Guía para entidades públicas y privadas en materia de Transparencia y Protección de Datos Personales para una Inteligencia Artificial responsable

Versión preliminar

AAIP
Transparencia
Acceso a la Información
Protección de Datos Personales

Autoridades

Beatriz Anchorena
Titular de la AAIP

Luciana Carpinacci
Directora Nacional de Evaluación de Políticas de Transparencia
Catalina Byrne
Directora de Transparencia Activa
Emiliano Arena
Director de Evaluación y Participación Ciudadana

Violeta Paulero
Directora Nacional de Protección de Datos Personales
Anastasia Dozo
Directora de Promoción del Derecho a la Privacidad
Maximiliano Rey
Director de Fiscalización y Regulación

Dirección Nacional de Políticas de Acceso a la Información
Luciano Acevedo
Director de Contenido y Normativa de Acceso a la Información
Estefanía Pinetta Biro Alemán
Directora de Gestión y Control de Acceso a la Información

Programa Nacional de Transparencia y Protección de Datos Personales en el uso de la Inteligencia Artificial

En un contexto de crecimiento exponencial de los procesos de integración de la Inteligencia Artificial (IA) a las soluciones tecnológicas en múltiples ámbitos, la Agencia de Acceso a la Información Pública (AAIP) busca acompañar esta situación y poner especial atención a los posibles riesgos sociales, económicos, culturales y ambientales que deben ser estudiados y atendidos para evitar sesgos y discriminación.

Este programa fue desarrollado por la Dirección Nacional de Protección de Datos Personales y Dirección Nacional de Evaluación de Políticas de Transparencia e impulsa procesos de análisis, regulación y fortalecimiento de capacidades para acompañar el desarrollo y uso de la IA, en el sector público y privado. Está destinado a instituciones públicas y privadas que utilicen sistemas de decisiones automatizadas con aplicaciones de IA.

El Programa cuenta con las siguientes líneas de acción: **Observatorio sobre IA**, para dar seguimiento a los avances regionales y globales en materia de regulación de los desarrollos tecnológicos basados en IA y promover la transparencia algorítmica; **Gobernanza y participación social**, para promover el trabajo conjunto con órganos gubernamentales de gobernanza tecnológica y jurídica en la materia; y **Fortalecimiento de capacidades**, para brindar capacitaciones y asistencia técnica en materia de transparencia y protección de datos personales en el uso de la IA y campañas de sensibilización, alfabetización mediática e informacional.

Coordinación general
Beatriz Anchorena

Coordinación metodológica
Luciana Carpinacci
Violeta Paulero

Equipo de trabajo:
Cecilia Buffa
Jorge Orovitz
Silvana Rica
Agustina Sirven

El equipo del Programa agradece los comentarios y aportes recibidos de: Anastasia Dozo, Catalina Byrne, Emiliano Arena y Agustín Pérez Aleda.

Cómo citar este documento:
Agencia de Acceso a la Información Pública (AAIP). Guía de recomendaciones para entidades públicas y privadas en materia de transparencia y protección de datos personales para una Inteligencia Artificial responsable. Buenos Aires, junio de 2024.

Índice

Introducción 6

01

Objetivo 9

02

Alcance 10

03

Definiciones y características de los sistemas de Inteligencia Artificial 11

04

Principales problemas y desafíos de la IA 16

05

Principios de Transparencia y Protección de Datos Personales 20

06

Recomendaciones de Transparencia y Protección de Datos Personales en el ciclo de vida de Sistemas de IA 26

Etapa 1: *Diseño del Sistema* 27
Etapa 2: *Verificación y validación* 36
Etapa 3: *Implementación* 39
Etapa 4: *Operación y mantenimiento* 42

07

Consideraciones finales 49

08

Acrónimos 50

09

Referencias bibliográficas 51

Anexo I *Antecedentes y evolución de estándares internacionales* 55

Introducción

El desarrollo y el uso de los sistemas que integran tecnologías de Inteligencia Artificial ha crecido exponencialmente en los últimos años. El mundo se encuentra ante una cuarta revolución industrial potenciada por estos nuevos sistemas que generan cambios aceleradamente y de una forma mucho más abrupta que las revoluciones industriales anteriores. La IA tiene el potencial de mejorar el bienestar de las personas, aumentar la innovación y la productividad y ayudar a responder a desafíos globales clave[1].

Algunas ventajas de la incorporación de esta tecnología tanto para el sector público como para el privado son: la personalización de la experiencia de los usuarios y la automatización de los procesos evitando labores repetitivas; la reducción de los tiempos de respuesta; la toma de decisiones basadas en evidencia, y la mejora en la precisión de los resultados; su contribución a la detección preventiva del fraude y a los procesos de análisis y diseño de soluciones tecnológicas, entre muchos otros aspectos[2].

Hoy en día la Inteligencia Artificial se aplica a los campos de la salud, favoreciendo el diagnóstico temprano de enfermedades; la industria y los servicios, automatizando tareas repetitivas o como asistentes de atención automatizada, liberando a los trabajadores para enfocarse en actividades de mayor valor y creatividad; transporte y logística; educación y muchos otros campos.

Si bien los sistemas de IA pueden aportar importantes beneficios a los usuarios y a la sociedad, provocan una serie de preocupaciones y riesgos, entre los que pueden mencionarse:

- Sesgos y discriminación: los sistemas de IA pueden reproducir e incluso amplificar sesgos existentes en los datos con los que son entrenados, lo que puede llevar a decisiones injustas y discriminatorias en áreas como contratación de personal, la justicia penal, la equidad de género, los servicios financieros, o también puede profundizar la discriminación étnica o religiosa.

[1] OCDE. Recommendation of the Council on Artificial Intelligence, 2019. [2024]. Adoptado el 21/05/2019, corregido el 02/05/2024. **https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449**

[2] Recomendaciones para una Inteligencia Artificial Fiable. Secretaría de Innovación Pública, Jefatura de Gabinete de Ministros, 2023. **https://www.argentina.gob.ar/sites/default/files/2023/06/recomendaciones_para_una_inteligencia_artificial_fiable.pdf**

- Violaciones a la privacidad: la IA requiere en la etapa de programación grandes cantidades de datos, entre los cuales pueden incluirse datos personales. Una mala gestión de estos datos puede llevar a violaciones de la privacidad y a la explotación de información sensible sin consentimiento de las personas.

- Falta de transparencia: también se han detectado riesgos en relación a la fiabilidad de la información que producen los sistemas de IA. La opacidad con la cual se desarrollan dificulta, en muchos casos, la posibilidad de acceder a las reglas de toma de decisiones con las que funcionan los modelos, de modo que es dificultoso para las personas realizar reclamos debido a falta de trazabilidad de los sistemas.

Estos riesgos podrían tener un mayor impacto en personas que se encuentran en una situación especial de vulnerabilidad, tales como, niños, niñas y adolescentes, géneros y disidencias, personas con discapacidades, minorías étnicas y raciales, personas de la tercera edad, personas en situación de pobreza, entre otras.

Haciéndose eco de la necesidad de abordar preventivamente estos riesgos, la comunidad internacional viene desarrollando diversos esfuerzos regulatorios y recomendaciones que abogan por un uso ético de la IA centrado en la protección de los derechos humanos, con el objetivo de mitigar los efectos no deseados desde etapas tempranas. En este sentido la Organización de Naciones Unidas (ONU) a través de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) emitió la Recomendación sobre la Ética de la Inteligencia Artificial, a la que adhirieron todos los países miembros en la Asamblea General de noviembre de 2021, entre los cuales se encuentra la República Argentina. En Anexo I se presenta una síntesis de los antecedentes y evolución de principios y estándares internacionales sobre el uso de la IA.

Para acompañar estos esfuerzos desde el ámbito de sus competencias, la AAIP creó en 2023 el Programa de Transparencia y Protección de Datos Personales en el uso de la Inteligencia Artificial. Esta línea de acción tiene el propósito de impulsar procesos de análisis, regulación y fortalecimiento de capacidades estatales necesarias para acompañar el desarrollo y uso de la Inteligencia Artificial garantizando el efectivo ejercicio de los derechos de la ciudadanía en materia de transparencia y protección de datos personales[3].

[3] Agencia de Acceso a la Información Pública. Resolución 161/2023. Programa de transparencia y protección de datos personales en el uso de la inteligencia artificial, 26/08/2023. RESOL-2023-161-APN-AAIP. **https://www.boletinoficial.gob.ar/detalleAviso/primera/293363/20230904**

En particular, uno de los componentes del Programa de trabajo de la AAIP consiste en el desarrollo de capacidades en los actores del sector público y privado para comprender los desafíos que implica la incorporación de la Inteligencia Artificial en los sistemas de toma de decisiones automatizadas, brindando herramientas para su abordaje en el marco de los principios éticos desarrollados por los instrumentos internacionales.

Además, la AAIP participa en las redes internacionales -Red Iberoamericana de Protección de Datos (RIPD), Red Iberoamericana de Transparencia y Acceso a la Información (RTA) y Asamblea Global de Privacidad (GPA)- ámbitos donde se desarrollan acciones de cooperación para generar instrumentos que permitan fortalecer las capacidades nacionales para regular y contener las prácticas abusivas en el uso de la Inteligencia Artificial. Asimismo, participa en carácter de observador del Comité de Inteligencia Artificial del Consejo de Europa en el cual se adoptó la Convención sobre Inteligencia Artificial (IA), Derechos Humanos, Democracia y Estado de Derecho[4].

[4] El instrumento, adoptado el 17 de mayo de 2024, en la 133rd Session of the Committee of Ministers, se abrirá a la firma el 5 de septiembre de 2024. En el siguiente vínculo se encuentra disponible el texto adoptado https://shorturl.at/d7XWg

01
Objetivo

La presente guía hace foco en las implicancias de las tecnologías basadas en sistemas de decisión automatizada, y en particular, aquellos que incorporan la Inteligencia Artificial, en relación a sus consecuencias sobre los derechos fundamentales y cómo abordar desde el punto de vista normativo e institucional estos desafíos.

A partir de los lineamientos incorporados en la Guía "Recomendaciones para una Inteligencia Artificial Fiable" producida por la administración pública nacional, este instrumento se propone profundizar las recomendaciones asociadas con la incorporación de los principios de transparencia y protección de datos personales[5]. Ambas guías reconocen y proponen estrategias para la incorporación de los principios incorporados por la UNESCO en la Recomendación sobre la Ética de la Inteligencia Artificial (2021).

Con este instrumento, nos proponemos acompañar a los actores, tanto del sector público como del privado, ofreciendo lineamientos que los ayuden a incorporar la transparencia y protección de datos personales, como dos dimensiones sustantivas y transversales en los proyectos de desarrollo tecnológico que implementen sistemas de Inteligencia Artificial, de modo que se garantice el efectivo ejercicio de los derechos de la ciudadanía.

Para ello, se brindan recomendaciones y pautas de responsabilidad proactiva para prevenir riesgos desde etapas tempranas y promover la integración de una Inteligencia Artificial responsable durante todo el ciclo de vida del sistema, de modo que se aseguren estándares básicos de protección.

[5] Recomendaciones para una Inteligencia Artificial fiable, 2023.

02
Alcance

Esta guía se dirige a un universo amplio de destinatarios, públicos y privados de distintas profesiones, actores implicados en el desarrollo e implementación de sistemas que incorporan Inteligencia Artificial. En este sentido, está destinada a organizaciones proveedoras y desarrolladoras de soluciones que integran sistemas de Inteligencia Artificial, Gobiernos y decisores de la implementación de políticas para el fomento del uso de este tipo de tecnologías, organismos y empresas que implementan Inteligencia Artificial en sus procesos y/o productos, instituciones académicas que investigan el impacto de estos sistemas, como así también organizaciones sociales que velan por la protección de los derechos de la ciudadanía ante el avance acelerado de la IA.

03

Definición y características de los sistemas de Inteligencia Artificial

En este apartado, se presentan algunas definiciones clave sobre un ámbito complejo y diverso como es el desarrollo tecnológico que cada vez más recurre a la Inteligencia Artificial para optimizar sus productos.

Los sistemas de Inteligencia Artificial se inscriben dentro de la categoría de Sistemas de Decisiones Automatizadas, pero es preciso establecer una diferenciación: un Sistema de Decisión Automatizada (SDA) puede tratarse de decisiones basadas en un conjunto predefinido de reglas o algoritmos, que opera de manera rígida y es diseñado para seguir instrucciones específicas y realizar tareas repetitivas o rutinarias. Este tipo de sistemas, cuando incorporan la Inteligencia Artificial (IA), si bien siguen utilizando conjuntos predefinidos de reglas y algoritmos, adoptan características más complejas, capaces de reconocer patrones, que les permiten la comprensión del lenguaje natural, la toma de decisiones y el aprendizaje adaptativo. A diferencia de los sistemas de decisión automatizada tradicionales, la IA utiliza técnicas como el aprendizaje automático (*machine learning*) y redes neuronales para mejorar su rendimiento con el tiempo, lo que implica que puede aprender de los datos y adaptarse a nuevas situaciones, mejorando su rendimiento y precisión.

Por ejemplo, un sistema de Inteligencia Artificial en finanzas puede utilizar modelos de aprendizaje automático para evaluar el riesgo de crédito basándose en el historial financiero de un cliente. La automatización se vuelve más efectiva, pues permite una mejora continua basada en el análisis de datos en tiempo real. En el campo de la salud, los sistemas de *machine learning* se utilizan para analizar imágenes médicas -como radiografías, resonancias magnéticas y mamografías- para detectar enfermedades como cáncer, fracturas óseas y anomalías cardíacas. Las plataformas digitales utilizan algoritmos de *machine learning* para recomendar productos, películas, series y música basándose en el historial de comportamiento del usuario. En síntesis, los sistemas de Inteligencia Artificial pueden definirse como aquellos sistemas informáticos que realizan acciones con cierto grado de autonomía para lograr objetivos específicos[6].

[6] Garrido, Romina; Lapostol, José Pablo y Paz Hermosilla, María. Transparencia algorítmica en el sector público. Consejo para la transparencia de Chile y Gob_Lab de la Universidad Adolfo Ibáñez, 2021.
https://www.consejotransparencia.cl/wp-content/uploads/estudios/2021/10/ESTUDIO-TRANSPARENCIA-ALGORITMICA-EN-EL-SECTOR-PUBLICO-GOBLAB-cambio-tablas-1.pdf

La UNESCO los define como "tecnologías de procesamiento de la información que integran modelos y algoritmos que producen una capacidad para aprender y realizar tareas cognitivas, dando lugar a resultados como la predicción y la adopción de decisiones en entornos materiales y virtuales. Los sistemas que utilizan la IA están diseñados para funcionar con diferentes grados de autonomía, mediante la modelización y representación del conocimiento y la explotación de datos y el cálculo de correlaciones"[7].

A su vez, la recientemente aprobada Ley de Inteligencia Artificial de la Unión Europea[8], la define como "un sistema basado en una máquina diseñado para funcionar con diferentes niveles de autonomía y que pueden mostrar adaptabilidad después del despliegue y que, por razones explícitas u objetivos implícitos, infiere, a partir de los insumos que recibe, cómo generar resultados, tales como predicciones, contenidos, recomendaciones o decisiones que puedan influir física o entornos virtuales"[9].

También es posible ver a la IA como una disciplina que intenta replicar y desarrollar la inteligencia natural, de humanos y seres vivos, a través de computadoras o máquinas. En este sentido, la Fundación Sadosky (2022) plantea que "la IA debe entenderse como un campo de investigación y aplicación que incluye más que enfoques técnicos. Este área debe pensarse en su diseño desde la interdisciplina y se construye desde la lógica, la estadística, la filosofía, las ciencias de la computación, las matemáticas, neurociencias, la lingüística, la psicología, la economía y otras ramas de estudios sociales. Vincula lo técnico, lo científico, las prácticas sociales, gubernamentales e industriales".

