.UBAderecho

IALAB

Propuestas de regulación y recomendaciones de inteligencia artificial en el mundo

Síntesis de principales aspectos

THOMSON REUTERS
LA LEY™

Corvalan, Juan Gustavo
Propuestas de regulación y recomendaciones de inteligencia artificial en el mundo : síntesis de principales aspectos / Juan Gustavo Corvalan. - 1a ed. - Ciudad Autónoma de Buenos Aires : La Ley ; Ciudad Autónoma de Buenos Aires : Universidad de Buenos Aires. Facultad de Derecho, 2023.
Libro digital, PDF

Archivo Digital: descarga
ISBN 978-987-03-4613-5

1. Derecho. I. Título.
CDD 341

Tucumán 1471 (C1050AAC) Buenos Aires
Queda hecho el depósito que previene la ley 11.723

Dirección:
Juan Gustavo **Corvalán**

Coordinación:
Mariana **Sánchez Caparrós**
Melisa **Raban**

Equipo de Investigación:
Antonella **Stringhini**
Carina Mariel **Papini**
Giselle **Heleg**
Valentín **Bonato**

.UBAderecho IALAB

ÍNDICE

A. RESUMEN EJECUTIVO

B. CUADRO COMPARATIVO: PARLAMENTO EUROPEO, CHILE,

C. RESUMEN DE NORMATIVAS Y RECOMENDACIONES

1. PARLAMENTO EUROPEO

a. Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se e establecen normas armonizadas en materia de inteligencia artificial adoptada por la Comisión Europea

b. PARLAMENTO EUROPEO. Orientación General adoptada por Consejo con r relación al proyecto de Reglamento por el que se establecen normas armonizadas en materia de inteligencia artificial

c. PARLAMENTO EUROPEO. Posición negociadora aprobada por el Parlamento Europeo con relación al proyecto de Reglamento por el que se establecen normas armonizadas en materia de inteligencia artificial

2. OCDE

3. CHILE

4. CALIFORNIA

5. CHINA

a. Medidas Administrativas para los Servicios de Inteligencia Artificial Generativa (Borrador para Comentarios)

b. Medidas Provisionales para la Gestión de Servicios de Inteligencia Artificial Generativa

6. BRASIL

7. ARGENTINA

a. Recomendaciones para una IA fiable. Disposición N° 2/2023 Subsecretaría de Tecnologías de la Información

b. Banco Central de la República Argentina. Comunicación A 7724 “Requisitos mínimos para la gestión y control de los riesgos de tecnología y seguridad de la información”

8. NUEVA YORK

9. OTROS ENFOQUES. Documentos relacionados

a. Un enfoque favorable a la innovación para la regulación de la IA en Reino Unido de Gran Bretaña e Irlanda del Norte

b. Libro Blanco de Inteligencia Artificial de Japón

D. CONCLUSIONES

RESUMEN EJECUTIVO

El presente documento contiene un resumen de varias recomendaciones y propuestas de regulación de inteligencia artificial que actualmente se debaten en la región y en Europa.

Para seleccionar qué documentos analizar tuvimos en cuenta, en primer lugar, la necesidad de conocer cuáles son los principales aspectos del Reglamento que se debate en el ámbito del Parlamento Europeo, ya que se trata de una regulación que alcanzará a los 26 países que conforman la Unión Europea. Y además, desde hace varios años esta institución se encuentra a la vanguardia en el desarrollo de diversos instrumentos que buscan posicionar a Europa como centro de la inteligencia artificial fiable en el mundo.

En segundo término, se seleccionaron proyectos regulatorios de países de la región como Chile y Brasil, para conocer los principales puntos que se abordan en el espacio geográfico más cercano a nosotros.

En relación a Argentina, se relevaron dos documentos de rango infra legal. La Disposición N° 2/2023 sobre Recomendaciones para una Inteligencia Artificial Fiable, cuyo ámbito de aplicación es el Sector Público Nacional, y la Comunicación "A" 7724 del Banco Central de la República Argentina "sobre Requisitos mínimos para la gestión y control de los riesgos de tecnología y seguridad de la información", destinada a las instituciones reguladas por aquella entidad.

Sin perjuicio de esta selección, entendimos importante divulgar las principales recomendaciones de la Organización para la Cooperación y el Desarrollo Económicos (OCDE), así como el borrador de medidas para gestionar los servicios de inteligencia artificial generativa que presentó China el pasado mes de abril, para ampliar la diversidad en la mirada de este documento.

Por el mismo motivo, y por su especificidad temática, también se aporta un breve resumen de la Ley 144/2021 de la Ciudad de Nueva York, que se aplica a las empresas que utilicen inteligencia artificial en los procesos de contratación de personal, como así también se incluye el proyecto de ley AB331 sobre herramientas de decisión automatizada de California.

Finalmente, se releva el contenido de dos documentos (del Reino Unido y de Japón, respectivamente), que estimamos oportuno incorporar dado que, si bien no contienen propuestas de regulación específicas de la inteligencia artificial, propician nuevos enfoques para el desarrollo de estrategias nacionales de inteligencia artificial, al esbozar una serie de principios y recomendaciones a ser tenidos en consideración en su desarrollo y aplicación.

Además de organizar la información en un índice temático, se diagramó un "Cuadro Comparativo de Regulaciones en materia de IA" que muestra cómo algunos países de la región y Europa pretenden regular algunas de las categorías jurídicas más relevantes en cuanto al desarrollo de la IA.

CUADRO COMPARATIVO DE REGULACIONES EN MATERIA DE IA

CATEGORIAS COMPARADAS		PARLAMENTO EUROPEO (*)	CHILE	BRASIL
PUNTOS RELEVANTES	PRÁCTICAS PROHIBIDAS	1. Técnicas intencionalmente manipuladoras o engañosas. 2. Grupo vulnerables. 3. Identificacion biométrica remota, en tiempo real, en espacios públicos, y a posteriori con la excepcion de usos policiales en delitos graves. 4. Sistemas de categorizacion biometrica que califa a las personas en funcion de sus atributos o caracteristicas sensibles (ej. raza, genero, etc.). 5. Calificacion social. 6.Sistemas policiales predictivos. 7. Base de datos s/reconocimiento facial (mediante la extraccion no dirigida de imagenes faciales de internet o circuito cerrado de tv). 8. Sistemas que interfieran en las emociones de las personas fisicas en la gestion de fronteras, lugar de trabajo y e instuciones educativas.	1. Daño fisico o psicologico. 2. Grupos Vulnerables. 3. Calificacion social. 4. Identificacion biometrica en espacios pùblicos (salvo cuestiones de seguridad publica).	1. Técnicas subliminales que tienen como objetivo (o producen efecto) inducir a una persona física a comportarse de forma dañina o peligrosa para su salud o seguridad(..); 2. Sistemas de IA que exploten las vulnerabilidades de grupos específicos de personas físicas, tales como las asociadas a su edad o discapacidad física o psíquicos, con el fin de inducirlos a comportarse de una manera nociva para su salud o la seguridad (..); 3. Uso de sistemas de IA por el gobierno, para evaluar, clasificar o jerarquizar las personas naturales, con base en su comportamiento social o atributos de su personalidad, a través de la puntuación universal, para el acceso a los bienes y servicios y políticas públicas, de forma ilegítima o desproporcionada.
	ALTO RIESGO	Perjuicio en la salud, seguridad, medio ambiente, derechos fundamentales. 1. Identificacion biometrica. 2. Gestion del trafico terrestre o areo y suministro de agua , electricidad y gas. 3. Educacion: evaluacion, admision; etc. 4. Empleo: admision, ascensos, seleccion, 5. Servicios escenciales: viviendia, salud, internet,evaluacion crediticio o seguros etc. 6. Utilizado por autoridades judiciales y policiales. 7. Asilo, migraciones, fronteras. 8. Administracion de Justicia y procesos democraticos.	1. Identificacion biometrica en espacios privados. 2. Suministro de agua, luz, gas. 3. Educacion. Asignacion y evaluacion. 4. Empleo: selección y contratacion, seguimiento y evaluacion. 5. Acceso a prestaciones sociales/calificacion crediticias. 6. Pretaciones enc casos de emergencias sanitarias. 7. Evaluacion previa en casos de delitos (incluye utilizacion en la investigacion. 8. Control de asilo, fronteras y migraciones. 9. Perjucio a la salud, seguridad o derechos fundamentales.	1. Gestion de la seguridad, transito, sumistro de agua, luz. 2. Educacion: ingreso/admision. 3. Empleo: reclutamiento, preseleccion, evalucion, promociones, control. 4. Evaluacion crediticia. 5.Emergencias medicas/incendios. 6. Administracion de justicia. 7. Vehiculos autonomos. 8. Area de salud. 9. Sistemas de identificacion biometrica. 10. Investigacion y evaluacion criminal y seguridad publica. 11. Gestion migratoria y fronteras.
	IA GENERATIVA	Obligaciones especiales (ademas de las comunes a todos los sistemas) 1. Transparencia. 2. Capacitar y diseñar de tal manera que se garantice la libertad de expresion y los derechos fundamentales. 3. Publicas base de datos que puedan comprometer derecho de autor.		
	SANDBOX REGULATORIO	Entorno controlado que fomente la innovación y facilite el desarrollo, la prueba y la validación de sistemas innovadores de IA.		Se podrá autorizar la operación del entorno regulatorio experimental a la innovación en inteligencia artificial (sandbox regulatorio) para las entidades que lo soliciten y completen ciertos requisitos.

(*) Segun posicion de negociacion aprobada por el parlamento en la sesion del 14 de junio de 2023.

RESUMEN DE REGULACIONES Y RECOMENDACIONES DE IA EN EL MUNDO

1.PARLAMENTO EUROPEO

a. Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial adoptada por la Comisión Europea[1]

Si bien el debate en torno al impacto y la regulación de la IA comenzó un tiempo antes, en abril de 2021, la Comisión Europea adoptó la propuesta de Reglamento por la que se establecen normas armonizadas en materia de inteligencia artificial, la cual fue presentada al Parlamento Europeo y al Consejo. Este proyecto también es conocido como ley o reglamento de inteligencia artificial.

Se trata de una propuesta cuyo objetivo está en desarrollar un ecosistema de confianza a través de la propuesta de un marco jurídico para lograr una IA fiable. La cuestión viene debatiéndose en el ámbito del Consejo y el Parlamento desde hace un tiempo.

A continuación una breve reseña de los aspectos más importantes del proyecto de Reglamento preparado por la Comisión Europea[2]:

(1) PRÁCTICAS DE INTELIGENCIA ARTIFICIAL PROHIBIDAS

1. Estarán prohibidas las siguientes prácticas de inteligencia artificial:

a) La introducción en el mercado, la puesta en servicio o la utilización de un sistema de IA que se sirva de técnicas subliminales que trasciendan la conciencia de una persona para alterar de manera sustancial su comportamiento de un modo que provoque o sea probable que provoque perjuicios físicos o psicológicos a esa persona o a otra.

b) La introducción en el mercado, la puesta en servicio o la utilización de un sistema de IA que aproveche alguna de las vulnerabilidades de un grupo específico de personas debido a su edad o discapacidad física o mental para alterar de manera sustancial el

[1] *"Propuesta de Reglamento de la Inteligencia Artificial"*, Parlamento Europeo, 21 de abril de 2021, disponible en: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206
[2] *"Propuesta de Reglamento de la Inteligencia Artificial"*, Parlamento Europeo, 21 de abril de 2021, disponible en: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206

comportamiento de una persona que pertenezca a dicho grupo de un modo que provoque o sea probable que provoque perjuicios físicos o psicológicos a esa persona o a otra.

c) La introducción en el mercado, la puesta en servicio o la utilización de sistemas de IA por parte de las autoridades públicas o en su representación con el fin de evaluar o clasificar la fiabilidad de personas físicas durante un período determinado de tiempo atendiendo a su conducta social o a características personales o de su personalidad conocidas o predichas, de forma que la clasificación social resultante provoque una o varias de las situaciones siguientes:

> i) un trato perjudicial o desfavorable hacia determinadas personas físicas o colectivos enteros en contextos sociales que no guarden relación con los contextos donde se generaron o recabaron los datos originalmente;
>
> ii) un trato perjudicial o desfavorable hacia determinadas personas físicas o colectivos enteros que es injustificado o desproporcionado con respecto a su comportamiento social o la gravedad de este.

d) El uso de sistemas de identificación biométrica remota «en tiempo real» en espacios de acceso público con fines de aplicación de la ley, salvo y en la medida en que dicho uso sea estrictamente necesario para alcanzar uno o varios de los objetivos siguientes:

> i) la búsqueda selectiva de posibles víctimas concretas de un delito, incluidos menores desaparecidos;
>
> ii) la prevención de una amenaza específica, importante e inminente para la vida o la seguridad física de las personas físicas o de un atentado terrorista;
>
> iii) la detección, la localización, la identificación o el enjuiciamiento de la persona que ha cometido o se sospecha que ha cometido alguno de los delitos mencionados en el artículo 2, apartado 2, de la Decisión Marco 2002/584/JAI del Consejo, para el que la normativa en vigor en el Estado miembro implicado imponga una pena o una medida de seguridad privativas de libertad cuya duración máxima sea al menos de tres años, según determine el Derecho de dicho Estado miembro.

2. El uso de sistemas de identificación biométrica remota «en tiempo real» en espacios de acceso público con fines de aplicación de la ley para conseguir cualquiera de los objetivos mencionados en el apartado 1, letra d), tendrá en cuenta los siguientes aspectos:

a) la naturaleza de la situación que dé lugar al posible uso, y en particular la gravedad, probabilidad y magnitud del perjuicio que se produciría de no utilizarse el sistema;

b) las consecuencias que utilizar el sistema tendría para los derechos y las libertades de las personas implicadas, y en particular la gravedad, probabilidad y magnitud de dichas consecuencias. Además, el uso de sistemas de identificación biométrica remota «en tiempo real» en espacios de acceso público con fines de aplicación de la ley para cualquiera de los objetivos mencionados en el apartado 1, letra d), cumplirá salvaguardias y condiciones necesarias y proporcionadas en relación con el uso, en particular en lo que respecta a las limitaciones temporales, geográficas y personales.

3. Con respecto al apartado 1, letra d), y el apartado 2, cualquier uso concreto de un sistema de identificación biométrica remota «en tiempo real» en un espacio de acceso público con fines de aplicación de la ley estará supeditado a la concesión de una autorización previa por parte de una autoridad judicial o una autoridad administrativa independiente del Estado miembro donde vaya a utilizarse dicho sistema, que la otorgarán previa solicitud motivada y de conformidad con las normas detalladas del Derecho interno mencionadas en el apartado

4. No obstante, en una situación de urgencia debidamente justificada, se podrá empezar a utilizar el sistema antes de obtener la autorización correspondiente, que podrá solicitarse durante el uso o después de este. La autoridad judicial o administrativa competente únicamente concederá la autorización cuando esté convencida, atendiendo a las pruebas objetivas o a los indicios claros que se le presenten, de que el uso del sistema de identificación biométrica remota «en tiempo real» es necesario y proporcionado para alcanzar alguno de los objetivos que figuran en el apartado 1, letra d), el cual se indicará en la solicitud. Al pronunciarse al respecto, la autoridad judicial o administrativa competente tendrá en cuenta los aspectos mencionados en el apartado 2. 4.

Los Estados miembros podrán decidir contemplar la posibilidad de autorizar, ya sea total o parcialmente, el uso de sistemas de identificación biométrica remota «en tiempo real» en espacios de acceso público con fines de aplicación de la ley dentro de los límites y en las condiciones que se indican en el apartado 1, letra d), y los apartados 2 y 3.

