

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

FACULTAD DE DERECHO



La Inteligencia Artificial y la Administración Pública en el Perú: La regulación de la
transparencia en el uso de tecnologías de Inteligencia Artificial por parte de la
Administración Pública

Tesis para obtener el título profesional de Abogado presentado por:

Hernández Lévano, Christian Gabriel

Tesis para obtener el título profesional de Abogada presentado por:

Molina Mendoza, Lucia Alejandra

Asesor:

Zegarra Valdivia, Diego Hernando

Lima, 2025

Informe de Similitud

Yo, **Diego Hernando Zegarra Valdivia**, docente de la **Facultad de Derecho** de la Pontificia Universidad Católica del Perú, asesor de la tesis titulada:

La Inteligencia Artificial y la Administración Pública en el Perú: La regulación de la transparencia en el uso de tecnologías de Inteligencia Artificial por parte de la Administración Pública

De los autores:

Lucía Alejandra Molina Mendoza
Christian Gabriel Hernández Lévano

Dejo constancia de lo siguiente:

- El mencionado documento tiene un índice de puntuación de similitud de 32%. Así lo consigna el reporte de similitud emitido por el software *Turnitin* el **07/08/2025**.
- He revisado con detalle dicho reporte y confirmo que cada una de las coincidencias detectadas no constituyen plagio alguno.
- Las citas a otros autores y sus respectivas referencias cumplen con las pautas académicas.

Lugar y fecha: Lima, 7 de agosto del 2025.

Apellidos y nombres del asesor / de la asesora: ZEGARRA VALDIVIA, DIEGO HERNANDO	
DNI: 07875295	Firma: 
ORCID: https://orcid.org/0000-0002-8901-1026	

*“He odiado las palabras, y las he amado;
y espero haber estado a la altura”*

Markus Suzak



AGRADECIMIENTOS

Le agradezco profundamente a Dios, quien me ha permitido llegar a este momento rodeada de todo lo que aprecio y me hace feliz.

A mis padres, a quienes les expreso mi más sincero agradecimiento por su amor y su apoyo incondicional, que han sido mi refugio a lo largo del tiempo y la distancia. A mi padre, Lizardo, por su sabiduría, y ejemplo de fortaleza y templanza, su presencia constante y su cariño han sido fundamentales, no solo en la realización de este trabajo, sino en cada etapa de mi vida. A mi madre, Frida, por su alegría y serenidad, por la confianza que ha depositado en mí y por su amor inagotable, sin condiciones ni reservas, que día a día me enseña entregar todo lo que soy.

Al doctor Diego Zegarra, sin quien este trabajo no habría sido posible. Le agradezco por la oportunidad de ser parte de su equipo de docencia por allá en el 2020, y por ser un modelo a seguir en el ámbito académico y un referente de excelencia y compromiso.

A mi abuela, Presentación, que festeja todos mis logros con la mayor ilusión; y a mi abuelo, Humberto, con la certeza de que, si le fuera posible, regresaría solo para compartir conmigo este instante.

A mis hermanos, Sebastián y Nicolás, por su paciencia, su cariño y por haberme enseñado el verdadero significado del amor fraternal, aquel que permanece inalterable más allá de las circunstancias.

A mis amigos, Karen Medina, Mayra Vargas, Jorge Mora, Marcelo Salas, André Montoya y Lucía López, quienes, de manera consciente o inadvertida, han sido parte esencial de este proceso. A Mayra Ramón, Cinthya Delgado de la Flor y Braylyn Paredes, por su amistad y su valiosa orientación en el GIDA. A Cecilia Dávila, Ariana Suárez y Cristina Díaz, por su constante apoyo y por las enseñanzas que han enriquecido mi crecimiento profesional.

Finalmente, a Christian Hernández, a quien agradezco profundamente por su amistad y con quien ha sido un privilegio compartir la autoría de este trabajo.

A todos y cada uno de ustedes, mi más sincero reconocimiento y gratitud.

Lucia Alejandra Molina Mendoza

Agradezco a Dios por la culminación de este trabajo y por mantener con salud y paz a quienes me rodean.

A mi mamá, Mónica, por representar un apoyo constante en mi camino profesional y personal. Su organización y perseverancia en el trabajo solo son superadas por su infinito amor a mis hermanos y a mí.

A mi papá, Christian Arturo, por ser mi inspiración para convertirme en abogado y por enseñarme que cada día es una nueva oportunidad para enfrentar la vida con alegría, optimismo y corazón.

A mis hermanos Daniel y Mónica Lucía, que algún día serán los mejores abogados del mundo. A mi hermano Daniel, por su bondad y generosidad, y por su deseo constante de superarse. A mi hermana Lucía, por su risa y amistad, su inteligencia y sus ganas de ser feliz.

A mis primos, en especial a Adela, David, Aharon y Joan, por regalarme siempre una sonrisa cuando la necesito y por ser un ejemplo para mí en mi crecimiento profesional.

Al doctor Diego Zegarra Valdivia, por asesorarnos en el presente trabajo. Le agradezco la oportunidad de formar parte de su equipo de docencia de derecho administrativo y del GIDA allá por el año 2021 y por ser un ejemplo a seguir para mí en lo académico y profesional.

A mis amigas, Almendra Fernández, Luciana Contti, Maria Fernanda Herrera, Marycielo Ivoska y Valeria Rojas, por ser parte de esto directa o indirectamente, las aprecio mucho.

A Mónica Nieva, Giuliana Gonzales, Bruno Bardellini y Mario Agramonte, por su mentoría, paciencia y enseñanzas, las cuales me acompañan siempre.

Finalmente, a Lucía Molina, con quien compartí buena parte de mi vida universitaria realizando actividad académica en el GIDA y como asistentes de docencia, le agradezco por su sincera amistad y por compartir la autoría de este estupendo trabajo.

A todos y cada uno de ustedes, mi más sincero reconocimiento y gratitud.

Christian Gabriel Hernández Lévano

RESUMEN

El acelerado avance tecnológico respecto a las nuevas tecnologías de Inteligencia Artificial (IA) no ha venido acompañado del desarrollo de un marco normativo que regule de manera efectiva el nuevo escenario en el que la Administración Pública emplea tecnologías de IA en sus actividades. En dicho contexto, el presente trabajo busca analizar las problemáticas generadas por la insuficiencia del marco normativo para garantizar la transparencia en un contexto novedoso que involucra tecnologías de funcionamiento complejo. En dicho sentido, el presente trabajo busca responder a la interrogante de si la normativa actual sobre IA y transparencia regula de forma efectiva y suficiente el nuevo contexto, contemplando las particularidades que existen y permiten preservar estándares aceptables de transparencia en un contexto de automatización de la actividad administrativa. Con dicho fin, se ha revisado la doctrina relevante para el presente problema, de origen nacional e internacional, así como los marcos normativos aplicables, nacionales y extranjeros. El presente trabajo concluye que el marco normativo actual relativo a IA y transparencia resulta insuficiente, pues su redacción no considera las particularidades del funcionamiento de los algoritmos de IA, y el dinamismo de estas tecnologías.

Palabras clave: Inteligencia Artificial – Transparencia – Administración Pública– Machine Learning – Deep Learning.

ABSTRACT

The rapid technological advancement in Artificial Intelligence (AI) has not been accompanied by the development of a regulatory framework capable of effectively governing the new landscape in which Public Administration employs AI technologies in its activities. In this context, the present study aims to analyse the challenges arising from the insufficiency of the regulatory framework in ensuring transparency within a novel setting that involves complex-functioning technologies. Accordingly, this study seeks to answer the question of whether the current AI and transparency regulations effectively and adequately govern this new context, taking into account its specific characteristics and ensuring acceptable transparency standards in an environment of administrative activity automation. In this sense, relevant national and international bibliographical sources have been reviewed, along with applicable national and foreign regulatory frameworks. This study concludes that the current regulatory framework concerning AI and transparency is insufficient, as its wording fails to consider the specificities of AI algorithms' functioning and the dynamic nature of these technologies.

Key words: Artificial Intelligence – Transparency – Public Administration – Machine Learning – Deep Learning.

ÍNDICE DE CONTENIDO

INTRODUCCIÓN.....	6
INTELIGENCIA ARTIFICIAL: DEFINICIÓN, CONCEPTOS ASOCIADOS, Y CICLO DE VIDA.....	8
1.1 Aproximación al concepto de inteligencia artificial.....	8
1.1.1 Antecedentes históricos.....	9
1.1.2 Definición de inteligencia artificial.....	14
a) Desde su equiparación con la inteligencia humana.....	15
b) Desde su capacidad de solución a problemas complejos.....	17
c) Desde sus componentes o capacidades.....	17
d) Desde el punto de vista tecnológico.....	19
e) Desde el punto de vista jurídico.....	20
1.2 Conceptos asociados a la inteligencia artificial.....	25
1.2.1 <i>Machine Learning</i>	26
1.2.2 Redes neuronales.....	27
1.2.3 <i>Deep learning</i>	28
1.2.4 Algoritmos.....	29
a) Aprendizaje supervisado.....	30
b) Aprendizaje no supervisado.....	32
i) Clustering.....	33
ii) Biclustering.....	34
1.2.5 Big data.....	34
1.3 Ciclo de vida de la IA.....	36
1.3.1 Diseño del proyecto.....	38
1.3.2 Evaluación y validación del proyecto.....	39
1.3.3 Despliegue y mantenimiento del proyecto.....	40
POLÍTICAS PÚBLICAS, MARCO JURÍDICO E INSTITUCIONALIDAD DE LA INTELIGENCIA ARTIFICIAL EN LA ADMINISTRACIÓN PÚBLICA.....	44
2.1 Políticas públicas sobre inteligencia artificial.....	44
2.1.1 Políticas públicas sobre inteligencia artificial en el mundo.....	44
a) Estados Unidos.....	45
b) China.....	50

c)	Unión Europea.....	52
2.1.2	Políticas públicas sobre inteligencia artificial en el Perú	53
a)	La Estrategia Nacional de Inteligencia Artificial	53
b)	Inteligencia Artificial: Desafíos y oportunidades para el Perú.....	59
c)	Política Nacional de Transformación Digital	61
2.2	Marco regulatorio sobre inteligencia artificial.....	65
2.2.1	¿Es necesario regular la inteligencia artificial?	68
2.2.2	La noción de regulación aplicable a la inteligencia artificial	71
2.2.3	Enfoques regulatorios de la inteligencia artificial	72
a)	Normas vinculantes o <i>hard law</i>	73
b)	Normas indicativas o <i>soft law</i>	73
2.2.4	Marco jurídico de la inteligencia artificial en el Perú	75
a)	Antecedentes legislativos	75
b)	Ley de inteligencia artificial peruana	79
c)	Reglamentación de la ley de inteligencia artificial peruana	82
2.3	Régimen institucional del uso de la IA en las actividades de la Administración Pública.....	88
LA GARANTÍA DE TRANSPARENCIA EN LA REGULACIÓN DEL USO DE INTELIGENCIA ARTIFICIAL DE LA ADMINISTRACIÓN PÚBLICA: DESAFÍOS Y PROPUESTA DE SOLUCIÓN		
91		
3.1	La garantía de transparencia en el uso de algoritmos por parte de la Administración Pública.....	92
3.1.1	Definición de transparencia algorítmica.....	92
3.1.2	Tipologías de la transparencia algorítmica.....	95
a)	Desde la disponibilidad de información	95
b)	Según la capacidad del sistema de inteligencia artificial	96
i)	Explicabilidad.....	97
ii)	Auditabilidad	99
iii)	Trazabilidad	100
3.1.3	La transparencia algorítmica en la regulación de transparencia y acceso a la información pública.....	102
a)	Regulación de la publicidad activa.....	102
b)	Regulación del derecho de acceso a la información pública respecto del uso de tecnologías de IA por parte de la Administración Pública	105
i)	La titularidad del derecho de acceso a la información pública en el marco del uso de tecnologías de IA por parte de la Administración Pública	105

ii)	Definición de información pública y causas de inadmisión en el marco del uso de tecnologías de IA por parte de la Administración Pública	107
3.2	Situación actual sobre la transparencia en el contexto del uso de inteligencia artificial	109
3.2.1	Experiencia internacional frente a la transparencia algorítmica.....	111
a)	Transparencia activa: Repositorios públicos de algoritmos	112
b)	Transparencia pasiva: Casos límite para la transparencia algorítmica	115
i)	Datos personales y propiedad intelectual	115
ii)	Contratación pública.....	118
iii)	Secreto fiscal o reserva tributaria	119
3.2.2	Experiencia nacional respecto de la transparencia algorítmica	120
a)	Implementación de sistemas de inteligencia artificial para el uso de las entidades de la Administración Pública del Perú	121
b)	Garantía de transparencia en el uso de inteligencia artificial por parte de la Administración Pública del Perú	127
i)	Principios de transparencia en el marco normativo de inteligencia artificial peruano	127
ii)	Obligaciones de la Administración Pública en el marco normativo de inteligencia artificial peruano	128
iii)	Mecanismos de transparencia específicos en el marco normativo de inteligencia artificial peruano	129
3.3	Desafíos para la garantía de transparencia en el uso de la inteligencia por parte de la Administración Pública en el Perú	129
3.4	Propuesta de solución	133
CONCLUSIONES.....		140
BIBLIOGRAFÍA		142

INTRODUCCIÓN

El creciente desarrollo tecnológico ha permitido grandes avances en la infraestructura tecnológica de las administraciones públicas, quienes cada vez incorporan a sus actividades nuevas tecnologías que permiten modernizar sus procesos y actividades. Dicho desarrollo tecnológico, pese a que viene teniendo lugar durante las últimas décadas, no ha sido acompañado de un desarrollo normativo que permita salvaguardar el cumplimiento y respeto de todos los principios y máximas con las que la Administración Pública opera.

El presente trabajo de investigación busca responder a la pregunta “¿El marco normativo peruano actual sobre IA y transparencia resulta suficiente para salvaguardar el principio de transparencia en las actividades de la Administración Pública?” cuya respuesta, a raíz de un análisis de la normativa y bibliografía pertinente, es negativa.

El análisis normativo para responder a la pregunta de la investigación se ha circunscrito principalmente a la regulación específica sobre IA y transparencia en el Perú, siendo ésta la Ley N.º 31814 – Ley que promueve el uso de Inteligencia Artificial en favor del desarrollo económico y social del país- (en adelante, “Ley de IA peruana”), la Ley N.º 27806 –Ley de Transparencia y Acceso a la Información Pública–, su Texto Único Ordenado (Decreto Supremo N.º 021-2019-JUS) y su Reglamento (Decreto Supremo N.º 007-2024-JUS).

Ciertamente, las propias características de las tecnologías de IA conducen a una oscuridad u opacidad algorítmica inherente, la cual la regulación peruana actual no ha podido atender. Aunado a ello, las propias características de las tecnologías de IA hacen difícil que la normativa pueda regularla en sus constantes reinventiones.

A fin de esclarecer la premisa planteada, la investigación realizada es de carácter dogmático jurídico, para la cual se ha realizado una revisión de fuentes bibliográficas y cuerpos normativos relevantes. Asimismo, se ha realizado un análisis cualitativo, empleando razonamientos deductivos e inductivos a fin de determinar la suficiencia o insuficiencia del marco normativo peruano aplicable.

En dicho sentido, la presente investigación se divide en tres capítulos. El primero, busca esclarecer el concepto de IA, proporcionando, analizando y clasificando las principales definiciones que al respecto ha producido la doctrina; asimismo, se busca brindar definiciones de conceptos relacionados, que resultan esenciales a fin de esclarecer la pregunta planteada. El segundo capítulo busca contextualizar las herramientas de IA en el marco de las actividades de la Administración Pública peruana, para lo cual se ha realizado un análisis del marco normativo e institucional vigentes, así como de bibliografía pertinente que esclarece el nuevo panorama de la infraestructura tecnológica de la Administración Pública que incorpora estas nuevas tecnologías. Finalmente, el tercer capítulo busca revisar la suficiencia de la normativa sobre IA, esta vez de cara a la normativa sobre transparencia, a fin de identificar si estas son suficientes para asegurar la transparencia en un contexto como el detallado.

CAPÍTULO 1

INTELIGENCIA ARTIFICIAL: DEFINICIÓN, CONCEPTOS ASOCIADOS, Y CICLO DE VIDA

El presente capítulo busca revisar los alcances del concepto de IA, el cual representa un eje importante en el desarrollo del presente trabajo. Con dicho fin, el primer capítulo brindará las aproximaciones al concepto de IA, revisando las principales definiciones aportadas por la doctrina. Asimismo, se revisarán los conceptos asociados, a fin de poder alcanzar un entendimiento pleno de los sistemas de IA, en el contexto en el que se desarrollan. Finalmente, se revisará el ciclo de vida de estas tecnologías, a fin de entender su proceso de desarrollo.

1.1 Aproximación al concepto de inteligencia artificial

Para poder determinar un concepto de IA en un sentido jurídico o relacionado con la regulación, primero debemos entender qué se entiende por IA. Brindar un concepto de IA que permita realizar un análisis jurídico posterior es de suma importancia para los objetivos del presente trabajo. La finalidad de ello es poder construir las bases y contar con las herramientas para realizar un análisis sobre la suficiencia o insuficiencia de la normativa actual para asegurar la transparencia en el escenario actual, en el que se avizora el uso de sistemas, algoritmos y herramientas de IA por parte de la Administración Pública con mayor rigor.

Para ello, antes de arribar a dicha definición, serán desarrollados los antecedentes históricos de la IA y las principales aproximaciones o definiciones que la literatura ha dado sobre la IA. Este entendimiento de la IA se postulará desde distintas perspectivas que la doctrina ha empleado para proporcionar un concepto a la IA.

1.1.1 Antecedentes históricos

Si bien la expresión “*inteligencia artificial*” se escuchó por primera vez en la conferencia *Dartmouth Summer Research Project on Artificial Intelligence* de John McCarthy en 1956, Sartor y Veiga colocan como el primer hito histórico de la IA a la propuesta de Warren McCulloch y Walter Pitts de un modelo matemático que tiene la finalidad de crear una red neuronal (2023, p. 97).

En 1950, serían Marvin Minsky y Dean Edmonds los primeros en construir un ordenador con un sistema que cuenta con un sistema de red neuronal. El mismo año, Alan Turing consolidó el trabajo de McCulloch y Pitts con su artículo *Computing Machinery and Intelligence*, en el que propuso el llamado *Test de Turing*: una prueba diseñada con el fin de determinar si una máquina es inteligente o no. Dicha prueba alcanzó tal fama que Alan Turing es considerado el padre de la IA¹ (Rebollo, 2023, p. 20). Herrera et al. señalan que el artículo de Turing “reproducía la conferencia que impartió en 1947 en el Laboratorio Nacional de Física del Reino Unido que comenzaba con una pregunta que continúa siendo el eje central del debate en nuestros días: ¿Pueden pensar las máquinas?” (2022, p. 55).

El Test de Turing consiste en una dinámica compuesta de preguntas y respuestas, en la que una persona sostiene una conversación con un interlocutor al que no puede ver e intenta identificar si se trata de un ser humano o de un ordenador, si este último puede hacer creer al interlocutor que está hablando con una persona, se podría determinar que estamos ante un ordenador inteligente (Herrera et al., 2022, p. 55).

¹ Rebollo menciona que Turing tuvo un trabajo anterior titulado “Números Calculables”, donde explica las bases teóricas para todas las ciencias de computación y que puede considerarse como el origen oficial de la informática teórica. Señala además que en este trabajo introduce el concepto de Máquina de Turing, la cual podía conceptualmente leer instrucciones de una cinta de papel perforada y ejecutar todas las operaciones críticas de un computador (2023, p. 19).

Los años posteriores al lanzamiento del *Test de Turing* fueron años bastante prósperos para el desarrollo de las tecnologías de IA, al punto de ser reconocidos como la *edad de oro de la IA*. Herrera et al. identifican que la IA transitó una serie de etapas, las cuales son resumidas a continuación:

Etapa	Años	Hechos
Edad de oro de la IA	1956-1972	<ul style="list-style-type: none"> – En 1956, Newell, Simon y Shaw crearon el primer programa para ordenador orientado y capaz de resolver problemas matemáticos, llamado <i>Logic Theorist</i>. – En 1957, Rosenblatt propone el perceptrón, que era un modelo neuronal que formaba parte de una investigación cuyo objetivo era diseñar un modelo de cerebro electrónico capaz de aprender patrones complejos y generalizar (Herrera et al., 2022, p. 57). – En 1959, mientras trabajaba en IBM, Arthur Samuel acuña el término “<i>machine learning</i>”. Años antes se le atribuyó el diseño de un <i>software</i> autónomo con la capacidad de jugar ajedrez (Sartor y Veiga, 2023, p. 97). – En los años sesenta aparecen los sistemas expertos, que permitían predecir soluciones con premisas. Dos de los más conocidos y pioneros son <i>Dentral</i>² y <i>Mycin</i>³. – En 1966, Josep Wizenbaum desarrolló en el Instituto Tecnológico de Massachusetts “<i>ELIZA</i>”, que según Herrera et al. podría ser considerado “el primer chatbot del mundo” (2022, p. 57).

² Desarrollado por Edward Feigenbaum y otros programadores de la Universidad de Stanford entre 1965-1975.

³ Desarrollado por Edward Shorliffe a principio de los setentas.

<p>Etapa de invierno de la IA</p>	<p>1973-1983</p>	<ul style="list-style-type: none"> - En 1969, Minsky y Papert publican su libro “Perceptron” en el que critican el Perceptrón de Rosenblatt, indicando que tenía deficiencias en tareas como diferenciar una figura de forma independiente respecto al fondo, diferenciar letras en distintos grados de rotación y ejecutar la función lógica de disyunción exclusiva, lo que provocó caídas en las investigaciones sobre redes neuronales (Herrera et al., 2022, p. 58). - En 1972, Hubert L. Dreyfus publica su obra <i>What computer still can't do. A critique of Artificial Intelligence</i> en donde discute que la computación en IA pudiera desarrollar sistemas inteligentes de actuación general con capacidad de realizar cualquier tarea humana. (Herrera et al., 2022, p. 58). - En 1982, David Parker propone un algoritmo que proporcionaba una solución práctica para su aprendizaje y entrenamiento, el algoritmo de <i>backpropagation</i>.
<p>Etapa de la IA moderna</p>	<p>1983-1999</p>	<ul style="list-style-type: none"> - En 1988, Feng-hsiung diseñó “<i>Deep Thought</i>”, un ordenador con hardware desarrollado específicamente para jugar ajedrez, el cual venció al gran maestro danés Bent Larsen. En 1989 se unió a IBM donde rebautizó su proyecto como “<i>Deep Blue</i>”⁴. En 1996 y 1997, Deep Blue venció al campeón mundial de ajedrez Gary Kaspárov. - En 1995, Ernst Dickmanns desarrolló un sistema autónomo que hizo que un coche modelo Mercedes Clase S hiciera 158 km sin intervención humana y alcanzando los 180 km/h. Los resultados de automóviles autónomos fueron avanzando en

⁴ Para 1997, Deep Blue era una máquina de procesamiento masivo paralelo con 30 nodos y 1,4 toneladas de peso con capacidad para calcular 200 millones de posiciones por segundo.

		2005, con el <i>Stanley</i> preparado por un equipo de la Universidad de Stanford. Dicho vehículo fue capaz de recorrer 212 km. Otro ejemplo más reciente de la conducción autónoma es el Honda Legend presentado en Japón en 2021. ⁵
--	--	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Posteriormente a la Edad Moderna -y hasta nuestros días- llega la “Era del *big data*”, ello al encontrarnos inmersos en la era de la información y registrarse diariamente grandes cantidades de datos. Es en esta época del *internet de las cosas* o *Internet of Things* (IoT) donde son fundamentales los algoritmos de IA derivados del aprendizaje automático distribuidos para el procesamiento masivo de datos. Nótese que el enfoque de los sistemas de IA basada en aprendizaje automático en función de datos no es nuevo, ya que fue vislumbrada por Turing y aplicada por Rosenblatt en los cincuenta con la primera red neuronal. Lo que sucede es que en dichos años no se disponían de los recursos computacionales ni la masividad de datos de ahora (Herrera et al., 2022, p. 61).

En relación con el *big data*, de acuerdo con Mena Roa (2021), la cantidad de datos a nivel mundial se ha multiplicado por más de treinta desde 2010, pasando de dos zetabytes en dicho año a 64 zetabytes en 2020. Según sus previsiones para el 2025 serían superados los 180 zetabytes⁶, lo que representaría un crecimiento medio anual de casi el 40% en cinco años. Esto último se daría por el auge del IoT, el desarrollo de la tecnología 5G y las redes sociales.

⁵ Según Herrera et al. el Honda Legend llega a una autonomía nivel 3, lo que significa una autonomía importante pero que aún reclama el control del conductor en ciertos momentos (2022, p. 60).

⁶ En 2024 el volumen de datos globales llegó a 149 zetabytes.

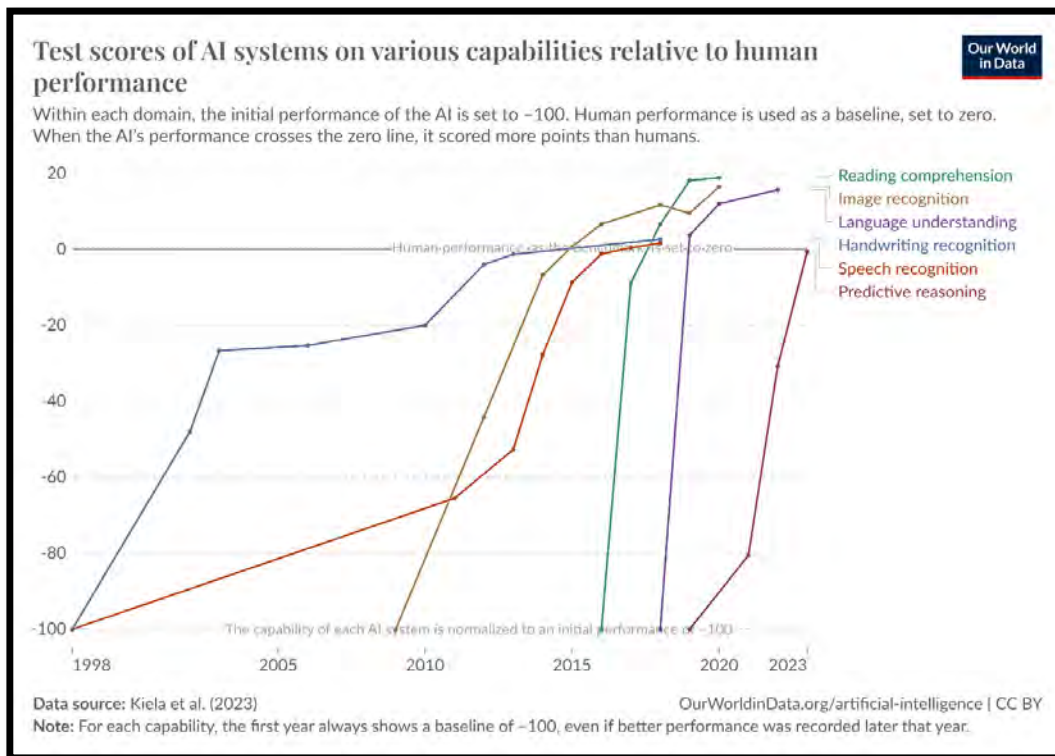
Así las cosas, es una realidad la era del *big data*, los datos alrededor del mundo siguen creciendo, lo que coloca el tema objeto de estudio como sumamente relevante, ya que la herramienta principal de la que se sirven las tecnologías IA son los datos.

Sobre los conceptos “*big data*” y IoT, el artículo 3° de la Ley N.º 31814, Ley que promueve el uso de la IA en favor del desarrollo económico y social del país y su Proyecto de Reglamento, publicado el 19 de noviembre de 2024, indican que se trata de tecnologías emergentes que involucran técnicas avanzadas de producción y operaciones tecnológicas, a fin de generar un ecosistema digital que contribuya al desarrollo social y económico.

Sobre esta referencia de los conceptos, nos parece que el proyecto de reglamento ha intentado agrupar distintos conceptos que guardan algún nivel de relación; sin embargo, consideramos que no son exactamente equiparables. El *big data*, como venimos comentando, refiere a un enorme conjunto de datos, mientras que el IoT hace referencia a una red física de objetos inteligentes “que están integrados con sensores, *software* y conectividad de red integrados, lo que les permite recopilar y compartir datos” (IBM, s.f.).

Por otro lado, como se ha podido observar en este breve recuento histórico, no solo ha tomado protagonismo el dato, sino que desde la obra de Turing se han desarrollado considerablemente las capacidades de la IA, siendo cada intento más sorprendente que el anterior.

Una forma de poder trazar el desarrollo antes comentado sobre la IA es a través de la siguiente gráfica, donde se visualiza cómo es que distintas capacidades humanas (comprensión lectora, reconocimiento de imágenes, entendimiento del lenguaje, razonamiento predictivo, reconocimiento de escritura y reconocimiento de voz) se han visto considerablemente superadas desde el año 2015 por sistemas de IA:



Nota. Adaptado de Plotting Progress in AI, por Kiela et al. (2023) con un procesamiento menor de Our World in Data.

El reconocer que tecnologías como las estudiadas en el presente trabajo se han venido desarrollando con rapidez en un contexto donde los datos han alcanzado una relevancia global y han alcanzado un potencial como el señalado en la gráfica demuestra la importancia de estudiarlas, ya que, si bien pueden ser beneficiosas por sus aplicaciones en distintas áreas, también pueden ser riesgosas en términos de transparencia.

1.1.2 Definición de inteligencia artificial

Para los fines del presente trabajo resulta imprescindible traer a colación un concepto claro de IA. Para ello, antes de presentar el concepto de IA que emplearemos en nuestro análisis, serán explorados distintos conceptos otorgados por la literatura más especializada sobre el tema.

a) Desde su equiparación con la inteligencia humana

Como ha sido explicado anteriormente, el primer hito en la elaboración del concepto de IA se remonta al año 1950, cuando Alan Turing sugiere que las máquinas pueden ser consideradas como pensantes si imitan el comportamiento humano llegando al punto de ser difícil distinguirlos de los seres humanos (1950, pp. 433-434).

Desde ese momento en que Turing propuso su *imitation game*, se pensaba que una IA era aquella máquina sobre la que se lograba percibir la capacidad de pensar o racionalizar humana, siendo este un entendimiento que persiste hasta nuestros días. Un ejemplo de ello es la posición de los autores Russell y Norvig, quienes optan por una definición de IA basada en el “agente racional”, el cual actúa de manera que busca el mejor resultado posible o esperado (2022, p. 22).

Sin perjuicio de ello, hasta el día de hoy la doctrina sigue proponiendo distintas definiciones para la IA, enfocándose en distintos aspectos y perspectivas, considerando cada vez más su uso dentro de las actividades usuales y diarias de las personas. Gran parte de la doctrina, al momento de conceptualizar la IA, coincide en considerarla en función de las similitudes que presenta respecto a la inteligencia humana.

En dicha línea de pensamiento, Gonzáles y Díaz recientemente han considerado a la IA como una disciplina que, a través de computadoras y otras máquinas, pretende replicar la inteligencia natural de los seres vivos (2023, p. 40). Por su parte, Rebollo propone también una conceptualización de la IA a partir de la emulación de la inteligencia humana. Dicha emulación no consistiría en replicar el proceso biológico y físico que realizan las mentes humanas al pensar, sino que se buscaría imitar la independencia con la que la mente humana puede procesar los datos, sin necesidad de que se establezca un flujo específico para ello. Y es que la IA es plasmada como una combinación de algoritmos

que buscan crear máquinas, *softwares* u otros que pretendan emular la capacidad humana de pensar (2023, p. 19).

También, en la misma línea de pensamiento, Moreno señala que la IA es la *inteligencia exhibida por máquinas*: con la programación de algoritmos, y proporcionando la data idónea, una máquina puede ser catalogada como inteligente al ser capaz de pensar como un humano, o incluso superar el rendimiento de una persona promedio (2022, p. 68).

