno y sobre los derechos de las personas; así como al tipo de sistema de IA contratado (por ejemplo, solución a medida, licencia o software potenciado por IA).

Los pliegos de bases y condiciones particulares deberían reconocer de forma explícita si en la contratación pueden existir posibles conflictos entre principios, y prever qué autoridades serán responsables de tomar decisiones ante dilemas, atendiendo a la modalidad de contratación, así como al nivel y área de impacto del sistema. A modo de ejemplo, se podrá ponderar el grado de afectación de derechos fundamentales, la sensibilidad de la información manejada, las restricciones presupuestarias y la criticidad del servicio o área de impacto del sistema de IA.

Es deseable que la toma de decisiones de este tipo sea asumida por un comité de conformación interdisciplinaria —integrado por áreas técnica, jurídica, de derechos humanos y de seguridad— encargado de analizar los valores en disputa, evaluar los riesgos y emitir una recomendación motivada sobre la priorización de principios, que se incorporará al pliego de bases y condiciones particulares.

El razonamiento que sustente cada elección debería quedar registrado en un anexo de decisiones, donde se documente el proceso de deliberación y se detalle el fundamento de la opción adoptada.

VI. Guía de estándares orientados a la Contratación Pública de sistemas de IA confiables, éticos y centrados en las personas

Considerando los principios éticos señalados, adaptados al contexto de la compra pública de sistemas de IA, se propone una **Guía de Estándares para la Contratación Pública de sistemas de IA Confiables, Éticos y Centrados en las Personas** (ver el Anexo I), que se compone de:

1. Un modelo de pliego de bases y condiciones generales.
2. Un menú de cláusulas contractuales base orientadas a la implementación de principios éticos en la compra pública de sistemas de IA.

Los documentos señalados operan como un **marco inicial de gobernanza y orientador para la contratación pública de sistemas de IA**, que busca promover su desarrollo, adquisición e implementación de forma ética, segura, centrada en las personas y respetuosa de los derechos humanos.

En esa línea, el pliego y los estándares sugeridos **no deben entenderse como estructuras rígidas**, sino como herramientas adaptables a la realidad de cada organización en pos de canalizar y aprovechar el potencial de la IA sin perder de vista la necesaria rendición de cuentas, la transparencia y la protección de los derechos de las personas y de los valores democráticos.

Por eso, considerando la diversidad de escenarios posibles en el ámbito de la compra pública de sistemas que incorporan IA —desde licencias de software comercial con componentes inteligentes hasta desarrollos a medida de sistemas de alto impacto y agentes de IA—, los estándares propuestos **contemplan márgenes de flexibilidad y adaptación**.

La posibilidad de adecuar esos estándares a las particularidades de distintas contrataciones en función del tipo de sistema (ej. a medida, licenciado o apoyado en modelos de propósito general de terceros), así como al grado de impacto de los

sistemas de IA sobre los derechos de las personas o áreas de aplicación, resulta clave para **evitar la sobrerregulación o la imposición de requisitos desproporcionados**.

La sobrerregulación podría generar requisitos excesivamente prescriptivos y engorrosos que terminen sofocando la innovación, limitando la adopción de soluciones de IA beneficiosas y aumentando los costos y la complejidad de los procesos de compra pública⁶⁷.

Asimismo, un enfoque regulatorio demasiado restrictivo puede llevar a que las entidades públicas se queden rezagadas en la adopción de tecnologías de punta, perdiendo oportunidades para mejorar la eficiencia, la eficacia y la calidad de los servicios que prestan a los ciudadanos⁶⁸.

Por el contrario, la **perspectiva flexible** propiciada desde este documento favorece la **innovación responsable**, reduce barreras de entrada para proveedores de diversos tamaños y asegura que el procedimiento de contratación sea ágil y efectivo, sin renunciar a los principios éticos que deben guiar el uso de la IA en el Sector Público.

