

Intervención, supervisión y control humano en las decisiones automatizadas*

Human Intervention, Oversight and Control in Automated Decision-Making

MICHELLE AZUAJE PIRELA**

Universidad Alberto Hurtado (Chile)

ALBERTO CODDOU MC MANUS***

Universidad Católica de Chile (Chile)

PABLO CONTRERAS VÁSQUEZ****

Universidad Central de Chile (Chile)

PABLO VIOLLIER BONVIN*****

Universidad Central de Chile (Chile)

Resumen: En este artículo se describe y analiza la supervisión y el control humano en el contexto de las decisiones automatizadas en Chile y la Unión Europea. Particularmente, se examinan normas sobre medidas de control y supervisión humana adoptadas (o en vías de adopción) en los ámbitos de la protección de datos personales, la regulación de la inteligencia artificial y el trabajo en plataformas. Asimismo, se exploran las distintas justificaciones para la intervención humana en dichas áreas, así como las limitaciones y los desafíos para su implementación.

Palabras clave: Supervisión humana, control humano, decisiones automatizadas, protección de datos personales, trabajo en plataformas, Ley de Inteligencia Artificial.

* Este trabajo es parte de la investigación de los proyectos: ANID Fondecyt Regular N.º 1230895: «Los límites a la adopción de decisiones automatizadas en la Administración del Estado» (2023-2026); ANID Fondecyt de Iniciación N.º 11220370: «Discriminación algorítmica en el Estado de Bienestar Digital y en las plataformas digitales de trabajo», y Ministerio de Ciencia e Innovación (España), Expediente N.º PID2021-123170OB-I00: «Claves para una justicia digital y algorítmica con perspectiva de género».

** Doctora en Derecho por la Universidad Autónoma de Chile (Chile). Coordinadora académica del Doctorado en Derecho e investigadora asociada del Centro Interdisciplinar de Políticas Públicas (CiPP) de la Universidad Alberto Hurtado (Chile).

Código ORCID: 0000-0002-8233-6928. Correo electrónico: mazuaje@uahurtado.cl

*** Doctor en Derecho por la University College of London (Reino Unido). Profesor College de la Escuela de Gobierno de la Pontificia Universidad Católica de Chile (Chile).

Código ORCID: 0000-0003-2041-2304. Correo electrónico: acoddou@uc.cl

**** Doctor en Derecho por la Northwestern University (Estados Unidos). Director del Doctorado en Derecho de la Universidad Central de Chile (Chile).

Código ORCID: 0000-0002-1131-182X. Correo electrónico: pablo.contreras@uccentral.cl

***** Abogado. Advanced LLM in Law and Digital Technologies por la Universidad de Leiden (Países Bajos). Estudiante del Doctorado en Derecho de la Universidad Central de Chile y la Universidad de Sevilla (España), financiado por ANID Subdirección de Capital Humano/Doctorado Nacional/2024-folio N.º 21241028.

Código ORCID: 0000-0001-9893-7974. Correo electrónico: pablo.viollier@alumnos.uccentral.cl

Abstract: This article describes and analyzes human oversight and control in the context of automated decisions in Chile and the European Union. It examines regulations on human oversight and control measures adopted (or in the process of being adopted) in the areas of personal data protection, regulation of artificial intelligence and platform work. The analysis explores the different justifications for human intervention in these areas, as well as the limitations and challenges to their implementation.

Keywords: Human oversight, human control, automated decisions, data protection, platform work, AI Act.

CONTENIDO: I. INTRODUCCIÓN.- II. DECISIONES AUTOMATIZADAS, SUPERVISIÓN Y CONTROL HUMANO ANTE LA PROTECCIÓN DE DATOS PERSONALES.- II.1. EL ARTÍCULO 22 DEL REGLAMENTO GENERAL DE PROTECCIÓN DE DATOS (RGPD).- II.2. EL ARTÍCULO 8 BIS DE LA LEY N.º 21.719, QUE REGULA LA PROTECCIÓN Y EL TRATAMIENTO DE DATOS PERSONALES Y CREA LA AGENCIA DE PROTECCIÓN DE DATOS PERSONALES.- III. LA SUPERVISIÓN HUMANA EN LA REGULACIÓN DE LA INTELIGENCIA ARTIFICIAL.- III.1. EL ARTÍCULO 14 DEL REGLAMENTO DE INTELIGENCIA ARTIFICIAL DE LA UNIÓN EUROPEA.- III.2. EL ARTÍCULO 8 DEL PROYECTO DE LEY QUE REGULA LOS SISTEMAS DE INTELIGENCIA ARTIFICIAL (BOLETÍN N.º 16821-19, REFUNDIDO CON: 15869-19).- IV. SUPERVISIÓN Y CONTROL HUMANO ANTE EL TRABAJO EN PLATAFORMAS.- IV.1. LOS ARTÍCULOS 10 Y 11 DE LA DIRECTIVA SOBRE EL TRABAJO EN PLATAFORMAS (DTP).- IV.2. EL ARTÍCULO 152 QUINQUES D DEL CÓDIGO DEL TRABAJO.- V. RECAPITULACIÓN: ¿DE QUÉ HABLAMOS CUANDO HABLAMOS DE INTERVENCIÓN, SUPERVISIÓN O CONTROL HUMANO EN EL CONTEXTO DE LAS DECISIONES AUTOMATIZADAS?.- VI. CONCLUSIONES.

Para que la participación humana sea la respuesta, la ley debe establecer exactamente cómo los humanos mejorarían los problemas con respecto a las predicciones algorítmicas en casos concretos. En lugar de eso, la ley se limita a señalar a un humano y dice: «Ey, hay un humano, así que todo está bien», aunque siga sin estar claro qué debe hacer el humano.

Hideyuki Matsumi y Daniel J. Solove (2023, p. 41)¹

I. INTRODUCCIÓN

El desarrollo de la inteligencia artificial (en adelante, IA) ha permitido contar con sistemas con diversos niveles de autonomía que realizan tareas como inferencias, predicciones, recomendaciones y resolución de problemas a partir del análisis de grandes cantidades de datos. Tales sistemas son particularmente atractivos para su empleo en la toma de decisiones automatizadas. En general, estas se refieren a decisiones «tomadas por» o «con el apoyo de» tecnologías o herramientas

¹ Traducción de los autores.

automatizadas en lugar de seres humanos; o, bien, con una intervención humana escasa. En muchos casos ellas implican el procesamiento de datos personales por medios digitales sin participación humana o en los que al menos una parte del proceso se realiza mediante algoritmos.

Las tecnologías digitales y, en específico, la IA, han generado una creciente dependencia de los sistemas de toma de decisiones automatizadas (SDA) en diversos sectores², lo que plantea profundas preguntas éticas, legales y filosóficas sobre el rol del control, supervisión o evaluación humana en los procesos algorítmicos. En el centro de estos debates se encuentra la idea del «humano en el circuito» o «humano en la cadena» (*human-in-the-loop*), que ha llegado, en algunas posturas, a considerarse como un derecho emergente «a la intervención humana» que busca mitigar los riesgos de opacidad, sesgo y deshumanización en decisiones que afectan a las personas (Adams-Prassl *et al.*, 2022, pp. 139-145). Esta idea, consagrada en marcos normativos como el Reglamento General de Protección de Datos (RGPD) de la Unión Europea (UE) y reflejado en propuestas emergentes como la Ley de IA de la UE, evidencia un reconocimiento creciente de que la rendición de cuentas en los SDA no puede depender exclusivamente de salvaguardas tecnológicas.