Para Morillas, la definición de IA pasa por el uso de diversas disciplinas para emular procesos cognitivos humanos:

“No existe un concepto pacífico respecto a lo que debe entenderse por inteligencia artificial y así son muchas las disquisiciones existentes al respecto, partiendo incluso de la propia definición de inteligencia. Ahora bien, lo que no debe negarse, en pleno año 2022, es la existencia, junto con la inteligencia de los seres humanos, de otros sistemas de inteligencia, en este caso particular artificial, entendiendo por semejante adjetivación el origen no intrínsecamente humano del citado pensamiento; esto es, la capacidad propia de pensamiento y su ejecución de una máquina, la cual, por un complejo procedimiento de algoritmos, puede realizar tareas similares a las personas, optimizando sus resultados tanto en un prisma temporal como cualitativo, constituyendo ese su principal objetivo fundacional, que no debe ser otro que realizar de manera satisfactoria determinadas actividades de forma precisa, acortando los plazos tradicionales y sin necesidad de una posterior intervención humana” (Morillas, 2023, p. 87).

No obstante, el concepto de IA no ha sido únicamente analizado en función de la emulación de la inteligencia humana, sino que se ha buscado dotar de contenido a estas definiciones desde diversas perspectivas que se analizarán a continuación, a fin de presentar una definición compleja que pueda ser ubicada en el contexto de la utilización de sistemas de IA por parte de la Administración Pública.

b) Desde su capacidad de solución a problemas complejos

Otro sector de la doctrina, más allá de enfocarse en la emulación de la inteligencia humana como característica principal, considera a la IA en función a su capacidad para solucionar problemas complejos a través de algoritmos y el procesamiento de datos.

En dicho tenor, Herrera et al. conceptualizan a la IA en función de su finalidad de simular procesos de inteligencia humana mediante algoritmos para resolver problemas complejos (2022, p. 51). En tal sentido, la complejidad de los problemas que la IA tiene por resolver se encuentra íntimamente vinculada a los datos que procesa y el resultado que de ellos se desprende. Es por ello que la IA se vincula con otros conceptos como la data o los algoritmos.

En la misma línea, la doctrina resalta la utilidad de la IA en situaciones de mayor complejidad, en tanto el ámbito de la IA no se enfoca únicamente en simular la inteligencia humana, sino que tiene también como finalidad el resolver problemas con cierto grado de complejidad (Serna, Serna & Acevedo, 2017, p. 167).

Esta perspectiva resalta porque no destina a la IA a gestionar tareas que pudieran considerarse sencillas e intrascendentes, sino que la reserva para las tareas más complejas, que pueden tener implicancias considerables.

c) Desde sus componentes o capacidades

Otro sector de la doctrina propone la definición de la IA en virtud de los componentes que se consideran fundamentales para poder catalogar una tecnología como IA.

En tal sentido, Herrera et al. también proponen que el concepto de IA se compone de tres elementos: sistemas informáticos, los cuales principalmente son algoritmos que

direccionan el procesamiento de los datos detrás de *softwares* y otras tecnologías; datos, los cuales se gestionan de manera automática y masiva para generar conocimiento y soluciones valiosas; y objetivo, el cual determina el sentido de la gestión de los datos a través de algoritmos (2022, p. 63).

Por su parte, Campos propone definir a la IA como cualquier sistema que pueda alcanzar objetivos específicos con cierto grado de autonomía, al manifestar un comportamiento inteligente y demostrar capacidad para analizar su entorno y realizar acciones con los fines preestablecidos (2022, p. 25). El mencionado concepto proporciona otro aspecto importante: las tecnologías que se encuentran dentro del concepto de IA persiguen objetivos específicos.

En concordancia con todo lo anteriormente mencionado, resalta el contexto que propone Palma respecto a los sistemas de IA, los cuales cataloga como:

“aquellos sistemas diseñados por seres humanos que, dado un objetivo complejo, actúan en la dimensión física o digital mediante la percepción de su entorno a través de la obtención de datos, y decidiendo la acción o acciones óptimas que deben llevar a cabo para lograr el objetivo establecido, y decidiendo la acción o acciones óptimas que deben llevar a cabo para lograr el objetivo establecido”.

(2022, p. 31)

La International Organization for Standardization (2022) ha señalado que se concibe a la IA como un campo científico y técnico que se orienta a los sistemas destinados a generar soluciones como contenido, pronósticos, recomendaciones, entre otros, basados en la data que se le proporciona.

Por su parte, la Comisión Europea introduce como componente fundamental de la IA el concepto de autonomía de la toma de decisiones, pues la define en función de sus características de sistema diseñado que actúa con cierto grado de autonomía para el procesamiento de data y la toma de decisiones para cumplir con determinados objetivos (2023, p. 16).

d) Desde el punto de vista tecnológico

Por otro lado, la doctrina ha conceptualizado el concepto de IA desde una perspectiva tecnológica, considerándola como compuesta por una serie de tecnologías diversas, ya sea por la presencia de algoritmos tecnológicos, o por la capacidad de procesamiento de *big data*.

En dicho tenor, Palma considera que la IA, lejos de englobar un único sistema o algoritmo, se compone de múltiples tecnologías que incluyen la inferencia automática, el reconocimiento de patrones, *machine learning*, y otras varias (2022, p. 39). La diferencia entre las tecnologías de IA radica en la complejidad de su programación, en la forma que tienen de procesar los datos o de generar conocimiento.

Por su parte, también se considera a la IA en función a su capacidad de procesamiento y gestión de datos:

“La IA conecta los datos a los procesos de acción y de aprendizaje en la medida que por un lado los utiliza para ofrecer resultados para decidir y para actuar ante una problemática considerada, pero al mismo tiempo utiliza dichos datos y resultados para su propio proceso de aprendizaje y mejora. Por ello resulta fundamental combinar datos, tecnología y procesos para generar sistemas integrados de IA en la organización” (Salvador, 2021, p.23).

Las definiciones de IA que parten del punto tecnológico hacen énfasis en aspectos técnicos de estas tecnologías y la complejidad de estas construcciones algorítmicas.

e) Desde el punto de vista jurídico

Resulta complejo definir en términos jurídicos un concepto que se origina en una disciplina como las ciencias de la computación. No obstante, en el presente acápite se considerarán las definiciones recogidas en cuerpos normativos, sean estos vinculantes o no. Un concepto que permite incluir varios de los enfoques mencionados en los acápites anteriores es aquel que señala que la IA es la tecnología que cuenta con la capacidad o habilidad de una máquina para funcionar con un propósito en específico, a través del pensamiento racional, e interactuar con el ambiente de manera eficaz, como se supone que idealmente deben hacer los humanos. No obstante, dado que el pensamiento y comportamiento humano no siempre implican racionalidad, las máquinas pueden pensar y actuar como humanos o pensar racionalmente (Deville et al., 2021, p. 2).

La complejidad que origina la IA al momento de pensar en su regulación es conocida, puesto que cuando se plantea la pregunta sobre ¿qué se debería regular? se sabe que no existe una definición universal. Además, la IA presenta diferentes problemáticas dependiendo de quién, en qué manera y por qué la use. Un ejemplo de ello es referirnos a los problemas sobre IA en armas automáticas y los problemas que presenta la IA cuando es usada en *trading*, publicidad personalizada o vehículos autónomos (Fierens et al., 2021, p. 57).

Ahora bien, una buena referencia de un concepto jurídico para la IA es el abordado desde la Unión Europea (UE), siendo esta última una fuente de referencias para lo que ha sido la regulación en el Perú, como será abordado posteriormente.

La Propuesta de Reglamento del Parlamento Europeo y del Consejo (en adelante, “Propuesta de Reglamento de IA de la UE”) tiene como objetivo establecer normas uniformes en el ámbito de IA, por lo que se modifican ciertas disposiciones legislativas de la UE. En dicho documento se reconoce la complejidad de definir con precisión la IA, por lo que sugiere que este último concepto se complemente con características esenciales funcionales del *software*.

En dicho sentido, el documento destaca la capacidad que tiene la IA para generar distintos tipos de información con base en objetivos que son definidos por humanos, tales como predicciones, recomendaciones, decisiones u otra información. Esta información de salida puede influir en el entorno digital o físico en el que opera el sistema. Asimismo, se enfatiza la necesidad de mantener actualizadas las estrategias y técnicas utilizadas en el desarrollo de la IA, considerando la evolución tecnológica y las exigencias del mercado (Considerando 6) (Berning, 2023, p.167).

En la Propuesta de Reglamento de IA de la UE se define el término *sistema de inteligencia artificial* como *software* en cuyo desarrollo intervienen una o varias técnicas concebidas en el Anexo I⁷ y que, respecto a un conjunto específico de objetivos planteados, es capaz de producir información de salida, tales como contenidos, predicciones, recomendaciones o decisiones, que influyan en los entornos con los que interactúa.

⁷ El Anexo I de la Propuesta de Reglamento de IA de la UE señala lo siguiente:

“Anexo I: TÉCNICAS Y ESTRATEGIAS DE IA mencionados en el artículo 3º, punto 1

- a) Estrategias de aprendizaje automático, incluidos el aprendizaje supervisado, el no supervisado y el realizado por refuerzo, que emplean una amplia variedad de métodos, entre ellos el aprendizaje profundo.

Estrategias basadas en la lógica y el conocimiento, especialmente la representación del conocimiento, la programación (lógica) inductiva, las bases de conocimiento, los motores de inferencia y deducción, los sistemas expertos y de razonamiento (simbólico).

Estrategias estadísticas, estimación bayesiana, métodos de búsqueda y optimización.”

Adicionalmente a lo señalado, el Reglamento del Parlamento Europeo y del Consejo de 13 de junio de 2024 (en adelante, “Reglamento de IA de la UE”) presenta una definición de IA que amerita ser abordada. Por el mencionado reglamento se establecen disposiciones normativas en materia de IA y se modifican los reglamentos (CE) N° 300/2008, (UE) N° 167/2013, (UE) N° 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144; y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828; y se define a los sistemas de IA en su artículo 3°, inciso 1) de la siguiente forma:

(...) un sistema basado en una máquina que está diseñado para funcionar con distintos niveles de autonomía y que puede mostrar capacidad de adaptación tras el despliegue, y que, para objetivos explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar resultados de salida, como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos físicos o virtuales (...)

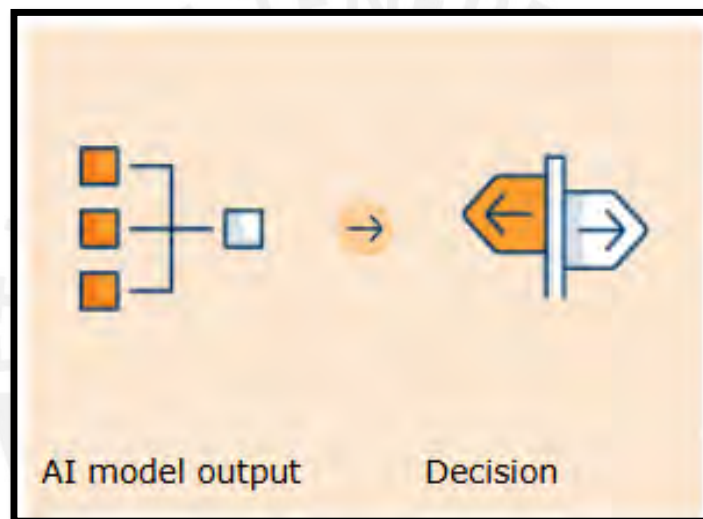
Como se puede apreciar de la cita, la versión final del Reglamento de IA de la UE apostó por definir a las tecnologías de IA desde una perspectiva que enfatiza su funcionamiento; esto es, explicando que todo sistema de IA funciona con las nociones de “información de entrada” o *input* y “resultados de salida” u *output*, pudiendo ser estas últimas predicciones, decisiones, entre otras.

La idea de *output* es explicada de manera sencilla por la Information Commissioner’s Office (ICO) y The Alan Turing Institute (2022), según la cual, el *output* de un modelo IA varía dependiendo del tipo de modelo que se use y su propósito. Señala que generalmente hay tres tipos de *output*: una predicción, una recomendación o una clasificación.

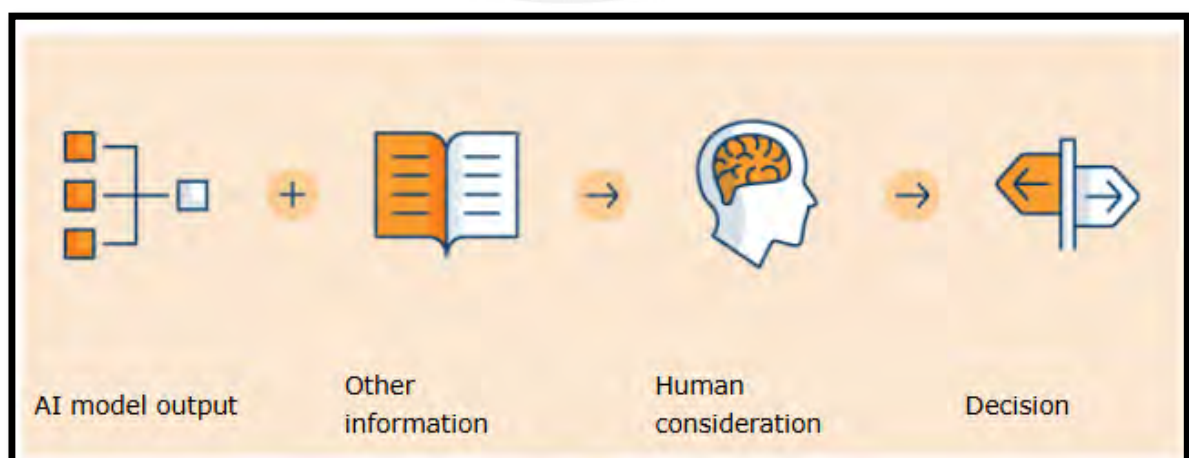
Además, dichos documentos indican que en algunos casos un sistema de IA puede estar totalmente automatizado en su despliegue cuando su *output* y cualquier acción que se tome como un resultado (decisión) sean implementados sin ningún tipo de interacción

humana o vigilancia. Menciona que esa situación cambia cuando, dentro de ese proceso de generar el resultado, el ser humano aparece y usa ese *output* generado por el sistema IA y, junto con otra información, toma una decisión. No obstante ello, cuando se hace referencia a una decisión generada por la IA, nos referimos indistintamente a aquella en la que participa o no un humano. Veamos algunos gráficos que resumen estos procedimientos:

Decisión automatizada con sistema IA



Decisión con sistema IA y participación humana



Nota. Adaptado de Explaining Decisions made with AI, por ICO y The Alan Turing Institute (2022).

Conforme a lo anterior, queda claro que la definición del Reglamento de IA de la UE ha apostado por una definición que “capta” la esencia de los sistemas de IA, la cual se circunscribe a las nociones de *input* y *output*. Es esta la definición jurídica sobre la que este trabajo analizará los siguientes capítulos.

Por otro lado, se trata de una definición acertada a nivel regulatorio, ya que la Propuesta de Reglamento de IA de la UE se remitía a un anexo al momento de definir los sistemas de IA, haciendo una referencia explícita a las tecnologías o técnicas con las que la IA logra sus objetivos.

Finalmente, resulta meritorio mencionar la definición que el ordenamiento jurídico peruano le ha atribuido a la IA, en las normas específicas, como la Ley de IA peruana, definiéndola como una tecnología emergente con el potencial de mejorar el bienestar de la sociedad, contribuir a una actividad sostenible, aumentar la innovación, productividad, entre otros.

Sobre esta definición, el Proyecto de Reglamento agrega que la IA también es considerada como una disciplina científica, por la cual se busca el desarrollo de sistemas, algoritmos o *softwares* en general que realicen actividades y procesen información conforme lo realizaría una mente humana, considerando premisas como el aprendizaje y el razonamiento lógico.

Nótese que estas definiciones no son similares a la que acabamos de comentar del Reglamento de IA de la UE. Ello es así porque la normativa específica peruana recoge un significado de IA que atiende a sus capacidades. Así, tanto la Ley de IA peruana como su

Proyecto de Reglamento, señalan que un sistema basado en IA es uno que tiene como finalidades establecidas la toma de decisiones, la generación de predicciones o de recomendaciones, con diferentes niveles de autonomía; por lo cual influye en entornos materiales y virtuales.

En cuanto a las definiciones de sistemas basados en IA que ha emitido el legislador peruano, nos parece que hace falta una precisión en la sección de definiciones como la anotada por el Reglamento de IA de la UE, haciendo referencia a los datos de entrada y de salida, ya que, como venimos comentando, es esa la esencia del concepto.

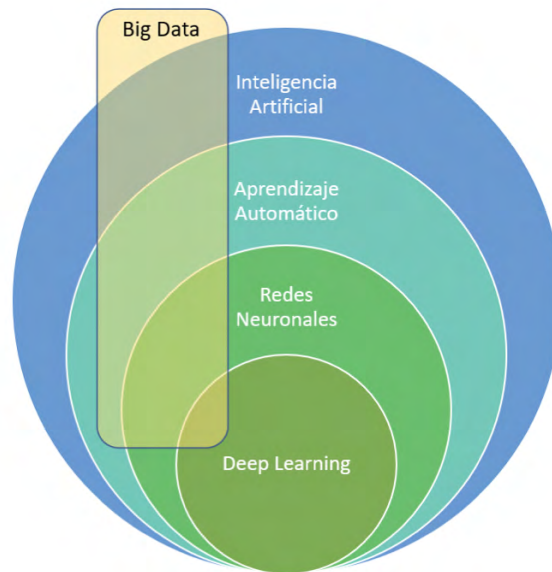
Señalamos lo anterior sin perjuicio que en el literal d) del artículo 15° del Proyecto de Reglamento de la Ley de IA peruana hay una referencia a los datos de entrada en las obligaciones específicas para aplicaciones de IA de riesgo alto.

Enfatizamos que consideramos preferible una definición general que pueda adaptarse a las distintas técnicas que existen a la fecha relacionadas a la IA, las cuales serán explicadas en el siguiente acápite.

1.2 Conceptos asociados a la inteligencia artificial

El presente apartado tiene como finalidad abordar los conceptos que se encuentran vinculados a la IA.

Los autores González y Díaz definen a las tecnologías de IA como una disciplina que pretende replicar la inteligencia natural, sea esta de seres humanos o seres vivos, mediante computadoras (2023, p. 39); y emplean el gráfico que se muestra a continuación para poder describir que la IA integra diversos “campos de estudio”:



Nota. Extraído de ¿Qué es la Inteligencia Artificial?, Gonzales y Díaz (2023, p. 39).

En tal sentido, corresponde ahora analizar conceptos aludidos en el gráfico, así como el aprendizaje automático o *machine learning* y las redes neuronales, entre otros, a fin de poder complementar el concepto de IA para los fines del presente trabajo.

1.2.1 *Machine Learning*

El *Machine Learning* (en adelante, ML) es definido por la doctrina como un área de estudio, una disciplina, por la cual se pretende emular la capacidad de aprender que es característica de la inteligencia natural; y es, por tanto, una disciplina que forma parte de la IA, la misma que pretende simular cualquier tipo de inteligencia humana o de cualquier ser vivo (Gonzales y Díaz, 2023, p. 40).

En tal contexto, aunque el ML forma parte de la IA según lo descrito en el apartado anterior, posee una característica definitiva: se centra en la capacidad que los algoritmos presentan para aprender de la data que le es proporcionada. Las tecnologías basadas en ML procesan estos datos y extraen conocimiento de ellos, sin necesidad de programación exhaustiva para cada aspecto de análisis.

El ML es una de las técnicas enmarcadas en la IA más utilizadas para el procesamiento y análisis automático de datos. Al respecto, Herrera et al. señalan lo siguiente:

“El aprendizaje automático es un área central de la IA actual. Se refiere al proceso de aprendizaje de sistemas de IA que se realiza a partir de los datos. En esencia, las técnicas de aprendizaje automático construyen o aprenden un modelo matemático o simbólico que representa un sistema real a partir de un conjunto de datos de entrenamiento. (2022, p. 64).

Por lo tanto, el concepto de ML imita la inteligencia humana en su capacidad para aprender, lo que permite trabajar con grandes volúmenes de datos y generar resultados de manera autónoma. Esto significa que los *softwares*, máquinas y sistemas que emplean ML adquieren la habilidad de aprender a partir de la información sin la necesidad de ser programados en cada nivel del proceso (Palma, 2022, p. 40).

Como técnica dentro de la IA, el ML abarca una variedad de conceptos que se analizarán en este trabajo. Una de las técnicas más destacadas dentro del ML son las redes neuronales, las cuales serán objeto de análisis en el siguiente apartado.

El Reglamento de IA de la UE tiene una pequeña mención sobre el aprendizaje automático en su considerando 12, indicando que una de las técnicas que permiten la inferencia al construir un sistema de IA son las estrategias de aprendizaje automático que aprenden de los datos cómo alcanzar determinados objetivos y estrategias basadas en procesos lógicos, y de conocimiento, las cuales se infieren de premisas codificadas o de representación simbólica de finalidades que se deben resolver.

1.2.2 Redes neuronales

Las redes neuronales representan una familia dentro del ML. En tal sentido, emulan el aprendizaje del ser humano mediante estructuras conceptuales similares a las neuronas y el cerebro humano, y es a raíz de ello que reciben el particular nombre de “redes neuronales” (Gonzales y Díaz, 2023, p. 41).

Estas redes neuronales están inspiradas en el funcionamiento del cerebro humano, que desarrolla el proceso de sinapsis, por el que las neuronas se transmiten señales entre ellas. Así, en las redes neuronales, cada neurona recibe una serie de entradas (valores numéricos) y determina su salida sumando las entradas (ponderando cada una de ellas con un peso) y aplicando en el resultado una función de activación (Herrera et al., 2022, p. 66).

La doctrina ha descrito a las redes neuronales como un modelo computacional de aprendizaje automático, mediante el cual los datos ingresados al sistema de IA recorren la red neuronal y se procesan a través de operaciones matemáticas establecidas, generando un resultado de salida. Este tipo de sistemas de IA realiza el señalado procesamiento mediante un conjunto de unidades a las que se les otorga el nombre de neuronas o neuronas artificiales, las mismas que se encuentran interconectadas y distribuidas en capas, emulando una red (González y Díaz, 2023, p. 58).

En diversos países, las tecnologías de redes neuronales se emplean por las oficinas de gestión documentaria, pues estos algoritmos pueden clasificar los documentos en función a su contenido, y pueden organizarlos de forma automática por categorías, lo cual optimiza la eficiencia de los despachos de las administraciones públicas que en sus actividades manejan grandes cantidades de documentos.

1.2.3 *Deep learning*

El *Deep Learning* (en adelante, DL) o aprendizaje profundo, se refiere a los sistemas que presentan una construcción compleja de capas ocultas de redes neuronales, los mismos que requieren una gran cantidad de datos de entrada para su entrenamiento, así como altos recursos computacionales (González y Díaz, 2023, p. 42).

En este sentido, el DL se considera una manifestación avanzada de las redes neuronales; siendo que “las redes neuronales profundas son el elemento fundamental de lo que se conoce como aprendizaje profundo” (Herrera et al., 2022, p. 67).

Aunque las tecnologías de DL están incluidas dentro del ML, tienen la particularidad de proporcionar resultados altamente precisos, pero que resultan bastante difíciles de explicar. Así, los sistemas de DL resultan complicados de interpretar, al punto de ser considerados como cajas negras o *black boxes* (Gonzales y Díaz, 2023, pp. 64-65).

Por lo tanto, los algoritmos basados en el DL no solo logran un aprendizaje más detallado y complejo, sino que también se caracterizan por la alta complejidad de su funcionamiento y la gran cantidad de datos que procesan.

Los conceptos anteriormente descritos ofrecen una visión más sofisticada de cómo se aplican las técnicas de IA en la práctica. Sin embargo, la IA suele concretarse a través de algoritmos, los cuales serán objeto de análisis en el siguiente apartado.

1.2.4 Algoritmos

Los algoritmos desempeñan un papel de suma relevancia en la IA, ya que son el medio a través del cual las técnicas de IA, previamente desarrolladas, se materializan en programas, *softwares*, que pueden ser aprovechados en diversos contextos. Por tanto, resulta esencial realizar un análisis de los conceptos propuestos por la doctrina, con el objetivo de complementar lo expuesto acerca de las técnicas de IA.

En tal sentido, Palma define el algoritmo como un sistema compuesto de operaciones lógicas e instrucciones, las cuales se entienden como reglas formales, que producen un resultado en virtud de la información de entrada proporcionada al sistema. En dicho sentido, se trata de un algoritmo que emite un resultado ante un *input*, al seguir un conjunto de normas previamente establecidas (2022, p. 37).

En líneas generales, se puede afirmar que los algoritmos pueden llegar a predecir nuestras necesidades y próximas acciones, pues analizan nuestras conductas y pueden advertir nuestros intereses (Rebollo, 2023, p. 27).

Resulta interesante que la normativa peruana relacionada a con la IA haya apostado por una definición de algoritmo. Así, tanto la Ley de IA peruana como su Proyecto de Reglamento lo definen como una secuencia de reglas y/o pasos para resolver un problema previamente establecido.

Como se puede apreciar, es una definición que va en línea con lo señalado por Palma. Sin embargo, desconocemos la necesidad de hacer una definición específica de esta palabra, ya que como hemos anotado anteriormente, es preferible tener una sola definición genérica que englobe distintos conceptos, como lo es la definición de sistema basado en IA del Reglamento de IA de la UE.

Por otro lado, aunque existen diversas clasificaciones doctrinales para los algoritmos, en este trabajo se abordarán específicamente los algoritmos de aprendizaje supervisado y los de aprendizaje no supervisado.

a) Aprendizaje supervisado

La doctrina define al aprendizaje supervisado como el escenario en el que, al momento de entrenar los datos aportados, se establece también la solución que se desea, a fin de

que el algoritmo pueda aprender. En este escenario, se dice que los datos se encuentran etiquetados (Gonzales y Díaz, 2023, p. 49).

En tal sentido, el algoritmo de aprendizaje supervisado cuenta con una solución o etiqueta para los datos que serán analizados, y a través de ello es que brinda soluciones útiles según la necesidad para la que fueron programados. “El aprendizaje supervisado, en esencia, trata de entrenar un modelo a partir de un conjunto de datos etiquetados, permitiendo así identificar patrones y realizar predicciones sobre nuevos datos” (Kelleher, 2022, p. 88).

La doctrina define al aprendizaje supervisado en función al conglomerado de datos de entrenamiento, en el que cada dato de entrada es acompañado de una salida esperada; lo cual permite a los algoritmos predecir salidas futuras (Russell y Norvig, 2022, p. 315). Estos algoritmos son hoy en día fundamentales en diversas aplicaciones modernas, para predecir datos no vistos (Aggarwal, 2023, p. 162). "En el aprendizaje supervisado, el algoritmo recibe datos etiquetados como entrada y ajusta su modelo en base a los errores cometidos en las predicciones, mejorando su precisión en futuras predicciones" (Alpaydin, 2023, p. 190).

Un ejemplo clásico de aprendizaje supervisado es el de los árboles de decisión, el cual constituye un árbol de reglas de decisión, en el cual se interconectan decisiones binarias a fin de obtener un resultado (González y Díaz, 2022, p. 51).

El Reglamento de IA de la UE hace una mención al aprendizaje supervisado cuando hace referencia a la documentación técnica de los sistemas de IA de alto riesgo. Esta disposición del artículo 11° de dicho Reglamento hace referencia al Anexo IV, donde se contemplan los elementos mínimos e indispensables que debe tener la documentación técnica de estos sistemas de riesgo alto.

Así, en el Anexo IV, cuando se solicita el detalle de los elementos del sistema de IA y el proceso seguido en su desarrollo, se indica que deben encontrarse incluidos los procedimientos de etiquetado, colocando como ejemplo a aquellos empleados para el aprendizaje supervisado.

b) Aprendizaje no supervisado

La característica principal de los algoritmos de aprendizaje no supervisado es que sus datos de entrenamiento no se encuentran etiquetados.

Así lo señalan Gonzáles y Díaz: en el caso del aprendizaje no supervisado, resalta que el sistema aprende de la información de entrada proporcionada sin que se le haya asignado una respuesta para el resultado de salida; así, el sistema aprende de la información sin contar con un resultado esperado, por lo que se dice que los datos de entrenamiento no están etiquetados (2023, p. 49). A similar conclusión llegan Russel y Norvig, indicando que el sistema de aprendizaje no supervisado busca patrones o agrupamientos deducidos de características inherentes a los datos, sin la necesidad de etiquetas (Russell y Norvig, 2022, p. 368).

Es esta característica, como apunta Aggarwal, la que hace que estos algoritmos se diferencien de los de aprendizaje supervisado, y, además, produce que el propio sistema deba explorar e identificar patrones en los datos, destacando técnicas como el *clustering* (Aggarwal, 2023, p. 145).

Kelleher también destaca al *clustering*, señalando que se trata de una de las técnicas más comunes en este tipo de algoritmos: “no se dispone de datos etiquetados, permitiendo a los algoritmos extraer estructuras y agrupamientos de manera autónoma, siendo el *clustering* una de las técnicas más comunes” (Kelleher, 2022, p. 102). Lo mismo hace

Alpaydin, indicando que dicho tipo de algoritmos prescinde de etiquetas predefinidas y que es el *clustering* la técnica más difundida en el agrupamiento de datos en función de sus similitudes, sin necesidad de una solución previamente definida (Alpaydin, 2023, p. 210).

En este contexto, los algoritmos de aprendizaje no supervisado operan con mayor autonomía, ya que analizan los datos sin intervención previa en cuanto a etiquetado y clasificación. A diferencia de los algoritmos de aprendizaje supervisado, en este caso el programador no proporcionará las respuestas esperadas, lo que permite al algoritmo descubrir patrones o estructuras ocultas en los datos y producir soluciones novedosas que no han sido predefinidas.

Sobre los algoritmos de aprendizaje no supervisado el Reglamento de IA de la UE solo hace una mención mínima en el considerando 97, indicando que es un método de entrenamiento de sistemas de IA.

Dentro de los algoritmos de aprendizaje no supervisado, resaltan dos grandes técnicas que es útil identificar a fin de ilustrar el funcionamiento de este tipo de algoritmos: el *clustering* (que ya habíamos mencionado) y el *biclustering*.

i) Clustering

El aprendizaje basado en *clustering* se basa en sistemas o algoritmos que buscan agrupar datos en conjuntos determinados, extrayendo información de los elementos y buscando similitudes entre estos; así, se dice que se forman clústers de los cuales se busca extraer información valiosa (González y Díaz, 2023, p. 54).

Así, los algoritmos de *clustering* permiten dividir datos no etiquetados en grupos homogéneos por sus características. Estos agrupamientos ayudan a entender la estructura

interna de los datos que se procesan, identificando relaciones que no son evidentes a simple vista (Alpaydin, 2023, p. 220).

Conforme a lo anterior, el *clustering* se refiere al proceso de agrupar datos, que no se encuentran etiquetados, en función de sus características o similitudes. En el marco de los algoritmos de aprendizaje no supervisado, el algoritmo agrupará datos en grupos o *clústers* en función a las relaciones inherentes entre ellos, sin que los datos hayan sido etiquetados previamente.

ii) **Biclustering**

El *biclustering* permite analizar de forma simultánea filas y columnas en los datos, identificando patrones de forma más completa que las técnicas de *clustering* tradicionales.

Así, las técnicas de *biclustering* se basan en procesar la información de manera bidireccional, agrupando los datos en filas y columnas, formando *biclusters* en función a la similitud que pueda presentarse en cada uno de los subconjuntos, dependiendo de sus variables y características (González y Díaz, 2023, pp. 56-57).

De este modo, el *biclustering* se refiere al agrupamiento simultáneo de datos en filas y columnas, para identificar patrones más complejos en comparación con el *clustering* convencional. Por ello, los *biclusters* formados son más numerosos y permiten un análisis bidimensional de los datos, ofreciendo una visión más detallada de las relaciones presentes.

Como se ha podido apreciar, los algoritmos son las herramientas a través de las cuales la IA obtiene sus objetivos.