A tal fin, tendrán que establecer en sus respectivos Derechos internos las normas detalladas necesarias aplicables a la solicitud, la concesión y el ejercicio de las

autorizaciones a que se refiere el apartado 3, así como a la supervisión de estas. Dichas normas especificarán también para cuáles de los objetivos enumerados en el apartado 1, letra d), y en su caso en relación con cuáles de los delitos indicados en su inciso iii), se podrá autorizar que las autoridades competentes utilicen esos sistemas con fines de aplicación de la ley

(2) SISTEMAS DE ALTO RIESGO

2.1 Reglas de clasificación para los sistemas de IA de alto riesgo

1. Un sistema de IA se considerará de alto riesgo cuando reúna las dos condiciones que se indican a continuación, con independencia de si se ha introducido en el mercado o se ha puesto en servicio sin estar integrado en los productos que se mencionan en las letras a) y b):

a) el sistema de IA está destinado a ser utilizado como componente de seguridad de uno de los productos contemplados en la legislación de armonización de la Unión que se indica en el anexo II, o es en sí mismo uno de dichos productos;

b) conforme a la legislación de armonización de la Unión que se indica en el anexo II, el producto del que el sistema de IA es componente de seguridad, o el propio sistema de IA como producto, debe someterse a una evaluación de la conformidad realizada por un organismo independiente para su introducción en el mercado o puesta en servicio.

2. Además de los sistemas de IA de alto riesgo mencionados en el apartado 1, también se considerarán de alto riesgo los sistemas de IA que figuran en el anexo III.

ANEXO III SISTEMAS DE IA DE ALTO RIESGO

Los sistemas de IA de alto riesgo con arreglo al artículo 6, apartado 2, son los sistemas de IA mencionados en cualquiera de los ámbitos siguientes:

1. Identificación biométrica y categorización de personas físicas: a) sistemas de IA destinados a utilizarse en la identificación biométrica remota «en tiempo real» o «en diferido» de personas físicas.

2. Gestión y funcionamiento de infraestructuras esenciales: a) sistemas de IA destinados a utilizarse como componentes de seguridad en la gestión y funcionamiento del tráfico rodado y el suministro de agua, gas, calefacción y electricidad.

3. Educación y formación profesional:

a) sistemas de IA destinados a utilizarse para determinar el acceso o la asignación de personas físicas a los centros de educación y formación profesional;

b) sistemas de IA destinados a utilizarse para evaluar a los estudiantes de centros de educación y formación profesional y para evaluar a los participantes en pruebas generalmente necesarias para acceder a centros de educación.

4. Empleo, gestión de los trabajadores y acceso al autoempleo:

a) sistemas de IA destinados a utilizarse para la contratación o selección de personas físicas, especialmente para anunciar puestos vacantes, clasificar y filtrar solicitudes o evaluar a candidatos en el transcurso de entrevistas o pruebas;

b) IA destinada a utilizarse para tomar decisiones relativas a la promoción y resolución de relaciones contractuales de índole laboral, a la asignación de tareas y al seguimiento y evaluación del rendimiento y la conducta de las personas en el marco de dichas relaciones.

5. Acceso y disfrute de servicios públicos y privados esenciales y sus beneficios:

a) sistemas de IA destinados a ser utilizados por las autoridades públicas o en su nombre para evaluar la admisibilidad de las personas físicas para acceder a prestaciones y servicios de asistencia pública, así como para conceder, reducir, retirar o recuperar dichas prestaciones y servicios;

b) sistemas de IA destinados a utilizarse para evaluar la solvencia de personas físicas o establecer su calificación crediticia, salvo los sistemas de IA puestos en servicio por parte de proveedores a pequeña escala para su uso propio;

c) sistemas de IA destinados a utilizarse para el envío o el establecimiento de prioridades en el envío de servicios de primera intervención en situaciones de emergencia, por ejemplo bomberos y servicios de asistencia médica.

6. Asuntos relacionados con la aplicación de la ley:

a) sistemas de IA destinados a utilizarse por parte de las autoridades encargadas de la aplicación de la ley para llevar a cabo evaluaciones de riesgos individuales de personas físicas con el objetivo de determinar el riesgo de que cometan infracciones penales o reincidan en su comisión, así como el riesgo para las potenciales víctimas de delitos;

b) sistemas de IA destinados a utilizarse por parte de las autoridades encargadas de la aplicación de la ley como polígrafos y herramientas similares, o para detectar el estado emocional de una persona física;

c) sistemas de IA destinados a utilizarse por parte de las autoridades encargadas de la aplicación de la ley para detectar ultra falsificaciones a las que hace referencia el artículo 52, apartado 3;

d) sistemas de IA destinados a utilizarse por parte de las autoridades encargadas de la aplicación de la ley para la evaluación de la fiabilidad de las pruebas durante la investigación o el enjuiciamiento de infracciones penales;

e) sistemas de IA destinados a utilizarse por parte de las autoridades encargadas de la aplicación de la ley para predecir la frecuencia o reiteración de una infracción penal real o potencial con base en la elaboración de perfiles de personas físicas, de conformidad con lo dispuesto en el artículo 3, apartado 4, de la Directiva (UE) 2016/680, o en la evaluación de rasgos y características de la personalidad o conductas delictivas pasadas de personas físicas o grupos;

f) sistemas de IA destinados a utilizarse por parte de las autoridades encargadas de la aplicación de la ley para la elaboración de perfiles de personas físicas, de conformidad con lo dispuesto en el artículo 3, apartado 4, de la Directiva (UE) 2016/680, durante la detección, la investigación o el enjuiciamiento de infracciones penales;

g) sistemas de IA destinados a utilizarse para llevar a cabo análisis sobre infracciones penales en relación con personas físicas que permitan a las autoridades encargadas de la aplicación de la ley examinar grandes conjuntos de datos complejos vinculados y no vinculados, disponibles en diferentes fuentes o formatos, para detectar modelos desconocidos o descubrir relaciones ocultas en los datos.

7. Gestión de la migración, el asilo y el control fronterizo:

a) sistemas de IA destinados a utilizarse por parte de las autoridades públicas competentes como polígrafos y herramientas similares, o para detectar el estado emocional de una persona física;

b) sistemas de IA destinados a utilizarse por parte de las autoridades públicas competentes para evaluar un riesgo, como un riesgo para la seguridad, la salud o relativo a la inmigración ilegal, que plantee una persona física que pretenda entrar o haya entrado en el territorio de un Estado miembro;

c) sistemas de IA destinados a utilizarse por parte de las autoridades públicas competentes para la verificación de la autenticidad de los documentos de viaje y los documentos justificativos de las personas físicas y la detección de documentos falsos mediante la comprobación de sus elementos de seguridad;

d) sistemas de IA destinados a ayudar a las autoridades públicas competentes a examinar las solicitudes de asilo, visado y permisos de residencia, y las reclamaciones asociadas con respecto a la admisibilidad de las personas físicas solicitantes.

8. Administración de justicia y procesos democráticos: sistemas de IA destinados a ayudar a una autoridad judicial en la investigación e interpretación de hechos y de la ley, así como en la aplicación de la ley a un conjunto concreto de hechos.

(3) REQUISITOS PARA LOS SISTEMAS DE IA DE ALTO RIESGO

Los sistemas de IA de alto riesgo cumplirán los requisitos que se definen en el presente capítulo. A la hora de verificar su cumplimiento se tendrán en cuenta la finalidad prevista del sistema de IA de alto riesgo y el sistema de gestión de riesgos al que se refiere el artículo 9.

Artículo 9 Sistema de gestión de riesgos

1. Se establecerá, implantará, documentará y mantendrá un sistema de gestión de riesgos asociado a los sistemas de IA de alto riesgo.

2. El sistema de gestión de riesgos consistirá en un proceso iterativo continuo que se llevará a cabo durante todo el ciclo de vida de un sistema de IA de alto riesgo, el cual requerirá actualizaciones sistemáticas periódicas. Constará de las siguientes etapas: a) la identificación y el análisis de los riesgos conocidos y previsibles vinculados a cada sistema de IA de alto riesgo; b) la estimación y la evaluación de los riesgos que podrían surgir cuando el sistema de IA de alto riesgo en cuestión se utilice conforme a su finalidad prevista y cuando se le dé un uso indebido razonablemente previsible; c) la evaluación de otros riesgos que podrían surgir a partir del análisis de los datos recogidos con el sistema de seguimiento posterior a la comercialización al que se refiere el artículo 61; d) la adopción de medidas oportunas de gestión de riesgos con arreglo a lo dispuesto en los apartados siguientes.

3. Las medidas de gestión de riesgos mencionadas en el apartado 2, letra d), darán la debida consideración a los efectos y las posibles interacciones derivados de la aplicación combinada de los requisitos estipulados en el presente capítulo 2. Asimismo, tendrán en cuenta el estado actual de la técnica generalmente reconocido, que, entre otras fuentes, está reflejado en las normas armonizadas o las especificaciones comunes pertinentes.

4. Las medidas de gestión de riesgos mencionadas en el apartado 2, letra d), considerarán aceptables los riesgos residuales asociados a cada peligro, así como el riesgo residual general de los sistemas de IA de alto riesgo, siempre que el sistema de IA de alto riesgo de que se trate se utilice conforme a su finalidad prevista o que se le dé un uso indebido razonablemente previsible. Se informará al usuario de dichos riesgos residuales. A la hora de determinar cuáles son las medidas de gestión de riesgos más adecuadas, se procurará:

> a) eliminar o reducir los riesgos en la medida en que sea posible mediante un diseño y un desarrollo adecuados;
>
> b) implantar, cuando proceda, unas medidas de mitigación y control apropiadas en relación con los riesgos que no puedan eliminarse;
>
> c) proporcionar la información oportuna conforme al artículo 13, en particular en relación con los riesgos mencionados en el apartado 2, letra b), del presente artículo y, cuando proceda, impartir formación a los usuarios. Cuando se eliminen o reduzcan los riesgos asociados a la

utilización del sistema de IA de alto riesgo, se tendrán en debida consideración los conocimientos técnicos, la experiencia, la educación y la formación que se espera que posea el usuario, así como el entorno en el que está previsto que se utilice el sistema.

5. Los sistemas de IA de alto riesgo serán sometidos a pruebas destinadas a determinar cuáles son las medidas de gestión de riesgos más adecuadas. Dichas pruebas comprobarán que los sistemas de IA de alto riesgo funcionan de un modo adecuado para su finalidad prevista y cumplen los requisitos establecidos en el presente capítulo.

6. Los procedimientos de prueba serán adecuados para alcanzar la finalidad prevista del sistema de IA y no excederán de lo necesario para ello.

7. Las pruebas de los sistemas de IA de alto riesgo se realizarán, según proceda, en cualquier momento del proceso de desarrollo y, en todo caso, antes de su introducción en el mercado o puesta en servicio. Los ensayos se realizarán a partir de parámetros y umbrales de probabilidades previamente definidos que sean adecuados para la finalidad prevista del sistema de IA de alto riesgo de que se trate.

(4) OBLIGACIONES DE TRANSPARENCIA PARA DETERMINADOS SISTEMAS DE IA

1. Los proveedores garantizarán que los sistemas de IA destinados a interactuar con personas físicas estén diseñados y desarrollados de forma que dichas personas estén informadas de que están interactuando con un sistema de IA, excepto en las situaciones en las que esto resulte evidente debido a las circunstancias y al contexto de utilización. Esta obligación no se aplicará a los sistemas de IA autorizados por la ley para fines de detección, prevención, investigación o enjuiciamiento de infracciones penales, salvo que estos sistemas estén a disposición del público para denunciar una infracción penal.
2. Los usuarios de un sistema de reconocimiento de emociones o de un sistema de categorización biométrica informarán del funcionamiento del sistema a las personas físicas expuestas a él. Esta obligación no se aplicará a los sistemas de IA utilizados para la categorización biométrica autorizados por la ley para fines de detección, prevención e investigación de infracciones penales.

3. Los usuarios de un sistema de IA que genere o manipule contenido de imagen, sonido o vídeo que se asemeje notablemente a personas, objetos, lugares u otras entidades o sucesos existentes, y que pueda inducir erróneamente a una persona a pensar que son auténticos o verídicos (ultrafalsificación), harán público que el contenido ha sido generado de forma artificial o manipulado. No obstante, el primer párrafo no se aplicará cuando el uso esté legalmente por la ley para fines de detección, prevención, investigación y enjuiciamiento de infracciones penales o resulte necesario para el ejercicio del derecho a la libertad de expresión y el derecho a la libertad de las artes y de las ciencias, garantizados por la Carta de los Derechos Fundamentales de la Unión Europea y supeditados a unas garantías adecuadas para los derechos y libertades de terceros.
4. Los apartados 1, 2 y 3 no afectarán a los requisitos y obligaciones dispuestos en el título III del presente Reglamento.

(5) SANDBOXES. Espacios controlados de pruebas para la IA

1. Los espacios controlados de pruebas para la IA establecidos por las autoridades competentes de uno o varios Estados miembros o por el Supervisor Europeo de Protección de Datos proporcionarán un entorno controlado que facilite el desarrollo, la prueba y la validación de sistemas innovadores de IA durante un período limitado antes de su introducción en el mercado o su puesta en servicio, en virtud de un plan específico. Esto se llevará a cabo bajo la supervisión y la orientación directa de las autoridades competentes con el fin de garantizar el cumplimiento de los requisitos establecidos en el presente Reglamento y, en su caso, en otras legislaciones de la Unión y de los Estados miembros supervisadas en el marco del espacio controlado de pruebas.
2. Los Estados miembros velarán por que, en la medida en que los sistemas innovadores de IA impliquen el tratamiento de datos personales o estén comprendidos dentro del ámbito de supervisión de otras autoridades nacionales o autoridades competentes que proporcionen o respalden el acceso a los datos, las autoridades nacionales de protección de datos y las demás autoridades nacionales estén ligadas al funcionamiento del espacio controlado de pruebas para la IA.
3. Los espacios controlados de pruebas para la IA no afectarán a las facultades de supervisión y correctoras de las autoridades competentes. Cualquier riesgo significativo para la salud, la seguridad y los derechos fundamentales detectado

durante el proceso de desarrollo y prueba de estos sistemas implicará la mitigación inmediata y, en su defecto, la suspensión del proceso de desarrollo y prueba hasta que se produzca dicha mitigación.

4. Los participantes en los espacios controlados de pruebas para la IA responderán de cualquier perjuicio infligido a terceros como resultado de la experimentación realizada en el espacio controlado de pruebas, con arreglo a la legislación aplicable de la Unión y de los Estados miembros en materia de responsabilidad.
5. Las autoridades competentes de los Estados miembros que hayan establecido espacios controlados de pruebas para la IA coordinarán sus actividades y cooperarán en el marco del Comité Europeo de Inteligencia Artificial. Presentarán informes anuales al Comité y a la Comisión sobre los resultados de la aplicación de dicho esquema, que incluirán buenas prácticas, enseñanzas extraídas y recomendaciones acerca de su configuración, y, en su caso, sobre la aplicación del presente Reglamento y otra legislación de la Unión supervisada en el marco del espacio controlado de pruebas.
6. Las modalidades y condiciones de funcionamiento de los espacios controlados de pruebas para la IA, incluidos los criterios de admisibilidad y el procedimiento de solicitud, selección, participación y salida del espacio controlado de pruebas, así como los derechos y obligaciones de los participantes, se determinarán en actos de ejecución. Dichos actos de ejecución se adoptarán de conformidad con el procedimiento de examen a que se refiere el artículo 74, apartado 2.

b. PARLAMENTO EUROPEO - Orientación General adoptada por Consejo con relación al proyecto de Reglamento por el que se establecen normas armonizadas en materia de inteligencia artificial[3]

El estudio de la propuesta de la Comisión en el Consejo lo llevó adelante el grupo TELECOM, a través de diversos talleres y reuniones que iniciaron en el año 2021. Estos debates dieron lugar a cuatro propuestas transaccionales presentadas entre ese año y el año 2022, hasta llegar a la versión definitiva del texto transaccional, que fue remitida al Consejo de Telecomunicaciones con vistas a la adopción de una orientación general.

[3] "Orientación general sobre la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (Reglamento de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión", *Comité de Representantes Permanentes*, 25 de noviembre de 2022, p. 2, en https://data.consilium.europa.eu/doc/document/ST-14954-2022-INIT/es/pdf (consultado el 28/06/2023)

Así fue que el 6 de diciembre del 2022 el Consejo adoptó su orientación general[4] en la que da cuenta del trabajo realizado por el Grupo TELECOM para el estudio de la propuesta de Reglamento, y efectúa una serie de propuestas sobre dicho documento.