Por su parte, la OCDE define a los sistemas de Inteligencia Artificial como "los sistemas basados en una máquina que, para objetivos explícitos o implícitos, infieren, a partir de los datos que reciben, cómo generar resultados tales como predicciones, contenidos, recomendaciones o decisiones que puedan influir en entornos físicos o virtuales. Los diferentes sistemas de Inteligencia Artificial varían en sus niveles de autonomía y adaptabilidad después del despliegue".

En relación con esta última definición, es importante entender los alcances de lo que se denomina la «inferencia de datos». Entendemos este concepto como el proceso de derivar conclusiones y predicciones a partir de un modelo de aprendizaje automático ya entrenado. Es el paso donde el modelo aplica el conocimiento aprendido durante el entrenamiento a nuevos datos para generar

[7] Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO). Recomendación sobre la ética de la inteligencia artificial. Adoptada el 23 de noviembre de 2021. **https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa**

[8] Ley de Inteligencia Artificial de la Unión Europea, aprobada el 21 de abril de 2021 y en vigor desde el 13 de marzo de 2024, establece un marco normativo y jurídico único para los sistemas de inteligencia artificial que operen en la Unión Europea. **https://artificialintelligenceact.eu/es/el-acto/**

[9] **https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_ES.pdf**

resultados útiles. Ejemplos de inferencias de datos que pueden mencionarse son: la inferencia de preferencias políticas a partir de los títulos de libros que una persona compra online, el análisis de sentimientos a partir de las publicaciones de una persona en redes sociales, o la posibilidad de acceder a programas sociales de acuerdo a datos proporcionados por fuentes externas a la persona como agencias gubernamentales, bases de datos de crédito, historiales laborales y de salud, etc.[10]

Como puede observarse en los casos mencionados, la inferencia de datos a partir de la utilización de una Inteligencia Artificial entrenada, puede tener implicancias directas y profundas sobre las personas tanto a nivel de sus derechos humanos, su privacidad y seguridad, así como a nivel de la sociedad, por el impacto que puede tener en las políticas públicas, en la confianza en las instituciones o la seguridad jurídica de una sociedad.

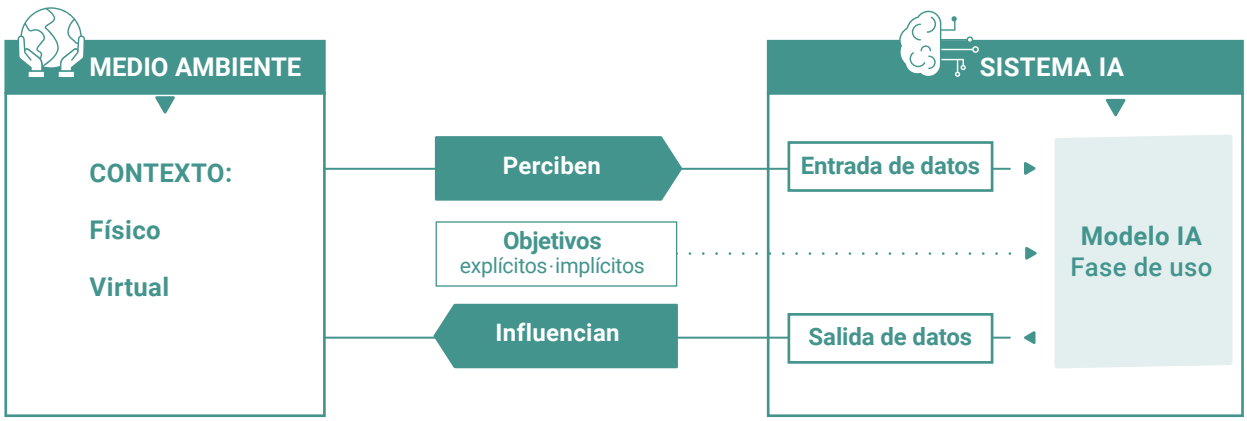
Existen muchos parámetros para caracterizar los impactos de estos sistemas, de acuerdo a qué derechos afecta, qué datos utiliza, cuál es su función y/o campo de aplicación, qué nivel de riesgo implica, qué incidencias tiene en procesos sociales críticos. Todo ello está siendo explorado y puesto en consideración en la discusión a nivel global.

A continuación, se caracteriza sintéticamente cómo funcionan los sistemas de IA.

Funcionamiento de un sistema de Inteligencia Artificial

Tal como se observa en la **Figura 1,** se trata de un sistema que interactúa con el entorno alimentándose de datos, procesándolos y emitiendo respuestas que impactan sobre ese mismo contexto.

Figura 1: Esquema de funcionamiento de un sistema de IA



Fuente: elaboración propia sobre la base de OCDE, 2023

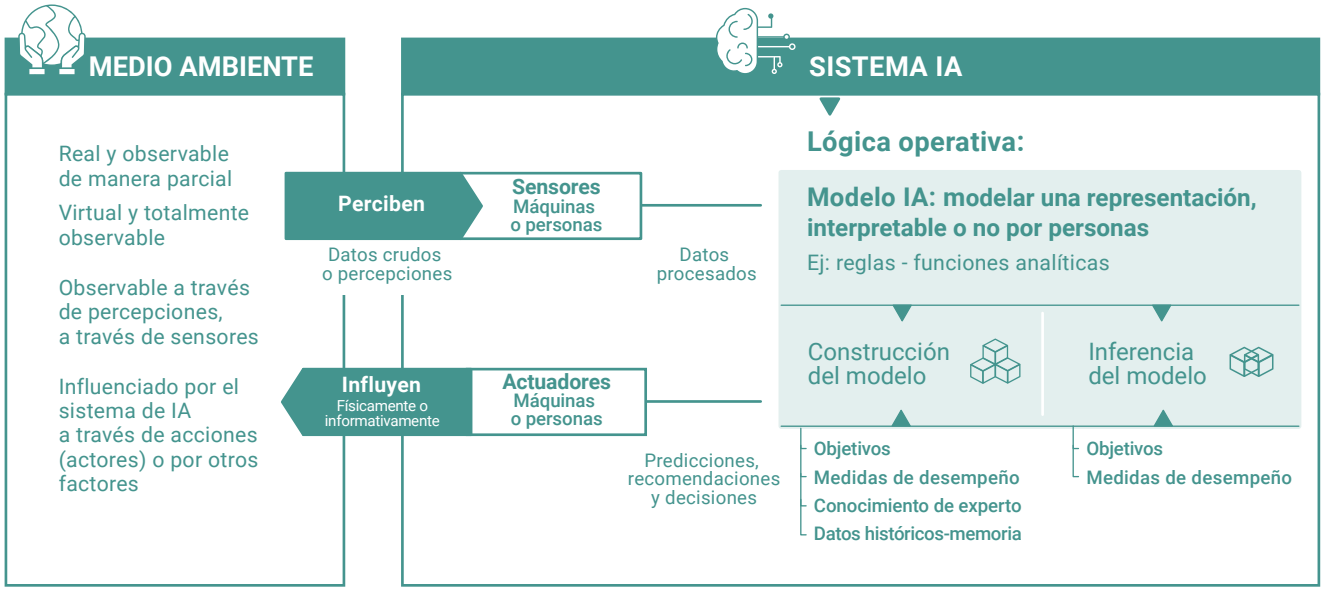
[10] Fundación Vía Libre, Informe: Protección legal de datos personales inferidos, 2023.

Si profundizamos la mirada al interior del sistema descrito, se incorporan conceptos adicionales, como se puede observar en la **Figura 2.**

Allí, se pueden observar los tres elementos principales que componen a un sistema de IA:

1. **Sensores:** recopilan datos brutos del entorno;
2. **Actuadores:** actúan para cambiar el estado del entorno a partir de la salida del modelo.
3. **Lógica operativa:** poder clave del sistema IA que reside en su lógica operativa. Para un conjunto dado de objetivos y basándose en datos de entrada de los sensores, la lógica operativa proporciona salida para los actuadores. Estos toman la forma de recomendaciones, predicciones o decisiones que pueden influir en el estado del entorno.

Figura 2: Principales componentes de la IA



Fuente: elaboración propia sobre la base de OCDE, 2023

Para cubrir diferentes tipos de sistemas de IA y diferentes escenarios, la lógica operacional se separa en tres partes:

1. el **modelo algorítmico** propiamente dicho;
2. la **construcción del modelo** (automatizado o no); y
3. la **inferencia del modelo:** proceso por el cual humanos y/o herramientas automatizadas derivan un resultado del modelo. Estos toman la forma de recomendaciones, predicciones o decisiones.

El entorno es el espacio observable a través de percepciones (a través de sensores) e influenciado a través de acciones (a través de actuadores). Los sensores y actuadores son máquinas o humanos. Estos pueden ser reales (como físicos, sociales, mentales) y generalmente sólo parcialmente observables, o virtuales (por ejemplo, juegos) y generalmente completamente observables.

Los objetivos (por ejemplo, variables de salida) y las medidas de rendimiento (por ejemplo, precisión, recursos para entrenamiento, representatividad del conjunto de datos) guían el proceso de construcción.

En algunos casos, un modelo puede ofrecer una única recomendación, mientras que en otros (por ejemplo, modelos probabilísticos), puede ofrecer una variedad. Estas recomendaciones están asociadas con diferentes niveles de, por ejemplo, medidas de rendimiento como nivel de confianza, robustez o riesgo. En algunos casos, durante el proceso de interpretación, es posible explicar por qué se hacen recomendaciones específicas. En otros casos, dada la complejidad del modelo, la explicación es casi imposible, lo que podría dificultar alcanzar niveles de transparencia y explicabilidad adecuados.

04
Principales problemas y desafíos de la IA

Dadas las particularidades de funcionamiento, se ha observado en repetidas situaciones que el uso no supervisado de la Inteligencia Artificial puede afectar derechos fundamentales, como la libertad de expresión -por falta de conocimiento acerca del funcionamiento del mismo-, la privacidad -por la utilización de datos sensibles sin consentimiento- y, consecuentemente, sobre la dignidad humana.

Entre los problemas y amenazas vinculadas con la transparencia y protección de datos personales, podemos enumerar los siguientes como los más relevantes[11]:

Sesgos y Discriminación. Los sistemas de IA pueden perpetuar o incluso amplificar sesgos y discriminación si no se diseñan, implementan y supervisan adecuadamente. En muchos casos los modelos con algoritmos faciales pueden implicar una identificación errónea por prejuicios raciales. También son frecuentes los sesgos de género, ya que estos sistemas a menudo reflejan los prejuicios presentes en los datos con los que fueron entrenados. Los sistemas de vigilancia basados en IA pueden evaluar injustamente el rendimiento de los empleados, favoreciendo a ciertos grupos sobre otros sin razones justificadas. Las herramientas de evaluación de riesgos en el sistema judicial pueden influir negativamente en decisiones sobre libertad bajo fianza o sentencias de algunas personas sobre otras.

Falta de calidad de los datos. Los sistemas de IA, suponen múltiples retos en cuanto a la legitimidad de la fuente de datos que utilizan estos sistemas para funcionar, dado que en muchos casos, los datos se obtienen de fuentes desactualizadas que, por ejemplo, pueden contener datos que no son de calidad o bien con errores o información falsa. Por ejemplo, el raspado de datos (*"data scraping"*)[12] es una forma de extraer datos de un sitio web de forma automática. Puede ser una herramienta eficiente para obtener información, pero implica riesgos a la privacidad, ya que puede ser utilizada para cometer fraudes, ataques de phishing, robo de datos,entre otros.

[11] OECD, Artificial Intelligence in Society, 2019.

[12] La Declaración Conjunta firmada en agosto de 2023 por la AAIP con otros países en el marco de la Asamblea Global de Privacidad (GPA), ofrece recomendaciones dirigidas a individuos y empresas para mitigar los principales riesgos de usar técnicas automáticas para extraer y procesar datos personales en redes sociales y sitios web. En particular afirma que "Las empresas de redes sociales que alojan datos personales de acceso público tienen la obligación de proteger la información personal en sus plataformas contra la extracción ilegal de datos". La Declaración completa se encuentra en: **https://ico.org.uk/media/about-the-ico/documents/4026232/joint-statement-data-scraping-202308.pdf**

Riesgo de violación de la privacidad. La IA requiere grandes cantidades de datos personales para entrenar modelos y realizar inferencias, lo que aumenta el riesgo de violación de la privacidad si estos datos se utilizan de manera inapropiada o se divulgan sin consentimiento. Algunas aplicaciones de IA, como el reconocimiento facial y el análisis de comportamiento, pueden ser intrusivas y recopilar información sensible sin el conocimiento o consentimiento de las personas.

Riesgos de seguridad. Relacionado por la violación a la privacidad, los sistemas de IA pueden ser vulnerables a ataques cibernéticos que pueden comprometer su funcionamiento y podría dar como resultado la divulgación no autorizada de datos personales, en particular de datos sensibles, manipular los resultados del modelo o comprometer la integridad del sistema.
Otro de los riesgos es la reidentificacion de datos anónimos. Esto es que, a pesar de los esfuerzos por anonimizar los datos personales, la IA puede ser capaz de identificar o reidentificar a individuos a partir de conjuntos de datos aparentemente anónimos, lo que aumenta el riesgo de violación de la privacidad.

Uso inadecuado de datos personales. Los datos personales recopilados para un propósito específico pueden ser utilizados para otros fines, violando principios de protección de datos. También existe el riesgo de la comercialización de datos personales, si las empresas venden o comparten datos personales con terceros sin el conocimiento o consentimiento del individuo, lo que puede resultar en la explotación comercial de la información personal.

Fraude de identidad. Existen técnica como las *"deepfakes"* (videos, imágenes o audios que imitan la apariencia y el sonido de una persona) que, mediante el uso de técnicas de IA (particularmente redes neuronales profundas, un modelo de IA que imita la estructura y el funcionamiento del cerebro humano), pueden sustituir la identidad y afectar la privacidad de las personas, así como a terceros mediante engaños para obtener información sensible o acceso a sistemas. Las *"deepfakes"* puede dañar la reputación personal y profesional de las personas, así como contribuir a la desinformación.

Vigilancia sin consentimiento. La IA incrementa las posibilidades de hacer un seguimiento y un análisis de las costumbres cotidianas de las personas. Por ejemplo, existe el riesgo potencial que se recurra a la IA para la vigilancia masiva del comportamiento de empleados y la creación de perfiles de rendimiento sin el consentimiento de las personas implicadas.

Falta de transparencia. Los modelos de IA pueden ser complejos y opacos, lo que dificulta la comprensión de cómo se toman las decisiones. Los modelos utilizados pueden manejar una gran cantidad de variables simultáneamente, complicando la explicación de cómo cada variable influye en la decisión final. La falta de transparencia puede socavar la confianza de los ciudadanos en el proceso y aumentar el riesgo de mal uso o abuso de la tecnología. La transparencia se ve dificultada si no pueden conocerse el tipo de datos utilizados, los algoritmos aplicados y los procesos internos del modelo.

Los sesgos de múltiple origen

"Los sistemas de Inteligencia Artificial pueden generar resultados sesgados de múltiples tipos, como los que se detallan a continuación[13]:

Sesgo de percepción	Sesgo técnico	Sesgo de modelado	Sesgo de activación
Los datos recopilados representan en exceso o en defecto a una determinada población y hacen que el sistema funcione mejor (o peor) para esa población en comparación con otras.	La propia tecnología introduce sesgos o imprecisiones debido por ejemplo a algoritmos que funcionan mejor con ciertas variables o características del sistema que con otras.	El diseño manual de un modelo por parte de expertos no tiene en cuenta algunos aspectos del entorno, ya sea consciente o inconscientemente.	Se produce cuando las salidas del sistema se utilizan en el entorno de manera sesgada.

Sesgos étnicos y de género

Muchas investigaciones han demostrado que los sistemas de reconocimiento facial suelen tener una mayor tasa de error en la identificación de mujeres, especialmente mujeres de piel más oscura, en comparación con hombres de piel clara. Estadísticamente es más probable que las palabras femeninas o los nombres propios de mujer se relacionen con la vida personal, las tareas domésticas, lo artístico o emocional, mientras que lo masculino se asocia con lo profesional, lo científico y el conocimiento. **En el procesamiento de grandes datos se dan estereotipos muy significativos en los que basa su aprendizaje la IA.**

Para profundizar en este tema se puede consultar: Benítez Eyzaguirre, Lucía. Ética y transparencia para la detección de sesgos algorítmicos de género. Revista Estudios sobre el Mensaje Periodístico. Ediciones Complutenses, 2019. ISSN-e: 1988-2696. **https://doi.org/10.5209/esmp.66989**

[13] Secretaría de Innovación Pública. Recomendaciones para una Inteligencia Artificial Fiable, 2023.