1.2.5 **Big data**

El término *big data* ha adquirido relevancia en los últimos años, y su relación con la IA es evidente, ya que los datos son el insumo esencial que ésta utiliza para generar resultados valiosos. Algunos autores, como Carbonero, han identificado al *big data* como *la clave*: “la clave de todo sistema de IA es el *big data*. A más cantidad y calidad en los datos objeto de tratamiento, más precisión y fiabilidad en los sistemas de IA” (2022, p. 131).

La doctrina ha subrayado la relevancia de los datos en el funcionamiento de los algoritmos, señalando que los algoritmos derivados del aprendizaje automático son imprescindibles en esta época del *big data*, donde predominan los algoritmos de aprendizaje automático distribuidos para el procesamiento masivo de datos, los cuales son capaces de convertir grandes cantidades de datos en conocimiento (Herrera et al., 2022, p. 61).

En tal sentido, se puede definir al *big data* como una serie de técnicas que procesan grandes volúmenes de datos, en distintos formatos y estructuras, a alta velocidad (Palma, 2022, p. 44). Por su parte, González y Díaz coinciden con esta visión, al afirmar que el Big Data requiere de programas o ampliaciones informáticas novedosas que procesen conjuntos de datos, grandes y complejos, de forma adecuada (2023, p.42).

Por su parte, Palma amplía el concepto de *big data* al incluir elementos adicionales que deben considerarse: volumen, velocidad, veracidad y valor. El señalado autor considera que el *big data* analiza un conjunto significativo y voluminoso de datos, a gran velocidad (pese a que los datos pueden presentarse no siempre en el mismo formato), logrando detectar información relevante, de la que rescata conocimiento valioso (2022, p.44).

La trascendencia del *big data* para el funcionamiento de los algoritmos es evidente, en tanto los datos hacen posible la creación de modelos más precisos y complejos, alimentan

los algoritmos de aprendizaje, y constituyen elementos que construyen la información, y la trasladan (Herrera et al., 2022, p. 63).

Asimismo, nos reiteramos en nuestros comentarios efectuados en el acápite vinculado a “antecedentes históricos” donde mencionamos que la normativa peruana no debió equiparar su definición con la del IoT.

Como se ha podido apreciar, tanto el Reglamento de IA de la UE como la normativa peruana sobre IA hacen referencia al concepto marco sobre IA (sistemas basados en IA), no obstante, son pocas las referencias jurídicas específicas que se tienen en estos cuerpos normativos sobre los conceptos asociados a IA, tales como el aprendizaje automatizado, aprendizaje supervisado, entre otros.

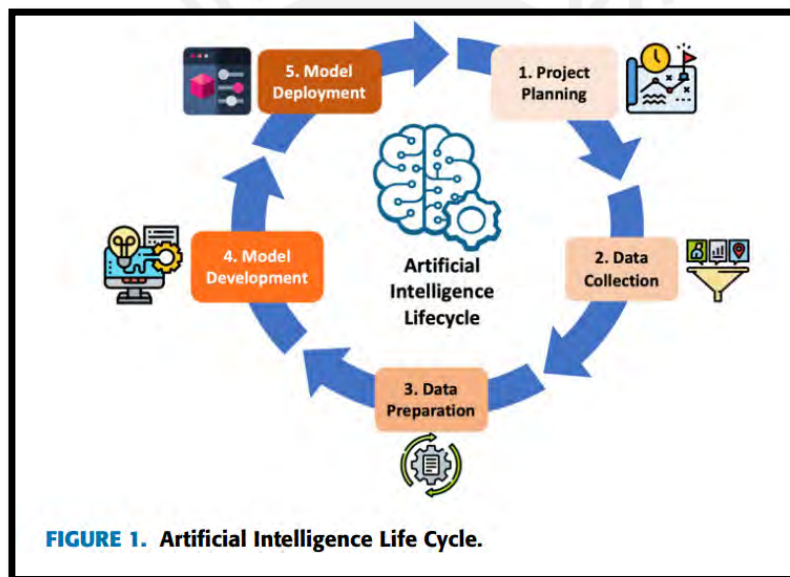
Al respecto, consideramos que la falta de mención hacia estos conceptos no es peligrosa ni involucra un mayor riesgo siempre y cuando el concepto “marco” -que viene a ser el de sistema de IA- se encuentre correctamente definido. A nuestro parecer, el concepto que existe en el Reglamento de IA de la UE es lo suficientemente preciso sobre el funcionamiento de las técnicas de la IA. Sin embargo, como ya hemos anotado anteriormente, la normativa peruana deja espacios que no son claros y que deben ser precisados para poder incorporar la lógica de datos de entrada y salida con mayor claridad.

1.3 Ciclo de vida de la IA

El ciclo de vida de la IA se compone de distintas etapas o fases, diferenciadas unas de las otras; y ha sido objeto de estudio por parte de la doctrina, en tanto cada fase del ciclo de vida de la IA resulta crucial a fin de conseguir que las nuevas tecnologías de IA no solo sean útiles y funcionales, sino que cumplan con los estándares básicos para minimizar riesgos y brindar garantías básicas a sus usuarios.

Si bien no existe consenso respecto al número de fases que conforman el ciclo de vida de la IA, el presente trabajo considerará tres fases específicas, que, a su vez, se subdividen en otras fases o hitos cruciales en el desarrollo de tecnologías de IA.

Sin perjuicio de ello, es meritorio mencionar las principales posturas respecto a las etapas del ciclo bajo comentario. Así, para autores como Shahriar, Allana, Hazratifard y Dara, el ciclo de vida de la IA se divide en cinco fases: (i) el planeamiento del proyecto, (ii) la recolección de data, (iii) la preparación de la data, (iv) el desarrollo del modelo, y (v) el despliegue del modelo (2023, p. 61832).



Fuente: Survey of Privacy Risks and Mitigation Strategies in the AI Life Cycle, Shahriar et al. (2023, p. 61831).

Por otro lado, De Silva y Alahakoon consideran tres grandes fases en el ciclo de vida de la IA, las cuales, a su vez, se componen de otras subfases específicas: (i) diseño, (ii) desarrollo y (iii) despliegue (Silva y Alahakoon, 2022, p. 4).



Fuente: An artificial intelligence life cycle: From conception to production, De Silva, D. & Alahakoon, D. (2022, p. 4)

Más allá de lo señalado, el presente apartado revisará el ciclo de vida de la IA dividido en tres grandes fases: (i) el diseño del proyecto; (ii) la evaluación y validación del proyecto; y (iii) el despliegue y mantenimiento del proyecto.

1.3.1 Diseño del proyecto

La fase o etapa de diseño en el ciclo de vida de la IA suele ser vinculada a aspectos más teóricos que prácticos; pues en esta etapa se formula e identifica el problema; se consulta literatura técnica sobre algoritmos, su aplicación y antecedentes de modelos preentrenados; y se prepara, explora y se adquiere data, de ser necesario (De Silva y Alahakoon, 2022, p. 4).

En esta fase del ciclo de vida, se pretende recolectar información útil sobre los objetivos del proyecto, lo cual resulta crucial, pues una incorrecta recopilación de información inevitablemente desencadenará en demoras en la implementación, y puede tener impacto negativo en el desempeño del sistema o algoritmo de IA.

Asimismo, la etapa de diseño comprende la definición de las características y el conjunto de habilidades que el sistema de IA debe tener. En esta fase, incluir especialistas en materias como privacidad, ética, y otros, puede facilitar el desarrollo del sistema, a fin de garantizar la calidad y la estricta adecuación a los estándares necesarios (Shahriar et. al., 2023, pp. 61832-61833).

Si bien el presente trabajo le ha otorgado a la abordada fase el nombre de “Diseño del proyecto”, la doctrina también le otorga el nombre de “Planeamiento del proyecto”, “Inicio del proyecto”, entre otros.

1.3.2 Evaluación y validación del proyecto

En la fase de evaluación del proyecto, se preprocesa y aumenta la data; se construye el modelo inicial de IA; se desarrollan los puntos de referencia; se evalúan las primeras métricas, y se definen aspectos a fin de asegurar la explicabilidad del modelo (De Silva y Alahakoon, 2022, p. 4).

En esta fase del ciclo de vida de la IA, la data es la gran protagonista, e incluso algunos autores consideran fases adicionales en el ciclo de vida de la IA destinadas a la recolección y preparación de la data.

La doctrina resalta la necesidad de recolectar data relevante y relacionada a los objetivos que persigue el sistema de IA en desarrollo, sin vulnerar parámetros éticos o principios de privacidad, con el fin de garantizar un correcto performance del programa. Dicha data

debe ser integrada a fin de fijar su utilidad para el proyecto (Shahriar et. Al., 2023, p. 618333).

Aunado a ello, se considera que en esta etapa se empieza a desarrollar el modelo de IA en base a un algoritmo que considere las aplicaciones y capacidades que debe presentar el sistema. Asimismo, se revisan las primeras métricas, las cuales persiguen que el desarrollo del algoritmo siga parámetros éticos y legales. De igual forma, es en la fase de evaluación que se busca darle al sistema de IA la explicabilidad que se requiere en aras de mantener la IA transparente (De Silva y Alahakoon, 2022, pp. 6-8).

Algunos autores consideran que las actividades comprendidas en la fase de evaluación y validación del proyecto pertenecen aún a una etapa de diseño del proyecto, mientras que otros la consideran por separado y aislada del desarrollo del proyecto. Asimismo, como ya se ha mencionado, también la doctrina ha propuesto separar los hitos relacionados a la data de las demás actividades mencionadas.

1.3.3 Despliegue y mantenimiento del proyecto

En el despliegue y mantenimiento del proyecto, se evalúan las métricas secundarias; se despliega el modelo de IA y se realizan las primeras revisiones de riesgos; se revisa el modelo post despliegue; se operacionalizan las conexiones del modelo empleando IA; se hiper automatizan los procesos y sistemas; y se monitorea y evalúa el desempeño (De Silva y Alahakoon, 2022, p. 4).

En el despliegue y mantenimiento del proyecto, se comprende la materialización del proyecto diseñado y evaluado previamente. En esta fase, se entrena al algoritmo con la data recolectada para este fin. Asimismo, no puede dejarse de lado en esta etapa del ciclo de vida de la IA la evaluación del modelo, con el fin de que los desarrolladores ajusten

los patrones para maximizar el performance del algoritmo (Shahriar et. Al., 2023, p. 618334).

Aunado a ello, esta etapa del ciclo de vida de la IA comprende subfases destinadas al monitoreo del algoritmo, en aras de incrementar la automatización del modelo, y resulta crucial a fin de determinar si el modelo contribuye a sus fines establecidos, y cómo lo hace (De Silva y Alahakoon, 2022, p. 9).

La doctrina suele coincidir con que la fase final del ciclo de vida de la IA corresponde al despliegue del modelo, en la cual se suele incluir las actividades de monitoreo del sistema de IA. No obstante, sí existen posturas que consideran una fase posterior, que implica el retiro del proyecto, que comprende la desactivación del sistema de IA desarrollado, y la gestión final de todos los impactos que el desarrollo del nuevo modelo de IA pueda implicar.

Por otro lado, existen menciones sobre el ciclo de vida de la IA tanto en el Reglamento de IA de la UE, como en la normativa peruana, tal como se explicará a continuación.

En relación al ciclo de vida de sistemas IA de alto riesgo, el Reglamento de IA de la UE indica lo siguiente:

- El sistema de gestión de riesgos implica un proceso continuo, planificado y ejecutado durante el ciclo del sistema IA de alto riesgo, que requerirá revisiones y actualizaciones sistemáticas periódicas (Artículo 9.2°).
- Los sistemas de IA de alto riesgo permitirán técnicamente el registro automático de acontecimientos (archivos de registro) a lo largo de todo el ciclo de vida del sistema (Artículo 12.1°).

- Los sistemas de IA de alto riesgo se diseñarán y desarrollarán de modo que alcancen un nivel adecuado de precisión, solidez y ciberseguridad y funcionen de manera uniforme en esos sentidos durante todo su ciclo de vida (Artículo 15.1°).
- La documentación técnica de los sistemas IA de alto riesgo (Artículo 11.1°) debe incluir una descripción de los cambios pertinentes realizados por el proveedor en el sistema a lo largo de su ciclo de vida (Anexo IV, numeral 6).

Por su parte, el Proyecto de Reglamento de la Ley de IA peruana indica sobre el ciclo de vida de IA lo siguiente:

- Principio rector de protección de la dignidad humana: En todas las fases del ciclo de vida de las aplicaciones de IA se promueven mecanismos para garantizar el respeto, protección de la dignidad y autonomía humana, el estado de derecho, los valores democráticos y centrados en el ser humano, así como los derechos humanos. (Artículo 4°, literal a).
- Principio rector de respeto de derechos de autor y derechos conexos: A lo largo del ciclo de vida de un sistema basado en IA se deben respetar los derechos de autor, tanto morales como patrimoniales, de los creadores respecto de sus obras originales de acuerdo con la normativa nacional y los tratados internacionales en la materia (Artículo 4°, literal h).
- Principio rector de seguridad, proporcionalidad y fiabilidad: Las aplicaciones de IA deben ser fiables y seguras durante todo su ciclo de vida para que, en condiciones de uso, según su finalidad prevista, funcionen adecuadamente y minimicen los riesgos del uso indebido u otras condiciones adversas vinculadas a dichos sistemas (Artículo 4°, literal i).

- Principio rector de sostenibilidad: En todas las fases del ciclo de vida de las aplicaciones de IA se tienen en cuenta y evalúan las repercusiones humanas, sociales, económicas, culturales y ambientales, en línea con la satisfacción de las políticas nacionales y los Objetivos de Desarrollo Sostenible (ODS) de la Organización de Naciones Unidas. También se debe tener en cuenta los potenciales impactos ambientales y promover soluciones que minimicen el consumo de energía y recursos (Artículo 4º, literal l).
- La entidades de la Administración Pública, en el ámbito de sus competencias, cumplen la obligación de asegurar el cumplimiento de los principios rectores en todas las fases del ciclo de vida de una aplicación de IA. (Artículo 8º, literal b).

Como se puede apreciar, el Reglamento de IA de la UE tiene respecto del ciclo de vida de la IA un conjunto de disposiciones específicas sobre los sistemas de IA de alto riesgo.

Ello no sucede en el caso de la normativa peruana, ya que según el texto propuesto en el Proyecto de Reglamento, se nota un tenor más inclinado hacia los principios que se desean proyectar en el ciclo de vida de la IA.

Lo anterior resulta en un problema más que una virtud, ya que los principios reseñados no muestran, al menos en su redacción actual, algún nivel de concreción en el resto del texto del Proyecto de Reglamento, ni tampoco se identifican mecanismos específicos que permitan su implementación efectiva. Para poder mejorar ello, el Proyecto de Reglamento debería proponer medidas concretas orientadas a la materialización de dichos principios, incluyendo referencias expresas al ciclo de vida de los sistemas de IA, especialmente en las disposiciones relativas a la gestión de riesgos asociados a su uso.

CAPÍTULO 2

POLÍTICAS PÚBLICAS, MARCO JURÍDICO E INSTITUCIONALIDAD DE LA INTELIGENCIA ARTIFICIAL EN LA ADMINISTRACIÓN PÚBLICA

En el presente acápite serán desarrollados los siguientes aspectos sobre los sistemas de IA. En primer lugar, se desarrollarán las políticas públicas sobre la IA en el mundo y en el Perú. En segundo lugar, se desarrollará el marco jurídico relacionado con la IA, tanto a nivel internacional como nacional. Para ello, serán respondidas las preguntas sobre si es necesaria una regulación de la IA, cuál es la noción de regulación aplicable a la IA y cuáles son los enfoques regulatorios. Finalmente, será desarrollada la institucionalidad de la IA en la Administración Pública.

2.1 Políticas públicas sobre inteligencia artificial

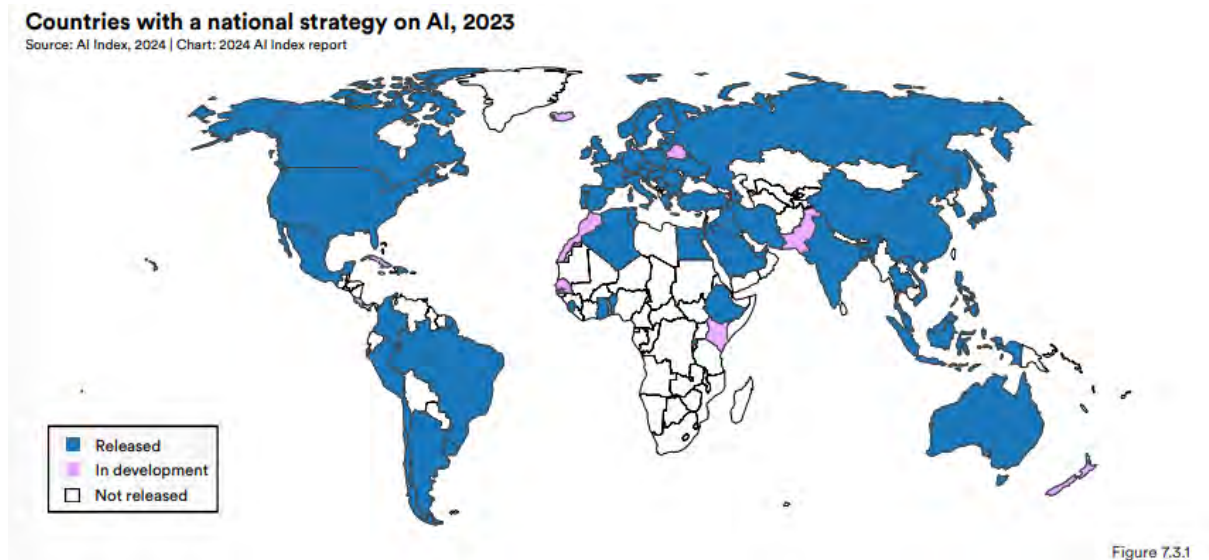
A continuación, desarrollaremos las políticas públicas sobre la IA en el mundo y en el Perú, realizando un particular énfasis en las entidades de la Administración Pública.

2.1.1 Políticas públicas sobre inteligencia artificial en el mundo

Si tuviéramos que analizar la IA basándonos exclusivamente en su aplicación o uso, podríamos delimitar dos aspectos: el público y el privado. El público hace referencia al uso que le dan los gobiernos a las IAs, a la implementación de éstas a través de sus servicios públicos. Mientras que el privado hace sentido con el trabajo que pueden dar las industrias a las tecnologías relacionadas con IA.

Al enmarcarse nuestro trabajo en el ámbito de la IA más relacionado con el uso “público” que se le pueda dar, desarrollaremos a continuación algunas estadísticas relevantes que sitúan a esta tecnología como una tecnología de particular interés para los gobiernos a nivel mundial.

Como un primer dato se tiene la siguiente gráfica donde puede observarse un planisferio que colorea las partes del mundo en donde se ha emitido una estrategia nacional sobre IA:



Nota. Adaptado de Artificial Intelligence Index Report 2024, por Nestor Maslej et al.
(2024)

Como se puede apreciar, la mayor parte del mundo ha emitido un plan estratégico sobre la IA. Si bien se trata de documentos de corte político, estos plantean la agenda legislativa en cada uno de los países donde se promulgan, sirviendo a todos los gobiernos como una guía o referencia del alcance o uso que se le quiere dar a la tecnología en cuestión.

Sobre la base de lo anterior, resulta importante explorar las políticas públicas llevadas a cabo en los países que principalmente han desarrollado tecnologías de IA, nos referimos a Estados Unidos, China y la UE.

a) Estados Unidos

El 11 de febrero de 2019 fue publicada una Orden Ejecutiva 13859 del Presidente de los Estados Unidos, llamada “*Maintaining American Leadership in AI*”. Dicho documento

define cinco (5) principios clave para desarrollar políticas de Estado en el campo del desarrollo de la IA:

- i) Estados Unidos debe impulsar avances tecnológicos sobre IA en todo el gobierno federal, la industria y el mundo académico para promover el descubrimiento científico, la competitividad económica y la seguridad nacional.
- ii) Estados Unidos debe promover estándares técnicos apropiados y reducir las barreras para la prueba y el despliegue seguros de tecnologías de IA para permitir la creación de nuevas industrias relacionadas con la IA y la adopción de la IA por las industrias actuales.
- iii) Estados Unidos debe capacitar a las generaciones actuales y futuras de trabajadores estadounidenses con las habilidades para desarrollar y aplicar tecnologías de IA a fin de prepararlos para la economía actual y los empleos del futuro.
- iv) Estados Unidos debe fomentar la confianza pública en las tecnologías de IA y proteger las libertades civiles, la privacidad y los valores estadounidenses en su aplicación para aprovechar al máximo el potencial de las tecnologías de IA.
- v) Estados Unidos debe promover un entorno internacional que respalde la investigación e innovación en IA estadounidense y abra mercados para las industrias estadounidenses de IA, al tiempo que proteja su ventaja tecnológica en IA y protege nuestras tecnologías de IA de la adquisición por parte de competidores estratégicos y naciones adversarias.

Sobre la base de esta orden, la Administración del Presidente de los Estados Unidos aprobó el 12 de noviembre de 2020 el memorándum para los jefes de los departamentos ejecutivos y agencias. Este documento determina los principios que deberán ser seguidos por las autoridades para desarrollar aproximaciones normativas y no normativas a la

implementación de sistemas IA tanto de forma general como por sector. Los principios se resumen tal como sigue según Kuteynikov et al. (2022):

- i) Crear confianza pública en el uso de los sistemas de IA.
- ii) Participación pública en la toma de decisiones.
- iii) Uso de información objetiva y con base científica en las actividades de la autoridades ejecutivas.
- iv) Aplicar un enfoque basado en el riesgo en la regulación del uso de los sistemas de IA.
- v) Considerar los beneficios y los costos de utilizar sistemas de IA.
- vi) Uso de enfoques flexibles y orientados a los resultados para la regulación.
- vii) Analizar las posibles consecuencias discriminatorias del usos de sistemas IA.
- viii) Garantizar la transparencia y la rendición de cuentas de las decisiones tomadas por los sistemas de IA.
- ix) Asegurar el funcionamiento de los sistemas de IA en todas las etapas.
- x) Implementar la coordinación interdepartamental entre los diversos organismos gubernamentales.

Después de la entrada en vigor de la Orden Ejecutiva antes mencionada, la Cámara de los Representantes adoptó una resolución titulada “*Supporting the development of guidelines for Ethical Development of AI*” el 27 de febrero de 2019.

Esta contiene tareas que se correlacionan con las disposiciones de la Orden cuya implementación debe garantizarse, tales como la interacción de la industria, el gobierno, la comunidad académica y la sociedad civil; la transparencia y “capacidad de expansión” de los sistemas de IA. Además, el empoderamiento de las mujeres y los grupos subrepresentados o marginados de la sociedad; la confidencialidad de la información y la

protección de datos personales; las oportunidades de carrera para las diferentes clases sociales; la rendición de cuentas y la supervisión de todos los sistemas automáticos de toma de decisiones; la educación permanente en ingeniería, ciencias sociales y humanidades; el acceso equitativo a los servicios tecnológicos; la investigación interdisciplinaria de IA segura y útil; la seguridad y el control de los sistemas de IA. (Kuteynikov et al., 2022, p. 697).

En 2021, más de diez (10) estados introdujeron proyectos de ley o resoluciones parlamentarias relacionadas con la regulación de la IA.

Todas ellas abordan cuestiones como el desarrollo de la IA (Alabama); el uso de métodos que minimicen el riesgo de consecuencias adversas causadas por los sistemas de toma de decisiones automatizados implementados por organismos gubernamentales (California); beneficios fiscales (Hawái); el establecimiento de requisitos para garantizar la equidad y transparencia de los sistemas de toma de decisiones automatizados utilizados por agencias gubernamentales, así como la confidencialidad de los datos de los consumidores (Massachusetts); la verificación de algoritmos y fórmulas lógicas utilizadas por la agencia de desempleo (Michigan); la prohibición de la discriminación contra ciertas categorías de la población por parte de los sistemas de toma de decisiones automatizados (Nueva Jersey); la creación de una comisión para supervisar el impacto de la tecnología en el mercado laboral y en la economía estatal en su conjunto (Nueva York); el establecimiento de un grupo asesor para eliminar los sesgos en el *software* gubernamental (Vermont); y el desarrollo de directrices para la contratación pública y el uso de sistemas automatizados de toma de decisiones con el fin de proteger a los consumidores y aumentar la transparencia del mercado (Washington) (Kuteynikov et al., 2022, p. 698).

Nótese que en Estados Unidos no existen leyes transversales que establezcan un enfoque legal unificado para la regulación de la IA tanto a nivel federal como estatal. Esta es una política del Estado basada en el principio de regulación *ad hoc*, que presupone la regulación de las relaciones sociales a medida que se desarrollan de manera natural. Así, los actos normativos adoptados en Estados Unidos son fragmentados y carecen de uniformidad desde una perspectiva semántica. Ello genera que la regulación legal solo afecte a ciertas áreas o a cuestiones específicas relacionadas con el desarrollo y funcionamiento de los sistemas impulsados por IA. Desde un punto de vista económico, esto es positivo; sin embargo, existe el riesgo de que ocurra un número significativo de incidentes con sistemas basados en IA antes de que se establezca un marco regulatorio sólido y suficiente en Estados Unidos. (Kuteynikov et al., 2022, p. 698).

Más allá de todos estos avances a nivel de políticas públicas en los Estados Unidos, la perspectiva que tiene el gobierno sobre la IA es bastante liberal y en contra de una excesiva regulación.

Recientemente, en la Cumbre sobre IA en París celebrada los días 10 y 11 de febrero de 2025, James David Vance, el vicepresidente de los Estados Unidos, señaló que no se deben colocar restricciones a una industria transformadora como la IA; además, indicó que como país la administración Trump buscará el desarrollo de la IA y se opondrá a su empleo en contextos de vigilancia y amenazas (DW Español, 2025).

Asimismo, Vance hizo referencia a la orden ejecutiva de Donald Trump emitida el 23 de enero de 2025, señalando que se tenía que equilibrar la regulación sin frenar la innovación. Indicó que se conocía de algunos gobiernos que pensaban imponer restricciones más estrictas a empresas tecnológicas estadounidenses con presencia internacional (DW Español, 2025).

b) China

China ha alcanzado resultados considerables en el campo de la protección de datos. En 2016, la Ley de Ciberseguridad de China fue adoptada y estableció requerimientos regulatorios como los existentes en la UE y en Estados Unidos. Dado que China es un Estado con un sistema político autoritario, la privacidad y confidencialidad de los datos está más alineada con las decisiones de las autoridades del Estado que con la creación de un marco regulatorio que sea reforzado por decisiones judiciales. (Kuteynikov et al., 2022, p. 700).

En julio de 2017, el Consejo de Estado de la República Popular de China anunció una estrategia para el desarrollo de IA llamada “Plan de Desarrollo de IA para la nueva generación”. Esta estrategia determina como meta nacional convertirse en el líder global de IA para el 2030 y en el desarrollo de marcos regulatorios, éticos y estándares relacionados a la IA. Este plan representa solo un modelo general y objetivos de una futura regulación (Kuteynikov et al., 2022, p. 699).

El enfoque de China en la regulación de la IA tiene un doble énfasis: la promoción de la innovación y el control estatal. China emplea leyes específicas para abordar problemas individuales relacionados con la IA, adoptando un enfoque regulatorio vertical. Hasta ahora, su regulación ha abordado desafíos como los algoritmos de recomendación y las herramientas de síntesis profunda (a menudo utilizadas para crear *deepfakes*). Las normativas exigen que los proveedores de servicios limiten la discriminación, mitiguen la propagación de información negativa, etc. Las leyes sobre herramientas de síntesis profunda establecen que dicho contenido debe cumplir con los controles de información

y etiquetarse como generado sintéticamente, además de implementar medidas adicionales para prevenir su uso indebido (Walter, 2024, p. 8).

A pesar del uso de la IA en la aplicación de la ley y la vigilancia en China, se han introducido regulaciones para controlar el uso de esta tecnología por parte de entidades no gubernamentales. Estas normativas estipulan los propósitos específicos para los cuales se pueden utilizar las herramientas de reconocimiento facial, haciendo hincapié en la seguridad pública en espacios públicos (Walter, 2024, p. 8).

Lo anterior se diferencia del enfoque de Estados Unidos y la UE. Por el lado de Estados Unidos, debido a que este país adopta un enfoque más descentralizado, centrado en aplicaciones específicas de la IA. En cuanto a la UE, esta última ha implementado un enfoque integral basado en el nivel de riesgo. En definitiva, la combinación de promoción de la innovación, control estatal e influencia social en China refleja sus actitudes políticas, como el comunismo y el colectivismo. (Walter, 2024, p. 8).

Desde la perspectiva de Esquivel Fernández (2024), China supera a Estados Unidos en lo que respecta a la IA en tres (3) aspectos centrales:

- Integración de los datos: China se caracteriza por concentrar la mayor cantidad de información en manos del Estado. Ello es débil en el caso de Estados Unidos y la UE, puesto que sus grandes empresas tecnológicas acaparan grandes cantidades de datos sin una cooperación efectiva entre ellas o debido a rivalidades existentes.
- Aplicación en el uso cotidiano: China usa IA en sus labores públicas, específicamente en temas de seguridad nacional e internacional.

- Cantidad de datos obtenidos: La población china representa 4 veces más de información que su contraparte americana.

c) Unión Europea

La política pública en la UE ha sido materializada a través de la EU AI Act, que se considera como la primera ley integral sobre IA en el mundo. Propuesta por la Comisión Europea en abril de 2021, esta clasifica los sistemas de IA en diferentes niveles de riesgo. La EU AI Act es parte de la estrategia digital de la UE y tiene como objetivo garantizar mejores condiciones para el desarrollo y el uso de tecnologías de IA, promoviendo mejoras en sectores como la salud, el transporte y la manufactura. (Walter, 2024, p. 4).

Asimismo, el Parlamento Europeo destaca la importancia de que los sistemas IA en la UE sean seguros, transparentes, rastreables, no discriminatorios y respetuosos con el medio ambiente. La postura de la UE radica en la prioridad de proteger a las personas de los posibles riesgos de la IA mientras fomenta sus beneficios. (Walter, 2024, p. 4).

Sobre la postura de la UE sobre IA, Ursula von der Leyen, presidenta de la Comisión Europea, ha señalado en la Cumbre sobre IA en París que se debe desarrollar una IA de confianza, lo cual requiere pasar por una regulación (FRANCE 24 Español, 2025).

Con relación al enfoque político de la UE, vale la pena rescatar la opinión de Estrada, quien señala que la estrategia de esta última consistió en considerar una ventaja el crear el primer estándar regulatorio global. Sin embargo, ello depende de que el Reglamento se desarrolle e implemente eficazmente respecto de la mitigación de daños, sin quedarse obsoleto ante los avances de la IA. Así, la UE cree que la creación de un ecosistema de confianza es una ventaja competitiva al aportar certidumbre a las empresas y seguridad a los ciudadanos. No obstante, también es consciente de que no es suficiente y que se

encuentra por detrás en la industria tecnológica, lo que ha quedado en evidencia desde el 2018 con la iniciativa “Inteligencia artificial para Europa”, donde uno de los pilares de actuación fue potenciar la capacidad tecnológica e industrial de la UE e impulsar la adopción de la IA en todos los ámbitos de la economía, tanto en el sector privado como en el público” (2024, pp. 41-42).

De la experiencia internacional citada, resalta la predilección en el mundo por regular el uso de IA desde una perspectiva de mitigación de riesgos, estableciendo principios y normas avocadas a la protección de la información, los datos personales, la mitigación de sesgos discriminatorios, y la mitigación de la oscuridad en los resultados.

Como se detallará a continuación, la normativa promulgada sobre el uso de IA resulta bastante escueta como para apreciar un enfoque en la mitigación de riesgos. Si bien los riesgos que el uso de IA representa no han sido dejados de lado en el contexto peruano, sí resalta una predilección por la promoción del uso de estas nuevas tecnologías, antes que por la mitigación de riesgos en este contexto.

En los siguientes apartados, se revisarán las políticas públicas relativas al uso de IA en el contexto peruano.