En otras palabras, es crucial que los gobiernos encuentren un **equilibrio cuidadoso entre mitigar los riesgos potenciales de la IA y fomentar un ecosistema que promueva la innovación y la competencia** en el mercado de soluciones de IA.

Con ese propósito, **los documentos propuestos en el Anexo I contemplan explícitamente la posibilidad de flexibilizar sus cláusulas**, especialmente en aquellos casos donde no se trate de sistemas de IA de **alto impacto**⁶⁹.

⁶⁷OHCHR. (2025). Submission on the use of artificial intelligence: State submission – Lithuania. Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos. <https://www.ohchr.org/sites/default/files/documents/issues/tech/ai-submissions/subm-use-artificial-intelligence-sta-lithuania.docx>

⁶⁸ NIGP. (s.f.). Artificial intelligence: The impact on public procurement – Opportunity & risk. National Institute of Governmental Purchasing. Consultado el 22 de mayo de 2025 en: <https://www.nigp.org/blog/ai-and-public-procurement>

⁶⁹ Executive Office of the President of the United States. (2025, febrero). Memorandum for the heads of executive departments and agencies: Accelerating Federal use of AI through innovation, governance, and public trust (M-25-21), pp. 14, 19. <https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-21-Accelerating-Federal-Use-of-AI-through-Innovation-Governance-and-Public-Trust.pdf>

Esta alternativa reconoce que **no todas las soluciones basadas en IA presentan el mismo nivel de efectos o incidencia sobre derechos fundamentales, activos o recursos estratégicos del Estado**, por lo que imponer requerimientos excesivos en situaciones de bajo o mediano impacto podría desalentar la adopción tecnológica o generar cargas administrativas innecesarias. Atento a ello, se alienta a las entidades contratantes a modular los requisitos éticos, técnicos y documentales conforme a los resultados del sistema, para garantizar una contratación más eficiente, proporcional y orientada al valor público, sin renunciar a los principios éticos desarrollados más arriba.

Adicionalmente, al momento de fijar las condiciones de la contratación, también deberían contemplarse las **particularidades** que se pueden dar con relación a:

- a) La contratación de **sistemas de IA de propósito general**.
- b) Los **flujos agénticos y agentes** impulsados por IA generativa.
- c) La contratación de **software potenciado por IA** para mejorar sus funcionalidades.

A los fines de este documento, se entiende por **sistema de IA de propósito general** al **sistema de IA basado en un modelo de IA de propósito general**⁷⁰, con capacidad para servir a diversos fines, tanto para su uso directo como para su integración en otros

⁷⁰ El Reglamento de IA de la Unión Europea define a los modelos de IA de uso general como “...un modelo de IA, incluso cuando dicho modelo de IA se entrene con una gran cantidad de datos utilizando la autosupervisión a escala, que muestre una generalidad significativa y sea capaz de realizar de forma competente una amplia gama de tareas distintas, independientemente de la forma en que el modelo se comercialice, y que pueda integrarse en una variedad de sistemas o aplicaciones posteriores, excepto los modelos de IA que se utilicen para actividades de investigación, desarrollo o creación de prototipos antes de su comercialización...”.

Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas sobre inteligencia artificial, art. 3, ap. 63, DOUE L 12.7.2024, p. 1. Disponible en: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>.

sistemas⁷¹. Esta definición comprende, por ejemplo, las aplicaciones que integran modelos de lenguaje grande, como ChatGPT, Copilot, Grok, Gemini, entre otros.

Los **flujos agénticos y agentes impulsados por IA generativa**, por su parte, son sistemas de IA que se apoyan en modelos de IA generativa como motor principal para su funcionamiento y los combinan con herramientas como la memoria y el acceso a internet, lo que les permite operar con distinto nivel de autonomía y adaptabilidad, para tomar decisiones y realizar acciones para completar flujos de trabajo completos⁷².