Los enfoques sociotécnicos asumen que el avanzado desarrollo de las tecnologías digitales no es capaz de comprender, abordar o mitigar todos los riesgos que derivan de la implementación o aplicación práctica de herramientas de IA en diversos sectores del ámbito público o privado. Estos enfoques, marcados por la idea de que el control o supervisión humanas son garantías adecuadas ante los riesgos derivados del diseño y aplicación de los SDA, están detrás de diversos intentos regulatorios que han abordado este fenómeno. De algún modo, se asume que la inteligencia humana es capaz de detectar los errores o sesgos algorítmicos, y de tomar una decisión que pueda corregirlos o enmendarlos. Asimismo, se ha estimado que en algunos contextos debiera establecerse «una reserva para la toma de ciertas decisiones a humanos», y de ahí que también se hable de una «reserva de humanidad» (Ponce, 2019, p. 13; 2024).

En este escenario, las preguntas sobre la supervisión o el control humano deben abordar ciertas premisas de las que depende un diseño regulatorio adecuado. En primer lugar, podemos abordar la pregunta de si acaso la inteligencia humana puede mostrar un mejor desempeño que la IA con respecto a ciertos criterios objetivos. Diversos estudios recientes

INTERVENCIÓN,
SUPERVISIÓN Y
CONTROL HUMANO
EN LAS DECISIONES
AUTOMATIZADAS

HUMAN
INTERVENTION,
OVERSIGHT
AND CONTROL
IN AUTOMATED
DECISION-MAKING

2 En el sector privado existen casos de uso de la IA para la toma de decisiones automatizadas en áreas tales como el reclutamiento de candidatos (Coddou & Padilla, 2024, pp. 1-34), y el análisis de riesgos y solvencia para el otorgamiento de créditos (Ashwini, 2024, p. 5), entre otras. En el sector público, existen casos de uso en la asistencia y seguridad social, la administración tributaria (Faúndez-Ugalde & Mellado-Silva, 2023; Ossandón, 2021, pp. 123-156), la administración de justicia (Barona, 2023, pp. 14-45), la salud y asistencia sanitaria (Ledford, 2019), etc.

han realizado análisis comparados del rendimiento de SDA en contraste con la actividad humana, mostrando que las máquinas pueden ser más precisas y eficientes en algunos ámbitos (Coddou & Padilla, 2024, p. 6).

En otros casos, si bien los SDA pueden mostrar desempeños equivalentes o levemente inferiores a los seres humanos con respecto a ciertas tareas, la reducción de costos o la radical ganancia en términos de la eficiencia decisional pueden ser incentivos suficientes para automatizar ciertos procesos. Así, la decisión de automatizar un proceso de reclutamiento y selección de personal o de robotizar una determinada tarea, previamente realizada por seres humanos, depende no solo de cuestiones técnicas o de rendimientos comparados, sino del contexto económico, por ejemplo, en lo relativo a la disponibilidad de trabajo humano a bajo costo que haga innecesaria una innovación tecnológica (Acemoglu & Johnson, 2023).

En segundo lugar, los SDA, a pesar de su mayor eficiencia y precisión en ciertas tareas, operan dentro de los límites de la lógica programada y la optimización basada en datos, cuestión que los hace vulnerables a no considerar casos excepcionales o consideraciones que agregan variables que modificarían la decisión algorítmica. En estos casos, resulta necesario que la supervisión humana pueda intervenir en casos excepcionales y justificados, donde el resultado del proceso algorítmico parece no coincidir con el objetivo o finalidad que lo justifica. Estos ejemplos son comunes para los operadores jurídicos, acostumbrados a operar sobre la base de los denominados casos difíciles, donde la aplicación de la norma a un caso concreto no se reduce a un silogismo fácilmente susceptible de ser automatizado (Atria, 2016).

En tercer lugar, la idea de la intervención humana podría derivar de la idea de dignidad humana y de la capacidad humana de razonar moralmente de un modo cualitativamente distinto a la IA, incluso en contextos donde los SDA muestran mejores rendimientos u observan un incremento radical de la eficiencia (Le Moli, 2022). En algunos casos, la supervisión o el control humano se justifica porque necesitamos proveer de legitimidad a decisiones que afectan derechos fundamentales u otros intereses de particular relevancia. Asimismo, el hecho de que dichas decisiones las adopte un ser humano no se justifica meramente por consideraciones funcionalistas (Connor, 2004), sino por cuestiones de legitimidad democrática, como se observa en la justificación de los juicios por jurados; o por la necesidad de mostrar empatía entre funcionarios y administrados, como se justifica en diversas corrientes de la Administración pública (Ponce, 2024; Ranchordas, 2021).

Tal como han sostenido diversos expertos, la restauración de la agencia humana es central para la flexibilidad y la empatía. Tomar decisiones fundamentales para la vida de las personas sin agencia humana,

sea en ámbitos públicos o privados, supone un riesgo de alienación y violaciones a la dignidad individual. Eliminar la relación personal entre el trabajador y el jefe, entre el funcionario y el administrado, entre el ejecutivo y el cliente, suponen un cuestionamiento mismo a la idea del cuidado debido y la responsabilidad (Ponce, 2024). La automatización de procesos, como se ha estudiado en diversos contextos, a veces ha implicado una negación o difusión de la responsabilidad que dificulta la rendición de cuentas que depende, en parte importante, de la evaluación de la agencia humana (Kellog *et al.*, 2020).

Por último, incorporar la agencia humana dentro de los procesos o SDA refuerza la confianza o legitimidad social en la tecnología al abordar las preocupaciones sobre la automatización irresponsable, particularmente cuando los sistemas operan de manera opaca como «cajas negras». Como se ha demostrado en diversos estudios (López *et al.*, 2025), la intervención humana puede justificarse por consideraciones consecuencialistas, pues la técnica no es capaz por sí sola de proveer su propia legitimidad.

Muchas de estas preguntas surgen a partir de problemas derivados de la falta de reflexión crítica en las fases iniciales de los ciclos de vida de SDA, cuando muchas veces se olvida preguntar si acaso es plausible, necesario o adecuado automatizar un proceso. En ciertos contextos, la presión por automatizar procesos o sistemas de toma de decisión deriva en una ausencia de reflexión o razonamiento crítico acerca de los fundamentos. Peor aún, en la actualidad aún existe un vacío de normas técnicas sectoriales que nos permitan hacer análisis comparados que entreguen certeza acerca de cuándo la IA supera el desempeño de los seres humanos en ciertas tareas (Hacker, 2023, p. 16). Otras preguntas, en cambio, derivan de la falta de reflexión filosófica o normativa acerca de los límites que podemos imponer, como sociedades organizadas a través del derecho, a la automatización de ciertos procesos.

En este escenario, los intentos regulatorios por articular la idea, el estándar o el derecho a la intervención humana se están desarrollando a través de reglas sectoriales o fragmentadas que impiden analizar este fenómeno a partir de un enfoque integral. Algunos autores han propuesto la necesidad de desarrollar estándares a partir de la idea de un «debido proceso algorítmico» (Citron & Pasquale, 2014), una suerte de justicia procedimental que permita responder las complejas y diferentes cuestiones que surgen al automatizar procesos que hasta hace poco eran adoptados por personas humanas. Sin embargo, los intentos regulatorios que abordaremos aquí no solo parecen estar cruzados por cuestiones procedimentales. Como señalamos antes, a veces la justificación de la idea de una intervención humana se relaciona con complejas cuestiones filosóficas que subyacen a nuestros arreglos normativos; cuestiones sustantivas que no solamente apuntan hacia un ser humano como

garantía ante la alienación derivada de procesos de tecnificación, sino hacia diversas concepciones acerca de la filosofía de la mente, como el estatus de la conciencia (Wieczorek, 2021).