2.1.2 Políticas públicas sobre inteligencia artificial en el Perú

A continuación, nos referiremos a los principales antecedentes políticos sobre IA en el Perú.

a) La Estrategia Nacional de Inteligencia Artificial

La Estrategia Nacional de Inteligencia Artificial (ENIA) es un documento de trabajo elaborado por la Secretaría de Gobierno y Transformación Digital de la Presidencia del

Consejo de Ministros (en adelante, “SGTD”) en 2021 cuyo propósito es facilitar las oportunidades que generan las nuevas tecnologías como la IA y reducir los riesgos derivados como consecuencia de su implementación. Precisa además que el desarrollo que pudiera darse en tecnologías de IA no debería ser solo una preocupación del sector privado, sino que también debe ser abordada por el sector público; el cual es el llamado a atender las nuevas necesidades de una ciudadanía digital (SGTD, 2021, p. 50).

Si bien el documento tiene múltiples alcances, nuestro propósito es comentar los ejes y objetivos estratégicos que esta presenta. Los ejes estratégicos que se presentan en la ENIA son los siguientes:

Eje estratégico	Nombre	Detalle
1	Formación y captación de talento	Formación de personal profesional que reúna competencias de investigación, desarrollo y uso de la IA en el país.
2	Modelo económico	Promover la IA y su adopción como medio de impulso del desarrollo, innovación y el bienestar del país.
3	Infraestructura tecnológica	Hacer posible la creación y el fortalecimiento de la infraestructura digital y la relacionada con telecomunicaciones como soporte de la IA.

4	Datos	Hacer posible la infraestructura de datos para poner a disposición datos públicos de alta calidad en formato abierto, reutilizable y accesible.
5	Ética	Adoptar lineamientos éticos para el uso sostenible, transparente y replicable de la IA con nociones claras sobre responsabilidad y protección de datos.
6	Colaboración	Desarrollar un ambiente colaborativo de IA a nivel nacional e internacional.

Nota. Adaptado de ENIA, por SGTD, 2021.

Como se puede apreciar, no todos los ejes estratégicos de la ENIA se encuentran vinculados a las actividades de la Administración Pública, por lo que en este caso nos ocuparemos principalmente del Eje Estratégico 2 sobre modelo económico, que tiene entre sus objetivos estratégicos el fortalecer en los organismos públicos la implementación de la IA en sus operaciones y servicios a la ciudadanía (SGTD, 2021, p. 65).

Dicho objetivo estratégico tiene además los siguientes objetivos específicos:

Objetivo estratégico 2.2	Impulsar en organismos públicos, la implementación de IA en sus operaciones y servicios a la ciudadanía
A.2.2.1	Realizar cursos y capacitaciones en línea para funcionarios públicos en la implementación, uso y beneficios de la IA con ejemplos locales.

A.2.2.2	Promocionar el uso de asistentes virtuales (chatbots) en la Administración, priorizando el <i>software</i> de código abierto.
A.2.2.3	Priorizar, mediante el Centro de Innovación e Inteligencia Artificial, el desarrollo de casos de uso donde la IA pueda generar soluciones prácticas como las propuestas en varias investigaciones alineadas a los ODS, como la eliminación de la pobreza, del hambre, la educación de calidad, el agua y energía limpias, las ciudades sostenibles, la buena salud, la reducción de las brechas sociales y otros.
A.2.2.4	Promover la participación ciudadana en eventos de IAckathon, Datathon o Hackathon organizados por la Administración Pública para crear nuevos servicios o mejoras en la atención del ciudadano, o para resolver problemáticas del país, tales como la detección de la corrupción en el ámbito público.
A.2.2.5	Promover en el sector público el uso de tecnologías de código abierto para el desarrollo de casos de uso en base a la IA.

Nota. Adaptado de ENIA, por SGTD, 2021.

Como se puede apreciar de la tabla, el objetivo estratégico 2.2 se encuentra orientado a promocionar el uso de asistentes virtuales en la Administración, así como el uso del código abierto para el desarrollo de IA (SGTD, 2021, pp. 67-68).

Estos subobjetivos, como han sido descritos, son de suma utilidad para brindar a los usuarios de un portal web (por ejemplo, de una entidad cualquiera de la Administración Pública) información útil sobre trámites administrativos o procedimientos que se necesiten realizar. Así, lo que se promueve es la creación de ventanas de usuario que

permitan en aplicaciones web “conversar” con la computadora por chat y tener una respuesta rápida y eficiente. De este modo, se reemplaza la necesidad de llamar a la central telefónica de las entidades, las cuales suelen estar congestionadas debido a la gran cantidad de administrados con consultas de todo tipo.

No obstante, si bien del documento se aprecian ideas interesantes sobre la promoción de aplicaciones como el *chatbot* y el *software* de tipo “*open source*”, se trata de un objetivo que a la fecha ya puede considerarse satisfecho en diversas entidades importantes del país, las cuales cuentan con un sistema chatbot integrado en su página web abierta al público que la necesite.

Por otro lado, el documento comparte los resultados de una encuesta realizada en enero del 2021 sobre el status de la IA en el ámbito público peruano:

Resultados de la pregunta: ¿Estás usando IA o ciencia de datos en tu institución?

Tipo de Institución	No	Sí	Total general
Empresa del Estado	12		12
Gobierno Local	375	29	404
Gobierno Regional	28	2	30
Organismos Autónomos	7	1	8
Poder Ejecutivo	17	6	23
Poder Judicial	1		1
Programa Social	3		3
Universidad	6	1	7
Total general	449	39	488

(*) Respondieron 454 instituciones públicas

Nota. Recogido de ENIA, por SGTD, 2021.

Al respecto, cabe indicar que según el documento, dicha encuesta fue respondida solo por el 20% de las entidades públicas, de las cuales apenas el 7% afirmaron usar algún tipo de IA en sus instituciones (SGTD, 2021, p. 41).

Se critica la metodología por la que fue realizada la encuesta, ya que el 20% de entidades públicas a nivel de un país no resulta ser una cifra o dato representativo. Por lo tanto, los resultados de esta encuesta deben tomarse con cautela, ya que son meramente referenciales. No obstante, la cifra del 7% nos demuestra que la cantidad de entidades públicas en el Perú que usan IA en su día a día es, por lo menos, menor a la mitad del total.

Pese a la metodología empleada, notamos algunos datos interesantes: ninguna de las empresas estatales ni el Poder Judicial emplean en sus tareas IA.

La encuesta también consultó “¿En qué casos aplicas IA o ciencia de datos en tu organización?” Sobre la misma se respondió que se usaba IA o ciencia de datos en la predicción de delitos, el reconocimiento facial, la identificación por medio de imágenes satelitales, la supervisión de gas natural y electricidad, los programas sociales, el control de usuarios de luz y agua, las aplicaciones tributarias, la gestión de riesgos y los asistentes virtuales (SGTD, 2021, p. 41).

Vemos cómo son múltiples las utilidades que la IA tiene para la Administración Pública desde el 2021. Sin embargo, este dato debe tomarse con la debida cautela, ya que la pregunta es abierta, pudiendo referirse a aplicaciones basadas en IA como a aquellas relacionadas con ciencias de datos.

Finalmente, otro dato interesante que trae el referido documento es el desarrollo de algunos aplicativos que se sirvieron de IA durante la pandemia causada por el COVID-19:

- Aplicativo “El Perú en tus manos”: Se trata de una aplicación desarrollada en colaboración con empresas privadas expertas en el rubro de aplicaciones, IA y

analítica de datos. Fue creada con el fin de alertar a la ciudadanía ubicada en las zonas del Perú que presentaban una mayor posibilidad de contagio (SGTD, 2021, p. 44).

- Proyecto “COVID: Control de virus dinámico”: Fue el proyecto ganador que convocó el Concytec. Este aplicativo detecta posibles infectados de COVID-19 y realizar, seguidamente, un prediagnóstico del virus (SGTD, 2021, p. 45).
- Sistema Integrado Inteligente para registrar, informar, alertar y asistir a las personas sintomáticas de COVID-19 (SIAMA): Dicho sistema contaba con un asistente de voz que realizaba preguntas periódicas y que, además, le informaba al personal de asistencia más cercano del infectado. SIAMA emplea DL, usando un *chatbot* con lenguaje natural (SGTD, 2021, p. 46).
- Smart Security Office: Fue el proyecto que ganó el reto InnovaCovid-19 de Innóvate Perú. Su fin es la gestión y el control del acceso de los colaboradores y visitas a los centros laborales, a través de la confirmación de identidad, la detección de mascarilla, la medición de la temperatura corporal y el triaje de prevención del COVID-19. Implementa un sensor de temperatura y un *chatbot* (SGTD, 2021, p. 47).

b) Inteligencia Artificial: Desafíos y oportunidades para el Perú

Después de la publicación de la ENIA, el Centro Nacional de Planeamiento Estratégico (CEPLAN) realizó un estudio cuyo fin es informar sobre las tecnologías que comprenden lo que se entiende como IA, sobre sus usos en sectores significativos del Perú, su potencial impacto económico, social y ambiental, así como los primeros pasos que vienen dando diversos países para aprovechar su potencial (CEPLAN, 2021, p. 8).

En relación con dicho documento, nos interesa particularmente su acápite 5.3 titulado “Situación y perspectivas de la IA en el Perú”. En dicho acápite señala que la IA se encuentra aún en una etapa de desarrollo en Perú, sin haber sido explorada totalmente en los distintos sectores e industrias. Precisa además que los sectores industrial, académico y estatal reconocen como carencias algunos factores habilitantes de la tecnología, tales como la infraestructura, la conectividad, una cultura basada en datos y el acceso a datos abiertos de calidad.” (CEPLAN, 2021, p. 49).

Dentro del anterior acápite, precisamente en el subacápite “5.3.3. Estado” se recogen, al igual que en la ENIA, algunos casos específicos de vital interés:

Nombre de la aplicación	Entidad de la Administración Pública y otros actores responsables	Descripción
AYNI LAB	Ministerio de Desarrollo e Inclusión Social, Universidad Peruana Cayetano Heredia y Ministerio de la Producción	Fue pensada para la anemia. Funciona con una fotografía de la conjuntiva ocular del menor que es enviada a un sistema de redes neuronales que procesa la imagen para determinar el nivel de hemoglobina y, así, la presencia o no de anemia (Mont et al., 2020, como se citó en CEPLAN, 2021, p. 53).
Open Clinic	Gobierno Regional de Cajamarca	Se trata de una IA que permite el análisis de historias clínicas, el

		reconocimiento de síntomas de un paciente, la realización de un diagnóstico automatizado a partir de imágenes y, con ello, la identificación de una posible enfermedad. (CEPLAN, 2021, p. 53).
--	--	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Más allá de las aplicaciones sobre la base de la IA que son usadas por la Administración Pública, este documento concluye que la normativa que fuera a establecerse sobre el desarrollo de la IA debe ser revisada y analizada con profesionales de distintas disciplinas principalmente para temas de privacidad y sesgos (CEPLAN, 2021, p. 56). Así, se señalan como principales problemas para la regulación de la IA a la discriminación y a la protección de datos personales.

c) Política Nacional de Transformación Digital

La agenda pública del Perú en lo que respecta a la IA es relativamente nueva, ya que fue recién por medio del Decreto Supremo N.º 085-2023-PCM, publicado el 28 de julio de 2023, que se aprobó la Política Nacional de Transformación Digital al 2030 (en adelante, “PNTD”).

La PNTD tiene la premisa de la aparición de tecnologías disruptivas, las cuales generan la necesidad de una transformación digital transversal en el país. El problema público identificado en la PNTD es el ejercicio de la ciudadanía digital.

Según la PNTD, el ejercicio de la ciudadanía digital es la capacidad de una persona para participar como agente activo en la sociedad, ejerciendo sus derechos y obligaciones en

un entorno digital seguro, lo cual se logra a través de la adquisición de una serie de capacidades y libertades (SGTD, 2023, p. 52).

En relación con la IA, el documento indica que uno de los problemas que ocasionan bajo ejercicio de ciudadanía digital es la incipiente cultura sobre un conjunto de temas, entre ellos, la innovación, la gestión segura, la ética e inteligencia de datos, así como también las tecnologías digitales y la IA. Sobre esta última indica que el desconocimiento o resistencia sobre el uso de herramientas basadas en IA podría llevar a su rechazo, privando a los administrados de muchos beneficios o utilidades que le podrían brindar su uso. Coloca el documento algunos ejemplos como la falta de confianza o comodidad al interactuar con un *chatbot*, o el de no confiar en el resultado de sistemas expertos basados en algoritmos de IA (por ejemplo, la detección de la presencia o ausencia de alguna enfermedad) (SGTD, 2023, p. 1106).

Es importante el apunte que hace la PNTD sobre los efectos de lo que identifica como problema público (el bajo ejercicio de la ciudadanía digital): i) la limitación al ejercicio de derechos, ii) el debilitamiento de la participación ciudadana, iii) el debilitamiento de la transparencia, iv) el entorpecimiento del proceso de inclusión, v) el obstáculo al diálogo intercultural, vi) genera poca eficiencia en las compañías y vii) obstaculiza en la ciudadanía el acceso a las herramientas para su desenvolvimiento integral, así como el desarrollo de las competencias digitales ciudadanas debido a la retroalimentación que se genera.

De estos efectos, resultan importantes para los fines del presente trabajo el primero y el tercero. Respecto a la limitación del ejercicio de derechos, la PNTD reconoce que los deberes establecidos y su cumplimiento se encuentran relacionados con la interacción que tiene, a través de trámites administrativos, el ciudadano con la Administración. Por

otro lado, respecto al debilitamiento de la transparencia, el documento se refiere a interacciones de la ciudadanía con la Administración en las que no existen medios que hagan efectiva una rendición de cuentas (SGTD, 2023 p. 140).

Vemos entonces cómo se materializan algunos riesgos de la IA dentro de la PNTD, como la falta de transparencia. Sin embargo, no deja de ser un documento con un alcance político, por lo que debemos observar cuál es el planteamiento más afín con la IA.

La PNTD se desarrolla a partir de “Objetivos prioritarios”, los cuales se resumen como sigue:

- Objetivo prioritario 1: Asegurar el acceso inclusivo, seguro y de calidad al entorno digital a la ciudadanía.
- Objetivo prioritario 2: Relacionar la economía digital con los procesos productivos sostenibles.
- Objetivo prioritario 3: Asegurar la disponibilidad de servicios públicos digitales inclusivos, predictivos y empáticos con los administrados.
- Objetivo prioritario 4: Fortalecer el talento digital en todas las personas.
- Objetivo prioritario 5: Reforzar la seguridad y confianza digital en la ciudadanía.
- Objetivo prioritario 6: Asegurar el uso ético y la adopción de nuevas tecnologías y la innovación en la sociedad.

Se puede apreciar que el objetivo prioritario más afín con la regulación de la IA es el Objetivo Prioritario 5, ya que en el desarrollo de este se indica que existe una dinamicidad originada por las nuevas tecnologías (IA, ML, análisis de datos. *blockchain*, 3D, etc.) la

cual dificulta el diseño de un ecosistema digital de confianza, que haga posible -entre otros- la participación, el manejo seguro de datos y el respeto a la privacidad. Señala además que las amenazas y los riesgos se incrementan, no pudiendo ser estos un obstáculo para la interacción de los administrados con su entorno digital, debiendo ser un factor clave la gestión de los riesgos de seguridad, privacidad y protección de la ciudadanía en las políticas de los países, a fin de que los administrados, las grandes compañías y organizaciones estén aún más preparadas para la gestión y el abordaje de los riesgos de seguridad digital (SGTD, 2023, pp. 228-229).

Vemos como la gestión de riesgos se vuelve una prioridad para generar ese entorno seguro a la ciudadanía. El propio trabajo de la SGTD así lo reconoce cuando se refiere específicamente sobre el Objetivo Prioritario 5, y señala que algunas de las acciones propuestas para llevar a cabo dicho objetivo fueron que las entidades de la Administración debían involucrar a sus áreas de riesgos en el marco de seguridad y en la estandarización de la metodología de gestión de riesgos (SGTD, 2023, p. 249).

Así también lo refieren los lineamientos del referido Objetivo Prioritario 5:

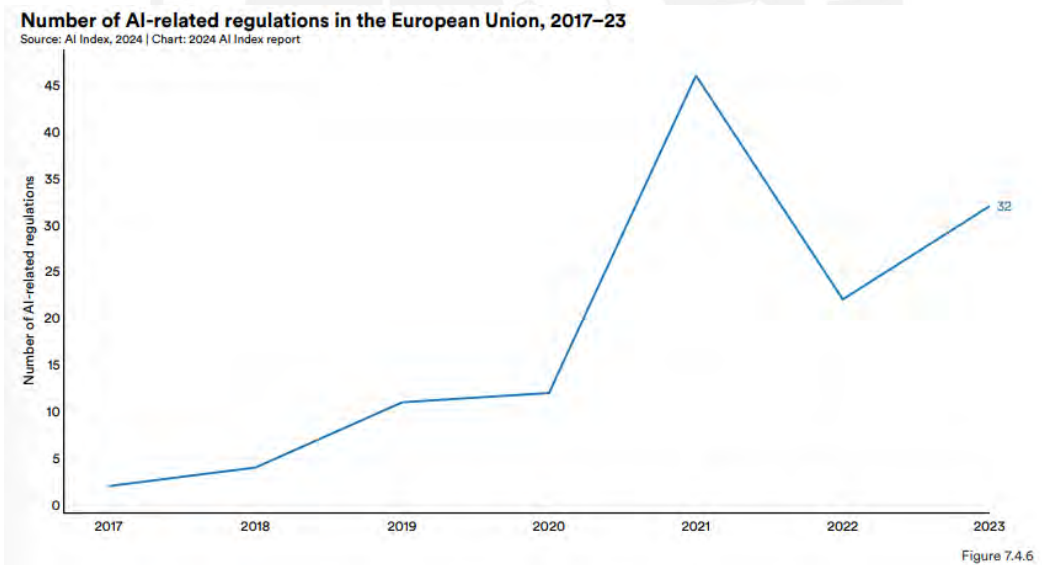
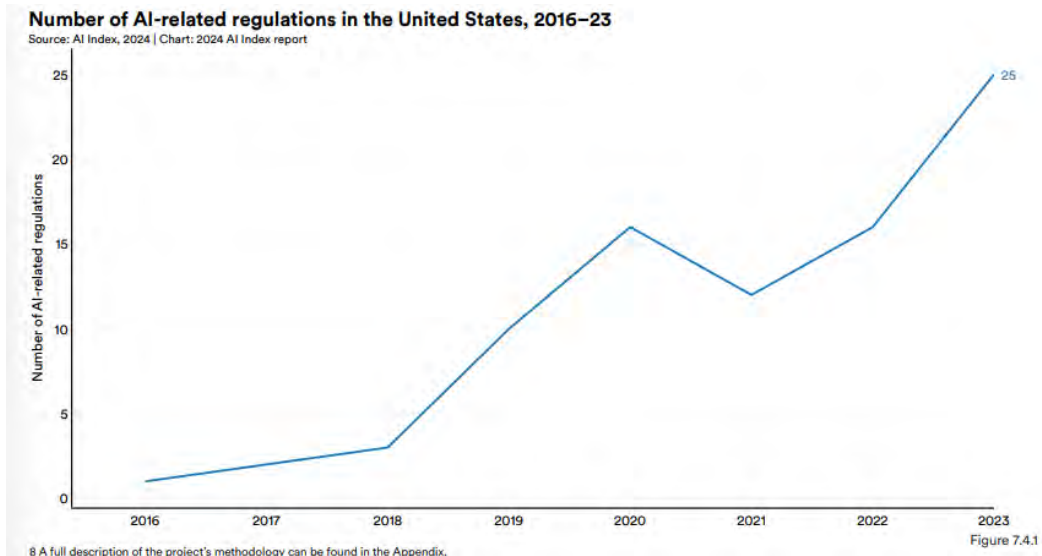
Lineamiento	Nombre	Detalle
5.1	Asegurar el adecuado funcionamiento del Sistema Nacional de Transformación Digital (SNTD) en	Indica que la confianza digital tiene como aspectos, la transparencia, la protección de datos personales, la ética, la seguridad digital y la protección del consumidor. Añade que se encuentra incluida la ciberseguridad, en la medida en que hace posible preservar el funcionamiento de redes, activos y sistemas informáticos, así

	el país, elaborando y articulando las estrategias nacionales con énfasis en seguridad y confianza digital	como protegerlos ante amenazas y vulnerabilidades. Se encuentra incluido dentro del lineamiento la mitigación de riesgos que puedan afectar a personas, así como también a la actividad económica (privada y pública), considerando además la implementación de dichas medidas que aseguren la confianza digital (SGTD, 2023, pp. 249-250).
5.2	Fomentar una cultura de gestión de riesgos y confianza digitales, con énfasis en la protección de niños, niñas y adolescentes	Señala que debe proveerse una cultura para proteger los derechos de la ciudadanía en el ámbito digital. Indica que todo ecosistema digital es un espacio permanente de riesgos y posibles amenazas, de forma que se debe sembrar confianza en toda transacción digital. (SGTD, 2023, p. 250).

Como se puede apreciar, la lectura de los lineamientos del Objetivo Prioritario 5 lleva a determinar que la PNTD, en lo que a IA se refiere, se encuentra -entre otros- orientada a la gestión de riesgos, de modo que sea posible el diseño de un entorno seguro para proteger los derechos de las personas.

2.2 Marco regulatorio sobre inteligencia artificial

En lo que regulación se refiere, podemos observar que, a nivel mundial, existe un gran número de propuestas legislativas relacionadas con IA. Esto puede verse reflejado en las siguientes gráficas sobre los casos norteamericano y europeo:



Nota. Adaptado de Artificial Intelligence Index Report 2024, por Nestor Maslej et al.

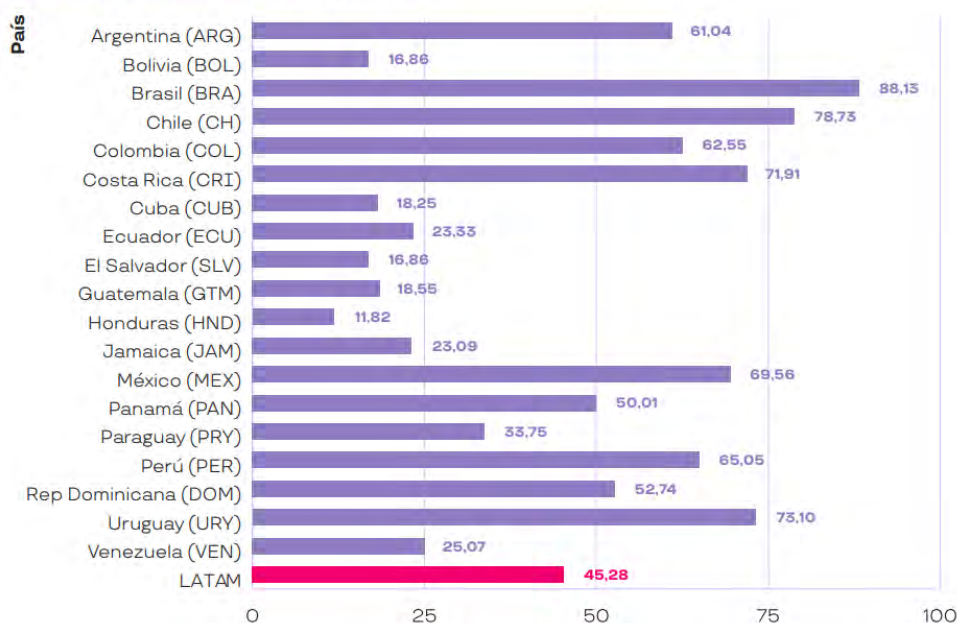
(2024)

Como se observa en las gráficas, en el caso de Estados Unidos, el número de regulaciones relacionadas con IA ha crecido exponencialmente desde el 2016 exponencialmente, año

en el que solo se conocía una regulación con dichas características. Por otro lado, en el caso de Europa, se tiene la particularidad que en el año 2021 pasaron por la UE alrededor de 46 regulaciones relacionadas con IA. (Stanford University Human-Centered Artificial Intelligence, 2024).

En relación con la región latinoamericana, a la fecha se cuenta con el Índice Latinoamericano de Inteligencia Artificial (ILIA), el mismo que reporta el estado de preparación de la IA en los 19 países de la región y es desarrollado por la Comisión Económica para América Latina y el Caribe (CEPAL) y el Centro Nacional de Inteligencia Artificial de Chile (CENIA)⁸. En relación con la regulación, este reporte identifica lo siguiente:

Gráfico 3: Subdimensión de Regulación



Nota. Adaptado de ILIA 2024, por CEPAL y CENIA (2024)

⁸ Este índice evalúa la preparación de los países desde tres (3) dimensiones: factores habilitantes, investigación, desarrollo y adopción, y gobernanza. Básicamente lo que se intenta abarcar es la infraestructura tecnológica, el desarrollo del talento humano y los marcos de gobierno necesarios para integrar la IA en las economías nacionales.

La estadística anterior es lo más cercano que se tiene a una información cuantitativa sobre la regulación que existe sobre IA a nivel latinoamericano. No obstante ello, debemos leer la misma con un poco de cautela, ya que proyecta lo que el ILIA 2024 entiende como la “subdimensión de regulación”, la cual mide “la madurez de los sistemas normativos que exceden las acciones concretas del poder ejecutivo y organismos internacionales” (2024, p. 236). Según el ILIA 2024, dicha subdimensión se compone de 3 indicadores: regulación sobre IA, ciberseguridad y ética y sostenibilidad.

Regresando a la gráfica, es Brasil quien lidera la subdimensión con 88,13 puntos. Le sigue Chile con 78,73 puntos. Se ubican muy por encima del promedio de la región en su conjunto, que tiene un puntaje de 45,28 puntos.

Como se puede apreciar, el Perú es un país que tiene un buen puntaje en lo que a regulación se refiere, siendo considerado inclusive por el propio ILIA 2024 como un país de regulación avanzada (65,05 puntos).

Sobre la base de este contexto, en el presente acápite se dilucidarán algunas cuestiones sobre la regulación de la IA. En primer lugar, se comentará sobre la necesidad de regular la IA. En segundo lugar, se desarrollará una noción o concepto de regulación que se adapte al fenómeno de la IA. Finalmente, se comentarán los principales enfoques regulatorios de *hard law* y *soft law* aplicables a la IA.

2.2.1 ¿Es necesario regular la inteligencia artificial?

Antes de pasar por la pregunta sobre: ¿cómo regular la IA? Debe afrontarse la pregunta de si es realmente necesario una nueva regulación asociada con nuevas tecnologías o si, por el contrario, son suficientes los instrumentos clásicos de regulación ya existentes (Fenwick et al., 2017; Leenes, 2019, citado en Fierens et al., 2021, p. 58).

La formulación de esta pregunta preliminar no es un tema menor, puesto que se trata de evaluar si es que existen suficientes razones para una regulación nueva. No obstante, para los referidos autores la respuesta a dicha pregunta ya ha sido ampliamente debatida y es claro que la IA trae consigo riesgos nuevos y desconocidos que nuestras leyes actuales no pueden enfrentar (Fierens et al., 2021, p. 58).

Es importante la idea de los autores sobre la regulación actual y que ésta no es suficiente para soportar los nuevos y desconocidos riesgos de la IA. A una comprensión similar llega Covarrubias con otros autores tomando en cuenta lo señalado por la Comisión Europea:

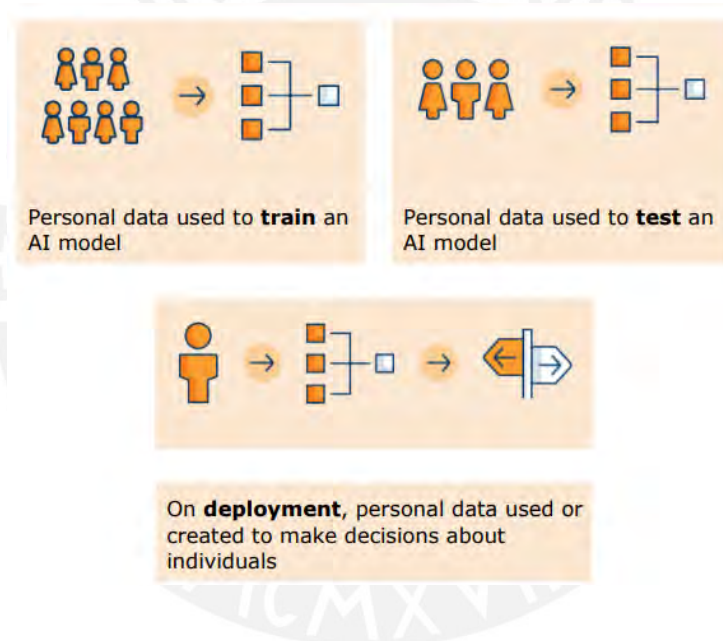
Por lo tanto, aunque los desarrolladores e implementadores de la IA ya se encuentran dentro del ámbito de aplicación de regulación de protección de datos, privacidad, no discriminación, protección de consumidores, seguridad de los productos y responsabilidad civil, algunas características específicas de la IA pueden dificultar la aplicación y el cumplimiento legal, y, aunque esto no descarta contemplar la legislación actual para hacer frente a los riesgos de la IA, sí supone la probabilidad de crear una nueva legislación. (Comisión Europea, 2020, citado en Covarrubias et al., 2022, p. 36)

Al igual que la Unión Europea, no negamos que sea posible paliar o mitigar algún riesgo generado o producido por la IA sobre la base de alguna regulación ya existente a la fecha, como puede ser la regulación aplicable a los datos personales.

Espinoza (2024) ya comentaba que en el caso peruano hacia abril de 2024 existían denuncias interpuestas y fiscalizaciones iniciadas por el uso indebido de datos personales por parte de desarrolladores de sistemas de IA. Enfatizaba, sobre la base de la publicación del Proyecto de Ley N.º 7033/2023-CR de fecha 13 de febrero de 2024, que resultaba

importante y necesario que no exista duplicidad de obligaciones entre las normas sobre datos personales y las que pretendían regular la IA, ya que una sobreregulación podría involucrar un desincentivo para los desarrolladores de ésta.

Sobre la base de lo explicado en el primer capítulo, nos queda claro que, al usar datos (entre ellos, datos personales), resulta perfectamente posible que se apliquen, ante un mismo hecho, una regulación específica de datos personales y una regulación específica de sistemas de IA. Para ser más específicos sobre la intersección que tendrían estas regulaciones, resultan bastante gráficos los siguientes tres (3) supuestos:



Nota. Adaptado de Explaining decisions made with AI, por The Alan Turing Institute y de Information Commissioner's Office (ICO), 2022.

Nótese que, pese a que los datos personales pueden ser utilizados o creados en varias de las etapas del ciclo de vida de la IA, este tipo de datos no son los únicos que pueden ser empleados en dichos sistemas. Es por dicha razón que las regulaciones antes mencionadas

no son excluyentes entre sí, sino que deben buscar la complementariedad y coherencia normativa⁹.

La complementariedad de la que comentamos debe tomar como premisa que, si bien existen regulaciones no especiales sobre IA que podrían resultar aplicables, los sistemas de IA tienen riesgos específicos propios de su tecnología tal como se comentó en el capítulo anterior.

2.2.2 La noción de regulación aplicable a la inteligencia artificial

Así como la IA tiene un problema de definición, la regulación en sí misma también lo tiene. La doctrina especializada propone tres definiciones al respecto. La primera señala que la regulación es la promulgación de reglas por parte del gobierno acompañadas de mecanismos de seguimiento y aplicación (llevados a cabo a través de una agencia pública especializada). La segunda toma como regulación cualquier forma de intervención estatal directa en la economía. La tercera entiende por regulación todos los mecanismos de control o influencia social que afectan todos los aspectos del comportamiento desde cualquier fuente, sean estos intencionales o no (Black, 2002, citado en Fierens et al., 2021, pp. 55-56).

