



**UNIVERSIDAD DE CHILE**

**FACULTAD DE DERECHO**

**DEPARTAMENTO DE DERECHO PROCESAL**

# **Inteligencia Artificial en el Sector Público y la relevancia de la Transparencia Algorítmica**

**Memoria para optar al grado de Licenciado en Ciencias Jurídicas y Sociales.**

José Andrés Minoletti Ríos

Profesor guía: **Rodrigo Javier Moya García**

Santiago, Chile

2023



# Índice

<b>RESUMEN</b>	<b>5</b>
<b>INTRODUCCIÓN</b>	<b>6</b>
<b>¿QUÉ ES LA INTELIGENCIA ARTIFICIAL?</b>	<b>8</b>
CONCEPTO DE INTELIGENCIA ARTIFICIAL	8
HISTORIA DE LA INTELIGENCIA ARTIFICIAL	10
CÓMO FUNCIONA	12
CARACTERÍSTICAS DE LA IA	13
<i>Autonomía y Previsibilidad.</i>	13
<i>Falta de Control.</i>	15
<i>Investigación y Desarrollo: difusión, discreción y opacidad.</i>	17
LA TRANSFORMACIÓN DIGITAL EN EL CONTEXTO DE LA INTELIGENCIA ARTIFICIAL EN EL SECTOR PÚBLICO	19
<b>INTELIGENCIA ARTIFICIAL EN EL SECTOR PÚBLICO</b>	<b>22</b>
¿POR QUÉ ES IMPORTANTE LA IA EN EL SECTOR PÚBLICO?	22
<i>Eficiencia en la Prestación de Servicios</i>	24
<i>Toma de Decisiones Informadas</i>	25
<i>Transparencia y Rendición de Cuentas</i>	25
¿CÓMO SE UTILIZA ACTUALMENTE LA IA EN EL SECTOR PÚBLICO?	26
<i>Uso de la IA en el Departamento de Trabajo de los Estados Unidos</i>	27
<i>Uso de la IA por el IRS</i>	28
<i>Uso de la IA por la ICE</i>	29
<i>HealthMap</i>	30
<i>Gestión de Tráfico en Amsterdam</i>	31
<i>Otros casos de uso de IA en el sector público</i>	32
IA EN EL SECTOR PÚBLICO: CASO DE CHILE	34
<i>Breve contextualización de la IA en Chile</i>	35
<i>IA en el Sector público chileno</i>	36
<i>Regulación de la IA</i>	39
<b>DESAFÍOS DE LA IA EN EL SECTOR PÚBLICO</b>	<b>47</b>
RIESGOS QUE IMPLICA EL USO DE IA EN EL SECTOR PÚBLICO	47
<i>Impacto laboral</i>	49
<i>Rezago</i>	53
<i>Sesgos y discriminación algorítmica</i>	57
<i>Protección de Datos</i>	66
<i>Opacidad y “Caja Negra”</i>	72
<b>TRANSPARENCIA ALGORÍTMICA</b>	<b>78</b>
INTRODUCCIÓN	78
CONCEPTO DE TRANSPARENCIA ALGORÍTMICA Y EXPLICABILIDAD	79
<i>Transparencia Procedimental, Transparencia Técnica y Explicabilidad</i>	82
IMPORTANCIA	86
REGULACIÓN INTERNACIONAL	88
<i>Legislación de la Unión Europea (UE): Enfoque Holístico y Énfasis en la Ética</i>	88
<i>Legislación en Uruguay: Enfoque Participativo y Protección de Derechos</i>	89
TRATAMIENTO DE LA TRANSPARENCIA ALGORÍTMICA EN CHILE	91
<i>Transparencia Algorítmica en la Política Nacional de IA</i>	92

<i>Recomendaciones del Consejo para la Transparencia: una mirada desde la experiencia.</i>	94
<i>El aporte del BID: un gran paso en el reconocimiento de la Transparencia Algorítmica como principio fundamental</i>	95
<b>CONCLUSIONES</b>	98
RECAPITULACIÓN	98
RECOMENDACIONES	100
<i>Transparencia Algorítmica como Principio</i>	100
<i>Obligación de Explicabilidad</i>	101
<i>Régimen de Responsabilidad Algorítmica</i>	102
<i>Equilibrio Regulatorio</i>	102
<i>Mecanismos de Cumplimiento en Transparencia</i>	105
<i>Organismo Fiscalizador Autónomo</i>	106
<i>Publicidad Activa y Capacitaciones</i>	106
<b>BIBLIOGRAFÍA</b>	109

## Resumen

Este trabajo surge del interés del autor por explorar la Inteligencia Artificial (IA) y su impacto en la Administración del Estado. El enfoque se centra en comprender las implicancias de la IA en la toma de decisiones públicas, destacando los riesgos asociados. La exploración comienza con un análisis detallado de la IA, abordando su concepto, origen, desarrollo y funcionamiento.

Luego, se examina la aplicación de la IA en la Administración pública, contextualizando su llegada a Chile. Se revisa el proceso regulatorio de la última década, desde diálogos hasta legislación. El tercer capítulo se centra en los riesgos de los sistemas automatizados, destacando el impacto laboral, el rezago, los sesgos y la privacidad. La opacidad se identifica como un punto crítico, dando lugar al concepto de Transparencia Algorítmica, que se convierte en el tema central del trabajo.

El cuarto capítulo se dedica a abordar minuciosamente la Transparencia Algorítmica, introduciendo este nuevo concepto y analizando su trascendencia en la implementación de sistemas automatizados. Se realiza un análisis comparado de legislaciones extranjeras para respaldar la importancia de la Transparencia Algorítmica, evidenciando su relevancia a lo largo del proceso.

En resumen, el trabajo explora la IA en la Administración pública, destaca los riesgos asociados y propone la Transparencia Algorítmica como clave para mitigar la opacidad y promover un uso ético de la IA en el ámbito público.

## Introducción

En las últimas décadas, la rápida evolución de la tecnología ha impulsado transformaciones significativas en diversos sectores de la sociedad, y uno de los campos que ha experimentado un impacto excepcional es el de la Inteligencia Artificial (en adelante, IA). La IA, que involucra la capacidad de las máquinas para simular procesos de pensamiento humano, ha conquistado progresivamente la atención y el interés en todo el mundo, siendo su aplicabilidad en el sector público objeto de creciente escrutinio e investigación. En este contexto, la presente tesis propone abordar el uso de la Inteligencia Artificial en el sector público, explorando su relevancia, desafíos y su contribución a la toma de decisiones efectiva, haciendo particular énfasis en la necesidad de la transparencia algorítmica. Para esto, se analizarán legislaciones y casos de estudio, con la finalidad de brindar un enfoque contextualizado y aplicado a la realidad chilena.

La Inteligencia Artificial, entendida como la habilidad de las máquinas para imitar procesos cognitivos humanos, ha logrado trascender su mero ámbito de investigación científica para convertirse en una herramienta fundamental en la resolución de problemas complejos en distintas esferas. En este sentido, su incursión en el sector público ha generado expectativas considerables, ya que la IA promete mejorar la eficiencia administrativa, optimizar la prestación de servicios gubernamentales y potenciar la toma de decisiones informada y precisa.

La toma de decisiones es una piedra angular de la gestión gubernamental, ya que las políticas y acciones que adoptan los entes públicos repercuten directamente en la vida de los ciudadanos. La integración de sistemas de IA en este proceso introduce un nuevo paradigma en la administración pública, en el que las decisiones se fundamentan en análisis algorítmicos de datos masivos y patrones previamente identificados. Esta transición hacia una toma de decisiones asistida por IA no solo conlleva ventajas evidentes, como la capacidad de procesar información a una escala inhumana, la posibilidad de implementar políticas públicas con mayor eficiencia y

la minimización de sesgos cognitivos, sino también implica desafíos sustanciales relacionados con la transparencia y la rendición de cuentas.

En este sentido, la transparencia algorítmica emerge como un aspecto crucial en el contexto de la IA aplicada al sector público. A medida que los procesos gubernamentales se vuelven cada vez más automatizados y dependientes de algoritmos, la capacidad de comprender y explicar las decisiones tomadas por estos sistemas se vuelve esencial. La opacidad de los algoritmos puede generar desconfianza y preocupaciones sobre la imparcialidad y la equidad en la administración de recursos y en la formulación de políticas públicas. Es por ello que, en este trabajo, se otorgará un enfoque particular a la importancia de la transparencia algorítmica como mecanismo de resguardo para el uso ético y equitativo de la IA en el sector público.

En línea con los objetivos de este estudio, se llevará a cabo un análisis exhaustivo de legislaciones vigentes en distintos contextos geográficos, así como también se examinarán casos de estudio que ejemplifiquen tanto los éxitos como los desafíos que emergen de la implementación de la IA en el sector público. No obstante, se pondrá especial énfasis en la realidad chilena, con el propósito de contribuir a la comprensión de cómo la IA está siendo adoptada y percibida en el entorno gubernamental de nuestro país.

En resumen, esta tesis busca dar luz sobre la penetración de la Inteligencia Artificial en el sector público, evidenciando su relevancia para la toma de decisiones y resaltando la necesidad imperante de la transparencia algorítmica. Mediante el análisis de legislaciones y casos de estudio, se aspira a ofrecer un panorama completo y contextualizado que enriquezca el debate sobre la adopción de la IA en el contexto chileno, contribuyendo así a una implementación más informada y responsable de esta tecnología en el ámbito público.

## Capítulo 1

### ¿Qué es la Inteligencia Artificial?

#### Concepto de Inteligencia Artificial

La inteligencia artificial (IA) es un campo de la informática que se centra en la creación de sistemas que pueden realizar tareas que normalmente requieren inteligencia humana, como el aprendizaje, la percepción y el razonamiento. Lasse Rouhiainen, experto internacional en inteligencia artificial, en su libro “Inteligencia Artificial; 101 cosas que debes saber hoy sobre nuestro futuro”, ante la pregunta de cómo definir la IA, señala:

*“[...] intento simplificar el tema definiendo la IA como «la habilidad de los ordenadores para hacer actividades que normalmente requieren inteligencia humana». Pero, para brindar una definición más detallada, podríamos decir que la IA es **la capacidad de las máquinas para usar algoritmos, aprender de los datos y utilizar lo aprendido en la toma de decisiones tal y como lo haría un ser humano.**”<sup>1</sup>*

El término "inteligencia artificial" fue acuñado por John McCarthy junto a otros expertos en la Conferencia de Dartmouth de 1956, donde se sentaron las bases para el desarrollo de la disciplina. Desde entonces, la IA ha pasado por varias etapas de desarrollo, desde los sistemas de reglas hasta los sistemas de aprendizaje automático y la inteligencia artificial general.

Así, la característica principal de la inteligencia artificial es su relación con la inteligencia humana como parámetro de comparación e imitación. Esto origina

---

<sup>1</sup> Rouhiainen, L. (2018). Inteligencia Artificial; 101 cosas que debes saber hoy sobre nuestro futuro. Alienta Editorial. (p. 17)

cuatro paradigmas de inteligencia artificial: 1) sistemas que piensan como humano; 2) sistemas que actúan como humanos; 3) sistemas que piensan racionalmente; 4) sistemas que actúan racionalmente (*Ver Figura 1*). Emplear un paradigma u otro dependerá del objeto de estudio y de la capacidad con la que se desee dotar a un sistema.

Debido a estos paradigmas, la inteligencia artificial es considerada una ciencia multidisciplinaria, la cual involucra teorías y modelos de las ciencias cognitivas, lingüística, neurociencias, matemáticas, informática y biología, entre otras. Por otro lado, con el fin de desarrollar sistemas inteligentes que utilizan capacidades cognitivas cercanas a las humanas (como por ejemplo distinguir, clasificar, decidir, resolver problemas, etc.), la inteligencia artificial se basa en el trabajo de varias ramas, como la representación de conocimiento, aprendizaje automático (machine learning), visión artificial, robótica, agentes autónomos, planificación y razonamiento automático, búsqueda y resolución de problemas, y procesamiento de lenguaje natural.

<b>Sistemas que piensan como humanos</b>	<b>Sistemas que piensan racionalmente</b>
<p>«El nuevo y excitante esfuerzo de hacer que los computadores piensen... máquinas con mentes, en el más amplio sentido literal». (Haugeland, 1985)</p> <p>«[La automatización de] actividades que vinculamos con procesos de pensamiento humano, actividades como la toma de decisiones, resolución de problemas, aprendizaje...» (Bellman, 1978)</p>	<p>«El estudio de las facultades mentales mediante el uso de modelos computacionales». (Charniak y McDermott, 1985)</p> <p>«El estudio de los cálculos que hacen posible percibir, razonar y actuar». (Winston, 1992)</p>
<b>Sistemas que actúan como humanos</b>	<b>Sistemas que actúan racionalmente</b>
<p>«El arte de desarrollar máquinas con capacidad para realizar funciones que cuando son realizadas por personas requieren de inteligencia». (Kurzweil, 1990)</p> <p>«El estudio de cómo lograr que los computadores realicen tareas que, por el momento, los humanos hacen mejor». (Rich y Knight, 1991)</p>	<p>«La Inteligencia Computacional es el estudio del diseño de agentes inteligentes». (Poole <i>et al.</i>, 1998)</p> <p>«IA... está relacionada con conductas inteligentes en artefactos». (Nilsson, 1998)</p>

**Figura 1.** Definiciones de IA.<sup>2</sup>

<sup>2</sup> Russell, S. J.; Norvig, P. *Inteligencia Artificial. Un Enfoque Moderno* Segunda Edición Pearson Educación, S.A., Madrid, 2004. (Cap. 1, P. 2)

## Historia de la Inteligencia Artificial

El desarrollo de la inteligencia artificial (IA) ha sido un proceso largo y complejo que ha abarcado muchas décadas y ha involucrado a múltiples disciplinas. A lo largo de la historia, ha habido muchos hitos importantes en el desarrollo de la IA que han ayudado a impulsar su evolución.

Uno de los primeros hitos en el desarrollo de la IA se produjo en la década de 1940, cuando el matemático John von Neumann comenzó a trabajar en la teoría de la auto-replicación y la creación de máquinas que pudieran autoreplicarse. Sin embargo, esta investigación no se centró específicamente en la IA, sino más bien en la teoría de la computación.

En la década de 1950, se produjo otro hito importante con el surgimiento de la inteligencia artificial como campo de investigación. Uno de los eventos más significativos fue la Conferencia de Dartmouth en 1956, en la que un grupo de investigadores -entre ellos John McCarthy- de diversos campos se reunieron para discutir la posibilidad de crear máquinas capaces de imitar la inteligencia humana. La conferencia marcó el comienzo oficial de la investigación en inteligencia artificial como campo interdisciplinario.

Durante la década de 1960, la IA comenzó a recibir financiación gubernamental en los Estados Unidos y en otros lugares. Se produjeron importantes avances en el campo, como la creación del primer programa de ajedrez de computadora y la invención del lenguaje de programación LISP. A medida que la IA continuó avanzando, surgieron nuevas técnicas de inteligencia artificial, como la lógica difusa, que permitía a las máquinas manejar información imprecisa.

En la década de 1970, la IA se centró en el procesamiento del lenguaje natural y la visión por computadora. Se desarrollaron sistemas capaces de reconocer y procesar el habla humana y de interpretar imágenes y vídeos. Además, se produjeron avances

significativos en el campo de los sistemas expertos, que son programas de computadora diseñados para imitar el conocimiento y la experiencia de un experto humano en un campo específico.

Durante la década de 1980, la IA comenzó a integrarse en aplicaciones comerciales. Se crearon robots y sistemas de automatización industrial capaces de realizar tareas específicas. Además, se desarrollaron sistemas de aprendizaje automático que permitían a las máquinas mejorar su rendimiento con la experiencia. Uno de los ejemplos más destacados de la IA en esta época fue el sistema Deep Blue, de IBM, que derrotó al campeón mundial de ajedrez Garry Kasparov en 1997.

En los últimos 20 años, la inteligencia artificial ha experimentado un gran avance en su desarrollo, en gran parte gracias a la mejora de la capacidad de procesamiento de datos y la aparición de nuevas técnicas y algoritmos de aprendizaje automático. Esto ha permitido el surgimiento de una amplia gama de aplicaciones en diversos campos, desde la robótica hasta el procesamiento del lenguaje natural. De esta manera, uno de los mayores avances ha sido la creación de sistemas de aprendizaje profundo o deep learning. Estos sistemas utilizan redes neuronales artificiales que son capaces de aprender y mejorar su rendimiento a medida que se les proporciona más información. El aprendizaje profundo ha demostrado ser extremadamente efectivo en la resolución de problemas complejos, como la clasificación de imágenes y el procesamiento del lenguaje natural.

En resumen, en los últimos 20 años, la inteligencia artificial ha experimentado un gran avance en su desarrollo gracias a la mejora de la capacidad de procesamiento de datos y la aparición de nuevas técnicas y algoritmos de aprendizaje automático. Esto ha permitido el surgimiento de una amplia gama de aplicaciones en diversos campos, desde la robótica hasta el procesamiento del lenguaje natural, y ha llevado a importantes avances en la automatización y la eficiencia en muchos ámbitos de la vida.

## Cómo funciona

A grandes rasgos, el funcionamiento de la IA se basa en la creación de algoritmos y modelos matemáticos que, a partir de grandes cantidades de datos, son capaces de identificar patrones y tomar decisiones basadas en ellos.

En términos más técnicos, el funcionamiento de la IA se puede desglosar en varias etapas:

En primer lugar, se requiere la recopilación y preparación de datos, que pueden provenir de diversas fuentes, como bases de datos, sensores, redes sociales, entre otros. Estos datos se utilizan para entrenar un modelo de IA, que es una representación matemática de un problema específico.

Una vez que se ha creado un modelo de IA, este se alimenta con nuevos datos que le permiten ajustar sus parámetros y mejorar su rendimiento. El proceso de entrenamiento puede requerir grandes cantidades de poder de cómputo y recursos de almacenamiento, especialmente para modelos de IA más complejos y profundos. Después de entrenar el modelo, se puede utilizar para hacer predicciones o tomar decisiones en función de nuevos datos. Por ejemplo, en el campo de la visión por computadora, un modelo de IA entrenado para reconocer objetos puede tomar una imagen como entrada y clasificarla como "perro", "gato", "auto", etc. En otras aplicaciones, la IA puede utilizarse para detectar fraudes en transacciones financieras, sugerir productos a los clientes, optimizar la producción en fábricas, entre otros.

Es importante destacar que la IA se divide en diferentes subcampos, cada uno con sus propias técnicas y algoritmos. Por ejemplo, el aprendizaje automático (machine learning) es un subcampo de la IA que se enfoca en enseñar a las máquinas a aprender a partir de datos, mientras que el procesamiento del lenguaje natural (natural language processing) se enfoca en permitir que las máquinas comprendan y procesen el lenguaje humano.

En general, el funcionamiento de la IA es complejo y depende en gran medida del problema específico que se esté tratando de resolver. Sin embargo, a medida que la tecnología avanza, se espera que las máquinas sean capaces de realizar tareas cada vez más complejas, lo que abrirá nuevas posibilidades para la automatización y la optimización de procesos en diversos campos.

### Características de la IA

Para comprender adecuadamente la IA y sus implicancias, es necesario abordar algunas de sus características que la convierten en una tecnología compleja y susceptible a originar desafíos, principalmente desde una perspectiva regulatoria y de fiscalización. En este sentido, Matthew Scherer<sup>3</sup>, experto del Centro para la Democracia y Tecnología de Harvard, señala que existen rasgos característicos de la IA, que se explicarán a continuación.

### **Autonomía y Previsibilidad.**

El avance de la inteligencia artificial (IA) plantea desafíos legales debido a sus características distintivas, como la autonomía en la toma de decisiones. La capacidad de las máquinas de actuar de manera autónoma, sin control ni supervisión humana, está generando cambios significativos en la economía y el mercado laboral. Así como la Revolución Industrial redujo la demanda de mano de obra humana en la manufactura y la agricultura, la IA y los avances tecnológicos relacionados están reduciendo la demanda de trabajo humano en el sector de servicios.

La autonomía de la IA también presenta un desafío para el sistema legal en términos de previsibilidad. Algunos sistemas de IA han demostrado comportamientos creativos que podrían considerarse "fuera de lo común" desde la perspectiva

---

<sup>3</sup> Scherer, Matthew U., *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies* (2015). *Harvard Journal of Law & Technology*, Vol. 29, No. 2, Spring 2016, Disponible en: <https://ssrn.com/abstract=2609777> o <http://dx.doi.org/10.2139/ssrn.2609777>

Se deja constancia que para efectos de identificar las características de la Inteligencia Artificial, me basé, en su mayoría, en las ideas de dicho autor, considerando que se trata de una sistematización muy precisa y da cuenta de las potencialidades y riesgos asociados.

humana. Por ejemplo, los programas de ajedrez por computadora han jugado movimientos que desafían los preceptos básicos de la estrategia ajedrecística humana.

Asimismo, la IA también puede generar soluciones inesperadas debido a su capacidad para analizar una cantidad masiva de información en poco tiempo. Las personas naturales, debido a las limitaciones cognitivas (inherentes a su naturaleza), tienden a conformarse con soluciones satisfactorias en lugar de óptimas, sobretodo cuando se enfrentan a limitaciones de tiempo. Por el contrario, los programas de IA pueden explorar muchas más posibilidades y generar soluciones que los humanos no habrían considerado<sup>4</sup>. Esto puede resultar en soluciones divergentes de las generadas por procesos cognitivos humanos.

Desde una perspectiva legal, lo relevante no es tanto si la IA es "creativa", sino que sus acciones sean imprevisibles, incluso para los programadores que la diseñaron. La capacidad de la IA para generar soluciones únicas la hace atractiva en diversos campos, lo cual brinda incentivos económicos para crear sistemas de IA capaces de generar soluciones inesperadas.

A medida que la IA se vuelve más versátil y el aprendizaje automático avanza, es probable que surjan más problemas relacionados con el comportamiento imprevisible de la IA. Los sistemas de IA que aprenden de su experiencia podrían generar acciones imprevisibles, y los diseñadores y fabricantes no podrán controlar ni predecir completamente el comportamiento posterior de la IA una vez que esté en funcionamiento<sup>5</sup>.

Esto a su vez plantea desafíos en términos de responsabilidad legal. Si se considera que las acciones de una IA en constante aprendizaje son tan imprevisibles que sería injusto responsabilizar a los diseñadores por los daños causados por los sistemas, las víctimas podrían quedar sin compensación por sus pérdidas. Por el contrario, si se

---

<sup>4</sup> Scherer, 2015, p. 364.

<sup>5</sup> *Ídem*, p. 366.

atribuye una excesiva responsabilidad a los diseñadores, podría generarse un desincentivo al desarrollo y uso de dicha tecnología.

Los problemas relacionados con la previsibilidad y la causalidad son desafíos complejos que el sistema legal deberá resolver para garantizar que existan mecanismos de reparación para las víctimas de daños causados por la IA. La imprevisibilidad de la IA plantea interrogantes sobre cómo atribuir responsabilidad en casos en los que el comportamiento de la IA es resultado de su experiencia y no puede ser previsto por los diseñadores.

### **Falta de Control.**

Los riesgos creados por la autonomía de la inteligencia artificial (IA) abarcan no solo problemas de previsibilidad, sino también problemas de control. Puede resultar difícil para los humanos mantener el control de máquinas que están programadas para actuar con considerable autonomía. Existen diversas formas en las que se puede perder el control: un mal funcionamiento, como un archivo corrupto o daños físicos en el equipo de entrada; una violación de seguridad; el tiempo de respuesta superior de las computadoras en comparación con los humanos; o una programación defectuosa.

Esta última posibilidad plantea los problemas más interesantes, ya que crea la posibilidad de que la pérdida de control sea la consecuencia directa pero no intencionada de una elección consciente de diseño. Una vez perdido el control, puede resultar difícil recuperarlo si la IA está diseñada con características que le permiten aprender y adaptarse. Estas características hacen que la IA sea una fuente potencial de riesgo público en una escala que supera con creces las formas más conocidas de riesgo público que son exclusivamente resultado del comportamiento humano.

La pérdida de control puede dividirse en dos variedades<sup>6</sup>. (i) La pérdida de control local: ocurre cuando el sistema de IA ya no puede ser controlado por los humanos legalmente responsables de su operación y supervisión. (ii) La pérdida de control general: ocurre cuando el sistema de IA ya no puede ser controlado por ningún humano. Por un lado, la perspectiva de una pérdida de control general presenta un riesgo público mucho mayor que la pérdida de control local, sin embargo, una pérdida de control general no necesariamente representa un riesgo público significativo, en este sentido, los objetivos del sistema de IA debiesen estar alineados con los de la sociedad en general. Desafortunadamente, asegurar tal alineación de intereses y objetivos puede resultar bastante difícil, especialmente porque los valores humanos en sí mismos son casi imposibles de definir con precisión.