Mientras tanto, a nivel nacional e internacional se han generado nuevas directrices, principios, regulaciones y estándares que buscan garantizar ese pretendido control, la supervisión o al menos la presencia humana en el proceso de toma de decisiones automatizadas. En la UE, el RGPD consagra en su artículo 22 el derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado. Aunque existen casos de excepción, en ellos se deberán adoptar medidas para garantizar a los interesados «como mínimo el derecho a obtener intervención humana por parte del responsable, a expresar su punto de vista y a impugnar la decisión» (RGPD, art. 22, num. 3, lit. c).

También en la UE, el artículo 14 del Reglamento 2024/1689, por el que se establecen normas armonizadas en materia de IA (en adelante, RIA), se refiere a la «supervisión humana», y establece que los sistemas de IA de alto riesgo deben diseñarse y desarrollarse de modo que puedan ser vigilados efectivamente por personas físicas durante el periodo en el que estén en uso.

Más recientemente, la Directiva sobre el trabajo en plataformas (en adelante, DTP) regula en su artículo 10 la «supervisión por humanos de los sistemas automatizados de seguimiento y de los sistemas automatizados de toma de decisiones», mientras que el artículo 11 se refiere a la «revisión por humanos».

Asimismo, en la actualidad diversos Estados se encuentran en proceso de importantes reformas para hacer frente a estos nuevos desafíos. De hecho, ya existen varias normas o propuestas que contienen obligaciones de control o supervisión humanas y permiten aproximar lo que se espera de esa intervención en algunas áreas sensibles. El derecho chileno no ha sido la excepción y recientemente actualizó su normativa sobre protección de datos personales, incluyendo una regulación sobre el diseño e implementación de procesos de tratamiento automatizado de datos personales. Además, habiendo sido pionero en la región para la implementación de una Política Nacional de IA, actualmente se discute un proyecto de ley que regula el uso de sistemas de IA y que sigue parcialmente el modelo europeo. Ambos instrumentos incluyen una regulación de la intervención o el control humano en SDA que debe integrarse al ordenamiento existente, cuestión que exige una reflexión más profunda, según lo hemos señalado.

Aunque no existe una única definición sobre el alcance de los distintos términos que utilizan los intentos regulatorios, la idea general es que con ellos se busca introducir a seres humanos en el proceso de decisión

automatizado para garantizar una mejor decisión final debido a la supervisión humana del sistema y a la intervención de seres humanos para prevenir, mitigar o, eventualmente, corregir los errores que este podría cometer (Agudo *et al.*, 2024).

La supervisión humana puede ser *ex ante*, lo que implica asegurarse de que la IA está programada correctamente, o *ex post*, en cuyo caso implica revisar las sugerencias de la IA antes de aplicarlas o revisar sus decisiones para definir si se recurren o plantean problemas (Maxwell, 2023, p. 6). En este escenario de experimentación regulatoria, surgen diversas preguntas: ¿en qué casos debe garantizarse? ¿Quién es el llamado a hacerlo? ¿Cuál es la forma? ¿Cuáles son los aspectos tecnológicos o sociales que deben tenerse en cuenta? (Beck & Burri, 2024; pp. 105-132; Lazcoz, 2024, pp. 570-586; Matsumi & Solove, 2023; Robbins, 2023). Estas y otras preguntas constituyen la motivación principal de este artículo.

Así, el objetivo principal de este artículo es describir y analizar diferentes intentos regulatorios sobre la supervisión, el control o la evaluación humana en SDA que han sido seleccionados con el objeto de realizar una reconstrucción o un análisis crítico sobre sus premisas o supuestos principales. En específico, nos interesa analizar el modo en que la regulación europea, por una parte, y la regulación chilena, por la otra, abordan la idea de la supervisión, el control o la intervención humana.

En ese sentido, la contribución de este artículo a la disciplina del derecho y las tecnologías es doble. En primer lugar, se trata de una contribución a un área no suficientemente explorada y a una discusión que recién comienza (Obregón & Lazcoz, 2021, p. 4). En efecto, la idea de la «intervención humana» surge en un contexto de rápidos y constantes cambios, no siempre acompañados de reflexiones, en el que los términos se repiten sin que queden claras sus implicaciones (Robbins, 2023). Por eso, en segundo lugar, el artículo ofrece una visión de conjunto que resume y permite visualizar, en concreto, las implicaciones de la intervención, la supervisión o el control humano en el contexto de las decisiones automatizadas en los ámbitos de la protección de datos personales, la regulación de la IA y el trabajo en plataformas, así como las limitaciones y los desafíos para su implementación.

INTERVENCIÓN,
SUPERVISIÓN Y
CONTROL HUMANO
EN LAS DECISIONES
AUTOMATIZADAS

HUMAN
INTERVENTION,
OVERSIGHT
AND CONTROL
IN AUTOMATED
DECISION-MAKING

II. DECISIONES AUTOMATIZADAS, SUPERVISIÓN Y CONTROL HUMANO ANTE LA PROTECCIÓN DE DATOS PERSONALES

II.1. El artículo 22 del Reglamento General de Protección de Datos (RGPD)

Uno de los principales impulsos regulatorios en materia de supervisión y control humano en el contexto de las decisiones automatizadas se materializó con la entrada en vigor del RGPD en el año 2018³. En particular, su artículo 22 representa el intento más sofisticado de regular la toma automatizada de decisiones basada en el tratamiento automatizado de datos personales de los titulares⁴.

En sus aspectos más sustantivos, el artículo 22 establece el derecho de los titulares a no ser objeto de toma de decisiones automatizadas basadas únicamente en el tratamiento automatizado de sus datos, cuando dicha decisión produzca efectos jurídicos en aquellos o les afecte significativamente de modo similar. Este tipo de decisiones automatizadas solo podrán realizarse en la medida en que la decisión se encuentre en alguna de las hipótesis contenidas en el inciso 2 del artículo 22, a saber: a) ser «necesaria para la celebración o la ejecución de un contrato», b) «está autorizada por el Derecho de la Unión o de los Estados miembros que se aplique al responsable del tratamiento», y c) «se basa en el consentimiento explícito del interesado».

Adicionalmente, en esos casos, los responsables están obligados a implementar «medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado», además de «como mínimo garantizar» los siguientes derechos: a) «obtener intervención humana por parte del responsable», b) «a expresar su punto de vista» y c) «a impugnar la decisión»⁵. Como última medida de resguardo, el inciso final del artículo 22 establece que este tipo de decisiones no se podrán basar en las categorías especiales de datos personales, salvo que: a) el interesado haya entregado su «consentimiento explícito» (art. 9,

3 Para un análisis pormenorizado del artículo 22 del RGPD desde una perspectiva europea, véase Palma (2019, pp. 1-23) y Medina (2022, pp. 141-171).

4 Esta redacción no ha estado exenta de problemas interpretativos, entre los que podemos considerar: a) si el artículo establece una prohibición general a la toma de decisiones automatizadas (excepto cuando se cumplan ciertas condiciones) o si se trata de un derecho *ex post* de oposición de parte del titular, b) qué implica el uso del término «basada únicamente en el tratamiento automatizado» y si esto permitiría que una intervención humana simbólica haga inaplicables las garantías del artículo, y c) cuál es el nivel necesario de afectación al titular para que operen las garantías contenidas en el artículo y que puedan ser consideradas como una decisión que afecta al titular «significativamente de modo similar». Muchos de estos problemas interpretativos fueron solucionados por la dictación de las Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679 (WP251rev.01) por parte del Grupo de trabajo sobre protección de datos del artículo 29.