Finalmente, se habla de **software potenciado con IA** para hacer referencia al software que incluye funcionalidades sustentadas en modelos o servicios de IA (p. ej. asistentes de texto, análisis de datos, reconocimiento de voz), pero que no constituye en sí mismo un sistema de IA con autonomía propia.

Reconocer las diferencias que presentan estos tres tipos de sistemas de IA, con relación al sistema de IA tradicional, diseñado como una solución a medida del organismo, no es meramente técnico, sino que tiene implicancias normativas y contractuales relevantes.

Por ello, se proponen adaptaciones específicas destinadas a modular los requisitos conforme a las características técnicas de estos sistemas. Esta flexibilización normativa favorece la integración de soluciones innovadoras, sin perder de vista la necesidad de garantizar una contratación ética, responsable y alineada con los derechos fundamentales.

Finalmente, y en todos los casos, para la compra pública de sistemas de IA, se sugiere superar la visión tradicional de la compra pública centrada exclusivamente en el precio

⁷¹ Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) n.º 300/2008, (UE) n.º 167/2013, (UE) n.º 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de Inteligencia Artificial). Disponible en: https://noticias.juridicas.com/base_datos/Laboral/928500-reglamento-ue-2024-1689-del-parlamento-europeo-y-del-consejo-de-13-de-junio.html#cpa3 (consultado el 22 de mayo de 2025).

⁷² Corvalán, J. G., & Sánchez Caparrós, M. (2025). Agentes de Inteligencia Artificial y Workflows Agénticos. La nueva frontera de la automatización. p. 38.

al momento de evaluar la conveniencia de una oferta, para adoptar un **enfoque de valor público**, en el cual la **oferta más conveniente** se determine considerando no solo el costo económico del sistema de IA, sino también factores cualitativos y éticos que reflejen la sostenibilidad, la inclusión y la confiabilidad de la solución ofrecida por el proveedor.

Este **enfoque ampliado** permite ponderar aspectos como la calidad técnica, la experiencia del proveedor, los tiempos de entrega y, especialmente, el cumplimiento de principios éticos como la equidad, la transparencia, la protección de datos, la explicabilidad, la supervisión humana y la responsabilidad.

Incorporar estos elementos en la evaluación contribuye a una visión más precisa del valor público generado, favorece decisiones más alineadas con los derechos fundamentales y permite orientar la inversión pública hacia soluciones sostenibles e innovadoras.

En suma, contar con pautas de contratación específicas aplicables a procesos de compras relacionados con sistemas de IA, con las particularidades y reglas de flexibilización señaladas, colaborará a proporcionar certeza, no sólo al organismo público contratante, sino también al proveedor tecnológico respecto a sus obligaciones, y a los usuarios y partes interesadas acerca del alcance y las características del sistema cuya adquisición se pretende.

VII. Conclusiones y recomendaciones finales

La integración de estándares contractuales como los que se plantean en los documentos agregados al Anexo I de este informe, emerge como un factor crítico para la **materialización de una contratación responsable y ética de la IA en el Sector Público**.

Estos estándares, alineados con principios éticos con elevado consenso internacional⁷³, establecen un marco de derechos y obligaciones que protege tanto a las entidades gubernamentales contratantes como a los ciudadanos impactados por los sistemas automatizados, a la vez que brinda claridad y seguridad jurídica al proveedor.

Para instituciones públicas, que a menudo lideran o asesoran en procesos de contratación tecnológica a gran escala, este repertorio de pautas generales flexibles representa una **herramienta de alto valor práctico y aplicabilidad inmediata**. Su implementación estratégica puede elevar significativamente la calidad, confiabilidad y rendición de cuentas de las soluciones de IA adquiridas por el Sector Público, mitigando riesgos potenciales y fomentando la confianza ciudadana.