Vale la pena mencionar que la primera definición es muy cercana a la realidad peruana, en la que existe la noción de un “sector” en la economía, donde actúa típicamente un ente rector que dicta las normas y establece los procedimientos que se relacionen con él. No obstante, hoy en día podemos afirmar que la regulación o la capacidad de influir en el comportamiento de los actores en el mercado no depende solo del Estado, sino de la

⁹ A similar conclusión llegan Fierens et al. (2021), al indicar que mucha legislación es potencialmente aplicable a sistemas de IA, poniendo como principal ejemplo al Reglamento General de Protección de Datos de la Unión Europea.

actuación de un conjunto de cosas (similar a la tercera definición antes mencionada). Así, para Black, la regulación vendría a ser un proceso que implica el intento sostenido y enfocado de alterar el comportamiento de otros con la intención de producir resultados identificables (Black, 2002, citado en Fierens et al., 2021, p. 56).

La definición de Black hace eco a la idea de “*smart regulation*” desarrollada ampliamente por Gunningham y Sinclair desde 1998 y, hoy en día, definida como una forma de pluralismo regulatorio que abarca formas flexibles, imaginativas e innovadoras de control social. Así, la *smart regulation* aprovecha tanto a los gobiernos como a las empresas y a terceros, abarcando la autorregulación y la correulación, utilizando intereses comerciales y Organizaciones No Gubernamentales (ONGs) como sustitutos regulatorios, junto con la mejora de la eficacia y eficiencia de formas más convencionales de regulación gubernamental directa. (Gunningham y Sinclair, 2017, p. 133).

Pareciera que la esencia de la *smart regulation* es lo que más coincide con la conceptualización de Black, ya que la óptica sigue siendo que el Estado no es el único que termina influyendo en el comportamiento de los actores que empleen o desarrollen IA. A nuestro parecer, esta es la perspectiva que debe tener todo operador jurídico, quedando bastante claro que no sirve tener una concepción “clásica de regulación”, sino que debe ampliarse un poco el espectro considerando la posible influencia que tengan otros actores relevantes que puedan influir en esa regulación.

2.2.3 Enfoques regulatorios de la inteligencia artificial

La doctrina más reciente ha estudiado los enfoques regulatorios que serían aplicables para la IA. A continuación, un desarrollo de estos.

a) Normas vinculantes o *hard law*

Respecto de las normas vinculantes, es preciso señalar que la doctrina considera como instrumentos de *hard law* las sentencias, legislaciones, jurisprudencia, o tratados, los cuales tienen fuerza legal, al estar compuestas por normas vinculantes, sanciones fijas y procedimientos. Las normas vinculantes son más fáciles de cumplir al encontrarse codificadas, pero también difíciles de ejercer en el tiempo por su velocidad e inflexibilidad (Covarrubias et al., 2022, p. 38).

Al respecto, vale la pena mencionar que la vinculatoriedad dependerá de muchos factores. Así, por ejemplo, resultará muy distinto hablar de “vinculatoriedad” en un país con una tradición jurídica anglosajona y un país como el peruano, toda vez que en este último caso no todo precedente resultará vinculante. Por último, queda claro que el impacto es más beneficioso en la mayoría de los países si es que hay algún tipo de base legal puesto que el tema de la vinculatoriedad no será puesto en duda, salvo por el ámbito de aplicación de la base legal en cuestión.

b) Normas indicativas o *soft law*

Existen un grupo de mecanismos regulatorios que se contraponen con las normas vinculantes conocido como *soft law*. A continuación, una explicación de cada uno de ellos:

Mecanismo	Definición
-----------	------------

Autoestructuración	La conducta del sujeto se rige por objetivos preestablecidos por él mismo, estableciendo una serie de reglas internas que permiten que la organización gobierne su propia conducta. Un ejemplo de esto es la guía <i>AI at Google: Our Principles</i> (Pichai, 2018, citado en Covarrubias et al. 2022, p. 39).
Reglas autoimpuestas	Las compañías cumplen voluntariamente con requisitos no vinculantes, establecidos de manera colectiva o reconocidos en las áreas de tecnologías, pero al ser no vinculantes, sus sanciones solo son sociales (Hoffmann-Riem, 2020, citado en Covarrubias et al. 2022, p. 39).
Autorregulación de la empresa	Las reglas que se crean para la empresa persiguen un propósito regulatorio y se han convertido en vinculantes, pero han sido creadas sin la participación de las autoridades, ya sean códigos de conducta o estándares técnicos (Hoffmann-Riem, 2020, citado en Covarrubias et al. 2022, p. 39).
Autorregulación regulada	Las autoridades confían la solución de problemas a los esfuerzos regulatorios autónomos por los miembros de la sociedad, sobre todo de las empresas que se adhieren a la idea del mandato legal, que puede ser motivado por incentivos y materializado en certificaciones o auditorías por parte de organismos privados acreditados (Hoffmann-Riem, 2020, citado en Covarrubias et al. 2022, p. 39).
Regulación híbrida	También conocida como correulación. Sucede cuando el organismo gubernamental y los reguladores privados cooperan en instituciones conjuntas (Kleinstauber, 2004, citado en Covarrubias et al. 2022, p. 40).
Regulación dictada por las autoridades	Se encuentra relacionada con los instrumentos internacionales no vinculantes, conocidos por su carácter orientador o moral. Son una fuente material que influye en la creación de normas de derecho. (Paúl, 2021, citado en Covarrubias et al. 2022, p. 40).
Tecno-regulación	Se refiere al mecanismo por el cual las tecnologías pueden usarse como herramientas para influir en los comportamientos de las personas mediante la implementación de valores, normas y estándares regulatorios en dispositivos tecnológicos (Hoffman-Riem, 2020, citado en Covarrubias et al. 2022, p. 40).

Como se verifica en el cuadro anterior, son muchas las formas de regulación que no tienen una injerencia “directa” o “absoluta” del Estado. Si bien todas estas pueden relacionarse con sistemas de IA, la que resulta más interesante de analizar para los fines del presente trabajo es la de “regulación dictada por las autoridades” con el enfoque de instrumentos

internacionales, toda vez que existen instrumentos internacionales emitidos por instituciones de reconocido prestigio que han servido de insumo a las legislaciones de IA, siendo una de ellas la de nuestro país.

2.2.4 Marco jurídico de la inteligencia artificial en el Perú

El presente apartado tiene como finalidad abordar los avances normativos que se han dado en Perú, respecto al uso de la IA en las actividades de la Administración Pública. Si bien no se pretende realizar un análisis de cada norma promulgada que involucre la utilización de tecnologías como la señalada en las actividades de la Administración Pública, se busca brindar un panorama completo de los cuerpos normativos promulgados que forman parte del régimen jurídico del uso de IA en las actividades de la Administración Pública en el Perú.

a) Antecedentes legislativos

- **Constitución Política del Perú:** Los artículos 4.2° y 14° de la Constitución indican que el Perú impulsa el desarrollo científico y tecnológico y la utilización de tecnologías de la información. Por su parte, el artículo 44° señala que es de vital importancia para el Estado asegurar la vigencia de los derechos humanos, proteger a la ciudadanía de los riesgos contra su seguridad, impulsar el bienestar general basado en la justicia y el desarrollo integral y equilibrado de la nación.
- **Política 35 del Acuerdo Nacional:** Indica en su literal b) que el Perú fomentará el ejercicio y el respeto de los derechos humanos en el entorno digital. Además, menciona en su literal h) indica que el Perú fomentará el uso transversal de las TIC en diversos aspectos, tales como gobierno abierto, defensa nacional, innovación, investigación, transferencia de conocimiento, sectores productivos y

sociales, educación, salud, conservación del ambiente, seguridad ciudadana y prevención de riesgos de desastres.

- **Decreto Legislativo N.º 1412 – Decreto Legislativo que aprueba la Ley de Gobierno Digital:** Publicado el 13 de septiembre de 2018. Establece el marco de gobernanza del gobierno digital para hacer posible una adecuada gestión de la identidad digital, los servicios digitales, la arquitectura digital, la interoperabilidad, la seguridad digital y datos, así como el régimen jurídico aplicable al uso transversal de tecnologías digitales en el proceso de digitalización de procesos y servicios digitales brindados por la Administración. Asimismo, en su artículo 8º se establece que la PCM, a través de la SGTD, es el ente rector en lo que gobierno digital se refiere, lo cual implica las tecnologías digitales, la identidad digital, la interoperabilidad, los servicios digitales, los datos, la seguridad y la arquitectura digital.
- **Decreto de Urgencia N.º 006-2020 – Decreto de Urgencia que crea el SNTD:** Publicado el 9 de enero de 2020. Crea el SNTD, también indica que es la PCM, a través de SGTD, su ente rector. Conforme a ello, este último es la autoridad técnico-normativa a nivel nacional sobre la materia, con objetivos como el impulso de la innovación y el fortalecimiento de una sociedad digital inclusiva, así como el ejercicio de una ciudadanía digital que posea deberes y derechos digitales de los ciudadanos, el fortalecimiento del acceso y la inclusión a las tecnologías digitales en el país y la confianza digital, impulsando la seguridad, la transparencia, la protección de datos personales y la gestión ética de las tecnologías en el entorno digital.

- **Decreto de Urgencia N.º 007-2020 – Decreto de Urgencia que Aprueba el Marco de Confianza Digital y Dispone Medidas para su Fortalecimiento:** Publicado el 9 de enero de 2020. Su artículo 12º establece que la Administración y las organizaciones del sector privado impulsan y garantizan un uso ético de tecnologías digitales, incluyendo el uso intensivo de datos, tales como IoT, IA, ciencia de datos, analítica y procesamiento de grandes volúmenes de datos.
- **Decreto Supremo N.º 157-2021-PCM – Decreto Supremo que aprueba el Reglamento del Decreto de Urgencia N.º 006-2020:** Publicado el 25 de septiembre de 2021. Señala en el literal h) de su artículo 3º que las tecnologías de tipo emergente pueden generar soluciones innovadoras como la robótica, la analítica, la IA, las tecnologías cognitivas, la nanotecnología y el internet de las cosas (IoT), lo cual genera un impacto en el entorno digital, las organizaciones y las personas. Asimismo, en su artículo 18º señala como instrumento para fortalecer la confianza en el entorno digital el Centro Nacional de Innovación Digital e IA, el cual promueve acciones para el despliegue y uso de nuevas tecnologías -en especial la IA- con la finalidad de colaborar con el fortalecimiento de la seguridad y confianza digital en el Perú.
- **Decreto Supremo N.º 029-2021-PCM – Decreto Supremo que aprueba el Reglamento del Decreto Legislativo N.º 1412, Decreto Legislativo que aprueba la Ley de Gobierno Digital, y establece disposiciones sobre las condiciones, requisitos y uso de las tecnologías y medios electrónicos en el procedimiento administrativo:** Publicado el 19 de febrero de 2021. Su reglamentación regula las actividades de gobernanza y gestión de las tecnologías digitales en las entidades de la Administración materia del Gobierno Digital, lo

que implica la identidad digital, la interoperabilidad, los servicios digitales, los datos, la seguridad digital y también la arquitectura digital. Además, establece el marco jurídico aplicable al uso transversal de nuevas tecnologías en el proceso de digitalización presente en la Administración, en línea con el Decreto Legislativo N.º 1412. Asimismo, determina las condiciones, los requisitos y el uso de las tecnologías y medios electrónicos en el procedimiento administrativo, así como los criterios, condiciones, mecanismos y plazos de implementación de la casilla única electrónica, en línea con el Texto Único Ordenado de la Ley N.º 27444, Ley del Procedimiento Administrativo General, aprobado mediante Decreto Supremo N.º 004-2019-JUS (TUO de la LPAG). Por otro lado, dicho reglamento señala lo siguiente en relación con la IA:

- Artículo 75º: Indica en dicho artículo que la SGTD es responsable de emitir los lineamientos, procedimientos, estándares y estrategias para asegurar la calidad, el uso adecuado y el aprovechamiento de datos gubernamentales en materia de IA, entre otros.
- Artículo 120º: Dispone la creación de la Plataforma Nacional de Gobierno de Datos (Datos Perú), la cual comprende, de manera no limitativa, la Plataforma Nacional de Datos Abiertos, la Plataforma GEOPERU y las fuentes disponibles en datos espaciales o georreferenciados. Asimismo, dispone la implementación de tecnologías digitales para la analítica de grandes volúmenes de datos, IA u otras tecnologías emergentes a fin de facilitar la disponibilidad de tableros de gestión y la toma de decisiones e intervenciones estratégicas en la Administración Pública.
- Trigésima Quinta Disposición Complementaria Final: Indica que la PCM a través de la SGTD aprueba una Estrategia Nacional de Gobierno de Datos e IA en un

plazo no mayor a ciento ochenta (180) días hábiles posteriores a la publicación del Reglamento. Señala además que debe actualizarse cada dos años y elaborarse con la participación del sector público, la academia, el sector privado y la sociedad civil.

b) Ley de inteligencia artificial peruana

El 5 de julio de 2023 fue publicada la Ley de IA peruana. El propósito de dicha ley es impulsar el uso de la IA en el marco del proceso de transformación digital, priorizando a la persona y el respeto a los derechos, con la finalidad de promover el desarrollo económico y social del país; ello en un espacio en el que existan garantías sobre su uso ético, sostenible, transparente, replicable y responsable.

La Ley de IA peruana señala en su Título Preliminar que los siguientes son los principios para el desarrollo y uso de la IA:

Literal	Principio	Desarrollo
a)	Estándares de seguridad basados en riesgos	Se promueve un enfoque basado en riesgos para el uso y desarrollo de la IA.
b)	Enfoque de pluralidad de participantes	Se promueve la participación de personas naturales y jurídicas u organizaciones e instituciones públicas y privadas en el debate para el desarrollo de políticas orientadas a la regulación sobre el uso de la IA en el país.

c)	Gobernanza de internet	Se promueve el desarrollo y aplicación de principios, normas, reglas, procedimientos de toma de decisión y programas determinantes para la evolución y el uso de internet por parte del Estado, instituciones del sector privado, y la sociedad civil, participando desde sus respectivos roles.
d)	Sociedad digital	Se valora la información y el conocimiento obtenidos mediante el acceso, uso y desarrollo de tecnologías digitales en todas sus dimensiones, y se impulsa la seguridad, la confianza, la economía digital, la conectividad digital, el talento, la innovación, la educación y la identidad digital, así como el aprovechamiento de las tecnologías emergentes en favor del bienestar social y económico de la ciudadanía.
e)	Desarrollo ético para una IA responsable	Se considera que la ética constituye la base fundamental para identificar de forma precisa el marco de responsabilidades en el uso de estos sistemas que conforman la industria 4.0.
f)	Privacidad de la IA	La inteligencia artificial no debe transgredir la privacidad de las personas, debe operar de manera segura para lograr un impacto positivo y de bienestar en los ciudadanos.

Sobre los principios, nos parece el más importante el de estándares de seguridad basados en riesgos, ya que es el hilo conductor y la esencia de la regulación peruana, que ha seguido esencialmente el estándar europeo materializado a la fecha por el ya presentado Reglamento de IA de la UE.

No obstante lo anterior, indicamos que es “en esencia”, ya que adelantamos que la regulación de la IA a nivel europeo es mucho más precisa y concreta en lo que gestión del riesgo se refiere, como se verá en el siguiente apartado donde se desarrolla un poco lo avanzado en la reglamentación.

Además, la Ley de IA en su artículo 3° muestra las siguientes definiciones:

Literal	Término	Explicación
a)	Inteligencia artificial	Tecnología emergente de propósito general que tiene el potencial de mejorar el bienestar de las personas, contribuir a una actividad económica global sostenible y positiva, aumentar la innovación y la productividad, y ayudar a responder a los desafíos globales clave.
b)	Sistema basado en inteligencia artificial	Sistema electrónico-mecánico que puede, para una serie de objetivos definidos por humanos, hacer predicciones, recomendaciones o tomar decisiones, influenciando ambientes reales o virtuales. Está diseñado para funcionar con diferentes niveles de autonomía.
c)	Tecnologías emergentes	Tecnologías digitales capaces de generar soluciones innovadoras tales como la robótica, la analítica, la

		inteligencia artificial, las tecnologías cognitivas, la nanotecnología, el IoT y similares, que conforman la industria 4.0 que combina técnicas avanzadas de producción y operaciones con tecnología, generando impacto en el ecosistema digital, en las organizaciones y en las personas.
d)	Algoritmo	Secuencia de instrucciones y conjuntos ordenados finitos de pasos para resolver un problema o tomar una decisión.

Sobre estos conceptos, ya han sido realizadas algunas observaciones y críticas en el Capítulo 1, principalmente sobre el término de sistema basado en IA, el cual consideramos que debe integrar la idea de datos de entrada y salida, en armonía con el Reglamento de IA de la UE.

c) **Reglamentación de la ley de inteligencia artificial peruana**

El 2 de mayo de 2024 se publicó la Resolución Ministerial N.º 132-2024-PCM, la cual dispuso el Proyecto de Reglamento de la Ley de IA peruana. Posteriormente, el 19 de noviembre de 2024, la PCM publicó un nuevo proyecto del referido reglamento.

A continuación, un desarrollo de sus principales disposiciones.

- **Objeto:** Tiene por finalidad el establecimiento de normas generales que permitan impulsar el desarrollo y el uso responsable de la IA en el ámbito público como privado. También tiene por finalidad estipular principios y condiciones que impulsen el desarrollo, la implementación y el uso de la IA en un entorno de

respeto a los derechos fundamentales protegidos por la Constitución y los tratados internacionales suscritos por el Perú.

➤ **Ámbito de aplicación:** Las regulaciones del reglamento aplican a las entidades establecidas en el artículo I del título preliminar del TUO de la LPAG y a las empresas del Estado que se encuentran en el Fondo Nacional de Financiamiento de la Actividad Empresarial del Estado (FONAFE), así como las empresas que son parte de gobiernos regionales y locales. También aplica al uso de la IA que realicen las organizaciones de la sociedad civil, ciudadanos, academia y el sector privado que integra el SNTD.

➤ **Definiciones destacadas**

- **Algoritmo:** Secuencia de instrucciones y de conjuntos ordenados y finitos de pasos para resolver un problema o tomar una decisión.
- **Desarrollador:** Toda persona natural o jurídica que desarrolle o programe una aplicación de software sobre IA.
- **Implementador:** Toda persona, ya sea una persona natural o una persona jurídica que despliegue un sistema IA.
- **IA:** Tecnología de tipo emergente con propósito general con potencial de mejorar el bienestar de las personas y la actividad económica global sostenible, aumentar la innovación y la productividad, y ayudar a responder a los retos globales actuales.

- Privacidad de la IA: La IA no debe transgredir la privacidad de las personas en concordancia con la Ley de Protección de Datos Personales y su reglamento.
- Rendición de cuentas: La aplicación de este principio rector se ejecuta a través de los mecanismos implementados en la gestión pública y conforme al ordenamiento nacional vigente.
- Sistema basado en IA: Sistema electrónico-mecánico que puede, para una serie de objetivos definidos por humanos, hacer predicciones, recomendaciones o tomar decisiones, influenciando ambientes reales o virtuales. Está diseñado para funcionar con diferentes niveles de autonomía.
- Transparencia y previsibilidad: Capacidad de anticipar y entender los resultados que los modelos de IA pueden predecir.
- Tecnologías emergentes: Son las tecnologías digitales capaces de generar soluciones innovadoras tales como, la robótica, la analítica, la IA, las tecnologías cognitivas, la nanotecnología, el internet de las cosas (IoT), *Blockchain*, *Cloud Computing*, *big data*, industria 4.0 que combinan técnicas avanzadas de producción con operaciones con tecnología, generando impacto en el ecosistema digital, las organizaciones y las personas.
- Uso indebido: Uso de una aplicación en IA, que no corresponde con la finalidad prevista en su diseño ni con el objeto del presente Reglamento, pudiendo darse de forma intencional o involuntaria.
- Uso indebido razonablemente previsible: Uso inadecuado o ilícito pero que puede deducirse de un comportamiento humano o una interacción con otros

sistemas, incluidos otros sistemas basados en IA, de forma razonablemente previsible.

- Usuario final: Persona natural o persona jurídica que utiliza una aplicación de IA y es diferente del desarrollador y del implementador.

➤ **Principios destacados**

- Centrado en la protección de la dignidad: En todas las fases del ciclo de vida de las aplicaciones de IA se promueven medios que permitan asegurar el respeto, la protección de la dignidad y la autonomía humanas, el Estado de Derecho, la democracia (sus valores) y los derechos humanos.
- Desarrollo ético para una IA responsable: La ética es la base fundamental para identificar con precisión el marco de responsabilidades en el uso de la IA y otras tecnologías emergentes.
- Estándares de seguridad basados en riesgos: Se promueve un enfoque basado en riesgos para el desarrollo y uso de la IA. El enfoque basado en riesgos busca una comprensión profunda de los riesgos específicos y sus implicancias. Se identifican amenazas, vulnerabilidades y posibles consecuencias negativas, así como, la evaluación de la probabilidad de ocurrencia y el impacto asociado a cada riesgo.
- Respeto del derecho de autor y derechos conexos: En todo el ciclo, se deben respetar los derechos de autor, tanto morales como patrimoniales, de los creadores respecto de sus obras originales de acuerdo con la normativa nacional y los tratados internacionales en la materia.
- Seguridad, proporcionalidad y fiabilidad: Las aplicaciones de IA deben ser fiables y seguras durante todo su ciclo de vida para que, en condiciones de

uso, según su finalidad prevista, funcionen adecuadamente y minimicen los riesgos del uso indebido u otras condiciones adversas vinculados a dichos sistemas.

- Sostenibilidad: En todo el ciclo se tienen en cuenta y evalúan los impactos humanos, sociales, económicos, culturales y ambientales, alineados con las políticas nacionales y los ODS. Además, deben ser consideradas las posibles afectaciones ambientales e impulsar soluciones que minimicen el consumo de energía y recursos.

➤ **Obligaciones de las entidades de la Administración Pública:** El proyecto de reglamento señala las siguientes obligaciones que deben cumplir las entidades de la Administración Pública dentro del ámbito de sus competencias.

- Conformar equipos técnicos multidisciplinarios para brindar soluciones tecnológicas basadas en IA.
- Asegurar el cumplimiento de los principios rectores definidos el presente Reglamento en todas las fases del ciclo de vida de una aplicación de IA.
- Generar, recolectar, consolidar y obtener datos y evidencias de las problemáticas públicas que requieran de una posible solución tecnológica basada en IA.
- Brindar e intercambiar información relevante y confiable que sirva de base para la toma de decisiones en el uso de IA.
- Capacitar y desarrollar las habilidades digitales del personal de la Administración en temas relacionados con las aplicaciones de IA.

- Cumplir con las disposiciones establecidas para el uso, adquisición y adecuación del software en las entidades de la Administración Pública.
 - Cumplir con las disposiciones sobre los estándares técnicos para la evaluación de *software* y las disposiciones sobre procesos de su ciclo de vida.
 - Realizar evaluaciones del uso y aplicaciones del *software* de IA.
 - Cumplir con los estándares técnicos sobre gestión de seguridad de la información.
 - Publicar los algoritmos de IA basados en software libre en la Plataforma Nacional de Software Público Peruano.
 - Garantizar que sus servidores y funcionarios cumplen con las normas del Código de Ética de la Función Pública, su Reglamento y las normas de Integridad, en el diseño, desarrollo, implementación y uso de aplicaciones de IA.
- **Gestión de riesgos en el uso de la IA:** El Proyecto de manera similar al modelo de la UE clasifica los riesgos generados en riesgo inaceptable, alto, medio y bajo. Precisa que en el caso de entidades de la Administración Pública los niveles de riesgo deben enmarcarse en los atributos de uso del modelo de calidad de software referido en el Informe Técnico Previo de Evaluación de Software, regulado por la Ley N.º 28612, ley que norma el uso, adquisición y adecuación del software en la Administración Pública, su Reglamento aprobado mediante Decreto Supremo N.º 024-2006-PCM y la “Guía Técnica sobre Evaluación de Software para la Administración Pública”, aprobada mediante la Resolución Ministerial N.º 139-2004-PCM u otra norma que la sustituya.

2.3 Régimen institucional del uso de la IA en las actividades de la Administración Pública

El desarrollo institucional en materia de IA ha sido incipiente, ya que recientemente ha sido publicada la Ley de IA peruana, la cual brinda algunos alcances sobre la organización institucional.

Según el artículo 4° de la referida ley, la autoridad técnico-normativa a nivel nacional responsable de dirigir, evaluar y supervisar el uso y la promoción del desarrollo de la IA y las tecnologías emergentes, a fin de alcanzar los objetivos del país en materia de transformación digital y los ODS conforme a la normativa vigente es la PCM a través de la SGTD.

Asimismo, el mismo artículo señala que las acciones a impulsar por dicha entidad son las siguientes:

- El desarrollo de la IA y su adopción como una herramienta que impulse el desarrollo y el bienestar del país.
- La formación de profesionales con competencia para el aprovechamiento, desarrollo y usos de la IA en el país.
- La creación y el fortalecimiento de la infraestructura digital como habilitadora para el desarrollo de la IA.
- El desarrollo de una infraestructura de datos a fin de poner a disposición datos públicos de alta calidad, reutilizable y accesible.
- La adopción de lineamientos éticos para un uso sostenible, transparente y replicable de la IA.

- Un ecosistema de colaboración de IA a nivel nacional e internacional.

Nótese que las actividades descritas anteriormente implican solo un “impulso” o “promoción”, por lo que los operadores de la Administración y la ciudadanía en su conjunto deberán esperar documentos o pronunciamientos adicionales de la SGTD para poder conocer los siguientes avances sobre la regulación de la IA.

El Proyecto de Reglamento reitera también, en su artículo 5º, que la Presidencia del Consejo de Ministros, a través de la SGTD, como ente rector del SNTD, es la autoridad a nivel nacional encargada de dirigir, evaluar y supervisar el uso y promoción del desarrollo de la IA.

El referido Proyecto de Reglamento indica, además, que los actores de la gobernanza para el diseño, el desarrollo, la implementación y el uso de la IA son:

- La PCM, a través de la SGTD, como ente rector del SNTD.
- La Autoridad Nacional de Protección de Datos Personales del Ministerio de Justicia y Derechos Humanos, la cual es responsable del adecuado tratamiento de los datos personales cuando se utilizan en el diseño, desarrollo, implementación y uso de la IA; así como fiscalizar y sancionar el incumplimiento de la Ley de Protección de Datos Personales y su reglamento.
- Los Comités de Gobierno y Transformación Digital de las entidades de la Administración Pública, responsables de la gestión de actividades y/o proyectos de aplicaciones de IA, en el marco de sus competencias, siguiendo lo dispuesto en el presente Reglamento y las normas complementarias que emita la SGTD.
- Las organizaciones del sector privado, la sociedad civil, la academia, u otros actores relevantes para una sociedad digital, participan activamente mediante

sesiones públicas, mesas de trabajo u otros mecanismos de articulación físicos y/o virtuales que establezcan las entidades de la Administración Pública en el marco de sus competencias, para impulsar el diseño, desarrollo, implementación y uso de aplicaciones de IA, que garanticen el respeto de los derechos humanos, en favor del desarrollo económico y social del país.



CAPÍTULO 3

LA GARANTÍA DE TRANSPARENCIA EN LA REGULACIÓN DEL USO DE INTELIGENCIA ARTIFICIAL DE LA ADMINISTRACIÓN PÚBLICA: DESAFÍOS Y PROPUESTA DE SOLUCIÓN

El presente capítulo tiene por objeto revisar la garantía de transparencia en el marco normativo del uso de tecnologías de IA por parte de la Administración Pública, así como identificar los desafíos que en este contexto nuevo se presentan para el marco normativo peruano, y las propuestas de solución que pueden ser planteadas para asegurar la garantía de transparencia en un contexto en el que las instituciones de la Administración Pública emplean IA.

En dicho sentido, se abordará la garantía de transparencia en el uso de algoritmos por parte de las distintas entidades de la Administración Pública. Así, se buscará brindar alcances sobre la transparencia algorítmica, las tipologías de transparencia, y su tratamiento en el marco normativo peruano sobre transparencia.

Asimismo, se presentará la situación actual o el *statu quo* respecto a la transparencia en el contexto del uso de IA, para lo cual se revisará la experiencia internacional en materia de transparencia algorítmica, así como la experiencia nacional frente a las afectaciones generadas por su falta o deficiencia.

Del mismo modo, se abordarán las dificultades que se presentan para la garantía de transparencia en el contexto del uso de herramientas de IA por parte de las instituciones de la Administración Pública, para finalmente culminar con la propuesta que postula el presente trabajo para garantizar la transparencia en el uso de IA por parte de la Administración Pública en el ordenamiento jurídico peruano.

Es preciso señalar que el mencionado análisis se realizará con base en lo desarrollado en los capítulos precedentes, así como sobre la base de las normas citadas y analizadas.

3.1 La garantía de transparencia en el uso de algoritmos por parte de la Administración Pública

La garantía de transparencia, como es de esperarse, enfrenta nuevos desafíos en un contexto en el que las instituciones de la Administración emplean IA para desarrollar sus actividades. En tal sentido, la transparencia algorítmica cobra un notorio protagonismo, pues las características de estas nuevas tecnologías hacen que el conocimiento y explicación de su funcionamiento se caracterice por presentar cierto grado de oscuridad.

En dicho sentido, el presente apartado busca analizar la garantía de transparencia en el marco del uso de IA por parte de la Administración Pública, para lo cual se brindarán alcances sobre la conceptualización de la transparencia algorítmica, sus tipologías, y cómo es que la normativa sobre transparencia la incorpora.

3.1.1 Definición de transparencia algorítmica

La doctrina ha propuesto diversas definiciones del concepto de transparencia algorítmica, enfatizando distintos aspectos según su enfoque. Sin embargo, convergen en considerar que dicha transparencia algorítmica implica elementos clave como la publicidad activa y el derecho de acceso a la información respecto de las finalidades, programación, estructura, diseño, datos, y demás componentes del algoritmo.

Es preciso señalar que algunos autores vinculan necesariamente a la transparencia con la comprensibilidad del algoritmo, no obstante, el presente trabajo considera que la explicabilidad del algoritmo no es una característica necesaria e imprescindible para que exista transparencia algorítmica en un contexto específico. En ese sentido, la

transparencia puede materializarse incluso cuando la lógica interna del algoritmo no es completamente comprensible, siempre que se garantice el acceso a la información relevante sobre su funcionamiento y propósito.

Tal como lo señala Castellano, la transparencia algorítmica se encuentra próxima a conceptos como la explicabilidad, lo cual puede llevar a confusiones terminológicas. Si bien existen posturas doctrinarias que tienden a equipararlos, es importante distinguir con claridad ambos conceptos. En efecto, es preciso considerar que la explicabilidad busca explicar en términos comprensibles el funcionamiento y resultados de las tecnologías de IA. Por tanto, no es posible realizar una equiparación entre ambos conceptos, ni establecer una relación de causalidad entre ambos, pues la existencia de explicabilidad no implica que exista transparencia algorítmica en un contexto en específico (2023, p. 181).

En dicho sentido, se entiende por transparencia algorítmica la posibilidad de acceder a información sustantiva sobre el funcionamiento, la operación y el rendimiento de los sistemas, herramientas, programas y otros que involucren algoritmos de IA en su funcionamiento (David-Gutiérrez y Castellanos-Sánchez, 2023, p. 8). Al respecto, se hace necesario entender que el objetivo perseguido por la transparencia algorítmica es generar confianza y permitir recurrir o impugnar las decisiones tomadas por algoritmos o sistemas de IA en base a detalles técnicos publicitados con anticipación (Castellano, 2023, p. 181). Así, la transparencia va más allá de la explicabilidad, que pretende brindar explicaciones respecto a una decisión algorítmica determinada y específica.

En concordancia con lo mencionado, la doctrina resalta la utilidad de la transparencia algorítmica para alcanzar objetivos y fines gubernamentales, siendo que la rendición de cuentas por parte de quienes emplean los sistemas de IA, así como de parte de los desarrolladores y adquirientes de estos sistemas, solo es posible en un escenario en el que

la información sobre los algoritmos de IA se encuentra a disposición, con el fin de poder encontrar la explicación de los resultados (David-Gutiérrez y Castellanos-Sánchez, 2023, p. 9).

De igual manera, se entiende que la publicidad de la información sobre los algoritmos es crucial para generar confianza en los administrados, y les permite cuestionar los resultados o decisiones emitidos por las tecnologías de IA (Castellano, 2023, p. 180).