Neurodatos y derecho a la privacidad

Los neurodatos -datos obtenidos de la actividad del sistema nervioso, como la actividad eléctrica del cerebro, imágenes cerebrales, y otros tipos de señales del sistema nervioso- pueden revelar información profunda sobre el estado mental, las emociones, y las capacidades cognitivas de una persona. En el ámbito de la salud, el crecimiento de las aplicaciones de neurotecnología mediante sistemas de Inteligencia Artificial, pueden predecir o inferir información de las personas.

Si bien este tipo de avances podrían generar beneficios como el tratamiento de las personas afectadas por trastornos neurológicos y enfermedades mentales, también podrían ocasionar problemas éticos y de privacidad. Existe el riesgo de que los neurodatos sean recolectados y utilizados sin el consentimiento explícito e informado de los individuos, lo que puede resultar en una invasión de la privacidad, pues estas tecnologías pueden recolectar información íntima sobre pensamientos, emociones y predisposiciones, que tradicionalmente han sido privados y personales. Permiten incluso no solo acceder y evaluar sino también manipular la estructura del cerebro actuando sobre nuestras identidades y emociones.

Existen problemas de transparencia asociados a los neurodatos. Las personas pueden no ser plenamente conscientes de cómo se utilizan sus neurodatos y para qué fines, lo que socava la transparencia y la confianza.

[13] Secretaría de Innovación Pública. Recomendaciones para una Inteligencia Artificial Fiable, 2023.

05

Principios de Transparencia y Protección de Datos Personales

La transparencia y la protección de datos personales son fundamentales para garantizar los derechos de las personas y el uso responsable de la Inteligencia Artificial (IA). Aseguran que las personas comprendan y controlen cómo se recopilan, utilizan y protegen sus datos, y que los sistemas de IA operen de manera comprensible y justa.

La transparencia y la protección de datos personales son dos aspectos claves a considerar en todo el ciclo de producción de soluciones tecnológicas y van de la mano en pos de lograr una IA responsable.

Principios relacionados con la protección de datos personales

La preservación de la intimidad y la protección de los datos personales constituyen derechos esenciales para la protección de la dignidad, la autonomía y la capacidad de actuar de los seres humanos, y debe ser respetada, protegida y promovida a lo largo del ciclo de vida de los sistemas de IA.

Es importante que los datos para los sistemas de IA se recopilen, utilicen, compartan, archiven y supriman de forma consistente con el derecho internacional y acorde a los valores y principios de la ley.

Algunos de los principios fundamentales que plantea la Ley 25326 son:

Licitud: el tratamiento de datos personales debe ser lícito y respetar todos los principios que establece la ley y las reglamentaciones vigentes.

Consentimiento: el tratamiento de datos personales requiere el consentimiento libre, expreso e informado del titular de los datos.

Finalidad: Los datos personales sólo pueden ser tratados para fines específicos y legítimos, previamente informados al titular.

Calidad de los Datos: los datos deben ser exactos, completos y actualizados.

Seguridad: se deben implementar medidas de seguridad adecuadas para proteger los datos personales contra su alteración, pérdida, acceso no autorizado y cualquier otro tratamiento indebido.

Confidencialidad: se debe garantizar la confidencialidad de los datos personales tratados, incluso después de finalizada la relación con el titular de los datos.

Minimización: solo se deben recopilar y procesar los datos personales que sean estrictamente necesarios para cumplir con un propósito específico y legítimo.

Algunos de los derechos que garantiza la Ley 25326 son:

Derecho de acceso: los titulares tienen derecho a acceder a sus datos personales y solicitar información sobre el tratamiento realizado.

Rectificación, actualización o supresión: Los titulares pueden solicitar la rectificación, actualización o supresión de sus datos personales si son inexactos, incompletos, desactualizados o irrelevantes.

Oposición: los titulares tienen derecho a oponerse al tratamiento de sus datos personales para fines publicitarios.

Revocación del consentimiento: los titulares pueden revocar su consentimiento para el tratamiento de sus datos personales en cualquier momento.

Principios relacionados con la transparencia

La transparencia y la explicabilidad de los sistemas de IA, por su parte, suelen ser condiciones previas fundamentales para garantizar el respeto, la protección y la promoción de los derechos humanos, las libertades fundamentales y los principios éticos. Las personas deben tener la oportunidad de solicitar explicaciones e información al responsable de la IA o a las instituciones correspondientes del sector público. Dichos responsables deben informar a los usuarios cuando un producto o servicio se proporcione directamente o con la ayuda de sistemas de IA de manera adecuada y oportuna.

El concepto de transparencia contempla dos dimensiones: como **"proceso"** que implica un conjunto de estrategias, prácticas, instrumentos y procedimientos que las organizaciones deben llevar a cabo de manera transversal en función de disponibilizar información a la ciudadanía y a las autoridades que correspondan. Por otro lado, la transparencia como **"resultado"** de organizaciones capaces de hacer visible los objetivos que persiguen, la gestión de los recursos que manejan, las acciones que realizan y los resultados que logran.

Las acciones de transparencia buscan acercar a la ciudadanía de forma comprensible y accesible la información pertinente respecto al accionar de las organizaciones para facilitar el ejercicio de los derechos individuales, sociales, económicos, ambientales y culturales de la sociedad. En virtud del ejercicio de estos derechos, las personas ciudadanas tienen derecho a solicitar información al organismo responsable de implementación de un sistema que integre Inteligencia Artificial.

Datos personales

Un dato personal consiste en **todo dato que identifica directa o indirectamente a una persona por uno o varios elementos característicos de su identidad,** como por ejemplo su domicilio, profesión, teléfono, situación crediticia. También son datos personales el número de tarjeta de crédito o débito, el historial crediticio, los ingresos salariales de una persona, la dirección IP de la computadora personal, las cookies y otros identificadores en línea, la información de empleo o de un contrato y el historial laboral, las calificaciones en evaluaciones de desempeño, la historia clínica, la ubicación personal en tiempo real y el historial de esa ubicación, entre otros cientos de ejemplos que son utilizados cotidianamente por los sistemas de Inteligencia Artificial y que pueden ser vinculados a una única persona.

Datos sensibles

Por su parte, **los datos sensibles son una categoría de datos que están especialmente protegidos por la normativa porque hacen a la esfera íntima de la persona y poseen potencialidad discriminatoria.** Por ejemplo, datos que puedan revelar el origen étnico; creencias o convicciones religiosas, filosóficas y morales; afiliación sindical u opiniones políticas; salud, discapacidad, orientación sexual, identidad de género y datos genéticos o biométricos cuando puedan revelar datos adicionales cuyo uso pueda resultar potencialmente discriminatorio para la persona y que estén dirigidos a identificarla de manera unívoca.

Asimismo, cabe resaltar que los datos genéticos y los datos biométricos que en muchos casos son empleados en sistemas de Inteligencia Artificial, como por ejemplo a través de técnicas de reconocimiento facial, son considerados de carácter sensible, siempre que identifiquen de manera unívoca a una persona física y de ellos se pueda desprender o revelar información cuyo uso pueda resultar potencialmente discriminatorio para su titular, al menos en Argentina en virtud de la definición aportada por esta AAIP en la Resolución Nº 4/2019, Anexo I, criterio 4.

En el tratamiento de datos sensibles se debe implementar la "responsabilidad reforzada" que implica, entre otras características, mayores niveles de seguridad, confidencialidad, restricciones de acceso, uso y circulación. Ampliaremos este concepto cuando lleguemos al punto de las recomendaciones.

Ejemplos de datos personales y sensibles:

Datos personales, permiten identificar a una persona	**Datos sensibles,** categoría especial de datos personales
• **nombre y apellido,** • **profesión,** • **datos de contacto,** • **correo electrónico,** • **información financiera** • **o cualquier otro dato que permita identificar a una persona**	• **opiniones políticas,** • **origen étnico,** • **convicciones religiosas,** • **información referente a la salud,** • **información de orientación sexual,** • **datos genéticos y biométricos.**

Marco normativo de la protección de datos personales

En Argentina, la protección de datos personales está regulada principalmente por la Ley Nº 25.326, conocida como "Ley de Protección de Datos Personales" del año 2000. Esta ley establece los principios, derechos y obligaciones relacionados con el tratamiento de datos personales en el país, para garantizar el derecho al honor y a la intimidad de las personas.
A su vez, la garantía de habeas data, por el que toda persona tiene derecho a conocer los datos a ella referidos y su finalidad, se encuentra reconocido en el artículo 43 de la Constitución Nacional.

La Agencia de Acceso a la Información Pública (AAIP) es la autoridad de aplicación de la ley 25.326 e implementa y produce normativa referida al tema de su competencia.
La Ley 25326 puede ser consultada en:
https://servicios.infoleg.gob.ar/infolegInternet/anexos/60000-64999/64790/texact.htm

Marco normativo de la Transparencia

La transparencia es una obligación del Sector Público Nacional en virtud de la Ley 27.275 de Derecho de Acceso a la Información Pública. Dicha norma tiene por objeto garantizar el efectivo ejercicio del derecho de acceso a la información pública, promover la participación ciudadana y la transparencia de la gestión pública. Uno de los principios establecidos en la Ley es el de Transparencia y máxima divulgación, que establece que “toda la información en poder, custodia o bajo control del sujeto obligado debe ser accesible para todas las personas”.

La t**ransparencia algorítmica** implica el proceso de disponibilizar información sobre el funcionamiento del sistema, aportando a la explicabilidad sobre su desempeño, para que los productos que integran IA sean comprensibles e interpretables por los usuarios y la ciudadanía en general.
A la vez, supone que el proveedor y/o desarrollador del producto ponga a disposición medios de comunicación y reclamo para poder cuestionar determinados resultados arrojados por el sistema y conocer los potenciales riesgos de su uso.

Decisiones automatizadas y derecho de los titulares

En el año 2019, la Agencia de Acceso a la Información Pública elaboró los "Criterios orientadores e indicadores de mejores prácticas en la aplicación de la Ley Nº 25.326", en los que se destaca que en caso que el responsable de una base de datos tome decisiones basadas en el «tratamiento automatizado» de datos que afecten significativamente al titular de lo datos de forma negativa, el mismo tendrá derecho a solicitar una explicación sobre la lógica aplicada en aquella decisión. Este principio es fundamental en la protección de los derechos de los individuos en el contexto de sistemas de Inteligencia Artificial y toma de decisiones automatizadas. Este derecho reconoce la importancia de que las personas comprendan cómo se toman las decisiones que les afectan, especialmente cuando son tomadas por algoritmos o sistemas de IA sin intervención humana directa.

El Convenio 108 modernizado del Consejo de Europa, que fue ratificado en 2023 por Argentina y está próximo a entrar en vigor, refuerza y actualiza el marco de protección de datos para enfrentar los desafíos planteados por los avances tecnológicos, incluyendo la Inteligencia Artificial y las decisiones automatizadas.

El mismo exige a los responsables del tratamiento el deber de proporcionar información suficiente sobre el procesamiento automatizado, incluyendo la lógica subyacente, sus propósitos y las consecuencias para las personas; que las personas comprendan cómo y por qué se toman decisiones automatizadas que les afectan; que las personas tengan el derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado de datos que les afecte significativamente; y el derecho de obtener una intervención humana, expresar su punto de vista y disputar la decisión.

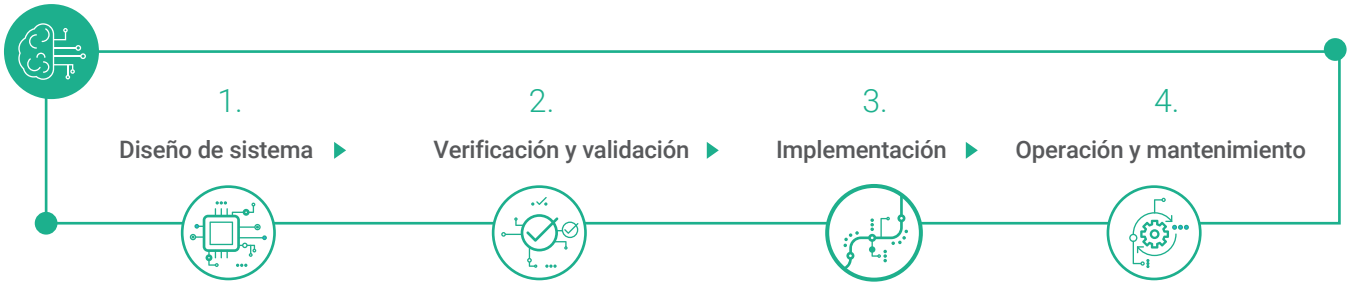
Para más información consultar la Ley 27699 de Protección de las personas con respecto al tratamiento automatizado de datos de carácter personal, 2022.
https://servicios.infoleg.gob.ar/infolegInternet/anexos/375000-379999/375738/norma.htm

06

Recomendaciones de Transparencia y Protección de Datos Personales en el ciclo de vida de Sistemas de IA

Tal como se analizó en la sección anterior, son numerosas las consecuencias y problemáticas que pueden generarse para las personas y la sociedad, si no se garantizan estándares de transparencia y de privacidad al desarrollar y adoptar aplicaciones de IA, especialmente cuando se utilizan en procesos de toma de decisiones que afectan los derechos de las personas.

En principio, es importante aclarar que, cuando se hace referencia al ciclo de vida de los sistemas de IA, se hace alusión a la importancia crucial de realizar un enfoque integral, que aplique a todas las etapas del proyecto desde su diseño hasta su implementación para abordar los riesgos y los posibles impactos adversos en los derechos humanos. Más allá del cumplimiento normativo, esta guía promueve la implementación del uso ético de estas nuevas tecnologías.



Este ciclo de vida está comprendido por las siguientes 4 (cuatro) etapas, las cuales fueron definidas tomando como punto de partida las definiciones de la OCDE[14].

- **Diseño del sistema:** en esta etapa se planifica y diseña el sistema, considerando la selección y procesamiento de los datos de entrada, diseño del modelo algorítmico, entrenamiento del modelo.

- **Verificación y validación:** una vez ya definido el sistema, se procede a testear su correcto funcionamiento.

[14] OECD, Artificial Intelligence in Society, 2019.

- **Implementación:** comprende poner el sistema en funcionamiento. Para esto, es necesario previamente contar con la infraestructura donde irá montado.

- **Operación y monitoreo:** en esta etapa el sistema ya está en plena operatividad, abierto al uso y en constante monitoreo de su funcionamiento.

Criterios aplicables a las actividades dentro del ciclo de vida de los sistemas de inteligencia artificial

La transparencia y la explicabilidad de los sistemas de IA, suelen ser condiciones previas fundamentales para garantizar el respeto, la protección y la promoción de los derechos humanos, las libertades fundamentales y los principios éticos. Las personas deben tener la oportunidad de solicitar explicaciones e información al responsable de la IA o a las instituciones correspondientes del sector público. Dichos responsables deben informar a los usuarios cuando un producto o servicio se proporcione directamente o con la ayuda de sistemas de IA de manera adecuada y oportuna.

Se recomienda la aplicación de los siguientes criterios sugeridos para que las organizaciones -sean públicas o privadas- los implementen internamente desde la etapa de diseño y en las siguientes, con flexibilidad según el contexto y estructura organizativa particular[15].

Etapa 1: Diseño de sistema

Esta etapa implica una serie de actividades que van desde la planificación y el diseño del sistema, a la selección y procesamiento de los datos de entrada, el entrenamiento del sistema y la construcción del modelo algorítmico para lograr los resultados esperados según los objetivos planteados.