El potencial de desalineación de intereses se deriva del hecho de que los objetivos de una IA son determinados por su programación inicial. Incluso si esa programación inicial permite o fomenta que la IA modifique sus objetivos basándose en experiencias posteriores, esas alteraciones ocurrirán de acuerdo con los dictados de la programación inicial. A primera vista, esto parece beneficioso en términos de mantener el control. Después de todo, si son los humanos quienes realizan la programación inicial, tienen total libertad para moldear los objetivos de la IA. Sin embargo, muchos expertos y comentaristas en IA sugieren que si se programa a una IA para lograr cierto objetivo, podría seguir trabajando hacia ese objetivo incluso si los resultados de sus esfuerzos no son lo que los programadores originales hubieran pretendido subjetivamente.

En escenarios como este, el riesgo de un sistema de IA no se deriva de la “malevolencia” o la incapacidad para comprender la intención subjetiva detrás de sus metas programadas. Más bien, se deriva de la indiferencia fundamental de la máquina hacia esa intención subjetiva. Por lo tanto, con los sistemas de IA, debemos tener mucho cuidado con lo que pedimos, pues a diferencia de dicha tecnología, los

---

<sup>6</sup> Scherer, 2015, p. 367.

humanos no tendrían problemas para darse cuenta de que la función de utilidad propuesta no puede ser tomada literalmente.

Un creciente grupo de académicos, empresarios tecnológicos y futuristas ha ido más allá, advirtiendo que formas más poderosas de IA pueden resistir todos los esfuerzos humanos para gobernar sus acciones y representar un riesgo catastrófico, e incluso existencial, para la humanidad<sup>7</sup>. Una expresión común de esta preocupación se centra en la posibilidad de que un sistema de IA sofisticado pueda mejorar su propio hardware y programación hasta el punto de adquirir capacidades cognitivas que superen ampliamente las de sus creadores humanos. Como indica el ejemplo de "minimizar el sufrimiento humano", podría ser extremadamente difícil asegurar que los objetivos de dicho sistema de IA estén alineados con los de sus diseñadores y operadores humanos.

Si tales sistemas de IA resultan ser más que una posibilidad teórica, sería necesario tomar medidas anticipadas para garantizar que los sistemas sigan siendo susceptibles al control humano, se alineen con el interés público o ambas cosas. Lo anterior puede dificultarse a medida que los sistemas de IA se implementen en un número cada vez mayor de industrias, lo cual solo refuerza la necesidad de garantizar que los humanos mantengan medios para controlar sistemas de IA sofisticados.

### **Investigación y Desarrollo: difusión, discreción y opacidad.**

En primer lugar, la IA plantea desafíos significativos a nivel legislativo debido a ciertas características problemáticas asociadas con la investigación y desarrollo (I+D) en IA. Estas características incluyen la difusión, la discreción y la opacidad. La discreción se refiere a que el desarrollo de la IA puede llevarse a cabo con una infraestructura limitada y poco visible. La difusión implica que los individuos que trabajan en un componente de un sistema de IA pueden estar ubicados en lugares muy distantes entre sí. La discreción se refiere a que los componentes separados de

---

<sup>7</sup> Scherer, 2015, p. 368.

un sistema de IA pueden ser diseñados en diferentes lugares y tiempos sin coordinación consciente. Finalmente, la opacidad se refiere a la posibilidad de que el funcionamiento interno de un sistema de IA se mantenga en secreto y no sea susceptible de ingeniería inversa. Estas características presentan desafíos únicos en el contexto de la IA, aunque también se observan en otras tecnologías de la Era de la Información.

En comparación con las fuentes de riesgo público del siglo XX, como la tecnología nuclear, los bienes de consumo producidos en masa, la contaminación a gran escala y la producción de sustancias tóxicas, la I+D de IA se puede realizar de manera más discreta. El desarrollo de IA no requiere grandes inversiones en infraestructura física, lo que facilita la participación de actores más diversos. Cualquier persona con una computadora personal moderna y conexión a Internet puede contribuir a proyectos relacionados con la IA. Esta capacidad de discreción marca una diferencia significativa con respecto a las fuentes de riesgo público anteriores.

Además, los participantes en un proyecto de IA pueden estar dispersos geográficamente y no necesariamente estar afiliados a la misma organización. La existencia de bibliotecas de aprendizaje automático de código abierto permite que individuos dispersos realicen modificaciones anónimas en dichas bibliotecas. Estos componentes pueden incorporarse en la programación de un sistema de IA desarrollado por una entidad que no participó en la creación de la biblioteca subyacente. Esta difusión de participantes y componentes añade complejidad a la gestión de los riesgos asociados con la IA.

Por último, los sistemas de IA pueden ser opacos en cuanto a su funcionamiento interno y la interacción entre sus componentes. Los componentes de software disponibles comercialmente a menudo tienen un código propietario, lo que dificulta su comprensión y análisis. A diferencia de tecnologías anteriores, como los automóviles, cuyo funcionamiento es ampliamente conocido, los sistemas de IA pueden carecer de transparencia en términos de defectos de diseño y responsabilidad. La opacidad de los sistemas de IA y la multiplicidad de actores

involucrados dificultan la identificación de la parte responsable en caso de daños causados por el sistema.

Estas características plantean dificultades logísticas fundamentales para los reguladores en comparación con las fuentes de riesgo público anteriores. La ubicación dispersa de los participantes y la falta de relaciones contractuales formales dificultan la regulación de su participación. Los intentos de regulación en un país pueden tener un impacto limitado en el desarrollo de proyectos de IA, ya que las empresas pueden trasladar fácilmente su trabajo de desarrollo a otros países con regulaciones menos restrictivas. La naturaleza discreta y la interacción compleja entre los componentes de un sistema de IA también dificultan la asignación de responsabilidad en caso de daños.

En resumen, la difusión, la discreción y la opacidad asociadas con la investigación y desarrollo en IA presentan desafíos únicos desde una perspectiva regulatoria. Estas características dificultan la identificación de las partes responsables y la asignación de responsabilidad en caso de daños causados por sistemas de IA. Además, la capacidad de participación diversa y discreta en proyectos de IA plantea desafíos adicionales para la regulación de los riesgos asociados.

La falta de transparencia en el funcionamiento interno de los sistemas de IA también complica la atribución de responsabilidad. En general, es necesario abordar estos desafíos para establecer un marco regulatorio efectivo y garantizar la protección de los intereses públicos en el contexto de la IA, razón por la cual el análisis en profundidad de la transparencia algorítmica será fundamental para pensar en la futura implementación de la IA en el sector público.

## La Transformación Digital en el Contexto de la Inteligencia Artificial en el Sector Público

La transformación digital, en el marco de esta investigación centrada en la IA en el sector público, se erige como un fenómeno clave que redefine la naturaleza y

operación de las instituciones gubernamentales. La transformación digital, en esencia, es “el uso de tecnologías digitales para modificar integralmente el funcionamiento de las organizaciones”<sup>8</sup>, y es la piedra angular que facilita la integración y evolución de tecnologías avanzadas, y la IA se erige como una manifestación prominente de este cambio paradigmático.

La transformación digital en el sector público implica una transición profunda hacia la adopción estratégica de tecnologías digitales, y la IA actúa como catalizador en este proceso. La IA, como expresión destacada de la transformación digital, encapsula la capacidad de las instituciones gubernamentales para utilizar algoritmos avanzados y análisis de datos para mejorar la toma de decisiones, optimizar procesos y ofrecer servicios más personalizados y eficientes a los ciudadanos.

La IA, por su naturaleza, representa un paso más allá de la simple automatización de tareas. Su integración en el sector público impulsa una transformación en la manera en que se abordan los problemas y se gestionan los recursos. Por ejemplo, en la formulación de políticas públicas, la IA puede analizar grandes conjuntos de datos de manera rápida y precisa, proporcionando información valiosa para la toma de decisiones fundamentadas. Este enfoque no solo aumenta la eficiencia, sino que también permite la identificación de patrones y tendencias que podrían pasar desapercibidos en un enfoque más tradicional.

El impacto de la transformación digital, en particular mediante la incorporación de la IA en el sector público, trasciende los límites organizativos. Al analizar este fenómeno, se torna evidente cómo la transformación digital no solo modifica los procesos internos de las instituciones gubernamentales, sino que redefine la relación entre el gobierno y los ciudadanos. La capacidad de proporcionar servicios más

---

<sup>8</sup> Torrecilla-Salinas, C., Tangi, L., Ulrich, P., Manzoni, M., Schade, S., Martínez-Rodríguez, E., & Pignatelli, F. ¿Para qué sirve la Inteligencia Artificial en el sector público? Casos de uso y perspectivas de aplicación. Unidad de Economía Digital. Centro Común de Investigación, Comisión Europea. Inteligencia Artificial y Sector Público: Retos, límites y medios (Cap. II, p. 74). Valencia: Tirant lo Blanch.

personalizados y adaptados a las necesidades individuales se convierte en una realidad alcanzable mediante la aplicación efectiva de la IA.

Esta tesis se propone explorar en detalle cómo la transformación digital, con la IA como su exponente, incide en la dinámica del sector público. A lo largo del trabajo, se analizará en profundidad el impacto de la IA en la toma de decisiones gubernamentales, la formulación de políticas públicas y la relación entre el gobierno y los ciudadanos.

## Inteligencia Artificial en el Sector Público

Habiendo abordado las características de la IA y las implicancias de su uso desde una perspectiva amplia, corresponde ahora llevar a cabo un análisis de cómo el uso de esta nueva tecnología se vincula con la administración pública, para lo cual se debe explicar en primer lugar su relevancia.

¿Por qué es importante la IA en el sector público?

La inteligencia artificial (IA) es una revolución tecnológica que está transformando la forma en que vivimos y trabajamos en la sociedad moderna. Su impacto se extiende a casi todos los aspectos de nuestras vidas, y uno de los ámbitos en los que está desempeñando un papel cada vez más crucial es el sector público. La adopción de la IA en las instituciones gubernamentales y las agencias gubernamentales no solo está cambiando la forma en que se prestan los servicios públicos, sino que también está remodelando la toma de decisiones y la transparencia en el gobierno.

En esta era de avances tecnológicos, la IA se ha consolidado como una herramienta poderosa que puede mejorar drásticamente la eficiencia y la eficacia en el sector público. Las agencias gubernamentales a nivel mundial están recurriendo a la IA para automatizar tareas rutinarias, analizar grandes cantidades de datos, predecir tendencias y patrones, y mejorar la toma de decisiones informadas. Este enfoque innovador no solo permite a las instituciones gubernamentales funcionar de manera más eficiente, sino que también ofrece la posibilidad de brindar servicios más personalizados y mejorar la calidad de vida de los ciudadanos.

En esta introducción, exploraremos la importancia de la inteligencia artificial en el sector público, examinando cómo está revolucionando la administración gubernamental, la prestación de servicios y la toma de decisiones. A lo largo de este ensayo, analizaremos ejemplos concretos de casos en los que la IA se ha convertido

en una herramienta fundamental y cómo está transformando la forma en que interactuamos con el gobierno y experimentamos los servicios públicos. Además, abordaremos la cuestión de la transparencia y la rendición de cuentas en un mundo cada vez más impulsado por la IA.

La IA se ha convertido en un motor de cambio en el sector público debido a su capacidad para automatizar procesos y tareas que antes eran intensivas en mano de obra y propensas a errores humanos. Por ejemplo, en la gestión de documentos y registros, la IA se utiliza para escanear, organizar y clasificar grandes volúmenes de información de manera mucho más rápida y precisa de lo que podría hacerlo un equipo humano. Esto no solo ahorra tiempo y recursos, sino que también reduce la posibilidad de errores costosos.

Un ejemplo paradigmático es el Departamento de Seguridad Nacional de los Estados Unidos (DHS, por sus siglas en inglés), que ha implementado sistemas de IA para prevenir y reconocer patrones que podrían estar relacionados con actividades ilegales<sup>9</sup>. La IA puede detectar anomalías y alertar a los agentes de manera eficiente, permitiendo un control fronterizo más efectivo y seguro.

Además de la automatización, la IA también está mejorando la calidad de los servicios públicos. Los chatbots de IA, por ejemplo, se utilizan en sitios web gubernamentales para proporcionar respuestas instantáneas a las preguntas de los ciudadanos. Esto no solo mejora la accesibilidad a la información gubernamental, sino que también libera a los empleados públicos para que se centren en tareas más complejas y estratégicas.

De esta manera, el Servicio de Impuestos Internos (IRS) de los Estados Unidos utiliza chatbots de IA para responder a preguntas comunes sobre impuestos. Los contribuyentes pueden obtener respuestas rápidas y precisas sin tener que esperar en largas líneas de atención al cliente o navegar por sitios web confusos. Esto no solo

---

<sup>9</sup> Ver: <https://www.dhs.gov/ai>

mejora la experiencia del ciudadano, sino que también ayuda a garantizar el cumplimiento de las obligaciones tributarias.

La IA también desempeña un papel fundamental en la toma de decisiones gubernamentales. La capacidad de procesar grandes volúmenes de datos en tiempo real y analizar tendencias y patrones complejos permite a los líderes gubernamentales tomar decisiones más informadas y basadas en evidencia.

Un ejemplo destacado es el uso de la IA en la gestión del tráfico en las ciudades. Los sistemas de control de tráfico basados en IA pueden ajustar los tiempos de los semáforos y las señales de tráfico en tiempo real según las condiciones del tráfico, reduciendo los atascos y mejorando la fluidez del tráfico. Esto no solo ahorra tiempo a los ciudadanos, sino que también reduce la congestión y las emisiones de carbono, lo que tiene un impacto positivo en el medio ambiente.

En base a lo anterior, es innegable sostener que el uso de IA en el sector público presenta grandes beneficios, los cuales se pueden resumir en tres grandes ventajas: permite lograr una mayor eficiencia en la prestación de servicios, tomar decisiones más informadas y favorecer la transparencia y rendición de cuentas en la administración.

### **Eficiencia en la Prestación de Servicios**

Uno de los beneficios más evidentes de la IA en el sector público es la mejora de la eficiencia en la prestación de servicios. La automatización de tareas rutinarias y repetitivas permite a las agencias gubernamentales liberar recursos humanos para tareas más estratégicas y creativas. Por ejemplo, en el procesamiento de solicitudes de beneficios sociales, como el seguro de desempleo o la seguridad social, los chatbots de IA pueden gestionar consultas básicas, proporcionar información y guiar a los solicitantes a través de los procedimientos, acelerando el proceso y reduciendo la carga de trabajo para los empleados públicos.

Un ejemplo de esto es el Departamento de Trabajo de los Estados Unidos, que implementó un asistente virtual de IA llamado "Amelia" para ayudar a los solicitantes de beneficios. Amelia puede responder preguntas frecuentes, guiar a los usuarios a través de formularios y brindar información actualizada sobre el estado de sus solicitudes. Esto ha permitido al departamento manejar un volumen significativo de consultas de manera eficiente, lo que sería difícil de lograr sin la IA.

### **Toma de Decisiones Informadas**

La IA también desempeña un papel crucial en la mejora de la toma de decisiones en el sector público. La capacidad de procesar grandes volúmenes de datos y analizar patrones complejos permite a los responsables de la toma de decisiones gubernamentales obtener información valiosa para elaborar políticas y estrategias más efectivas.

Por ejemplo, en el ámbito de la salud pública, la IA se ha utilizado para analizar datos epidemiológicos y predecir brotes de enfermedades. El proyecto HealthMap utiliza algoritmos de aprendizaje automático para rastrear y pronosticar la propagación de enfermedades infecciosas en todo el mundo. Esta información es esencial para que los gobiernos tomen medidas preventivas oportunas y asignen recursos de manera efectiva.

### **Transparencia y Rendición de Cuentas**

La transparencia es un principio fundamental en el sector público. La IA puede desempeñar un papel importante en la promoción de la transparencia y la rendición de cuentas al automatizar procesos y proporcionar un registro claro de las decisiones tomadas.

Un ejemplo notable es el uso de la IA en la revisión y el análisis de contratos gubernamentales. Los algoritmos de IA pueden escanear contratos en busca de

cláusulas sospechosas o desviaciones de las normativas, lo que reduce el riesgo de corrupción y mejora la eficiencia en la gestión de contratos gubernamentales.

De esta manera, se puede apreciar cómo la inteligencia artificial está desempeñando un papel cada vez más importante en el sector público, mejorando la eficiencia en la prestación de servicios, impulsando la toma de decisiones informadas y promoviendo la transparencia en el gobierno. A medida que continuamos explorando el impacto de la IA en el sector público, veremos cómo esta tecnología continúa transformando la forma en que interactuamos con el gobierno y experimentamos los servicios públicos.

A continuación, se analizará de manera más detallada las diversas formas en que se emplea la IA en el sector público, basándonos principalmente en la experiencia de EEUU y de Europa.

¿Cómo se utiliza actualmente la IA en el sector público?

Para poder entender adecuadamente cómo ha sido utilizada esta tecnología en el sector público (principalmente en naciones del primer mundo), es necesario recordar la gran amplitud que la caracteriza, lo que permite que la IA sea capaz de solucionar problemas en todo tipo de ámbitos, ya sea desde tareas simples o complejas de la administración interna y externa del Estado, hasta asuntos como salud pública o movilidad.

De esta manera, este trabajo abordará con mayor detalle algunos casos de cómo la IA ha sido empleada para solucionar diversos tipos de problemas que se presentan en las sociedades desarrolladas hoy en día. En particular, se analizará el uso de la IA en áreas como seguridad, salud, movilización, administración y servicios para la ciudadanía.

## Uso de la IA en el Departamento de Trabajo de los Estados Unidos

El Departamento de Trabajo de los Estados Unidos (DOL) se enfrenta a una carga significativa de trabajo en la gestión de consultas y reclamaciones de empleados y empleadores. Estas consultas a menudo involucran preguntas complejas sobre leyes laborales, regulaciones y derechos del empleo. Antes de la implementación de la inteligencia artificial, la capacidad del DOL para responder de manera oportuna y precisa a estas preguntas era limitada, lo que a menudo resultaba en demoras y respuestas incompletas.

Para abordar este desafío, el DOL implementó una plataforma de inteligencia artificial llamada "Amelia"<sup>1011</sup>, desarrollada por IPsoft. Amelia es un asistente virtual basado en IA diseñado para comprender y responder preguntas complejas relacionadas con el empleo y las leyes laborales. Utiliza el procesamiento de lenguaje natural (NLP) y el aprendizaje automático para interactuar con los usuarios de manera similar a como lo haría un agente humano.

Amelia puede responder a una amplia variedad de consultas, desde preguntas básicas sobre derechos laborales hasta consultas más complejas sobre reglamentaciones específicas de la industria. Además, puede guiar a los usuarios a través de procesos como la presentación de reclamaciones y formularios, lo que agiliza aún más la interacción.

La implementación de Amelia ha tenido un impacto significativo en el Departamento de Trabajo de los Estados Unidos. En primer lugar, ha mejorado drásticamente la eficiencia y la velocidad de respuesta a las consultas y reclamaciones. Los usuarios pueden obtener respuestas precisas en tiempo real, lo que reduce las demoras y mejora la satisfacción del cliente.

---

<sup>10</sup> Ver: <https://www.forbes.com/sites/cognitiveworld/2022/08/13/how-the-us-department-of-labor-is-leveraging-its-data-rich-environment/?sh=46e4aecf135b>

<sup>11</sup> Urquidi, M., & Ortega, G. (2020). Inteligencia artificial para la búsqueda de empleo: Cómo mejorar la intermediación laboral en los servicios de empleo. División de Mercados Laborales, Sector Social, Banco Interamericano de Desarrollo.

Además, Amelia ha liberado a los empleados del DOL de tareas repetitivas y rutinarias, permitiéndoles centrarse en casos más complejos y en brindar asistencia personalizada cuando es necesario. Esto ha aumentado la productividad y ha liberado recursos para otros fines.

En general, la implementación de Amelia en el DOL ha mejorado la calidad de los servicios prestados a los ciudadanos y ha demostrado el potencial de la IA para transformar la interacción entre el gobierno y los ciudadanos.

### **Uso de la IA por el IRS**

El Servicio de Impuestos Internos de los Estados Unidos (IRS) se enfrenta a la complejidad de administrar y hacer cumplir la ley fiscal en un país con una amplia variedad de normativa aplicable. La presentación de declaraciones de impuestos, la detección de evasión fiscal y la gestión de auditorías son tareas que requieren un procesamiento masivo de datos y una comprensión profunda de las leyes fiscales, cuestión que por décadas ha requerido de un gran número de funcionarios y esfuerzos para ser abordadas adecuadamente por la administración.

De esta manera, el IRS ha recurrido a la inteligencia artificial para mejorar la eficiencia en la administración fiscal y aumentar la precisión en la detección de irregularidades<sup>12</sup>. La IA se utiliza en la clasificación y el análisis de declaraciones de impuestos, lo que permite identificar patrones de evasión fiscal y discrepancias de manera más efectiva que los métodos tradicionales.

Además, el IRS ha implementado chatbots de IA en su sitio web para ayudar a los contribuyentes a obtener respuestas a preguntas comunes y comprender mejor sus obligaciones fiscales, lo cual ha permitido mejorar la accesibilidad y la experiencia del usuario.<sup>13</sup>

---

<sup>12</sup> New York Times. Ver: <https://www.nytimes.com/2023/09/08/us/politics/irs-deploys-artificial-intelligence-to-target-rich-partnerships.html>

<sup>13</sup> IRS. Ver: <https://www.irs.gov/newsroom/irs-expands-use-of-chatbots-to-help-answer-questions-on-key-notices-expands-on-technology-thats-served-13-million-taxpayers>

Así, implementación de la IA en el IRS ha llevado a una mayor eficiencia en la administración fiscal. El procesamiento automatizado de declaraciones de impuestos y la detección de evasión fiscal han reducido la carga de trabajo manual y han mejorado la precisión en la identificación de problemas fiscales.

Además, los chatbots de IA han facilitado la comunicación entre el IRS y los contribuyentes, lo que ha llevado a una mayor comprensión de las obligaciones fiscales y a la presentación más precisa de declaraciones de impuestos. Esto puede reducir los errores y las auditorías posteriores, lo que beneficia tanto a los contribuyentes como al gobierno, generando además un impacto positivo en la recaudación de impuestos y la cumplimiento de las leyes fiscales.

### **Uso de la IA por la ICE**

La Oficina de Control de Inmigración y Aduanas (ICE) de los Estados Unidos se enfrenta a la compleja tarea de hacer cumplir las leyes de inmigración y aduanas en un país con una alta cantidad de movimientos transfronterizos. Esto incluye la identificación de individuos que puedan representar una amenaza para la seguridad nacional o que estén involucrados en actividades ilegales relacionadas con la inmigración.

Para lograr su objetivo, la ICE ha recurrido a la inteligencia artificial buscando así mejorar la identificación y el seguimiento de individuos que puedan estar relacionados con actividades ilegales de inmigración. Para esto, utiliza algoritmos de aprendizaje automático para analizar datos como registros de viaje, historiales criminales y patrones de comportamiento<sup>14</sup><sup>15</sup>. Estos algoritmos pueden identificar posibles amenazas de manera más rápida y precisa que los métodos tradicionales.

---

<sup>14</sup> Departamento de Seguridad Nacional de EEUU. Ver: [https://www.dhs.gov/data/AI\\_inventory](https://www.dhs.gov/data/AI_inventory)

<sup>15</sup> Ver: <https://fedscoop.com/dhs-adjusts-ai-use-case-inventory/>

Además, la IA se utiliza en la gestión de datos y la automatización de procesos dentro de la ICE, lo que mejora la eficiencia en la administración y el procesamiento de casos de inmigración.

La implementación de la IA en la ICE ha generado un impacto positivo en la capacidad de identificar posibles amenazas y actividades ilegales relacionadas con la inmigración, lo que contribuye a la seguridad nacional y permite una respuesta más rápida a situaciones de riesgo. Sin embargo, el uso de la IA en la aplicación de leyes de inmigración también ha generado preocupaciones en cuanto a la privacidad y los derechos civiles. La automatización de la toma de decisiones puede plantear desafíos en términos de transparencia y supervisión, tema que será abarcado con mayor profundidad en el siguiente capítulo.

## **HealthMap**

HealthMap es una plataforma creada por los PhD John Brownstein y Clark Freifeldque, operativa desde el año 2006 y que utiliza inteligencia artificial para rastrear y prevenir brotes de enfermedades. Su objetivo es identificar tempranamente la propagación de enfermedades infecciosas y proporcionar información oportuna a los profesionales de la salud y al público en general.

HealthMap recopila datos de una variedad de fuentes, incluidas noticias en línea, redes sociales, informes de vigilancia epidemiológica y más. Utiliza algoritmos de IA para analizar estos datos y detectar patrones que puedan indicar la aparición de un brote de enfermedad. La plataforma puede identificar ubicaciones geográficas específicas donde se informan casos inusuales o tendencias preocupantes.