La doctrina considera como premisa fundamental la utilidad de la transparencia algorítmica no solo para conseguir los fines y objetivos que la Administración Pública persigue, sino también su papel trascendental al habilitar el ejercicio de otros derechos.

Así se refieren algunos autores, indicando que la transparencia de los algoritmos permite la actuación de derechos como la protección de los datos personales, la privacidad, el acceso a la información, entre otros. Por un lado, concluyen que este tipo de transparencia algorítmica da pie a que se audite y supervisen los datos de entrada, su procesamiento y cómo son leídos en un sistema de toma de decisiones automatizadas. La transparencia que pueda existir en estos procesos minimizaría el riesgo que muchas veces se presenta en el uso de algoritmos: sesgos discriminatorios, condicionamiento arbitrario de las decisiones de las personas, y el establecimiento de responsabilidad ante resultados no conclusivos, con datos incorrectos o injustos de los algoritmos. Es preciso resaltar que la transparencia también es una herramienta útil para identificar la responsabilidad ante un resultado que merme la cadena de valor en el sistema de un algoritmo público (Contreras y Pefaur, 2020; Diakopoulos, 2020; citado en David-Gutiérrez y Castellanos-Sánchez, 2023, p. 9).

En tal sentido, la transparencia algorítmica abarca la transparencia activa y el derecho de acceso a la información de manera general respecto a los algoritmos que operan las

tecnologías de IA. Esta noción es clave para conocer las fortalezas y debilidades del marco normativo nacional respecto al uso de sistemas de IA por parte de las entidades. A continuación, se analizarán las tipologías de la transparencia algorítmica.

3.1.2 Tipologías de la transparencia algorítmica

Una vez abordado el término de transparencia algorítmica, es preciso revisar las tipologías que se califican como parte de este término. En tal sentido, en primer lugar, se revisará la transparencia algorítmica desde la disponibilidad de la información. En segundo lugar, se revisarán las tipologías de la transparencia algorítmica según la capacidad del sistema de IA, por lo que se abordarán principalmente los conceptos de explicabilidad, auditabilidad y trazabilidad.

a) Desde la disponibilidad de información

El acceso a la información puede ser un efecto de dos tipos de transparencia: la pasiva y la activa. La primera es el resultado de las solicitudes de acceso a información que realizan las personas, a través de derechos de petición o solicitudes de acceso a información pública. La segunda se corresponde con la puesta a disposición del público interesado de información de forma periódica sin que medie un requerimiento de antemano (Porumbescu et al., 2022; citado en David-Gutiérrez y Castellanos-Sánchez, 2023, p. 8).

Así las cosas, la información de algoritmos puede encontrarse accesible debido a que las entidades lo informen en respuesta a solicitudes que hayan sido formuladas (transparencia pasiva) o porque el hecho de que dichas entidades u otras la ofrezcan permanente o periódicamente a través de sus distintos canales, que pueden ser registros públicos de algoritmos. Garrido et al. (2021) dividen a la transparencia algorítmica en clases distintas

pero parecidas a la ya comentada: la jurisdiccional y la registral. Esta última puede producirse porque voluntariamente la Administración Pública brinde la información o porque la ley así lo señala (legislaciones de transparencia y acceso a la información o disposiciones que establezcan obligaciones puntuales de transparencia) (David-Gutiérrez y Castellanos-Sánchez, 2023, pp. 8-9).

Por su parte, Sánchez señala que la transparencia activa opera de cara a los administrados, de forma general, o frente a quien desee acceder a información relativa a los usos de tecnologías de IA y cómo es que estas tecnologías influyen en su esfera jurídica (2024, p. 170).

b) Según la capacidad del sistema de inteligencia artificial

El presente apartado abordará la transparencia algorítmica en sus tipologías según la capacidad que presenta el sistema de IA.

Si bien autores como Castellano presentan como tipologías de la transparencia algorítmica conceptos como la descomponibilidad, la simulabilidad, la auditabilidad, la legibilidad, y la publicidad activa de los resultados, proyecciones, tests y comprobaciones (2023, p. 182); el presente trabajo busca revisar con énfasis los conceptos de explicabilidad, trazabilidad y auditabilidad, que son conceptos ampliamente reconocidos por la doctrina como tipologías de la transparencia algorítmica.

El uso de IA genera una serie de nuevas situaciones que muchas veces no se encuentran previstas por la normativa vigente, y con los cambios que el uso de algoritmos por parte del sector público acarrea, se prevé una mejora en el desempeño de las instituciones públicas, de cara a los ciudadanos (Ortiz de Zárate, 2022, p. 329).

No obstante, esta nueva realidad debe venir acompañada de un marco de desarrollo que siga estándares éticos, técnicos y jurídicos, que salvaguarden los derechos de los administrados en sus interacciones con la Administración Pública.

En dicho sentido, es pertinente revisar conceptos como la explicabilidad, auditabilidad y trazabilidad, los cuales representan estándares para salvaguardar la transparencia en un escenario en el que la esfera jurídica de los administrados puede verse afectada por el contenido de las decisiones que la tecnología empleada por la Administración Pública pueda arrojar.

i) Explicabilidad

La explicabilidad es una garantía directamente relacionada con la transparencia, y ha sido considerada por la Comisión Europea como uno de los cuatro principios éticos en el marco de una IA fiable. En dicho sentido, la Comisión Europea considera en los siguientes términos a la explicabilidad de la IA:

“Esto significa que los procesos han de ser transparentes, que es preciso comunicar abiertamente las capacidades y la finalidad de los sistemas de IA y que las decisiones deben poder explicarse —en la medida de lo posible— a las partes que se vean afectadas por ellas de manera directa o indirecta. Sin esta información, no es posible impugnar adecuadamente una decisión. No siempre resulta posible explicar por qué un modelo ha generado un resultado o una decisión particular (ni qué combinación de factores contribuyeron a ello). Esos casos, que se denominan algoritmos de «caja negra», requieren especial atención. En tales circunstancias, puede ser necesario adoptar otras medidas relacionadas con la explicabilidad (por ejemplo, la trazabilidad, la auditabilidad y la comunicación transparente sobre las prestaciones del sistema), siempre y cuando el sistema en su conjunto respete los derechos fundamentales. El grado de necesidad de

explicabilidad depende en gran medida del contexto y la gravedad de las consecuencias derivadas de un resultado erróneo o inadecuado” (2019, p. 16).

En tal sentido, la explicabilidad busca no solo asegurar que las decisiones tomadas por los sistemas de IA puedan ser explicables, sino también su lógica, finalidades y funcionamiento.

La doctrina no presenta un concepto uniforme respecto a la explicabilidad; no obstante, las definiciones propuestas suelen girar en torno a una premisa base: la importancia de poder explicarle a los ciudadanos la lógica detrás de los criterios empleados en una decisión tomada con sistemas de IA, para lo cual es crucial que la información proporcionada sea comprensible, legible y entendible (Ortiz de Zárate, 2022, p. 334).

En la misma línea, la doctrina también considera que la explicabilidad es útil para la mejora del algoritmo que la adopta, pues los sistemas explicables pueden ser depurados y supervisados de manera óptima (Cotino, 2022, p. 18).

Un concepto que, en opinión de los autores del presente trabajo, reúne los aspectos principales de la explicabilidad es el que proporciona la Unesco en el punto 40° de las Recomendaciones sobre la ética de la IA:

“La explicabilidad supone hacer inteligibles los resultados de los sistemas de IA y facilitar información sobre ellos. La explicabilidad de los sistemas de IA también se refiere a la inteligibilidad de la entrada, salida y funcionamiento de cada componente algorítmico y la forma en que contribuye a los resultados de los sistemas. Así pues, la explicabilidad está estrechamente relacionada con la transparencia, ya que los resultados y los subprocesos que conducen a ellos deberían aspirar a ser comprensibles y trazables, apropiados al contexto. Los actores de la IA deberían comprometerse a velar por que los algoritmos desarrollados sean explicables. En el caso de las aplicaciones de IA cuyo

impacto en el usuario final no es temporal, fácilmente reversible o de bajo riesgo, debería garantizarse que se proporcione una explicación satisfactoria con toda decisión que haya dado lugar a la acción tomada, a fin de que el resultado se considere transparente” (Unesco, 2022, p. 23).

En dicho sentido, resalta la necesidad de garantizar la explicabilidad en un sentido amplio, que abarque la entrada, salida y procesos del sistema de IA; sobre todo en escenarios en los que el resultado del funcionamiento del sistema de IA tiene alguna influencia en la esfera de las personas.

En el escenario que el presente trabajo busca abordar, uno en el que las instituciones de la Administración Pública emplean IA en sus actividades, la explicabilidad de los algoritmos empleados es crucial a fin de garantizar la transparencia, sobre todo en un escenario en el que el resultado del funcionamiento de la IA puede influir en la esfera jurídica de los administrados.

ii) Auditabilidad

Por auditabilidad, se entiende la posibilidad de un algoritmo o sistema de IA de ser evaluado, a fin de que se determine el cumplimiento de sus procesos de diseño, datos, entre otros (Ortiz de Zárate, 2022, p. 334). Como resulta evidente, la auditabilidad también guarda relación con la transparencia, y, por ende, con la explicabilidad del algoritmo.

Con la auditabilidad de los algoritmos, se pretende evitar riesgos inasumibles de manera anticipada, prohibiendo el ingreso al mercado de sistemas de IA que puedan representar contingentes importantes (Fernández-Aller y Serrano, 2022, p. 328). Así, la finalidad detrás de la auditabilidad se remite a la mitigación de riesgos relacionados al uso de IA; no obstante, la auditabilidad puede encontrar barreras al momento de aplicarse a

algoritmos de machine learning que, por su complejidad, presentan espacios oscuros de difícil explicación, o *black boxes*.

A fin de consultar el desarrollo normativo en el extranjero, resalta que la Carta de Derechos Digitales española considera que, en el desarrollo y ciclo de vida de los sistemas de IA, se asegurarán la transparencia, la auditabilidad, la explicabilidad y la trazabilidad.

Como se ha podido apreciar, el concepto de auditabilidad no ha sido dejado de lado por la doctrina, que analiza los riesgos asociados para la transparencia en el uso de IA, sobre todo cuando esta es empleada por la Administración Pública.

iii) Trazabilidad

Respecto de la trazabilidad, la doctrina señala que esta busca identificar el razonamiento detrás de una decisión errónea de un sistema de IA, y prevenir los futuros errores, para lo cual se presenta el deber de realizar una recopilación rigurosa de los datos, procesos y algoritmos empleados (Cotino, 2022, p. 21). Al respecto, resalta que la trazabilidad es relativa, sobre todo, al proceso de toma de decisiones de los algoritmos de IA.

Autores como Corvalán consideran que la trazabilidad implica que cualquier sistema de IA, que considere que mantiene observancia de derechos humanos, debería poder explicar, de inicio a fin, el proceso que desarrolla (2018, p. 312). Sobre ello, nuevamente se encuentran limitaciones al momento de exigir trazabilidad en tecnologías de machine learning, que tienen como característica que su complejidad implica un grado de oscuridad en su funcionamiento.

Es correcto señalar que, así como la explicabilidad y la auditabilidad, la trazabilidad busca salvaguardar aspectos como la rendición de cuentas y responsabilidad en el desarrollo de los sistemas de IA; para lo cual se busca que tanto usuarios como autoridades puedan

comprender el funcionamiento, y, por ende, los impactos que estas puedan tener en la sociedad (Martínez, 2023, p. 12).

Tanto el Reglamento de IA de la UE como la normativa peruana hacen referencia al término transparencia. Sin embargo, solo el Reglamento de IA de la UE hace referencia explícita a las garantías de trazabilidad y explicabilidad cuando define la transparencia:

(...) Por «transparencia» se entiende que los sistemas de IA se desarrollan y utilizan de un modo que permita una trazabilidad y explicabilidad adecuadas, y que, al mismo tiempo, haga que las personas sean conscientes de que se comunican o interactúan con un sistema de IA e informe debidamente a los responsables del despliegue acerca de las capacidades y limitaciones de dicho sistema de IA y a las personas afectadas acerca de sus derechos (...) (Considerando 27).

Sin perjuicio de ello, el Reglamento de IA de la UE contempla disposiciones específicas sobre trazabilidad:

- Para garantizar un nivel de trazabilidad del funcionamiento del sistema de IA de alto riesgo que resulte adecuado para la finalidad prevista del sistema, las capacidades de registro permitirán que se registren los acontecimientos pertinentes señalados en el Reglamento (Artículo 12.2°).
- En relación con la declaración UE de conformidad que es redactada por el proveedor para cada sistema de IA de alto riesgo, señala que deberá contener el nombre y tipo del sistema de IA, y toda referencia inequívoca adicional que permita la identificación y trazabilidad del sistema de IA (Anexo V – Declaración UE de conformidad, numeral 1).
- En relación con la información que debe presentarse para la inscripción en el registro de sistemas de IA de alto riesgo, señala que debe informar y actualizar el

nombre comercial del sistema de IA y toda referencia inequívoca adicional que permita su identificación y trazabilidad (Anexo VIII – Información que debe presentarse para la inscripción en el registro de sistemas de IA de alto riesgo de conformidad con el artículo 49°).

Nótese que las tecnologías de IA presentan un avance acelerado y constante, que necesariamente debe ir acompañado de aspectos como la explicabilidad, la auditabilidad y la trazabilidad, a fin de preservar la transparencia en este nuevo escenario.

3.1.3 La transparencia algorítmica en la regulación de transparencia y acceso a la información pública

En el presente apartado, se busca proporcionar alcances sobre las consideraciones debatidas al momento de regular aspectos como transparencia y acceso a la información pública en el contexto del empleo de algoritmos de IA por las entidades de la Administración Pública.

En dicho sentido, se abordarán las consideraciones que existen respecto a la regulación de aspectos como la publicidad activa, el derecho a la información pública, la titularidad de los derechos, la definición de información pública y las causas de inadmisión; todo ello concebido en el contexto de uso de tecnologías de IA por parte de las entidades públicas.

a) Regulación de la publicidad activa

Es importante traer al presente trabajo la noción de “transparencia regulatoria”. Al respecto, los autores Lodge y Stirton (2001) brindan en su trabajo *“Regulating in the interest of the citizen: Towards a Single Model of Regulatory Transparency?”* un importante aporte sobre la transparencia en regímenes regulatorios. Así, dichos autores

señalan los siguientes mecanismos de transparencia en la regulación: información, elección, representación y voz. Indican que los cuatro (4) mecanismos funcionan como instrumentos de política regulatoria *ex ante* y *ex post*.

Es el mecanismo de transparencia sobre información el que se encuentra directamente relacionado con la idea de publicidad activa. En efecto, los autores antes mencionados indican que una de las razones claves para la regulación es la existencia de asimetrías informativas. Sin información, los consumidores no pueden realizar buenas decisiones. Además, los consumidores bien informados exigen servicios decentes y promueven la competencia de las empresas (2001, p. 108).

Para los fines del presente acápite resulta interesante tomar en cuenta el trabajo de Guichot, quien indica que si bien no hay en la normativa básica estatal de transparencia española ninguna previsión expresa de publicidad de los algoritmos, pues ello no quiere decir que no sean de aplicación las categorías generales (Guichot, 2024, p. 84).

Guichot realiza su trabajo sobre la base de la Ley de Transparencia Acceso a la Información Pública y Buen Gobierno (LTAIPBG) de España. Adelantamos que su trabajo se encuentra inclinado principalmente a analizar el tema de la transparencia algorítmica desde dicha ley; esto es, desde una normativa general y transversal sobre transparencia en el territorio español.

A tales efectos, Guichot menciona que la LTAIPBG prevé como regla general la publicidad activa en los portales de transparencia de información cuyo conocimiento sea relevante para garantizar la transparencia de su actividad relacionada con el funcionamiento y control de la actuación pública, por lo que no cabría duda de que, en el caso del uso de IA, una explicación de su funcionamiento bien puede reputarse de una relevancia tal (Guichot, 2024, p. 84).

No obstante ello, señala que si bien se prevé la publicidad de información de relevancia jurídica, en ella no hay mención específica a los algoritmos, de modo que depende de la naturaleza que se les otorgue (Guichot, 2024, p. 84).

Guichot señala que, en el caso español, las obligaciones de la LTAPBG deben entenderse sin que estas representen afectación a la aplicación de la normativa autonómica correspondiente, o de otras disposiciones normativas que sean aplicables en materia de publicidad de información. Como de hecho sucede en el caso de la Ley 1/2022, de 13 de abril, de transparencia y buen gobierno de la Comunidad Valenciana, que ha incluido sin entrar en categorizaciones jurídicas entre la información legalmente relevante sobre la relación entre los sistemas algorítmicos o de inteligencia artificial y su impacto en los procedimientos administrativos o la prestación de servicios públicos, incluyendo una descripción clara y comprensible de su diseño y funcionamiento, la evaluación de los riesgos que representan y el contacto disponible para consultas en cada caso, todo ello conforme a los principios de transparencia y explicabilidad (Guichot, 2024, pp. 84-85).

Es interesante el ejemplo que coloca el autor Guichot, ya que el Proyecto de Reglamento de la Ley de IA peruana señala como una obligación de las entidades de la Administración Pública publicar los algoritmos de IA basados en *software* libre en la Plataforma Nacional de Software Público Peruano.

Nótese que no hay ningún tipo de anexo que nos sugiera siquiera intuir qué tipo de información sobre el algoritmo de código abierto se encontrará publicada. Sin embargo, de la información disponible web con la que se cuenta por el resto de softwares, podríamos intuir que es bastante básica ya que se colocan datos muy generales como versión, descripción, funcionalidad, especificaciones, entidad que lo usa, contacto, correo y teléfono. Esta problemática será abordada en nuestra propuesta de solución.

b) Regulación del derecho de acceso a la información pública respecto del uso de tecnologías de IA por parte de la Administración Pública

El uso intensivo y generalizado de las TIC en los gobiernos ha traído beneficios en los últimos años y en la coyuntura actual. No obstante, también ha planteado desafíos significativos para el derecho de acceso a la información, ya que han surgido brechas derivadas de las diferencias en el acceso y uso de estas tecnologías entre distintos sectores de la población. Además, la ausencia de un marco normativo que garantice confianza sobre el manejo de la información recopilada ha llevado el debate hacia los dilemas éticos que surgen con la implementación de estas tecnologías por parte del gobierno (Arguelles y Amaro, 2022, p. 92).

El inciso 5 del artículo 2° de la Constitución Política del Perú establece que toda persona tiene derecho a solicitar sin expresión de causa la información que requiera y a recibirla de cualquier entidad pública en el plazo legal, con el costo que suponga el pedido. Señala además la referida disposición que se exceptúan las informaciones que afectan la intimidad personal y las que expresamente sean excluidas por ley o por razones de seguridad nacional.

A continuación, se comentarán algunos aspectos sobre el derecho de acceso a la información, haciendo énfasis en la regulación peruana y la aplicabilidad de ésta a la IA que pudiera utilizar las entidades de la Administración Pública.

i) La titularidad del derecho de acceso a la información pública en el marco del uso de tecnologías de IA por parte de la Administración Pública

La LTAIPBG y sus homólogas en el derecho comparado contemplan una titularidad que podríamos calificar como universal del derecho, la cual se desvincula del criterio del

interesamiento. De este modo, cualquier ciudadano puede ejercer el derecho respecto del acceso a los algoritmos (Guichot, 2024, p. 85).

Así, en el caso peruano debe tomarse en cuenta que este acceso a la información se encuentra regulado por el TUO de la Ley N.º 27806 – Ley de Transparencia y Acceso a la Información Pública, publicado mediante Decreto Supremo N.º 021-2019-JUS.

El artículo 7º del referido TUO indica que toda persona tiene derecho a solicitar y recibir información de cualquier entidad de la Administración Pública y que en ningún caso se exige expresión de causa para el ejercicio de dicho derecho. Por lo que ratificamos que esta es una legitimación a título universal en nuestro ordenamiento jurídico.

La experiencia del Consejo de Transparencia y Buen Gobierno (CTBG) determina que el ejercicio de este derecho en lo que algoritmos se refiere ha sido ejercido principalmente por una organización no gubernamental especializada, la Fundación Civio, lo que desvirtúa la afirmación referida a que lo importante es la inteligibilidad del algoritmo y no tanto el acceso al código fuente mismo (que resulta incomprensible para el común de los ciudadanos), ya que para algunos interesados, como una ONG especializada sí podría serlo (Guichot, 2024, p. 85).

Por otro lado, es importante tomar en cuenta que pedir acceso a un algoritmo utilizado en un determinado procedimiento individual es hacerlo a un documento contenido en el procedimiento (Guichot, 2024, pp. 85-86).

Al respecto, en la normativa peruana lo anterior tiene sus matices, puesto que según el artículo 17.3º del TUO de la Ley de Transparencia y Acceso a la Información Pública el derecho de acceso a la información pública no puede ser ejercicio respecto de la información vinculada a investigaciones en trámite referidas al ejercicio de la potestad

sancionadora de la Administración Pública, en cuyo caso la exclusión del acceso termina cuando la resolución que pone fin al procedimiento queda consentida o cuando transcurren más de seis (6) meses desde que se inició el procedimiento administrativo sancionador, sin que se haya dictado resolución final.

Si bien la anterior excepción se encuentra referida al procedimiento administrativo sancionador, debe tomarse en cuenta que según el artículo 17.2º del mismo cuerpo normativo hay información sobre la cual el derecho de acceso a la información pública no podrá ser ejercido debido a que se encuentra protegida por el secreto bancario, tributario, comercial, industrial, tecnológico y bursátil.

Estas excepciones al ejercicio del derecho en realidad no nos parecen un obstáculo a la garantía de la transparencia algorítmica desde una perspectiva teórica. Señalamos esto porque el interés que consideramos tiene los administrados no es el de la información confidencial en particular de un sujeto, sino más bien en el funcionamiento del algoritmo que usa la entidad de la Administración Pública. Por lo que a nuestro parecer el *quid* de la cuestión se encuentra en establecer cuál es la definición de información pública que aplicaremos.

ii) Definición de información pública y causas de inadmisión en el marco del uso de tecnologías de IA por parte de la Administración Pública

En Europa, las normas de acceso a la información pública tienen por objeto el documento y no a la información. Así, la doctrina española discute si hay relación entre la cuestión de si el concepto de documento engloba o no el acceso a los algoritmos y cuál la naturaleza jurídica de estos (Guichot, 2024, p. 86).

Sobre el papel que tienen los algoritmos en la idea de información pública hay varios enfoques. Por un lado, existe el enfoque que niega la condición de información pública a las aplicaciones informáticas, por tratarse de meras herramientas técnicas que no aportan datos y por tanto no contienen información. Por otro lado, se encuentra el enfoque que señala que sí es información pública, ya que estas herramientas sí tienen un rol configurando o condicionando la decisión administrativa. Existe un tercer enfoque que analiza primero la función del tipo de programa o aplicación en cuestión. Bajo este último, la clave sería comprobar si conocer el código fuente de la aplicación informática utilizada sirve al objetivo de la legislación de transparencia, esto es, al control por parte de la opinión pública de las decisiones adoptadas por las diferentes Administraciones, pues de lo contrario, se trataría tan solo de un instrumento técnico, instrumental, auxiliar, para una más eficaz gestión administrativa, jurídicamente irrelevante. (Guichot, 2024, p. 86).

Sobre los enfoques mencionados, nos parece que el que se adecúa a la normativa peruana de transparencia es el segundo, ya que el artículo 10° del TUO de la Ley de Transparencia y Acceso a la Información Pública indica que la Administración Pública posee el deber de proveer la información requerida por el interesado si se refiere a la contenida en documentos escritos, fotografías, grabaciones, soporte magnético o digital, o en cualquier otro formato, siempre que haya sido creada u obtenida por ella o que se encuentre en su posesión o bajo su control. Además, el mismo artículo precisa que para los efectos de dicha ley, se considera como información pública cualquier tipo de documentación financiada por el presupuesto público que sirva de base a una decisión de naturaleza administrativa, así como las actas de reuniones oficiales.

Como se puede apreciar, la normativa peruana es amplia y si bien no coloca a los algoritmos que usan las entidades de la Administración Pública como objeto explícito del derecho de acceso, ello no significa que puedan ser los mismos considerados.

Sobre la base de lo anterior la discusión pareciera encontrarse en el contenido del algoritmo a ser entregado o en todo caso su “alcance”. Si tomamos en cuenta la legislación de transparencia en los términos en los que se encuentra redactada, vemos que hay una intención de otorgar la información que la entidad disponga respecto del pedido del solicitante.

No obstante, la doctrina es clara al establecer que las administraciones tienen el deber de proporcionar no solo todas las instrucciones sobre el funcionamiento del algoritmo, garantizando que su operatividad sea comprensible para el ciudadano común, sino también el lenguaje informático resultante, conocido como código fuente, del sistema algorítmico (Sassi, 2019; citado en Carloni, 2020, p. 9).

Sobre el código fuente, vale recalcar que el acceso a este brinda la posibilidad de comprobar la conformidad del sistema algorítmico con las previsiones normativas que se pretenden aplicar, además, es la única ventana que permite analizar el razonamiento llevado a cabo por el sistema y detectar eventuales errores. (Carullo, 2021; Citron, 2008; citado en Rkioua, 2023, p. 178).

Nótese que Carloni señala lo anterior pensando en el algoritmo que sirve de base para un acto administrativo resultante de un procedimiento previo. Sin embargo, desde nuestro punto de vista, lo que señala es válido desde el derecho de acceso a la información pública.

3.2 Situación actual sobre la transparencia en el contexto del uso de inteligencia artificial

Existe una razón de fondo que soporta y mantiene relevante un estudio como este que busca garantizar la transparencia sobre el uso de los sistemas de IA por parte de las

entidades de la Administración Pública y es el riesgo que generan dichas tecnologías al ser inherentemente “opacas”.

Cuando se habla de opacidad de algoritmos, se habla de la falta de transparencia motivada por la existencia de una especie de caja negra que carece de capacidad explicativa y dificulta su correspondiente entendimiento. Así, se desconoce por completo su funcionamiento, pero la experiencia permite constatar que cuando el algoritmo adopta una resolución, esta se convierte en una especie de resultado ante el cual no es posible objetar ni recurrir, a pesar de que esa decisión haya podido alimentarse de datos parciales, eventualmente incorrectos o insuficientes, o que han podido ser malinterpretados, sin revelar nada a cambio. Por tanto, cuando un cliente solicita un préstamo al banco, los responsables de la gestión deberían explicar por qué motivos rechazan concederlo (Felzmann, H., et al., 2020; citado en Blázquez, 2022, p. 267).

La opacidad algorítmica resalta la importancia de dotar a los sistemas de aprendizaje profundo de significado y capacidad explicativa. Para ello, es fundamental incorporar módulos que clarifiquen el proceso mediante el cual los algoritmos llegan a sus resultados y conclusiones. Esto se debe a que la capacidad de explicación es una característica esencial e innegociable de cualquier sistema verdaderamente inteligente (López de Mántaras, 2019; citado en Blázquez, 2022, p. 268).

Es importante señalar que la opacidad algorítmica puede originarse por tres razones no excluyentes: (1) puede ser intencionalmente provocada por los desarrolladores o implementadores del sistema, por ejemplo, para proteger secretos comerciales; (2) puede derivarse de la brecha digital, que dificulta la comprensión de la información por parte del receptor; y (3) puede ser intrínseca al sistema, debido a que opera a través de

instrucciones en lenguajes matemáticos complejos. (Garrido et al., 2021, p. 21; citado en David-Gutiérrez y Castellanos-Sánchez, 2023, p. 9).

Es importante considerar que la falta de transparencia u opacidad no es el único riesgo asociado a estas tecnologías, sino que existen diversos peligros derivados de sus características inherentes. No obstante, la opacidad agrava estos problemas. Entre ellos se encuentran: la vulneración de la protección de datos personales, cuando los titulares desconocen las finalidades para las cuales se utilizan sus datos; la discriminación algorítmica, que ocurre cuando los sesgos de la realidad se replican o amplifican en el entorno digital; y la afectación al debido proceso, cuando un sistema algorítmico influye en una decisión sobre una persona investigada sin que esta pueda conocer plenamente la justificación de la determinación estatal que la afecta (Contreras y Pefaur, 2020; Gutiérrez, 2020; UNESCO, 2023b; citado en David-Gutiérrez y Castellanos-Sánchez, 2023, p. 9).

A pesar de lo señalado, si nos enfocamos en la falta de transparencia y partimos de la premisa de que estas tecnologías son inherentemente opacas, es relevante destacar que instituciones como el Parlamento Europeo fomentan la transparencia y la rendición de cuentas en el uso de algoritmos. En este sentido, se han establecido obligaciones de transparencia algorítmica en virtud del derecho de acceso a la información. Existen diversas formas de garantizar dicha transparencia, entre ellas: el derecho de acceso a la información pública, la obligación de motivar las decisiones, la prohibición de la arbitrariedad, el derecho a una buena administración y las garantías del debido proceso, entre otras (Castellanos Claramunt, 2023, p. 283).

3.2.1 Experiencia internacional frente a la transparencia algorítmica

El presente apartado abordará las experiencias internacionales referentes a la transparencia algorítmica activa y pasiva. Con dicho fin, se revisarán las experiencias con repositorios de algoritmos públicos (transparencia algorítmica activa), y las experiencias internacionales relacionadas a casos límite de la transparencia algorítmica (transparencia algorítmica pasiva), para lo cual se revisarán los casos de los datos personales, la contratación pública y el secreto fiscal o reserva tributaria, frente a la garantía de transparencia.

a) Transparencia activa: Repositorios públicos de algoritmos

Tal como se ha abordado a lo largo del presente trabajo, resulta evidente que la transparencia activa es deseable en un contexto como en el que nos encontramos, en el que las tecnologías de IA se encuentran cada vez más presentes, incluso en las actividades de las entidades de la Administración Pública.

Respecto de lo señalado, poner a disposición de los ciudadanos información relativa a los algoritmos y sistemas empleados por la Administración Pública puede darse de distintas formas. Al día de hoy, la experiencia internacional ha mostrado preferencia por adoptar medidas como la creación de repertorios que recopilen datos estructurados en formato de archivos planos que se traduzcan en documentos descargables disponibles en línea, fáciles de consultar en medios digitales (David-Gutiérrez y Castellanos-Sánchez, 2023, p. 13).

Nombre del repositorio	Entidad a cargo	Disponibilidad	No. de sistemas registrados
Dashboard de seguimiento del Marco Ético de Inteligencia Artificial	DAPRE	Inaccesible desde mayo de 2023	6
Portal de Datos Abiertos	MinTIC	Accesible a noviembre de 2023	16
Ejercicios de Innovación Basados en Inteligencia Artificial	MinTIC	Accesible a noviembre de 2023	6

Nota. Adaptado de Gutierrez y Muñoz-Cadena (2023), por David-Gutiérrez y Castellanos-Sánchez, 2023.

Con el fin de citar experiencias en otras legislaciones, resalta el trabajo realizado en Colombia, donde se ha optado por recopilar los datos relativos a los algoritmos y sistemas de IA, y ponerla a disposición de los administrados a través de repositorios disponibles en línea. La mencionada medida fue implementada por el Departamento Administrativo de la Presidencia de la República (Dapre) y el Ministerio de Tecnologías de la Información y las Comunicaciones (MinTIC).