Como recomendación final, se sugiere tener en cuenta las siguientes estrategias para su implementación:

Integración proactiva y adaptativa:

- Incorporar estos requerimientos desde las etapas iniciales de los procesos de compras (pliegos de licitación, términos de referencia) para comunicar claramente las expectativas de IA responsable a los proveedores.
- Establecer un mecanismo de revisión y actualización periódica de las cláusulas, que permita su adaptación ágil a la evolución del panorama regulatorio y los avances tecnológicos en el campo de la IA. Esto asegurará la continua relevancia y eficacia del marco contractual.
- Promover la aplicación voluntaria de cláusulas clave, incluso en casos de IA de bajo impacto, como una estrategia para fortalecer la fiabilidad de las soluciones y demostrar un compromiso proactivo con la ética en la IA.

Desarrollo de capacidades y conciencia ética:

⁷³ Sánchez Caparrós, Mariana. "Principios éticos para una inteligencia artificial antropocéntrica: consensos actuales desde una perspectiva global y regional". En: Corvalán, Juan G. (dir.), Tratado de Inteligencia Artificial y Derecho, 2.ª ed., Thomson Reuters La Ley, 2023.

- Implementar programas de capacitación y sensibilización dirigidos a gestores públicos, equipos de adquisiciones y asesores legales, enfocados en la comprensión tanto de los aspectos técnicos como de las implicaciones éticas de la incorporación de IA.
- Fomentar la creación de comunidades de práctica y espacios de diálogo interdisciplinario para facilitar el intercambio de conocimientos y la resolución colaborativa de dilemas éticos relacionados con la IA.
- Promover el uso de un lenguaje claro y accesible en la documentación contractual y los procesos de capacitación, para asegurar que los responsables no técnicos puedan comprender la importancia de conceptos como la explicabilidad, la mitigación de sesgos y la privacidad.

Construcción de alianzas colaborativas con proveedores:

- Enfocar la presentación de las cláusulas contractuales como un marco que genera valor tanto para el organismo adquirente como para el proveedor, clarificando expectativas y promoviendo la confianza mutua.
- Desarrollar programas de acompañamiento y apoyo a proveedores, especialmente aquellos con menor experiencia, ofreciendo recursos como plantillas, ejemplos y acceso a expertos en ética de la IA para facilitar la adecuación de sus propuestas.
- Establecer mecanismos de retroalimentación continua con los proveedores para identificar áreas de mejora en las cláusulas contractuales y promover la adopción de mejores prácticas en el desarrollo de IA ética.

Implementación de un enfoque de supervisión del ciclo de vida:

- Diseñar e implementar mecanismos institucionales robustos para el seguimiento y la verificación del cumplimiento de las cláusulas contractuales a lo largo de todo el ciclo de vida de los sistemas de IA. Esto puede incluir la designación de un responsable interno de seguimiento de IA, la realización de

auditorías periódicas y la implementación de sistemas de alerta temprana para la detección de desviaciones o incidentes.

- Establecer protocolos claros para la gestión y resolución de incidentes, incluyendo la definición de plazos de respuesta, la asignación de responsabilidades y la implementación de medidas correctivas y preventivas.
- Promover la transparencia activa hacia la ciudadanía mediante la publicación de información relevante sobre los sistemas de IA adquiridos y utilizados por el organismo, resultados de auditorías, evaluaciones de impacto y métricas de desempeño (siempre respetando la seguridad y los secretos comerciales). Esto fortalecerá la confianza pública y fomentará un círculo virtuoso de mejora continua.

Por último, cabe señalar que la propuesta contenida en el presente documento, en lo referido al instrumento de pliego y cláusulas modelo para la contratación de sistemas de inteligencia artificial, reviste carácter meramente orientativo y no vinculante. Su finalidad es servir como guía de referencia para los Estados de la región, los cuales podrán adaptarla conforme a sus marcos normativos, institucionales y procedimientos administrativos vigentes.