En concreto:

- Restringe la definición del artículo 3, apartado 1, a los sistemas desarrollados a través de estrategias de aprendizaje automático y estrategias basadas en la lógica y el conocimiento, para garantizar que la definición de IA proporcione criterios claros para distinguirlos de otros sistemas de software más clásicos.
- Suprime el anexo I y los correspondientes poderes de la Comisión para la adopción de actos delegados para actualizar la definición de sistemas de IA.
- Añade los considerandos 6 bis y 6 ter para aclarar los conceptos de «estrategia de aprendizaje automático» y «estrategias basadas en la lógica y el conocimiento».
- Añade en el artículo 4 la posibilidad de adoptar actos de ejecución para especificar y actualizar las técnicas de las estrategias de aprendizaje automático y de las estrategias basadas en la lógica y el conocimiento, con el objetivo de garantizar que el Reglamento de Inteligencia Artificial sea flexible y pueda adaptarse a las transformaciones futuras.
- En relación a las prácticas prohibidas, el artículo 5 amplía a los agentes privados la prohibición de utilizar la IA con fines de puntuación ciudadana; y en la disposición por la que se prohíbe el uso de sistemas de IA que aprovechan las vulnerabilidades de grupos específicos de personas se incluye a las personas vulnerables por su situación social o económica.
- En relación a la prohibición de uso de sistemas de identificación biométrica remota «en tiempo real» en espacios de acceso público por parte de las autoridades encargadas de la aplicación de la ley, enumera objetivos "...para cuya consecución se considera que dicho uso es estrictamente necesario con fines de aplicación de la ley y qué a las autoridades encargadas de la aplicación de la ley se les debe, por tanto, permitir excepcionalmente el uso de dichos sistemas…".
- Suprime tres supuestos de uso de IA de alto riesgo del listado del Anexo III (la detección de ultra falsificaciones por parte de las autoridades encargadas de la aplicación de la ley, la realización de análisis penales y la comprobación de la autenticidad de los documentos de viaje); agrega dos nuevos (las infraestructuras digitales críticas y los seguros de vida y salud) y matiza otros.

[4] Una orientación general es un tipo de acuerdo político a nivel del Consejo que se adopta a la espera de la posición del Parlamento, que busca facilitar el acuerdo entre ambas instituciones.

- Modifica el artículo 7, apartado 1, previendo la posibilidad no solo de añadir a la lista de supuestos de uso de alto riesgo mediante actos delegados, sino también de suprimirlos. Especifica cuáles son las condiciones en las que puede adoptarse un acto delegado.
- Para la clasificación de la IA como de alto riesgo se incluyen una serie de criterios de clasificación horizontales que se añaden a la clasificación que figura en el anexo III. Con ello buscan que no se incluyan en la clasificación de sistemas de alto riesgo a sistemas de IA que probablemente no provoquen violaciones graves de los derechos fundamentales u otros riesgos significativos.
- Se aclaran y precisan muchos de los requisitos de los sistemas de IA de alto riesgo que se establecen en el título III, capítulo 2 de la propuesta para que sean más viables desde lo técnico.
- Se incluyen varios cambios que añaden claridad a la asignación de responsabilidades y funciones.
- Se incorpora un nuevo título I BIS para que los sistemas de IA pueden utilizarse con muchos fines diferentes (IA de uso general) y para regular las situaciones en las que pueden darse circunstancias en las que la tecnología de IA de uso general se integra en otro sistema que pueda convertirse en un sistema de alto riesgo.
- Se prevé que determinados requisitos de los sistemas de IA de alto riesgo también podrían aplicarse a los sistemas de IA de uso general.
- En cuanto al ámbito de aplicación del Reglamento, se dispone expresamente que se excluyen de su ámbito de aplicación las actividades militares o las relacionadas con la defensa y la seguridad nacional.
- Asimismo, se aclara que el Reglamento no debe aplicarse a los sistemas de IA, incluida su información de salida, que se utilicen únicamente con fines de investigación y desarrollo, ni a las obligaciones de las personas que utilicen IA con fines no profesionales, excepto en lo que respecta a las obligaciones de transparencia.
- Adicionalmente se han pulido algunas de las definiciones del artículo 3 como las de los términos «sistema de identificación biométrica remota» y «sistema de identificación biométrica remota “en tiempo real”».
- Se suman una serie de aclaraciones y simplificaciones de las disposiciones sobre los procedimientos de evaluación de la conformidad y de las disposiciones relativas a la vigilancia del mercado.
- Se ha revisado exhaustivamente el artículo 41 para limitar la discrecionalidad de la Comisión.

- Se han revisado los artículos 56 y 58 con el fin de reforzar el papel del Comité de IA para que pueda prestar mejor apoyo a los Estados miembros en la aplicación y cumplimiento del Reglamento.
- En cuanto a las sanciones por el incumplimiento de las disposiciones del Reglamento de Inteligencia Artificial, el artículo 71 del texto transaccional establece límites más proporcionados al importe de las multas administrativas que pueden imponerse a las pymes y a las empresas emergentes y en el artículo 71, apartado 6, se han añadido cuatro criterios más para decidir el importe de las multas administrativas con el fin de proteger en mayor medida su proporcionalidad global.
- Varios cambios se introducen para reforzar la transparencia en relación con el uso de sistemas de IA de alto riesgo y para que las personas que tengan motivos para considerar que se ha producido una infracción de lo dispuesto en el Reglamento de Inteligencia Artificial puedan presentar una reclamación.
- Modifican sustancialmente las medidas de apoyo a la innovación, con aclaraciones respecto de los espacios controlados (art. 53) y la prueba de sistemas de IA en condiciones reales no supervisadas, siempre que se den determinadas circunstancias y se ofrezcan determinadas garantías (art. 54 bis y 54 ter).

c. PARLAMENTO EUROPEO - Posición negociadora aprobada por el Parlamento Europeo con relación al proyecto de Reglamento por el que se establecen normas armonizadas en materia de inteligencia artificial[5]

En línea con ello, el 14 de junio de 2023, el Parlamento Europeo aprobó su posición negociadora sobre el Reglamento de IA, por 499 votos a favor, 28 en contra y 93 abstenciones.

En sus enmiendas a la propuesta de la Comisión, los eurodiputados pretenden garantizar que los sistemas de IA sean supervisados por personas, sean seguros, transparentes, rastreables, no discriminatorios y respetuosos con el medio ambiente. También quieren tener una definición uniforme de IA diseñada para ser tecnológicamente neutral, de modo que pueda aplicarse a los sistemas de IA de hoy y del mañana.

La adopción de la orientación general por parte del Consejo y la aprobación de la posición del Parlamento permitirán que se inicie el ciclo de diálogos tripartitos en busca de

[5] *"La Eurocámara, lista para negociar la primera ley sobre inteligencia artificial"*, Noticias Parlamento Europeo, 14 de junio de 2023, disponible en: https://www.europarl.europa.eu/news/es/press-room/20230609IPR96212/la-eurocamara-lista-para-negociar-la-primera-ley-sobre-inteligencia-artificial (consultado el 18/06/2023).

alcanzar un acuerdo sobre la propuesta, para que el reglamento finalmente se apruebe y entre en vigor[6].

A continuación, se ofrece una síntesis de la posición negociadora votada por el Parlamento Europeo el 14 de junio:

(1) PRINCIPIOS GENERALES APLICABLES A TODOS LOS SISTEMAS DE IA

a) «Intervención y vigilancia humanas»: los sistemas de IA se desarrollarán y utilizarán como una herramienta al servicio de las personas, que respete la dignidad humana y la autonomía personal, y que funcione de manera que pueda ser controlada y vigilada adecuadamente por seres humanos.

b) «Solidez y seguridad técnicas»: los sistemas de IA se desarrollarán y utilizarán de manera que se minimicen los daños imprevistos e inesperados, así como para que sean sólidos en caso de problemas imprevistos y resistentes a los intentos de modificar el uso o el rendimiento del sistema de IA para permitir una utilización ilícita por parte de terceros malintencionados.

c) «Privacidad y gobernanza de datos»: los sistemas de IA se desarrollarán y utilizarán de conformidad con las normas vigentes en materia de privacidad y protección de datos, y tratarán datos que cumplan normas estrictas en términos de calidad e integridad.

d) «Transparencia»: los sistemas de IA se desarrollarán y utilizarán facilitando una trazabilidad y explicabilidad adecuada, haciendo que las personas sean conscientes de que se comunican o interactúan con un sistema de IA, informando debidamente a los usuarios sobre las capacidades y limitaciones de dicho sistema de IA e informando a las personas afectadas de sus derechos.

e) «Diversidad, no discriminación y equidad»: los sistemas de IA se desarrollarán y utilizarán incluyendo a diversos agentes y promoviendo la igualdad de acceso, la igualdad de género y la diversidad cultural, evitando al mismo tiempo los efectos discriminatorios y los sesgos injustos prohibidos por el Derecho nacional o de la Unión.

[6] Ver en "El Reglamento de Inteligencia Artificial", *Future of Life Institute*, 2023, en https://artificialintelligenceact.eu/about/ [acceso el 18/4/2023].

f) «Bienestar social y medioambiental»: los sistemas de IA se desarrollarán y utilizarán de manera sostenible y respetuosa con el medio ambiente, así como en beneficio de todos los seres humanos, al tiempo que se supervisan y evalúan los efectos a largo plazo en las personas, la sociedad y la democracia.

(2) ALFABETIZACIÓN EN MATERIA DE IA

Se promoverán medidas para el desarrollo de un nivel suficiente de alfabetización en materia de IA, en todos los sectores y teniendo en cuenta las distintas necesidades de los grupos de proveedores, implementadores y personas afectadas de que se trate, también a través de la educación y la formación y de programas de capacitación y de mejora de capacidades, garantizando al mismo tiempo un equilibrio adecuado en materia de género y de edad, con vistas a permitir un control democrático de los sistemas de IA.

Los proveedores e implementadores de sistemas de IA adoptarán medidas para garantizar un nivel suficiente de alfabetización en materia de IA entre su personal y otras personas que se ocupen en su nombre de la operación y el uso de sistemas de IA, teniendo en cuenta sus conocimientos técnicos, su experiencia, su educación y su formación, así como el contexto en que se utilizarán los sistemas de IA y las personas o los grupos de personas con los que se utilizarán.

En particular, dichas medidas de alfabetización consistirán en la enseñanza de nociones y capacidades básicas sobre sistemas de IA y su funcionamiento, incluidos los distintos tipos de productos y usos, sus riesgos y sus beneficios.

Un nivel suficiente de alfabetización en materia de IA es un nivel que contribuye de modo necesario a la capacidad de los proveedores e implementadores para garantizar el cumplimiento y la aplicación del Reglamento.

(3) PRÁCTICAS PROHIBIDAS

1. Sistema de IA que se sirva de técnicas subliminales que trasciendan la conciencia de una persona o de técnicas deliberadamente **manipuladoras o engañosas** con el objetivo o el efecto de alterar de manera sustancial el comportamiento de una persona o un grupo de personas mermando de manera apreciable su capacidad para adoptar una decisión informada y

causando así que la persona tome una decisión que de otro modo no habría tomado, de un modo que provoque o sea probable que provoque perjuicios significativos a esa persona o a otra persona o grupo de personas. La prohibición de los sistemas de IA que se sirvan de las técnicas subliminales a las que se hace referencia en el párrafo primero no se aplicará a los sistemas de IA destinados a ser utilizados para fines terapéuticos autorizados sobre la base de un consentimiento informado específico de las personas expuestas a ellos o, en su caso, de su tutor legal.

2. La introducción en el mercado, la puesta en servicio o la utilización de un sistema de IA que aproveche alguna de las **vulnerabilidades de una persona** o un grupo específico de personas —incluidas las características conocidas o predichas de los rasgos de personalidad o la situación social o económica de esa persona o grupo, la edad y la capacidad física o mental— con el objetivo o el efecto de alterar de manera sustancial el comportamiento de dicha persona o de una persona que pertenezca a dicho grupo de un modo que provoque o sea probable que provoque perjuicios significativos a esa persona o a otra.

   B. bis) La introducción en el mercado, la puesta en servicio o la utilización de sistemas de **categorización biométrica** que clasifiquen a personas físicas con arreglo a atributos o características sensibles o protegidos, o sobre la base de una inferencia de dichos atributos o características. Esta prohibición no se aplicará a los sistemas de IA destinados a ser utilizados para fines terapéuticos autorizados sobre la base de un consentimiento informado específico de las personas expuestas a ellos o, en su caso, de su tutor legal.

3. La introducción en el mercado, la puesta en servicio o la utilización de sistemas de IA con el fin de evaluar o clasificar a las personas físicas o grupos de personas físicas a efectos de su **calificación social** durante un período determinado de tiempo atendiendo a su comportamiento social o a características personales o de su personalidad conocidas, inferidas o predichas, de forma que la puntuación ciudadana resultante provoque una o varias de las situaciones siguientes:

   I) Tratamiento perjudicial o desfavorable de ciertas personas físicas o grupos de éstas en contextos sociales que no están relacionados con los contextos en los que los datos se generaron o recopilaron originalmente;

II) tratos lesivos o desfavorables a determinadas personas físicas o grupos de ellas que sean injustificados o desproporcionados con respecto a su comportamiento social o a su gravedad;

4. El uso de sistemas de **identificación biométrica** remota «en tiempo real» en espacios de acceso público.
5. Sistemas de IA para llevar a cabo **evaluaciones de riesgo** de personas físicas o grupos de personas físicas con el objetivo de determinar el riesgo de que estas personas cometan delitos o infracciones o reincidan en su comisión, o para predecir la comisión o reiteración de un **delito o infracción administrativa** reales o potenciales, mediante la elaboración del perfil de personas físicas o la evaluación de rasgos de personalidad y características, en particular la ubicación de la persona o las conductas delictivas pasadas de personas físicas o grupos de personas físicas.
6. Sistemas de IA que creen o amplíen bases de datos de **reconocimiento facial** mediante la extracción no selectiva de imágenes faciales a partir de internet o de imágenes de circuito cerrado de televisión.
7. Sistemas de IA para inferir las **emociones** de una persona física en los ámbitos de la aplicación de la ley y la **gestión de fronteras**, en lugares de trabajo y en centros educativos.
8. Sistemas de IA para el análisis de imágenes de vídeo grabadas de espacios de acceso público que empleen sistemas de identificación biométrica remota «en diferido», salvo que estén sujetos a una autorización judicial previa de conformidad con el Derecho de la Unión y sean estrictamente necesarios para una búsqueda selectiva destinada a fines de aplicación de la ley y relacionada con un delito grave (según la definición del artículo 83, apartado 1, del TFUE) concreto que ya se haya cometido.

(4) SISTEMA DE ALTO RIESGO

Enumeración de Sistemas de Alto Riesgo (Anexo III del proyecto de Reglamento)

1. Sistemas biométricos y basados en la biometría:
   A) sistemas de IA destinados a utilizarse en la identificación biométrica de personas físicas, con la excepción de los mencionados en el artículo 5;

a bis) sistemas de IA destinados a ser utilizados para extraer conclusiones sobre las características personales de las personas físicas a partir de datos biométricos o basados en la biometría, incluidos los sistemas de reconocimiento de emociones, a excepción de los mencionados en el artículo 5. El punto 1 no incluirá los sistemas de IA destinados a utilizarse con fines de verificación biométrica cuya única finalidad sea confirmar que una persona física concreta es la persona que afirma ser.

2. Gestión y operación de infraestructura crítica:
   A) sistemas de IA destinados a utilizarse como componentes de seguridad en la gestión y funcionamiento del tráfico rodado, ferroviario y aéreo, a menos que estén regulados en la legislación de armonización o en la normativa sectorial;

   a bis) sistemas de IA destinados a utilizarse como componentes de seguridad en la gestión y funcionamiento del suministro de agua, gas, calefacción, electricidad e infraestructuras digitales críticas.