5 Estos tres derechos no serán aplicables en la medida en que las decisiones se tomen en la hipótesis del artículo 22, inciso 2, literal b); es decir, cuando esté autorizada por el derecho de la Unión o de los Estados miembros que se aplique al responsable del tratamiento.

apart. 2, lit. a) y b) el tratamiento sea «necesario por razones de un interés público esencial, sobre la base del Derecho de la Unión o de los Estados miembros» (lit. g); así como la obligación de demostrar que «se hayan tomado medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado».

Las garantías más relevantes para concretar la supervisión o el control humano son el derecho a obtener una intervención humana por parte del responsable, a que el titular exprese su punto de vista y a impugnar la decisión.

En este contexto, el derecho a obtener una intervención humana tiene como objetivo garantizar al titular que fue afectado por una decisión automatizada el derecho a exigir que durante el proceso de la toma de decisiones exista un ser humano supervisando la operación. En definitiva, se trata de un derecho que busca garantizar que, en última instancia, el proceso esté guiado por el criterio humano y no solamente con base en el tratamiento automatizado de datos. Los requisitos que la intervención humana debe cumplir para satisfacer lo establecido en el artículo 22 fueron desarrollados por el Grupo de trabajo sobre protección de datos del artículo 29 (en adelante, WP29)⁶ y resumidos por la Agencia Española de Protección de Datos (2024) de la siguiente manera:

- *Competencia y autoridad*: el dependiente del responsable debe contar con la autoridad o la tarea asignada que le permita alterar el resultado de la decisión automatizada.
- *Preparación y formación*: el dependiente del responsable debe contar con la capacidad y las aptitudes para evaluar la decisión y los factores que determinan esa decisión con relación al contexto del tratamiento y el sistema automatizado empleado en sus capacidades y limitaciones.
- *Independencia y diligencia en el ejercicio de su competencia*: el dependiente del responsable debe estar libre de presiones desde la organización o desde fuera de esta que condicionen la disputa de la decisión por parte de la persona.
- *Medios para poder ejercer su competencia y cualificación*: el dependiente del responsable debe contar con: a) la capacidad de intervenir en el momento adecuado, antes de aplicar la decisión al individuo; b) la información necesaria para poder ejercer su cualificación; c) los recursos necesarios para poder ejercer su cualificación; y d) el tiempo necesario para ejercer

6 El WP29 fue establecido por la Directiva 95/46/EC y reemplazado por el Comité de Protección de Datos Personales (EDPB, por sus siglas en inglés) de la Unión Europea el día 25 de mayo de 2018, fecha en la que entró en vigor el RGPD.

su cualificación en cada una de las decisiones que sean de su competencia.

Del mismo modo, el revisor debe ser capaz de evaluar todos los datos pertinentes, incluyendo la información adicional facilitada por el titular, y tener la capacidad de entender la manera en que el sistema toma las decisiones de forma automatizada (WP29, 2017, p. 30). De lo contrario, se corre el peligro de que la intervención humana solo se transforme en un mecanismo para legitimar el sistema de toma de decisiones automatizadas, «blanqueando» su funcionamiento con una participación humana simbólica o *token* en un fenómeno que los comentaristas han denominado el «síndrome *computer says yes*»⁷.

El derecho a expresar su punto de vista garantiza al titular la capacidad de proporcionar al responsable del sistema de decisiones automatizadas información adicional que este deberá tener en consideración al momento de supervisar la toma de decisiones. Puesto que muchos de los sesgos, errores y arbitrariedades que estos sistemas cometen se deben a la información que utilizaron para el procesamiento o la ponderación que se le asignó, contar con el punto de vista y la información provista por el titular se transforma en un mecanismo para verificar que el titular se cerciore de que todos los antecedentes están siendo tomados en consideración al momento de tomar la decisión. Del mismo modo, el titular podrá aportar con elementos que vayan más allá de solo proporcionar información adicional, como lo será en aquellos casos en que dé cuenta de su disconformidad con el funcionamiento del sistema, los parámetros utilizados o las medidas de seguridad de elementos que el responsable siempre deberá tomar en consideración.

Por último, el derecho a impugnar la decisión implica la capacidad del titular a objetar jurídicamente el resultado de la decisión automatizada y a exigir su revisión, ya sea ante el mismo responsable del sistema o ante la agencia de protección de datos respectiva, con el objetivo de obtener su revocación o modificación. Sin embargo, como correctamente ha señalado el WP29, la condición necesaria para que el titular sea capaz de impugnar la decisión e incluso expresar su punto de vista es que este pueda entender plenamente cómo se ha tomado la decisión y sobre qué base (WP29, 2017, p. 30). Esto es particularmente cierto si tenemos en consideración las enormes asimetrías de poder (Breidbach, 2024, p. 3) e información (Dowding & Taylor, 2024, p. 68) que pueden existir entre un desarrollador o administrador de un sistema de inteligencia artificial avanzado y un titular promedio de datos personales.

⁷ El síndrome del *computer says yes* se refiere al riesgo de que las personas que han trabajado durante mucho tiempo con sistemas automatizados se vuelvan naturalmente menos inclinadas a cuestionar la precisión de este tipo de sistemas (Google, 2024). Para un análisis en profundidad de los factores que afectan la confianza de las personas en la automatización, véase Hoff y Masooda (2015, pp. 407-434).

Al respecto, uno de los problemas más importantes que han surgido en torno a la interpretación del artículo 22 se relaciona con la existencia del llamado «derecho a una explicación» (Kaminski, 2019, pp. 209-217). Este problema surge porque el considerando 71 del RGPD establece que, junto a los tres derechos antes citados, el titular tiene el derecho a «recibir una explicación de la decisión tomada después de tal evaluación». No obstante, dicho derecho está notablemente ausente del texto del inciso tercero del artículo 22. Esto ha llevado a algunos comentaristas a sugerir que no existe el derecho a una explicación (en términos significativos) en el RGPD (Wachter *et al.*, 2017, p. 41), en particular puesto que los recitales no tienen una operación positiva por sí mismos y no pueden permitir que emerja una expectativa legítima de derecho (Klimas & Vaiciukaite, 2008, p. 32). Si bien no lo hacen de forma directa, las directrices del WP29 parecen indicar que el derecho a una explicación estaría implícitamente contenido no en el artículo 22, sino en las obligaciones generales de información de los artículos 13 y 14. Estos, a su vez, se cumplirían por parte del responsable al entregar información significativa sobre la «lógica aplicada», así como las eventuales «consecuencias previstas» del tratamiento automatizado al titular.

Sin embargo, esto implicaría consagrar una versión particularmente débil del derecho a una explicación, puesto que parece descartar la posibilidad de una explicación *ex post* —esto es, respecto de una decisión específica— en relación con casos individuales de decisión, haciendo del derecho a una explicación solo un nombre dado a un deber de información más intenso que también cubre las consecuencias potenciales del procesamiento automatizado de datos personales. Es decir, se trataría más bien de una explicación sobre el funcionamiento general del sistema a futuro, en contraposición de una explicación de una decisión específica que ya tuvo lugar respecto del titular particular que reclama dicha explicación. Es interesante notar que el borrador de las Directrices del WP29 incluía la posibilidad de que se hiciera una explicación matemática más detallada sobre cómo funciona el sistema algorítmico o de aprendizaje automático, y «si es necesario para permitir a los expertos verificar más a fondo cómo funciona el proceso de toma de decisiones» (Veale & Edwards, 2018, p. 399), pero esa sección fue posteriormente abandonada en la versión final de las Directrices, confirmando aún más la interpretación restrictiva de la información que debe ser revelada a los interesados.