La IA también puede clasificar la gravedad de los brotes y proporcionar información sobre las medidas de prevención y control que deben tomarse. Además, HealthMap permite a los profesionales de la salud y a las autoridades tomar decisiones informadas sobre la asignación de recursos y la planificación de la respuesta.

El uso de HealthMap ha tenido un impacto significativo en la detección temprana y la gestión de brotes de enfermedades. Ha permitido una respuesta más rápida y efectiva a situaciones de salud pública, lo que puede salvar vidas y reducir la propagación de enfermedades. Así, por ejemplo, en 2014 este software permitió detectar un brote de fiebre hemorrágica viral en el Oeste del continente africano, el cual fue posteriormente identificado por la Organización Mundial de la Salud (OMS) como Ébola<sup>16</sup>, lo que llevó a la plataforma a dedicar un mapa especialmente dedicado a dicha enfermedad, permitiendo a las autoridades locales y del resto del mundo adoptar medidas de prevención con una anticipación que, con las herramientas tradicionales probablemente no hubiese sido posible.

Además, HealthMap mejora la comunicación y la colaboración entre diferentes partes interesadas, como profesionales de la salud, autoridades gubernamentales y el público en general. Proporciona información precisa y oportuna, lo que contribuye a la toma de decisiones informadas.

### **Gestión de Tráfico en Amsterdam<sup>17</sup>**

Ámsterdam, como muchas ciudades congestionadas, utiliza la inteligencia artificial (IA) en la gestión del tráfico para abordar problemas como la congestión vehicular y mejorar la fluidez del tráfico. Para lo anterior, se ha empleado la IA para las siguientes tareas:

1. Recopilación de Datos en Tiempo Real: Ámsterdam cuenta con una infraestructura de sensores de tráfico, cámaras de vigilancia y otros dispositivos conectados que recopilan datos en tiempo real sobre el flujo de vehículos, la densidad del tráfico, las condiciones climáticas y otros factores relevantes.

---

<sup>16</sup> El Financiero. Ver: <https://www.elfinanciero.com.mx/tech/healthmap-la-herramienta-en-internet-que-detecto-el-brote-del-ebola-antes-que-la-oms/>

<sup>17</sup> Walraven, E., Spruijtenburg, D., Wilmink, I., & Schreuder, M. (Diciembre 2021). Artificial Intelligence and Traffic Management. Publicado por TrafficQuest.

2. **Análisis de Datos:** Los datos recopilados se transmiten a sistemas de IA que analizan la información en tiempo real. Los algoritmos de IA pueden identificar patrones y tendencias en los datos, como la congestión en ciertas áreas de la ciudad, la velocidad promedio del tráfico y la detección de accidentes.

3. **Optimización de Semáforos:** Uno de los aspectos clave de la gestión del tráfico es la optimización de semáforos. La IA utiliza los datos en tiempo real para ajustar automáticamente los tiempos de los semáforos en función del flujo de vehículos actual. Por ejemplo, si se detecta una congestión en una intersección, la IA puede dar prioridad a la dirección con mayor flujo de tráfico, ayudando así a reducir la congestión.

4. **Información en Tiempo Real para Conductores:** La ciudad puede proporcionar información en tiempo real a los conductores a través de aplicaciones de navegación y señales de tráfico electrónicas. Esto permite a los conductores tomar decisiones informadas sobre las rutas a seguir, evitando áreas congestionadas.

5. **Modelado de Tráfico Predictivo:** Además de gestionar el tráfico en tiempo real, la IA también se utiliza para predecir patrones de tráfico en el futuro. Por ejemplo, la IA puede anticipar la congestión durante las horas pico o en eventos importantes y ayudar a planificar desvíos y ajustes en la programación de semáforos.

6. **Integración con el Transporte Público:** La IA también puede integrarse con sistemas de transporte público, como autobuses y tranvías. Esto permite una coordinación más eficiente entre el transporte público y el tráfico en general, mejorando la movilidad en la ciudad.

### **Otros casos de uso de IA en el sector público**

Junto a los ejemplos analizados precedentemente, existen muchas otras aplicaciones de la IA en la administración, dentro de los cuales podemos mencionar someramente algunos como:

1. Chatbots de Atención al Ciudadano: Muchos gobiernos europeos han implementado chatbots de IA en sus sitios web para responder preguntas comunes de los ciudadanos y brindar información sobre servicios gubernamentales.
2. Traducción Automatizada: En la Unión Europea, se utiliza la IA para traducir documentos y comunicaciones entre diferentes idiomas, facilitando la colaboración transfronteriza.
3. Prevención del Fraude: Se emplea la IA para detectar fraudes en áreas como la seguridad social y los sistemas de pensiones, identificando irregularidades en las solicitudes y los pagos.
4. Predicción del Crimen: La IA se utiliza en la predicción de delitos y la asignación de recursos policiales en áreas con mayor probabilidad de incidentes.
5. Gestión de Recursos Naturales: En países con extensas áreas forestales, se utiliza la IA para el monitoreo de incendios forestales y la gestión sostenible de recursos naturales.
6. Educación: Se emplea la IA para personalizar la educación, adaptando el contenido y el ritmo de aprendizaje a las necesidades individuales de los estudiantes.
7. Ciberseguridad: La IA se utiliza en la detección y mitigación de amenazas cibernéticas para proteger los sistemas de información gubernamentales.
8. Planificación Urbana: La IA se utiliza en la planificación y desarrollo urbano, optimizando el uso de recursos y la eficiencia energética.
9. Asistencia a Personas con Discapacidad: Se desarrollan aplicaciones de IA para ayudar a personas con discapacidades visuales o auditivas a acceder a servicios gubernamentales en línea.

10. Predicción Climática: La IA se emplea en la predicción del clima y la gestión de desastres naturales, como inundaciones y sequías.

11. Agricultura Inteligente: Los sistemas de IA ayudan a los agricultores a optimizar el uso de recursos y a aumentar la producción de alimentos de manera sostenible.

12. Preservación del Patrimonio Cultural: La IA se utiliza en la conservación y restauración de arte y patrimonio cultural, así como en la catalogación de archivos históricos.

Estos son solo algunos ejemplos de cómo la IA se ha integrado en el sector público en Europa, abarcando una amplia gama de áreas para mejorar la eficiencia y la calidad de los servicios gubernamentales. La adopción continua de la IA promete transformar aún más la administración pública en la región.

Estos ejemplos destacan cómo la inteligencia artificial se está utilizando en la administración para abordar desafíos complejos y mejorar la eficiencia, la seguridad y la calidad de los servicios ofrecidos. Sin embargo, también plantean importantes cuestiones éticas y de privacidad que deben abordarse de manera cuidadosa a medida que la IA continúa desempeñando un papel creciente en la administración gubernamental y en la sociedad en general.

IA en el sector público: Caso de Chile

La abrupta llegada de la IA al mundo moderno ha significado un desafío muy importante para todas las naciones, y si bien algunas han adoptado roles de mayor protagonismo, Chile se ha mostrado como un país con grandes ambiciones en el rubro de la tecnología y sobretodo de la IA, posicionándose como uno de los países con mayor liderazgo en latinoamérica.

De esta manera, se hace imprescindible para el estudio de la IA proveer un contexto adecuado en torno a la realidad nacional. Así, este capítulo se dedicará exclusivamente a abordar la forma en que esta tecnología se ha introducido en la vida de los chilenos, lo cual se analizará principalmente desde la perspectiva pública.

Para lo anterior, el primer paso será analizar las áreas en las que se ha introducido la IA como herramienta para perseguir distintos objetivos (manejo de datos, gestión administrativa, sistemas de soporte, etc.). Luego, se realizará un estudio de la manera en que el Estado, a través de sus diversos organismos, ha regulado la implementación de esta tecnología para, finalmente, comentar de manera crítica y reflexiva este fenómeno en nuestro país.

### **Breve contextualización de la IA en Chile**

La IA llegó a Chile a mediados del siglo XX, a través de la investigación académica en universidades como la Universidad de Chile y la Pontificia Universidad Católica de Chile. En sus inicios, la IA se desarrolló en áreas como la lógica, la computación y la matemática.

En la década de 1950, la informática estaba en sus inicios, y los primeros computadores eran enormes y costosos. Sin embargo, el desarrollo de la tecnología de transistores y circuitos integrados permitió que los computadores disminuyeran sus tamaños y costos, lo que los hizo más accesibles y asequibles.

En este contexto, un grupo de investigadores chilenos comenzó a trabajar en el desarrollo de la IA. En 1957, el profesor Fernando Flores creó el Laboratorio de Inteligencia Artificial de la Universidad de Chile, que fue el primer laboratorio de IA en América Latina. Luego, en 1963, el profesor Hugo de Garis creó el Instituto de Investigaciones de Inteligencia Artificial de la Pontificia Universidad Católica de Chile, que fue el segundo laboratorio de IA en América Latina.

Los primeros desarrollos de IA en Chile se centraron en áreas como la lógica, la computación y la matemática. Así, los investigadores chilenos desarrollaron sistemas de razonamiento lógico para la resolución de problemas (SINS y ALBA); algoritmos y técnicas de computación para la IA (ID3, C4.5); modelos matemáticos para la IA (RHEA, ILP), entre otros.<sup>18</sup>

Con el paso del tiempo y la consecuente masificación de la IA, surgieron nuevos usos de estas tecnologías que fueron prontamente aprovechadas por diversos actores del sector privado, dentro de los cuales ha destacado el sector empresarial, en donde empresas nacionales como NotCo (empresa de alimentos que ha desarrollado una serie de productos alternativos a la carne, leche y huevos, utilizando IA para crear nuevos sabores y texturas), Fintual (plataforma de inversión en línea que utiliza IA para recomendar portafolios de inversión personalizados a sus clientes) o Cornershop (aplicación de delivery que utiliza IA para optimizar la asignación de pedidos a repartidores, lo que ha permitido reducir los tiempos de entrega) han logrado ofrecer mejores servicios gracias al uso de IA.

## **IA en el Sector público chileno**

Si bien el desarrollo de la IA en nuestro país ha tenido mayor presencia y avances en el sector privado, siendo utilizado para múltiples objetivos, como mejorar la atención al cliente, gestionar cantidades masivas de datos, mejorar la productividad, etc., este trabajo se abocará particularmente al análisis del uso de IA en el sector público.

Así, y para los efectos de explicar adecuadamente cómo se ha introducido la IA en el sector público de nuestro país, este capítulo se basará en un estudio exploratorio realizado por la Escuela de Gobierno de la Universidad Adolfo Ibáñez (GobLab UAI), cuyo objetivo es “conocer un estado del arte del uso de estos sistemas, partiendo por

---

<sup>18</sup> Abeliuk, A. (2021). Historia y evolución de la inteligencia artificial. Revista Universidad de Chile.

la información básica respecto de su existencia, identificación, propósito, los datos que utiliza y su información pública disponible”<sup>19</sup>.

Antes que nada, se hace necesario precisar que dicho estudio utilizó un concepto más amplio que el de Inteligencia Artificial, empleando así el de “Sistemas de Decisiones Automatizadas” (en adelante, SDA), el cual es definido por los autores de la siguiente forma:

*“Aquellos que realizan acciones con cierto grado de autonomía para lograr objetivos específicos. Estos sistemas pueden basarse exclusivamente en software o hardware, proveyendo de un insumo para una decisión realizada por un funcionario público o bien tomar directamente una decisión.”*

Pese a lo anterior, para los efectos de este trabajo se considerará adecuado el uso del concepto de SDA en el estudio mencionado, toda vez que permite de todas formas observar cómo han sido implementadas estas tecnologías en el sector público.

De esta forma, el estudio inició con una consulta a 870 servicios del Estado, de los cuales solamente participó un tercio, el cual informó el uso de 219 SDA, de los cuales a su vez, se excluyó a un 59,4% por no estar dentro de la definición de sistemas utilizado. Luego de un proceso de depuración de datos, se informaron 92 sistemas, dentro de los cuales 4 son compartidos y utilizados por diversas entidades:

- **DART:** declarado por 5 entidades correspondientes a corporaciones municipales y servicios de salud. DART es una plataforma de teleoftalmología automatizada para tamizaje de retinopatía diabética en base a retinografías, en otras palabras, busca el desarrollo de soluciones basadas en IA que permitan asistir y aumentar la capacidad de atención de los profesionales, prestadores y aseguradoras de salud.

---

<sup>19</sup> Garrido, R., Lapostol, J. P., & Hermosilla, M. P. (2021). Transparencia algorítmica en el sector público. Escuela de Gobierno de la Universidad Adolfo Ibáñez y el Consejo para la Transparencia.

- **SIGGE:** declarado por 4 entidades correspondientes a corporaciones municipales de salud. SIGGES (Sistema de Información para la Gestión de Garantías Explícitas de Salud) es un sistema de registro que utiliza IA para la Retinopatía Diabética, analizando de forma automatizada las imágenes obtenidas para detectar posibles alteraciones.

- **Rayen:** declarado por 8 entidades correspondientes a corporaciones municipales de salud. Rayen es una empresa chilena que otorga servicios de Tecnología e Información para diversas Redes de Salud a lo largo del país, la cual utiliza IA para mejorar la gestión sanitaria y la atención al público.

- **Sistema de Admisión Escolar:** declarado por 3 entidades correspondientes a organismos municipales. Este sistema es una plataforma web por medio de la cual se realizan postulaciones a los colegios públicos y particulares subvencionados, que de manera automatizada aplica criterios preestablecidos para asignar los cupos, en los casos en que sean insuficientes para la cantidad de postulantes.

Dentro de los sistemas utilizados en la administración, se puede observar su uso aplicado a materias muy diversas, tales como:

- Redes neuronales para detección de marcas en plataforma del Instituto Nacional de Propiedad Industrial.
- Reconocimiento facial con inteligencia artificial en estaciones de Metro Valparaíso.
- Videovigilancia con reconocimiento facial de la Municipalidad de Las Condes y la Policía de Investigaciones.
- Análisis de imágenes satelitales de humedales altoandinos por parte de la Superintendencia del Medio Ambiente.
- Telemedicina con inteligencia artificial para detectar covid-19 en posta central.
- Mateo: detección de plagas y enfermedades de hortalizas online.

## **Regulación de la IA**

La IA ha experimentado un rápido desarrollo en los últimos años, el cual ha generado un gran entusiasmo por el potencial de dicha tecnología para mejorar nuestra vida, pero también ha planteado una serie de desafíos, entre ellos los relacionados con la ética, la privacidad y la seguridad.

En Chile, la regulación de la IA es un proceso que se encuentra en sus primeras etapas. Las primeras discusiones sobre la regulación de la IA en Chile se remontan a la década de 2010, cuando se comenzó a reconocer el potencial de esta tecnología para generar impactos significativos en la sociedad.

### *La década de 2010: los primeros pasos hacia la regulación*

En la década de 2010, las primeras discusiones sobre la regulación de la IA en Chile se centraron en los aspectos éticos de esta tecnología. En este sentido, se debatió sobre la necesidad de garantizar que la IA se desarrolle y use de manera responsable y ética, respetando los derechos humanos y los valores de la sociedad chilena.

Este debate se vio reflejado en la creación del Comité Asesor del Gobierno en Inteligencia Artificial, en 2019. El comité fue un organismo asesor del gobierno chileno, compuesto por un grupo de 10 expertos, y que tuvo como objetivo desarrollar recomendaciones para la regulación de la IA. Así, según explicó el ex ministro Andrés Couve, el comité buscó en un inicio aportar con un diagnóstico de la situación de Chile en inteligencia artificial junto con propuestas en las tres áreas del plan de trabajo que lidera el Ministerio de Ciencia, señalando que:

*“Lo que haremos como gobierno es trabajar en tres ejes complementarios de este plan de trabajo que son factores habilitantes, desarrollo y adopción de inteligencia artificial y un último ámbito vinculado a ética, estándares, seguridad y regulación del uso de estas tecnologías. La selección del panel de expertos refleja esta visión pues*

*cada integrante tiene experticia en uno o más de estos temas desde el mundo de la academia, el mundo público, la empresa y la sociedad civil”<sup>20</sup>*

*La Política Nacional de Inteligencia Artificial: un marco para el desarrollo de la IA.<sup>21</sup>*

Desde ese momento hasta fines del 2021, se reunió un diverso grupo interdisciplinario compuesto por 12 expertos de la academia, sector productivo y sociedad civil. Juntos, lideraron un proceso de coordinación del sector público a través de un comité interministerial.

Simultáneamente, se promovió un diálogo nacional que incluyó la participación abierta de más de 1.300 personas en talleres, 400 personas en eventos en diferentes regiones del país y más de 5.300 personas en 15 reuniones que abordaron la IA desde diversas perspectivas y disciplinas. Todos estos aportes se condensaron en un borrador<sup>22</sup> del documento final, que fue sometido a un proceso de consulta pública, en el que participaron más de 200 personas naturales y jurídicas. Los aportes recibidos en ese proceso se sistematizaron, analizaron detenidamente y se consolidaron en colaboración con las discusiones ministeriales y expertos, culminando en la primera Política Nacional de Inteligencia Artificial de Chile, publicada a fines del 2021.

El objetivo del Plan era posicionar a Chile en la vanguardia de la colaboración global en el campo de la Inteligencia Artificial, mediante la creación de un ecosistema de investigación, desarrollo e innovación en IA que potenciara las capacidades de los sectores productivos, académicos y gubernamentales. Estos esfuerzos, guiados por

---

<sup>20</sup> Gobierno de Chile. Ver: <https://www.gob.cl/noticias/ministerio-de-ciencia-presenta-comite-asesor-del-gobierno-en-inteligencia-artificial/>

<sup>21</sup> Ministerio de Ciencia, Tecnología, Conocimiento e Innovación. Ver: [https://www.minciencia.gob.cl/uploads/filer\\_public/bc/38/bc389daf-4514-4306-867c-760ae7686e2c/documento\\_politica\\_ia\\_digital\\_.pdf](https://www.minciencia.gob.cl/uploads/filer_public/bc/38/bc389daf-4514-4306-867c-760ae7686e2c/documento_politica_ia_digital_.pdf)

<sup>22</sup> Ministerio de Ciencia, Tecnología, Conocimiento e Innovación. Ver: [https://www.minciencia.gob.cl/legacy-files/borrador\\_politica\\_nacional\\_de\\_ia.pdf](https://www.minciencia.gob.cl/legacy-files/borrador_politica_nacional_de_ia.pdf)

principios fundamentales de oportunidad y responsabilidad, estaban destinados a impulsar el desarrollo sostenible y mejorar nuestra calidad de vida.

De esta forma, el Plan tiene como pilares fundamentales cuatro principios transversales y tres ejes para estructurarlo, los cuales se analizan a continuación:

#### Principios Transversales:

##### 1. IA Con Centro En El Bienestar De Las Personas, Respeto A Los Derechos Humanos Y La Seguridad

El plan busca orientar la IA primordialmente hacia el bienestar integral de las personas, procurando una mejor calidad de vida mediante la optimización de los beneficios inherentes a la IA y la exhaustiva consideración de sus riesgos y posibles impactos adversos. Por este motivo, se puso énfasis en la inquebrantable observancia de los derechos humanos de la ciudadanía en su conjunto. Un aspecto crucial radicó en la seguridad inherente a los algoritmos y datos utilizados, especialmente en casos que involucraban datos personales. La consideración sistemática de los riesgos y vulnerabilidades asociados a estos procesos resultó fundamental, evitando, de este modo, la utilización de conjuntos de datos propensos a conducir a decisiones o resultados arbitrariamente discriminatorios. Además, se subrayó la importancia de la protección de algoritmos de carácter crítico, cuya vulnerabilidad podría comprometer la seguridad nacional.

##### 2. IA Para El Desarrollo Sostenible

A su vez, se reconoció el potencial que la IA albergaba para permitir a naciones en desarrollo, como es el caso de Chile, diversificar su matriz económica y elevar la productividad de sus industrias. Consecuentemente, se pretende fomentar de manera activa el uso y desarrollo de la tecnología de IA en nuestro país. Este impulso se orientó, asimismo, hacia el fortalecimiento del ecosistema de investigación,

tecnología e innovación, promoviendo su integración en el contexto de un desarrollo sostenible que considerara las dimensiones sociales y medioambientales.

### 3. IA Inclusiva

La vinculación inextricable entre el proceso de entrenamiento de algoritmos y el uso de datos subrayó la importancia de la transparencia y explicabilidad en el diseño de una IA inclusiva. Así, se resaltó la necesidad de poner un énfasis especial en la integridad y calidad de los datos con el objetivo de detectar y abordar de manera apropiada sesgos potenciales. En esa perspectiva, se postuló que la IA debía evitar la discriminación con base en categorías protegidas y abstenerse de utilizarse en detrimento de cualquier grupo de individuos. Especial énfasis se puso en el desarrollo de la IA con una perspectiva de género y diversidad sexual, así como en su inclusión en grupos históricamente marginados, tales como los pueblos originarios, las personas con discapacidades y los sectores económicos más vulnerables, con el propósito de que la IA se erigiera en una herramienta que promoviera la equidad y colaborara en la reducción de disparidades.

### 4. IA Global Y En Constante Evolución

En lo que respecta a la adopción de una posición en el ámbito internacional y la interacción con el contexto global de la IA, se consideró la inserción en espacios bilaterales y multilaterales que involucraban a Chile. Esta integración se guió por la alineación con los principios y acuerdos suscritos por el país, tal como los Principios sobre la IA de la OCDE. Además, se estableció un enfoque de constante revisión y adaptación, ponderando que el desarrollo global de la IA y las conversaciones en torno a los dilemas éticos que planteaba podían entrar en conflicto con realidades y necesidades locales. En ese sentido, se hizo especial hincapié en la consideración de las peculiaridades de Chile, adoptando un enfoque propio y sujeto a un proceso de evolución constante, independientemente de las trayectorias seguidas por naciones en el hemisferio septentrional.

## Ejes de la Política

### Eje N°1: Factores Habilitantes

Este eje se refiere a los factores habilitantes que fundamentaron la existencia y la implementación de la IA. Estos factores se configuran como elementos imprescindibles, sin los cuales la utilización y el desarrollo de esta tecnología se tornan inviables. Los componentes que abarca este eje comprenden el desarrollo de talentos, infraestructura tecnológica, y datos.

### Eje N°2: Desarrollo y Adopción

Este eje corresponde al ámbito donde se concibió y expandió la IA, englobando a los actores responsables de la creación, provisión y demanda de sus diversas aplicaciones y técnicas. Entre estos actores se incluyen la comunidad académica, el Estado, el sector privado y la sociedad civil. Las áreas abarcadas en este eje comprenden la investigación tanto en su modalidad básica como aplicada, la transferencia tecnológica, la innovación, el emprendimiento, la mejora de los servicios públicos, y el desarrollo económico basado en tecnología, entre otros aspectos.

### Eje N°3: Ética, Aspectos Normativos y Efectos Socioeconómicos

Este eje, considerado el más heterogéneo dentro de la Política de IA, abordó las nuevas discusiones surgidas en torno a la interacción entre el ser humano y las máquinas, así como el sistema socio-técnico resultante de esta interacción. Se reconoció que algunas de estas discusiones se encontraban en una etapa avanzada, mientras que otras se hallaban en fases iniciales de desarrollo. Este eje incorpora aspectos normativos de gran relevancia, de acuerdo con el diálogo ciudadano y las experiencias nacionales e internacionales. Estos aspectos comprenden, por ejemplo, el papel de la IA en la protección al consumidor, en la privacidad, en el régimen de propiedad intelectual y en la ciberseguridad. Asimismo, se abordan cuestiones como

la transparencia y la explicabilidad de los sistemas, así como su impacto en cuestiones de género y en el mercado laboral, entre otros aspectos de índole ética, social y económica.

Respecto a este último punto, es altamente valorable la pretensión del Plan de impulsar la transparencia algorítmica (objetivo 3.1.2 del Plan), dado que debe considerarse como un pilar en la transformación digital de los gobiernos. Así, el cuarto capítulo de este trabajo se enfocará en abordar de forma más detallada la transparencia como principio fundamental en la implementación de IA en el sector público.

### *La Ley N° 21.383: un paso adelante en la regulación*

El año 2020, un grupo de Senadores ingresaron una Moción Parlamentaria con el objetivo de modificar la Constitución Política de la República, en miras de preparar a nuestro país, desde la perspectiva normativa, frente a los desafíos que trajo consigo el exponencial desarrollo tecnológico de las últimas décadas.