3. Educación y formación profesional:
   A) Sistemas de IA destinados a utilizarse para determinar el acceso o para influir sustancialmente en las decisiones relativas a la admisión o la asignación de personas físicas a los centros de educación y formación profesional;
   B) Sistemas de IA destinados a utilizarse para evaluar a los estudiantes de centros de educación y formación profesional y para evaluar a los participantes en pruebas generalmente necesarias para acceder a estos centros.

   b bis) sistemas de IA destinados a utilizarse con el fin de evaluar el nivel adecuado de educación de una persona y de influir sustancialmente en el nivel de educación o formación profesional que la persona vaya a recibir o al que pueda acceder;

   b ter) sistemas de IA destinados a ser utilizados para hacer seguimiento y detectar comportamientos prohibidos de los estudiantes durante los exámenes celebrados en el contexto o en el seno de instituciones de educación y formación profesional;

4. Empleo, gestión de trabajadores y acceso al autoempleo:
   a) sistemas de IA destinados a utilizarse para la contratación o selección de personas físicas, especialmente para publicar anuncios de empleo específicos, clasificar y filtrar solicitudes o evaluar a candidatos en el transcurso de entrevistas o pruebas;

b) sistemas de IA destinados a utilizarse para tomar decisiones o influir sustancialmente en ellas que afecten a la iniciación, promoción y resolución de relaciones contractuales de índole laboral, a la asignación de tareas basada en la conducta individual o en rasgos o características personales, o al seguimiento y evaluación del rendimiento y la conducta de las personas en el marco de dichas relaciones;

5. Acceso y disfrute de los servicios privados esenciales y de los servicios y prestaciones públicas:

A) Sistemas de IA destinados a ser utilizados por las autoridades públicas o en su nombre para evaluar la admisibilidad de las personas físicas para acceder a prestaciones y servicios de asistencia pública, incluidos los servicios de asistencia sanitaria y servicios esenciales, entre ellos, vivienda, electricidad, calefacción/refrigeración e internet, así como para conceder, reducir, retirar, aumentar o recuperar dichas prestaciones;

B) Sistemas de IA destinados a utilizarse para evaluar la solvencia de personas físicas o establecer su calificación crediticia, salvo los sistemas de IA utilizados al objeto de detectar fraudes financieros;

b bis) sistemas de IA destinados a utilizarse para la toma de decisiones, o para influir sustancialmente en esta, sobre la admisibilidad de las personas físicas para acceder a seguros de salud y de vida;

C) Sistemas de IA destinados a utilizarse para la evaluación y la clasificación de las llamadas de emergencia realizadas por personas físicas o para el envío o el establecimiento de prioridades en el envío de servicios de primera intervención en situaciones de emergencia, por ejemplo policía y fuerzas de seguridad, bomberos y servicios de asistencia médica, y sistemas de triaje de pacientes para la asistencia sanitaria de emergencia.

6. Cumplimiento de la ley:

A) Sistemas de IA destinados a utilizarse por parte de las autoridades encargadas de la aplicación de la ley o en su nombre, o por agencias, órganos u organismos de la Unión en apoyo de las autoridades encargadas de la aplicación de la ley como polígrafos y herramientas similares, siempre que el Derecho de la Unión y el derecho nacional pertinente permitan su uso;

B) Sistemas de IA destinados a utilizarse por parte de las autoridades encargadas de la aplicación de la ley o en su nombre, o por agencias, órganos u organismos de la Unión en apoyo de las autoridades encargadas de la aplicación de la ley para evaluar la fiabilidad de las pruebas durante la investigación o el enjuiciamiento de infracciones penales;

C) Sistemas de IA destinados a utilizarse por parte de las autoridades encargadas de la aplicación de la ley o en su nombre o por agencias, órganos u organismos de la Unión en apoyo de las autoridades encargadas de la aplicación de la ley para la elaboración de perfiles de personas físicas, de conformidad con lo dispuesto en el artículo 3, apartado 4 de la Directiva (UE) 2016/680, durante la detección, la investigación o el enjuiciamiento de infracciones penales o, en el caso de las agencias, los órganos y los organismos de la Unión en el artículo 3, apartado 5 del Reglamento (UE) 2018/1725;

D) Sistemas de IA destinados a utilizarse por parte de las autoridades encargadas de la aplicación de la ley o en su nombre o por agencias, órganos u organismos de la Unión en apoyo de las autoridades encargadas de la aplicación de la ley para llevar a cabo análisis sobre infracciones penales en relación con personas físicas que permitan examinar grandes conjuntos de datos complejos vinculados y no vinculados, disponibles en diferentes fuentes o formatos, para detectar modelos desconocidos o descubrir relaciones ocultas en los datos.

7. Gestión de migraciones, asilo y control de fronteras:

A) Sistemas de IA destinados a utilizarse por parte de las autoridades públicas competentes o en su nombre, o por agencias, órganos u organismos de la Unión como polígrafos y herramientas similares, siempre que el Derecho de la Unión y nacional pertinente permita su uso

B) Sistemas de IA destinados a utilizarse por parte de las autoridades públicas competentes, o en su nombre, o por agencias, órganos u organismos de la Unión para evaluar un riesgo, como un riesgo para la seguridad, la salud o relativo a la inmigración ilegal, que plantee una persona física que pretenda entrar o haya entrado en el territorio de un Estado miembro;

C) Sistemas de IA destinados a utilizarse por parte de las autoridades públicas competentes o en su nombre, o por agencias, órganos u organismos de la Unión para la verificación de la autenticidad de los documentos de viaje y los documentos justificativos de las personas físicas y la detección de documentos falsos mediante la comprobación de sus elementos de seguridad;

D) Sistemas de IA destinados a ser utilizados por las autoridades públicas competentes o en su nombre o por agencias, órganos u organismos de la Unión para ayudar a las autoridades públicas competentes a examinar y evaluar la veracidad de las pruebas en relación con las solicitudes de asilo, visado y permisos de residencia, y las reclamaciones asociadas con respecto a la admisibilidad de las personas físicas solicitantes;

d bis) sistemas de IA destinados a ser utilizados por las autoridades públicas competentes o en su nombre, o por agencias, órganos u organismos de la Unión en la gestión de la inmigración, el asilo y el control fronterizo para supervisar, vigilar o tratar datos en el contexto de las actividades de gestión de fronteras con el fin de detectar, reconocer o identificar a personas físicas;

d ter) sistemas de IA destinados a utilizarse por parte de las autoridades públicas competentes o en su nombre, o por agencias, órganos u organismos de la Unión en la gestión de la migración, el asilo y el control fronterizo para la previsión o predicción de tendencias relacionadas con los movimientos migratorios y los cruces de fronteras.

8. Administración de justicia y procesos democráticos:

A) sistemas de IA destinados a ser utilizados por una autoridad judicial o un órgano administrativo, o en su nombre, para ayudar a una autoridad judicial o un órgano administrativo en la investigación e interpretación de hechos y de la ley, así como en la aplicación de la ley a un conjunto concreto de hechos, o a ser utilizados de forma similar en una resolución alternativa de litigios;

a bis) sistemas de IA destinados a ser utilizados para influir en el resultado de una elección o referéndum o en el comportamiento electoral de personas físicas en el ejercicio de su voto en elecciones o referendos. No incluye los sistemas de IA a cuya información de salida no están directamente expuestas las personas físicas, como las herramientas utilizadas para organizar, optimizar y estructurar campañas políticas desde un punto de vista administrativo y logístico;

a ter) sistemas de IA destinados a ser utilizados por plataformas de redes sociales designadas como plataformas en línea de muy gran tamaño en el sentido del artículo 33 del Reglamento (UE) 2022/2065, en sus sistemas de recomendación para recomendar al destinatario del servicio sobre el contenido generado por los usuarios disponible en la plataforma.

(5) SISTEMA DE GESTIÓN DE RIESGOS

El sistema de gestión de riesgos consistirá en un proceso iterativo continuo que se llevará a cabo durante todo el ciclo de vida de un sistema de IA de alto riesgo, el cual requerirá revisiones y actualizaciones periódicas del proceso de gestión del riesgo, a fin de garantizar su eficacia continua y la documentación de cualquier decisión y medidas significativas adoptadas sujetas al presente artículo. Constará de las siguientes etapas:

a) la identificación, la estimación y la evaluación de los riesgos conocidos y razonablemente previsibles que el sistema de IA de alto riesgo pueda plantear para la salud o la seguridad de las personas físicas, sus derechos fundamentales, incluida la igualdad de acceso y oportunidades, la democracia y el Estado de Derecho o el medio ambiente, cuando el sistema de IA de alto riesgo se utilice de conformidad con su finalidad prevista y en condiciones de uso indebido razonablemente previsible;
b) la evaluación de los riesgos significativos emergentes descritos en la letra a) y detectados a partir del análisis de los datos recogidos con el sistema de seguimiento posterior a la comercialización a la que se refiere el artículo 61;
c) la adopción de medidas apropiadas y específicas de gestión de riesgos diseñadas para abordar los riesgos detectados de conformidad con las letras a) y b) del presente apartado con arreglo a lo dispuesto en los apartados siguientes.

Los sistemas de IA de alto riesgo se diseñarán y desarrollarán de modo que sean vigilados de manera efectiva por personas físicas, lo que incluye dotarlos de una herramienta de interfaz humano-máquina adecuada, entre otras cosas, de forma proporcionada a los riesgos asociados a dichos sistemas. Las personas físicas encargadas de garantizar la vigilancia humana tendrán un nivel suficiente de alfabetización en materia de IA de conformidad con el artículo 4 ter y contarán con el apoyo y la autoridad necesarios para ejercer esa función durante el período en que los sistemas de IA estén en uso y para permitir una investigación exhaustiva tras un incidente.

El objetivo de la vigilancia humana será prevenir o reducir al mínimo los riesgos para la salud, la seguridad, los derechos fundamentales o el medio ambiente que pueden surgir cuando un sistema de IA de alto riesgo se utiliza conforme a su finalidad prevista o cuando se le da un uso indebido razonablemente previsible, en particular cuando dichos riesgos persisten a pesar de aplicar otros requisitos establecidos en el presente capítulo, y cuando las decisiones basadas únicamente en el procesamiento automatizado por parte de

sistemas de IA producen efectos jurídicos o significativos de otro tipo para las personas o grupos de personas con los que se deba utilizar el sistema.

(6) DATOS (y datos sintéticos) y SESGOS

Los conjuntos de datos de entrenamiento y, cuando se utilicen, los conjuntos de datos de validación y prueba, incluidas las etiquetas, deben ser pertinentes, suficientemente representativos, debidamente evaluados en lo que respecta a los errores y ser tan completos como sea posible en vista de la finalidad prevista. Asimismo, tendrán las propiedades estadísticas adecuadas, también en lo que respecta a las personas o los grupos de personas en relación con los que se pretenda utilizar el sistema de IA de alto riesgo, cuando proceda. Los conjuntos de datos reunirán estas características individualmente para cada conjunto de datos o para una combinación de esto.

Los conjuntos de datos tendrán en cuenta, en la medida necesaria en función de su finalidad prevista o los usos indebidos razonablemente previsibles del sistema de IA, las características o elementos particulares del marco geográfico, contextual, conductual o funcional específico en el que se pretende utilizar el sistema de IA de alto riesgo.

En la medida en que sea estrictamente necesario para garantizar la detección y la corrección de los sesgos negativos asociados a los sistemas de IA de alto riesgo, los proveedores de dichos sistemas podrán tratar excepcionalmente las categorías especiales de datos personales que se mencionan en el artículo 9, apartado 1, del Reglamento (UE) 2016/679; el artículo 10 de la Directiva (UE) 2016/680, y el artículo 10, apartado 1, del Reglamento (UE) 2018/1725, ofreciendo siempre las salvaguardias adecuadas para los derechos y las libertades fundamentales de las personas físicas, lo que incluye establecer limitaciones técnicas a la reutilización y la utilización de las medidas de seguridad y protección de la privacidad más recientes.

En particular, se aplicarán todas las condiciones siguientes para que se produzca este tratamiento:

> a) el tratamiento de datos sintéticos o anonimizados no permita alcanzar eficazmente la detección y corrección de sesgos;
>
> b) los datos sean seudonimizados;

c) el proveedor tome las medidas técnicas y organizativas adecuadas para garantizar que los datos tratados a efectos del presente apartado esten asegurados y protegidos, con sujeción a las garantías adecuadas, y que solo las personas autorizadas tengan acceso a dichos datos con las obligaciones de confidencialidad adecuadas;

d) los datos tratados a efectos del presente apartado no serán transmitidos, transferidos ni consultados de otro modo por otras partes;

e) los datos personales tratados a efectos del presente apartado se protejan por medio de las medidas técnicas y organizativas adecuadas y se eliminen una vez se ha corregido el sesgo o cuando los datos personales lleguen al final de su período de conservación;

f) se hayan adoptado medidas eficaces y adecuadas para garantizar la disponibilidad, la seguridad y la resiliencia de los sistemas y servicios de tratamiento frente a incidentes técnicos o físicos;

g) se hayan adoptado medidas eficaces y adecuadas para garantizar la seguridad física de los lugares en los que se almacenan y procesan los datos, la gobernanza y la gestión de sistemas informáticos internos y de sistemas de seguridad informática, la certificación de procesos y productos. Los proveedores que recurran a esta disposición elaborarán documentación que explique por qué el tratamiento de categorías especiales de datos personales era necesario para detectar y corregir sesgos.

(7) PRINCIPALES OBLIGACIONES DE LOS PROVEEDORES E IMPLEMENTADORES DE SISTEMAS DE IA DE ALTO RIESGO Y DE OTRAS PARTES

a) velarán por que sus sistemas de IA de alto riesgo cumplan los requisitos definidos en el capítulo 2 del presente título antes de introducirlos en el mercado o ponerlos en servicio;

a bis) indicarán su nombre, su nombre comercial registrado o marca registrada y su dirección e información de contacto en el sistema de IA de alto riesgo o, cuando no sea posible, en la documentación que lo acompañe, según proceda;

a ter) garantizarán que las personas físicas a las que se asigna la supervisión humana de los sistemas de IA de alto riesgo sean, en concreto, conscientes del riesgo de sesgo de automatización o confirmación;

a quater) proporcionarán especificaciones relativas a los datos de entrada, o a cualquier otra información pertinente en relación con los conjuntos de datos usados, en particular sus limitaciones, teniendo en cuenta la finalidad prevista y los usos indebidos previsibles o razonablemente previsibles del sistema de IA;

b) elaborarán y conservarán la documentación técnica del sistema de IA de alto riesgo
c) cuando estén bajo su control, conservarán los archivos de registro que sus sistemas de IA de alto riesgo generen automáticamente, que son necesarios para garantizar y demostrar el cumplimiento del presente Reglamento, de conformidad con el artículo 20;
d) se asegurarán de que los sistemas de IA de alto riesgo sean sometidos al procedimiento de evaluación de la conformidad oportuno antes de su introducción en el mercado o puesta en servicio, de conformidad con el artículo 43;
e) adoptarán las medidas correctoras necesarias.

(8) EVALUACIÓN DEL IMPACTO EN LOS DERECHOS FUNDAMENTALES PARA LOS SISTEMAS DE IA DE ALTO RIESGO

Antes de poner en uso un sistema de IA de alto riesgo, (con excepción de los sistemas de IA destinados a utilizarse en el ámbito 2 del anexo III), los implementadores llevarán a cabo una evaluación del impacto de los sistemas en el contexto específico de uso.

Esta evaluación abarcará, como mínimo, los siguientes elementos:

A. una descripción clara de la finalidad prevista para la que se utilizará el sistema;
B. una descripción clara del ámbito geográfico y temporal previsto de utilización del sistema;
C. las categorías de personas físicas y grupos que puedan verse afectados por la utilización del sistema;
D. una verificación de que la utilización del sistema es conforme al derecho de la unión y al derecho nacional pertinente en materia de derechos fundamentales;
E. el impacto razonablemente previsible en los derechos fundamentales de poner en uso el sistema de ia de alto riesgo;

F. los riesgos de perjuicio específicos que puedan afectar a personas marginadas o a grupos vulnerables;
G. las repercusiones negativas razonablemente previsibles del uso del sistema en el medio ambiente;
H. un plan detallado sobre cómo se mitigarán los perjuicios y el impacto negativo en los derechos fundamentales;
I. el sistema de gobernanza que pondrá en marcha el implementador, incluida la vigilancia humana, la tramitación de reclamaciones y las vías de recurso.