En ese sentido, se recomienda avanzar en la elaboración de modelos estandarizados de pliegos y cláusulas contractuales específicas para sistemas de IA, que puedan ser utilizados como base común en los procesos de adquisición pública. La disponibilidad de este tipo de modelos contribuirá a agilizar los procedimientos de contratación, al reducir los tiempos de diseño documental y permitir su adecuación al caso concreto con simples ajustes contextuales. Esta estrategia no solo incrementará la eficiencia operativa del Estado en materia de compras públicas tecnológicas, sino que fortalecerá la responsabilidad, transparencia y ética de dichas contrataciones.

Asimismo, cada país deberá evaluar y definir el nivel normativo más adecuado para institucionalizar estas recomendaciones, pudiendo optar por su incorporación a través de directivas administrativas, resoluciones ministeriales, reglamentaciones específicas

o incluso normas con rango legal, según el esquema jurídico y la madurez institucional de cada jurisdicción.

VIII. Glosario de términos

a) **Sistema de Inteligencia Artificial (Sistema de IA):** sistema basado en máquinas diseñado para operar con diversos niveles de autonomía que puede, para objetivos explícitos o implícitos, inferir cómo lograr dichos objetivos a partir de la entrada que recibe, y producir resultados –como predicciones, contenidos, recomendaciones o decisiones– que influyan en entornos físicos o virtuales.

b) **Sistemas de IA de Propósito General:** sistemas de IA basado en un modelo de IA de propósito general, con capacidad para servir a diversos fines, tanto para su uso directo como para su integración en otros sistemas.

c) **Modelo de IA de propósito general:** modelo de IA entrenado con una vasta variedad y fuentes de datos, que puede adaptarse a una amplia gama de tareas distintivas e integrarse en una variedad de sistemas o aplicaciones posteriores, como sucede con los modelos de lenguaje grande⁷⁴.

d) **Software potenciado con IA:** sistema que incorpora módulos o servicios basados en IA para mejorar sus funcionalidades (p. ej. software de escritura que integra IA generativa), pero que no constituye un sistema de IA con autonomía propia.

e) **Agentes de IA y flujos agénticos:** son sistemas de IA que se apoyan en modelos de inteligencia artificial generativa como motor principal para su funcionamiento y los combinan con herramientas (ej. memoria y acceso a internet) para tomar decisiones y realizar acciones que les permitan completar flujos de trabajo completos con distintos

⁷⁴ El Reglamento de IA de la Unión Europea define a los modelos de IA de uso general como “...un modelo de IA, incluso cuando dicho modelo de IA se entrene con una gran cantidad de datos utilizando la autosupervisión a escala, que muestre una generalidad significativa y sea capaz de realizar de forma competente una amplia gama de tareas distintas, independientemente de la forma en que el modelo se comercialice, y que pueda integrarse en una variedad de sistemas o aplicaciones posteriores, excepto los modelos de IA que se utilicen para actividades de investigación, desarrollo o creación de prototipos antes de su comercialización...”.

Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas sobre inteligencia artificial, art. 3, ap. 63, DOUE L 12.7.2024, p. 1. Disponible en: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>.

niveles de autonomía y adaptabilidad. Los agentes impulsados por IA generativa pueden emplear un único agente o múltiples agentes con roles específicos para comprender solicitudes, planificar flujos de trabajo, coordinar entre agentes con otros roles específicos, agilizar acciones, colaborar con humanos y validar resultados⁷⁵.

f) **Proveedor:** toda persona humana o jurídica, pública o privada, nacional o extranjera, inscrita o pre inscrita en el [ej. Sistema de Información de Proveedores (SIPRO)], que participe como oferente o resulte adjudicataria en procedimientos de contratación del organismo.

g) **Ciclo de Vida del Sistema de IA:** comprende todas las etapas desde la concepción inicial, pasando por el diseño, desarrollo, entrenamiento, validación, implementación, operación, mantenimiento, actualización, hasta el retiro o desmantelamiento del sistema.