[15] Estos principios se formularon tomando como referencia la Ley N° 25326 de Protección de Datos Personales, las buenas prácticas y estándares internacionales que se incorporaron en el Proyecto de Ley de Protección de Datos Personales remitido al Congreso de la Nación (Mensaje 83/2023) y el Convenio 108 modernizado aprobado por nuestro país mediante la Ley N° 27.699. En lo que respecta a Transparencia, se tuvieron en cuenta los estándares internacionales y las obligaciones de transparencia establecidas en la Ley de Acceso a la Información N°27.275 para el sector público. El Proyecto de ley puede ser consultado en: **https://www.argentina.gob.ar/sites/default/files/mensajeyproyecto_leypdp2023.pdf**

Por las razones anteriores, se sugiere que en esta instancia se definan los objetivos del proyecto, los integrantes del equipo que llevará a cabo todo el proceso, los responsables de cada área, los costos y recursos necesarios, la viabilidad del proyecto, la gestión de riesgos, la forma de rendición de cuentas de cada etapa, los requerimientos éticos, la adaptación al contexto y si es posible, la construcción de un sistema inteligente a validar y verificar en la etapa siguiente.

En cuanto a la creación del modelo algorítmico, esto incluye el diseño del modelo y/o adaptación de un modelo algorítmico preexistente, su calibración y entrenamiento, así como la interpretación de los resultados, el cálculo de errores y la precisión del sistema[16].

Protección de datos personales

Principio de responsabilidad proactiva y demostrada

El principio de responsabilidad proactiva y demostrada se ha transformado en los últimos años en un pilar fundamental en la protección de datos personales y está en el corazón de muchas regulaciones de privacidad en todo el mundo. En nuestro país, este principio es un eje fundamental del Proyecto de Ley de Protección de Datos Personales.

Este principio implica que las organizaciones no solo deben cumplir con las leyes y regulaciones de protección de datos, sino también deben ser capaces de demostrar esta conformidad de manera continua y efectiva.

Se trata de un principio transversal a todas las etapas del ciclo de un sistema de Inteligencia Artificial, e incluye la evaluación de impacto y gestión de riesgos, monitoreo y documentación detallada de las acciones, formación del personal, equipos técnicos especializados (también llamados Delegados de protección de datos), gestión de incidentes, auditorias y revisiones, entre otras medidas.

El Proyecto de Ley puede ser consultado en:
https://www.argentina.gob.ar/sites/default/files/mensajeyproyecto_leypdp2023.pdf

[16] OECD, Artificial Intelligence in Society, 2019.

Evaluación de Impacto

La Evaluación de Impacto en la Protección de Datos en la etapa de diseño, es una herramienta fundamental para identificar y mitigar riesgos asociados al tratamiento de datos personales en un sistema de Inteligencia Artificial desde el principio.

Entre los riesgos que permite detectar está la reidentificación de personas, sesgos algorítmicos, uso indebido de datos, entre otros, antes de que el sistema esté desarrollado.

Por lo general aquellas organizaciones que deberían realizar una evaluación de impacto son las que tratan grandes volúmenes de datos, que realizan operaciones con datos sensibles, o que realizan una indagación sistemática y exhaustiva de aspectos personales de las personas que se base en un tratamiento automatizado y semiautomatizado, como la elaboración de perfiles. Pueden ser grandes empresas pero también pequeñas startup cuyo procesamiento de datos pueda conllevar riesgo para la privacidad de las personas y de las instituciones del sector público.

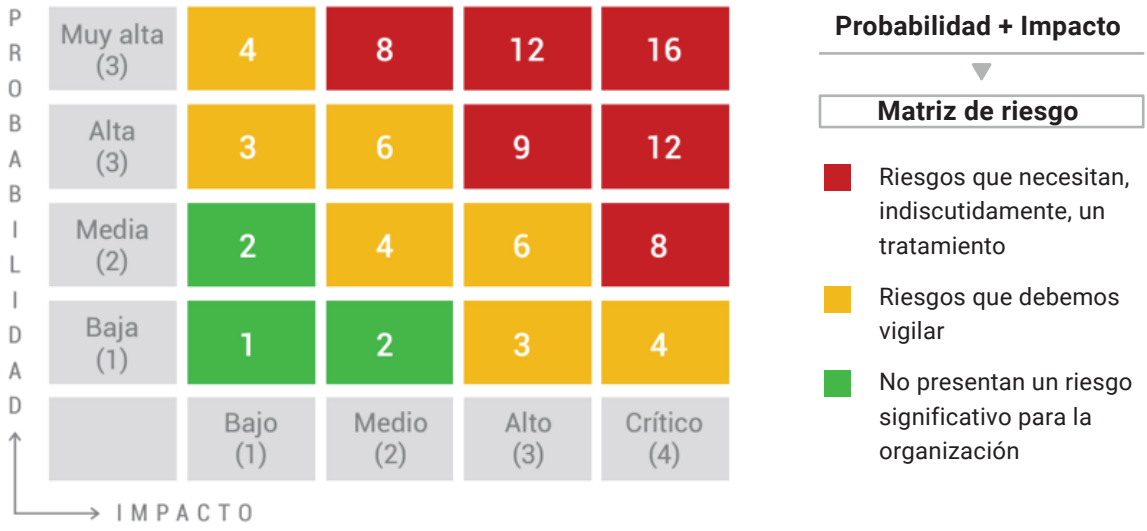
Entre las etapas de una evaluación de impacto está la "gestión de riesgos", proceso mediante el cual se identifica, analiza y valora la probabilidad e impacto de las ocurrencias de amenazas. El objetivo es establecer cuáles son las hipótesis de riesgo para, luego, en una etapa posterior, definir el plan de tratamiento necesario para minimizarlos. En la Guía de Evaluación de Impacto en la Protección de Datos, elaborada por las agencias de Argentina y Uruguay, se propone una escala para la medición de probabilidad de impacto, con cuatro niveles: probabilidad baja, media, alta y crítica[17].

La relación entre los riesgos y la probabilidad de impacto permite construir una matriz de riesgo que permitirá anticipar si el sistema que se está diseñando es de alto, medio o bajo riesgo. Un proyecto que en el diseño de como resultado una matriz de alto riesgo probablemente deba ser desestimado y rediseñado. En la figura 3 se brinda un ejemplo de una matriz de riesgo.

[17] Agencia de Acceso a la Información Pública Argentina y Unidad Reguladora y de Control de Datos Personales Uruguay. Guía de Evaluación de Impacto en la Protección de Dato, 2020.
https://www.argentina.gob.ar/noticias/argentina-y-uruguay-lanzan-la-guia-evaluacion-de-impacto-en-la-proteccion-de-datos

Figura 3: Matriz de riesgos

Si combinamos la probabilidad y el impacto obtenemos la matriz de riesgo, tal como se ilustra a continuación



Fuente: Guía de Evaluación de Impacto en la Protección de Datos, 2020.

Guías y recomendaciones para la evaluación de impacto en un sistema de IA

Para mayor detalle, se recomienda tomar como referencia la evaluación de impacto ética publicada por la UNESCO, la que principalmente tiene como objetivo evaluar si los algoritmos se encuentran alineados con los valores, principios y orientaciones establecidos por su Recomendación. Esta evaluación debería garantizar la transparencia siendo abierta al público la información sobre estos sistemas y cómo se desarrollaron.

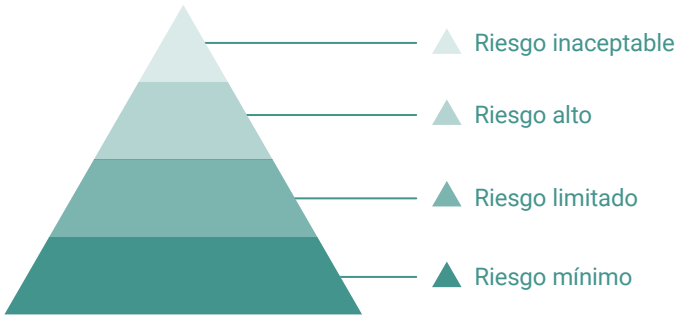
Oportunamente, la AAIP de Argentina y la Unidad Reguladora y de Control de Datos Personales de Uruguay, publicaron una Guía "Evaluación de Impacto en la Protección de Datos". Este instrumento no fue desarrollado específicamente para la evaluación de impacto de sistemas que utilicen la Inteligencia Artificial, pero es aplicable en la medida que estos sistemas utilicen datos personales.

Este instrumento repasa los distintos principios de privacidad y realiza preguntas que el responsable de tratamiento debería considerar para poder abordar adecuadamente los riesgos de privacidad desde instancias tempranas, tal como se detalla en el cuadro a continuación en relación al principio de finalidad.

UNESCO. *Ethical impact assesment. A tool of the Recomendation on the Ethics of Artificial Intelligence.* Disponible en: **https://unesdoc.unesco.org/ark:/48223/pf0000386276**
Evaluación de Impacto en la Protección de Datos, 2020. Disponible en: **https://www.argentina.gob.ar/sites/default/files/guia_final.pdf**

Niveles de riesgo según la ley europea

El Reglamento de Inteligencia Artificial de la Unión Europea aprobado en marzo de 2024, aborda los riesgos que presentan los sistemas de IA, clasificándolos en varios niveles según su potencial impacto en la seguridad:



Fuente: Elaboración propia sobre la base de esquema de Shaping Europe's digital future disponible en: digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

Riesgo inaceptable

Técnicas manipuladoras o engañosas para influir en decisiones; evaluación o clasificación de personas según su comportamiento social, datos personales, que puedan ser usadas para discriminar o perjudicar a la persona.

Ejemplos:
Evaluaciones penales basadas solo en perfiles de personalidad; Inferencia de emociones en entornos laborales o educativos; categorización de personas físicas sobre la base de sus datos biométricos para deducir o inferir su raza, opiniones políticas, entre otros.

Riesgo alto

Riesgo significativo para la salud, la seguridad o los derechos fundamentales. Los sistemas que pueden presentar este tipo de riesgo no están prohibidos, pero deben cumplir estrictos requisitos antes de ser desplegados.

Ejemplos:
El equipamiento médico, vehículos, infraestructuras críticas, etc., para evaluar la solvencia de personas y/o establecer su calificación crediticia.

Riesgo limitado

Sistemas que interactúan con usuarios y notifican información. Tienen requisitos específicos relacionados principalmente con la transparencia y el uso responsable.

Ejemplos:
Asistentes chatbots y asistentes virtuales, entre otros.

Riesgo mínimo

Estos sistemas representan un riesgo bajo o nulo para los derechos fundamentales y la seguridad. No tienen requisitos adicionales más allá del cumplimiento de las normativas generales.

Ejemplos:
IA utilizada en videojuegos o aplicaciones de entretenimiento; filtros de spams en correos electrónicos, entre otros que no afectan derechos fundamentales.

Asimismo, cabe resaltar que los datos genéticos y los datos biométricos que en muchos casos son empleados en sistemas de Inteligencia Artificial, como por ejemplo a través de técnicas de reconocimiento facial, son considerados de carácter sensible, siempre que identifiquen de manera unívoca a una persona física y de ellos se pueda desprender o revelar información cuyo uso pueda resultar potencialmente discriminatorio para su titular, al menos en Argentina en virtud de la definición aportada por esta AAIP en la Resolución Nº 4/2019, Anexo I, criterio 4.

En el tratamiento de datos sensibles se debe implementar la "responsabilidad reforzada" que implica, entre otras características, mayores niveles de seguridad, confidencialidad, restricciones de acceso, uso y circulación. Ampliaremos este concepto en la sección de las recomendaciones.

Ejemplos de impacto material y moral

Alto

Los titulares de datos son afectados de manera significativa, causándoles dificultades y requiriéndoles soluciones complejas.

Ejemplos de impacto material	Ejemplos de impacto moral
• Adjudicación errónea del dinero del titular de datos a otra persona sin compensación. • Dificultades financieras a medio o largo plazo • Pérdida del trabajo • Separación o divorcio • Daño a la propiedad • Pérdida financiera como resultado de un frade	• Daños psicológicos serios (depresión, paranoía, desarrollo de una fobia) • Sensación de invasión de la privacidad con daño irreversible • Sensación de vulnerabilidad por tener que intervenir en un procedimiento judicial • Sensación de violación de los derechos fundamentales (discriminación, libertad de expresión) • Sufrimiento de extorsiones o escraches • Cyberbullying y acoso

Fuente: Guía de Evaluación de Impacto en la Protección de Datos, 2020.

Principio de protección por diseño y defecto

La privacidad “por diseño" es parte de un enfoque integral y de responsabilidad proactiva para la protección de datos personales que se inserta en el diseño y operación de sistemas desde el inicio del proyecto. De esta manera se garantiza que la privacidad y la protección de datos no sean un añadido posterior, sino parte sustancial desde el comienzo.

La privacidad “por defecto” garantiza que desde el propio diseño solo se recopilan, procesan y comparten los datos personales que son estrictamente necesarios y que se opta, en cada nivel, por la opción más protectiva entre todas las alternativas posibles. También deben ser tenidas en cuenta las opciones de configuración abogando por un procesamiento mínimamente intrusivo.

Principio de Licitud

Se debe garantizar la legalidad de la recolección y el procesamiento de los datos que alimentan los sistemas de IA, por lo tanto, debe cumplir con una base jurídica adecuada, en nuestro caso, con la Ley 25326 de Protección de Datos Personales y las normas complementarias.

La base fundamental de un tratamiento legítimo es el consentimiento libre, expreso e informado. Sin embargo, hay excepciones que permiten el tratamiento de datos personales cuando se trata del ejercicio de funciones propias de los poderes del Estado o en virtud de una obligación legal, deriven de una relación contractual, científica o profesional, o las operaciones que realicen las entidades financieras, entre otras.

Además, los archivos y bases de datos que permitan obtener información sobre las personas y alimenten un sistema de IA, sea público o privado, deben estar inscriptos en el Registro Nacional de Bases de Datos Personales[18].

Medidas de seguridad

La seguridad en el tratamiento de datos personales que se utilizan para alimentar un sistema de IA, implica la implementación de medidas técnicas y organizativas adecuadas para proteger los datos personales contra el acceso no autorizado, la divulgación, la alteración y la destrucción.

[18] Toda la información sobre el Registro Nacional de Base de Datos puede encontrarse en: https://www.argentina.gob.ar/aaip/datospersonales

En la etapa de diseño del proyecto, estas medidas deben ser pensadas a partir de la evaluación de impacto, y entre las medidas que suelen estar a disposición están:

- Implementación de controles en los ambientes de desarrollo
- Minimización de datos
- Cifrado de datos
- Control de acceso y de cambios
- Respaldo y recuperación
- Gestión de vulnerabilidades
- Protocolo de conservación y almacenamiento de la información
- Anonimización y seudonimización
- Monitoreo y auditoría
- Protocolo de gestión de incidentes de seguridad
- Protocolo de destrucción de la información

Los daños no deseados deben prevenirse y eliminarse durante todo el ciclo de vida de los sistemas de IA. Las "Medidas de seguridad recomendadas para el tratamiento y conservación de los Datos Personales en medios informatizados" elaborados por la AAIP pueden ser de utilidad al sistematizar etapas y medidas específicas para reducir los riesgos y amenazas a la integridad y preservación de los datos personales[19].

Plazos de conservación

La ley vigente sostiene que los datos deben ser destruidos cuando hayan dejado de ser necesarios o pertinentes a los fines para los cuales hubiesen sido recolectados. En la legislación regional y en gran parte de la internacional existen artículos específicos dedicados a los plazos de conservación de los datos, puesto que cuanto más se conservan sin un fin o propósito de tratamiento legítimo, más riesgos aparecen para la privacidad de las personas. Por lo tanto, ya en el proceso de diseño de un sistema de IA, junto a la o las finalidades propuestas y a la base legal, debe poder establecerse el plazo de conservación de los datos que se van a recolectar.

Principio de calidad

Establecer criterios claros para la calidad de los datos, definiendo las características que deben cumplir los datos recolectados (precisión, integridad, relevancia, etc.).