Si no es posible definir un plan detallado para mitigar los riesgos descritos en el transcurso de la evaluación contemplada en el apartado 1, el implementador se abstendrá de poner en uso el sistema de IA de alto riesgo e informará sin demora indebida al proveedor y a las autoridades nacionales de supervisión pertinentes. Las autoridades nacionales de supervisión, con arreglo a los artículos 65 y 67, tendrán esta información en cuenta al investigar sistemas que presenten un riesgo a nivel nacional.

La obligación descrita con arreglo al apartado 1 se aplicará al primer uso del sistema de IA de alto riesgo. En casos similares, el implementador podrá recurrir a una evaluación de impacto sobre los derechos fundamentales realizada previamente o a una evaluación existente realizada por los proveedores. Si, durante el uso del sistema de IA de alto riesgo, el implementador considera que ya no se cumplen los criterios enumerados en el apartado 1, llevará a cabo una nueva evaluación de impacto en materia de derechos fundamentales.

En el transcurso de la evaluación de impacto, el implementador, con excepción de las pymes, notificará a las autoridades nacionales de supervisión y a las partes interesadas pertinentes e incluirá, en la medida de lo posible, la participación de los representantes de las personas o grupos de personas que probablemente se vean afectadas por el sistema de IA de alto riesgo, según se determina en el apartado 1, incluidos, entre otros, los organismos de igualdad, los organismos de protección de los consumidores, los interlocutores sociales y las agencias de protección de datos, con vistas a recibir su contribución a la evaluación de impacto. El implementador concederá a estos organismos un plazo de seis semanas para responder. Las pymes podrán aplicar voluntariamente las disposiciones establecidas en el presente apartado. En el caso contemplado en el artículo 47, apartado 1, las autoridades públicas podrán quedar exentas de estas obligaciones.

El implementador que sea una autoridad pública o una empresa contemplada en el artículo 51, apartado 1 bis, letra b), publicará un resumen de los resultados de la evaluación de impacto como parte del registro de uso con arreglo a su obligación en virtud del artículo 51, apartado 2.

Cuando el implementador ya deba llevar a cabo una evaluación de impacto relativa a la protección de datos en virtud del artículo 35 del Reglamento (UE) 2016/679 o del artículo 27 de la Directiva (UE) 2016/680, la evaluación de impacto en materia de derechos fundamentales prevista en el apartado 1 se llevará a cabo junto con la evaluación de impacto relativa a la protección de datos. La evaluación de impacto relativa a la protección de datos se publicará como apéndice.

(9) OBLIGACIONES DE TRANSPARENCIA

1. Los proveedores garantizarán que los sistemas de IA destinados a interactuar con personas físicas estén diseñados y desarrollados de forma que el propio proveedor o el usuario informen de manera clara, inteligible y oportuna a dichas personas físicas expuestas a un sistema de IA de que están interactuando con un sistema de IA, excepto en las situaciones en las que esto resulte evidente debido a las circunstancias y al contexto de utilización. Cuando proceda y sea pertinente, esta información también revelará qué funciones se encuentran habilitadas por la IA, si existe vigilancia humana y quién es responsable del proceso de toma de decisiones, así como los derechos y procesos existentes que, de conformidad con el Derecho de la Unión y nacional, permiten a las personas físicas o a sus representantes oponerse a que se les apliquen dichos sistemas y solicitar reparación judicial contra las decisiones adoptadas por los sistemas de IA o los perjuicios causados por ellos, incluido su derecho a solicitar una explicación. Esta obligación no se aplicará a los sistemas de IA autorizados por la ley para fines de detección, prevención, investigación o enjuiciamiento de infracciones penales, salvo que estos sistemas estén a disposición del público para denunciar una infracción penal.
2. Los usuarios de un sistema de reconocimiento de emociones o de un sistema de categorización biométrica que no esté prohibido con arreglo al artículo 5 informarán de manera oportuna, clara e inteligible del funcionamiento del sistema a las personas físicas expuestas a él y obtendrán su consentimiento

antes del tratamiento de sus datos biométricos y otros datos personales de conformidad con el Reglamento (UE) 2016/679, el Reglamento (UE) 2016/1725 y la Directiva (UE) 2016/280, según proceda. Esta obligación no se aplicará a los sistemas de IA utilizados para la categorización biométrica autorizados por la ley para fines de detección, prevención e investigación de infracciones penales.

3. Los usuarios de un sistema de IA que genere o manipule contenido de texto, sonido o visual que pueda inducir erróneamente a pensar que tal contenido es auténtico o verídico y que presente representaciones de personas que parecen decir o hacer cosas que no han dicho ni hecho, sin su consentimiento (ultrafalsificación), harán público de manera adecuada, oportuna, clara y visible que el contenido ha sido generado de forma artificial o manipulado, así como, cuando sea posible, el nombre de la persona física o jurídica que lo generó o manipuló. Por hacer público se entenderá etiquetar el contenido de un modo que informe que el contenido no es auténtico y que resulte claramente visible para su destinatario. Para etiquetar el contenido, los usuarios tendrán en cuenta el estado de la técnica generalmente reconocido y las normas y especificaciones armonizadas pertinentes.

> *3 bis.* El apartado 3 no se aplicará cuando el uso de un sistema de IA que genere o manipule contenido de texto, sonido o visual esté legalmente autorizado por las infracciones penales o resulte necesario para el ejercicio del derecho a la libertad de expresión y el derecho a la libertad de las artes y de las ciencias, garantizados por la Carta de los Derechos Fundamentales de la Unión Europea y supeditados a unas garantías adecuadas para los derechos y libertades de terceros. Cuando el contenido forme parte de una obra cinematográfica claramente creativa, satírica, artística o ficticia, de imágenes de videojuegos y de obras o formatos análogos, las obligaciones de transparencia establecidas en el apartado 3 se limitarán a revelar la existencia de tales contenidos generados o manipulados de una manera adecuada, clara y visible, que no obstaculice la presentación de la obra, y a revelar los derechos de autor aplicables, cuando proceda. Asimismo, no impedirá que las autoridades encargadas de la aplicación de la ley utilicen sistemas de IA destinados a detectar ultrafalsificaciones y a prevenir, investigar y enjuiciar las infracciones penales relacionadas con su uso.

(10) SANDBOX DE IA

Los Estados miembros establecerán al menos un espacio controlado de pruebas para la IA a escala nacional, que estará operativo a más tardar el día de la entrada en vigor del Reglamento. Este espacio controlado de pruebas también podrá establecerse conjuntamente con uno o varios otros Estados miembros. También podrán establecerse otros espacios controlados de pruebas para la IA a escala regional o local o conjuntamente con otros Estados miembros;

- Las autoridades que creen dichos espacios asignarán recursos suficientes para cumplir lo dispuesto en el presente artículo de manera efectiva y oportuna;

- Los espacios controlados de pruebas para la IA, de conformidad con los criterios establecidos en el artículo 53 bis, proporcionarán un entorno controlado que fomente la innovación y facilite el desarrollo, la prueba y la validación de sistemas innovadores de IA durante un período limitado antes de su introducción en el mercado o su puesta en servicio, con arreglo a un plan específico acordado entre los proveedores potenciales y las autoridades creadoras de tales espacios;

El establecimiento de espacios controlados de pruebas para la IA tendrá el fin de contribuir a los siguientes objetivos:

A. que las autoridades competentes proporcionen orientaciones a los proveedores potenciales de sistemas de IA para lograr el cumplimiento reglamentario del presente Reglamento o, en su caso, de otra legislación aplicable de la Unión y de los Estados miembros;
B. que los proveedores potenciales permitan y faciliten la prueba y el desarrollo de soluciones innovadoras relacionadas con los sistemas de IA; c) el aprendizaje reglamentario en un entorno controlado.

Las autoridades que creen espacios controlados de pruebas proporcionarán orientación y supervisión en el marco del espacio controlado de pruebas con vistas a detectar los riesgos, en particular para los derechos fundamentales, la democracia y el Estado de Derecho, la salud y seguridad, y el medio ambiente, probar y demostrar las medidas de mitigación para los riesgos detectados, así como su eficacia, y garantizar el cumplimiento de los requisitos del presente Reglamento y, cuando proceda, de otra legislación de la Unión y de los Estados miembros;

(11) DERECHO A EXPLICACIÓN DE LA TOMA INDIVIDUAL DE DECISIONES

1. Toda persona afectada sujeta a una decisión adoptada por el implementador sobre la base de la información de salida de un sistema de IA de alto riesgo que produzca efectos jurídicos o que le afectan significativamente de una manera que considera que perjudica a su salud, seguridad, derechos fundamentales, bienestar socioeconómico o cualquier otro de sus derechos derivados de las obligaciones establecidas en el presente Reglamento, tendrá derecho a solicitar al implementador una explicación clara y significativa, de conformidad con el artículo 13, apartado 1, sobre el papel del sistema de IA en el procedimiento de toma de decisiones, los principales parámetros de la decisión adoptada y los datos de entrada correspondientes.

2. El apartado 1 no se aplicará a la utilización de sistemas de IA para los que el Derecho nacional o de la Unión prevea excepciones o restricciones a la obligación prevista en el apartado 1 en la medida en que dichas excepciones o restricciones respeten la esencia de los derechos y libertades fundamentales y sean una medida necesaria y proporcionada en una sociedad democrática.

3. El presente artículo se aplicará sin perjuicio de lo previsto en los artículos 13, 14, 15 y 22 del Reglamento (UE) 2016/679.

(12) IA GENERATIVA

Los proveedores de modelos básicos utilizados en sistemas de IA destinados específicamente a generar, con distintos niveles de autonomía, contenido como texto, imágenes, audio o video complejos ("IA generativa") y proveedores que especializan un modelo básico en un sistema de IA generativa, deberán (además):

1. Cumplir con las obligaciones de transparencia (ver art. 52 1);
2. Capacitar y, en su caso, diseñar y desarrollar el modelo de base de manera que se garanticen las garantías adecuadas contra la generación de contenido que infrinja el Derecho de la Unión, de conformidad con el estado de la técnica generalmente reconocido, y sin perjuicio de los derechos fundamentales, incluidos la libertad de expresión;

3. Documentar y poner a disposición del público un resumen suficientemente detallado del uso de los datos de formación protegidos por la ley de derechos de autor.

2. OCDE[7]

El 22 de mayo de 2019, treinta y seis países que integran la OCDE —Organización para la Cooperación y el Desarrollo Económicos—, entre los que se encuentra la República Argentina, suscribieron el primer conjunto de directrices de políticas intergubernamentales sobre IA, con el objetivo de promover desarrollos robustos, seguros, imparciales y fiables, reconociendo la vigencia de cinco principios complementarios en orden a la promoción de una IA fiable.

(1) PRINCIPIOS

. Crecimiento inclusivo, desarrollo sostenible y bienestar;

. Valores centrados en el ser humano y equidad;

. Transparencia y explicabilidad;

. Robustez, seguridad;

. Protección;

. Rendición de cuentas.

(2) RECOMENDACIONES A LOS RESPONSABLES DE FORMULACIÓN DE POLÍTICAS PÚBLICAS

. Invertir en investigación y desarrollo de IA;

. Fomentar un ecosistema digital para la IA;

. Dar forma a un entorno político propicio para la IA;

. Desarrollar la capacidad humana y prepararse para la transformación del mercado laboral;

. Cooperación internacional para una IA confiable.

[7] OCDE, *"Recomendación del Consejo sobre inteligencia artificial"*, 21 de mayo de 2019, disponible en: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449 (consultado el 28/06/023)

3. PROYECTO DE LEY DE CHILE SOBRE ROBÓTICA, INTELIGENCIA ARTIFICIAL Y TECNOLOGÍAS CONEXAS[8]

En abril de 2023, se presentó en el Congreso de Chile un proyecto de ley que busca regular los sistemas de IA. Esta iniciativa legislativa realiza una clasificación de los sistemas de IA basada en el riesgo, dividiéndolos en sistemas de riesgo inaceptable y sistemas de alto riesgo, de manera similar al enfoque tomado por la Propuesta de Regulación de la Inteligencia Artificial del Parlamento Europeo.

(1) OBJETIVOS PRINCIPALES

a. Lograr una legislación unificada y coherente que sistematice los procesos de desarrollo, distribución, comercialización y utilización de los sistemas de IA, para brindar certeza y seguridad jurídica;

b. Regular la responsabilidad civil y los derechos de propiedad intelectual relacionados con los sistemas de IA;

c. Regular la utilización de los sistemas de IA en ámbitos como el penal, educativo, cultural y audiovisual;

d. Proteger a los consumidores en general, y en particular, regular el tratamiento de datos personales en el contexto de los sistemas de IA;

e. Evitar la discriminación y abordar la discriminación algorítmica.

(2) SISTEMAS DE IA DE RIESGOS INACEPTABLE

1. Sistemas de IA capaces de alterar de manera sustancial el comportamiento de una persona de modo que provoque o sea probable que provoque perjuicios físicos o psicológicos;
2. Sistema de IA que aprovechen vulnerabilidades de una persona, debido a su edad o discapacidad física o mental, para alterar su comportamiento de modo que provoque o pueda provocar perjuicios físicos o psicológicos;

[8] Cámara de Diputadas y Diputados de Chile, *"Proyecto de Ley: Regula los sistemas de inteligencia artificial, la robótica y las tecnologías conexas, en sus distintos ámbitos de aplicación"*, 24 de abril de 2022, disponible en: https://www.camara.cl/legislacion/proyectosdeley/tramitacion.aspx?prmID=16416&prmBOLETIN=15869-19

3. Sistema de IA utilizado por las autoridades para evaluar o clasificar la fiabilidad de personas según su conducta, características personales o de su personalidad, que pueda provocar un trato perjudicial o desfavorable que no guarde relación con el contexto o que sea desproporcionado;
4. Sistema de identificación biométrica remota, en tiempo real o diferido, en espacios de acceso público, salvo para la búsqueda de posibles víctimas de un delito; la prevención de una amenaza para la vida o un atentado terrorista; la localización, identificación o enjuiciamiento de una persona que cometió o se sospecha que cometió un delito.

(3) SISTEMA DE IA DE ALTO RIESGO

1. La identificación biométrica remota en tiempo real o diferido de personas en espacios privados;
2. La utilización en gestión del suministro de agua, electricidad y gas;
3. La asignación y determinación del acceso a establecimientos educacionales y la evaluación de estudiantes;
4. La selección y contratación de personas en trabajos;
5. La asignación de tareas y el seguimiento y evaluación del rendimiento y la conducta de los trabajadores;
6. La evaluación de las personas para acceder a prestaciones y servicios de asistencia pública;
7. La evaluación de la solvencia de personas o su calificación crediticia;
8. La utilización en situaciones de emergencia y desastre (ejemplo, bomberos o ambulancias);
9. La utilización para determinar el riesgo de personas que cometen infracciones penales o reinciden, así como el riesgo para las potenciales víctimas de delitos;
10. La utilización en cualquier etapa de investigación e interpretación de hechos que pudieran constituir un delito;
11. La utilización para gestión de la migración, el asilo y el control fronterizo;
12. Sistemas de IA que puedan causar un perjuicio para la salud y la seguridad o para los derechos fundamentales.

(4) REQUISITOS PREVIOS A LA AUTORIZACIÓN DE SISTEMAS DE ALTO RIESGO

1. Contar con un plan de gestión de riesgos que permita eliminar o reducir riesgos e implementar medidas de mitigación y control;
2. Contar con un plan de gestión de datos de entrada que incluya la evaluación de sesgos y la detección de lagunas o deficiencias en los datos, así como la forma de subsanarlas;
3. Contar con mecanismos que aseguren que los datos de entrada ingresados sean pertinentes para la finalidad prevista;
4. Contar con un sistema de registro automático de eventos, de gestión de calidad e instrucciones de uso, medidas que permitan su vigilancia efectiva por personas;
5. Demostrar un nivel adecuado de precisión, solidez y ciberseguridad;
6. Demostrar ser resistente a los intentos de terceros de alterar su uso o funcionamiento aprovechando las vulnerabilidades del sistema.

4. CALIFORNIA: PROYECTO DE LEY AB331 SOBRE HERRAMIENTAS DE DECISIÓN AUTOMATIZADA[9]

Este proyecto de ley busca incorporar un capítulo relacionado con la inteligencia artificial al Código de Negocios Profesiones vigente y regula los siguientes aspectos.