II.2. El artículo 8 bis de la Ley N.º 21.719, que regula la protección y el tratamiento de datos personales y crea la Agencia de Protección de Datos Personales

En Chile, la nueva Ley N.º 21.719, que regula la protección y el tratamiento de datos personales y crea la Agencia de Protección de Datos Personales, publicada el 14 de diciembre de 2024, contiene en su artículo 8 bis un equivalente al artículo 22 del RGPD. Si bien el contenido de ambas disposiciones es casi idéntico, el legislador chileno modificó ciertas redacciones para evitar los problemas de interpretación identificados en Europa. Entre estos cambios podemos mencionar: a) se recalcó que el titular tiene derecho a no ser objeto, pero también a oponerse a las decisiones automatizadas, para enfatizar que se trata de una prohibición general y no un derecho que debe ser ejercido por el titular; b) se eliminó la exigencia de que el tratamiento fuese «únicamente» automatizado para evitar que una participación humana simbólica dejase fuera de aplicación el artículo; c) se rebajó el nivel de afectación exigido al eliminar el término «de forma similar» al nivel de afectación significativo requerido por el artículo 22 del RGPD; y, por último, pero más relevante que los anteriores, d) se consagró expresamente el derecho a una explicación dentro del catálogo de los derechos otorgados al titular^{8 y 9}.

Mientras que el artículo 8 bis no ha sido estrenado, el artículo 22 RGPD sí ha sido aplicado y su interpretación sigue en evolución. Tres antecedentes son especialmente relevantes en este punto. En primer lugar, la interpretación del Tribunal de Justicia de la Unión Europea (TJUE) en el caso *Schufa* (Sentencia C-634/21, 7 de diciembre de 2023); en segundo lugar, por la Opinión de 2024 del European Data Protection Board (EDPB, por sus siglas en inglés) sobre determinados aspectos de la protección de datos relacionados con el tratamiento de datos personales en el contexto de los modelos de IA; y, en tercer lugar, la reciente sentencia del TJUE en el caso *Din & Bradstreet* (Sentencia C-203/22, febrero de 2025) respecto al contenido del derecho a una explicación.

En el primer caso, el TJUE debía resolver sobre los derechos de una persona a la que se le había negado un crédito con base a una puntuación de *score* crediticio por un buró de crédito (*Schufa*). La sentencia decide que la calificación crediticia es una «decisión» a efectos del artículo 22 RGPD y se trata de una actividad que constituye una elaboración

8 Para un análisis sobre las interpretaciones del artículo 22 del RGPD, véase Viollier y Fischer (2023, pp. 151-169).

9 Si bien se consagró explícitamente el derecho a una explicación en el artículo 8 bis, el texto de la ley no entrega detalles sobre el contenido normativo de dicho derecho, el que quedará sujeto a la interpretación administrativa que le entregue la nueva Agencia de Protección de Datos Personales y, eventualmente, al control jurisdiccional de las cortes de justicia.

de perfiles en virtud de una generación automatizada de un valor de probabilidad a partir de los datos de una persona (Sentencia C-634/21, 7 de diciembre de 2023, §§ 47-48). En este punto, lo central es que el valor de probabilidad generado por una agencia de información comercial y comunicado a un banco desempeña un papel determinante en la concesión de un crédito, siendo aplicable el artículo 22.1 RGPD (§ 50). Bajo dicha interpretación, el RGPD amplía el radio de protección más allá de decisiones exclusivamente adoptadas por un sistema de decisión automatizada, permitiendo a los titulares de datos ejercer sus derechos «cuando de ese valor de probabilidad dependa de manera determinante que un tercero, al que se comunica dicho valor, establezca, ejecute o ponga fin a una relación contractual con esa persona» (§ 75). Esta interpretación reduce la distancia entre el artículo 8 bis LPDP y el artículo 22 RGPD en cuanto al alcance de las decisiones cubiertas por ambas normas.

Respecto del segundo caso, el EDPB (2024) examinó un aspecto central de modelos de IA y su impacto en el tratamiento de datos personales: bajo qué circunstancias un modelo de IA que ha sido entrenado utilizando datos personales puede considerarse «anónimo» (§ 35). En este caso, se determinaron dos criterios normativos centrales. Por un lado, se debe considerar la probabilidad de extracción directa (incluso probabilística) de datos personales relativos a las personas cuyos datos personales se utilizaron para entrenar el modelo (§ 43). Por el otro, se debe atender a la probabilidad de obtener, intencionadamente o no, dichos datos personales a partir de consultas que deberían ser insignificantes para cualquier interesado (§ 43).

A estos efectos, el EDPB consideró que se debe evaluar caso a caso si las medidas aplicadas por el responsable del tratamiento para garantizar y probar que un modelo de IA es anónimo son adecuadas y eficaces (§ 45). Para determinar la evaluación de las medidas y la supervisión del modelo, el EDPB brinda un listado no exhaustivo de criterios. Con relación a la fase de diseño del sistema, en primer lugar, se requiere evaluar las medidas adoptadas para evitar o limitar la recogida de datos personales (§ 50). En segundo lugar, con relación a la fase de entrenamiento de un sistema, se debe velar por la implementación del principio de minimización del tratamiento y procurar la anonimización de los datos personales (§ 51). En tercer lugar, se requiere evaluar las opciones metodológicas que pueden reducir significativamente o eliminar la identificabilidad de titulares de datos (§ 52). Finalmente, hay que examinar los métodos que puedan disminuir la probabilidad de obtener datos personales relacionados con los datos de entrenamiento de las consultas (§ 53). Cabe añadir que el EDPB agrega un estándar de documentación de los procesos de auditoría para limitar las posibilidades

de identificación de titulares, así como del entrenamiento del modelo en general (§§ 53 y 56).

En el tercer caso mencionado, el TJUE, en su sentencia de febrero de 2025, dictada en el asunto *Dun & Bradstreet* (C-203/22), de Austria, se pronunció respecto al contenido del derecho a una explicación. En dicho pronunciamiento, el Tribunal estableció que el artículo 15, numeral 1, literal h del RGPD impone al responsable del tratamiento la obligación de proporcionar al interesado una explicación clara, inteligible y suficiente de los procedimientos y criterios empleados en la adopción de una decisión que le concierna, de manera que el titular de los datos se encuentre en condiciones de comprender el proceso decisorio para, con dicha base, ejercer adecuadamente su derecho a impugnarla. Esta sentencia puede ser entendida como la consolidación, en sede jurisprudencial, del reconocimiento del derecho a una explicación como un derecho sustantivo en el marco del RGPD. Sin embargo, la lógica del fallo parece dar cuenta de que este derecho sustantivo se encontraría contenido en el artículo 15 (deberes de información) y no en el artículo 22, el cual regula la toma de decisiones automatizadas.

III. LA SUPERVISIÓN HUMANA EN LA REGULACIÓN DE LA INTELIGENCIA ARTIFICIAL

III.1. El artículo 14 del Reglamento de Inteligencia Artificial de la Unión Europea

El 12 de julio de 2024 se publicó en el *Diario Oficial de la Unión Europea*¹⁰ el Reglamento 2024/1689, por el que se establecen normas armonizadas en materia de inteligencia artificial o RIA. Este ha sido catalogado como la primera norma de carácter general e integral para hacer frente a los desafíos de la IA y se espera que ejerza una influencia significativa en su gobernanza y regulación a nivel mundial, o que llegue a producir un nuevo Efecto Bruselas (Bradford, 2019, pp. 25-66). Esta última idea se ve reforzada por el hecho de que se aplicará a proveedores con independencia de si están establecidos o ubicados en la Unión o en un tercer país (López-Tarruella, 2024a; RIA, 2024, art. 2, num. 1, lit. a).