La implementación del señalado repertorio de dio en un contexto en el cual Colombia buscaba generar las condiciones necesarias para incluir herramientas de IA en las actividades de las instituciones pública, tal como lo señalan David-Gutiérrez y Castellanos-Sánchez:

“El DAPRE puso en marcha un repositorio de sistemas de IA en el 2022 con el fin de contribuir a hacer seguimiento al cumplimiento de los principios y recomendaciones del “Marco Ético de Inteligencia Artificial”. El Marco Ético había sido publicado por el DAPRE en octubre de 2021, con el apoyo de consultores externos financiados por la CAF – Banco de Desarrollo de América Latina y el Caribe. Su objetivo es Proporcionar una guía de *soft law* para entidades públicas de recomendaciones y sugerencias, para abordar la formulación y gestión de los proyectos que incluyan el uso de IA...El repositorio denominado “Dashboard de Seguimiento del Marco Ético de Inteligencia Artificial” estaba disponible en una plataforma en línea y la información era cargada por las propias entidades públicas encargadas de los sistemas. Para guiar a los funcionarios en dicha tarea, la “Consejería Presidencial para la transformación digital y gestión y cumplimiento” publicó junto con el Banco Interamericano de Desarrollo un “Manual de usuario formulario proyectos de inteligencia artificial (tablero de seguimiento marco ético)” de 37 páginas. La visualización de los datos de cada uno de los sistemas de IA operaba a partir de la herramienta *Power BI*. Sin embargo, la administración del presidente Gustavo Petro eliminó la “Consejería Presidencial para la transformación

digital y gestión y cumplimiento” a comienzos de 2023, que apoyaba las actividades asociadas al marco ético, y desde mayo la información dejó de ser accesible. Una ex funcionaria del DAPRE, que solicitó preservar su anonimidad, nos indicó que el gobierno no había renovado la licencia *Power BI*, el servicio de *Microsoft* usado para visualizar los datos, razón por la cual la información dejó de estar disponible públicamente.” (David-Gutiérrez y Castellanos-Sánchez, 2023, p. 13).

La citada experiencia no es la única con la que hoy en día se cuenta. No obstante, el impulso para este tipo de medidas no siempre viene desde el sector público, sino desde el privado. En Chile, el GobLab de la Universidad Adolfo Ibáñez ha desarrollado el Repositorio de Algoritmos Públicos, una plataforma que recopila y pone a disposición información sobre sistemas de soporte o de toma de decisiones automatizadas implementadas por instituciones públicas chilenas.

Dentro de las características que presenta el citado repositorio, resalta la exploración de algoritmos, pues la plataforma permite a los usuarios explorar y descargar información detallada sobre diversos algoritmos empleados por la Administración Pública. Asimismo, en el marco de esta iniciativa, el GobLab publica informes anuales que analizan el uso e implementación de los sistemas de IA que acopla la Administración Pública. El informe anual del año 2024 presentó un análisis de 101 sistemas automatizados y semiautomatizados empleados por la Administración Pública. De igual forma, emplean una metodología transparente: brinda información detallada sobre la metodología utilizada para recopilar y clasificar la información sobre algoritmos, a fin de asegurar la confiabilidad y transparencia.

El señalado repertorio se implementó con el fin de otorgar visibilidad a los algoritmos de IA que se involucran en el ámbito público chileno, a fin de lograr un panorama con una mejor gobernanza y apertura de datos.

Los dos ejemplos mencionados ilustran una manera práctica de promover la transparencia activa, recopilando datos que no solo son de utilidad para los administrados, quienes tienen derecho de contar con alcances sobre la información de los algoritmos de IA; sino para los operadores de la Administración que emplean estas herramientas. Si bien la mencionada no es la única medida que se puede emplear para promover la transparencia activa, es el ejemplo que el presente trabajo desea proponer para este fin.

b) Transparencia pasiva: Casos límite para la transparencia algorítmica

En el presente apartado, se abordarán tres casos límite para la transparencia algorítmica: los datos personales y la propiedad intelectual, la contratación pública, y el secreto fiscal o reserva tributaria. Dicha regulación permitirá revisar la experiencia internacional con respecto a la transparencia algorítmica pasiva.

i) Datos personales y propiedad intelectual

Sobre este tema resulta interesante traer a colación el caso de CoronApp. CoronApp era una aplicación móvil que lanzó el gobierno nacional de Colombia durante la pandemia del COVID-19 y que –según el gobierno– tenía la intención de “impactar positivamente en la gestión de la pandemia” (Venturini, 2021; citado en David-Gutiérrez y Castellanos-Sánchez, 2023, p. 14).

Con fecha 6 de agosto de 2020, Juan Carlos Upegui-Mejía presentó una solicitud de información pública ante la Agencia Nacional Digital (AND) con el fin de tener acceso al código fuente de CoronApp o subsidiariamente al histórico de sus versiones. Su solicitud fue denegada, indicándose que la información requerida estaba sujeta a reserva, indicándose que la publicación del código podía comprometer la seguridad de los datos personales de los usuarios. El solicitante interpuso un recurso de reposición explicando

que no estaba solicitando los datos personales, sino el código fuente. Al respecto, la AND confirmó su decisión inicial argumentando que permitir el acceso al código fuente podría generar riesgos en materia de ciberseguridad y chocar con derechos de propiedad intelectual.

El solicitante presentó una acción de tutela en contra de la AND para salvaguardar su derecho de acceso a la información pública. Sin embargo, los jueces de primera y segunda instancia no accedieron a tutelar los derechos invocados por Upegui, quien solicitó la revisión de su caso a la Corte Constitucional.

El planteamiento de Upegui dirigido a la Corte Constitucional partió de categorizar el código de fuente de CoronApp como información pública utilizada por el gobierno nacional para la gestión y control epidemiológico del COVID-19. Al respecto, señaló que el código fuente de una aplicación de una entidad pública es información pública, más aún cuando es un instrumento de política pública que guarda interacción con la ciudadanía. Indicó además que el código fuente es el equivalente en tecnologías digitales a los protocolos de atención de las administraciones públicas (Upegui-Mejía, 2020; citado en David-Gutiérrez y Castellanos-Sánchez, 2023, pp. 14-15).

Finalmente, señaló que denegar el acceso al código fuente es denegar el conocimiento de cómo se está realizando parte de la gestión epidemiológica de la COVID-19 (ocultando un protocolo) y denegar la información relativa al tratamiento de nuestros datos personales (ignorando la normativa de habeas data). Preciso que la apertura del código fuente era el único modo de permitir que la comunidad científica colombiana y la sociedad civil tengan la posibilidad de identificar posibilidades de mejora y ayudar al desarrollo y buen funcionamiento de CoronApp (Upegui-Mejía, 2020; citado en David-Gutiérrez y Castellanos-Sánchez, 2023, p. 15).

El caso CoronApp fue resuelto a favor del administrado, lográndose no solo que se le entregue el código fuente de la aplicación CoronApp, con todo el historial y control de versiones de dicho código fuente desde su versión original hasta el momento en que se tuvo uso, sino que además se ordenó a la AND y a las oficinas designadas por la Procuraduría General de la Nación y a la Defensoría del Pueblo que expidan lineamientos que permitan la maximización de los estándares de transparencia, confianza y acceso a la información pública.

Dichos lineamientos deberían señalar, según la Sentencia T-067 de 2025, como mínimo, las garantías de protección de la transparencia algorítmica en su faceta pasiva y las obligaciones del Estado para asegurar que en los próximos desarrollos de herramientas digitales se garantice la faceta activa de este principio, al disponer de forma abierta y periódica de información que permita comprender el diseño y funcionamiento de los algoritmos públicos.

En la experiencia española se ha señalado que hay una “suerte desigual” al solicitar información sobre algoritmos públicos. En lo que propiedad intelectual se refiere, se tiene el caso Bosco como emblemático, en donde fue denegado hasta en tres (3) ocasiones el acceso al código fuente del algoritmo utilizado para conceder un bono social eléctrico. En este caso se habrían priorizado argumentos como la seguridad informática y la protección de la propiedad intelectual, sin realizar una ponderación adecuada con el principio de transparencia ni con los derechos de los administrados, lo que consolida un precedente negativo que debilita el control democrático sobre los sistemas automatizados

utilizados por la Administración Pública y refleja una tendencia preocupante en la que las excepciones legales se convierten en barreras sistemáticas al acceso (Cotino, 2024)¹⁰.

ii) Contratación pública

Las organizaciones de madres comunitarias locales protestaron en contra del Instituto Colombiano de Bienestar Familiar (ICBF) en 2021, señalando supuestas irregularidades de contratación asociadas a la herramienta BETTO (WRadio, 2021; citado en David-Gutiérrez y Castellanos-Sánchez, 2023, p. 16).

BETTO fue implementado por el ICBF en el año 2020 con el objetivo de garantizar transparencia, objetividad y excelencia en la contratación de los servicios de primera infancia. El algoritmo seleccionaba y evaluaba a los mejores operadores en rankings, con base en los criterios objetivos que habían sido establecidos por el ICBF. BETTO era usado para calificar a los operadores que previamente habían prestado servicios contratados por el ICBF, como las asociaciones de madres trabajadoras. (ICBF, 2021; citado en David-Gutiérrez y Castellanos-Sánchez, 2023, p. 16).

Un año después de su implementación, en 2021, la preferencia por la elección de operadores ajenos al territorio de la prestación del servicio, supuestamente asociados a dinámicas clientelistas, con poca experiencia y/o trayectoria en el campo, generó preocupaciones en torno al funcionamiento de BETTO. Al poco rato iniciaron protestas nacionales en lugares como La Guajira, Bucaramanga y Valledupar, donde las organizaciones locales de madres comunitarias alegaban cumplir con los criterios objetivos (a pesar de lo indicado por los resultados de BETTO); además, en otros casos

¹⁰ El autor se refiere a la sentencia de 30 de abril de 2024, de la Sala de lo Contencioso-Administrativo, Sección Séptima, de la Audiencia Nacional, en el recurso de apelación 51/2022.

afirmaban que llevaban varios años prestando el servicio (Colmenares, 2021; Pineda, 2021; Reyes, 2021; citado en David-Gutiérrez y Castellanos-Sánchez, 2023, pp. 16-17).

Los problemas continuaron y fue en noviembre de 2023 que el ICBF anunció que dejó de usar BETTO tras haber realizado una auditoría de control interno al sistema, cuyos hallazgos fueron el generar dudas importantes en la transparencia y los resultados del proceso de selección a través de la plataforma (David-Gutiérrez y Castellanos-Sánchez, 2023, p. 17).

Resulta interesante como es que un caso como el retratado ha necesitado que una parte de la sociedad civil organizada genere en entidades de la Administración Pública acciones internas que hagan posible la rendición de cuentas sobre un sistema algorítmico.

iii) Secreto fiscal o reserva tributaria

Tanto el acceso a la información como la transparencia tributaria han sido un tabú en España, lo que ha permitido a su Administración Tributaria la posibilidad de delimitar su sistema tributario de forma muchas veces discrecional, apoyándose en la existencia del secreto fiscal. Un ejemplo es la Resolución 825/2019, de 13 de febrero de 2020, del Consejo de Transparencia y Buen Gobierno, la cual se pronunció acerca de una petición de información de un ciudadano que solicitaba a la Agencia Estatal de Administración Tributaria (AEAT) que le facilitara datos relativos a las aplicaciones, contenido, función, etc., que utilizaba, declarando la insuficiencia y carácter genérico de la respuesta proporcionada por la Administración tributaria. La reclamación fue estimada y se instó a la AEAT a que remitiese el «contenido y función de los programas/aplicaciones “ZÚJAR, TESEO, INEX, INTER, DÉDALO, PROMETEO y GENIO” (Sánchez, 2024, p. 174).

En España, la actividad de gestión del riesgo fiscal de incumplimiento de las obligaciones tributarias se lleva a cabo mediante la aplicación HERMES que produce informes de riesgo estandarizados y que permite la selección de contribuyentes para llevar a cabo actuaciones de gestión, inspección y recaudación. La aplicación se nutre de otra llamada ZUJAR que facilita el tratamiento de la información mediante la verificación cruzada de datos de contribuyentes para desarrollar actuaciones en los diferentes procedimientos de aplicación de los tributo. (Soto, 2021, p. 98).

Ahora bien, si bien el funcionamiento de las aplicaciones antes mencionadas puede conllevar, en ciertos supuestos, un riesgo de vulneración del derecho a la intimidad y del derecho a la protección de los datos personales de los contribuyentes; estos deben ser configurados de una manera adecuada para garantizar la protección de estos derechos (Soto, 2021, p. 100).

No obstante ello, de cara a los administrados, la garantía de la transparencia en el ámbito tributario representa un reto en lo que algoritmos se refiere, ya que es necesario que se regule y prevea una explicación clara y precisa del modo en que se utiliza la información para elaborar los algoritmos y los perfiles de riesgos, así como la necesidad de su motivación y la exteriorización de las razones que sirvieron de base al acto emitido por la Administración Tributaria, siendo ello una expresión del principio de proporcionalidad y transparencia (Sánchez, 2024, p. 173).

3.2.2 Experiencia nacional respecto de la transparencia algorítmica

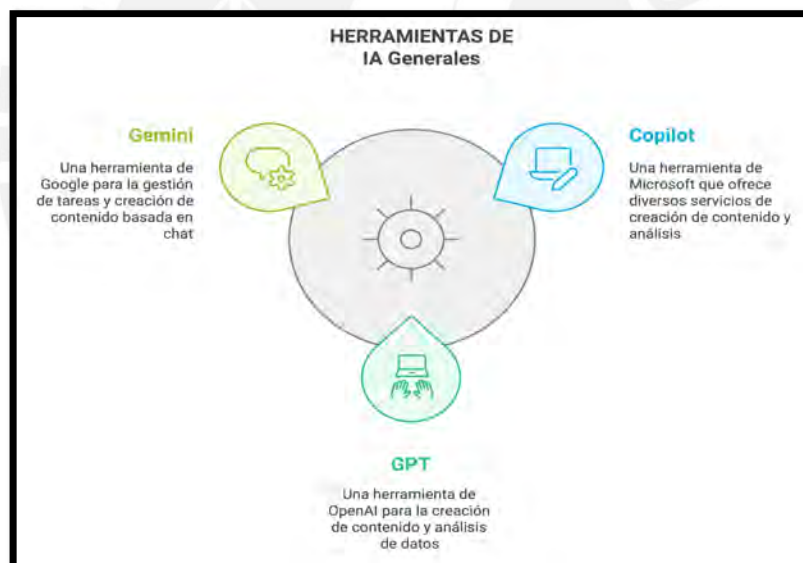
El presente apartado revisará la experiencia nacional respecto a la transparencia algorítmica, para lo cual se revisarán la situación actual respecto de la implementación de programas o sistemas de IA para el uso de estas por parte de las entidades de la Administración Pública. Asimismo, se abordará la situación de la garantía de

transparencia en el contexto del uso de herramientas, sistemas o algoritmos de IA por parte de la entidades de la Administración Pública, para lo cual se analizarán los alcances que existen al respecto en el marco normativo de transparencia.

a) Implementación de sistemas de inteligencia artificial para el uso de las entidades de la Administración Pública del Perú

En el caso del Perú, la Administración Pública, se tiene registro de que se viene utilizando y ha implementado los siguientes aplicativos de IA para optimizar su sistema de gestión, los cuales pueden ser clasificados de la siguiente manera:

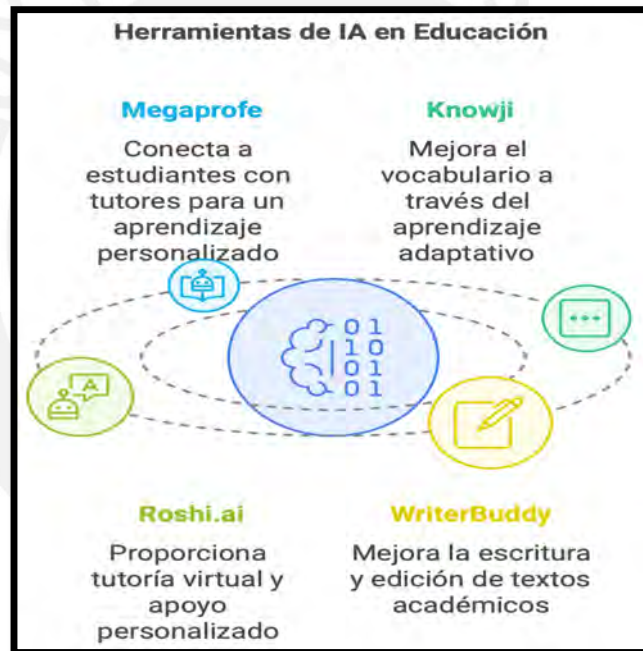
➤ **IA generales**



➤ **IA para diseño y creatividad**



➤ **IA en Educación**



➤ **IA en Documentación y Administración**



Asimismo, se han implementado los siguientes programas en los que se incorpora IA:

- **SIAMA**

El Sistema Integrado Inteligente para registrar, informar, alertar, monitorear y asistir a las personas sintomáticas de COVID-19 (SIAMA) es una herramienta de IA desarrollada en Perú durante los años de pandemia de COVID-19, con el fin de gestionar y monitorear individuos que presenten síntomas de la enfermedad.

SIAMA utiliza técnicas de DL, e integra un chatbot para interacción con el usuario. El sistema registra la información brindada por el usuario, y genera alertas automatizadas, para conocer la evolución de los síntomas, y notifica al personal de salud más cercano en caso de ser necesario.

La implementación práctica de SIAMA se realizó en la Clínica Alivari de la Universidad Católica de Santa María en Arequipa.

- **Perú en tus manos**

Perú en tus manos es un aplicativo móvil desarrollada por el gobierno peruano en 2020 como parte de la estrategia digital para hacerle frente a la pandemia de COVID-19. Esta

herramienta incorporó componentes de IA para brindar a los ciudadanos en tiempo real información sobre zonas de riesgo y facilitar el seguimiento de casos.

La aplicación mostraba mapas de calor que identificaban, a nivel de manzana, las áreas con mayor probabilidad de contagio. De igual forma, brindaba la posibilidad de realizar un triaje digital, mediante el cual los usuarios podían realizar una autoevaluación de síntomas mediante un cuestionario integrado. La aplicación también permitía realizar seguimiento de los casos reportados; y, mediante un trazador de contactos, empleando tecnologías de GPS y Bluetooth, podía estimar la probabilidad de que un usuario hubiera estado en contacto cercano con una persona infectada.

La aplicación contaba con una política de privacidad que establecía el consentimiento de los ciudadanos para el uso de sus datos personales, la cual exponía la finalidad del tratamiento de información, garantizando la eliminación de los datos recibidos una vez culminada la emergencia sanitaria.

- **MAIA**

MYPE Asesor IA (MAIA) es una herramienta de IA desarrollada en el Perú para brindar asesoría personalizada y gratuita a las micro y pequeñas empresas. Esta herramienta emplea modelos avanzados de lenguaje, e interactúa con emprendedores a través de WhatsApp, ofreciendo consejos y soluciones adaptadas a las necesidades específicas de cada negocio. El proyecto es una iniciativa de Hacer Perú, una organización dedicada a mejorar el desempeño y productividad de las MYPES en el Perú, y cuenta con el apoyo de donantes internacionales.

- **ChatBot contra el dengue**

ChatBot contra el dengue es una herramienta implementada por el Seguro Social de Salud del Perú. Se trata de un asistente virtual, accesible a través de WhatsApp, y tiene como objetivo prevenir y detectar casos de dengue en todo el país.

Se trata de una herramienta accesible, que brinda información educativa respecto al dengue, y permite que los usuarios realicen una autoevaluación de síntomas para identificar posibles casos de dengue, y alerta sobre cuándo se presume que es necesario acudir a un establecimiento de salud lo antes posible.

La citada herramienta se reviste de relevancia, en un país que presenta indicadores considerables de contagio de dengue, en varias regiones de su territorio. Asimismo, resalta por su labor educativa, en la medida en que proporciona información en circunstancias en las que el acceso a información verás al respecto puede ser limitada por factores sociales y geográficos.

- **Open Clinic**

Open Clinic es una plataforma de código abierto diseñada para la gestión hospitalaria y la implementación de historias clínicas electrónicas. Esta herramienta fue implementada por el Gobierno Regional de Cajamarca en hospitales como el Hospital Simón Bolívar y el Hospital General de Jaén.

Dentro de los beneficios que se presentan con el uso de Open Clinic, resaltan su accesibilidad y bajo costo, al ser una plataforma de código abierto que no requiere incurrir en gastos por licencias.

- **Ayni Lab Social**

Finalmente, es preciso mencionar el laboratorio de innovación social del Ministerio de Desarrollo e Inclusión Social del Perú, creado en el 2016. Esta herramienta tiene como objetivo el promover las innovaciones costo efectivas en políticas sociales, con el propósito de mejorar la calidad de vida de la población.

Esta plataforma incluye tecnologías avanzadas, dentro de las cuales resaltan las herramientas de IA, que se emplean para el diagnóstico no invasivo de anemia en niños, pues analiza imágenes de la conjuntiva ocular tomadas con un Smartphone para estimar los niveles de hemoglobina.

De igual modo, en materia de seguridad, el Sector Interior ha implementado los siguientes sistemas con IA:



Tal como se puede observar, el uso de programas o sistemas de IA dentro de la Administración Pública es un hecho que trasciende en la transformación digital de todos los gobiernos. Las aplicaciones, programas y sistemas identificados, evidencian la progresiva implementación por la Administración Pública en el Perú de herramientas con IA como parte de sus procesos ordinarios.

b) Garantía de transparencia en el uso de inteligencia artificial por parte de la Administración Pública del Perú

El presente apartado busca brindar alcances respecto de las garantías de transparencia en el uso de IA que se conciben en el marco normativo de IA en el Perú, que claramente le son aplicables a un contexto en el que se emplean herramientas o sistemas de IA por parte de las entidades de la Administración Pública.

En dicho sentido, se analizará el marco normativo compuesto por la Ley de IA peruana así como el Proyecto de Reglamento de dicha ley.

i) Principios de transparencia en el marco normativo de inteligencia artificial peruano

Como primera garantía de transparencia presente en la normativa sobre uso de IA, se debe mencionar la inclusión del principio de transparencia. En dicho sentido, el artículo 4° de la Ley de IA peruana establece que es PCM la autoridad encargada de evaluar, supervisar y dirigir el uso de IA; lo cual debe ser analizado sin perder de vista que el artículo IV del Título Preliminar de la Ley N.º 29158, Ley Orgánica del Poder Ejecutivo, establece el principio de transparencia en la gestión del Poder Ejecutivo. En tal sentido, se debe entender que el principio de transparencia rige también en el contexto en el que se emplean tecnologías de IA.

Asimismo, el artículo 1° de la Ley de IA peruana dispone que el abordado cuerpo normativo busca garantizar un entorno seguro, ético, sostenible y transparente, confirmando que en el uso de tecnologías de la IA es plenamente aplicable el principio de transparencia.

Finalmente, es meritorio mencionar que la Ley de IA peruana promueve en su artículo 4° el desarrollo de infraestructura de datos públicos de alta calidad, accesibles y reutilizables, contribuyendo con la premisa de datos abiertos y accesibilidad.

ii) Obligaciones de la Administración Pública en el marco normativo de inteligencia artificial peruano

En segundo lugar, es preciso reparar en las obligaciones de las entidades públicas que concibe el marco normativo vigente, en pro de la transparencia. En tal sentido, el artículo 4° de la Ley N.° de IA peruana postula que la PCM, a través de la SGTD, supervisa y evalúa el uso de IA. Asimismo, el artículo 5° del mismo cuerpo normativo establece que la PCM, a través de la SGTD, remite un informe anual al Congreso de la República sobre la implementación de la IA. Así, se identifica la obligación de supervisión y rendición de cuentas para la Administración Pública.

En el mismo sentido, el artículo único del Título Preliminar de la Ley de IA peruana establece como principios para el desarrollo y uso de la IA el desarrollo ético para una IA responsable, por el cual se deben establecer lineamientos éticos para garantizar el uso sostenible, transparente y replicable; y la privacidad de la IA, por el cual no se debe transgredir la privacidad de los administrados. En dicho sentido, es obligación de la Administración Pública el respetar los lineamientos éticos y de seguridad que impone el marco normativo vigente.

Finalmente, es oportuno mencionar que el artículo único del Título Preliminar de la Ley de IA peruana establece como principio para el desarrollo y uso de la IA los estándares de seguridad basados en riesgos. Asimismo, el artículo 5° de la citada norma dispone que, de identificarse amenazas graves o vulneración de la ciberseguridad nacional, la SGTD debe informar inmediatamente a la Comisión de Inteligencia del Congreso de la

República. En tal sentido, se identifica la obligación para la Administración Pública de prevención de riesgos y control de impactos en el uso de IA.

iii) Mecanismos de transparencia específicos en el marco normativo de inteligencia artificial peruano

Al respecto, se debe tener en consideración que la Ley de IA peruana establece mecanismos de transparencia específicos que, si bien resultan beneficiosos, son inexactos, en un contexto en el que se requiere normativa que establezca mecanismos de transparencia de forma explícita y concreta.

En tal sentido, resalta el artículo 4º, literal d) de la Ley de IA peruana, mediante el cual se dispone que la SGTD debe promover el desarrollo de infraestructura de datos a fin de poner a disposición datos públicos de alta calidad, reutilizables y accesibles. Así, se advierte una alusión a lo que podría entenderse como la publicación obligatoria de información sobre algoritmos y modelos de IA.

Asimismo, el artículo único del Título Preliminar de la citada norma, en su literal e), establece que la ética es clave para identificar responsabilidades en el uso de IA. Así, se reconoce una alusión a la exigencia de explicabilidad en las decisiones automatizadas, sobre todo en el marco del uso de tecnologías de IA por parte de la Administración Pública.

Finalmente, el artículo 4º de la Ley de IA peruana establece, en su literal f), dispone que la SGTD debe impulsar un ecosistema de IA a nivel nacional e internacional. De dicha premisa puede advertirse una alusión a la implementación de auditorías y control interno de los algoritmos y sistemas de IA.

3.3 Desafíos para la garantía de transparencia en el uso de la inteligencia por parte

de la Administración Pública en el Perú

En el presente apartado, se pretende abordar los desafíos a nivel normativo que se advierten para la transparencia en el Perú, en un contexto en el que las entidades de la Administración Pública emplean sistemas de IA en sus actividades. Los desafíos señalados, en un momento posterior, sustentarán la necesidad de implementar medidas a fin de promover la transparencia en el contexto mencionado.

- **Opacidad algorítmica y explicabilidad**

El primer desafío que se enfrenta al momento de regular el empleo de herramientas de IA por parte de las entidades de la Administración Pública es la alta opacidad algorítmica y falta de explicabilidad en los sistemas de IA que emplea la Administración Pública. Si bien el artículo 3° de la Ley N.º 27806 concibe el principio de publicidad, mediante el cual toda información que posee el Estado se presume pública, salvo específicas excepciones; en los casos en los que se ve involucrada la tecnología de IA, se presentan complejidades adicionales.

En el caso de decisiones automatizadas, los algoritmos que se encuentran detrás emplean modelos de DL cuyo funcionamiento es complejo de comprender por el público en general y por los funcionarios encargados de su supervisión. Este factor sin duda dificulta en la realidad el cumplimiento pleno del principio de publicidad concebido en el artículo 3° de la Ley N.º 27806 en los casos en los que la Administración Pública emplea tecnologías de IA, ya que muchas veces las decisiones basadas en IA pueden no ser explicables en términos accesibles para la ciudadanía.

Aunado a ello, no existe normativa actual que imponga la obligación específica de divulgar el funcionamiento de los algoritmos utilizados por la Administración Pública.

Esta situación, sin duda, representa un desafío para el cumplimiento de la garantía de transparencia en el Perú.

- **Datos y transparencia**

El segundo desafío que se advierte en el marco normativo aplicable para la transparencia se encuentra íntimamente vinculado con la data empleada por los algoritmos de IA. En principio, el Decreto Supremo N.º 0007-2024-JUS establece que la información de acceso público incluye documentos en soportes digitales y electrónicos. Asimismo, el artículo 10º de la Ley N.º 27806 dispone que las entidades públicas se encuentran obligadas a brindar información sobre los documentos en su posición, siempre que estos no se encuentren protegidos por excepciones legales.

En tal sentido, es necesario considerar que las bases de datos empleadas para entrenar sistemas de IA pueden contener información sensible. Al respecto, no hay disposiciones normativas específicas que regule el uso de datos en las tecnologías de IA de la Administración Pública, ni los criterios de anonimización y encriptación.

Si bien, como se ha sostenido en el presente trabajo, la publicidad de la información relativa a las tecnologías de IA empleadas por la Administración Pública es deseable en este nuevo contexto, es innegable que la ausencia de parámetros sobre qué datos y/o documentos deben ser compartidos en el marco normativo aplicable representa un riesgo para los titulares de la data que procesan los algoritmos de IA.

- **Acceso a la información sobre algoritmos utilizados por la Administración Pública**

El tercer desafío para la transparencia surge al momento de revisar el acceso por parte de los administrados a los algoritmos utilizados por la Administración Pública. En tal

sentido, el artículo 5° de la Ley N.º 27806 obliga a las entidades de la Administración Pública a mantener habilitado un portal de transparencia estándar que cuente con información relevante sobre sus actividades, que considere también información sobre la adquisición de bienes y servicios y procedimientos administrativos.

No obstante, resalta que no se menciona, de forma explícita, la obligación de publicar información sobre los algoritmos y modelos de IA empleados por las entidades de la Administración Pública.

La ausencia de un marco normativo específico que regule el uso de IA por parte de la Administración Pública representa un desafío para la transparencia, dado que los administrados no cuentan con base legal específica que les permita acceder información sobre los algoritmos que emplean las instituciones de la Administración Pública, ni con qué propósito o sobre los impactos que generan las decisiones públicas en las que algoritmos de IA se involucran.

Aunado a ello, tampoco existe un mandato normativo explícito y específico que le requiera a las entidades públicas documentar y divulgar las características y el funcionamiento de los sistemas de IA que emplean.

En relación al Proyecto de Reglamento de la Ley de IA peruana, llama la atención el literal k) del inciso 8.1 del mismo, en donde se establece la obligación de publicar los algoritmos de IA basados en software libre en la Plataforma Nacional de Software Público Peruano, ya que a la fecha se encuentra en debate a nivel europeo la regulación del código abierto u *open source*.

En efecto, como señala Velasco (2024):

(...) aunque pueda parecer que [el código abierto] es una medida que fomenta la transparencia, lo es solo a medias, puesto que el control —social, académico, comunitario— de los modelos de IA de código abierto dista mucho de poderse realizar fácilmente en la práctica, lo que redundaría en que las empresas que opten por este tipo de modelos tengan una ventaja regulatoria que, en realidad, no se cohonestaría con los verdaderos riesgos que estos modelos pueden generar (p. 65).

En ese sentido, la publicación de algoritmos de dicho tipo no garantiza por sí sola la transparencia ni la auditabilidad efectiva de los sistemas de IA utilizados por la Administración Pública, lo que evidencia la insuficiencia del marco normativo peruano para enfrentar los desafíos que plantean estas tecnologías.

- **No regulación de auditorías algorítmicas**

El cuarto desafío que se presenta para la transparencia en el contexto abordado en la ausencia de regulación específica que requiera el desarrollo de auditorías algorítmicas periódicas a los algoritmos de IA que emplea la Administración Pública.

Si bien el marco normativo sobre transparencia contempla la supervisión del cumplimiento de los principios de transparencia por parte de la Autoridad Nacional de Transparencia y Acceso a la Información Pública, no existe un mecanismo claro que imponga la normativa actual para auditar y supervisar los algoritmos de IA, en especial respecto de aspectos relativos a oscuridad, funcionamiento y transparencia; ni se exige que las entidades públicas sometan los sistemas de IA que emplean a alguna prueba o supervisión de cualquier tipo.

3.4 Propuesta de solución

En virtud a todo lo desarrollado en el presente trabajo, corresponde proponer la implementación de medidas que, a consideración de los autores del presente trabajo, pueden contribuir con la transparencia en un contexto en el que la IA se encuentra presente en las actividades diarias de las entidades de la Administración Pública.

En tal sentido, las siguientes medidas encuentran sustento no solo en la doctrina, sino también en la experiencia de otros países que han adoptado estas premisas, a fin de garantizar la transparencia en el contexto que el presente trabajo aborda.

a) Incorporar una sección sobre IA en la Ley N.º 27806 y su Reglamento

Incorporar en la Ley N.º 27806, y su Reglamento, un apartado específico que regule la obligación de documentar y publicar información sobre los algoritmos de IA que emplea la Administración Pública resulta imprescindible para asegurar la transparencia en el presente contexto, ello en la medida en que, como se ha mencionado ya, la regulación actual resulta insuficiente.