(1) DECISIONES DE CONSECUENCIAS SIGNIFICATIVAS

Se define como una decisión o juicio que tiene un efecto legal, material o significativo en la vida de una persona:

(1) Empleo, gestión de trabajadores o trabajo por cuenta propia
(2) Educación y formación profesional
(3) Vivienda o alojamiento
(4) Servicios públicos esenciales, incluyendo electricidad, calefacción, agua, acceso a internet o telecomunicaciones, o transporte
(5) Planificación familiar, incluyendo servicios de adopción, servicios reproductivos o servicios de protección infantil
(6) Cuidado de la salud o seguro médico

[9] Californian Legislature, *"Automated decision tools"*, 18 de mayo de 2023, disponible en: https://legiscan.com/CA/text/AB331/2023#:~:text=This%20 bill%20would%20 define%20%E2%80%9c Deployer%2ccosts%20mandated%20by%20the%20state

(7) Servicios financieros

(8) El sistema de justicia penal, incluyendo:

(A) Evaluaciones de riesgo para audiencias previas al juicio

(B) Sentencias

(C) Libertad condicional.

(9) Servicios legales

(10) Votación

(11) Acceso a beneficios o servicios o asignación de sanciones

(2) EVALUACIÓN DE IMPACTO

Tanto implementadores como desarrolladores deberán llevar a cabo evaluaciones de impacto para minimizar los riesgos. Estas deberán comprender:

(1) Propósito de la herramienta de decisión automatizada y sus beneficios, usos y contextos de implementación previstos

(2) Descripción de los resultados y cómo se utilizan para tomar, o ser un factor determinante en la toma de decisiones

(3) Un resumen del tipo de datos personales recopilados y procesados

(4) Un análisis de posibles impactos adversos en función del sexo, raza, color, etnia, religión, edad, origen nacional, dominio limitado del inglés, discapacidad, condición de veterano o información genética debido al uso de la herramienta

(5) Una declaración de hasta qué punto el uso de la herramienta es coherente o difiere de la declaración de usos previstos

Además de ello, el desarrollador deberá contemplar:

(1) Descripción de las medidas tomadas para mitigar el riesgo conocido

(2) Una descripción de cómo la herramienta puede ser utilizada por una persona o cuando se utiliza para tomar, o ser un factor determinante en la toma de decisiones de consecuencias significativas

5. REPÚBLICA POPULAR CHINA[10]

a.Medidas Administrativas para los Servicios de Inteligencia Artificial Generativa (Borrador para Comentarios)

La inteligencia artificial generativa deberá cumplir con los requisitos de las leyes y reglamentos, respetar la moral social, el orden público y las buenas costumbres, y cumplir con los siguientes requisitos:

(1) CUESTIONES POLÍTICAS

El contenido debe reflejar los valores fundamentales del socialismo y no debe contener la subversión del poder estatal, el derrocamiento del sistema socialista, la incitación a dividir el país, socavar la unidad nacional, promover el terrorismo, el extremismo, y promover el odio étnico y la discriminación étnica, la violencia, la información obscena y pornográfica, la información falsa y el contenido que pueda perturbar el orden económico y social.

(2) DISCRIMINACION

En el proceso de diseño de algoritmos, selección de datos de entrenamiento, generación y optimización de modelos y prestación de servicios, se deben tomar medidas para prevenir la discriminación basada en raza, etnia, creencia, país, región, género, edad, ocupación, etc.

(3) ÉTICA COMERCIAL

Respetar los derechos de propiedad intelectual y la ética comercial, y no utilizar ventajas tales como algoritmos, datos y plataformas para implementar la competencia desleal.

(4) CONTENIDO VERDADERO

El contenido debe ser exacto, y se deben tomar medidas para evitar la generación de información falsa.

(5) RESPETAR DERECHOS Y PREVENIR DAÑOS

Respetar los intereses legítimos de los demás, prevenir daños a la salud física y mental de los demás, dañar los derechos de imagen, los derechos de reputación y la privacidad personal, e infringir los derechos de propiedad intelectual. Se prohíbe la adquisición, divulgación y uso ilegales de información personal, privacidad y secretos comerciales.

---

[10] Administración de Ciberespacio de China, *"Medidas Administrativas para los Servicios de Inteligencia Artificial Generativa (Borrador para Comentarios)"*, 11 de abril de 2023, disponible en: http://www.cac.gov.cn/2023-04/11/c_1682854275475410.htm

(6) RESPONSABILIDAD

Organizaciones e individuos deben asumir la responsabilidad del productor del contenido generado por el producto. Si se trata de información personal, asumir la responsabilidad legal del procesador de información personal y cumplir con la obligación de proteger la información personal.

(7) SEGURIDAD

Antes de utilizar productos de inteligencia artificial generativa, se deberá presentar una evaluación de seguridad para llevar a cabo los procedimientos de presentación, modificación y presentación de cancelación de algoritmos.

(8) DATOS DE ENTRENAMIENTO

Los datos de entrenamiento de la inteligencia artificial generativa deben cumplir los siguientes requisitos:

- No contener contenido que infrinja los derechos de propiedad intelectual;
- Si los datos contienen información personal, deberá obtener el consentimiento o satisfacer otras circunstancias estipuladas por las leyes y reglamentos administrativos;
- Poder garantizar la autenticidad, exactitud, objetividad y diversidad de los datos.

(9) ETIQUETADO MANUAL

El proveedor deberá formular reglas de etiquetado claras, específicas y operables. Brindar la capacitación necesaria para el personal de etiquetado y realizar una verificación aleatoria de la corrección del contenido etiquetado.

(10) PROTECCIÓN DE DATOS

- Se requerirá que los usuarios proporcionen información de identidad real de conformidad con las disposiciones;
- Los proveedores deben aclarar y divulgar los grupos, ocasiones y usos aplicables de sus servicios, y tomar las medidas para evitar que los usuarios confíen demasiado en el contenido generado o se entreguen a él;
- El prestador asume la obligación de proteger la información ingresada por el usuario y los registros de uso. No debe retener ilegalmente información que

pueda inferir la identidad del usuario y no debe proporcionar la información ingresada por el usuario a otros.

(11) USUARIOS Y CONSUMIDORES

El proveedor deberá:

- Establecer un mecanismo para gestionar las quejas de los usuarios relativas a su privacidad personal o secretos comerciales;
- Proporcionar la información que pueda afectar la confianza y la elección de los usuarios, incluidas descripciones como la fuente, la escala, el tipo y la calidad del pre-entrenamiento y datos de entrenamiento optimizados, reglas de etiquetado manual, la escala y el tipo de datos etiquetados manualmente, algoritmos básicos y sistemas técnicos, etc.;
- Guiar a los usuarios a comprender científicamente y utilizar racionalmente el contenido generado por inteligencia artificial generativa, a no utilizar el contenido generado para dañar la imagen, la reputación y otros derechos e intereses legítimos de otros, y a no participar en exageraciones comerciales o comercialización inadecuada.

b. Actualización: Medidas Provisionales para la Gestión de Servicios de Inteligencia Artificial Generativa[11]

Las "Medidas provisionales para la gestión de los servicios de inteligencia artificial generativa" han sido revisadas y aprobadas en la 12ª reunión de la Oficina Estatal de Información de Internet el 23 de mayo de 2023. A su vez, han sido aprobadas por la Comisión Nacional de Desarrollo y Reforma, el Ministerio de Educación, el Ministerio de Ciencia y Tecnología, el Ministerio de Informática, el Ministerio de Seguridad Pública y la Administración Estatal de Radio y Televisión y entrarán en vigencia el 15 de agosto de 2023.

(1) CUESTIONES POLÍTICAS

El Estado se adhiere a los principios de otorgar igual importancia al desarrollo y la seguridad, promover la innovación y combinar la gobernanza basada en el derecho, adoptar medidas efectivas para fomentar la innovación y el desarrollo de la inteligencia

[11] Administración de Ciberespacio de China, *"Medidas Provisionales para la Gestión de Servicios de Inteligencia Artificial Generativa"*, 13 de julio de 2023, disponible en: http://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm?_x_tr_sl=zh-CN&_x_tr_tl=en&_x_tr_hl=en&_x_tr_pto=sc&_x_tr_sch=http

artificial generativa, e implementar la prudencia inclusiva y la supervisión clasificada y jerárquica de servicios de inteligencia artificial generativa.

El contenido debe respetar la moral y la ética social y adherirse a los valores fundamentales del socialismo.

(2) DISCRIMINACIÓN

En el proceso de diseño de algoritmos, selección de datos de entrenamiento, generación y optimización de modelos y provisión de servicios, se deberán tomar medidas efectivas para prevenir la discriminación basada en etnicidad, creencia, país, región, género, edad, ocupación, salud, etc.

(3) ÉTICA COMERCIAL

Se deben respetar los derechos de propiedad intelectual, la ética comercial, guardar secretos comerciales y no utilizar algoritmos, datos, plataformas y otras ventajas para implementar monopolios y competencia desleal.

(4) TRANSPARENCIA

Con base en las características del tipo de servicio, se deberán tomar medidas efectivas para mejorar la transparencia de los servicios de inteligencia artificial generativa y mejorar la precisión y confiabilidad del contenido generado.

(5) RESPETAR DERECHOS Y PREVENIR DAÑOS

Se deben respetar los derechos e intereses legítimos de los demás. No se debe poner en peligro la salud física y mental de los demás, ni se debe infringir los derechos de imagen, derechos de reputación, derechos de honor, derechos de privacidad y derechos de información personal de otros.

(6) ETIQUETADO DE DATOS

Cuando el etiquetado de datos se lleve a cabo durante la investigación y el desarrollo de tecnología de inteligencia artificial generativa, el proveedor deberá formular reglas de etiquetado claras, específicas y operables que cumplan con los requisitos de estas medidas. Además, se deberá realizar una evaluación de la calidad del etiquetado de datos; se deberá llevar adelante una verificación de la exactitud del contenido del etiquetado; realizar la capacitación necesaria para el personal de etiquetado y guiarlo para que lleve a cabo el trabajo de etiquetado de manera estandarizada.

(7) DESARROLLO TECNOLÓGICO Y GOBERNANZA

Se debe fomentar la aplicación innovadora de tecnología de inteligencia artificial generativa en diversas industrias y campos, generar contenido positivo, saludable y de alta calidad, explorar y optimizar escenarios de aplicación y construir un ecosistema de aplicación.

Además, se debe apoyar a las organizaciones de la industria, las empresas, las instituciones educativas y de investigación científica, las instituciones culturales públicas y las instituciones profesionales relevantes para colaborar en la innovación de tecnología de inteligencia artificial generativa, la construcción, transformación y aplicación de recursos de datos y la prevención de riesgos.

(8) OBLIGACIONES DE LOS PROVEEDORES

. Asumirán legalmente la responsabilidad de los productores de contenido de información de la red y cumplirán con las obligaciones de seguridad de la información de la red. Si se trata de datos personales, la responsabilidad del procesador de datos personales se asumirá de conformidad con la ley y se cumplirá con la obligación de proteger los datos personales.

. Deberán suscribir contratos de servicios con los usuarios de servicios de inteligencia artificial generativa que registren sus servicios, aclarando los derechos y obligaciones de ambas partes.

. Aclararán y divulgarán la población aplicable, las ocasiones y los usos de sus servicios, orientarán a los usuarios para que comprendan y utilicen científica y racionalmente la tecnología de inteligencia artificial generativa de conformidad con la ley,

. Tomarán medidas efectivas para evitar que los usuarios menores de edad confíen excesivamente en la inteligencia artificial generativa

. Cumplirán con las obligaciones de protección de la información de los usuarios y utilizarán los registros de acuerdo con la ley; no recopilarán información personal no esencial; no retendrán ilegalmente la información de entrada ni utilizarán registros que puedan identificar a los usuarios.

. Aceptará y procesará las solicitudes de la persona para revisar, copiar, corregir, complementar y eliminar su información personal de manera oportuna de conformidad con la ley.

. Marcarán las imágenes, los videos y otros contenidos generados de conformidad con el "Reglamento sobre la administración de síntesis profunda de los servicios de información de Internet".

. Cuando un proveedor descubra contenido ilegal, deberá tomar medidas de eliminación de inmediato, cómo detener la generación, detener la transmisión y eliminación, adoptar medidas como la optimización del modelo para la rectificación e informar a la autoridad competente pertinente.

. Si el proveedor descubre que el usuario está utilizando el servicio de inteligencia artificial generativa para realizar actividades ilícitas, tomará las advertencias, restringirá las funciones, suspenderá o dará por terminada la prestación de los servicios y demás medidas de eliminación conforme a la ley, llevará los registros correspondiente e informará a la autoridad competente correspondiente.

. Deberá establecer y mejorar el mecanismo de denuncias, establecer un portal conveniente de denuncias, anunciar el proceso de tramitación y el límite de tiempo de respuesta, aceptar y tramitar las denuncias de manera oportuna y dar retroalimentación sobre la tramitación.

(9) USUARIOS

Si los usuarios encuentran que los servicios de inteligencia artificial generativa no cumplen con las leyes, los reglamentos administrativos y las disposiciones de estas medidas, tienen derecho a reclamar y denunciar ante las autoridades competentes correspondientes.

(10) LICENCIAS

Cuando las leyes y reglamentos administrativos establezcan que la prestación de servicios de inteligencia artificial generativa los proveedores deberán obtener las licencias administrativas correspondientes de conformidad con la ley.

6. BRASIL

6.1 Proyecto de Ley 2338/2023[12]

Este proyecto fue presentado en el año 2023 y es producto del trabajo de una Comisión de Juristas presidida por el Sr. Ricardo Villas Bôas Cueva, creada por la Ley del Presidente del Senado N° 4 de 2022, destinada a subsidiar la elaboración de un proyecto sustitutivo para instruir la apreciación de los proyectos de ley N° 5.051, de 2019; N° 21, de 2020, y N°

[12]Cámara de Diputados de Brasil, "Proyecto de ley 2338/23" https://www25.senado.leg.br/web/atividade/materias/-/materia/157233

872, de 2021, que tienen como objetivo establecer principios, reglas, directrices y bases para regular el desarrollo y la aplicación de la inteligencia artificial en Brasil.

(1) DESARROLLO E IMPLEMENTACIÓN BASADA EN:

I. Centralidad de la persona humana;
II. Respeto a los derechos humanos y de los valores democráticos;
III. Libre desarrollo de la personalidad;
IV. Protección del medio ambiente y desarrollo sustentable;
V. Igualdad, no discriminación, pluralidad y respeto de los derechos laborales;
VI. Privacidad, protección de datos y autodeterminación en beneficio de la persona humana, del régimen democrático y del derechos laborales;
VII. Innovación y desarrollo tecnológico;
VIII. Libre iniciativa, libre competencia y protección al consumidor;
IX. Privacidad, protección de datos y autodeterminación informativa;
X. Promoción de la investigación y el desarrollo con el objetivo de estimular la innovación en los sectores productivos y en el poder público;
XI. Acceso a la información y educación, así como a la concientización sobre sistemas de inteligencia artificial y sus aplicaciones.

(2) PRINCIPIOS

I. Buena fe;
II. Crecimiento inclusivo, desarrollo sostenible y bienestar;
III. Libre determinación y libertad de decisión y elección;
IV. Participación humana en el ciclo de la inteligencia artificial y supervisión humana efectiva;
V. No discriminación;
VI. Justicia, equidad e inclusión;
VII. Transparencia, explicabilidad, inteligibilidad y auditabilidad;
VIII. Confiabilidad y robustez de los sistemas de inteligencia artificial;
IX. Debido proceso legal, impugnable y contradictorio;
X. Trazabilidad en las decisiones durante el ciclo de vida de los sistemas de inteligencia artificial como medio de rendición de cuentas y atribución de responsabilidades a una persona física o jurídica;
XI. Rendición de cuentas y reparación integral del daño;

XII. Prevención, prevención y mitigación de riesgos sistémicos derivados de usos previstos y efectos imprevistos de sistemas de inteligencia artificial;

XIII. No maleficencia y proporcionalidad entre métodos empleados y los fines declarados y legítimos de la inteligencia artificial.