La normativa europea busca garantizar sistemas éticos, seguros y fiables. Para ello, adopta un enfoque de clasificación de riesgos. En ese sentido, define cuatro niveles de riesgo vinculados con el uso de sistemas de IA: riesgo inaceptable, alto riesgo, riesgo limitado, y riesgo mínimo o nulo.

¹⁰ En vigor desde el 1 de agosto de 2024, pero algunas disposiciones se aplicarán progresivamente: los capítulos I y II se aplican desde el 2 de febrero de 2025; la sección 4 del capítulo III, el capítulo V, el capítulo VII, el capítulo XII y el artículo 78 a partir del 2 de agosto de 2025, a excepción del artículo 101; y el artículo 6, apartado 1, y las obligaciones correspondientes serán aplicables a partir del 2 de agosto de 2027.

En el nivel de riesgo inaceptable se incluyen las «prácticas de IA prohibidas» enumeradas en el artículo 5. Aunque con algunas excepciones, en el listado se incluyen prácticas tales como «aquellas que usan la IA para manipular personas o alterar deliberadamente su comportamiento»; «sistemas de IA para evaluar o clasificar a personas físicas o a colectivos de personas durante un período determinado de tiempo atendiendo a su comportamiento social o a características personales o de su personalidad conocidas, inferidas o predichas»; «sistemas de IA que creen o amplíen bases de datos de reconocimiento facial mediante la extracción no selectiva de imágenes faciales de internet o de circuitos cerrados de televisión»; y «sistemas de categorización biométrica que clasifiquen individualmente a las personas físicas sobre la base de sus datos biométricos para deducir o inferir su raza, opiniones políticas, afiliación sindical, convicciones religiosas o filosóficas, vida sexual u orientación sexual» (RIA, 2024, art. 5), entre otros.

Los sistemas de alto riesgo se establecen en el capítulo III, artículo 6 del RIA. Son aquellos que pueden afectar negativamente a la salud, la seguridad o los derechos fundamentales agrupados en dos criterios: a) la pertenencia a productos regulados en anexos legislativos vinculados con la seguridad de los productos (anexo I), o b) su aplicación en áreas o ámbitos de operación específicos (anexo III). Dentro de estas áreas se encuentran, por ejemplo, la biometría, la gestión y explotación de infraestructura crítica, el acceso y disfrute de servicios privados esenciales y de servicios y prestaciones públicas, el empleo y gestión de trabajadores, la educación y formación profesional, entre otras. La situación considerada de alto riesgo puede ocurrir cuando un sistema de IA de estas características es utilizado conforme a su finalidad prevista o bien cuando se le da un uso indebido razonablemente previsible.

Estos sistemas serán evaluados y están sujetos a obligaciones estrictas antes de su comercialización y a lo largo de su ciclo de vida (arts. 26-27). Asimismo, los ciudadanos tendrán derecho a presentar reclamaciones sobre ellos ante autoridades nacionales específicas¹¹.

Los sistemas de riesgo limitado se categorizan como tales por su falta de transparencia, de modo que, en cuanto a ellos, se establecen requisitos de información y transparencia. Sobre esto, el artículo 50 dispone diversos supuestos que pueden agruparse en cuatro categorías: a) sistemas destinados a interactuar directa o indirectamente con personas físicas (como los *chatbots*); b) sistemas de uso general que producen contenido sintético como audio, imagen, video o texto;

11 La sección 4 establece como «vías de recurso»: el derecho a presentar una reclamación ante una autoridad de vigilancia del mercado (art. 85), el derecho a explicación de decisiones tomadas individualmente (art. 86) y la denuncia de infracciones (art. 87). Para un análisis crítico de estas disposiciones, véase López-Tarruella (2024b, pp. 757-780).

c) sistemas de reconocimiento de emociones y sistemas de categorización biométrica que no se encuentren prohibidos¹²; y d) las ultrasuplantaciones o *deepfakes*.

Los sistemas de riesgo mínimo o nulo son todos los demás que no se encuentren en los supuestos anteriores y, por lo tanto, no están sometidos a requisitos imperativos (Barrio, 2024, p. 45).

En el RIA (2024) la supervisión humana es un principio general presente en varios apartados. En efecto, en el considerando 27, esta se incluye entre los principios éticos que contribuyen a garantizar una IA fiable. En el mismo considerando, se rescata la definición de «acción y supervisión humanas» contenida en las directrices del Grupo independiente de expertos de alto nivel sobre IA, de acuerdo con las cuales «los sistemas de IA se desarrollan y utilizan como herramienta al servicio de las personas, que respeta la dignidad humana y la autonomía personal, y que funciona de manera que pueda ser controlada y vigilada adecuadamente por seres humanos». Igualmente, como se detalla a continuación, las obligaciones de supervisión humana pueden hallarse principalmente en los artículos 14, 16, 26 y 99 del RIA (este último en relación con las sanciones).

En ese sentido, en lo que se relaciona con esta investigación, es clave la relación que existe entre los sistemas de alto riesgo y el contenido del artículo 14 porque, de acuerdo con este, la supervisión humana es una obligación específica para estos. Así, se dispone que los sistemas de IA de alto riesgo deben diseñarse y desarrollarse de forma que puedan ser vigilados efectivamente por personas físicas durante el periodo de uso, y esto incluye dotarlos de herramientas de interfaz humano-máquina adecuadas (RIA, 2024, art. 14, num. 1). En opinión de Beck y Burri (2024), en el RIA la supervisión humana «es sólo uno de los componentes de un enfoque más amplio» y se establecen diversas obligaciones que, si se cumplen correctamente, equivaldrían al «control humano de la IA» (p. 9).

En ese sentido, es un requisito de obligado cumplimiento principalmente para los proveedores de sistemas de IA de alto riesgo, quienes deberán garantizarla desde el diseño. No obstante, no debe perderse de vista que, dado que se extiende por todo el ciclo de vida del sistema, en algunos casos conlleva también obligaciones para los responsables del despliegue, quienes deberán utilizarlos de acuerdo con las instrucciones de uso del proveedor.

¹² De acuerdo con el artículo 50.3, se exceptúan los sistemas de IA utilizados para la categorización biométrica y el reconocimiento de emociones que hayan sido autorizados por ley para detectar, prevenir e investigar delitos, con sujeción a las garantías adecuadas para los derechos y libertades de terceros y de conformidad con el derecho de la Unión.

En el contexto del RIA el proveedor es cualquier entidad, ya sea persona natural o jurídica, organización e, incluso, autoridad pública, que desarrolla un sistema o modelo de IA de uso general. Esto incluye a quienes ponen el sistema en el mercado, independientemente de si lo hacen con fines lucrativos o de forma gratuita (RIA, 2024, art. 3, num. 3), y sus obligaciones específicas se establecen principalmente en el artículo 16 del RIA. En tanto que los responsables del despliegue son «una persona física o jurídica, o autoridad pública, órgano u organismo que utilice un sistema de IA bajo su propia autoridad, salvo cuando su uso se enmarque en una actividad personal de carácter no profesional» (num. 4), sus principales obligaciones se establecen en el artículo 26 del RIA.

Una de las formas de intervención humana previstas en el artículo 14, numeral 2 del RIA busca integrar a seres humanos para prevenir o reducir al mínimo los riesgos para la salud, la seguridad o los derechos fundamentales en el uso de los sistemas de IA de alto riesgo.

La supervisión debe ser «efectiva» desde el diseño y desarrollo del sistema, lo que significa que debe garantizarse antes de comercializarlo. En lo que se refiere al «cómo hacerlo», el apartado 3 del artículo 14 señala que la intervención humana puede hacerse a través de medidas que «resulten proporcionales a los riesgos», «al nivel de autonomía» y «al contexto de uso del sistema», y que pueden consistir en: a) «medidas que el proveedor defina y que integre, cuando sea técnicamente «viable [...] antes de su introducción en el mercado o su puesta en servicio»; y/o b) aquellas «que el proveedor defina antes de la introducción del sistema de IA de alto riesgo en el mercado o de su puesta en servicio», y «que sean adecuadas para que las ponga en práctica el responsable del despliegue» (art. 14, num. 3)¹³.