Es preciso reconocer que la Ley N.º 27806 establece en su artículo 3º el principio de publicidad, en virtud del cual toda información que posee el Estado es pública, salvo excepciones expresamente señaladas. Si bien esta premisa también le es aplicable a un escenario como el descrito, y le sería aplicable a las características de los algoritmos de IA, es innegable que una premisa así de general resulta ineficiente para garantizar la transparencia en un escenario en el que la Administración Pública emplea tecnologías de IA.

Ciertamente, los algoritmos de IA, por sus características propias, no pueden ser regulados de manera general a fin de transparentar sus características y procesos. Como se ha mencionado con anterioridad, las tecnologías de IA no solo son complejas y adolecen

de oscuridad algorítmica a causa de su compleja programación, sino también existe limitantes en cuando a compartir información que puede ser catalogada como secretos comerciales.

Asimismo, se trata de tecnologías que constantemente se reinventan, por lo que es necesario regular en específico este fenómeno, considerando acápite exclusivos en los cuerpos normativos peruanos sobre transparencia, a fin de que la garantía de transparencia sea posible en la realidad, y no exista únicamente como un principio considerado en el marco normativo.

La doctrina considera que, para elaborar disposiciones normativas que gobiernen el uso de tecnologías de IA, es necesario abordar en específico las cuestiones relacionadas a la transparencia y las herramientas de IA en uso por parte de la Administración Pública. Si bien las características de los sistemas de IA dificultan la elaboración de normativa aplicable, es imprescindible que exista normativa que aborde de manera específica y exclusiva los supuestos que involucren el empleo de herramientas de IA por parte de las entidades de la Administración Pública, a fin de garantizar la transparencia y explicabilidad (Writz, Weyerer y Geyer, 2019, p. 605). Asimismo, autores como Floridi y Cowsls consideran que un marco normativo sobre transparencia en la IA, específico y condensado, es beneficioso para la preservación de los principios éticos relacionados al empleo de tecnologías de IA por parte de las instituciones del sector público (2019, p. 9).

Respecto de lo señalado, se propone como medida que contribuye a la garantía de transparencia en el contexto del uso de IA por parte de la Administración Pública, la inclusión de un capítulo sobre el uso de IA por las instituciones públicas en la Ley N.º

27806, con requisitos de transparencia y acceso a la información, de publicidad, trazabilidad y acceso ciudadano.

Es preciso mencionar que la Unión Europea ha establecido un marco regulatorio específico y condensado, dedicado a regular la IA, que clasifica los sistemas de IA, y en función a su nivel de riesgo, estipula requisitos mayores o menores de transparencia y auditoría.

b) Creación de un registro público de algoritmos

Establecer un registro obligatorio en el que se detallen qué algoritmos se encuentran en uso por parte de la Administración Pública, considerando los propósitos y su influencia en las decisiones públicas, resulta crucial para la condensar la información sobre los algoritmos de IA que la Administración Pública emplea, y facilita la disposición de esta para los administrados.

En Francia, se implementó el Répertoire des Algorithmes, que recopila información como los algoritmos empleados en cualquier toma de decisión a nivel gubernamental, la fuente de datos utilizados, y la finalidad y criterios de diseño de los modelos. Dicho registro es accesible al público, con la finalidad de propiciar la transparencia y el control ciudadano sobre el uso de herramientas o sistemas de IA por parte de las entidades del sector público. Asimismo, resulta útil recordar las dos experiencias citadas anteriormente: la chilena y la colombiana.

Las mencionadas experiencias nos permiten notar que otras legislaciones consideran valioso el poner a disposición de los administrados la información sobre los programas y herramientas de IA, a fin de salvaguardar la garantía de transparencia. En dicho sentido, recopilar la información sobre las herramientas de IA empleadas por la entidades del

sector públicos y presentársela al público como un registro público es una medida relativamente sencilla de implementar, y directamente permite el acceso a esta información por parte de los Administrados.

La doctrina también abordó la implementación de registros de las herramientas de IA que la Administración Pública emplea. Autores como Veale, Van Kleek y Binns resaltan la ausencia de registros centralizados confiables de tecnologías de IA que son empleadas por las instituciones públicas, y considera los intentos del gobierno inglés de compartir registros como incompletos e inexactos (2018, p. 3). Aunado a ello, se considera que existe deficiencias en la documentación y supervisión de los proveedores de IA, a pesar de la generalizada aceptación de la necesidad de registros y procesos claros para evaluar los algoritmos empleados por la Administración Pública (Ada Lovelace Institute. 2020, p. 3).

En el Perú, se cuenta con el Catálogo de Herramientas de Inteligencia Artificial, que es una iniciativa del gobierno peruano que recopila una variedad de aplicaciones y plataformas diseñadas para automatizar procesos en ámbitos como el trabajo, el estudio y la creación de contenido innovado. Este catálogo cuenta con las categorías generales, diseño y creatividad, marketing, contenido multimedia y detección de contenido.

Si bien esta medida es deseable y aplaudible en el marco de la búsqueda de transparencia en el contexto del uso de sistemas de IA, la propuesta que se aborda en el presente punto busca la obligatoriedad del registro de las herramientas. La obligatoriedad del registro de las herramientas de IA es una medida proporcional, teniendo en cuenta que cualquier algoritmo o herramienta que se involucre en las actividades de las entidades del sector público puede influir en la esfera jurídica de los administrados, de forma directa, o indirecta.

En tal sentido, la creación de un registro nacional obligatorio de algoritmos públicos resulta una medida idónea para garantizar la transparencia en este nuevo contexto. Se hace imprescindible que las instituciones públicas documenten y publiquen información sobre las herramientas de IA utilizadas en sus procesos en un portal de transparencia disponible para los administrados.

Cabe indicar que la implementación de un repositorio público de algoritmos no excluye la necesidad de mantener vigente el derecho de acceso a la información pública, sino que entiende que ambos mecanismos son complementarios. Así, mientras que el repositorio responde a la lógica de transparencia activa, permitiendo el acceso general y sistemático a información importante sobre los algoritmos que use la Administración Pública, el derecho de acceso a la información pública garantiza que los ciudadanos puedan solicitar información adicional o específica que no se encuentre publicada. En ese orden de ideas, se fortalece la rendición de cuentas y se genera una supervisión más efectiva.

c) Implementación de auditorías algorítmicas

Incorporar en la normativa sobre transparencia una obligación de auditoría de IA, que evalúe aspectos como la precisión del algoritmo, resulta necesario, a fin de propiciar el óptimo funcionamiento de las tecnologías, que se verían expuestas a revisiones obligatorias necesarias y periódicas, transparentando sus características y funcionamiento.

A fin de ilustrar la presente propuesta de solución, resulta oportuno mencionar el Automated Decision Systems Task Force, implementado en la ciudad de Nueva York, cuenta con un equipo de auditoría independiente que revisa el impacto de la IA en las decisiones públicas, los cuales persiguen problemáticas que pueden presentarse en un contexto en el que se emplea IA.

La doctrina define el término de auditoría algorítmica en función a su propósito: detectar sesgos y errores sistemáticos en los sistemas de IA utilizados en la toma de decisiones públicas; y resaltan su capacidad para impulsar a los proveedores de IA a mejorar la transparencia en sus sistemas (Raji y Buolamwini, 2019, p. 1). En dicho sentido, las auditorías algorítmicas ayudan a detectar si algún sistema incumple con requisitos ya previamente establecidos por la legislación, por lo que regular la exigencia de este tipo de auditorías permite garantizar la transparencia en un momento posterior a la elaboración de legislación (Selbst y Barocas, 2018, p. 11).

En tal sentido, incorporar en la Ley de Transparencia y Acceso a la Información Pública un mandato para que los sistemas de IA sean auditados resulta crucial para garantizar la transparencia de manera continua, como medida frente a la constante innovación y mutabilidad de los algoritmos. Si bien el proyecto de reglamento de la Ley de IA peruana señala que se podrán realizar pruebas de experimentación y pruebas controladas de sistemas de IA, esto resulta insuficiente para los propósitos que ya se han descrito, pues las auditorías algorítmicas van más allá de la observancia de resultados de pruebas experimentales como las mencionadas.

CONCLUSIONES

La IA tiene su primer antecedente en el año 1950. Desde dicho año se han desarrollado diferentes tecnologías que han intentado reproducir la inteligencia humana en una máquina. Habiendo pasado por una serie de etapas, hoy en día la IA se encuentra en la era del *big data*, donde es común que se procesen grandes volúmenes de datos por diferentes técnicas de IA.

No existe a la fecha un concepto unívoco de IA, la literatura especializada ha ido formulando distintas aproximaciones sobre la base de su equiparación con la inteligencia humana, su tecnología, sus capacidades y sus componentes. Sin embargo, desde una perspectiva jurídica o regulatoria, la IA ha sido concebida sobre la noción de sistemas de IA, que funcionan con cierto nivel de autonomía y son capaces de mostrar capacidad de adaptación, empleando información de entrada (*input*) que recibe para generar resultados de salida (*output*), pudiendo ser estos últimos de decisión, predicción o recomendación.

La IA hoy en día es conocida como una disciplina, la misma que tiene asociada a sí misma diferentes conceptos conocidos como técnicas de IA, las que explican el funcionamiento de los algoritmos sobre la base de datos (ML, DL, Redes Neuronales, entre otros). Sobre la IA es imprescindible además tomar en cuenta su ciclo de vida, en tanto las fases de este último son de importante estudio para la verificación del cumplimiento de los estándares básicos para la minimización de riesgos a través de garantías básicas, como las postuladas por el Proyecto de Reglamento de IA como el principio rector de dignidad humana, de respeto de derechos de autor y derechos conexos, de seguridad, proporcionalidad y fiabilidad y de sostenibilidad.

A nivel internacional, las políticas públicas sobre IA están en permanente discusión. En el caso de Estados Unidos y China, estas se proyectan desde el desarrollo de la IA como

un *gadget* tecnológico, mientras que la UE se encuentra más ceñida a la búsqueda de una regulación unificada y funcional que permita la minimización de riesgos generados por IA. En el ámbito nacional se han emitido documentos como la ENIA y el informe del CEPLAN y la PNTD, los cuales demuestran los avances de la agenda política en el Perú.

La regulación sobre IA en el Perú ha tenido un avance significativo en lo que respecta a su marco jurídico. A la fecha se cuenta con una Ley de IA peruana y un Proyecto de Reglamento que han regulado o tratado de forma inadecuada los temas relacionados a la transparencia algorítmica como garantía que mitiga la opacidad intrínseca de los algoritmos.

La garantía de transparencia en usos de IA por parte de la Administración puede ser salvaguardada desde diferentes enfoques, como la transparencia activa o el ejercicio legítimo del derecho de acceso a la información pública (transparencia pasiva).

Sin embargo, la falta de adecuación que tiene el marco normativo peruano respecto a la garantía de la transparencia no permite que esta última se encuentre protegida y que, además, exista incertidumbre sobre los alcances del derecho de acceso a la información pública en lo que IA se refiere cuando se trata de Administración Pública.

Por ello, el presente trabajo recomienda sobre la base del marco político y jurídico peruano, así como la literatura especializada y la experiencia comparada revisada, incorporar una sección específica sobre IA y conceptos asociados en la normativa peruana relacionada al derecho de acceso a la información pública, así como la creación de un registro público de algoritmos y la implementación de auditorías algorítmicas.

BIBLIOGRAFÍA

Ada Lovelace Institute. (2020). *Examining the black box: Tools for assessing algorithmic systems*. Technical report, AdaLovelace Institute. <https://ico.org.uk/media/about>

Aggarwal, C. C. (2023). *Machine learning for text* (2nd ed.). Springer.

Alpaydin, E. (2023). *Introduction to machine learning* (4th ed.). MIT Press.

Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., & Mané, D. (2016). *Concrete problems in AI safety*. Google Brain, OpenAI, Stanford University, UC Berkeley.

Argüelles, E., & Amaro, M. (2023). Preocupaciones éticas en el uso de inteligencia artificial, transparencia y derecho de acceso a la información. El caso de los chatbots en el gobierno de México, en el contexto de la COVID-19. *Revista Estudios en Derecho a la Información*, 15, 85-111. <https://revistas.juridicas.unam.mx/index.php/derecho-informacion/issue/archive>

Azuaje, M., & Finol, D. (2020). Transparencia algorítmica y la propiedad intelectual e industrial: Tensiones y soluciones. *Revista La Propiedad Inmaterial*, 30, 11

Bartley, K. (2024, 11 de diciembre). *Big data statistics: How much data is there in the world?* Rivery.

Berning Prieto, A. D. (2023). El uso de sistemas basados en inteligencia artificial por las Administraciones públicas: estado actual de la cuestión y algunas propuestas ad futurum para un uso responsable. *Revista De Estudios De La Administración Local Y Autonómica*, (20), 165–185. <https://doi.org/10.24965/reala.11247>

Blanco García, A. I. (2024). Retos para una inteligencia artificial inclusiva de los colectivos vulnerables. *Actualidad Jurídica Iberoamericana*, 21, 360-383.

Blazquez Ruiz, F. Javier. (2022). The Paradox of Transparency in AI: Opacity and Explainability, Allocation of Responsibility. *Revista Internacional de Pensamiento Politico*, 17, 261-272.

Brundage, M., Avin, S., Wang, J., Belfield, H., Krueger, G., Hadfield, G., et al. (2020). *Toward trustworthy AI development: Mechanisms for supporting verifiable claims*. arXiv. <https://arxiv.org/abs/2004.07213v2>

Campos, C. (2022). Inteligencia artificial en la gestión pública: ¿De las tres leyes de Asimov a la anarquía desreguladora? En G. Vestri (Dir.), *La disrupción tecnológica en la administración pública. Retos y desafíos de la inteligencia artificial* (pp. 23-40). Thomson Reuters Aranzadi.

Carbonero Redondo, J. J. (2022). *Inteligencia artificial y administraciones públicas*. En F. Herrera Triguero, A. Peralta Gutiérrez, & L. S. Torres López (Coords.), *El derecho y la inteligencia artificial* (pp. 131-151). Editorial Universidad de Granada.

Centro Nacional de Inteligencia Artificial y Comisión Económica para América Latina. (2024). *Índice Latinoamericano de Inteligencia Artificial*. https://indicelatam.cl/wp-content/uploads/2025/01/ILIA_2024_020125_compressed.pdf

Centro Nacional de Inteligencia Artificial. (2024). *Índice Latinoamericano de Inteligencia Artificial 2024* [PDF]. https://indicelatam.cl/wp-content/uploads/2025/01/ILIA_2024_020125_compressed.pdf

Centro Nacional de Planeamiento Estratégico. (2021). Inteligencia artificial: Desafíos y oportunidades para el Perú. <https://observatorio.ceplan.gob.pe/publicacion/detalle/1>

Cerrillo i Martínez, A. (2024). Retos y oportunidades de la regulación española de la inteligencia artificial. En A. Cerrillo i Martínez, F. Di Lascio, I. Martín Delgado & C. I. Velasco Rico (Dirs.), *Inteligencia artificial y administraciones públicas: una triple visión en clave comparada* (pp. 101–117). Madrid: Iustel.

Centro Nacional de Planeamiento Estratégico (CEPLAN). (2023). *Perú 2040: Estrategia Nacional de Desarrollo* [Monografía]. Observatorio CEPLAN. <https://observatorio.ceplan.gob.pe/publicacion/detalle/1>

Comisión Europea. (2021). Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, por el que se establecen normas armonizadas en materia de inteligencia artificial (ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión. <https://eur-lex.europa.eu/legal-content/ES/ALL/?uri=CELEX:52021PC0206>

Comisión Europea. (2021). *Propuesta de reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión* [Propuesta de reglamento]. EUR-Lex. <https://eur-lex.europa.eu/legal-content/ES/ALL/?uri=CELEX:52021PC0206>

Comisión Europea. (2022). *Commission Staff Working Document: Executive Summary of the Impact Assessment Report accompanying the document Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive)* (SWD (2022) 320 final). Bruselas.

Comisión Europea. (2023). *Artificial intelligence white paper: On fostering a European approach to excellence and trust*. [White Paper on Artificial Intelligence: a European approach to excellence and trust](#)

Corvalán, J. G. (2018). Inteligencia artificial: retos, desafíos y oportunidades – Prometea: la primera inteligencia artificial de Latinoamérica al servicio de la Justicia. *Revista De Investigações Constitucionais*, 5(1), 295–316. <https://doi.org/10.5380/rinc.v5i1.55334>

Cotino, L. (2021). *Hacia la transparencia 4.0: El uso de la inteligencia artificial y big data para la lucha contra el fraude y la corrupción y las (muchas) exigencias constitucionales*. En *Repensando la administración digital y la innovación pública*. Instituto Nacional de Administración Pública.

Cotino, L., & Claramunt, C. (2022). Transparencia y explicabilidad de la inteligencia artificial y “compañía” (comunicación, interpretabilidad, inteligibilidad, auditabilidad, testabilidad, comprobabilidad, simulabilidad...). Para qué, para quién y cuánta. En *Transparencia y explicabilidad de la inteligencia artificial* (pp. 29-70).

Cotino, L. (2024, mayo 17). *Caso Bosco, a la tercera tampoco va la vencida: Mal camino en el acceso a los algoritmos públicos*. Diario La Ley. <https://diariolaley.laleynext.es/dll/2024/05/17/caso-bosco-a-la-tercera-tampoco-va-la-vencida-mal-camino-en-el-acceso-los-algoritmos-publicos>

Covarrubias, J., Enriquez, O. & Guerrero, M. (2022). Enfoques regulatorios para la IA. *Revista Chilena de Derecho*, 49(3), 31–62. <https://doi-org.ezproxybib.pucp.edu.pe/10.7764/R.493.2>

Decreto Legislativo N. ° 1412, Decreto Legislativo que aprueba la Ley de Gobierno Digital (13 de setiembre de 2018). <https://www.gob.pe/institucion/pcm/normas-legales/289706-1412>

Decreto de Urgencia N. ° 006-2020, Decreto de Urgencia que crea el Sistema Nacional de Transformación Digital (9 de enero de 2020). <https://www.gob.pe/institucion/mpfn/informes-publicaciones/1678070-decreto-de-urgencia-n-006-2020>

Decreto de Urgencia N. ° 007-2020, Decreto de Urgencia que Aprueba el Marco de Confianza Digital y Dispone Medidas para su Fortalecimiento (9 de enero de 2020). <https://www.gob.pe/institucion/pcm/normas-legales/395322-007-2020>

Decreto Supremo N. ° 157-2021-PCM, Decreto Supremo que aprueba el Reglamento del Decreto de Urgencia N. ° 006-2020 (25 de septiembre de 2021). <https://www.gob.pe/institucion/pcm/normas-legales/2174363-157-2021-pcm>

Decreto Supremo N. ° 029-2021-PCM – Decreto Supremo que aprueba el Reglamento del Decreto Legislativo N. ° 1412, Decreto Legislativo que aprueba la Ley de Gobierno Digital, y establece disposiciones sobre las condiciones, requisitos y uso de las tecnologías y medios electrónicos en el procedimiento administrativo (19 de febrero de 2021). <https://www.gob.pe/institucion/pcm/normas-legales/1705101-029-2021-pcm>

De Silva, D., & Alahakoon, D. (2022). An artificial intelligence life cycle: From conception to production. *Patterns*, 3(7), 100489. <https://doi.org/10.1016/j.patter.2022.100489>

Devillé, R. et al. (2021). Basic Concepts of AI for Legal Scholars. In J. De Bruyne & C. Vanleenhove (Eds.), *Artificial Intelligence and the Law* (pp. 1–22). Chapter, Intersentia.

DW Español. (2025, 11 de febrero). *Cumbre en París aborda el impacto global de la IA*.

YouTube. <https://www.youtube.com/watch?v=5zKi--7dR9U>

Espinoza, A. (30 de abril de 2024). Suplemento Jurídica: Reflexiones sobre la regulación de la inteligencia artificial y protección de datos. El Peruano.

<https://elperuano.pe/noticia/242525-suplemento-juridica-reflexiones-sobre-la-regulacion-de-la-inteligencia-artificial-y-proteccion-de-datos>

Estrada, E. (2024). *Aproximación al Marco Regulatorio de la Inteligencia Artificial en la Unión Europea*. Universidad Pontificia Comillas.

<https://repositorio.comillas.edu/xmlui/handle/11531/80406>

Fernández, J. (2024). Inteligencia artificial y Compliance: Tecnología, Ética y Ley.

Revista Buen Gobierno, 36. https://revistabuengobierno.org.mx/wp-content/uploads/2024/08/BG36_5.pdf

Fernández-Aller, C., & Serrano Pérez, M. M. (2022). ¿Es posible una inteligencia artificial respetuosa con la protección de datos? *Doxa. Cuadernos de Filosofía del Derecho*, 45, 307-336. <https://doi.org/10.14198/DOXA2022.45.11>

Fierens, M. et al. (2021). Setting the Scene: On AI Ethics and Regulation. In J. De Bruyne & C. Vanleenhove (Eds.), *Artificial Intelligence and the Law* (pp. 49–72). Chapter, Intersentia.

FRANCE 24 Español. (2025, 11 de febrero). *Informe desde París: posiciones encontradas sobre la IA en cumbre en Francia*. YouTube.

https://www.youtube.com/watch?v=0KxG6nlRz_Q

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1), Summer 2019. <https://doi.org/10.1162/99608f92.8cd550d1>

Gobierno de España. (2021). *Carta de derechos digitales*. Secretaría de Estado de Digitalización e Inteligencia Artificial. <https://www.digitaloportunidad.gob.es/carta-de-derechos-digitales>

González, F., y Díaz, N. (2023). ¿Qué es la Inteligencia Artificial? En E. Gamero Casado & F. L. Pérez Guerrero (Eds.), *Inteligencia artificial y sector público: Retos, límites y medios* (pp. 37-91). Editorial Tirant lo Blanch.

Grupo de Expertos de Alto Nivel sobre Inteligencia Artificial. (2019). *Directrices éticas para una IA fiable*. Comisión Europea. https://ec.europa.eu/digital-strategy/hleg-expert-group-artificial-intelligence_en

Gunningham, N. y Sinclair, D. (2017). Smart regulation. In P. Drahos (Ed.), *Regulatory Theory: Foundations and applications* (pp. 133-148). The Australian National University. <https://doi.org/10.22459/RT.02.2017.08>

Herrera, et al. (2022). Una visión actual de la inteligencia artificial: Recorrido histórico, datos y aprendizaje, y responsabilidad en el diseño y uso. En F. Herrera, A. Peralta & L. Torres (Coords.), *El derecho y la inteligencia artificial* (pp. 51-79). Editorial Universidad de Granada.

IBM. (s.f.). *¿Qué es el Internet de las cosas (IoT)?* Recuperado de <https://www.ibm.com/mx-es/topics/internet-of-things>

Information Commissioner's Office y The Alan Turing Institute. (2022). Explaining decisions made with AI. <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/explaining-decisions-made-with-artificial-intelligence/>

International Organization for Standardization. (2022). *What is artificial intelligence (AI)?* ISO. <https://www.iso.org/artificial-intelligence/what-is-ai>

Kelleher, J. D. (2022). *Deep learning* (2nd ed.). MIT Press.

Kiela, D. et al. (2023) *Plotting Progress in AI*. Contextual AI Blog. <https://contextual.ai/blog/plotting-progress-in-ai/#contact>

Kuteynikov, D. (2022). Legal regulation of artificial intelligence and robotic systems: review of key approaches. *Cuestiones Políticas*, 40(72), 690-703

Ley N.º 31814, Ley que promueve el uso de la IA en favor del desarrollo económico y social del país (5 de julio de 2023). <https://www.gob.pe/institucion/congreso-de-la-republica/normas-legales/4565760-31814>

Li, Y.y Goel, S. (2024). Making It Possible for the Auditing of AI: A Systematic Review of AI Audits and AI Auditability. *Inf Syst Front* (2024). <https://doi.org/10.1007/s10796-024-10508-8>

Llinares, F. M. (2018). Inteligencia artificial y justicia penal: más allá de los resultados lesivos causados por robots. *Revista de Derecho Penal y Criminología*, (20), 87-130.

Lodge, M., & Stirton, L. (2001). Regulating in the Interest of the Citizen: Towards a Single Model of Regulatory Transparency? *Social & Economic Studies*, 50(2), 103–137.

Martínez Espín, P. (2023). La propuesta de marco regulador de los sistemas de inteligencia artificial en el mercado de la UE. *Revista CESCO de Derecho de Consumo*, (46), Artículo e3322. https://doi.org/10.18239/RCDC_2023.46.3322

McCulloch, W., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5, 115-133.

Mena Roa, M. (2021). *El Big Bang del Big Data*. Statista.

Moreno, M. (2022). La propuesta de Reglamento de la Unión Europea sobre Inteligencia Artificial y las Directrices Éticas para una Inteligencia Artificial fiable: una oportunidad para la Administración Pública española. En G. Vestri (Dir.), *La disrupción tecnológica en la Administración Pública. Retos y desafíos de la inteligencia artificial* (pp. 67-81). Editorial Aranzadi, S.A.U.

Morillas, L. (2023). *Implicaciones de la inteligencia artificial en el ámbito del Derecho Penal*. En Roma Tre Press (Ed.), *Derecho Penal, Inteligencia Artificial y Nuevas Tecnologías* (pp. 59-91). Roma Tre Press.

Nestor Maslej et al. (2024). *The AI Index 2024 Annual Report*. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University. https://aiindex.stanford.edu/wp-content/uploads/2024/05/HAI_AI-Index-Report-2024.pdf

Nilsson, N. (1980). *Principles of artificial intelligence*. Tioga Publishing Company.

Obreque, R., & Gomez, P. (2024). Aplicación de la inteligencia artificial por parte de las administraciones tributarias: Transparencia algorítmica y gestión de sesgos. *Revista*

Chilena de la Administración del Estado, 11, Primer semestre 2024.
<https://doi.org/10.57211/revista.v11i1.186>

OECD. (2022). *Artificial intelligence: A transformative technology*. Organisation for Economic Co-operation and Development. <https://www.oecd.org/en/topics/policy-issues/artificial-intelligence.html>

Ordoñez-Iturralde, D. D., & Proaño-Piedra, C. X. (2024). Gobernanza y desafíos éticos en la adopción empresarial de la inteligencia artificial. *Innovación Empresarial*, 4(2), e633-e633.

Ortiz de Zárate Alcarazo, L. (2022). Explicabilidad (de la inteligencia artificial). *Eunomía. Revista en Cultura de la Legalidad*, (22), 328-344.
<https://doi.org/10.20318/eunomia.2022.6819>

Palma, A. (2022). El ciclo de vida de los sistemas de inteligencia artificial: Aproximación técnica de las fases presentes durante el diseño y despliegue de los sistemas algorítmicos. En L. Cotino Hueso & M. Bauzá Reilly (Coords.), *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas* (pp. 29-51). Thomson Reuters Aranzadi.

Parlamento Europeo y Consejo Europeo. (2024). Reglamento del Parlamento Europeo y del Consejo de 13 de junio de 2023 por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) N.º 300/2008, (UE) N.º 167/2013, (UE) N.º 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828.
<https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689>

Pérez, F. (2024). La autorregulación de la Inteligencia Artificial. *Cuadernos de la Guardia Civil: Revista de seguridad pública*, (71), 73-94.

Presidencia del Consejo de Ministros. Nuevo Proyecto de Reglamento de la Ley de IA (19 de noviembre de 2024). <https://www.gob.pe/institucion/pcm/informes-publicaciones/6197119-nuevo-proyecto-de-reglamento-de-la-ley-de-inteligencia-artificial>

Raji, I., & Buolamwini, J. (2019). Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial AI products. En *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 429-435).

Raji, I., & Dobbe, R. (2023). *Concrete problems in AI safety, revisited*. AI Now Institute, New York University. Publicado en la conferencia ICLR 2020. <https://arxiv.org/abs/2401.10899v1>

Rebollo, L. (2023). *Inteligencia artificial y derechos fundamentales*. (1ª ed.) Editorial Dykinson.

Resolución Ministerial N. ° 132-2024-PCM (2 de mayo de 2024). <https://www.gob.pe/institucion/pcm/normas-legales/5516872-132-2024-pcm>

Rodríguez, J. D. G. (2020). Retos éticos de la inteligencia artificial en el proceso judicial. En *Derecho Procesal. #NuevasTendencias: XLI Congreso Colombiano de Derecho Procesal* (pp. 499-516). Instituto Colombiano de Derecho Procesal.

Romero, L. y Villarroel, I. (2024). Transparency of artificial intelligence in public administration: review of international standards. *Brazilian Journal of International Law*, 21(1), 168-192.

Russell, S., & Norvig, P. (2022). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

Saghafian, M., et al. (2025). Understanding automation transparency and its adaptive design implications in safety-critical systems. *Safety Science*, 184, 106730.

Sartor, P., & Veiga, L. (2023). Inteligencia Artificial: Su Evolución Y Futuro Parte 1. *Hacer Empresa*, 123, 96–98.

Secretaría de Gobierno y Transformación Digital de la Presidencia del Consejo de Ministros. (2021). Estrategia Nacional de Inteligencia Artificial. <https://precaetano.pe/cursos/wp-content/uploads/2023/05/Estrategia-Nacional-de-Inteligencia-Artificial.pdf>

Secretaría de Gobierno y Transformación Digital de la Presidencia del Consejo de Ministros. (2023). Política Nacional de Transformación Digital. <https://www.gob.pe/institucion/pcm/normas-legales/4471543-085-2023-pcm>

Selbst, A., & Barocas, S. (2018). The intuitive appeal of explainable machines. *Fordham Law Review*, 87, 1085.

Serna, M. S. (2021). Inteligencia artificial y gobernanza de datos en las administraciones públicas: Reflexiones y evidencias para su desarrollo. *Gestión y Análisis de Políticas Públicas*, 26, 20-32. Instituto Nacional de Administración Pública (INAP).

Serna, A., Acevedo, E., & Serna, E. (2017). Principios de la inteligencia artificial en las ciencias computacionales. En E. Serna M. (Ed.), *Desarrollo e innovación en ingeniería* (2ª ed., pp. 161-172). Editorial Instituto Antioqueño de Investigación.

Shahriar, S., Allana, S., Hazratifard, S. M., & Dara, R. (2023). A survey of privacy risks and mitigation strategies in the artificial intelligence life cycle. *IEEE Access*, *11*, 58885–58910. <https://doi.org/10.1109/ACCESS.2023.3287195>

Soto, L. (2024). *La inteligencia artificial y la gestión del riesgo fiscal de los contribuyentes en la aplicación del sistema tributario*. En G. Sánchez-Archidona Hidalgo (Coord.), *La tributación del comercio electrónico: Modelos de negocio altamente digitalizados* (pp. 436-473).

Turing, A. (1950). Computing machinery and intelligence. *Mind*, *59*(236), 433–460.

UNESCO. (2022). *Recomendación sobre la ética de la inteligencia artificial*. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>

Veale, M., Van Kleek, M., & Binns, R. (2018). Fairness and accountability design needs for algorithmic support in high-stakes public sector decision-making. En *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-14). ACM.

Velasco, C. (2024). Implementación de sistemas algorítmicos en el sector público, preferencias ciudadanas y obligaciones de transparencia a la luz del Reglamento de IA de la Unión Europea. En A. Cerrillo, F. Di Lascio, I. Martín & C. Velasco (Dirs.), *Inteligencia artificial y administraciones públicas: Una triple visión en clave comparada* (pp. 47–70). Iustel.

Vestri, G. (2021). La inteligencia artificial ante el desafío de la transparencia algorítmica: Una aproximación desde la perspectiva jurídico-administrativa. *Revista aragonesa de Administración pública*, (56), 368-398.

Villoria, M. (2021). ¿Qué condiciones favorecen una transparencia pública efectiva? Artículo de revisión. *Revista de Estudios Políticos*, 194, 213-247. <https://doi.org/10.18042/cepc/rep.194.08>

Walter, Y. (2024). Managing the race to the moon: Global policy and governance in Artificial Intelligence regulation—A contemporary overview and an analysis of socioeconomic consequences. *Discover Artificial Intelligence*, 4(1), 14.

Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2018). Artificial intelligence and the public sector—Applications and challenges. *International Journal of Public Administration*.