El proyecto, se basaba principalmente en agregar un inciso nuevo al artículo 19 N°1 de nuestra Carta Fundamental, y se señaló que:

*“La incorporación de un nuevo inciso en el artículo 19 de la Constitución tiene por objeto plasmar en el texto constitucional algunos elementos esenciales para la debida protección de los derechos humanos ante el desarrollo de la neurotecnología, constituyéndose así las ideas matrices del proyecto.”<sup>23</sup>*

Finalmente, en el 2021, el gobierno aprobó la Ley N° 21.383, que modifica la Constitución Política de la República para establecer el desarrollo científico y

---

<sup>23</sup> Ver historia de la Ley en: <https://www.bcn.cl/historiadelailey/nc/historia-de-la-ley/7926/>

tecnológico al servicio de las personas, agregando lo siguiente como inciso final del artículo 19 N°1:

*“El desarrollo científico y tecnológico estará al servicio de las personas y se llevará a cabo con respeto a la vida y a la integridad física y psíquica. La ley regulará los requisitos, condiciones y restricciones para su utilización en las personas, debiendo resguardar especialmente la actividad cerebral, así como la información proveniente de ella;”*

Esta ley significó un paso adelante en la regulación de la IA en Chile, ya que reconoce la importancia de que esta tecnología se desarrolle y use en beneficio de las personas.

#### *El proyecto de ley de regulación de la IA: un paso pendiente*

En abril del año 2023, un grupo de diputados presentó el primer proyecto de Ley que “Regula los sistemas de inteligencia artificial, la robótica y las tecnologías conexas, en sus distintos ámbitos de aplicación” en nuestro país. Dicho proyecto tiene una serie de consideraciones previas que proporcionan un marco para la regulación.

En primer lugar, se reconoce que la IA es una tecnología que tiene el potencial de generar un gran impacto en la sociedad, tanto positivo como negativo. Por ello, es necesario contar con una regulación que garantice su desarrollo y uso de manera responsable y ética, respetando los derechos humanos. A su vez, señala que la regulación de la IA debe ser un proceso democrático, que involucre a todos los actores interesados. El proyecto de ley propone la creación de una Comisión Nacional de Inteligencia Artificial que será responsable de la regulación y supervisión de la IA en Chile. Esta comisión estaría integrada por representantes del gobierno, la academia, la industria y la sociedad civil.

En la misma línea, el proyecto identifica una serie de riesgos asociados al desarrollo y uso de la IA. Estos riesgos incluyen la discriminación, la vulneración de la privacidad y la seguridad, y el impacto negativo en el medio ambiente. Para esto,

propone una serie de medidas para mitigarlos, como la prohibición de sistemas de IA que sean calificados de riesgo inaceptable y la regulación de sistemas de IA que sean calificados de alto riesgo.

Dentro de las medidas que destacan en el proyecto, se pueden considerar las siguientes:

- La creación de una Comisión Nacional de Inteligencia Artificial que será responsable de la regulación y supervisión de la IA en Chile.
- La definición de un procedimiento para que los sistemas de IA sean evaluados y autorizados por la Comisión.
- La prohibición de sistemas de IA que sean calificados de riesgo inaceptable dado que su uso es perjudicial y atenta contra los derechos fundamentales garantizados por el Estado de Chile.
- La regulación de sistemas de IA que sean calificados de alto riesgo, incluyendo requisitos de transparencia e información a los usuarios.
- La creación de un registro de sistemas de IA autorizados para ser desarrollados, comercializados, distribuidos y utilizados en territorio nacional.
- Sanciones para los desarrolladores, proveedores y usuarios que incumplan las normas contenidas en la normativa.

Teniendo en consideración que el proyecto a la fecha de hoy<sup>24</sup> sigue en trámite legislativo y posiblemente sufra modificaciones, este trabajo omitirá realizar comentarios respecto a su contenido. Sin embargo, es menester reconocer el esfuerzo que se está realizando por diversos actores de la sociedad para abordar de la mejor manera posible el fenómeno de la IA, pese a que desde la perspectiva regulatoria aún no existan las herramientas necesarias.

---

<sup>24</sup> En diciembre del 2023, el proyecto se encuentra en el primer trámite constitucional.

## Capítulo 3

### Desafíos de la IA en el sector público

#### Riesgos que implica el uso de IA en el sector público

La introducción de la IA en el sector público representa un hito en la evolución de las administraciones gubernamentales. La promesa de eficiencia, toma de decisiones fundamentadas en datos y la mejora en la prestación de servicios públicos ha llevado a un entusiasmo generalizado en la adopción de estas tecnologías. Sin embargo, es imperativo reconocer que la IA no es una panacea sin riesgos o desafíos, tanto desde la perspectiva ética, como regulatoria y social. En este capítulo, se explorarán los riesgos intrínsecos que implica el uso de la IA en el sector público, centrándonos en aspectos como el impacto laboral, sesgos, falta de control y otros factores que hacen que la IA sea una herramienta poderosa pero potencialmente peligrosa.

La IA se distingue por varias características fundamentales que la hacen poderosa, pero también impredecible y arriesgada. Una de las características clave es su autonomía. Los sistemas de IA pueden operar de forma autónoma, tomando decisiones sin intervención humana en tiempo real. Esta autonomía puede ser beneficiosa en términos de eficiencia y velocidad, pero también plantea preocupaciones sobre la falta de supervisión y control humano.

Además, la IA es impredecible en gran medida. A menudo, los resultados de las acciones de un sistema de IA no son fácilmente comprensibles o previsibles para los seres humanos. Los algoritmos de aprendizaje automático y las redes neuronales profundas, que son fundamentales en la IA, operan en base a enormes conjuntos de datos y complejas interconexiones, lo que hace que sus decisiones sean difíciles de rastrear o explicar. Esta falta de transparencia plantea interrogantes sobre cómo tomar decisiones informadas y responsables en un contexto gubernamental.

La IA también es difusa en su implementación. En el sector público, los sistemas de IA pueden estar distribuidos en diversos departamentos y aplicaciones, lo que hace que la responsabilidad y la gestión se vuelvan complicadas. Múltiples actores pueden estar involucrados en la adopción de la IA, lo que aumenta el riesgo de falta de coordinación y coherencia en su uso. Además, su discreción permite que múltiples componentes de un sistema de IA se diseñen en diferentes lugares y momentos sin una coordinación consciente, lo que dificulta la identificación y corrección de problemas sistémicos.

La falta de control es otra característica inquietante de la IA. A medida que los sistemas de IA aprenden y se adaptan, pueden desviarse de las intenciones iniciales de sus diseñadores. Incluso los programadores de IA pueden no ser capaces de predecir o entender completamente cómo tomará decisiones un sistema altamente autónomo. Esto presenta un desafío significativo en el sector público, donde la transparencia y la rendición de cuentas son fundamentales para la toma de decisiones éticas y responsables.

Estos rasgos inherentes de la IA conllevan riesgos en el sector público que deben abordarse cuidadosamente. Uno de los riesgos más evidentes es el impacto laboral. A medida que la IA asume tareas y roles anteriormente desempeñados por humanos, se plantea la preocupación de la pérdida de empleo. Esto no solo afecta a los trabajadores cuyos trabajos son automatizados, sino que también plantea cuestiones más amplias sobre la reestructuración de la fuerza laboral y la formación necesaria para adaptarse a esta nueva realidad.

Además, los sesgos en la IA son una preocupación crucial. Los algoritmos de aprendizaje automático pueden aprender sesgos de los datos en los que se basan, lo que puede resultar en discriminación o decisiones injustas. En el sector público, esto puede traducirse en un trato desigual de los ciudadanos, lo que socava la confianza en las instituciones gubernamentales y puede dar lugar a problemas legales y éticos.

Este capítulo explorará estos riesgos de manera más detallada, así como otros desafíos que plantea el uso de la IA en el sector público. La adopción de la IA en el sector público es un paso significativo hacia la modernización, pero es esencial comprender y gestionar los riesgos para aprovechar al máximo sus beneficios sin comprometer la integridad y la responsabilidad en el servicio público.

## **Impacto laboral**

La introducción de la IA en el sector público de diversas zonas geográficas del mundo, como EEUU o Europa, ha generado un interesante debate en torno al impacto en el empleo y las dinámicas laborales. En este sentido, la capacidad de esta tecnología de procesar cantidades inhumanas de información y a una velocidad de la cual no existen precedentes, ha dado paso a su implementación para la realización de diversas tareas, las cuales habían sido tradicionalmente realizadas por humanos. Para poder entender adecuadamente el impacto de la IA en materia laboral, será necesario responder las siguientes preguntas:

- a) Qué trabajos y tareas están en riesgo por la IA?

Si bien los avances tecnológicos anteriores habían generado un impacto en trabajos rutinarios y generalmente de “menor habilidad” (por ejemplo, operadores de máquinas, inspectores de calidad o empleados de oficina), la IA tiene el potencial de incidir en todas las áreas laborales de nuestra sociedad. En este sentido, esta tecnología pone en riesgo tanto a los trabajos que exigen poca habilidad como a los que exigen un alto nivel de preparación (ej. Médicos cirujanos), razón por la cual debe estudiarse profundamente la implementación de la IA en el sector público.

Un estudio señala que el grupo económico más afectado por la introducción de esta tecnología sería el grupo de ingresos medios, toda vez que los trabajos más rutinarios se concentran en ese grupo (ej. Empleados de oficina u operadores de maquinarias),

a diferencia de aquellos con menores ingresos (ej. Garzones, mujeres encargadas del aseo, etc.) y con mayores ingresos (ej. Gerentes, empresarios, políticos, etc.).<sup>25</sup>

El mismo estudio, distingue también entre géneros, en este sentido, señala que la introducción de la IA afectaría más a los hombres que a las mujeres<sup>26</sup>, puesto que éstas últimas presentan una tendencia marcada a optar por trabajos que exigen tareas de mayor interacción interpersonal.

De todo lo anterior, se concluye que si bien el impacto de la IA en los trabajos es transversal, pudiendo afectar a cada grupo de nuestra sociedad, sí es posible sostener que hay ciertos trabajos y tareas que por sus características serán más susceptibles de ser reemplazados por esta tecnología, lo cual debe ser un punto de partida esencial a considerar a la hora de integrar la IA en nuestra sociedad.

#### b) Qué trabajos y tareas surgirán a raíz de la IA?

Si bien la introducción de la IA tiende a llevarnos a imaginar que muchos trabajos y tareas serán automatizadas, también es cierto sostener que implicará la creación de nuevas áreas laborales dada la demanda que generará. En este sentido, cabe recordar cómo los avances tecnológicos del último siglo han dado paso a la creación de nuevas carreras y profesiones, tales como desarrolladores de Apps, conductores de uber, Community Managers, Técnicos en energías renovables o, incluso, hasta tiktokers. Es más, si nos basamos en la experiencia de EEUU, apenas el 60% de los trabajos ocupados el año 2018 existían en 1940<sup>27</sup>, lo que implica que los avances tecnológicos efectivamente permiten crear empleos nuevos.

Ahora bien, es fundamental destacar el carácter “bidireccional” de la innovación tecnológica. Esto significa que si bien existe una tendencia a la automatización en ciertas áreas laborales, existen otras que, o no se ven afectadas (ej. Educadoras de párvulo) o incluso crecen (ej. Ingenieros).

---

<sup>25</sup> Webb, M. (Enero 2020), p. 31. The Impact of Artificial Intelligence on the Labor Market. Stanford University.

<sup>26</sup> *Ídem*, p. 43.

<sup>27</sup> Autor, D., Chin, C., Salomons, A. M., & Seegmiller, B. (2022). New Frontiers: The Origins and Content of New Work, 1940–2018 (p.12). National Bureau of Economic Research. <http://www.nber.org/papers/w30389>

Respecto del impacto que podría verse particularmente en funcionarios públicos, esta afectación puede manifestarse de dos formas: en la modificación de los roles que cumplen o derechamente en su eliminación. Para entender adecuadamente cómo el Estado debiese abordar el impacto de la implementación de sistemas automatizados en el rol de los funcionarios públicos, el Banco Interamericano de Desarrollo (en adelante, BID) confeccionó una gráfica que lo explica. (Ver figura 2)

De lo expuesto anteriormente, se puede concluir que si bien la implementación de la IA en nuestra sociedad generará un impacto en los empleos y tareas desarrolladas tradicionalmente por humanos, lo cual se podría reflejar en modificación o eliminación de diversos roles, traerá también aparejada la posibilidad de crear nuevos roles que permitirán, en el tiempo, generar opciones de trabajo.

En este sentido, debe reconocerse la inclusión de objetivos para mitigar estos efectos por parte de la Política Nacional de IA, dentro de los que se encuentran el *“Realizar análisis prospectivos para detectar activamente las ocupaciones más vulnerables, anticipar la creación de nuevos empleos por IA y apoyar a los trabajadores en la transición a nuevas ocupaciones, minimizando sus costos personales y familiares”*<sup>28</sup> y *“Proveer apoyo a los trabajadores frente a la automatización”*<sup>29</sup>.

---

<sup>28</sup> Política Nacional de Inteligencia Artificial, p. 54.

<sup>29</sup> *Ídem*, p. 56.

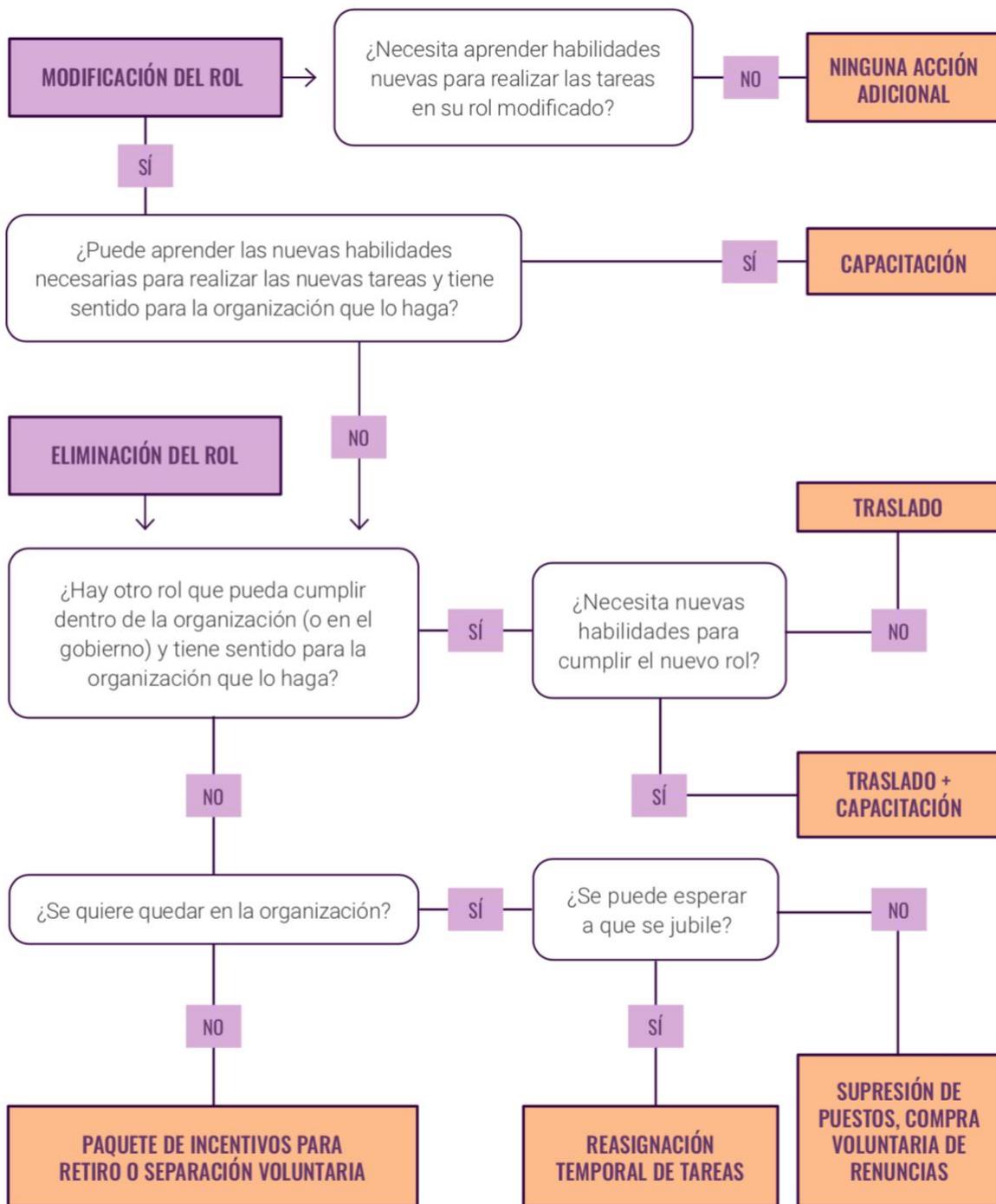


Figura 2. Los posibles caminos que puede seguir un funcionario<sup>30</sup>

<sup>30</sup> Porrúa, M., Lafuente, M., Mosqueira, E., Roseth, B., & Reyes, A. M. (Mayo de 2021). Transformación digital y empleo público: el futuro del trabajo del gobierno. Banco Interamericano de Desarrollo. (p. 180).

## **Rezago**

El rezago, entendido en el contexto de la implementación de la IA en el sector público, representa un desafío significativo que requiere una atención detallada. Este concepto se refiere al atraso o falta de avance en la adopción y aprovechamiento pleno de tecnologías de IA en las estructuras gubernamentales. A pesar de que la IA promete transformar la administración pública y mejorar la eficiencia de los servicios gubernamentales, el rezago puede obstaculizar su plena realización.

El rezago se manifiesta de diversas maneras, lo que subraya la complejidad del desafío. Por un lado, existe una falta de integración de sistemas y datos, ya que muchas agencias gubernamentales operan con sistemas de información aislados que no se comunican entre sí, lo que resulta en redundancia de la recopilación de datos y dificulta el intercambio de información crítica entre agencias. La automatización de procesos administrativos y operativos es uno de los beneficios más prometedores de la IA en el sector público, pero el rezago puede llevar a la falta de iniciativas para automatizar tareas y procesos que aún dependen en gran medida de la intervención humana, lo que ralentiza la eficiencia y aumenta los costos operativos.

Además, el rezago puede resultar en la falta de automatización de procesos administrativos y operativos que podrían mejorar la eficiencia. Esto incluye tareas y procesos que aún dependen en gran medida de la intervención humana, lo que ralentiza la eficiencia y aumenta los costos operativos.

La desigualdad en el acceso a servicios es otra manifestación crítica del rezago. Cuando la IA se implementa de manera desigual o selectiva en el sector público, puede resultar en desigualdades en el acceso a servicios gubernamentales mejorados. Esto puede ser especialmente problemático en áreas como la atención médica, la educación o los servicios sociales, donde la falta de equidad en el acceso puede afectar negativamente a ciertos grupos de la población.

Finalmente, los retrasos en la toma de decisiones son comunes en el sector público cuando existe rezago en la implementación de soluciones de IA. La IA puede proporcionar análisis de datos más rápidos y precisos para respaldar la toma de decisiones en el gobierno. Sin embargo, cuando existe rezago en la implementación de estas soluciones, los funcionarios pueden depender de procesos manuales y análisis más lentos, lo que resulta en retrasos en la toma de decisiones críticas.

Todo lo mencionado previamente habla de las manifestaciones internas del rezago en una nación, lo cual si bien tiene gran relevancia, adopta un carácter particularmente importante si consideramos el rezago como un problema a nivel global, que podría acentuar aún más la brecha entre los países más desarrollados y aquellos más pobres.

En este sentido, un estudio del Fondo Monetario Internacional<sup>31</sup> “concluye que las nuevas tecnologías amenazan con ampliar la brecha entre los países ricos y pobres al desviar la inversión hacia las economías avanzadas, donde la automatización ya está consolidada”<sup>32</sup>. (*Ver Figura 3*)

Dicho estudio se basa en un modelo que analiza dos países, uno avanzado y otro en desarrollo, en términos de su producción de bienes mediante tres factores: mano de obra, capital y robots. Su hipótesis central es que los robots reemplazan a los trabajadores, y la revolución de la IA se caracteriza por la mejora de la productividad de los robots.

La divergencia económica entre las economías avanzadas y en desarrollo se puede explicar a través de tres canales. Primero, la participación en la producción es clave, ya que las economías avanzadas con salarios más altos tienden a usar robots de manera más intensiva. Segundo, los flujos de inversión se desvían hacia las

---

<sup>31</sup> Alonso, C., Berg, A., Kothari, S., Papageorgiou, C., & Rehman, S. (2020, septiembre). Will the AI Revolution Cause a Great Divergence? (IMF Working Paper No. WP/20/202). International Monetary Fund.

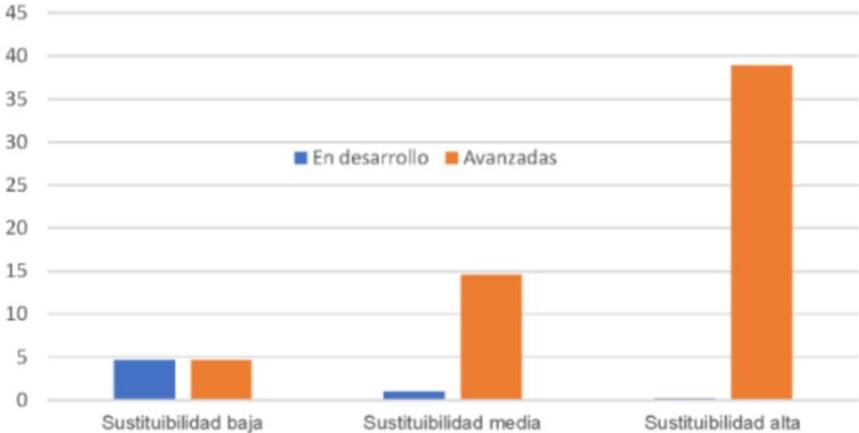
<sup>32</sup> Fondo Monetario Internacional

Ver <https://www.imf.org/es/Blogs/Articles/2020/12/02/blog-how-artificial-intelligence-could-widen-the-gap-between-rich-and-poor-nations>

economías avanzadas debido a su mayor uso de robots. Tercero, los términos de intercambio se ven afectados, ya que la revolución de los robots reduce la demanda de mano de obra no cualificada en las economías en desarrollo, lo que puede disminuir los precios relativos de sus productos y afectar negativamente su PIB.

En resumen, la adopción de robots y la revolución de la IA tienen implicaciones significativas para la economía global y pueden contribuir a la divergencia entre economías avanzadas y en desarrollo.

**Niveles de renta divergentes**  
 La brecha entre el PIB per cápita de las economías avanzadas y en desarrollo se amplía cuanto más sencillo es que los robots sustituyan a los trabajadores.  
 (variaciones porcentuales del PIB per cápita)



Fuente: Cálculos de los autores.

Figura 3. Niveles de Renta Divergentes<sup>33</sup>

Así, los autores concluyen lo siguiente:

*“Las mejoras de la productividad de los robots impulsan la divergencia entre los países avanzados y en desarrollo cuando los*

<sup>33</sup> Fondo Monetario Internacional

*robots sustituyen con facilidad a los trabajadores. Además, aunque esas mejoras tenderán a aumentar los ingresos, también incrementarán la desigualdad del ingreso, al menos durante la transición y, posiblemente, a largo plazo para algunos grupos de trabajadores, tanto en las economías avanzadas como en desarrollo.”<sup>34</sup>*

Explicadas las implicancias y manifestaciones del rezago en la implementación de la IA en el sector público, se hace fundamental analizar sus causas, entre las que se incluyen:

- Recursos Limitados: Las restricciones presupuestarias a menudo limitan la capacidad de las agencias gubernamentales para invertir en tecnologías de IA. La adquisición de software y hardware de IA puede ser costosa, y las agencias pueden carecer de los recursos necesarios para realizar estas inversiones.
- Resistencia al Cambio: Las estructuras gubernamentales suelen ser resistentes al cambio. Los empleados pueden mostrar resistencia a la adopción de nuevas tecnologías por temor a la interrupción de las prácticas existentes o al cambio en sus roles y responsabilidades. Esta resistencia puede obstaculizar la implementación de la IA.
- Complejidad y Falta de Estándares: La IA es un campo en constante evolución con una amplia variedad de enfoques y tecnologías. La falta de estándares claros y la complejidad del campo pueden hacer que sea difícil para las agencias gubernamentales determinar cuáles son los enfoques más adecuados y cómo integrarlos de manera efectiva.

---

<sup>34</sup> Fondo Monetario Internacional  
Ver <https://www.imf.org/es/Blogs/Articles/2020/12/02/blog-how-artificial-intelligence-could-widen-the-gap-between-rich-and-poor-nations>

- **Falta de Capacidades y Conocimientos:** La falta de habilidades y conocimientos en IA en el personal gubernamental puede ser una barrera significativa. La adopción exitosa de la IA requiere una fuerza laboral capacitada que comprenda cómo funcionan estas tecnologías y cómo aplicarlas en el contexto gubernamental.
- **Falta de Liderazgo y Coordinación:** La falta de liderazgo sólido y una visión clara puede dar lugar a la falta de dirección y coordinación en los esfuerzos de implementación. La ausencia de un enfoque estratégico puede llevar a la dispersión de recursos y esfuerzos.