En ese sentido, las medidas descritas se complementarán con las «instrucciones de uso» a las que hace referencia el artículo 13, numeral 3, literal d del RIA y que deben contener información sobre las medidas técnicas establecidas para facilitar la interpretación de los resultados de salida de los sistemas de IA de alto riesgo por parte de los responsables del despliegue.

El artículo 26, numeral 2 del RIA establece que los responsables del despliegue deben encomendar la supervisión humana a personas físicas que tengan la competencia, la formación y la autoridad necesarias. Asimismo, de modo similar a lo previsto en el artículo 22 del RGPD, de acuerdo con lo dispuesto en el artículo 26, numeral 11 del RIA, los responsables del despliegue de los sistemas de IA de alto riesgo contenidos en el anexo III, que tomen decisiones o ayuden a tomar

¹³ Sobre el particular, véase Lazcoz (2024, pp. 573-574).

decisiones relacionadas con personas físicas, deberán informar a estas que están expuestas a la utilización de dichos sistemas.

Según el artículo 14, numeral 4 del RIA, las medidas de supervisión humana para sistemas de IA de alto riesgo requieren (según proceda de acuerdo con el contexto y de manera proporcionada) que los supervisores humanos: a) comprendan las capacidades y limitaciones del sistema; b) reconozcan el sesgo de automatización; c) interpreten correctamente los resultados de salida del sistema (así como los métodos y herramientas de interpretación disponibles); y d) que puedan decidir anularlos, descartarlos, revertirlos, invalidarlos, no utilizar el sistema e, incluso, detenerlo si es necesario. Para definir y adoptar tales medidas y que sean proporcionales es fundamental determinar y gestionar los riesgos (Míguez & Torres, 2024, p. 68).

Finalmente, el artículo 14, numeral 5 del RIA dispone que, en el caso de los sistemas de IA de alto riesgo mencionados en el anexo III, punto 1, literal a (esto es, sistemas de identificación biométrica remota, no prohibidos), debe garantizarse también que el responsable del despliegue no actúe ni tome ninguna decisión basándose en la identificación generada por el sistema, a menos que dos personas físicas «con la competencia, formación y autoridad necesarias han verificado y confirmado por separado dicha identificación». Este requisito

no se aplicará a los sistemas de IA de alto riesgo utilizados con fines de garantía del cumplimiento del Derecho, de migración, de control fronterizo o de asilo cuando el Derecho nacional o de la Unión considere que la aplicación de este requisito es desproporcionada.

En cuanto a la relación entre el RIA y el RGPD¹⁴, es importante destacar que el propio texto del RIA aclara en su considerando 10 que esta nueva regulación no busca alterar la aplicación del derecho que regula el tratamiento de datos personales y que se encuentra en vigor, así como tampoco modificar las obligaciones de los proveedores y responsables del despliegue de sistemas de IA en lo que respecta a la protección de datos personales. En ese sentido, los interesados mantendrán todos sus derechos y garantías conforme a dicha legislación, lo que incluye a aquellos relacionados con decisiones individuales completamente automatizadas, como la elaboración de perfiles (a la que se hizo referencia en el epígrafe anterior de este trabajo). Con la nueva normativa se busca entonces que las normas armonizadas para sistemas de IA faciliten el ejercicio efectivo de los derechos de los ciudadanos en materia de protección de datos y otros derechos fundamentales.

¹⁴ Para un análisis más detallado véase a Jiménez (2024), quien reflexiona sobre las interacciones entre el RIA y el RGPD buscando aclarar cómo se complementan, relacionan, superponen o modifican ambos textos.

No obstante, existen múltiples interacciones entre el RGPD y el RIA que no siempre son fáciles de comprender. Una de las más importantes es que, en opinión de algunos comentaristas, el contenido del artículo 14 del RIA y los requisitos con los que deben cumplir los sistemas de IA de alto riesgo (cap. III, sec. 2) harían inaplicable el supuesto de hecho previsto en el artículo 22 del RGPD, en particular en los casos en que la decisión individual automatizada sería posible por incluirse en alguno de los casos de su apartado 2 (Jiménez, 2024, p. 127).

Otros autores, como Lazcoz (2024), entienden que la norma del RIA es de carácter general y establece una obligación para que los proveedores comercialicen sistemas de IA de alto riesgo «que puedan ser supervisados de manera efectiva por personas físicas. No obstante, el tipo de supervisión deberá determinarse en función de la normativa aplicable (por ejemplo, el artículo 22 del RGPD) y a discreción del propio responsable del despliegue» (p. 576). De ahí que resulte clave para entender la supervisión humana en el RIA separar y comprender las obligaciones del proveedor (RIA, 2024, art. 16) y las del responsable del despliegue (art. 26).

En cambio, Radtke (2025) considera que el artículo 22 del RGPD debe interpretarse de manera amplia, especialmente a la luz de la sentencia del TJUE en el asunto *Schufa* (previamente analizada). En ese sentido, abarcaría muchas formas de integración de sistemas de IA con una influencia sustancial en los procesos de toma de decisiones, por lo que, en su opinión, no era necesaria una disposición de este tipo en el RIA. Sin embargo, aquella fue incluida y también el artículo 86, que establece el derecho a una explicación de las decisiones tomadas individualmente (p. 95), aunque en su opinión no está claro que este pueda cumplirse en la práctica, ya que conlleva sus propios desafíos. Así «supervisar eficazmente» y, por tanto, «explicar», implica comprender una decisión basada en un mecanismo que es completamente diferente y difícil de entender para los humanos.

Aunque la aprobación del RIA es reciente y su entrada en vigor es progresiva en la literatura, ya existen opiniones tanto a favor como en contra de la intervención humana en el contexto de los procesos de toma de decisiones automatizadas relacionadas con sistemas de IA (Lazcoz, 2024, pp. 572-573). Por ejemplo, autores como Laux (2024) opinan que los seres humanos pueden carecer de competencia o recibir incentivos perjudiciales, y que esto puede dificultar la eficacia de la supervisión humana (p. 2854). Asimismo, existe evidencia de que a los operadores humanos les resulta difícil juzgar la calidad de los resultados algorítmicos de forma fiable y, por tanto, incluso si un humano decide anular un resultado algorítmico, no habrá garantías de que con ello mejore la decisión (Beck & Burri, 2023, p. 16).

Todo ello plantea la duda sobre si este tipo de supervisión tendrá realmente el efecto que se espera; es decir, si podrá cumplir con la finalidad para la que ha sido prevista (Lazcoz, 2024, pp. 582-583; Radtke, 2025, pp. 91-109). En efecto, en el estudio realizado por Gaudel *et al.* (2024) se concluyó que la supervisión humana puede no evitar resultados discriminatorios, no solo por posibles sesgos de automatización, sino también por casos de aversión a los algoritmos. En ese sentido, estiman que los esfuerzos para prevenir los resultados discriminatorios deben centrarse tanto en programar sistemas que respeten normas de imparcialidad como en establecer sistemas de supervisión que garanticen que los usuarios no introducen nuevos sesgos en los resultados de la recomendación que arroja la IA (p. 1073).