[19] La Resolución 47/2018 de la Agencia de Acceso a la Información Pública, aunque no esta pensada específicamente para las tecnologías de Inteligencia Artificial, son recomendaciones que puede ser de utilidad para sistematizar una estrategia de seguridad de datos personales.
Las "Medidas de seguridad recomendadas para el tratamiento y conservación de los Datos Personales en medios informatizados", se encuentran a disposición en: **https://servicios.infoleg.gob.ar/infolegInternet/anexos/310000-314999/312662/norma.htm**

Respecto a la recolección y el procesamiento de los datos de entrada, es importante diseñar cómo será el control de calidad de los datos, la documentación de metadatos y las características de la base de datos y cómo se mantendrá en el tiempo.

Minimización

Aunque en la etapa de diseño lo más probable es que se trabaje con datos anonimizados, es importante establecer el tipo de dato que se necesitará recolectar. Al haber establecido el principio de calidad de los datos, podemos diseñar el modelo sobre la base de la menor cantidad de datos posibles.

Se debe minimizar la cantidad de datos personales procesados por las aplicaciones de IA, reduciendo datos innecesarios y/o redundantes durante el desarrollo y fases de entrenamiento. Los datos personales deben ser tratados de manera que sean adecuados, pertinentes y limitados a lo necesario en relación con los fines para los que fueron recolectados. Aplicar este principio ayuda a reducir los riesgos asociados con el manejo de grandes volúmenes de datos personales y a proteger la privacidad de los individuos. En particular, es importante establecer la exclusión total de datos sensibles innecesarios.

Principio de transparencia y explicabilidad

Las buenas prácticas exigen asegurar niveles razonables de apertura y claridad en la gobernanza de las actividades dentro del ciclo de vida de los sistemas de IA. Esto incluye la comprensión y accesibilidad de las partes interesadas a la información sobre los procesos de toma de decisiones, los algoritmos y el funcionamiento interno de los modelos de IA.

Para ello, resulta primordial garantizar la trazabilidad y auditabilidad de los sistemas a lo largo de todo el ciclo de vida, es decir, que se registren y controlen las consideraciones como la procedencia de los datos, la validez de las fuentes de datos, los esfuerzos de mitigación de riesgos, los procesos y las decisiones implementadas para ayudar a la comprensión integral y rendición de cuentas de cómo se derivan los resultados del sistema de IA.

La información vinculada al tratamiento de los datos será transparente si las personas son informadas adecuadamente sobre los detalles de los datos utilizados para producir el sistema, así como informar cuando interactúan directamente con un sistema de Inteligencia Artificial o cuando sus datos personales sean procesados por dichos sistemas.

A su vez, se debe promover la transparencia algorítmica a través de la adopción de medidas para garantizar condiciones adecuadas de explicabilidad de los sistemas, es decir la capacidad de la organización para proporcionar explicaciones claras y comprensibles sobre cómo llega un modelo algorítmico a sus predicciones y decisiones específicas para que las personas comprendan los resultados arrojados y tengan la oportunidad de pedir explicaciones e información a la entidad responsable de la IA.

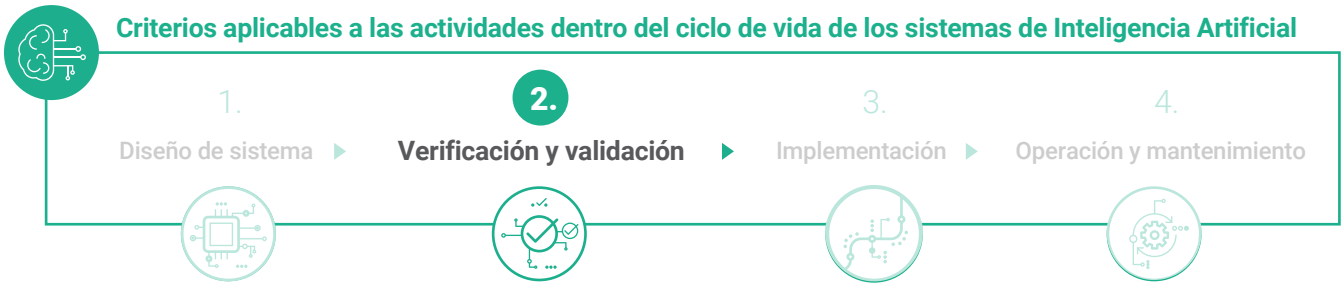
Al respecto, la diversidad de conocimientos y de perspectivas en los equipos humanos es fundamental para resolver estos desafíos, comprender las implicaciones sociales, priorizar soluciones centradas en el usuario, evitar sesgos y discriminación. Es por esto que, es fundamental que el equipo a cargo de estas tareas sea diverso y multidisciplinario, comprendiendo los aspectos éticos básicos involucrados, y en su caso, se consulten comités independientes de expertos de una variedad de campos y/o se involucren con instituciones académicas independientes.

También es deseable la inversión en sensibilización, investigación y formación para garantizar un buen nivel de comprensión de la Inteligencia Artificial y sus posibles efectos.
La etapa de diseño construye los cimientos para llevar a cabo con éxito las etapas posteriores, de allí que resulta crucial incorporar aspectos relacionados a la transparencia y protección de datos personales, considerando la planificación de las etapas venideras.

Síntesis de las recomendaciones de la etapa 1

- Invertir en investigación para conocer buenas prácticas, en formación de los equipos de trabajo y en sensibilización de los actores en general. para garantizar un buen nivel de comprensión de la IA.
- Evaluar el impacto del proyecto en la protección de datos personales desde el inicio, incluyendo riesgos e impactos adversos.
- Considerar la privacidad por diseño y por defecto para que sean objeto de tratamiento solo aquellos datos personales necesarios.
- Garantizar que la fuente de datos que alimentan los sistemas de IA sea lícita.
- Asegurar el tratamiento de datos personales aplicando técnicas de minimización, anonimización, cifrado, control de acceso, conservación, destrucción, entre otras.
- Determinar los plazos de conservación de los datos a recolectar.
- Controlar la calidad de datos y metadatos, aplicando medidas para garantizar que sean ciertos, exactos y actualizados.
- Garantizar la trazabililidad y auditabilidad de los sistemas de IA.
- Asegurar que los datos personales sean para la finalidad que motivaron su obtención.
- Promover la explicabilidad de los sistemas y garantizar la transparencia algorítmica.

Etapa 2: Verificación y validación



En esta instancia se deben realizar las verificaciones y validaciones de los diseños construidos en la etapa anterior, lo que implica ejecutar y ajustar modelos con pruebas para evaluar el rendimiento y el impacto en diversas dimensiones. Se debe verificar el diseño, los datos y los modelos involucrados, aplicando los criterios definidos en la etapa anterior.

Adicionalmente, como una segunda buena práctica para transparentar las decisiones tomadas durante esta fase, se sugiere la publicación de un acta de compromiso ético del proyecto de IA que refleje el compromiso ético del equipo desarrollador, asentando su comprensión y responsabilidad frente a los aspectos éticos mínimos necesarios.

Dicha acta debería detallar los responsables de roles críticos del proyecto como por ejemplo:

- la autoridad a cargo de la decisión de implementar un sistema IA,
- el/la responsable de validar los datos,
- el responsable de monitorear manualmente el comportamiento del sistema,
- el responsable de responder consultas y/o reclamos de información sobre el sistema y de responder ante la ocurrencia de posibles incidentes.

Protección de datos personales

Monitoreo y registro del testeo

El monitoreo y registro del testeo en la etapa de validación de un sistema de IA son procesos importantes para garantizar la calidad y seguridad de los datos, y permite identificar y mitigar riesgos potenciales.

Entre otras acciones posibles, se puede testear mediante:

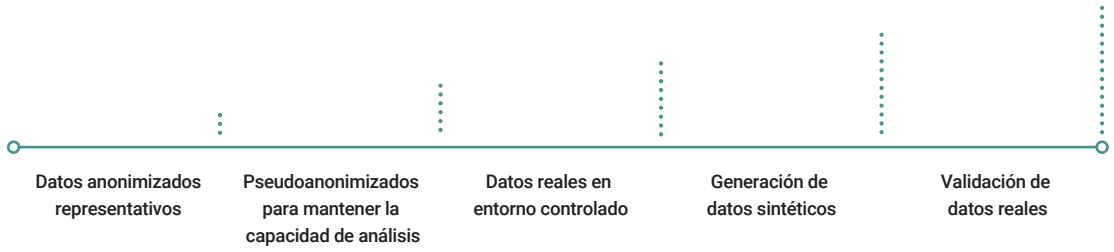
- Métricas de rendimiento
- Monitoreo en tiempo real
- Registro de eventos
- Pruebas de seguridad
- Revisión de sesgos
- Documentación e informes

Medidas de seguridad

Este es el momento preciso donde, además de monitorear y testear las medidas diseñadas, se pueden tomar acciones para remediar vulnerabilidades de seguridad antes del despliegue del sistema y su puesta en operación. Es el momento donde se pueden ajustar los tipos de datos que se utilizan, el tipo de control de acceso, el tipo de cifrado o configurar los entornos virtuales, entre otras medidas.

Anonimización y seudoanonimización

Si en la etapa de diseño se utilizan mayoritariamente datos anonimizados, a medida que se escala en el proyecto, por ejemplo, en la etapa de validación y siguientes, dependiendo del tipo de proyecto, se puede utilizar un sistema escalonado como el siguiente:



Principio de calidad

En la etapa de evaluación es fundamental calibrar la relevancia y precisión de los datos, pertinencia y patrones, y detectar posibles sesgos y errores en el algoritmo y en los resultados.
En cuanto a los sesgos de los datos de entrada de los algoritmos, se deben evaluar en detalle (datos parciales, insuficientes, no actualizados o manipulados), la pertinencia (datos irrelevantes, inconsistentes o incompletos), como así también los patrones (sesgos en la lógica de programación, manipulación de tendencias, inclusión de funciones no previstas) y los errores (condiciones de la operación que reflejan un funcionamiento diferente al previsto y atentan contra las premisas del diseño planteado).

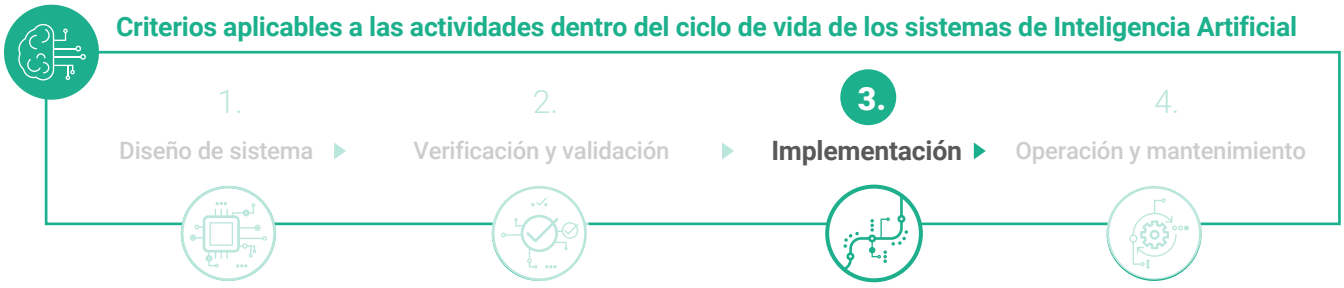
A este fin, se recomienda:

- Llevar un registro de procedencia de los datos;
- Realizar auditorías de los sets de datos utilizados en la creación de algoritmos que ayuden a corregir errores o limitaciones;
- Contar con sets de datos separados para entrenar la máquina, probar y validar el proceso de toma de decisiones;
- Documentar procesos para probar conjuntos de datos contra sesgos y otros resultados inesperados y en su caso, desarrollar un procedimiento para documentar cómo se resolvieron los problemas de calidad de los datos durante el proceso de diseño.

Síntesis de las recomendaciones de la etapa 2

- Publicar un acta de compromiso ético del sistema.
- Evaluar si los algoritmos se encuentran alineados con los valores, principios y orientaciones establecidos.
- Calibrar la relevancia y precisión de los datos e identificar los sesgos, pertinencia, patrones y errores del modelo algorítmico y los resultados.

Etapa 3: Implementación



Para efectivizar la implementación del sistema se requiere determinar la infraestructura tecnológica en la cual se instalará, la cual puede ser contratada a un proveedor externo que brinde servicios de nube o construida con capital propio o bien, una combinación de ambas opciones.

Lo principal al momento de montar el sistema, es asegurar que la infraestructura permita un grado adecuado de seguridad de la información a la vez que todas las acciones y decisiones del sistema sean trazables con identificación de las personas que las llevaron a cabo.

Esto es fundamental para facilitar la tarea de auditorías ante incidentes o como forma de prevención de posibles daños, sesgos u otros errores[20]. En el caso de utilizar servicios tercerizados es importante comprender, previo a la contratación, los niveles de trazabilidad, seguridad de la información y facilidades de auditoría ofrecidas por el prestador para poder evaluar con anticipación si se cumplen los estándares requeridos. El prestador debe comprometese, una vez terminado el contrato, a la devolución de los datos a la institución pública o garantizar su eliminación.

A su vez, otra cuestión a tener en cuenta en esta etapa es la facilidad del usuario para acceder y utilizar el sistema, por lo que se recomienda realizar una evaluación de la experiencia de los usuarios para asegurar el mayor nivel de accesibilidad TIC[21] posible y los estándares de usabilidad y seguridad del usuario.

En esta etapa en particular, la seguridad de la información incluye tener en cuenta la protección de datos personales y respecto a la transparencia, los niveles de trazabilidad y auditabilidad, los que permitirán transparentar todo lo necesario a la ciudadanía y ante autoridades en caso de que sea requerido.

En cuanto a la evaluación con usuarios, es importante dejar documentados los resultados de tales pruebas, así como las medidas a tomar para mejorar dichos sistemas, de acuerdo a los resultados obtenidos.

En todos los casos, previo a lanzar el sistema, es necesario que los resultados y las decisiones estén debidamente identificadas mediante rótulo o etiquetado, tal como las marcas de agua en imágenes, para permitir la detección y transparencia del contenido generado por los sistemas de IA, de manera tal de notificar al usuario y proteger los derechos de autor.

Protección de datos personales

Información

Los sistemas deben contar con una «política de privacidad». Se trata de un documento redactado de manera clara y sencilla que un responsable, público o privado, produce y divulga para informar a los titulares de los datos cómo se recopilan, utilizan, procesan y protegen sus datos personales, así como los derechos que los asisten y otra información que también debe estar

[20] OECD, Artificial Intelligence in Society, 2019.

[21] En Argentina rige la Ley N° 26.653 de "Accesibilidad Web" que promueve la accesibilidad de la información, facilitando especialmente el acceso a todas las personas con discapacidad. **https://www.argentina.gob.ar/normativa/nacional/175694/texto**

disponible para las personas. La política de privacidad es fundamental para cumplir con las regulaciones y establecer una base de confianza con los titulares de los datos. Es también el medio por excelencia, por su naturaleza, para recoger su consentimiento.

La información debe incluir, entre otros datos que sean pertinentes, las finalidades del tratamiento, los datos que se recolectan, los derechos que asisten a los titulares y cómo se pueden ejercer, los plazos de conservación, si las medidas de seguridad son acordes a la finalidad y el tipo de tratamiento, la inscripción en el Registro de la base de datos, las cesiones a terceros que se hagan, si se realizan transferencias internacionales, el derecho de iniciar un reclamo ante la Agencia de Acceso a la Información Pública y los medios y datos de contacto para hacerlo , entre otros[22].

Monitoreo

El monitoreo continuo sobre seguridad, sesgos y transparencia es la continuidad del monitoreo y testeo realizados en la etapa de verificación y validación. Sin embargo, en la etapa de implementación se agregan otros elementos y se amplía el alcance para adaptarse a un entorno de producción. Por ejemplo, se incorporan la detección y respuesta a incidentes, el reentrenamiento y ajuste continuo, la supervisión sobre regulación y cumplimiento, además de la interacción con los usuarios para recibir un feedback y alimentar la mejora continua.