(3) DERECHOS (DE LAS PERSONAS AFECTADAS POR SISTEMAS DE IA)

I. Derecho a la información previa sobre sus interacciones con sistemas de inteligencia artificial;

II. Derecho a una explicación sobre la decisión, recomendación o previsión asumidos por los sistemas de inteligencia artificial;

III. Derecho a impugnar decisiones o predicciones de IA que produzcan efectos jurídicos que impacten significativamente en los intereses de los afectados;

IV. Derecho a la determinación y participación humana en las decisiones de sistemas de IA, teniendo en cuenta en contexto y el estado del arte del desarrollo tecnológico;

V. Derecho a la no discriminación y corrección de prejuicios, discriminacion directa, indirecta, ilegal o abusiva;

VI. Derecho a la privacidad y protección de datos personales.

(4) CATEGORIZACIÓN DEL RIESGO

Revisión preliminar: Antes de su comercialización o uso en servicio, cada sistema de IA se someterá a una evaluación preliminar realizada por el proveedor, para clasificar su grado de riesgo (..)

4.1 RIESGO EXCESIVO. PROHIBICIONES

I. Uso de técnicas subliminales que tienen como objetivo (o producen efecto) inducir a una persona física a comportarse de forma dañina o peligrosa para su salud o seguridad(..);

II. Sistemas de IA que exploten las vulnerabilidades de grupos específicos de personas físicas, tales como las asociadas a su edad o discapacidad física o psíquicos, con el fin de inducirlos a comportarse de una manera nociva para su salud o la seguridad (..);

III. Uso de sistemas de IA por el gobierno, para evaluar, clasificar o jerarquizar las personas naturales, con base en su comportamiento social o atributos de su

personalidad, a través de la puntuación universal, para el acceso a los bienes y servicios y políticas públicas, de forma ilegítima o desproporcionada.

4.2 ALTO RIESGO

I. Aplicación de dispositivos de seguridad en la gestión y funcionamiento de la infraestructura crítica, como el control del tráfico y redes de suministro de agua y electricidad;

II. Educación y formación profesional, incluidos los sistemas de determinación del acceso a instituciones educativas y de formación profesional o para la evaluación y seguimiento de los alumnos;

III. Reclutamiento, preselección, filtrado, evaluación de candidatos, toma de decisiones sobre promociones o terminaciones de relaciones contractuales de trabajo, división de tareas y control y evaluación del desempeño y comportamiento de las personas afectadas por dichas aplicaciones de inteligencia artificial en los ámbitos del empleo, la gestión de los trabajadores y el acceso al empleo por cuenta propia;

IV. Evaluación de los criterios de acceso, elegibilidad, concesión, revisión, reducción o revocación de los servicios públicos y privados que son considerados esenciales, incluidos los sistemas utilizados para evaluar la elegibilidad de las personas físicas para la prestación de los servicios públicos de asistencia y seguridad;

V. Evaluación de la capacidad de endeudamiento de las personas físicas o el establecimiento de su calificación crediticia;

VI. Despacho o establecimiento de prioridades para servicios de respuesta de emergencia, incluida la asistencia médica y de incendios;

VII. Administración de justicia, incluidos los sistemas que ayudan a las autoridades judiciales en la investigación de los hechos y la aplicación de la ley;

VIII. Vehículos autónomos, cuando su uso pueda implicar riesgos para la integridad física de las personas;

IX. Aplicaciones en el área de la salud, incluidas las destinadas a auxiliar diagnósticos y procedimientos médicos;

X. Sistemas de identificación biométrica;

XI. Investigación criminal y seguridad pública, especialmente para evaluaciones de riesgos individuales por parte de las autoridades competentes a fin de determinar el riesgo de una persona de delinquir o reincidir, o el riesgo para las posibles víctimas de delitos penales o para evaluar las características de

personalidad y características o comportamiento delictivo pasado de personas naturales o grupos;

XII. Estudio analítico de los delitos relacionados con las personas naturales, permitir a las autoridades encargadas de hacer cumplir la ley la búsqueda de grandes conjuntos de datos complejos, relacionados o no relacionados, disponibles en diferentes fuentes de datos o en diferentes formatos de datos, con el fin de identificar patrones desconocidos o descubrimiento de relaciones ocultas en los datos;

XIII. Investigación de las autoridades administrativas para evaluar la credibilidad de la evidencia en el curso de la investigación o enjuiciamiento de delitos, para anticipar la ocurrencia o reincidencia de un infracción real o potencial basada en la elaboración de perfiles de personas;

XIV. Gestión migratoria y control de fronteras.

4.2.1 EVALUACIÓN DE IMPACTO ALGORÍTMICA

Se debe realizar en los sistemas de inteligencia artificial siempre que el sistema sea considerado de alto riesgo. Consiste en un proceso iterativo continuo, que se ejecuta a lo largo del ciclo de vida de los sistemas de IA de alto riesgo.

Deberá tener en cuenta:

a) Riesgos conocidos y previsibles asociados con la inteligencia artificial en el momento en que se desarrolló, así como los riesgos que razonablemente se pueden esperar de él;
b) Beneficios asociados al sistema de inteligencia artificial;
c) Probabilidad de consecuencias adversas, incluido el número de personas potencialmente afectadas;
d) Gravedad de las consecuencias adversas, incluido el esfuerzo necesario para mitigarlas;
e) Lógica de funcionamiento del sistema de inteligencia artificial;
f) Proceso y resultados de pruebas y evaluaciones y medidas de mitigación realizadas para verificar posibles afectaciones a derechos, con especial énfasis en los impactos discriminatorios potenciales;
g) Formación y acciones de sensibilización sobre los riesgos asociados a los sistema de inteligencia artificial;

h) Medidas de mitigación e indicación y justificación del riesgo residual del sistema de inteligencia artificial, acompañado de pruebas de control de calidad frecuentes;

i) Medidas de transparencia al público, especialmente a los potenciales usuarios del sistema, en cuanto a los riesgos residuales, especialmente cuando impliquen un alto grado de nocividad o peligro para la salud o la seguridad de los usuarios (...).

(5) RESPONSABILIDAD CIVIL

El proveedor u operador del sistema de inteligencia artificial que cause daño material, moral, individual o colectivo está obligado a repararlo íntegramente, independientemente del grado de autonomía del sistema:

a) Cuando se trata de un sistema de inteligencia artificial de alto riesgo o riesgo excesivo, el proveedor u operador responde objetivamente por el daño causado, en la medida de su participación en el daño;

b) Cuando no se trate de un sistema de inteligencia artificial de alto riesgo o riesgo excesivo, se presumirá la culpa del causante del daño, aplicándose la inversión de la carga de la prueba a favor de la víctima.

(6) SANDBOX REGULATORIO

Se podrá autorizar la operación del entorno regulatorio experimental a la innovación en inteligencia artificial (sandbox regulatorio) para las entidades que lo soliciten y completen ciertos requisitos.

Características que deben cumplir los proyectos:

a) innovación en el uso de la tecnología o en el uso alternativo de tecnologías existentes;

b) mejoras hacia ganancias de eficiencia, reducción de costos, mayor seguridad, reducción de riesgos, beneficios para la sociedad y consumidores, entre otros;

c) plan de discontinuidad, con previsión de medidas a tomar para asegurar la viabilidad operativa del proyecto una vez que el período de autorización regulatorio del sandbox haya finalizado.

6.2 Proyecto de Ley 21/20 (año 2020)[13]

Este proyecto de ley fue presentado por el diputado Eduardo Bismark en el año 2020. Es uno de los antecedentes que consolidó el proyecto de ley 2338/23. A los efectos informativos y comparativos se incluye un apartado con los principales puntos abordados.

(1) PRINCIPIOS:

- Finalidad
- Centralidad en el ser humano
- No discriminación
- Transparencia y explicabilidad
- Seguridad
- Rendición de cuentas

(2) DERECHOS

- Derecho a conocer la institución responsable del sistema de inteligencia artificial;
- Derecho a acceder a información clara y adecuada sobre los criterios y procedimientos utilizados por el sistema que les afecten negativamente, observando secretos comerciales e industriales;
- Derecho a acceder a información clara y completa sobre el uso de sus datos sensibles.

(3) OBLIGACIONES

- Divulgar públicamente la institución responsable del establecimiento del sistema;
- Obligación de responder por los daños causados;
- Proporcionar información clara y adecuada sobre los criterios y procedimientos utilizados por el sistema de IA;
- Asegurar que los datos utilizados por el sistema de inteligencia artificial cumplan con la Ley General de Protección de Datos Personales;
- Implementar un sistema de IA solo después de la debida evaluación de sus objetivos, beneficios y riesgos y, no habilitar o utilizar un sistema si su control humano no es posible;
- Responder por decisiones tomadas por un sistema de IA;

[13] Cámara de Diputados de Brasil, *"Proyecto de Ley 21/20"*, disponible en: https://www.camara.leg.br/propostas-legislativas/2236340

- Proteger continuamente los sistemas de IA contra las amenazas de ciberseguridad.

7. ARGENTINA

a. Recomendaciones para una IA fiable. Disposición N° 2/2023 Subsecretaría de Tecnologías de la Información[14]

(1) Aspectos generales. Alcance y objetivos

Las recomendaciones tienen impacto en la Administración Pública Nacional y están orientadas, exclusivamente, a quienes formen parte del sector público, ya sea liderando proyectos de innovación, desarrollando tecnologías, adoptando tecnologías desarrolladas por otros equipos técnicos/proveedores, formulando las especificaciones técnicas para esas adquisiciones.

El documento, principalmente, se basa en los estándares internacionales sobre los principios éticos de la IA sobre los cuales existe consenso internacional. Una síntesis del contenido de este documento, plasmada a modo de cuadro comparativo, realizada por UBA IALAB, se encuentra disponible en "Argentina. Recomendaciones para una IA fiable. Disposición N° 2/23 de la Subsecretaría de Tecnologías de la Información", publicado por La Ley el 8 de junio de 2023, disponible en: http://laley.thomsonreuters.com/nota/7622

La referida disposición se apoya, básicamente, en tres documentos internacionales:

- Recomendación sobre la Ética de la Inteligencia Artificial de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO);
- Conferencia de Asilomar organizada por el Instituto "Future of Life";
- Principios de la OCDE sobre IA.

[14] Jefatura de Gabinete de Ministros, Disposición 2/2023, 1 de junio de 2023, disponible en: https://www.argentina.gob.ar/normativa/nacional/disposici%C3%B3n-2-2023-384656/texto

b. Banco Central de la República Argentina. Comunicación A 7724 "REQUISITOS MÍNIMOS PARA LA GESTIÓN Y CONTROL DE LOS RIESGOS DE TECNOLOGÍA Y SEGURIDAD DE LA INFORMACIÓN"[15]

(1) Responsabilidades (alta gerencia) respecto de la Tecnología y Seguridad de la Información

- Diseñar estrategias y planes de tecnología de la información y definir el presupuesto necesario para cumplirlos;
- Conocer y comprender los riesgos relacionados con la tecnología y la seguridad de la información, asegurar que sean contemplados en los programas de gestión establecidos y definir planes de mitigación de los riesgos detectados;
- Diseñar estrategias, planes y medidas de seguridad de la información, y definir el presupuesto necesario para cumplirlos;
- Definir y asegurar la implementación y el mantenimiento de políticas de alto nivel;
- Definir los roles y responsabilidades necesarios para los procesos de tecnología y seguridad de la información de manera coordinada y eficaz;
- Establecer un marco de gestión de la seguridad de la información que permita asegurar la identificación, prevención, detección, respuesta y recuperación ante ciberincidentes;
- Implementar las prácticas de control interno y gestión de riesgos, y garantizar que las decisiones de tecnología de la información se tomen de acuerdo con el apetito de riesgo de la entidad;
- Delinear un marco de gestión de continuidad del negocio, sus documentos asociados y los informes resultantes;
- Definir e implementar un esquema de control y monitoreo continuo de los procesos, servicios y/o actividades delegadas en las terceras partes;
- Asegurar la gestión de los conocimientos, habilidades y capacidades de acuerdo con las tecnologías utilizadas;
- Establecer mecanismos de comunicación y coordinación entre las áreas de gestión de riesgos, tecnología y seguridad de la información para el cumplimiento de sus objetivos;

[15] Disponible en https://www.bcra.gob.ar/Pdfs/comytexord/A7724.pdf

- Asegurar la incorporación en los proyectos de tecnologías de la información el principio de seguridad desde el diseño;
- Asegurar la realización de evaluaciones de impacto y definición de sistemas de riesgo para la utilización de inteligencia artificial;
- Aprobar los protocolos de comunicación y las responsabilidades ante situaciones de escenarios de crisis y/o emergencia;
- Asegurar que los requerimientos vinculados a la protección de los usuarios de servicios financieros sean contemplados en los procesos de tecnología correspondientes;
- Aceptar los riesgos residuales derivados de la gestión de riesgos de tecnología y seguridad.

(2) Comité de gobierno de tecnología y seguridad de la información

Las entidades deberán definir al menos un comité de gobierno de tecnología y seguridad de la información. Este comité deberá estar integrado, al menos, por un miembro del Directorio o autoridad equivalente, miembros de la Alta Gerencia, y los responsables de las áreas de tecnología y seguridad de la información. A su vez, se deberá procurar la participación de funcionarios de alto nivel de las otras áreas de acuerdo con los temas a tratar.

Sus responsabilidades incluirán, como mínimo:

- Vigilar y evaluar el funcionamiento del marco de gestión de tecnología de la información y contribuir a la mejora de su efectividad;
- Vigilar y evaluar el funcionamiento del marco de gestión de seguridad de la información y la efectividad del mismo;
- Supervisar las definiciones, la priorización y el cumplimiento de los planes de tecnología y seguridad de la información;
- Supervisar la efectividad del marco de gestión de continuidad del negocio y los mecanismos que aseguren resiliencia tecnológica;
- Supervisar la ejecución de las acciones correctivas tendientes a regularizar o minimizar las observaciones surgidas de los informes de las auditorías sobre los aspectos de tecnología y seguridad de la información;
- Monitorear los resultados del marco de gestión de riesgos relacionados con tecnología y seguridad de la información y verificar que los planes de mitigación sean ejecutados de acuerdo con los cronogramas definidos;
- Supervisar la gestión integral de ciberincidentes y los reportes asociados;

- Mantener informado al Directorio de los temas tratados y las decisiones tomadas.

(3) Gestión de riesgos de tecnología y seguridad de la información

Las áreas de tecnología y seguridad de la información serán responsables de efectuar la identificación de los riesgos y la definición técnica e implementación de las medidas de tratamiento.

La entidad deberá asegurar que el marco para la gestión del riesgo esté sujeto a un proceso de auditoría interna y externa. Además, se podrá involucrar a otros terceros independientes debidamente calificados.

Dentro de los riesgos relacionados con la tecnología y seguridad de la información incluidos en la evaluación, se deberán considerar especialmente los relacionados con:

- Escenarios que afecten la resiliencia tecnológica;
- La obsolescencia de la tecnología y los sistemas;
- La gestión de la relación con terceras partes;
- El desarrollo y utilización de algoritmos de inteligencia artificial o aprendizaje automático;
- La adopción de tecnología nueva o emergente;
- Software o aplicaciones utilizadas por usuarios que no fueron formalmente autorizados;
- Los aspectos de protección de datos personales en el uso de tecnologías de registros distribuidos (Distributed Ledger Technology - DLT);
- Escenarios de ciberincidentes relacionados con datos personales.

Adicionalmente, se deberán realizar evaluaciones de riesgos específicas:

- Antes del lanzamiento de nuevos productos o servicios que originen cambios importantes en los sistemas de información, en los procesos, servicios y/o actividades de tecnología y seguridad de la información;
- Antes de la delegación en terceras partes de procesos, servicios y/o actividades.

(4) Gestión de la Tecnología de la Información: Inteligencia artificial o aprendizaje automático

Las entidades deberán identificar y documentar el objetivo del uso, por sí o por terceros, del software que utilice algoritmos de inteligencia artificial o aprendizaje automático en sus proyectos o procesos.

Deberán establecer roles y responsabilidades para la definición del contexto en que operan los sistemas de inteligencia artificial o aprendizaje automático, la identificación de los modelos, algoritmos y los conjuntos de datos utilizados y la definición de métricas y umbrales precisos para evaluar la confiabilidad de las soluciones implementadas.