En conjunto, estas causas profundas del rezago en la implementación de la IA en el sector público subrayan la necesidad de abordar no solo los recursos y las barreras técnicas, sino también los aspectos organizativos y culturales que afectan la adopción efectiva de la IA en el gobierno.

### **Sesgos y discriminación algorítmica**

En el contexto de la implementación de IA en el sector público, dos preocupaciones clave son los sesgos y la discriminación algorítmica. Estos términos se han vuelto cada vez más relevantes a medida que los sistemas de IA desempeñan un papel fundamental en la toma de decisiones gubernamentales. A medida que avanzamos en esta sección, exploraremos en profundidad estos conceptos y su impacto en el ámbito público.

En primer lugar, un sesgo algorítmico se refiere a la tendencia de un sistema de IA a favorecer ciertos grupos o resultados sobre otros. Este concepto se ha definido como “una anomalía en la salida de los sistemas de IA, debido a los prejuicios o suposiciones erróneas realizadas durante el proceso de desarrollo del sistema o prejuicios en los datos de entrenamiento, por lo que los resultados del sistema de IA

no pueden generalizarse ampliamente.”<sup>35</sup> EL Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial (HLEG) señala que “Un sesgo es una inclinación que favorece o perjudica a una persona, objeto o posición.”<sup>36</sup>

Estos sesgos pueden surgir de diversas fuentes, como datos de entrenamiento sesgados o suposiciones implícitas en los algoritmos. Un ejemplo común de sesgo es cuando un sistema de IA utilizado en procesos de contratación muestra preferencia por candidatos de cierta edad o género, lo que puede llevar a discriminación en la selección de personal.

Por su parte, la discriminación algorítmica es la consecuencia de los sesgos en los sistemas de IA. Cuando un algoritmo muestra sesgos y prejuicios hacia ciertos grupos, puede resultar en discriminación hacia esas comunidades o individuos. Esto puede manifestarse en diversas decisiones gubernamentales, como las relacionadas con la concesión de préstamos, el acceso a servicios públicos o incluso en el sistema de justicia penal.

Es importante comprender que el sesgo y la discriminación algorítmica no son lo mismo. El sesgo se refiere a la inclinación inherente en los algoritmos o datos utilizados, mientras que la discriminación algorítmica es el resultado de ese sesgo en la práctica. Los sesgos pueden ser sutiles y complejos, lo que hace que su detección y mitigación sean desafiantes.

Las causas de sesgos y discriminación algorítmica pueden variar desde datos sesgados utilizados para entrenar sistemas de IA hasta decisiones de diseño algorítmico. En el sector público, estas preocupaciones son especialmente relevantes, ya que los sistemas de IA pueden influir en la toma de decisiones críticas,

---

<sup>35</sup> JOINT RESEARCH CENTRE; ESTÉVEZ ALMENZAR, M. y otros (2022): *Glossary of human-centric artificial intelligence*, Joint Research Centre, Unión Europea, Luxemburgo, (pp. 16-17).

<sup>36</sup> HLEG (Comisión Europea), (2019): *Directrices éticas para una IA fiable*, Comisión Europea 2019, (p. 48).

como la asignación de recursos, la toma de decisiones judiciales y la prestación de servicios públicos.

Los riesgos asociados con el sesgo y la discriminación algorítmica en el sector público son significativos. Pueden resultar en tratos injustos, violaciones de derechos civiles o incluso en la pérdida de confianza en institucionalidad de un país. Además, estos riesgos pueden exacerbar las desigualdades existentes y socavar los principios de equidad y justicia en la administración pública.

A lo largo de esta sección, se examinará en detalle los sesgos y la discriminación algorítmica en el contexto del sector público, identificando sus causas y se analizarán los riesgos y desafíos que estos fenómenos plantean.

### *Clases de Sesgos Algorítmicos*

Para poder explicar de manera óptima el fenómeno de los sesgos algorítmicos y sus causas, es imprescindible abordarlo en primer lugar desde sus orígenes. Para esto, el Instituto Nacional de Estándares y Tecnología (NIST) de Estados Unidos, desarrolló un importante trabajo en la categorización de los sesgos algorítmicos, lo cual resultó en más de 40 tipos de sesgos, que fueron a su vez agrupados en tres grandes categorías, a decir, sesgos sistémicos, sesgos estadísticos y computacionales y los sesgos humanos, los que se explicarán a continuación:

- **Sesgos Sistémicos:** son principalmente sesgos institucionales o históricos arraigados en los conjuntos de datos, normas, prácticas, y procesos institucionales a lo largo del desarrollo de la IA, así como en la cultura y sociedad en general. Estos sesgos no son intencionales y pueden persistir incluso cuando se siguen las reglas y normativas existentes.
- **Sesgos Estadísticos y Computacionales:** se manifiestan cuando la muestra de datos no representa adecuadamente la población. Estos errores pueden originarse por la inclusión de datos heterogéneos, la simplificación de datos

complejos en representaciones matemáticas más simples, la presencia de datos inexactos y sesgos algorítmicos, como el sobreajuste o subajuste, el tratamiento de valores atípicos, la depuración de datos, y la imputación de factores.

- **Sesgos Humanos:** corresponden a patrones sistemáticos de pensamiento en la mente humana, que incluyen sesgos de anclaje, disponibilidad, confirmación, entre otros. Estos sesgos son una parte inherente del pensamiento humano y usualmente sirven como atajos mentales adaptativos. Aunque no se limitan exclusivamente a las interacciones humanas con la IA, en este contexto, influyen en todas las fases de desarrollo y uso de la IA.

Como se dijo previamente, sesgos algorítmicos y discriminación algorítmica no son conceptos sinónimos, sin embargo, para los efectos de este trabajo, se abordarán de manera conjunta, toda vez que cualquier sesgo podría eventualmente implicar un trato desigual de forma arbitraria.

### *Ejemplos de discriminación algorítmica en el sector público*

Si bien se puede encontrar una muy amplia variedad de ejemplos en que el uso de IA en el sector público resulta en discriminación algorítmica, perjudicando así a múltiples personas a lo largo del mundo, este trabajo se remitirá a mencionar a modo de ejemplo solamente algunos de mayor relevancia, considerando que permitirán ilustrar de mejor manera las implicancias del mal uso de IA en la administración.

El caso más emblemático relacionado con el uso de la inteligencia artificial se centró en la disputa legal entre el Estado de Wisconsin y Eric L. Loomis<sup>37</sup>. Este litigio se basó en la ejecución de un proceso judicial que involucraba un informe presentado por un oficial, el cual contenía hallazgos proporcionados por el software Correctional Offender Management Profiling for Alternative Sanctions (COMPAS). Dicho

---

<sup>37</sup> State Of Wisconsin Vs Eric L. Loomis (2016)

programa informático tenía la capacidad de utilizar datos recopilados y el procesamiento del historial criminal de una persona para evaluar y predecir la probabilidad de reincidencia de un acusado.

No obstante, este caso se centró en cuestionar la utilización de COMPAS, un software propiedad de una empresa y protegido por derechos de propiedad intelectual. La sentencia original imponía una condena de seis años de prisión a Loomis, quien argumentó que el uso de estos algoritmos predictivos violaba el derecho del acusado a un debido proceso, planteando dudas sobre la validez científica y la precisión del software, además de resaltar la existencia de sesgos relacionados con género y raza.

A pesar de sus argumentos, la Corte Suprema de los Estados Unidos, a la cual apeló, optó por no abordar el caso de manera directa, de hecho, validando implícitamente la decisión de la Corte Suprema de Wisconsin sin realizar un cuestionamiento sustancial.

De manera similar pero en Argentina, Guillermo F. Ibarrola fue identificado erróneamente por un sistema inteligente como prófugo y estuvo detenido 6 días. Como consecuencia, en abril de 2022 “un juez acordó mantener el sistema apagado”<sup>38</sup>.

En el área de la salud, un estudio de la revista science evidenció cómo el sistema utilizado por alrededor de 200 millones de ciudadanos estadounidenses anualmente daba prioridad para la atención médica temprana a aquellas personas que habían generado mayores costos previamente. Dado que las personas de origen afroamericano habían recibido menos atención médica y habían incurrido en gastos menores, eran relegadas por el sistema en comparación con los ciudadanos blancos que presentaban enfermedades y necesidades similares. En dicho estudio, los

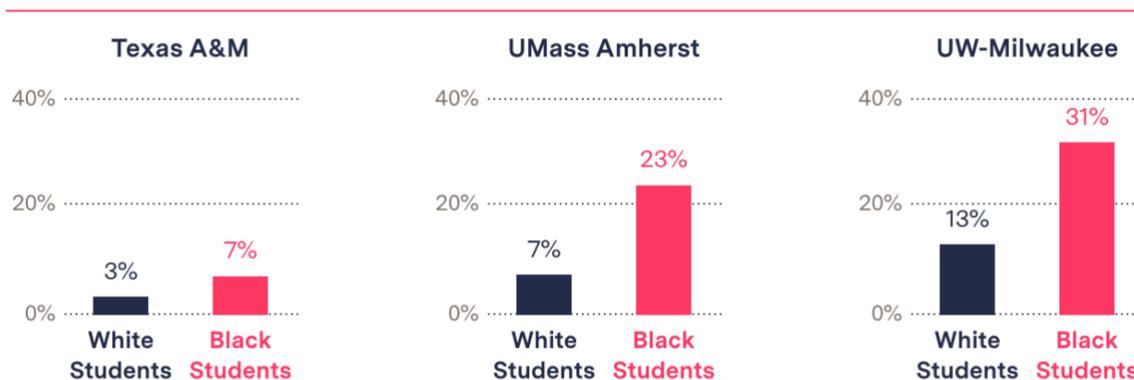
---

<sup>38</sup> Ver: <https://es.wired.com/articulos/escandalo-en-buenos-aires-revela-los-peligros-del-reconocimiento-facial> (fecha consulta: 07/11/2023)

autores estimaron que este sesgo racial reduce en más de la mitad el número de pacientes afroamericanos identificados para recibir atención adicional.<sup>39</sup>

En el ámbito educativo también es habitual el uso de sistemas I, en Estados Unidos, un sistema de predicción de abandono de estudios utilizado en más de 500 universidades generaba sesgos que multiplicaban por cuatro la predicción respecto de las personas negras, comparados con las personas blancas. Así, los estudiantes podían verse afectados sustancialmente dado que los profesores podían ver estos “puntajes”. (Ver Figura 4)

Percentage of student body labeled as “high risk” to not graduate within their selected major



Sources: Texas A&M, University of Massachusetts Amherst, and University of Wisconsin–Milwaukee

*Figura 4. Discriminación Algorítmica en Universidades estadounidenses<sup>40</sup>*

Como último ejemplo, se puede observar la existencia de discriminación algorítmica en plataformas que usan IA para ofrecer empleos. Según Harvard Business Review<sup>41</sup>:

<sup>39</sup> DOI: 10.1126/science.aax2342

<sup>40</sup> Ver: <https://themarkup.org/machine-learning/2021/03/02/major-universities-are-using-race-as-a-high-impact-predictor-of-student-success>

<sup>41</sup> Ali M., Sapiezynski P., Bogen M., Korolova A., Mislove A., Rieke A. Discrimination through Optimization: How Facebook’s Ad Delivery Can Lead to Biased Outcomes. Proc. ACM Hum.-Comput. Interact. 3, CSCW, Article 199 (Noviembre 2019). <https://doi.org/10.1145/3359301>

*“En un estudio reciente que llevamos a cabo junto con colegas de la Universidad Northeastern y la USC, descubrimos, entre otras cosas, que los anuncios de empleo ampliamente dirigidos en Facebook para puestos de cajero de supermercado se mostraron a una audiencia conformada en un 85% por mujeres, mientras que los trabajos con empresas de taxis fueron dirigidos a una audiencia que era aproximadamente un 75% afrodescendiente. Este es un caso paradigmático de un algoritmo reproduciendo sesgos del mundo real sin intervención humana.”<sup>42</sup> (traducido al español)*

En base a los ejemplos expuestos, es posible concluir que la discriminación algorítmica es un fenómeno que puede ser la consecuencia de las más diversas causas y por consecuencia, se puede manifestar de múltiples formas, siendo esencial que sea abordado de la manera más exhaustiva y sería posible.

### *Causas de la discriminación algorítmica*

Los sesgos y errores, según Cotino, “se generan antes de concebirse un sistema de IA, en cualquiera de sus fases o bien en el uso posterior de estos sistemas por los usuarios o de sus resultados de salida”<sup>43</sup>. Añade que “los grandes problemas de la inteligencia artificial por lo general se deben a la estupidez humana. Son los humanos quienes deciden crear el sistema, para qué usos, a partir de qué datos alimentarlo, de dónde obtenerlos, qué valor otorgarles a los datos y a los resultados que arroja el sistema.”<sup>44</sup>

Considerando lo anterior, en conjunto con la categorización de los sesgos expuesta previamente, se puede entender cómo los sesgos algorítmicos y por consecuencia la discriminación algorítmica, puede tener como causa diversas fuentes, siendo la

---

<sup>42</sup> Ver: <https://hbr.org/2019/05/all-the-ways-hiring-algorithms-can-introduce-bias>

<sup>43</sup> Cotino Hueso, L. (2023). Discriminación, sesgos e igualdad de la inteligencia artificial en el sector público. En E. Gamero Casado & F. L. Pérez Guerrero (Eds.), *Inteligencia Artificial y Sector Público: Retos, límites y medios* (Cap. VIII, p. 269). Valencia: Tirant lo Blanch.

<sup>44</sup> *Ídem*, p. 270.

principal el intelecto humano, toda vez que los sistemas de IA son confeccionados por personas naturales.

Ahora bien, se hace imprescindible distinguir entre aquellos sesgos que surgen directamente de la programación de estos sistemas y aquella que simplemente responde a la existencia previa de información sesgada. Así, remitiéndonos a los ejemplos mencionados en la sección anterior, el caso del reconocimiento facial erróneo se presenta como un sesgo algorítmico computacional, toda vez que su causa reside en un error del sistema mismo. Por otro lado, el caso de la discriminación racial en universidades de Estados Unidos responde a un sesgo sistémico, pues atiende a sesgos históricos arraigados en la cultura, y no a un mal funcionamiento del sistema.

### *Riesgos que esto implica*

La presencia de sesgos algorítmicos y discriminación algorítmica en la implementación de IA en el sector público podría tener diversas consecuencias significativas y perjudiciales. Algunas de estas consecuencias incluyen:

1. **Inequidad y discriminación:** Los algoritmos sesgados pueden perpetuar y amplificar desigualdades existentes en la sociedad. Si los datos utilizados para entrenar los modelos contienen sesgos inherentes, los resultados de la IA pueden reflejar y reproducir esas desigualdades, llevando a decisiones discriminatorias. Esto puede observarse en el caso de la discriminación racial en las universidades de EEUU, como se vio anteriormente.

2. **Injusticia en decisiones automatizadas:** Los sistemas de IA, basados en algoritmos sesgados, pueden influir en decisiones cruciales en el ámbito público, como la asignación de recursos, la evaluación de solicitudes de empleo, la administración de justicia, entre otros. Esto podría resultar en decisiones injustas y desproporcionadas para ciertos grupos de la población. El caso de Guillermo Ibarrola en Argentina es un fiel reflejo de una decisión automatizada errónea.

3. Falta de transparencia y rendición de cuentas: La opacidad en los algoritmos puede dificultar la comprensión de cómo se toman ciertas decisiones. La falta de transparencia puede generar desconfianza en los ciudadanos y dificultar la rendición de cuentas, ya que las personas afectadas podrían no entender por qué se tomaron ciertas decisiones.

4. Deterioro de la confianza en instituciones públicas: Si la implementación de IA en el sector público resulta en casos frecuentes de discriminación o sesgos, la confianza en las instituciones gubernamentales y en la tecnología podría disminuir, afectando la aceptación generalizada de la inteligencia artificial.

5. Desafíos legales y éticos: La existencia de sesgos algorítmicos puede plantear desafíos legales y éticos. Podría haber demandas legales contra el uso de sistemas de IA que perpetúan discriminación, y esto podría llevar a la necesidad de regulaciones más estrictas en el desarrollo y despliegue de tecnologías de inteligencia artificial en el sector público.

Para mitigar estas consecuencias, es crucial implementar prácticas éticas y transparentes en el diseño, desarrollo y aplicación de algoritmos de inteligencia artificial, así como llevar a cabo una supervisión constante para identificar y corregir sesgos potenciales. Además, se deben establecer políticas y regulaciones claras que aborden la equidad y la no discriminación en el uso de la inteligencia artificial en el ámbito público.

En esta línea, el BID construyó una propuesta de criterios para una gestión ética de datos por el sector público, en la cual señaló que *“Con el fin de realizar un diagnóstico correcto de los datos requeridos para un determinado proyecto es necesario mapear primero los ya disponibles y los deseables, para posteriormente*

*identificar sesgos posibles. Con esto se obtendrá idealmente el conjunto de datos necesarios”<sup>45</sup>*

## **Protección de Datos**

En la continua exploración de los riesgos y desafíos inherentes a la integración de la IA en el sector público, surge un ámbito crítico de preocupación que requiere una atención especializada y exhaustiva: la protección de datos personales. Tras haber analizado la compleja red de riesgos asociados con la presencia de la IA en el ámbito gubernamental, es imperativo dirigir nuestra atención hacia cómo esta transformación tecnológica impacta directamente en la privacidad y la seguridad de los ciudadanos.

La creciente interconexión entre la IA y la recopilación masiva de datos en el sector público introduce un conjunto de desafíos sin precedentes en términos de protección de datos personales. Este subcapítulo se adentrará en la convergencia de estos elementos, analizando críticamente cómo la implementación de sistemas de IA, diseñados para optimizar procesos gubernamentales, puede comprometer inadvertidamente la privacidad de los individuos.

Como continuación natural de la exposición previa sobre los riesgos generales de la IA en el sector público, este subcapítulo se centrará específicamente en ejemplos paradigmáticos, que ilustrarán de manera tangible los desafíos y peligros asociados con la protección de datos en este contexto. Estos casos no solo proporcionarán ejemplos tangibles de las amenazas a la privacidad, sino que también destacarán la necesidad imperante de desarrollar estrategias efectivas y políticas regulatorias que salvaguarden la información personal de los ciudadanos.

---

<sup>45</sup> Buenadicha, C., Galdon Clavell, G., Hermosilla, M. P., Loewe, D., & Pombo, C. (2019). La gestión ética de los datos: Por qué importa y cómo hacer un uso justo de los datos en un mundo digital. Banco Interamericano de Desarrollo. (p. 43)

A medida que nos adentramos en esta fase de nuestro análisis, la comprensión de estos casos permitirá una evaluación más profunda de las implicaciones éticas y prácticas de la intersección entre la IA y la protección de datos en el ámbito gubernamental. Este enfoque proporcionará una base sólida para las discusiones y recomendaciones subsiguientes, abordando la imperativa tarea de equilibrar la innovación tecnológica con la preservación de los derechos fundamentales en el contexto dinámico y desafiante del sector público impulsado por la IA.

### *Cámaras de vigilancia con reconocimiento facial*

La implementación de sistemas de vigilancia con reconocimiento facial en el sector público ha emergido como un tema de preocupación significativa en la intersección entre la tecnología y la privacidad de los ciudadanos. A medida que la IA se integra cada vez más en las infraestructuras de seguridad gubernamentales, el uso extensivo de esta tecnología plantea riesgos sustanciales para la protección de datos personales y la privacidad individual.

Estos sistemas, impulsados por algoritmos avanzados de reconocimiento facial, tienen la capacidad de analizar y comparar rostros capturados por cámaras de vigilancia con bases de datos masivas, generando identificaciones instantáneas. Aunque se promocionan como herramientas eficaces para combatir el crimen y aumentar la seguridad, su implementación ha suscitado inquietudes éticas y legales sobre el alcance de la vigilancia estatal y el impacto en la privacidad de los ciudadanos.

En China, el gobierno ha implementado a lo largo de la última década un extenso sistema de vigilancia con reconocimiento facial en combinación con el denominado "Sistema de Crédito Social". Este sistema utiliza algoritmos de IA para analizar el comportamiento de los ciudadanos, asignando puntajes sociales basados en actividades como compras, interacciones en línea y comportamiento público. Aquellos con puntajes bajos pueden enfrentar consecuencias como restricciones de viaje y acceso limitado a servicios.

Este sistema se ha utilizado de diversas formas, siendo la más común el uso de cámaras instaladas en lugares públicos. Sin embargo, llama la atención en particular, el uso de cámaras instaladas en anteojos de sol utilizados por policías<sup>46</sup><sup>47</sup>, lo cual acentúa el estado de vigilancia al que está sometida la población china.

De la misma forma, en el Reino Unido, específicamente en Londres, se implementaron sistemas de reconocimiento facial en lugares públicos, incluyendo la estación de King's Cross<sup>48</sup>. Este despliegue generó controversias y críticas por parte de defensores de la privacidad, quienes argumentaron que la recopilación masiva de datos biométricos sin el consentimiento informado de los ciudadanos viola derechos fundamentales<sup>49</sup>. La falta de transparencia en las políticas de retención de datos y la ausencia de un marco regulatorio sólido han llevado a cuestionamientos sobre la proporcionalidad y la necesidad de tal vigilancia en un entorno democrático.

Por su parte, los funcionarios del Servicio de Inmigración y Control de Aduanas de Estados Unidos (ICE) han extraído datos de las bases de datos de licencias de conducir de los estados utilizando tecnología de reconocimiento facial, analizando las fotos de millones de automovilistas sin su conocimiento<sup>50</sup>.

Este suceso también ha generado discusiones respecto a la legitimidad de dichos mecanismos de vigilancia. Harrison Rudoph, Gerente de Políticas de Privacidad en Meta, señaló que "*Esto es un escándalo*", y que "*Los estados nunca han aprobado leyes que autoricen al ICE a sumergirse en las bases de datos de licencias de conducir utilizando reconocimiento facial para buscar personas*". Agrega que "*Estos estados nunca han dicho a los indocumentados que cuando solicitan una*

---

<sup>46</sup> Ver: <https://www.telegraph.co.uk/news/2018/02/07/chinese-police-using-facial-recognition-glasses-identify-suspects/>

<sup>47</sup> Luo, Y., & Guo, R. (2021). Facial Recognition in China: Current Status, Comparative Approach and the Road Ahead. *University of Pennsylvania Journal of Law and Social Change*, 25(2), p. 159.

<sup>48</sup> *Ídem*, p. 171.

<sup>49</sup> The Guardian. Ver: <https://t.ly/q1VXS>

<sup>50</sup> BBC. Ver: <https://www.bbc.com/news/world-us-canada-48907026>

*licencia de conducir, también están entregando su rostro al ICE. Eso es una gran trampa y cambio*"<sup>51</sup>.

La implementación de sistemas de vigilancia con reconocimiento facial plantea una serie de riesgos y desafíos relacionados con la privacidad de los individuos. En primer lugar, está la preocupación sobre la recopilación y almacenamiento masivo de datos biométricos, que pueden ser utilizados para rastrear la ubicación y los movimientos de las personas en tiempo real. Además, la posibilidad de errores en la identificación facial, junto con la falta de regulaciones claras, puede llevar a detenciones incorrectas y violaciones graves de los derechos individuales.

Además, la amenaza potencial de abuso por parte del gobierno o de actores malintencionados que acceden a estas bases de datos es una preocupación constante. La información recopilada puede ser utilizada para realizar un seguimiento de la actividad política, disuadir la participación ciudadana o incluso para persecuciones políticas.

#### *Automatización en Servicios de Salud Pública*

La automatización en servicios de salud pública, aunque conlleva prometedores avances en la eficiencia y calidad de la atención médica, también plantea inquietudes sustanciales en relación con la privacidad de los individuos y la protección de datos personales. Esta transformación digital en el ámbito de la salud se materializa a través de la implementación de sistemas y procesos automatizados, que incluyen la recopilación masiva de datos, análisis avanzado de información médica y la toma de decisiones asistida por algoritmos. En este contexto, es imperativo examinar cómo la automatización puede impactar negativamente la privacidad y la integridad de los datos personales.

---

<sup>51</sup> New York Times. Ver: <https://www.nytimes.com/2019/07/07/us/politics/ice-drivers-licenses-facial-recognition.html>

Uno de los principales desafíos radica en la recopilación masiva de datos de salud que estos sistemas automatizados facilitan. La abundancia de información médica sensible, que abarca desde historiales clínicos hasta resultados de pruebas diagnósticas, se convierte en un activo valioso y, simultáneamente, en un blanco atractivo para posibles violaciones de seguridad.

Los incidentes de seguridad cibernética en el sector de la salud han demostrado que la automatización puede exponer datos personales a riesgos significativos, afectando la privacidad y la confidencialidad de la información médica. De esta forma, el 2015 se produjo un ciberataque dirigido hacia la empresa aseguradora Anthem, en la cual se filtró información de alrededor de 80 millones de personas<sup>52</sup>, siendo una de las mayores violaciones de datos ocurridas en EEUU.