Principio de seguridad y privacidad

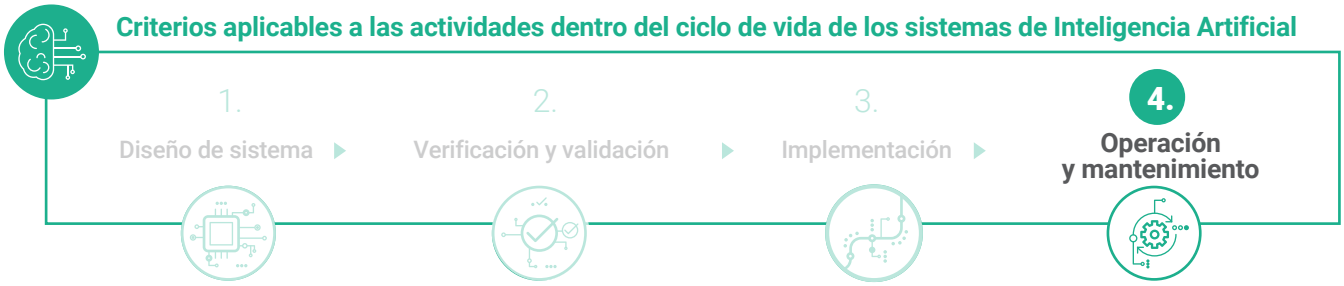
El principio de seguridad y privacidad en la etapa de implementación de un sistema de Inteligencia Artificial es fundamental para proteger los datos personales y garantizar la integridad del sistema. Este proceso puede llevarse a cabo de manera escalonada, incluyendo una fase piloto de implementación, y otra mediante el despliegue completo. En esta última etapa ya se deja de lado el entorno controlado y aislado y la seudoanonimización, para desplegar el sistema alimentado con datos reales y un entorno abierto, por lo que el monitoreo y el control de seguridad pasan a jugar un papel fundamental.

[22] En lo que respecta al sector público, recomendamos tomar como guía la Resolución 40/18 referida a la política modelo de protección de datos en el sector público. Dicha norma establece una serie de pautas para la elaboración de una política de privacidad. Disponible en: **https://www.argentina.gob.ar/normativa/nacional/resoluci%C3%B3n-40-2018-312130/texto**

Síntesis de las recomendaciones de la etapa 3

- Evaluar la experiencia de usuario para asegurar el mayor nivel de accesibilidad y los estándares de usabilidad
- Asegurar la información: protección de datos personales, transparencia, trazabilidad y auditabilidad
- Documentar resultados de las pruebas y medidas a tomar.
- Publicar una Política de Privacidad.

Etapa 4: Operación y mantenimiento



Una vez que el sistema se encuentra funcionando, se requerirán acciones de operación y de mantenimiento tanto de la infraestructura en donde se despliega la solución, así como del propio modelo, dado que muchas veces, se degradan y dejan de responder de manera correcta o quedan desactualizados[23]. En este último supuesto, se recomienda informar sobre nuevas versiones y actualizaciones a la ciudadanía y ante los órganos de autoridad de aplicación, si correspondiera.

Asimismo, durante esta fase es fundamental detectar y en su caso, comunicar cualquier incidente ético y/o de seguridad a la autoridad competente correspondiente teniendo en consideración el sector afectado en particular y los derechos implicados.

Los incidentes éticos pueden ser causados por diferentes motivos, tales como errores humanos involuntarios en alguna de las etapas del ciclo de vida que provoquen un mal funcionamiento del sistema, un uso intencional e indebido de una de las personas de la organización o de los usuarios finales, un ataque interno o externo a la seguridad de la organización o de la información, entre otros.

[23] Los incidentes éticos pueden ser causados por diferentes motivos, tales como errores humanos involuntarios en alguna de las etapas del ciclo de vida que provoquen un mal funcionamiento del sistema, un uso intencional e indebido de una de las personas de la organización o de los usuarios finales, un ataque interno o externo a la seguridad de la organización o de la información, entre otros.

Ficha de transparencia de los Sistemas de IA

En línea con los principios de transparencia algorítmica existe una serie de dimensiones que es posible publicar por parte de las entidades a cargo para darle mayor explicabilidad a los sistemas de Inteligencia Artificial de cara a la ciudadanía[24]. Este listado comprende una serie de características del sistema, que incluye aspectos generales del sistema, como así también de la organización que lo diseña y lo implementa. Asimismo, se brindan aspectos específicos que definen el tipo de tecnología e información útil dirigida al usuario que le corresponde como derecho fundamental para cumplir con los estándares básicos de transparencia.

Ficha del sistema, información a publicar:

1. Caracterización **general** del sistema

- Denominación del sistema
- Organismo, institución o empresa donde se implementa
- Objetivo y problema que resuelve el sistema
- Descripción de funcionalidades, modo de funcionamiento y productos del sistema
- Descripción de beneficios o derechos afectados por el sistema o las sanciones sobre las que interviene en los casos que correspondiese
- Medios de impugnación o reclamo puestos a disposición de la ciudadanía
- Formas de supervisión humana del sistema
- Nombre y contacto de persona responsable de la implementación del sistema
- Datos del proveedor/desarrollador del sistema

2. Caracterización **tecnológica** del sistema

- Tecnología de Inteligencia Artificial utilizada
- Nivel de riesgo del sistema
- Tipo de datos que alimentan el sistema
- Tipo de datos de salida que arroja el sistema
- Fuente de donde se obtienen los datos utilizados

[24] La AAIP ha diseñado una serie de estándares de calidad recomendables, sobre la base de lo trabajado en el marco de la Alianza para la Transparencia Algorítmica (ALTA), de la cual forma parte, y que se construyó como un espacio de colaboración entre el sector público, la academia y la sociedad civil para desarrollar la transparencia algorítmica en Latinoamérica.

Ficha del sistema, información a publicar:

2. Caracterización **tecnológica** del sistema

- Tipo de servidor donde se guardan los datos utilizados y los resultados generados por el sistema (servidores propios, tercerizados locales o tercerizados extranjeros)

3. **Transparencia** del producto en interacción **con la ciudadanía**

- Indicar si el sistema interactúa de algún modo con la ciudadanía
- Indicar si se da a conocer a la ciudadanía que se está interactuando con una máquina y la existencia de una alternativa de interacción con un ser humano, si correspondiese
- Especificar los canales de comunicación y reclamo para la ciudadanía
- Etiquetar los productos generados por el sistema con una leyenda que indique que se elaboraron con Inteligencia Artificial, si correspondiese

Canales de comunicación y reclamo

En conexión con los derechos antes expuestos y como medida de transparencia, es fundamental que las organizaciones responsables de la operación de estos sistemas, habiliten canales de comunicación para que los usuarios puedan realizar consultas y en su caso, interponer reclamos.

A nivel interno en la organización, es deseable que existan procedimientos para que se canalicen estas consultas en tiempo y forma. Asimismo, es importante que la comunicación sea lo más clara y completa utilizando un lenguaje comprensible por una persona no experta en IA. Para los casos dónde existan servicios brindados a través de tecnologías de IA, se sugiere establecer la comunicación a través de una vía humana para atender a las personas que por su perfil o su situación particular, no tengan acceso a los dispositivos y servicios tecnológicos básicos necesarios para ser usuarios de los servicios de IA, o prefieran la atención de una persona humana.

Se resalta como buena práctica -en línea con el principio de responsabilidad proactiva y demostrada- que, dichos procedimientos se encuentren documentados para el eventual supuesto que tenga que demostrarse ante la autoridad de aplicación correspondiente que se ha obrado de manera diligente ante las solicitudes y/o reclamos recepcionados.

Protección de datos personales

Responsabilidad proactiva y demostrada

La responsabilidad en la etapa de operación y mantenimiento, significa adoptar un enfoque continuo y sistemático para garantizar que el sistema opere de manera segura, ética y conforme a las normativas y estándares de privacidad y protección de datos personales. Este tipo de compliance se refiere a la adherencia a las leyes y regulaciones, políticas internas y procedimientos relacionados con la protección de datos personales. Para llevarlo adelante se requieren diversas medidas que, en gran parte, son continuidad de las acciones realizadas en las etapas previas, pero que requiere tareas específicas:

Monitoreo continuo y auditorías regulares

Este monitoreo tiene el objetivo de supervisar el funcionamiento del sistema para detectar y responder a problemas de seguridad, sesgos, y cualquier anomalía en tiempo real. incluye la elaboración de informes periódicos sobre los resultados del monitoreo manual de desempeño del sistema, evaluaciones de impacto, así como las acciones correctivas implementadas. Este monitoreo podría incluir auditorías internas y externas de control así como evaluaciones de impacto regulares, por lo general, anuales.

Personal especializado

Contar con un Delegado de Protección de Datos (DPD), es crucial en esta etapa, debido a la complejidad y la cantidad de tareas y aspectos a considerar en sistemas que manejan un volumen grande de datos personales o datos sensibles, desde el cumplimiento de las regulaciones, hasta la implementación de las mejores prácticas para garantizar la privacidad y seguridad de los datos. El DPD es importante porque el procesamiento de un sistema de IA, por lo general, involucra cesiones a terceros y contratos con encargados de tratamiento de datos que requieren un conocimiento y una actualización constante de todos los procedimientos, y transferencia internacional de datos. También porque procedimientos periódicos como evaluaciones de impacto o la gestión de riesgos o la capacitación del personal se vuelven tareas continuas que requieren especialistas formados que puedan actualizarse de manera continua en las prácticas y nuevas regulaciones.

Formación y concientización continua del personal

La formación continua apunta no solo a los operadores directos, sino a todo el personal de una empresa u organización con sistemas de IA y gestión de grandes volúmenes de datos personales, puesto que el objetivo no es sólo cumplir normas o procedimientos, sino fomentar una cultura de privacidad. Esta cultura genera conciencia y hábitos cotidianos que terminan siendo la mejor cobertura frente a brechas de seguridad, amenazas o errores humanos. A su vez, el equipo de *compliance* y el DPD deben desarrollar calendarios de capacitación y a su interior, promover las certificaciones y estudios especializados sobre la temática.

Principio de seguridad

En esta etapa, las medidas de seguridad se centran en garantizar que el sistema funcione de manera segura y eficaz, protegiendo los datos personales y manteniendo la confianza de los usuarios. Para ello, debe llevar a la práctica las medidas ya elaboradas en la etapa del diseño del proyecto y llevar a cabo el Plan de Respuesta a Incidentes, con procedimientos detallados para la identificación, evaluación, y respuesta a incidentes. Existen diversas herramientas de monitoreo continuo y alertas automáticas.

También se deben aplicar los protocolos para los incidentes de seguridad, donde la respuesta inmediata, la transparencia hacia los ciudadanos y una comunicación fluida al interior y exterior de la organización, ya que son cruciales para hacer frente al mismo. La Resolución 47/018 de la AAIP relativa a los incidentes de seguridad, estructura un modelo de respuesta a los incidentes, para su mejor detección, evaluación, contención y respuesta, como así también las actividades de escalamiento y corrección del entorno técnico y operativo[25].

Plazo de conservación

Si en la etapa de diseño, los plazos de conservación se definieron en función de los fines propuestos, asegurando que no se conserven datos más allá de lo necesario, en esta etapa es importante revisar el registro de plazos y chequear los fines propuestos, pues en ocasiones las modificaciones requieren su actualización.

Además, es importante la implementación de políticas de conservación y eliminación segura de los datos una vez que estos plazos se cumplan.

Derechos de los titulares de los datos

Los derechos de los titulares de los datos buscan asegurar que las personas tengan control sobre sus datos personales y puedan decidir cómo y cuándo se utilizan.

De acuerdo a la Ley 25326, se deben garantizar los siguientes derechos:

Derecho de acceso	Derecho de actualización	Derecho de rectificación	Derecho de Supresión
El titular tiene el derecho a saber si se están tratando sus datos personales y tener acceso a ellos.	El titular del dato tiene derecho a solicitar que sus datos sean actualizados si éstos fueron modificados.	El titular de los datos tiene derecho a la rectificación de sus datos personales cuando estos resulten ser inexactos, falsos, erróneos, incompletos o desactualizados.	El titular tiene derecho a solicitar la supresión de sus datos personales cuando ya no sean necesarios para la finalidad o cuando el titular haya revocado su consentimiento.

[25] AAIP. Resolución 47/2018 “Medidas de seguridad recomendadas para el tratamiento y conservación de los Datos Personales en medios informatizados”.

Nuevos derechos

En el Proyecto de Ley elaborado por la AAIP con aportes de la ciudadanía, y enviado por el Poder Ejecutivo al Congreso de la Nación (mensaje 87/2023)[26], se adicionan los derechos de limitación y portabilidad, y en particular los siguientes vinculados a las decisiones automatizadas:

Derecho a la Limitación del Tratamiento: Permite a las personas solicitar a las organizaciones que el tratamiento de sus datos personales sea limitado o restringido en ciertas circunstancias.

Derecho a la Portabilidad de los Datos: Permite a las personas recibir sus datos personales en un formato estructurado, de uso común y legible por máquina, y a transferirlos a otro responsable del tratamiento si así lo desean.

Derecho de oposición. El titular de los datos puede oponerse al tratamiento, o una finalidad específica de este, si no prestó su consentimiento o si tuvieran por objeto la publicidad, la prospección comercial o la mercadotecnia directa, incluida la elaboración de perfiles.

Derecho de no inferencia y revisión humana: El titular de los datos tiene derecho a no ser objeto de una decisión que le produzca efectos jurídicos perniciosos, lo afecte de forma negativa o tenga efectos discriminatorios, basada, única o parcialmente, en el tratamiento automatizado de datos, incluida la elaboración de perfiles e inferencias. Se entiende por decisiones parcialmente automatizadas o semiautomatizadas a aquellas en las que no hubo intervención humana significativa.

En relación con lo anterior, la persona tiene derecho a solicitar la revisión por una persona humana de las decisiones tomadas sobre la base del tratamiento automatizado o semiautomatizado que afecten a sus intereses, incluidas las decisiones encaminadas a definir sus aspectos personales, profesionales, de consumo, de crédito, de su personalidad u otros.

El Responsable debe proporcionar, siempre que se le solicite, información clara, completa y adecuada sobre los criterios y procedimientos utilizados para la decisión automatizada o semiautomatizada, con observancia de secretos comerciales e industriales.

En caso de no poder proporcionarla por estas razones, la Autoridad de Aplicación puede realizar auditorías para verificar, entre otros, aspectos discriminatorios o de contenido erróneo o sesgado en el tratamiento automatizado o semiautomatizado de información personal.

[26] AAIP, Proyecto de Ley de Protección de Datos Personales (2023) Disponible en: **https://www.argentina.gob.ar/sites/default/files/mensajeyproyecto_leypdp2023.pdf**

El Responsable de tratamiento debe adoptar las medidas adecuadas para salvaguardar los derechos de la persona Titular de los datos; como mínimo, el derecho a obtener intervención humana por parte del Responsable de tratamiento, a expresar su punto de vista y a impugnar la decisión.

El Responsable no puede llevar a cabo tratamientos automatizados o semiautomatizados de datos personales que tengan como efecto la discriminación en detrimento de las personas Titulares de los datos, particularmente si se encuentran basados en alguna de las categorías de datos sensibles[27].

Síntesis de las recomendaciones de la etapa 4

- Actualizar e informar a la ciudadanía el versionado del sistema.
- Documentar el desempeño del sistema, evaluaciones de impacto y acciones correctivas.
- Comunicar si existen incidentes éticos y/o de seguridad a la autoridad competente.
- Designar un (DPD) Delegado de Protección de Datos
- Habilitar canales de comunicación para consultas y reclamos, procedimientos internos y respuestas en lenguaje comprensible para personas no expertas en IA.

[27] Además, cabe remarcar el derecho de acceso a la información pública establecido por la Ley N° 27275, mediante la cual se garantiza el efectivo ejercicio a este derecho, promoviendo la participación ciudadana y la transparencia de la gestión pública. Esto implica que toda persona humana o jurídica, pública o privada, tiene derecho a solicitar a los sujetos obligados información pública y poder realizar reclamos ante la AAIP en caso de incumplimiento.