Los análisis de riesgos correspondientes deberán considerar, como mínimo:

- Los modelos adoptados, su entrenamiento y las posibles discrepancias con la realidad del contexto;
- Los datos utilizados para el entrenamiento, su volumen, complejidad y obsolescencia;
- La privacidad y la afectación a los usuarios en su calidad de consumidores;
- El nivel de madurez de los estándares de pruebas de software y las dificultades para documentar las prácticas basadas en IA.

Deberán implementar procesos que promuevan la confiabilidad en el uso de este tipo de algoritmos e incluyan al menos:

- Medidas para evitar la existencia de sesgos o discriminación contra grupos o segmentos de clientes o usuarios de los productos y/o servicios financieros;
- Documentación respecto de la transparencia, la explicabilidad de los modelos utilizados y la interpretabilidad de los resultados;
- La ejecución de revisiones periódicas de los resultados respecto de la tolerancia al riesgo definida;
- La comunicación al cliente cuando utilice servicios soportados por este tipo de tecnología.

(5) Gestión del ciclo de vida de “software”

Las entidades deberán establecer un marco para la gestión del ciclo de vida de desarrollo, adquisición y mantenimiento de software que contemple:

- Objetivos estratégicos del negocio;
- La arquitectura empresarial;

- La evaluación de los riesgos y la implementación de controles de mitigación de acuerdo con los lineamientos establecidos dentro de la sección 3. Gestión de riesgos de tecnología y seguridad de la información;
- Las metodologías de gestión de proyectos establecidas y los aspectos aplicables de la Sección 4. Gestión de Tecnología de la Información;
- Los aspectos aplicables de la Sección 5. Gestión de seguridad de la información;
- Los requerimientos de gestión de cambios definidos en la Sección 7;
- Infraestructura tecnológica y procesamiento.

Este marco de gestión deberá establecer como mínimo:

- La asignación de roles y responsabilidades;
- La documentación que describa las metodologías a utilizar en el ciclo de vida del software;
- Criterios para la evaluación de requerimientos;
- Procedimientos para la evaluación y selección de proveedores;
- Procedimientos de evaluación para la incorporación o integración de componentes de terceras partes en el ciclo de vida del software;
- Esto incluye código abierto, API y algoritmos de inteligencia artificial o aprendizaje automático;
- Criterios para la construcción y el uso de modelos de inteligencia artificial, los procesos de recolección y preparación de datos de entrenamiento, las actividades de verificación y validación de las respuestas;
- Estándares que establezcan buenas prácticas para el desarrollo y mantenimiento de software;
- Controles para asegurar la disponibilidad y actualización de los programas fuentes, y la documentación técnica y funcional;
- Procedimientos para la implementación del modelado de amenazas;
- Procedimientos que establezcan los criterios para la realización de pruebas de software y para la revisión de código;
- Planes de capacitación y concientización acordes a los roles y responsabilidades definidos;
- Procedimientos para las evaluaciones de seguridad en la adquisición de software y en la incorporación de componentes de software de terceras partes a los desarrollos propios;

- Procedimientos que establezcan criterios para la calidad de software y su aseguramiento

Las entidades deberán establecer procedimientos para el mantenimiento y control de los sistemas y aplicaciones que consideren:

- Evaluación y actualización de componentes obsoletos propios y de terceras partes;
- Cambios en los servicios de terceras partes consumidos por los sistemas de la entidad;
- Los resultados de la gestión de vulnerabilidades;
- La evolución de los sistemas y aplicaciones que utilizan algoritmos de inteligencia artificial y aprendizaje automático.

8. LEY 144/2021 DE LA CIUDAD DE NUEVA YORK[16]

Se trata de una ley promulgada, que aplica a empleadores que hacen uso de herramientas automatizadas para la selección y evaluación de personal. Es una ley con normas específicas para un uso de alto riesgo de la inteligencia artificial, puntualmente para decisiones de contratación y ascensos de personal.

La nueva ley:

- Obliga a las empresas que utilicen IA para clasificar candidatos a auditar sus herramientas de contratación en busca de sesgos;
- Exige que las empresas que usan softwares de IA para sus procesos de contratación de personal avisen a los candidatos que se está utilizando un sistema automatizado;
- Obliga a las empresas a tener auditores independientes que revisen la tecnología para verificar la ausencia de sesgos;
- Evalúa una "relación de impacto". Un cálculo del efecto que el uso de inteligencia artificial tendría en un grupo protegido de candidatos laborales. Es decir, no analiza

[16] Ley Local No. 144, Sección 1, Capítulo 5, Título 20, Subcapítulo 25. Disponible en: https://legistar.council.nyc.gov/LegislationDetail.aspx?ID=4344524&GUID=B051915D-A9AC-451E-81F8-6596032FA3F9&Options=Advanced&Search
Ampliar en: James, Barron, *"How New York Is Regulating A.I."*, The New York Times, 22 de junio de 2023, disponible en: https://www.nytimes.com/2023/06/22/nyregion/ai-regulation-nyc.html?searchResultPosition=17 (consultado el 25/06/2023)

cómo un algoritmo toma decisiones, sino que el enfoque se limita al resultado del algoritmo;
- Prevé que los candidatos podrán consultar acerca de los datos recopilados y cuál es su análisis. Se prevén multas en caso de infracción.

De esta manera, Nueva York se convierte en una ciudad pionera en la regulación de la IA a partir de una serie de reglas que desarrollan la ley promulgada en el año 2021, pero que esperan que influya en las prácticas a nivel nacional. La ciudad comenzará a hacer cumplir la ley el 5 de julio de 2023.

9. OTROS ENFOQUES. DOCUMENTOS RELACIONADOS

a.Un enfoque favorable a la innovación para la regulación de la IA en Reino Unido de Gran Bretaña e Irlanda del Norte[17]

El Departamento de Ciencia, Innovación y Tecnología del Reino Unido publicó el 29 de marzo de 2023 un Libro Blanco sobre Inteligencia Artificial que describe el enfoque de este país para regular la IA.

La propuesta busca crear un marco regulatorio favorable a la innovación, proporcionado, adaptable, claro y colaborativo, que promueva la confianza pública en la IA, mediante la creación de reglas proporcionales a los riesgos asociados con el uso de la IA en diferentes sectores. También se compromete a establecer una zona de pruebas regulatorias para comprender mejor cómo la regulación afecta las tecnologías emergentes de IA.

A diferencia de la Unión Europea (UE), el enfoque del Reino Unido hacia la IA no se centra en una nueva legislación a corto plazo. En su lugar, se centra en resultados, no crea reglas para sectores o tecnologías, sino que busca la creación de pautas para empoderar a los reguladores. La regulación específica del contexto se basará en los resultados que es probable que generen los usos específicos de la IA y solo se tomarán medidas legales cuando sea necesario.

El marco regulatorio cuenta con tres objetivos:

[17] https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper (consultado el 26/06/2023)

1. Impulsar el crecimiento y la prosperidad para facilitar la innovación responsable, reducir la incertidumbre regulatoria y obtener una ventaja de mercado a largo plazo en IA.
2. Aumentar la confianza pública en la IA, abordando sus riesgos y protegiendo los valores fundamentales lo que a su vez, impulsará la adopción de la IA.
3. Fortalecer la posición del Reino Unido como líder mundial en IA para que siga siendo atractivo para los innovadores e inversores y, al mismo tiempo, minimizar la fricción transfronteriza con otros enfoques internacionales.

Este marco regulatorio no afectará las cuestiones relacionadas con el acceso a los datos, la capacidad informática y la sostenibilidad o el "equilibrio de los derechos de los productores de contenido y los desarrolladores de IA".

El gobierno del Reino Unido se centra en cinco principios que cree que deberían regir la IA para fomentar el desarrollo y el uso responsable de la tecnología. La aplicación de estos cinco principios estará inicialmente a discreción de los reguladores y puede ser seguida por un deber legal que requiera que los reguladores tengan debidamente en cuenta los principios:

1. Seguridad, protección y solidez;
2. Transparencia y explicabilidad adecuadas;
3. Equidad;
4. Rendición de cuentas y gobernanza;
5. Contestabilidad y compensación.

Los usuarios y otras partes interesadas necesitan rutas claras para disputar cualquier daño causado por la IA. El gobierno espera que los reguladores aclaren las rutas existentes, alienten y guíen a las entidades reguladas para asegurarse de que las partes afectadas puedan impugnar claramente los resultados dañinos de la IA a través de canales informales o formales.

A partir de este nuevo enfoque, se propone que los reguladores de IA:

1. Adopten un enfoque proporcionado, a favor del crecimiento y a favor de la innovación que se centre en los riesgos específicos que plantea la IA específica;
2. Consideren medidas proporcionadas para abordar los riesgos priorizados, considerando las evaluaciones de riesgo realizadas por o para el gobierno;
3. Diseñen, implementen y hagan cumplir los requisitos reglamentarios apropiados que integren los nuevos principios reglamentarios de IA en los procesos existentes;

4. Desarrollen una guía conjunta para apoyar el cumplimiento de IA con los principios y requisitos relevantes;
5. Consideren cómo las herramientas, como las técnicas de aseguramiento y los estándares técnicos, pueden respaldar el cumplimiento;
6. Se comprometan con el monitoreo y la evaluación del marco por parte del gobierno.

Los nuevos cinco principios mencionados precedentemente serán implementados por los reguladores para adaptarlos al contexto y al uso de la IA. También colaborarán para identificar las barreras a la implementación de los principios, mientras que el gobierno asumirá un papel central de apoyo para garantizar que el marco funcione de manera proporcional y beneficie la innovación de la IA.

Solo en caso de resultar necesario, el gobierno introducirá nueva legislación para crear medidas adicionales que exijan a los reguladores implementar los principios relevantes para sus sectores o dominios.

Se espera que la naturaleza adaptable y proporcionada de este nuevo marco regulatorio ayude a establecer normas globales para la regulación de IA preparada para el futuro.

b. Libro Blanco de Inteligencia Artificial de Japón[18]

En Japón ha habido regulaciones analógicas que inhibieron el uso de tecnologías, incluida la IA. Desde el año 2019, Japón sostiene la idea básica de no contar con un marco regulatorio y vinculante para la IA.

A partir de aquel entonces, se han llevado a cabo discusiones que siguen el enfoque de no imponer regulaciones sobre el desarrollo y uso de la IA para evitar frenar la innovación. Por otra parte, se propone la formulación conjunta de herramientas de política entre los sectores públicos y privados a través de instrumentos de soft law.

A la luz del desarrollo de GPT y otros modelos básicos, sumado a los avances significativos logrados en el campo de la IA en estos últimos meses, en abril del 2023 el gobierno de Japón presentó una serie de recomendaciones para la elaboración de una nueva estrategia

[18] Japan's National Strategy in the New Era of AI, "The AI White Paper", abril de 2023, disponible en: https://www.linkedin.com/posts/hiroyuki-sanbe-32175922_the-japans-ai-white-paper-english-translaiton-activity-7055817735087292416-Z2FW/ (consultado el 28/6/2023)

nacional de inteligencia artificial, capaz de situar a Japón en una posición de ventaja competitiva internacional.

Por un lado, Japón propone el establecimiento de una política de IA que signifique una ampliación de su estructura dentro del gobierno, junto con el estudio inmediato y exhaustivo de medidas desde una amplia gama de perspectivas: investigación y desarrollo, estructura económica, infraestructura social, recursos humanos, desarrollo de recursos y garantía de seguridad nacional.
Por otro lado, la Agencia Digital de Japón trabaja actualmente en una revisión exhaustiva de las regulaciones analógicas y está en proceso de revisión de todas las leyes y reglamentos existentes.

Las recomendaciones efectuadas comprenden:

1. Desarrollo y fortalecimiento de la capacidad de desarrollo de IA en Japón
   -Redefinición de objetivos y enfoques de política para la investigación y desarrollo de modelos de IA, incluidos los modelos básicos;
   -Acelerar la investigación y el desarrollo;
   -Continuar invirtiendo y apoyando el desarrollo tecnológico;
   -Establecimiento de un "centro de IA" que recopile información sobre IA y sirva como punto de contacto entre empresas para apoyar la formación de una comunidad.
2. Acumulación y coordinación de recursos de datos
   -Crear un entorno que promueva el uso de datos del sector público y privado a través de la IA. Promover aún más el desarrollo de un modelo estándar de datos y aclarar su estructura y atributos;
   -Organizar las reglas y el formato para facilitar los datos a terceros.
3. Mejora y utilización de recursos computacionales
   -Apoyo por parte del gobierno para el desarrollo y la expansión de los recursos informáticos nacionales, a fin de apoyar la innovación y el crecimiento económico.
4. Promoción de la utilización activa de la IA en el servicio público
   -Investigar aplicaciones avanzadas de IA en organizaciones gubernamentales de otros países y directrices para tales aplicaciones, con el fin de informar la planificación e implementación de la introducción de IA en Japón.
   -Llevar adelante pruebas piloto con resultados visibles a corto plazo;

-Redacción de respuestas parlamentarias, asistencia en asuntos legislativos y análisis de estadísticas gubernamentales;

-Establecer un equipo especializado dentro del gobierno para relevar, analizar y compartir casos de uso de IA y apoyar la introducción de IA en organizaciones relevantes.

5. Apoyo a la promoción de ciudades inteligentes con IA
6. Políticas para fomentar y apoyar el uso de IA en el sector privado

   -Realizar estudios sobre el impacto de la IA en diversas industrias nacionales;

   -Fomentar la creación de empresas emergentes y nuevos negocios que utilicen IA.

   -Profundizar los debates sobre cómo se debe implementar la gobernanza de la IA, tanto para gestionar los riesgos, como para fomentar la innovación y, si resulta necesario, establecer pautas.
7. Nuevos enfoques para la regulación de la IA

   -Analizar el estado de las regulaciones de IA en la Unión Europea, Estados Unidos y otros países. Llevar a cabo estudios específicos sobre áreas en donde las medidas, incluyendo leyes y reglamentos, se consideren necesarias en la nueva era de la IA como violaciones graves de los derechos humanos, seguridad nacional, intervención injustificada en los procesos democráticos;

   -Participar activa y estratégicamente en la elaboración de normas internacionales;

   -Adaptación regulatoria ágil a la nueva era de la IA;

   -Combinación de métodos como la formulación de directrices y normas para garantizar actualizaciones oportunas y flexibles en respuesta a la evolución tecnológica;

   -Establecer un mecanismo para promover aún más la revisión de las regulaciones analógicas;

   -Mejorar la velocidad y la facilidad de uso de los procedimientos de desregulación actuales, como los sandboxes y la eliminación de zonas grises para crear y desarrollar un entorno en el que las empresas puedan tomar nuevos desafíos sin estar limitadas por la regulación existente;

   -Establecer pautas para la interpretación de leyes de propiedad intelectual en materia de IA generativa, que promuevan el progreso de la IA, mientras prevengan su uso abusivo y favorezcan el desarrollo de la industria de contenidos;
8. Organizar pautas para la utilización de IA en la educación

-Mejora en la alfabetización de IA en el plan de estudios de educación pública;
-Formular pautas para el manejo de IA en la educación pública.

CONCLUSIONES

De la lectura de la sistematización realizada en el presente documento se observa que la mayoría de los proyectos detallan los principios que rigen los desarrollos y la implementación de sistemas basados en inteligencia artificial. Algunos plantean un enfoque regulatorio basado en riesgos y proponen una técnica legislativa que regula prácticas prohibidas, como así también pautas específicas que se deben aplicar a los sistemas considerados de "alto riesgo". Se puede observar que en general todos tienen similares características.

Salvo la regulación de China y la enmienda realizada a la propuesta de regulación del Parlamento Europeo, son pocas las regulaciones que abordan expresamente los efectos de la IA generativa.

Creemos que será clave la aprobación por el Parlamento Europeo y el Consejo de la propuesta de Reglamento que se debate en aquel ámbito, pues significará el nacimiento de la primera norma vinculante, regulatoria de la IA, de amplio alcance. A partir de ello y con la entrada en vigencia del Reglamento se evidenciará en la práctica cuáles son los efectos que este tipo de previsiones (hard law) tienen en el campo de la inteligencia artificial.