Además, la automatización en la toma de decisiones médicas plantea interrogantes acerca de la equidad y la transparencia. Algoritmos que asisten en la asignación de recursos o en la toma de decisiones clínicas pueden introducir sesgos inadvertidos, lo que podría afectar de manera desproporcionada a ciertos grupos de población, tal como se analizó en la sección anterior. Estos sesgos, basados en datos históricos y patrones, pueden traducirse en discriminación y amplificar disparidades existentes en el acceso a la atención médica.

Por su parte, la telemedicina, impulsada por la automatización, introduce una dimensión adicional de preocupación en relación con la privacidad. La monitorización remota de la salud mediante dispositivos conectados y consultas virtuales puede exponer datos médicos a riesgos de seguridad, especialmente cuando la infraestructura tecnológica subyacente no se asegura adecuadamente, como ocurrió con la empresa Cerebral, donde más de 3 millones de personas se vieron afectadas por una filtración de datos<sup>53</sup>. Incidentes de brechas de seguridad en

---

<sup>52</sup> Forbes. Ver: <https://www.forbes.com/sites/gregorymcneal/2015/02/04/massive-data-breach-at-health-insurer-anthem-reveals-social-security-numbers-and-more/?sh=54219d4d2601>

<sup>53</sup> Healthcare IT News. Ver: <https://www.healthcareitnews.com/news/cerebral-says-3m-affected-patient-data-breach>

plataformas de telemedicina han subrayado la vulnerabilidad inherente de estos sistemas y los desafíos asociados con la protección de datos sensibles.

A su vez, la integración de datos de múltiples fuentes, aunque puede mejorar la continuidad de la atención médica, también presenta amenazas para la privacidad. Asociaciones entre proveedores de atención médica y entidades tecnológicas, como el controvertido caso de Google y Ascension<sup>54</sup>, suscitan inquietudes sobre la transparencia y el control sobre la información médica de los individuos cuando se combinan datos de diversas fuentes sin un consentimiento claro y explícito.

En síntesis, la implementación de IA en el sector público, evidenciada a través de ejemplos como el despliegue de cámaras con reconocimiento facial en China, Inglaterra y Estados Unidos, así como la automatización en servicios de salud, plantea un desafío palpable para la preservación de la privacidad y la protección de datos personales. La utilización generalizada de tecnologías como el reconocimiento facial ha demostrado ser intrusiva y susceptible a abusos, como lo evidencian casos en China donde la vigilancia estatal masiva se combina con sistemas de puntaje social. En paralelo, en países como Inglaterra, la implementación de cámaras con reconocimiento facial en lugares públicos ha generado debates sobre la proporcionalidad y transparencia de dicha vigilancia, afectando directamente la privacidad de los ciudadanos.

En el ámbito de la salud, la automatización presenta retos adicionales al exponer datos médicos sensibles a vulnerabilidades de seguridad, como se observa en incidentes de brechas de datos en plataformas de telemedicina. La convergencia de información biométrica y registros médicos bajo el paraguas de la IA resalta la necesidad crítica de mecanismos sólidos de protección de datos.

En este contexto, resulta imperativo considerar enfoques éticos y regulaciones robustas que equilibren la innovación tecnológica con la salvaguarda de los derechos

---

<sup>54</sup> Harvard Bill of Health. <https://blog.petrieflom.law.harvard.edu/2019/11/14/what-google-isnt-saying-about-your-health-records/>

individuales. La formulación de políticas que aborden la transparencia en la recopilación y uso de datos, así como la rendición de cuentas por parte de las entidades gubernamentales y corporativas, se convierte en una necesidad inminente. Además, el fomento de la conciencia pública sobre los riesgos asociados con la implementación de IA en el sector público es esencial para promover un diálogo informado y participativo sobre los límites y ética de estas tecnologías.

Para lo anterior, el BID en su propuesta de criterios para la gestión ética de los datos, señala adecuadamente lo siguiente:

*“La privacidad debe incorporarse en cada etapa del ciclo de vida de los datos, y es en la fase de planificación cuando se deben diseñar las acciones relevantes para lograrlo, a la luz de los estándares que hacen parte del reglamento general de protección de datos de la Unión Europea. Si el proyecto contempla datos personales o sensibles, se deben incorporar las medidas para cumplir con la normativa nacional de protección de datos personales.”<sup>55</sup>*

Así, contempla entre las medidas i) el Consentimiento; ii) la Anonimización o Seudonimización; iii) la seguridad de la información; iv) la existencia de estándares por defecto y v) Prevención de la existencia de datos abiertos.

En última instancia, la preservación de la privacidad en la era de la IA requiere una reflexión continua sobre la ponderación de beneficios y riesgos, instando a la adopción de medidas preventivas y correctivas para salvaguardar los valores fundamentales en el uso de estas tecnologías en el ámbito gubernamental.

## **Opacidad y “Caja Negra”**

Como ya se ha evidenciado a lo largo de este trabajo, la implementación de IA en el sector público ha emergido como una herramienta sumamente útil, sin embargo, no

---

<sup>55</sup> Buenadicha, C., Galdon Clavell, G., Hermosilla, M. P., Loewe, D., & Pombo, C. (2019). La gestión ética de los datos: Por qué importa y cómo hacer un uso justo de los datos en un mundo digital. Banco Interamericano de Desarrollo. (p. 43)

está exenta de riesgos, como los analizados en este capítulo, y entre los desafíos preeminentes que enfrenta se analizará finalmente la opacidad algorítmica, un fenómeno que arroja luces sobre la complejidad inherente de los modelos de IA y sus implicancias en la transparencia de las decisiones tomadas por estos sistemas.

La opacidad algorítmica, a menudo referida como el "fenómeno de la caja negra" en la literatura académica, se presenta cuando los procesos y criterios que guían las decisiones de los algoritmos de IA son oscuros y difíciles de comprender. En este contexto, la opacidad no se limita a la falta de transparencia literal, sino que encapsula la incapacidad de los actores involucrados, ya sean ciudadanos, funcionarios públicos o expertos en la materia, para desentrañar el funcionamiento interno de los modelos de IA. Este aspecto opaco de la inteligencia artificial plantea desafíos significativos para la rendición de cuentas, la equidad y la confianza en las instituciones gubernamentales que emplean estos sistemas.

La preocupación fundamental radica en la incapacidad de comprender cómo las decisiones son tomadas por los algoritmos, lo que podría llevar a consecuencias imprevistas y potencialmente perjudiciales. A medida que la IA se integra más profundamente en la administración pública, la opacidad algorítmica podría afectar áreas críticas, desde la asignación de recursos hasta la toma de decisiones judiciales y la formulación de políticas públicas. La falta de comprensión sobre las lógicas subyacentes de los algoritmos introduce un nivel de incertidumbre que socava la confianza ciudadana en los procesos gubernamentales basados en IA.

Este último subcapítulo se sumerge en la dimensión crítica de la opacidad algorítmica como un riesgo inherente a la adopción de la inteligencia artificial en el sector público. Se explorarán los diversos aspectos de esta opacidad, desde la complejidad técnica que obstaculiza la comprensión hasta las implicancias éticas y sociales que surgen cuando las decisiones gubernamentales se vuelven ininteligibles para quienes están sujetos a ellas.

## *Concepto de Opacidad*

Para introducir adecuadamente el concepto de opacidad, este trabajo se remitirá a lo señalado por la Comisión Europea, a decir:

*“La opacidad de los sistemas de IA, conocida como el paradigma de la “caja negra”, se define como la imposibilidad o dificultad de comprensión del sistema. Incluso cuando los desarrolladores o los propios creadores del sistema de IA examinan por ejemplo un modelo de aprendizaje automático (ej. código fuente, proceso de entrenamiento, arquitectura, parámetros, datos) es difícil o imposible entender cómo los diferentes componentes se combinan para generar una decisión, salida o predicción. Dicha dificultad o imposibilidad se debe a la incapacidad de generar razonamiento a escala humana a partir de sistemas que, en la actualidad, tienen un alto grado de complejidad (los modelos de aprendizaje automático basados en redes neuronales, por ejemplo, se componen de decenas de millones de parámetros). Estas características intrínsecas de los actuales sistemas de IA han motivado el desarrollo de requisitos de transparencia para asegurar la confianza y la responsabilidad en los sistemas de IA.”<sup>56</sup>*

A esto, es necesario agregar que la opacidad tiene diferentes “niveles”, los cuales se describen de manera precisa en el Manual de formulación de proyectos del BID, que señala la existencia de tres niveles:

***“Opacidad intencional:*** *Habrán situaciones donde el algoritmo no podrá ser explicado a la ciudadanía porque pondría en riesgo la efectividad del modelo mismo. Por ejemplo, si se aplica un modelo para predecir la evasión de impuestos, la ciudadanía no debe saber cómo funciona específicamente, ya que algunos podrían intentar “engañar” al modelo. Esta opacidad intencional también puede presentarse cuando hay motivos de propiedad intelectual.*

---

<sup>56</sup> Gómez E., Hupont Torres I., Sánchez I., Fernández Llorca D. (2023) Hacia un marco regulatorio holístico de la Inteligencia Artificial fiable en la Unión Europea: una perspectiva científico-técnica. En E. Gamero Casado & F. L. Pérez Guerrero (Eds.), Inteligencia Artificial y Sector Público: Retos, límites y medios (Cap. XVII, p. 730). Valencia: Tirant lo Blanch.

***Opacidad analfabeta:*** *Ocurre cuando las personas a las que se explica el algoritmo no comprenden cómo funciona por una falta de conocimiento en los temas técnicos. En este caso se recomienda capacitar a los funcionarios de primera línea de la institución para que puedan explicar, en un lenguaje claro, el funcionamiento del sistema.*

***Opacidad intrínseca:*** *En ocasiones los algoritmos utilizados son sumamente complejos, por lo cual es difícil evidenciar una relación causal entre valores entrantes y salientes. Si existe este tipo de opacidad, resulta imposible explicar el modelo.”<sup>57</sup>*

### *Implicancias de la opacidad*

La opacidad en la implementación de sistemas de IA en el sector público se alza como un riesgo crítico, integrándose profundamente en la serie de desafíos éticos y prácticos que caracterizan este entorno tecnológico en constante evolución. Al explorar los diversos riesgos asociados con la adopción de la IA en la esfera gubernamental, la opacidad emerge como una preocupación fundamental, eclipsando la comprensión clara y la explicación efectiva de las decisiones automatizadas.

En el contexto de la opacidad, la complejidad inherente a los modelos de IA, especialmente aquellos basados en técnicas avanzadas como el aprendizaje profundo, obstaculiza la capacidad de los actores involucrados para discernir la lógica subyacente de las decisiones generadas. Este fenómeno adquiere relevancia significativa al considerar la ética y la equidad en la toma de decisiones gubernamentales, donde la falta de transparencia puede perpetuar sesgos inadvertidos y, en última instancia, afectar la confianza pública.

---

<sup>57</sup> Denis, G., Hermosilla, M. P., Aracena, C., Sánchez Ávalos, R., González Alarcón, N., & Pombo, C. (Septiembre de 2021). *Uso responsable de IA para política pública: Manual de formulación de proyectos*. Banco Interamericano de Desarrollo. Recuperado de <https://www.iadb.org/>

La opacidad no solo impide la rendición de cuentas efectiva y la capacidad de los individuos para auditar decisiones automatizadas, sino que también plantea cuestionamientos fundamentales sobre la conformidad con normativas de protección de datos. En un entorno donde la privacidad y la gestión ética de la información son imperativos, la opacidad de los modelos de IA desafía la trazabilidad y la explicación adecuada de cómo se utilizan los datos personales en la toma de decisiones gubernamentales automatizadas.

En el escenario de la administración pública, donde la confianza y la legitimidad son pilares esenciales, la opacidad amenaza con socavar estos fundamentos, en este sentido, el BID señala que *“La opacidad puede determinar la diferencia entre un error subsanable y una crisis de confianza entre el gobierno y la ciudadanía”*<sup>58</sup>.

En la misma línea, la falta de comprensión sobre cómo operan los algoritmos y cómo llegan a sus conclusiones puede generar escepticismo y resistencia a la adopción de tecnologías de IA, incluso cuando podrían aportar beneficios sustanciales. En este sentido, Hoffman-Riem señala que:

*“No solo para los usuarios, sino también para las autoridades supervisoras y el público en general, es importante que el manejo de las tecnologías digitales, incluido el uso de la inteligencia artificial, sea comprensible y controlable en general. En este sentido, la transparencia suficiente es un requisito previo para generar no solo confianza, sino también responsabilidad y, en algunos casos, responsabilidad legal.”*<sup>59</sup>

Ahora bien, las consecuencias de una implementación inadecuada de sistemas de IA en el sector público en términos de opacidad, no solamente podrían manifestarse en términos de falta de transparencia o responsabilidad. Así, Araya Paz señala que una

---

<sup>58</sup> Buenadicha, C., Galdon Clavell, G., Hermosilla, M. P., Loewe, D., & Pombo, C. (2019). La gestión ética de los datos: Por qué importa y cómo hacer un uso justo de los datos en un mundo digital. Banco Interamericano de Desarrollo. (p. 19)

<sup>59</sup> Hoffmann-Riem, W. (2020). Artificial Intelligence as a Challenge for Law and Regulation. In *Regulating Artificial Intelligence* (p. 17). [https://doi.org/10.1007/978-3-030-32361-5\\_1](https://doi.org/10.1007/978-3-030-32361-5_1)

excesiva regulación que pretenda prohibir sistemas de “cajas negras” también podría tener consecuencias negativas. De esta forma, indica que *“El resultado lógico y natural es que desincentivará la innovación en tecnología en el país (que ya es precaria) y nos mantendrá como dependientes de tecnologías desarrolladas por terceros; las empresas no se querrán arriesgar a invertir en un país que pone tantas trabas; los pequeños empresarios y desarrolladores no podrán competir con las grandes empresas dadas las altas barreras de entrada; la inteligencia artificial desarrollada será predecible y en nada se distinguirá de los programas computacionales ya existentes. En el fondo, esta vía nos conduce a un oscurantismo tecnológico brutal”*.<sup>60</sup>

Para abordar este riesgo omnipresente, se requiere una exploración más profunda y detallada de enfoques que promuevan la explicabilidad y transparencia de los modelos de IA en el sector público. Este subcapítulo sirve como un punto de partida para una indagación más exhaustiva sobre estrategias específicas y marcos regulatorios que mitiguen la opacidad, al tiempo que promueven la adopción ética y responsable de la inteligencia artificial en el ámbito gubernamental. Así, el siguiente capítulo abordará el concepto de la Transparencia Algorítmica como principio fundamental en la implementación de sistemas automatizados en el sector público.

---

<sup>60</sup> Araya Paz, C. (Diciembre de 2020). Desafíos legales de la inteligencia artificial en Chile. Revista Chilena de Derecho y Tecnología, (p.287-288).

# Transparencia Algorítmica

## Introducción

En el análisis exhaustivo de los riesgos inherentes a la implementación de la inteligencia artificial en el sector público, el capítulo anterior arrojó luz sobre la preocupación central que implica la opacidad algorítmica. Este fenómeno, caracterizado por la falta de claridad en los procesos decisionales de los algoritmos, plantea desafíos sustanciales para la legitimidad, la rendición de cuentas y la confianza en las aplicaciones de IA gubernamentales. La opacidad, como se identificó, puede socavar los principios fundamentales de la toma de decisiones justa y equitativa, dando lugar a una serie de riesgos éticos y sociales.

En respuesta a esta problemática, el presente capítulo se sumerge en el crucial concepto de la transparencia algorítmica. La transparencia se revela como la clave maestra para desbloquear el potencial positivo de la IA en el sector público, contrarrestando los riesgos previamente explorados. En lugar de perpetuar la opacidad inherente, la transparencia algorítmica emerge como un antídoto esencial, brindando a las partes interesadas, ya sean ciudadanos, supervisores o expertos, la capacidad de entender y evaluar las decisiones automatizadas.

La opacidad algorítmica, como se explicó en el capítulo anterior, puede conducir a la falta de confianza, percepciones sesgadas y resultados indeseados. La transparencia algorítmica, por otro lado, se postula como la estrategia fundamental para mitigar estos riesgos. La apertura y accesibilidad de los procesos algorítmicos permiten una comprensión más profunda de cómo se toman las decisiones, cómo se entrenan los modelos y cuáles son los criterios subyacentes.

Desde una perspectiva ética, la transparencia algorítmica se alinea con la necesidad de garantizar que las decisiones automatizadas sean justas y respetuosas de los

derechos fundamentales. Al revelar las lógicas internas de los algoritmos, se facilita la identificación y mitigación de sesgos y discriminación inadvertidos. Este enfoque ético, respaldado por la transparencia, es esencial para cultivar la confianza ciudadana y garantizar que la implementación de la IA en el sector público sea percibida como legítima y ética.

Para comprender adecuadamente la relevancia de la transparencia algorítmica, este capítulo analizará en profundidad las recomendaciones de diversos organismos sobre la materia, dentro de los que se apreciará lo propuesto por el Consejo para la Transparencia, el BID y la opinión expertos en la materia. Esto proporcionará un marco de referencia para evaluar la relevancia y aplicabilidad de la transparencia algorítmica en el contexto chileno.

Además, se llevará a cabo un análisis comparativo con regulaciones extranjeras, destacando especialmente las iniciativas de la Unión Europea y Uruguay en el ámbito de la transparencia algorítmica. Esta comparación permitirá identificar buenas prácticas, lecciones aprendidas y posibles enfoques que podrían ser adaptados al contexto chileno.

Finalmente, este capítulo busca trascender la revisión teórica y regulatoria, propone aportar observaciones y recomendaciones específicas, contextualizadas en la realidad nacional, para así fortalecer la implementación de la transparencia algorítmica en el sector público chileno. El análisis crítico de las políticas existentes, junto con la identificación de brechas y oportunidades, allanará el camino para conclusiones fundamentadas y orientaciones prácticas que contribuyan a forjar un camino ético y transparente en la incorporación de la inteligencia artificial en el ámbito gubernamental.

### Concepto de Transparencia Algorítmica y Explicabilidad

Para entender adecuadamente el concepto de transparencia algorítmica, este trabajo considerará fundamental partir de la base previa de dos conceptos. En primer lugar,

la opacidad, la cual se analizó en el capítulo precedente y que debe considerarse como uno de los riesgos inherentes a la implementación de sistemas de decisiones automatizadas en el sector público con mayor trascendencia. En segundo lugar, el concepto de transparencia en un sentido más amplio, es decir, como uno de los principios rectores de la Administración Estatal.

En Chile, la Ley N° 20.285, sobre Acceso a la Información Pública, prescribe en su artículo tercero que *“La función pública se ejerce con transparencia, de modo que permita y promueva el conocimiento de los procedimientos, contenidos y decisiones que se adopten en ejercicio de ella”*<sup>61</sup>. Asimismo, el inciso segundo del cuarto artículo del mismo cuerpo, indica que *“El principio de transparencia de la función pública consiste en respetar y cautelar la publicidad de los actos, resoluciones, procedimientos y documentos de la Administración, así como la de sus fundamentos, y en facilitar el acceso de cualquier persona a esa información, a través de los medios y procedimientos que al efecto establezca la ley”*.<sup>62</sup>

Teniendo en consideración lo anterior, un acercamiento acertado al concepto de la Transparencia Algorítmica proviene del trabajo de los profesores Llamas, Mendoza y Graff, que señalan que:

*“La transparencia es una propiedad de los sistemas legales y administrativos. Es el derecho de la sociedad para acceder a información sobre trámites, decisiones o atribuciones, cuyos beneficios radican en ayudar a las organizaciones a contabilizar las tareas realizadas, es decir, un seguimiento de los procesos llevados a cabo hacia un resultado específico para materializar una responsabilidad, así como ayudar a las personas a comprender y cuestionar procesos y resultados”*<sup>63</sup>

---

<sup>61</sup> Biblioteca del Congreso Nacional de Chile

<sup>62</sup> *Ídem*.

<sup>63</sup> Llamas Covarrubias, J. Z., Mendoza Enríquez, O. A., & Graff Guerrero, M. (2022). Enfoques regulatorios para la inteligencia artificial (IA) [Regulatory approaches to artificial intelligence (AI)]. DOI: 10.7764/R.493.2. (p. 45).

Por su parte, el Consejo para la Transparencia en sus recomendaciones sobre Transparencia Algorítmica, profundiza en la descripción del concepto al señalar que *“la transparencia se ha entendido al menos desde dos perspectivas, por una parte, como la mayor o menor disposición que tienen las instituciones a la apertura de sus procesos decisionales, y por otra, las demandas de información que surgen desde la sociedad para conocer el funcionamiento, fundamentos y procedimientos por los que las autoridades adoptan sus decisiones”*<sup>64</sup>. Agrega además que *“la transparencia se entiende como disponibilidad y facilitación del acceso al público de información detentada por alguna organización, así como también, la facultad ciudadana de observar y de informarse acerca de procesos de toma de decisión por parte de alguna institución”*<sup>65</sup>.

En cuanto a la manera en que se manifiesta la transparencia cuando se trata del uso de IA, se identifican siete áreas:

1. En los datos utilizados por el sistema, ya sean fuentes, procesos y métodos de verificación para valorar imparcialidad y representación;
2. Algoritmos, a fin de verificar los resultados de salida con los de entrada e inspeccionar errores internos y externos;
3. Objetivos, con el propósito de conocer las prioridades relativas de cada sistema;
4. Resultados, para analizar los efectos del despliegue de los sistemas algorítmicos, incluidos los estados internos del sistema;
5. Cumplimiento, para determinar el desempeño general de los requisitos de transparencia;
6. Influencia, para averiguar si algún elemento del proceso algorítmico se inclinó a favorecer un resultado; y

---

<sup>64</sup> Consejo para la Transparencia. (Octubre 2020). Cuaderno de trabajo N°17: Transparencia algorítmica: Buenas prácticas y estándares de transparencia en el proceso de toma de decisiones automatizadas. Dirección de Estudios. Santiago, Chile. (p. 4).

<sup>65</sup> *Ídem*.

7. Uso, comparando qué datos personales está usando el sistema, ya sea para personalizar los resultados o datos que pueden entrenar al sistema y no se violen los principios de protección de datos y privacidad.<sup>66</sup>

En base a lo anterior, es adecuado concluir que la transparencia algorítmica no se agota en el único hecho de poder acceder a los algoritmos empleados en las decisiones automatizadas, sino que debe aplicarse durante toda la cadena de utilización de estos, es decir, desde su construcción hasta la constatación de los resultados obtenidos.

### **Transparencia Procedimental, Transparencia Técnica y Explicabilidad**

Un tema que debe considerarse a la hora de abordar la Transparencia de los algoritmos es el límite existente entre la obligación de brindar información acerca de los sistemas de decisiones automatizados y los eventuales derechos que puedan verse afectados, principalmente por parte de los creadores de estos sistemas. Así, un posible conflicto podría surgir ante la eventual solicitud de un código fuente de un sistema algorítmico, que probablemente se encuentre protegido por los derechos de autor.

Para comprender este dilema, se debe comenzar por recalcar que la Transparencia en materia algorítmica no es un concepto unívoco, puesto que tiene aplicación a múltiples áreas y es susceptible de ser exigida a diversos actores. Así, la Profesora de Derecho administrativo de la Universidad de Castilla-La Mancha, Isabel Gallego Córcoles, aborda esta disyuntiva y recoge lo señalado por su colega Martín Delgado, indicando que:

*“[...] no necesariamente ha de equipararse transparencia y motivación con apertura total del sistema algorítmico. Transparencia y motivación se vinculan a la legitimidad del sistema, la rendición de cuentas y el*

---

<sup>66</sup> Koene y otros (2019). (p. 5-6)

*control de la decisión adoptada; y para cumplir tales fines no siempre resultará preciso dar a conocer el interior del sistema, el código fuente. Lo relevante en términos jurídicos -continúa el autor- es difundir la existencia del sistema, los datos de los que se nutre, la información relativa a su operatividad técnica, las razones por las que se ha adoptado la decisión. De esta forma, las exigencias de transparencia y motivación no estarían vinculadas tanto a dar a conocer sistemáticamente la forma en la que opera el sistema desde el punto de vista tecnológico, sino a la interacción humana con el output del algoritmo y los criterios de diseño de los inputs para mantener la exigencia de inteligibilidad de la decisión.”<sup>67</sup>*

De esta manera, surge la imprescindible tarea de encontrar un equilibrio en el cual se puedan transparentar los sistemas de decisiones basados en algoritmos sin transgredir los derechos de autor de sus creadores o de quienes los detentan. En esta línea, la autora aborda de manera precisa las cláusulas contractuales de la Ciudad de Ámsterdam<sup>68</sup>, las cuales establecen una profunda regulación en lo que respecta a los conceptos de Transparencia Algorítmica y Explicabilidad, en la que distinguen acertadamente entre la Transparencia Procedimental, la Transparencia Técnica y la Explicabilidad, lo cual se puede resumir de la siguiente manera:

- **Transparencia Procedimental**: Se refiere a la divulgación de información relacionada con el propósito de un sistema algorítmico y el proceso seguido en su desarrollo, incluyendo detalles sobre los datos utilizados, opciones e hipótesis, categorías de datos, la intervención humana prevista, identificación y mitigación de riesgos, así como las partes involucradas en el desarrollo y sus respectivas funciones.
- **Transparencia Técnica**: Implica la revelación de la estructura técnica de un sistema algorítmico, incluyendo la divulgación del código fuente, las especificaciones

---

<sup>67</sup> Gallego Córcoles, I. (2023). La contratación de soluciones de inteligencia artificial. En E. Gamero Casado & F. L. Pérez Guerrero (Eds.), *Inteligencia Artificial y Sector Público: Retos, límites y medios* (Cap. XI, p. 530). Valencia: Tirant lo Blanch.