07
Consideraciones finales

En esta guía se brindaron definiciones y recomendaciones para el uso de una IA responsable, con el objeto de aportar lineamientos unificados sobre las principales cuestiones a tener en cuenta por las organizaciones y facilitar de este modo, su aplicación práctica durante todo el ciclo de vida de estos sistemas, focalizando en reforzar la transparencia y la protección de datos personales a cada paso.

Se espera que las presentes recomendaciones sirvan de orientación al sector privado y público, para que, previo a iniciar cada proyecto y durante su desarrollo, puedan adaptar e implementar los estándares, principios y aspectos detallados para alcanzar una IA responsable, aplicando en cada etapa de manera proactiva, los mayores esfuerzos para promover entre otras cosas, mayor legitimidad, prevención de riesgos y rendición de cuentas.

Estas medidas no solo favorecerán el efectivo ejercicio de derechos de los ciudadanos y el cumplimiento de la normativa, sino que también potenciarán a las organizaciones al comprender más de cerca sus decisiones, lo que mejorará la confianza de sus usuarios y su competitividad en el largo plazo, en esta temática que recién está dando sus primeros pasos y que todo indica, llegó para quedarse.

Si bien a lo largo del presente documento, se relevan distintos antecedentes y estándares, son muchos los debates globales y nacionales que hoy día continúan en curso para fomentar una IA responsable que maximice oportunidades y disminuya riesgos, debido a los constantes y rápidos avances que estos sistemas imprimen en la agenda tanto de los distintos Gobiernos, como de las empresas.

Conscientes del largo camino a transitar atendiendo los distintos desafíos derivados del uso de IA, el “Programa de Transparencia y protección de datos personales en el uso de la IA” y la presente guía, son sólo las primeras líneas de trabajo de fortalecimiento y de soporte a las organizaciones y a la ciudadanía, las que se irán complementando y actualizando.

08
Acrónimos

AAIP: Agencia de Acceso a la Información Pública es la autoridad de aplicación de protección de datos personales y de transparencia en la República Argentina.

ALTA: Alianza para la Transparencia Algorítmica

BID: Banco Interamericano de Desarrollo

CAHAI: Comité Ad hoc sobre Inteligencia Artificial del Consejo de Europa

CAI: Comité de Inteligencia Artificial

CAMIA: Centro Argentino Multidisciplinario de Inteligencia Artificial

DPD: Delegado de Protección de Datos

G7: Grupo de los 7 (Alemania, Canadá, Estados Unidos, Francia, Italia, Japón y Reino Unido)

GIFT: Grupo de investigación de Inteligencia Artificial, Filosofía y Tecnología

GIGAPP: Grupo de investigación en Gobierno, Administración y Políticas Públicas

GPA: Asamblea Global de Privacidad

GPAI: Global Partnership on Artificial Intelligence

IA: Inteligencia Artificial

ICDPPC: International Conference of Data Protection & Provacy Commissioners

OCDE: Organización para la Cooperación y Desarrollo Económico

ONU: Organización de las Naciones Unidas

RIPD: Red Iberoamericana de Protección de Datos Personales

RTA: Red Iberoamericana de Transparencia y Acceso a la Información

SDA: Sistemas de Decisiones Automatizadas

TIC: Tecnologías de la Información y la Comunicación

UE: Unión Europea

UNESCO: Organización de las Naciones Unidas para la educación, la ciencia y la Cultura

URCDP: Unidad Reguladora y de Control de Datos Personales (República Oriental del Uruguay)

09
Referencias bibliográficas

- Agencia de Acceso a la Información Pública Argentina. Mensaje 87/2023 - Proyecto de Ley de Datos Personales, 2023.
  https://www.argentina.gob.ar/sites/default/files/mensajeyproyecto_leypdp2023.pdf
- Agencia de Acceso a la Información Pública Argentina. Resolución Nº 4/2019.
  https://servicios.infoleg.gob.ar/infolegInternet/anexos/315000-319999/318874/norma.htm
- Agencia de Acceso a la Información Pública Argentina. Resolución Nº 40/2018. Política modelo de protección de datos personales para organismos públicos, 2018.
  https://www.argentina.gob.ar/normativa/nacional/resoluci%C3%B3n-40-2018-312130/texto
- Agencia de Acceso a la Información Pública Argentina. Resolución Nº 47/2018. Medidas de seguridad recomendadas para el tratamiento y conservación de los Datos Personales en medios informatizados, 2018.
  https://servicios.infoleg.gob.ar/infolegInternet/anexos/310000-314999/312662/norma.htm
- Agencia de Acceso a la Información Pública Argentina. Resolución Nº 161/2023. Programa de transparencia y protección de datos personales en el uso de la inteligencia artificial, 2023.
  https://www.boletinoficial.gob.ar/detalleAviso/primera/293363/20230904
- Agencia de Acceso a la Información Pública Argentina y Unidad Reguladora y de Control de Datos Personales Uruguay. Guía de Evaluación de Impacto en la Protección de Datos, 2020.
  https://www.argentina.gob.ar/noticias/argentina-y-uruguay-lan-zan-la-guia-evaluacion-de-impacto-en-la-proteccion-de-datos
- Argentina Ley Nacional N° 25.326 - “Protección de Datos Personales", 2000.
  https://servicios.infoleg.gob.ar/infolegInternet/anexos/60000-64999/64790/texact.htm
- Argentina Ley Nacional N° 26.653 - “Accesibilidad de la información en las páginas web", 2010.
  https://www.argentina.gob.ar/normativa/nacional/175694/texto
- Argentina Ley Nacional N° 27.275 - “Derecho de acceso a la información pública", 2016.
  https://servicios.infoleg.gob.ar/infolegInternet/anexos/265000-269999/265949/norma.htm

- Argentina Ley Nacional N° 27.699 - "Protección de las personas con respecto al tratamiento automatizado de datos de carácter personal", 2022.
  https://servicios.infoleg.gob.ar/infolegInternet/anexos/375000-379999/375738/norma.htm
- Benítez Eyzaguirre, Lucía. Ética y transparencia para la detección de sesgos algorítmicos de género, 2019.
  https://doi.org/10.5209/esmp.66989
- Comisión Europea. Declaración de los dirigentes del G7 sobre el proceso de la IA de Hiroshima.
  https://digital-strategy.ec.europa.eu/en/library/g7-leaders-statement-hiroshima-ai-process
- Comisión Europea. Libro blanco sobre la inteligencia artificial "Un enfoque europeo orientado a la excelencia y la confianza, 2020.
  https://commission.europa.eu/document/download/d2ec4039-c5-be-423a-81ef-b9e44e79825b_es?filename=commission-white-paper-artificial-intelligence-feb2020_es.pdf
- Consejo de Europa (Council of Europe). Convención sobre Inteligencia Artificial, Derechos Humanos, Democracia y Estado de Derecho (CAI), 2024.
  https://www.coe.int/en/web/artificial-intelligence/cai#{%22126720129%22:[],%22126720142%22:[1]
- Consejo de Europa (Council of Europe). Convenio 108+, 2018.
  https://rm.coe.int/convention-108-convention-for-the-protection-of-individuals-with-regar/16808b36f1
- Consejo de Europa (Council of Europe). Directrices éticas para una inteligencia artificial confiable, 2019.
  https://op.europa.eu/es/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1
- Consejo de Europa (Council of Europe). Hacia la regulación de sistemas de IA, 2020.
  https://rm.coe.int/prems-107320-gbr-2018-compli-cahai-couv-texte-a4-bat-web/1680a0c17a
- Escuela de Gobierno de la Universidad Adolfo Ibáñez de Chile. Repositorio Algoritmos Públicos. Informe Anual, 2023.
  https://goblab.uai.cl/informe-anual-2023-repositorio-algoritmos-publicos/
- Fundación Sadosky. Innovar con Ciencia de Datos en el Sector Público, 2022.
  https://innovacionpublicacondatos.fundacionsadosky.org.ar/descargar/HojaDeRuta.pdf
- Fundación Vía Libre. Protección legal de datos personales inferidos, 2023.
  https://www.vialibre.org.ar/datos-inferenciales/
- Future of Life Institute. Asilomar AI Principles, 2017.
  https://futureoflife.org/open-letter/ai-principles

- Garrido, Romina; Lapostol, José Pablo y Paz Hermosilla, María. Transparencia algorítmica en el sector público. Consejo para la transparencia de Chile y Gob_Lab de la Universidad Adolfo Ibáñez, 2021.
  https://www.consejotransparencia.cl/wp-content/uploads/estudios/2021/10/ESTUDIO-TRANSPARENCIA-ALGORITMICA-EN-EL-SECTOR-PUBLICO-GOBLAB-cambio-tablas-1.pdf
- Global Partnership on Artificial Intelligence (GPAI). Pacto Global de Inteligencia Artificial, 2022.
  https://www.actuia.com/english/declaration-of-the-ministers-of-the-gpai-members/
- Global Privacy Assembly (GPA). 40th International Conference of Data Protection and Privacy Commissioners. Declaración de ética y protección de datos en inteligencia artificial. Bruselas, 2018.
  http://globalprivacyassembly.org/wp-content/uploads/2018/10/20180922_ICDPPC-40th_AI-Declaration_ADOPTED.pdf
- Global Privacy Assembly (GPA). Joint statement on data scraping and the protection of privacy, 2023.
  https://ico.org.uk/media/about-the-ico/documents/4026232/joint-statement-data-scraping-202308.pdf
- Global Privacy Assembly (GPA). 45th Closed Session of the Global Privacy Assembly. Resolution on Generative Artificial Intelligence System, 2023.
  https://www.edps.europa.eu/system/files/2023-10/edps-gpa-resolution-on-generative-ai-systems_en.pdf
- Grupo de investigación en Gobierno, Administración y Políticas Públicas (GIGAPP). Gobernanza de datos e inteligencia artificial, 2023.
  https://gigapp.org/ewp/index.php/GIGAPP-EWP
- Hermosilla, María; González Alarcón, Natalia; Pombo, Cristina; Sánchez Ávalos, Roberto; Denis, Gabriela; Aracena, Claudio. Banco Interamericano de Desarrollo (BID). Uso responsable de IA para política pública: manual de formulación de proyectos, 2021.
  https://publications.iadb.org/es/uso-responsable-de-ia-para-politica-publica-manual-de-formulacion-de-proyectos
- Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO). Ethical impact assessment: a tool of the Recommendation on the Ethics of Artificial Intelligence, 2023.
  https://unesdoc.unesco.org/ark:/48223/pf0000386276
- Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO). Recomendación sobre la ética de la inteligencia artificial, 2021.
  https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa
- Organización para la Cooperación y el Desarrollo Económico (OCDE). Artificial Intelligence in Society, 2019.
  https://doi.org/10.1787/eedfee77-en

- Organización para la Cooperación y el Desarrollo Económico (OCDE). Recomendación del Consejo sobre Inteligencia Artificial, 2019-2024.
  https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449
- Parlamento Europeo. Reglamento de Inteligencia Artificial, 2024.
  https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_ES.pdf
- Red Iberoamericana de Protección de Datos Personales. Orientaciones específicas para el cumplimiento de los principios y derechos que rigen la protección de los datos personales en los proyectos de Inteligencia Artificial, 2019.
  https://www.redipd.org/sites/default/files/2020-02/guia-orientaciones-espec%C3%ADficas-proteccion-datos-ia.pdf
- Red Iberoamericana de Protección de Datos Personales. Recomendaciones generales para el tratamiento de datos en la inteligencia artificial, 2019.
  https://www.redipd.org/sites/default/files/2020-02/guia-recomendaciones-generales-tratamiento-datos-ia.pdf
- Secretaría de Asuntos Estratégicos de la Presidencia de la Nación Argentina. Resolución Nº 90. Programa de Inteligencia Artificial, 2021.
  https://www.boletinoficial.gob.ar/detalleAviso/primera/253666/20211130
- Secretaría de Innovación Pública de la Jefatura de Gabinete de Ministros de la Nación Argentina. Recomendaciones para una Inteligencia Artificial Fiable, 2023.
  https://www.argentina.gob.ar/sites/default/files/2023/06/recomendaciones_para_una_inteligencia_artificial_fiable.pdf
- Secretaría de Gobierno de Ciencia, Tecnología e Innovación Productiva, Argentina. Plan Nacional de Inteligencia Artificial (ArgenIA), 2019.
  https://oecd-opsi.org/wp-content/uploads/2021/02/Argentina-National-AI-Strategy.pdf
- The White House. Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence, 2023.
  https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/

ANEXO I

Antecedentes y evolución de estándares internacionales

A nivel internacional

En enero de 2017, se llevó a cabo la **Conferencia de Asilomar** organizada por el Instituto "Future of Life" con el objetivo de mostrar la visión de la academia e industria sobre las oportunidades y amenazas que crea la Inteligencia Artificial[28] (IA). En este marco, los participantes realizaron diversos aportes compilados en una lista de 23 principios sobre cómo se debe administrar la IA, basándose en tres ejes: (i) cuestiones de investigación, (ii) ética y valores y (iii) problemas a largo plazo. Algunos de ellos son: seguridad, transparencia, responsabilidad, alineación de valores humanos, libertad y privacidad, beneficio compartido y control humano.

El 8 de abril de 2019, el Grupo de expertos de alto nivel sobre IA del **Consejo de Europa** presentó las "Directrices éticas para una inteligencia artificial confiable", las que detallan que una IA confiable debería ser: (1) legal: respetando todas las leyes y regulaciones aplicables (2) ética: observando los principios y valores éticos y (3) robusta, tanto desde una perspectiva técnica como teniendo en cuenta su entorno social[29].

En mayo de 2019, el **Consejo de la Organización para la Cooperación y Desarrollo Económico (OCDE)** y los países socios, han adoptado formalmente el primer marco normativo universal sobre ética de la IA titulado "Recomendación del Consejo sobre Inteligencia Artificial"[30], un conjunto de directrices y principios para el desarrollo y despliegue de sistemas de IA alineados con los valores y derechos humanos, la transparencia y la rendición de cuentas, que incluye cinco principios clave para la gobernanza de la IA: crecimiento inclusivo, desarrollo sostenible y bienestar; valores y equidad centrados en el ser humano; transparencia y explicabilidad; robustez, seguridad y protección y responsabilidad.

Estas recomendaciones resaltan que la IA debe beneficiar a las personas y al planeta impulsando el crecimiento inclusivo, el desarrollo sostenible y el bienestar; que los sistemas de IA deben respetar el estado de derecho, los derechos humanos, los valores democráticos y la

---

[28] Future of Life Institute. Asilomar AI Principles, 2017. **https://futureoflife.org/open-letter/ai-principles**

[29] Consejo de Europa. Directrices éticas para una inteligencia artificial confiable, 2019. **https://op.europa.eu/es/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1**

[30] La República Argentina tiene en el ámbito general de la OCDE el status de país observador y se encuentra entre los países no miembros que han suscrito estos principios. Recommendation of the Council on Artificial Intelligence, 2019. **https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449**

diversidad; también deben ser transparentes para garantizar que las personas comprendan los resultados y puedan cuestionarlos; ser sólidos, seguros y protegidos durante todo su ciclo de vida; y las partes interesadas deben colaborar para maximizar los beneficios de la IA y minimizar sus riesgos.

En diciembre de 2020, el Comité Ad hoc sobre Inteligencia Artificial (CAHAI) del Consejo de Europa publicó un informe titulado "Hacia la regulación de sistemas de IA", documento que mapea iniciativas internacionales y marcos legales nacionales y lineamientos éticos, así como en el análisis de los riesgos y oportunidades que surgen de la Inteligencia Artificial, en particular su impacto en los derechos humanos, el Estado de derecho y democracia[31].

En 2020 se ha publicado el "Libro blanco sobre la Inteligencia Artificial: un enfoque europeo orientado a la excelencia y la confianza"[32] por parte de la Comisión Europea, el que enuncia siete requisitos esenciales para un ecosistema de confianza de IA: acción y supervisión humana; solidez técnica y seguridad; gestión de la privacidad y de los datos; transparencia; diversidad, no discriminación, equidad; bienestar social y medioambiental y rendición de cuentas. Dicho informe facilita una lista para que las empresas comprueben en la práctica si se cumplen los requisitos.