h) **Gobernanza de IA:** conjunto de políticas, procesos, roles, responsabilidades, métricas y controles establecidos por una organización para dirigir, gestionar y supervisar el desarrollo, despliegue y uso de sistemas de IA de manera ética, legal y responsable.

i) **Evaluación de Impacto Ético y de Protección de Datos (EIPD):** proceso sistemático para identificar, analizar, evaluar y mitigar los riesgos éticos y para la protección de datos personales derivados del desarrollo y uso de un sistema de IA.

j) **Sistema de IA de alto impacto:** clasificación asignada a un sistema de IA cuyo resultado sirve de base principal para tomar decisiones o ejecutar acciones con impacto legal o material significativo sobre los derechos civiles; la privacidad de una persona o entidad; el acceso de una persona o entidad a la educación, la vivienda, los seguros, el crédito, el empleo u otros programas; el acceso de una persona o entidad a recursos o servicios gubernamentales críticos; la salud y la seguridad humanas; la infraestructura crítica o la seguridad pública; o activos o recursos estratégicos del

⁷⁵ Corvalán, J. G., & Sánchez Caparrós, M. (2025). Agentes de Inteligencia Artificial y Workflows Agénticos. La nueva frontera de la automatización. p. 43.

gobierno, incluyendo propiedades de alto valor e información marcada como sensible o clasificada⁷⁶.

k) **Sesgo algorítmico:** carga de estereotipos, valores, creencias sociales, patrones de exclusión, desigualdad y/o discriminación estructural presente en un sistema de inteligencia artificial, que se origina en diversas circunstancias e, incluso en otros sesgos⁷⁷, y da lugar a resultados errados que perjudican o discriminan a ciertos grupos de personas porque provocan daños de asignación o representación, con base en características protegidas —por ejemplo, el género, el color de la piel, etc.—, determinadas características no protegidas ligadas a ellas —por ejemplo, el código postal del barrio que habitan, su nombre o apellido, etc.—, o bien, de manera aleatoria⁷⁸.

⁷⁶ Definición adoptada en el reciente Memorando para los Responsables de Departamentos Ejecutivos y Agencias, publicado por la Oficina Ejecutiva de la Presidencia de los EE.UU. Además, allí se provee un listado de casos de uso que se presumirá automáticamente como de alto impacto. Ver: <https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-21-Accelerating-Federal-Use-of-AI-through-Innovation-Governance-and-Public-Trust.pdf>, p. 14, 19.

⁷⁷ Puede tener origen en uno o varios sesgos sistémicos, estadísticos, computacionales y humanos, tanto individuales como grupales. En esta línea, resulta pertinente volver sobre la conceptualización propuesta por Schwartz, Reva; Vassilev, Apostol; Greene, Kristen; Perine, Lori; Burt, Andrew; Hall, Patrick, en el marco del informe del National Institute of Standards and Technology (NIST), donde se avanza hacia un estándar para la identificación y gestión del sesgo en sistemas de inteligencia artificial (NIST Special Publication 1270. Towards a Standard for Identifying and Managing Bias in Artificial Intelligence, 2022).

⁷⁸ Cfr. Sánchez Caparrós, Mariana, La discriminación provocada por la inteligencia artificial a la luz de la jurisprudencia de la Corte Suprema de Justicia de la Nación. Evaluación de la idoneidad de los estándares constitucionales de igualdad y no discriminación frente a los impactos de la inteligencia artificial, El Dial, en prensa, abril de 2025.

Anexo I - Modelo de resolución de aprobación de estándares para la contratación de sistemas de IA confiables, éticos y centrados en las personas.

Ingresar al documento mediante el siguiente link:

[!\[\]\(f8021f2863ff022ca9b09416c5c56134_img.jpg\) *Modelo de resolución y modelo de pliego_CAF*](#)

ubatec^{SA}

CAF BANCO DE DESARROLLO
DE AMÉRICA LATINA
Y EL CARIBE