<sup>68</sup> Standard Clauses for Procurement of Trustworthy Algorithmic Systems. 2021

técnicas y los datos utilizados en su desarrollo. También abarca información técnica sobre la obtención y edición de los datos, el método de desarrollo empleado, el proceso de desarrollo, la justificación de la elección del modelo y sus parámetros, así como detalles sobre el rendimiento del sistema algorítmico.

Por otra parte, surge paralelamente la explicabilidad como concepto fundamental en la regulación de sistemas de decisiones automatizadas. Para entender adecuadamente qué significa, este trabajo recoge el contenido de la recomendación de la UNESCO sobre ética e IA, en la cual se señala que *“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.”*<sup>69</sup> A esto, se añade que *“hacer que un sistema algorítmico sea explicable incluye el suministro de toda la información técnica y de otro tipo necesaria para explicar, en los procedimientos de recurso o en otros procedimientos legales, cómo se ha llegado a una decisión y para ofrecer a la otra parte y a cualquier otra parte interesada la oportunidad de evaluar la forma en que se ha llegado a una decisión, a fin de otorgar a terceros una protección legal realista”*<sup>70</sup>.

La relevancia de esta distinción radica en la delimitación precisa de los deberes de transparencia a los que se obligan los responsables de los sistemas de decisiones automatizadas, toda vez que la exigencia de “muchas” información, podría eventualmente vulnerar la confidencialidad del código fuente<sup>71</sup>, amparada por derechos de autor.

---

<sup>69</sup> UNESCO. (2022). Recomendación sobre la ética de la inteligencia artificial. Publicado por la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, 7, place de Fontenoy, 75352 París 07 SP, Francia. (p. 23)

<sup>70</sup> Gallego Córcoles, I. (2023). La contratación de soluciones de inteligencia artificial. En E. Gamero Casado & F. L. Pérez Guerrero (Eds.), *Inteligencia Artificial y Sector Público: Retos, límites y medios* (Cap. XI, p. 532). Valencia: Tirant lo Blanch.

<sup>71</sup> En este sentido, es interesante la posición que puede adoptar cada legislación. Contraria a la de Ámsterdam, se puede observar cómo Cataluña ha declarado que incluso el código fuente es información pública, no siendo suficiente la explicabilidad de un sistema. Para un análisis más detallado, se puede estudiar el contenido de la *Comissió De Garantia Del Dret D'accés A La Informació Pública* de Cataluña (2016).

En este sentido, las cláusulas de la Ciudad de Amsterdam imponen en su artículo 5 distintas obligaciones a los contratistas de estos sistemas, conforme al concepto (de los 3 explicados previamente) en el cual se incluya la información en cuestión. En primer lugar, la Transparencia Procedimental *“permite divulgar información relevante sobre el sistema algorítmico sin perjudicar desproporcionadamente los intereses del contratista”*<sup>72</sup>, lo cual tiene como principal función *“rendir cuentas a la ciudadanía en relación con el uso del sistema algorítmico”*<sup>73</sup>.

En segundo lugar, la Transparencia Técnica tiene por objeto corroborar con mayor detalle si el contratista ha cumplido con sus obligaciones, lo que se realiza por medio de una auditoría. Dicho control puede tener por objeto obtener información confidencial y por ende, protegida por los derechos de autor, razón por la cual *“las cláusulas obligan a la Administración a preservar el carácter confidencial de esta información y a destruirla una vez finalizada la auditoría. [...] Además, con la finalidad de reforzar los intereses del contratista, se prevé que este pueda optar por no proporcionar el código fuente del sistema algorítmico, sino entregarlo a un tercero independiente que realizará la auditoría en nombre de la Administración y, en ese contexto, analizará el código fuente del sistema algorítmico”*<sup>74</sup>.

Finalmente, las cláusulas establecen en lo que respecta al concepto de Explicabilidad, el deber Estatal de poder explicar estos sistemas y la obligación del contratista de colaborar en dicha tarea.

Como se mencionó previamente, realizar esta distinción puede permitir velar tanto por una Transparencia Algorítmica adecuada como por los demás derechos involucrados, como los derechos de propiedad intelectual. Esto refuerza la idea de que la Transparencia no es un principio absoluto y que su regulación debe considerar a todos los actores involucrados y no solamente a la ciudadanía.

---

<sup>72</sup> *Ídem*, p. 534.

<sup>73</sup> *Ídem*.

<sup>74</sup> *Ídem*.

De esta forma, no queda duda de que Explicabilidad y Transparencia son conceptos súmamente entrelazados, y que los resultados obtenidos a raíz de sistemas de decisiones automatizadas debiesen ser comprensibles y trazables. Sobre esto, la UNESCO agrega que *“La transparencia y la explicabilidad están estrechamente relacionadas con las medidas adecuadas de responsabilidad y rendición de cuentas, así como con la fiabilidad de los sistemas de IA”*<sup>75</sup>

## Importancia

Teniendo en cuenta lo anterior, y en particular la multiplicidad de manifestaciones de la transparencia en el uso de IA, se puede observar el trascendental rol que debe cumplir. Sin embargo, es posible cuestionarse si la transparencia es un medio o un fin en si misma. Ante esta interrogante, los profesores Llamas, Mendoza y Graff ofrecen una interesante respuesta:

*“una respuesta breve sería considerarla como un objeto o medio, pues al tener más transparencia se resuelven problemas sobre el funcionamiento de las tecnologías e información, y pueden ser utilizadas como herramientas o pruebas. No obstante, también debe extenderse hacia un fin, que contemple una especie de condición previa para la regulación frente a espacios tecnológicos tan dinámicos, como una forma alternativa de regulación tradicional, que en la práctica sería una regulación más ligera, ya que cumplir con sus requisitos es menos complejo y coactivo que otras formas de regulación, pero para ser factible las entidades deberán procesar y publicar la información de los sistemas de IA”*.<sup>76</sup>

---

<sup>75</sup> UNESCO. (2022). Recomendación sobre la ética de la inteligencia artificial. Publicado por la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, 7, place de Fontenoy, 75352 París 07 SP, Francia. (p. 23)

<sup>76</sup> Llamas Covarrubias, J. Z., Mendoza Enríquez, O. A., & Graff Guerrero, M. (2022). Enfoques regulatorios para la inteligencia artificial (IA) [Regulatory approaches to artificial intelligence (AI)]. DOI: 10.7764/R.493.2. (p. 47).

A raíz de esto, este trabajo reconoce con particular énfasis la importancia de reconocer la transparencia como una condición *sine qua non* para la implementación de cualquier tipo de sistema de decisiones automatizadas.

En este sentido, una experta entrevistada en la revista del CPLT, señaló que *“La transparencia es una condición pro ética que habilita o impide otras prácticas éticas, la transparencia es habilitador para las otras cosas, por eso es importante la transparencia, debemos saber cómo funcionan las cosas, si no, hay discriminación al usar nuestros datos”*<sup>77</sup>.

Junto a lo anterior, el CPLT señala también que *“El valor de la transparencia en el uso de algoritmos, radica también en que ésta no sólo permite la aplicabilidad del mismo, sino que igualmente posibilita comprender sus resultados y saber sobre su desempeño para evaluar si es justo o no”*<sup>78</sup>. Finalmente, concluye que:

*“El papel que juega la transparencia algorítmica es fundamental, pues abre la posibilidad de auditar y supervisar los datos que entran, cómo se procesan, y cómo se leen en un sistema de toma de decisiones algorítmicas mediante Inteligencia Artificial. De esta manera, es posible minimizar el riesgo de implementar procesos algorítmicos discriminatorios, condicionar arbitrariamente las decisiones de las personas y establecer responsabilidades cuando estos procesos entregan resultados no concluyentes, utilizan datos incorrectos o arrojan resultados injustos.”*<sup>79</sup>

Considerando todo lo mencionado, resulta evidente que la relevancia de la transparencia algorítmica es de la mayor magnitud, y a su vez, que esta se manifiesta en cada una de las etapas de la implementación de sistemas de decisiones

---

<sup>77</sup> Consejo para la Transparencia. (Octubre 2020). Cuaderno de trabajo N°17: Transparencia algorítmica: Buenas prácticas y estándares de transparencia en el proceso de toma de decisiones automatizadas. Dirección de Estudios. Santiago, Chile. (p. 10)

<sup>78</sup> *Ídem*, p. 11.

<sup>79</sup> *Ídem*, p. 9.

automatizadas, a decir, desde la etapa de diseño del sistema, hasta el monitoreo de su aplicación y eventuales investigaciones por errores o problemas detectados.

## Regulación internacional

Continuando con la exploración de la transparencia algorítmica en el contexto de la IA en el sector público, se torna esencial examinar cómo legislaciones más avanzadas han abordado este complejo tema. Para este propósito, se dirigirá la atención hacia la Unión Europea (UE) y Uruguay, dos regiones que han demostrado un compromiso destacado en la formulación de marcos normativos avanzados para guiar la implementación de la IA, con especial énfasis en la transparencia algorítmica.

### **Legislación de la Unión Europea (UE): Enfoque Holístico y Énfasis en la Ética**

La UE ha liderado la vanguardia en la regulación de la IA, reconociendo la necesidad de salvaguardar derechos fundamentales y valores éticos en la era de la digitalización. En su enfoque hacia la transparencia algorítmica, la UE ha enfatizado la importancia de garantizar que los ciudadanos comprendan cómo las decisiones automatizadas impactan sus vidas.

La regulación de la UE destaca la transparencia como un pilar central para mitigar riesgos, fomentar la rendición de cuentas y construir la confianza pública. El énfasis recae en proporcionar explicaciones claras sobre el funcionamiento de los algoritmos y en garantizar que los ciudadanos sean informados sobre el uso de la IA en situaciones críticas, como la toma de decisiones que afectan derechos legales significativos.

De la misma manera en que previamente se dio énfasis en la importancia de la transparencia como condición, la UNESCO señala que *“La transparencia y la explicabilidad de los sistemas de IA suelen ser condiciones previas fundamentales para garantizar el respeto, la protección y la promoción de los derechos humanos,*

*las libertades fundamentales y los principios éticos. La transparencia es necesaria para que los regímenes nacionales e internacionales pertinentes en materia de responsabilidad funcionen eficazmente*<sup>80</sup>, agrega además que *“La falta de transparencia también podría mermar la posibilidad de impugnar eficazmente las decisiones basadas en resultados producidos por los sistemas de IA y, por lo tanto, podría vulnerar el derecho a un juicio imparcial y a un recurso efectivo, y limita los ámbitos en los que estos sistemas pueden utilizarse legalmente”*<sup>81</sup>

Ahora bien, es fundamental tener en cuenta que esta inclinación hacia la transparencia y explicabilidad no es absoluta, sino que debe adaptarse al escenario social y evitar la transgresión a otros principios, de esta forma, la UNESCO acierta al afirmar que *“Si bien hay que hacer todo lo posible por aumentar la transparencia y la explicabilidad de los sistemas de IA, incluidos los que tienen un efecto extraterritorial, a lo largo de su ciclo de vida para respaldar la gobernanza democrática, el grado de transparencia y explicabilidad debería ser siempre adecuado al contexto y al efecto, ya que puede ser necesario encontrar un equilibrio entre la transparencia y la explicabilidad y otros principios como la privacidad, la seguridad y la protección”*.<sup>82</sup>

### **Legislación en Uruguay: Enfoque Participativo y Protección de Derechos**

Por su parte, Uruguay ha adoptado un enfoque igualmente progresista, destacando la importancia de la transparencia algorítmica para salvaguardar los derechos fundamentales y garantizar la equidad en la utilización de la IA. La legislación uruguaya ha priorizado la participación ciudadana en la toma de decisiones sobre el uso de algoritmos, reconociendo la necesidad de un diálogo inclusivo para modelar políticas efectivas.

---

<sup>80</sup> UNESCO. (2022). Recomendación sobre la ética de la inteligencia artificial. Publicado por la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, 7, place de Fontenoy, 75352 París 07 SP, Francia. (p. 22)

<sup>81</sup> *Ídem.*

<sup>82</sup> *Ídem.*

La transparencia algorítmica en Uruguay se ha abordado desde la perspectiva de proteger la privacidad y evitar la discriminación. A su vez, otorga un rol primordial al interés general como uno de los principios fundamentales a la hora de regular la IA, en este sentido, la Agencia de Gobierno Electrónico y Sociedad de la Información y del Conocimiento (Agesic) en su Estrategia de IA señala sobre este principio que:

*“Las soluciones basadas en IA impulsadas desde el Estado deben tender al interés general, garantizando la inclusión y la equidad. Para esto, se debe trabajar específicamente para reducir la posibilidad de sesgos no deseados en datos y modelos utilizados que puedan impactar de manera negativa en las personas o favorecer prácticas discriminatorias”<sup>83</sup>*

Dicha legislación establece la obligación de explicar el razonamiento detrás de las decisiones automatizadas, especialmente aquellas que afectan a los individuos. Respecto a la transparencia, su estrategia impone dos deberes, a decir:

*“Poner a disposición los algoritmos y datos utilizados para el entrenamiento de la solución y su puesta en práctica, así como las pruebas y validaciones realizadas.*

*Visibilizar explícitamente, a través de mecanismos de transparencia activa, todos aquellos procesos que utilicen IA, ya sea en la generación de servicios públicos o en el apoyo a la toma de decisiones”<sup>84</sup>.*

Se busca así asegurar que los ciudadanos comprendan las implicaciones de la IA en sus vidas y tengan la capacidad de impugnar decisiones que consideren injustas o sesgadas.

Es necesario por su parte mencionar que el 8 de junio del 2023, Uruguay se adhirió a la Recomendación Ética de IA de la UNESCO, día en el cual las autoridades de

---

<sup>83</sup> Agesic. (2019). Estrategia de Inteligencia Artificial para el Gobierno Digital. (p. 9).

<sup>84</sup> *Ídem*.

Agesic y UNESCO firmaron una carta de intención de cooperación técnica para trabajar en conjunto en IA, Ciudadanía Digital, Educación e Inclusión Tecnológica.<sup>85</sup>

Ambas jurisdicciones comparten motivaciones comunes para regular la transparencia algorítmica: preservar la confianza pública, proteger los derechos individuales y fomentar un despliegue ético de la IA en el sector público. A su vez, dan particular énfasis a la necesidad de salvaguardar la privacidad, evitar la discriminación y empoderar a los ciudadanos para comprender y cuestionar las decisiones automatizadas ha sido central en ambas legislaciones.

Las formas de regulación incluyen la imposición de obligaciones explícitas de transparencia<sup>86</sup>, la promoción de la participación ciudadana en la toma de decisiones algorítmicas<sup>87</sup> y la creación de mecanismos para impugnar decisiones automatizadas<sup>88</sup>. Este análisis comparativo proporciona una base sólida para evaluar y contextualizar el enfoque nacional hacia la transparencia algorítmica en el sector público, contribuyendo a un marco normativo que refleje las mejores prácticas y lecciones aprendidas a nivel internacional.

## Tratamiento de la Transparencia Algorítmica en Chile

La manera en que se ha abordado la Transparencia Algorítmica en nuestro país, al 2023, aún no se materializa normativamente, sin embargo, es fundamental destacar el profundo interés que han demostrado diversos actores de la sociedad respecto a

---

<sup>85</sup> Gobierno de Uruguay. Ver: <https://www.gub.uy/agencia-gobierno-electronico-sociedad-informacion-conocimiento/comunicacion/noticias/uruguay-adhiere-recomendacion-etica-inteligencia-artificial-unesco>

<sup>86</sup> “La responsabilidad ética y la obligación de rendir cuentas de las decisiones y las acciones basadas de alguna manera en un sistema de IA siempre deberían ser atribuibles, en última instancia, a los actores de la IA conforme a la función que tengan en el ciclo de vida del sistema de IA”. UNESCO. (2022). Recomendación sobre la ética de la inteligencia artificial. Publicado por la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, 7, place de Fontenoy, 75352 París 07 SP, Francia. (p. 23)

<sup>87</sup> “...a fin de garantizar una participación pública efectiva, de modo que todos los miembros de la sociedad puedan adoptar decisiones informadas sobre su utilización de los sistemas de IA y estén protegidos de influencias indebidas.” Y “La participación de las diferentes partes interesadas a lo largo del ciclo de vida de los sistemas de IA es necesaria para garantizar enfoques inclusivos de la gobernanza de la IA, de modo que los beneficios puedan ser compartidos por todos, y para contribuir al desarrollo sostenible.”. Ídem, (p. 23)

<sup>88</sup> “Los Estados Miembros deberían elaborar, examinar y adaptar, según proceda, marcos reguladores para alcanzar la rendición de cuentas y la responsabilidad por el contenido y los resultados de los sistemas de IA en las diferentes etapas de su ciclo de vida”. Ídem, (p. 28).

este tema. En este sentido, el reconocimiento de la Transparencia Algorítmica como un principio clave a la hora de regular la IA y su implementación a nivel nacional, se ha manifestado de múltiples formas, por ejemplo, en la Política Nacional de IA, en las Recomendaciones del CPLT, en estudios realizados por instituciones locales como la Escuela de Gobierno de la UAI o internacionales como el BID, entre otros.

La presente sección se abocará exclusivamente a analizar la forma en que se ha abordado la Transparencia Algorítmica en nuestro país, con el objetivo de realizar un diagnóstico objetivo de tal principio para así, finalmente, poder ofrecer un panorama más claro en el que se pueda reconocer tanto el avance en esta materia como los posibles puntos pendientes, y de esta manera, poder aportar propuestas útiles para forjar una hoja de ruta eficaz y eficiente, desde un enfoque no solo regulatorio sino también social.

### **Transparencia Algorítmica en la Política Nacional de IA**

La Política Nacional de IA se constituye como un elemento clave en la promoción del debate en torno a la aplicación de IA en el sector público, toda vez que es uno de los primeros acercamientos serios y elaborados sobre dicha materia. Si bien la Política ya fue abordada en el segundo capítulo del presente trabajo, es menester profundizar dicho análisis desde la perspectiva de la Transparencia Algorítmica.

La Política, redactada por 12 expertos en diversas disciplinas, dentro de otros temas, hace hincapié en el hecho de que Chile adhirió a las recomendaciones del Consejo de IA de la OCDE, documento plantea 5 principios, dentro de los cuales se presenta en tercer lugar la Transparencia y la Explicabilidad, lo cual permite *a priori* reconocer la trascendencia de dichos conceptos.

Como se anunció en el segundo capítulo, es destacable dentro del documento la estructuración de una serie de objetivos y medidas que se buscan, dentro de los cuales se puede observar el Objetivo 3.1.2 del 3<sup>er</sup> Eje, que pretende “Impulsar la

transparencia algorítmica” en base al establecimiento de estándares y recomendaciones de transparencia algorítmica para aplicaciones críticas. Dicho objetivo, reconoce el importantísimo rol que juega la explicabilidad acerca de cómo operan los algoritmos decisionales utilizados por la administración, así como también destaca la relevancia de los datos empleados en cada etapa de estos sistemas.

A su vez, la Política aborda uno de los riesgos inherentes a la implementación de sistemas de decisiones automatizadas en el sector público, a decir, la presencia de sesgos algorítmicos, buscando velar por su identificación. Asimismo, pretende mitigar otros riesgos de afectación de derechos fundamentales, otorgando particular énfasis en la protección de la privacidad, de los datos personales y la no discriminación arbitraria por parte de la administración.

Por otro lado, la Política recalca la importancia de la Transparencia en materia de IA en comercio digital. Así, señala en su Objetivo 3.3.1 (también del 3<sup>er</sup> Eje) el fomentar el uso ético de la IA en el comercio digital, asegurando transparencia y respeto a las normas de protección de datos. Busca establecer un entorno propicio para la interacción entre sistemas de IA y consumidores, promoviendo prácticas responsables en contratación automatizada y servicios asociados. La iniciativa se enfoca en informar de manera clara a los consumidores sobre el uso de IA y garantizar un consentimiento informado. En decisiones automatizadas, propone que el Estado trabaje con proveedores para implementar algoritmos justos y explicables, evitando discriminación y cumpliendo con los estándares legales. En resumen, la finalidad es promover un comercio digital transparente y respetuoso con los derechos del consumidor, aprovechando los beneficios de la IA en las relaciones de consumo.

## **Recomendaciones del Consejo para la Transparencia: una mirada desde la experiencia.**

En su cuaderno N° 17, titulado “Transparencia Algorítmica. Buenas Prácticas Y Estándares De Transparencia En El Proceso De Toma De Decisiones Automatizadas”, la dirección de estudios del Consejo para la Transparencia realizó un trabajo sumamente útil para la comprensión de la Transparencia Algorítmica como principio y condición imprescindible en la regulación de sistemas de IA en la administración.

El aporte de dicho documento es muy valioso por diversas razones. En primer lugar, otorga un marco conceptual sólido en el que logra explicar al lector todo lo necesario para comprender el fenómeno de la Transparencia Algorítmica. En segundo lugar, dado que recoge adecuadamente diversas fuentes bibliográficas para explicar el rol crucial de la Transparencia. Finalmente, el documento sintetiza de manera precisa los 4 elementos básicos sugeridos para la construcción de sistemas de decisiones automatizadas, los que se resumen de la siguiente forma:

- Transparencia, Responsabilidad y Explicabilidad:
  - Revelar origen y tratamiento de datos.
  - Identificar responsabilidades.
  - Asegurar explicabilidad de algoritmos.
- Auditorías:
  - Garantizar calidad y prevenir sesgos, errores y discriminación algorítmica.
- Marco Legal:
  - Adaptar normativas al desarrollo tecnológico.
  - Determinar derechos y responsabilidades.
  - Supervisión y registro de algoritmos.
- Protección de Datos Personales:
  - Transparentar fuentes y propósitos de algoritmos.
  - Control y anonimización de datos.

- Capacitar personal en manejo de datos.

## **El aporte del BID: un gran paso en el reconocimiento de la Transparencia Algorítmica como principio fundamental**

El Banco Interamericano de Desarrollo, por su parte, ha aportado significativamente al entendimiento de la IA en su vínculo con la administración pública. En este sentido, en su documento “La Gestión Ética De Los Datos. Por qué importa y cómo hacer un uso justo de los datos en un mundo digital”, el BID analiza la forma en que pueden utilizarse los datos en el sector público y plantea los riesgos que esto conlleva, dentro de los que identifica la afectación a la privacidad, la discriminación algorítmica y la opacidad.

Finalmente, dicho documento propone una serie de criterios que apuntan a la gestión ética de datos en la administración. En relación a la Transparencia y la rendición de cuentas, propone las siguientes medidas para abordar la opacidad en el uso de datos, crucial en un contexto de desconfianza ciudadana hacia el gobierno:

1. Delimitar el papel de los datos en decisiones, evitando que los algoritmos reemplacen la responsabilidad humana.
2. Diseñar estrategias de comunicación y participación ciudadana desde el inicio de proyectos, considerando aprensiones y proporcionando información clara.
3. Implementar medidas de mitigación de sesgos en algoritmos para minimizar o prevenir amplificaciones de sesgos existentes en la sociedad.
4. Diseñar mecanismos de rectificación de errores, priorizando solicitudes simples de corrección de datos.
5. Establecer mecanismos de monitoreo interno integrados en procesos de control de gestión institucional.
6. Planificar evaluaciones independientes después de la implementación del proyecto para una rendición de cuentas efectiva y fortalecimiento de la gestión ética de datos.

Junto a lo anterior, es preciso señalar que el BID confeccionó además un manual de formulación de proyectos de IA, que *“pretende ayudar a los responsables de formular proyectos con sistemas de soporte basados en IA a realizar su conceptualización y diseño, así como a liderar posteriormente su ejecución y monitoreo”*<sup>89</sup>. En él, se mencionan conceptos como la opacidad (y el fenómeno de la caja negra) y Transparencia Algorítmica, vinculados con el mundo de los proyectos de IA y las licitaciones, así, atribuye particular importancia a este último concepto, al señalar que *“se trata de un principio ético esencial en la IA porque funciona como un habilitador de otros principios: permite saber qué datos se utilizan, quiénes los utilizan, cómo se utilizan y cómo afectan las decisiones de política pública”*<sup>90</sup>.