La **Red Iberoamericana de Protección de Datos Personales,** foro regional constituido por 16 autoridades de protección de datos de 12 países de la región -incluido Argentina que integra su Comité Ejecutivo-, ha elaborado unas "Recomendaciones generales para el tratamiento de datos en la inteligencia artificial"[33], las que brindan sugerencias con enfoque preventivo a quienes desarrollan productos de IA, y de manera complementaria publicaron unas directrices en un documento titulado "Orientaciones específicas para el cumplimiento de los principios y derechos que rigen la protección de los datos personales en los proyectos de Inteligencia Artificial"[34].

En noviembre de 2021, la **UNESCO (United Nations Educational, Scientific and Cultural Organization)** elaboró la primera norma mundial sobre la ética de la IA: la "Recomendación sobre la ética de la inteligencia artificial"[35], marco adoptado por los 193 Estados miembros[36].

---

[31] Towards Regulation of AI Systems. Global perspectives on the development of a legal framework on Artificial Intelligence (AI) systems based on the Council of Europe's standards on human rights, democracy and the rule of law, 2020.
https://rm.coe.int/prems-107320-gbr-2018-compli-cahai-couv-texte-a4-bat-web/1680a0c17a

[32] Comisión Europea. Libro blanco sobre la inteligencia artificial. Un enfoque europeo orientado a la excelencia y la confianza, 2020.
https://commission.europa.eu/document/download/d2ec4039-c5be-423a-81ef-b9e44e79825b_es?filename=commission-white-paper-artificial-intelligence-feb2020_es.pdf

[33] Red Iberoamericana de Protección de Datos Personales. "Recomendaciones generales para el tratamiento de datos en la inteligencia artificial", 2019.
https://www.redipd.org/sites/default/files/2020-02/guia-recomendaciones-generales-tratamiento-datos-ia.pdf

[34] Red Iberoamericana de Protección de Datos Personales. Orientaciones específicas para el cumplimiento de los principios y derechos que rigen la protección de los datos personales en los proyectos de Inteligencia Artificial, 2019.
https://www.redipd.org/sites/default/files/2020-02/guia-orientaciones-espec%C3%ADficas-proteccion-datos-ia.pdf

[35] UNESCO. Recomendación sobre la ética de la inteligencia artificial, 2021. **https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa**

[36] Argentina es Estado miembro desde 1948. **https://www.unesco.org/es/countries**

Esta Recomendación trata las cuestiones éticas relacionadas con el ámbito de la Inteligencia Artificial y sirve de base para guiar a las sociedades para afrontar de manera responsable con acciones, los efectos conocidos y desconocidos de las tecnologías de la IA, con respecto a la gobernanza de datos, el medio ambiente, el género, la educación, la investigación, la salud y el bienestar social, entre otros.

Los principios que declara son los de proporcionalidad e inocuidad, seguridad y protección, equidad y no discriminación, derecho a la intimidad y protección de datos, supervisión y decisión humanas, transparencia y explicabilidad, responsabilidad y rendición de cuentas, sensibilización y educación, sostenibilidad y gobernanza y colaboración adaptativas de múltiples partes interesadas.

También existen numerosas iniciativas de autorregulación que han surgido también desde el sector privado, como el de ocho grandes empresas –GSMA, INNIT, Grupo Lenovo, LG AI Research, Mastercard, Microsoft, Salesforce y Telefónica– que han firmado un acuerdo con la UNESCO para construir una inteligencia artificial más ética, incorporando los valores y principios de la Recomendación de la UNESCO sobre la Ética de la Inteligencia Artificial en el diseño y despliegue de sus propios sistemas de IA.

Los países del G7 (Alemania, Canadá, Estados Unidos, Francia, Italia, Japón y Reino Unido) y la Unión Europea (UE) en mayo de 2023 realizaron una declaración acordando promover un uso "responsable" de la Inteligencia Artificial ante el avance de estos sistemas y su potencial uso para socavar la democracia y vulnerar la privacidad de las personas[37]. En la misma, se citan cinco principios clave para regular el uso de la IA y otras tecnologías emergentes: estado de derecho, garantías legales, democracia y respeto por los derechos humanos y aprovechamiento de oportunidades para promover la innovación.

Por otra parte, se resalta que, la AAIP, en el marco de su trabajo como miembro de la Asamblea Global de Privacidad, ha co-sponsoreado distintas resoluciones relativas a temas de Inteligencia Artificial.

En particular, se resalta la **"Declaración de ética y protección de datos en inteligencia artificial" de octubre 2018**[38] **y posteriormente en octubre del 2023**[39]**,** la que respalda distintos principios rectores para preservar los derechos humanos en el desarrollo de la Inteligencia Artificial tales como considerar las expectativas razonables de las personas garantizando

[37] Comisión Europea. G7 Leaders' Statement on the Hiroshima AI Proces, 2023. **https://digital-strategy.ec.europa.eu/en/library/g7-leaders-statement-hiroshima-ai-process**

[38] Global Privacy Assembly. 40th International Conference of Data Protection and Privacy Commissioners. Declaración de ética y protección de datos en inteligencia artificial. Bruselas, 2018. **http://globalprivacyassembly.org/wp-content/uploads/2018/10/20180922_ICDPPC-40th_AI-Declaration_ADOPTED.pdf**

[39] Global Privacy Assembly. 45th Closed Session of the Global Privacy Assembly. Resolution on Generative Artificial Intelligence System, 2023. **https://www.edps.europa.eu/system/files/2023-10/edps-gpa-resolution-on-generative-ai-systems_en.pdf**

que el uso de sistemas de Inteligencia Artificial sean coherentes con sus fines originales, delimitar y limitar determinados usos, garantizar la rendición de cuentas, establecer procesos de gobernanza demostrables para todos los actores relevantes, como así también proporcionar información adecuada sobre el propósito y sus efectos.

El 30 de octubre de 2023, Estados Unidos emitió una **Orden Ejecutiva sobre IA segura y confiable**[40], con el objetivo de que los desarrollos tecnológicos respeten medidas de seguridad, privacidad, derechos civiles, principios de los consumidores y trabajadores.

El 13 de marzo de 2024, el Parlamento Europeo aprobó una Ley de Inteligencia Artificial (más conocida como "IA Act"), cuyo objetivo es proteger los derechos fundamentales, la democracia, el Estado de derecho y la sostenibilidad medioambiental frente a la IA que entraña un alto riesgo, impulsando al mismo tiempo la innovación. El Reglamento fija una serie de obligaciones para la IA en función de sus riesgos potenciales y su nivel de impacto, por ejemplo, prohibiendo ciertas aplicaciones de Inteligencia Artificial que atentan contra los derechos de la ciudadanía, como los sistemas de categorización biométrica basados en características sensibles y la captura indiscriminada de imágenes faciales de internet o grabaciones de cámaras de vigilancia para crear bases de datos de reconocimiento facial[41].

En este punto, la AAIP ha participado activamente junto con otros países de las reuniones y negociaciones para arribar a un texto consensuado en el marco de la Convención sobre Inteligencia Artificial, Derechos Humanos, Democracia y Estado de Derecho (CAI)[42] del Consejo de Europa.

Por otra parte, cabe resaltar que existen múltiples trabajos de recomendaciones para instrumentar soluciones éticas.

A nivel de organismos internacionales de crédito en la región, es posible consultar los manuales del **Banco Interamericano de Desarrollo** "Uso responsable de IA para política pública: manual de formulación de proyectos"[43], en particular el Manual de proyectos y el Manual de Ciencia de Datos, que cuenta con una "Caja de herramientas humanísticas" desarrollada por el Grupo GIFT de Filosofía, documento que surge en el marco de GuIA, un espacio para el desarrollo de documentos que abordan la ética de la IA, su gobernanza y aplicaciones en América Latina y el Caribe.

---

[40] The White House. Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence, 2023. **https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/**

[41] Ley de Inteligencia Artificial de la Unión Europea, 2024.

[42] Consejo de Europa. Convención sobre Inteligencia Artificial, Derechos Humanos, Democracia y Estado de Derecho (CAI), 2024. **https://www.coe.int/en/web/artificial-intelligence/cai#{%22126720129%22:[],%22126720142%22:[1]**

[43] Banco Interamericano de Desarrollo (BID). Hermosilla, M. et al. Uso responsable de IA para política pública: manual de formulación de proyectos, 2021. **https://publications.iadb.org/es/uso-responsable-de-ia-para-politica-publica-manual-de-formulacion-de-proyectos**

A nivel regional, la **Escuela de Gobierno de la Universidad Adolfo Ibáñez de Chile** a través de su laboratorio de innovación pública GobLab, en conjunto con el Consejo para la Transparencia del gobierno chileno han realizado un estudio exploratorio sobre la existencia y uso de sistemas de decisiones automatizadas (SDA) en el sector público[44], cuyo objetivo es conocer un estado del arte del uso de estos sistemas, partiendo por la información básica respecto de su existencia, identificación, propósito, los datos que utiliza y su información pública disponible. No sólo se indagan sistemas de Inteligencia Artificial, sino que también consideran sistemas de algoritmos secuenciales que resuelven problemas con impacto público. El informe final del estudio cuenta con una parte específica donde propone un estándar de transparencia algorítmica acorde a la normativa chilena en la administración del Estado. A su vez se construyó un repositorio de algoritmos públicos publicado por GobLab, en conjunto con el Consejo para la Transparencia del gobierno chileno[45].

En Colombia, en el marco de la estrategia nacional de Inteligencia Artificial en 2022, se publicó un sitio web donde se presentaba información sobre algunos sistemas de IA implementados por el Gobierno. Posteriormente, se presentó un **trabajo de investigación,** que culminó con la publicación de una nueva base de datos que contiene información sobre 113 (ciento trece) sistemas de decisión automatizada (SDA) utilizados por las entidades públicas colombianas para automatizar o facilitar procesos de toma de decisiones[46].

A nivel nacional

El Plan Nacional de Inteligencia Artificial (ArgenIA, 2019) fue desarrollado por la Secretaría de Gobierno de Ciencia, Tecnología e Innovación Productiva, a través de coordinación de la Secretaría de Planeamiento y Políticas en Ciencia, Tecnología e Innovación Productiva (SPPCTI). Durante 2018 y 2019 se llevaron adelante reuniones, consultas y mesas de trabajo multisectoriales con diferentes instituciones y referentes en la temática.

En 2021, mediante la Resolución N°90, se creó el Programa de Inteligencia Artificial, dentro de la Secretaría de Asuntos Estratégicos de la Presidencia de la Nación. Su objetivo fue la promoción de inteligencia artificial. Se creó el Consejo Económico y Social para promover la IA, que en 2022 creó el Centro Argentino Multidisciplinario de Inteligencia Artificial (CAMIA).

---

[44] Consejo para la Transparencia Chile y Gob Lab UAI. Transparencia algorítmica en el sector público, 2021. **https://goblab.uai.cl/wp-content/uploads/2021/11/ESTUDIO-TRANSPARENCIA-ALGORITMICA-EN-EL-SECTOR-PUBLICO-GOBLAB-vf.pdf**

[45] Consejo para la Transparencia Chile y Gob Lab UAI. Transparencia algorítmica en el sector público. Informe Anual 2023 - Repositorio Algoritmos Públicos, 2023. **https://goblab.uai.cl/informe-anual-2023-repositorio-algoritmos-publicos/**

[46] **https://gigapp.org/ewp/index.php/GIGAPP-EWP**

En 2022 la Argentina firmó su adhesión al Pacto Global de Inteligencia Artificial (GPAI, por sus siglas en inglés), promovido a partir de las recomendaciones de la OCDE y donde participan gobiernos, organismos internacionales y asociaciones civiles y académicas. Al día de hoy incluye a 29 países con fuerte presencia europea y de Japón.

En el mismo año, la Fundación Dr. Manuel Sadosky, una institución público-privada cuyo objetivo es favorecer la articulación entre el sistema científico-tecnológico y la estructura productiva en todo lo referido a la temática de las Tecnologías de la Información y la Comunicación (TIC), publicó una guía denominada "Innovar con Ciencia de Datos en el sector público" [47], la cual propone una hoja de ruta para pensar las políticas y servicios públicos tanto desde su gestión, como desde el diseño y la implementación. También, presenta algunos disparadores para problematizar el uso de la Ciencia de Datos e Inteligencia Artificial e interrogantes para profundizar sobre la calidad de los datos y el impacto de estas herramientas. Fue desarrollado según el marco normativo argentino y tiene en cuenta las particularidades del país. Además, contiene secciones específicas donde detalla recomendaciones respecto a transparencia y protección de datos personales.

En 2023, la Subsecretaría de Tecnologías de la Información dependiente de la Secretaría de Innovación Pública de la Jefatura de Gabinete de Ministros aprobó, a través de la Disposición 2/2023, la guía de Recomendaciones para una Inteligencia Artificial Fiable (Secretaría de Innovación Pública, 2023)[48], cuyo objetivo es establecer reglas claras para garantizar que los beneficios de los avances tecnológicos sean aprovechados por todos los sectores de la sociedad, fortaleciendo el ecosistema científico y tecnológico argentino.

Este manual brinda un marco para la adopción tecnológica de sistemas de Inteligencia Artificial centrada en la ciudadanía y sus derechos. Para ello, realiza recomendaciones subdivididas en cada una de las etapas del ciclo de vida de la IA, es decir, para el diseño del sistema, la verificación y validación, la implementación y la operación y mantenimiento del sistema. Además, considera que existen aspectos transversales a todo el ciclo de la IA, dentro de los cuales se encuentran el derecho a la intimidad y a la protección de datos y por el otro, la transparencia y la explicabilidad.

[47] Fundación Sadosky (2022). Innovar con Ciencia de Datos en el Sector Público. **https://innovacionpublicacondatos.fundacionsadosky.org.ar/descargar/HojaDeRuta.pdf**

[48] Secretaría de Innovación Pública (Mayo 2023). Recomendaciones para una Inteligencia Artificial Fiable. Subsecretaría de Tecnologías de la Información, Jefatura de Gabinete de Ministros de Nación.

Acciones de la AAIP para una IA responsable

La Agencia de Acceso a la Información Pública, consciente del crecimiento exponencial de los procesos de integración de la Inteligencia Artificial a las soluciones tecnológicas en múltiples ámbitos, ha establecido distintas metas estratégicas para abordar los numerosos desafíos existentes en la actualidad.

- A nivel normativo, inició un proceso de debate participativo con diversos sectores de la sociedad que concluyó en una Propuesta de **nueva Ley de Protección de Datos Personales,** enviada al Congreso de la Nación en junio de 2023, que incorpora nuevos derechos y obligaciones para proteger a las personas ante tratamientos automatizados, elaboración de perfiles e inferencias de datos. **(https://www.argentina.gob.ar/sites/default/files/mensajeyproyecto_leypdp2023.pdf).**

- El 17 de abril de 2023 nuestro país depositó el Instrumento de Ratificación del Protocolo Modificatorio del Convenio para la Protección de las Personas con respecto al tratamiento automatizado de datos de carácter personal - **Convenio 108 +**, que incorpora protecciones reforzadas para los ciudadanos ante procesamientos de la información con sistemas de IA.

- La Agencia **participa activamente en el Comité de Inteligencia Artificial del Consejo de Europa,** en la Asamblea Global de Privacidad (GPA por sus siglas en inglés) **y en la Red Iberoamericana de Protección de Datos (RIPD),** entre otros.

- **Forma parte de la Alianza para la Transparencia Algorítmica (ALTA),** que se construyó como un espacio de colaboración entre el sector público, la academia y la sociedad civil en Latinoamérica.

- Asimismo, cabe mencionar que la AAIP lleva adelante una **estrategia federal de concientización y tratamiento de la temática** a través de la organización de diversos encuentros junto a expertos y referentes de las provincias y ciudadanía local.


AAIP
Transparencia
Acceso a la Información
Protección de Datos Personales


Presidencia
de la Nación
República Argentina