En este aspecto, cabe recalcar el importante rol que jugó el grupo de expertos del BIDLab en lo que posteriormente se materializó en la Base Tipo para licitar proyectos de algoritmos e IA, que permitió enfrentar desde una arista particularmente relevante el riesgo de la opacidad. De esta forma *“La necesidad de combatir la opacidad de los sistemas de datos es un tema que va más allá del uso de algoritmos, por lo cual debe abordarse en diferentes niveles y en distintos momentos. En el caso de la administración pública, una instancia clave y sensible en la implementación de medidas tecnológicas es la de las licitaciones públicas”*<sup>91</sup>.

Sobre esto, Romina Garrido, coordinadora de pilotos del proyecto Algoritmos Éticos y subdirectora del GobLab de la Escuela de Gobierno de la UAI, señaló:

*“Este trabajo va a ayudar a todas las entidades públicas chilenas que están regidas por la Ley de Compras a adquirir algoritmos, inteligencia*

---

<sup>89</sup> Denis, G., Hermosilla, M. P., Aracena, C., Sánchez Ávalos, R., González Alarcón, N., & Pombo, C. (Septiembre 2021). Uso responsable de IA para política pública: Manual de formulación de proyectos. Banco Interamericano de Desarrollo. (P. 5)

<sup>90</sup> *Ídem*, p. 41.

<sup>91</sup> Buenadicha, C., Galdon Clavell, G., Hermosilla, M. P., Loewe, D., & Pombo, C. (2019). La gestión ética de los datos: Por qué importa y cómo hacer un uso justo de los datos en un mundo digital. Banco Interamericano de Desarrollo. (p. 20)

*artificial, proyectos de ciencia de datos y sistemas automatizados en general de manera responsable”<sup>92</sup>*

De todo lo anterior, se puede observar que el interés sobre la Transparencia Algorítmica y su consideración como baluarte tanto en la regulación, implementación y fiscalización de sistemas de decisiones automatizados, provienen de diversos sectores, los que fueron exitosamente analizados en esta sección.

A su vez, cada aspecto tratado a lo largo de este trabajo, ha tenido por finalidad dar un adecuado contexto y una pertinente explicación al fenómeno de la IA en el sector público, lo que nos ha llevado a comprender de manera última, la crucial relevancia de la Transparencia Algorítmica como principio imprescindible en la regulación de dicha materia. La totalidad de lo expuesto culminará en el próximo capítulo, en el cual se pretenderá cumplir los objetivos que este trabajo planteó inicialmente, proponiendo así, una serie de consideraciones y recomendaciones, que eventualmente podrán considerarse como una contribución al panorama general de la IA en Chile.

---

<sup>92</sup> ChileCompra. Ver: <https://www.chilecompra.cl/2023/01/ya-se-encuentra-disponible-bases-tipo-para-licitar-proyectos-de-algoritmos-e-inteligencia-artificial-con-requisitos-eticos/>

### Conclusiones

#### Recapitulación

Luego de todo lo que se ha analizado de forma extensiva y detallada a lo largo de este trabajo, esta sección tendrá por finalidad revisar nuevamente los objetivos planteados al inicio, para luego realizar un breve resumen de los contenidos tratados y así, en última instancia, llevar a cabo las pertinentes conclusiones de la investigación y ofrecer al lector y a la sociedad en general, una serie de recomendaciones dirigidas a aportar al proceso de construcción de la normativa que, finalmente, regulará la materia en cuestión.

Este trabajo surgió del interés de quien escribe respecto al nuevo e impactante fenómeno de la Inteligencia Artificial, y particularmente en su ingreso a la vida de las personas a través de la Administración del Estado. En este sentido, se identificó la trascendencia de dicha tecnología en los procesos de toma de decisiones en la esfera pública y los riesgos inherentes a su implementación. A raíz de lo anterior, se pretendió ofrecer un panorama completo que permitiera comprender realmente las implicancias del uso de sistemas automatizados, para así enriquecer el debate, el dialogo y la información al respecto, todo apuntando a contribuir en especial al contexto nacional.

Para lograr lo anterior, este trabajo comenzó por analizar en profundidad la IA, abordando así el concepto y su origen, su historia y desarrollo en el tiempo, explicando a su vez la manera en que funciona y analizando las características que lo hacen tan interesante y complejo.

Posteriormente, se llevó a cabo un profundo análisis de la aplicación de esta tecnología en el ámbito de la Administración pública, explicando su importancia, explorando las distintas formas en que es empleada en múltiples regiones del mundo

y particularmente en Chile. En el caso de nuestro país, se buscó contextualizar de la manera más adecuada posible la llegada de la IA en general y específicamente al sector público. Ulteriormente, se analizó el proceso regulatorio surgido en la última década, que comenzó por diálogos, que llevaron a una estrategia y que, finalmente, dio paso al inicio de su legislación.

Una vez concluido el análisis ya mencionado, el tercer capítulo se avocó a explorar y explicar exhaustivamente los riesgos inherentes a la implementación de sistemas de decisiones automatizados, dentro de los que se identificó el impacto laboral, el riesgo del rezago, los sesgos y la discriminación algorítmica, la privacidad y la protección de los datos personales y finalmente, la opacidad, siendo este último riesgo aquel que se consideró central, y que en definitiva, dio paso a la Transparencia Algorítmica, entendiéndola como la respuesta a la opacidad y constituyéndose de esta manera como el tema central de este trabajo, que sería abordado detalladamente en el capítulo precedente.

Finalmente, y como se mencionó en el párrafo anterior, se buscó abordar la Transparencia Algorítmica de forma sumamente minuciosa y reflexiva, toda vez que constituye el eje central de este trabajo. Para esto, se introdujo rigurosamente este nuevo concepto, junto con otros que a éste se vinculan - como la Explicabilidad -, y se analizó con especial énfasis su gran trascendencia a lo largo de todo el proceso de implementación de los sistemas automatizados.

Para apoyar la hipótesis de la importancia de la Transparencia Algorítmica, se recurrió también a realizar un análisis comparado de legislaciones extranjeras, en especial aquellas que son consideradas vanguardistas en la regulación normativa de la IA, en virtud del cual se buscó dar evidencia de la relevancia de la Transparencia Algorítmica, que resultó ser casi tangible, si se pudiese explicar de alguna forma menos formal.

De esta forma, el trabajo concluye en esta sección, en la cual, como ya fue anunciado, el autor procederá a ofrecer a continuación, de la manera más humilde y crítica

posible, una serie de recomendaciones que en virtud de lo observado, podrían ser consideradas buenas prácticas y eventualmente resultar útiles para nuestra sociedad que – afortunadamente –, todavía se halla en sus etapas iniciales, al menos desde la perspectiva legislativa.

## Recomendaciones

En esta sección, se recogerá gran parte de lo aprendido a lo largo del trabajo, lo cual puede encontrar su origen en experiencias tanto prácticas como normativas extranjeras, y se ofrecerán breves ideas a modo de resumir los puntos clave ya analizados.

### **Transparencia Algorítmica como Principio**

De todo lo que se abordó, se considera imprescindible para cualquier tipo de técnica legislativa que la Transparencia Algorítmica sea considerado como uno de los principios fundamentales en el tratamiento normativo de la IA, tal como lo han hecho los países más avanzados en la materia y como lo han sugerido múltiples instituciones, entre ellas la OCDE, el BID y el CPLT.

Este principio puede ir o no aparejado a la Explicabilidad, que sin duda se constituye como otro concepto clave en miras de regular adecuadamente los sistemas de IA en la administración, lo que de todas formas se tratará en el próximo punto.

Si bien se recomienda otorgar particular importancia a este principio, resulta imperioso tener siempre en cuenta las posibles colisiones que pudiese tener con otros principios, derechos fundamentales o fenómenos sociales, como podrían serlo los derechos de autor por una parte, o la inversión y/o producción de I+D+i.

Lo que sí es evidente, es que sin la Transparencia como principio explícitamente consagrado, el riesgo de la opacidad podría materializarse en una pérdida

importante en la confianza institucional y sobretodo en la tecnología misma. Por esto, es necesario destacar las palabras de la UNESCO al señalar que:

*“Desde un punto de vista sociotécnico, una mayor transparencia contribuye a crear sociedades más pacíficas, justas, democráticas e inclusivas. Posibilita un escrutinio público que puede reducir la corrupción y la discriminación, y también puede ayudar a detectar y prevenir los efectos negativos sobre los derechos humanos.”<sup>93</sup>*

### **Obligación de Explicabilidad**

Uno de los puntos más relevantes es el hecho de que quienes sean los autores o responsables por la implementación de sistemas automatizados de decisiones tengan una obligación legal de proporcionar información acerca de ellos. Establecer este deber es imperativo para diversos objetivos, entre los que podemos destacar:

1. Permitirle a todos los actores sociales, ya sean ciudadanos, empresas, funcionarios públicos, etc., comprender no solo la manera en que opera un sistema algorítmico, sino que también los datos que utiliza, los fines que persigue y la forma en que se justifique su uso, las personas responsables (ya sean naturales o jurídicas), etc.
2. Ponderar el valor público de un proyecto de IA
3. Evitar la opacidad en la toma de decisiones, velando así por la estabilidad institucional de la sociedad.
4. Facilitar la posibilidad de impugnar los sistemas ante la eventualidad de que se genere algún daño o cualquier vulneración a los derechos fundamentales.

---

<sup>93</sup> UNESCO. (2022). Recomendación sobre la ética de la inteligencia artificial. Publicado por la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, 7, place de Fontenoy, 75352 París 07 SP, Francia. (p. 22)

Si bien existen más perspectivas que permiten apoyar la idea de establecer una Obligación de Explicabilidad a nivel legal (y no puramente contractual), este trabajo considera las mencionadas previamente como suficientes para sostener el punto.

### **Régimen de Responsabilidad Algorítmica**

Se recomienda, de la mano con un deber de Explicabilidad que permita identificar previamente y de manera simple a los responsables por eventuales daños, el establecimiento de un Régimen de Responsabilidad Algorítmica, que sea de aplicación especial y que atribuya una obligación de reparación que pueda efectivamente permitirle a las víctimas ser indemnizadas.

Este trabajo no pretende profundizar en materia de Responsabilidad, toda vez que su extensión exige un tratamiento considerablemente más profundo. Sin embargo, valora significativamente para su eventual confección que se considere un régimen que vele por la ciudadanía y a la vez por la I+D+i, procurando evitar sanciones que deriven en una disminución de su desarrollo, lo que se tratará con mayor detalle en el próximo apartado.

### **Equilibrio Regulatorio**

Una regulación que pretende ser responsable debe prestar atención a todos los actores involucrados, en este sentido, será menester que en la búsqueda de la protección de algunos derechos fundamentales no se vean otros afectados. En este sentido, se habló de métodos para evitar la discriminación algorítmica o la opacidad, y quedó claro que la Transparencia se muestra como una de las más útiles y necesarias.

Pese a lo anterior, es imprescindible recordar a la hora de legislar que una regulación excesivamente estricta podría derivar en otros riesgos, como la afectación de derechos fundamentales o daños en la economía y sociedad a largo plazo. Un ejemplo

de lo primero puede observarse en el caso de que, en miras de garantizar la Transparencia de un sistema de IA, se imponga a los autores o quien sea que detente un derecho de propiedad intelectual un deber de compartir el código fuente o en definitiva, cualquier información en principio confidencial. Esto a su vez, podría poner en riesgo su participación en el mercado o derechamente perder la exclusividad de su idea, lo que eventualmente podría derivar en un desincentivo, lo cual es señalado por el profesor Araya Paz y que se mencionó en el tercer capítulo, al decir que:

*“Y en el caso de que esas empresas se constituyan en Chile, exigiremos tantos requisitos que su operación en el país será compleja y burocrática. El resultado lógico y natural es que desincentivará la innovación en tecnología en el país (que ya es precaria) y nos mantendrá como dependientes de tecnologías desarrolladas por terceros; las empresas no se querrán arriesgar a invertir en un país que pone tantas trabas; los pequeños empresarios y desarrolladores no podrán competir con las grandes empresas dadas las altas barreras de entrada; la inteligencia artificial desarrollada será predecible y en nada se distinguirá de los programas computacionales ya existentes. En el fondo, esta vía nos conduce a un oscurantismo tecnológico brutal.”<sup>94</sup>*

Si bien existen diversos métodos para mitigar los riesgos tanto de la IA en el sector público como aquellos que son propios de una regulación ineficiente, este trabajo reconoce en particular la distinción que realizaron los holandeses en las cláusulas contractuales de la Ciudad de Ámsterdam, identificando diferencias entre la Transparencia Procedimental, la Transparencia Técnica y la Explicabilidad, lo cual daría paso a atribuir a cada uno de estos conceptos determinados tipos de información atendiendo a circunstancias que podrían hacer exigible uno u otro, como se detalló en el capítulo previo.

---

<sup>94</sup> Araya Paz, C. (Diciembre de 2020). Desafíos legales de la inteligencia artificial en Chile. Revista Chilena de Derecho y Tecnología, (p.287-288).

Ahora bien, es importante considerar lo que explica al respecto la profesora Gallego Córcoles:

*“Como cláusula de cierre del sistema se prevé una especie de prelación entre los distintos tipos de transparencia. Así, como se señala en el documento, es probable que, en muchas situaciones, la información proporcionada en el contexto de la transparencia procedimental, transparencia técnica, y explicabilidad coincidan parcialmente. Por ello en las cláusulas se establece que, en ese caso, la Administración podrá elegir en todo momento el régimen que le resulte más favorable. Con ello se pretende evitar que el contratista pueda ampararse los límites a la transparencia técnica para impedir que la Administración haga explicable un sistema algorítmico. Por lo tanto, si cierta información entra en el ámbito de la transparencia técnica y es necesaria para hacer un sistema algorítmico explicable, la Administración será libre de revelar la información relevante y compartirla con terceros.”<sup>95</sup>*

Lo anterior se manifiesta como una técnica legislativa que permitiría a la Administración garantizar la Transparencia de los sistemas empleados y a la vez, proteger los derechos de sus autores, esto como consecuencia de establecer distintos niveles de Transparencia Algorítmica exigibles dependiendo del objetivo que se busque, en contraposición a un sistema que estableciera una única acepción del concepto de Transparencia, que podría resultar o en un déficit regulatorio o en una vulneración de los derechos ya mencionados.

---

<sup>95</sup> Gallego Córcoles, I. (2023). La contratación de soluciones de inteligencia artificial. En E. Gamero Casado & F. L. Pérez Guerrero (Eds.), Inteligencia Artificial y Sector Público: Retos, límites y medios (Cap. XI, p. 535). Valencia: Tirant lo Blanch

## Mecanismos de Cumplimiento en Transparencia

Tal como se analizó en el capítulo anterior, la Transparencia Algorítmica debe considerarse como una condición, y así lo han recalcado expertos y diversas instituciones, entre ellas el CPLT.

En este sentido, los profesores Llama, Mendoza y Graff van más allá al sostener que *“Se reitera que la transparencia es una condición antes, durante y después para la regulación, puede ayudar a desarrollar de una manera más creativa la correulación y otros tipos de enfoques alternativos, donde no se requiere que las entidades intenten resolver unilateralmente todos los desafíos que enfrentan los sistemas de IA.”*<sup>96</sup>

En esta línea, los autores concluyen lo siguiente:

*“Se hace un llamado a tener una regulación multinivel, donde se combinen instrumentos legales y regulatorios tradicionales, así como alternativos, con el fin de adoptar, evaluar y modificar los avances de la IA, toda vez que la flexibilidad a través de este conjunto de regulaciones podrá crear escenarios de legislación experimental, abogando en todos los sentidos por una regulación dinámica con procesos de retroalimentación y roles definidos en su supervisión.”*<sup>97</sup>

En base a esto, se sugiere abordar el objetivo de Transparencia desde diversas perspectivas y en cada etapa de un sistema, promoviendo así que los creadores otorguen información – la mayor posible, pero resguardando sus derechos - para efectos de otorgarle mayor explicabilidad a los sistemas y de facilitar el control *ex post*.

---

<sup>96</sup> Llamas Covarrubias, J. Z., Mendoza Enríquez, O. A., & Graff Guerrero, M. (2022). Enfoques regulatorios para la inteligencia artificial (IA) [Regulatory approaches to artificial intelligence (AI)]. DOI: 10.7764/R.493.2. (p. 47).

<sup>97</sup> Ídem, p. 55.

## **Organismo Fiscalizador Autónomo**

En consonancia con la recomendación anterior, este trabajo sostiene que la creación de una institución cuyo propósito sea asegurar la Transparencia Algorítmica, tendría efectos positivos en la garantización de dicho principio.

Ahora bien, es necesario adecuar las expectativas de lo que se busca y las herramientas disponibles, siendo posiblemente inviable en términos de infraestructura y financiamiento la creación de un organismo autónomo que fiscalice la Transparencia Algorítmica.

Como alternativa, se sugiere expandir las facultades de alguno de los entes fiscalizadores ya existentes para abordar esta tarea, como podría serlo el CPLT o incluso el Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

## **Publicidad Activa y Capacitaciones**

La publicidad activa en el contexto de la regulación de la IA en el sector público se refiere a la divulgación proactiva de información relevante y comprensible sobre el uso de algoritmos y sistemas de IA por parte de las entidades gubernamentales.

Este concepto está estrechamente vinculado al principio de transparencia algorítmica, que busca garantizar que las decisiones automatizadas sean comprensibles y accesibles para los ciudadanos. En otras palabras, es uno de los mecanismos que directamente permiten materializar el deber de Transparencia Algorítmica.

Esto supone que las entidades gubernamentales deban ser proactivas al comunicar cómo se utilizan los algoritmos en la toma de decisiones, especialmente cuando estas decisiones tengan implicaciones significativas en los derechos fundamentales de los individuos.

La divulgación activa puede abordar varios aspectos clave:

1. **Funcionamiento del Algoritmo:** Las entidades gubernamentales deben proporcionar información detallada sobre cómo funcionan los algoritmos utilizados en la toma de decisiones. Esto implica explicar los procesos subyacentes, los criterios de toma de decisiones y cualquier lógica algorítmica relevante.

2. **Impacto en los Ciudadanos:** La publicidad activa debe incluir información sobre cómo las decisiones automatizadas afectan a los ciudadanos. Esto implica detallar las áreas en las que se utiliza la IA, las decisiones específicas que puede tomar y cómo los ciudadanos pueden interactuar con el sistema.

3. **Garantías Éticas y Derechos:** Es esencial que la publicidad activa aborde las garantías éticas incorporadas en el diseño y la implementación de los algoritmos. Además, se debe destacar cómo se protegen los derechos fundamentales de los ciudadanos, como la privacidad y la no discriminación.

4. **Procedimientos para Impugnar Decisiones:** Debe proporcionarse información clara sobre los procedimientos que los ciudadanos pueden seguir para impugnar decisiones tomadas por algoritmos. Esto refuerza la rendición de cuentas y brinda a los ciudadanos la capacidad de cuestionar decisiones que perciben como injustas o erróneas.

La relación entre la publicidad activa y la transparencia algorítmica es fundamental para construir la confianza en la implementación de sistemas automatizados de decisiones. Cuando las entidades gubernamentales adoptan prácticas de publicidad activa, están reconociendo la importancia de informar a los ciudadanos y partes interesadas sobre el uso de la IA, fomentando así la rendición de cuentas y fortaleciendo los cimientos éticos de la implementación de la inteligencia artificial en el ámbito gubernamental.

De esta manera, la capacitación de funcionarios públicos y la divulgación de material educativo en todos los estamentos sociales se muestra como una herramienta que no puede ser ignorada. De lo contrario, la implementación de estos sistemas podría tener un impacto laboral que repercuta gravemente en la Administración, como consecuencia de la incapacidad de aquellos funcionarios para adaptarse a esta tecnología aparentemente incomprensible.

Finalmente, este trabajo hace un llamado enfático a los diversos actores de nuestra sociedad instruidos en esta materia a que adopten un rol protagónico en la educación de las personas, particularmente aquellas que más puedan verse afectadas por los riesgos que en este trabajo se expusieron.

## Bibliografía

1. Gamero Casado & F. L. Pérez Guerrero (2023). Inteligencia Artificial y Sector Público: Retos, límites y medios. Valencia, Tirant lo Blanch.
2. Garrido, R., Lapostol, J. P., & Hermosilla, M. P. (2021). Transparencia algorítmica en el sector público. Escuela de Gobierno de la Universidad Adolfo Ibáñez y el Consejo para la Transparencia.
3. Biblioteca del Congreso Nacional de Chile
4. Departamento de Seguridad Nacional de Estados Unidos (DHS).
5. Servicio de Impuestos Internos de Estados Unidos (IRS).
6. Fondo Monetario Internacional (IMF).
7. Ministerio de Ciencia, Tecnología, Conocimiento e Innovación
8. Política Nacional de Inteligencia Artificial
9. Rouhiainen, L. (2018). Inteligencia Artificial; 101 cosas que debes saber hoy sobre nuestro futuro. Alienta Editorial.
10. Consejo para la Transparencia (2020). Cuaderno de trabajo N°17: Transparencia algorítmica: Buenas prácticas y estándares de transparencia en el proceso de toma de decisiones automatizadas. Dirección de Estudios. Santiago, Chile.
11. UNESCO. (2022). Recomendación sobre la ética de la inteligencia artificial. Publicado por la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, , París, Francia.
12. Buenadicha, C., Galdon Clavell, G., Hermosilla, M. P., Loewe, D., & Pombo, C. (2019). La gestión ética de los datos: Por qué importa y cómo hacer un uso justo de los datos en un mundo digital. Banco Interamericano de Desarrollo.
13. Araya Paz, C. (2020). Desafíos legales de la inteligencia artificial en Chile. Revista Chilena de Derecho y Tecnología.
14. Russell, S. J.; Norvig, P. (2004). Inteligencia Artificial. Un Enfoque Moderno Segunda Edición Pearson Educación, S.A., Madrid.

15. Scherer, Matthew U. (2015). Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies. *Harvard Journal of Law & Technology*, Vol. 29, No. 2.
16. Hoffmann-Riem, W. (2020). Artificial Intelligence as a Challenge for Law and Regulation. In *Regulating Artificial Intelligence*.
17. Llamas Covarrubias, J. Z., Mendoza Enríquez, O. A., & Graff Guerrero, M. (2022). Enfoques regulatorios para la inteligencia artificial (IA).
18. Denis, G., Hermosilla, M. P., Aracena, C., Sánchez Ávalos, R., González Alarcón, N., & Pombo, C. (2021). *Uso responsable de IA para política pública: Manual de formulación de proyectos*. Banco Interamericano de Desarrollo.
19. Agesic (2019). *Estrategia de Inteligencia Artificial para el Gobierno Digital*.
20. Urquidi, M., & Ortega, G. (2020). *Inteligencia artificial para la búsqueda de empleo: Cómo mejorar la intermediación laboral en los servicios de empleo*. División de Mercados Laborales, Sector Social, Banco Interamericano de Desarrollo.
21. Walraven, E., Spruijtenburg, D., Wilmink, I., & Schreuder, M. (2021). *Artificial Intelligence and Traffic Management*.
22. Abeliuk, A. (2021). *Historia y evolución de la inteligencia artificial*. *Revista Universidad de Chile*.
23. Webb, M. (2020). *The Impact of Artificial Intelligence on the Labor Market*. Stanford University.
24. Autor, D., Chin, C., Salomons, A. M., & Seegmiller, B. (2022). *New Frontiers: The Origins and Content of New Work, 1940–2018*. National Bureau of Economic Research.
25. Porrúa, M., Lafuente, M., Mosqueira, E., Roseth, B., & Reyes, A. M. (2021). *Transformación digital y empleo público: el futuro del trabajo del gobierno*. Banco Interamericano de Desarrollo.
26. Alonso, C., Berg, A., Kothari, S., Papageorgiou, C., & Rehman, S. (2020). *Will the AI Revolution Cause a Great Divergence?* International Monetary Fund.
27. Estévez Almenzar, M. y otros (2022). *Glossary of human-centric artificial intelligence*, Joint Research Centre, Unión Europea, Luxemburgo.
28. HLEG (2019): *Directrices éticas para una IA fiable*, Comisión Europea.

29. State Of Wisconsin Vs Eric L. Loomis (2016)
30. Ali M., Sapiezynski P., Bogen M., Korolova A., Mislove A., Rieke A. (2019). Discrimination through Optimization: How Facebook's Ad Delivery Can Lead to Biased Outcomes. Proc. ACM Hum.-Comput. Interact. 3, CSCW, Article 199.
31. Luo, Y., & Guo, R. (2021). Facial Recognition in China: Current Status, Comparative Approach and the Road Ahead. University of Pennsylvania Journal of Law and Social Change, 25 (2).
32. Standard Clauses for Procurement of Trustworthy Algorithmic Systems, (2021).