Garantizar que exista un *human in the loop* puede ser insuficiente para prevenir o mitigar las afectaciones a derechos fundamentales o los daños producidos por el uso de sistemas de IA, así como los perjuicios. Por ello, la regulación de estos sistemas de IA debe complementarse con otros tipos de medidas; por ejemplo, con normas de responsabilidad para hacer frente a aquellas situaciones complejas que devengan de las particularidades de las interacciones humano-máquina (Botero Arcila, 2024), además de con el derecho a una explicación.

III.2. El artículo 8 del Proyecto de Ley que Regula los Sistemas de Inteligencia Artificial (Boletín N.º 16821-19, refundido con: 15869-19)

En Chile no existe una ley general aplicable a los sistemas de IA; no obstante, a la fecha de elaboración de este trabajo se encuentra en discusión (en primer trámite constitucional) el Proyecto de Ley que Regula los Sistemas de Inteligencia Artificial (Boletín 16821-19, 7 de mayo de 2024)¹⁵. En el proyecto la supervisión humana se incluye entre los principios generales aplicables a los sistemas de IA que deberán ser observados por todos los operadores¹⁶ que entren en el ámbito de aplicación de la ley. En este orden de ideas, el artículo 4, literal a señala que los sistemas de IA «se desarrollarán y utilizarán como una herramienta al servicio del ser humano, que respete la dignidad humana y la autonomía personal, y que funcione de manera que pueda ser controlada y vigilada adecuadamente por seres humanos».

¹⁵ El análisis de este trabajo se realizó teniendo en cuenta el texto original del proyecto de ley, presentado el 7 de mayo de 2024 y que pudo haber sido objeto de modificaciones para la fecha de publicación final de este trabajo. La tramitación de este proyecto puede seguirse en el sitio web de la Cámara de Diputadas y Diputados a través del siguiente enlace: <https://www.camara.cl/legislacion/ProyectosDeLey/tramitacion.aspx?prmID=17429&prmBOLETIN=16821>

¹⁶ En el proyecto de ley se entenderá por «operador» al proveedor, el implementador, el representante autorizado, el importador y el distribuidor (Boletín 16821-19, 7 de mayo de 2024, art. 3, num. 10).

El proyecto de ley adopta un enfoque basado en riesgos similar al del RIA, aunque incorpora sus propias definiciones y mecanismos de control. Los niveles de riesgo se definen en el artículo 5, que distingue entre usos de sistemas de IA: a) de riesgo inaceptable o prohibido, esto es, aquellos considerados incompatibles con el respeto y garantía de los derechos fundamentales de las personas; b) de alto riesgo, que son aquellos sistemas autónomos o componentes de seguridad de productos que pueden tener un impacto negativo en la salud y seguridad de las personas, sus derechos fundamentales o el medio ambiente, así como en los derechos de los consumidores, especialmente si fallan o se usan de manera incorrecta; c) de riesgo limitado, los que se refieren a sistemas que presentan riesgos no significativos de manipulación, engaño o error en su interacción con personas físicas; y d) sin riesgo evidente, categoría en la cual se incluyen todos los sistemas de IA que no se encuadran en las anteriores.

El título III se refiere a los sistemas de IA de alto riesgo que son definidos en el artículo 7. Aunque el proyecto chileno distingue categorías similares a las del RIA, no proporciona el mismo nivel de detalle. Esto porque, conforme se establece en el artículo 15, literal a, se deja a un reglamento dictado por el Ministerio de Ciencia, Tecnología, Conocimiento e Innovación la creación de la lista específica de los sistemas de alto riesgo. Dicha lista se elaborará previa propuesta efectuada por el Consejo Asesor Técnico de IA (órgano de carácter consultivo y permanente cuya creación se propone en el artículo 14 de la propuesta).

Dichos sistemas, a su vez, estarán permitidos siempre que cumplan con una serie de reglas obligatorias previstas en el artículo 8. Estas implican contar con un sistema de gestión de riesgos, gobernanza de datos, documentación técnica, registros, mecanismos de transparencia y mecanismos de supervisión humana, entre otros.

El literal f del artículo 8 del proyecto de ley se refiere a algunos mecanismos para concretar la supervisión humana. En particular, dispone que los sistemas de IA de alto riesgo deben diseñarse y desarrollarse asegurando la supervisión de personas físicas «técnicamente capacitadas para esta función», considerando «el escenario de implementación en cuestión y de forma proporcionada a los riesgos asociados». Esto se da

con el objeto de prevenir o reducir al mínimo los riesgos para la salud, la seguridad, los derechos fundamentales, la democracia, y/o el medio ambiente, que puedan surgir cuando un sistema de IA de alto riesgo se utilice conforme a su finalidad prevista o cuando se le dé un uso indebido razonablemente previsible.

IV.1. Los artículos 10 y 11 de la Directiva sobre el Trabajo en Plataformas (DTP)

La Directiva de la UE 2024/2831 de 2024, conocida como Directiva sobre el Trabajo en Plataformas (DTP), busca mejorar las condiciones laborales en plataformas digitales. Establece una presunción de empleo para combatir la clasificación errónea de trabajadores, regula sistemas automatizados para garantizar transparencia y equidad, refuerza los derechos laborales colectivos, y asegura la aplicación efectiva mediante sanciones y otras herramientas procedimentales (por ejemplo, una regla de inversión de la carga de la prueba). En conjunto, la DTP establece un marco normativo para abordar los desafíos de la economía digital y promover un trabajo sostenible. Para diversos comentaristas, uno de los principales objetivos de la DTP es abordar las limitaciones del RGPD en la protección de los derechos en el ámbito laboral (Abraham, 2022, p. 187; Todolí, 2024, p. 42). En efecto, como analizaremos en esta sección, muchas de las dudas o cuestionamientos que derivan de la interpretación y aplicación práctica de las disposiciones del RGPD fueron consideradas como valiosos insumos para la discusión de la DTP (Aloisi *et al.*, 2023, p. 32).

Al otorgar derechos de gestión algorítmica y derechos colectivos a los trabajadores de plataformas que son genuinamente autónomos, la directiva amplió significativamente el ámbito personal de aplicación de los derechos laborales (art. 2, num. 2). En este sentido, la directiva ha sido considerada «como uno de los primeros pasos hacia el rediseño de los paradigmas normativos que rigen el derecho laboral» (Rainone & Aloisi, 2024, p. 1). Por otra parte, al establecer un marco integral para la gestión algorítmica y la gobernanza de datos, tanto a nivel individual como colectivo, esta directiva ha sido considerada como el primer paso hacia una regulación general del uso de tecnologías digitales en el ámbito laboral, más allá del trabajo en plataformas (Ponce del Castillo & Naranjo, 2022, p. 2).

Los artículos 10 y 11 de la DTP establecen un marco normativo para proteger los derechos de los trabajadores frente a decisiones automatizadas. El artículo 10 exige que las plataformas evalúen periódicamente el impacto de los sistemas automatizados con participación de los representantes de los trabajadores, implementen supervisión independiente con autoridad para anular decisiones y suspendan sistemas con alto riesgo de discriminación. Las evaluaciones deben compartirse con trabajadores y autoridades, promoviendo transparencia y responsabilidad. Además, cualquier decisión que afecte severamente la relación laboral debe ser tomada por un humano. En efecto, antes que un enfoque basado en el

ejercicio voluntario de derechos por parte de sus titulares, la DTP parte de la base de obligaciones o prohibiciones estrictas que aspiran a mejorar la condición laboral de este tipo de trabajadores.

A su vez, el artículo 11, denominado «Revisión por humanos», amplía garantías frente a decisiones automatizadas y no solo ante aquellas con impacto jurídico significativo. Reconoce el derecho a explicaciones claras, con acceso a personal cualificado y revisiones formales, abordando las lagunas del RGPD (art. 22). A diferencia del RGPD, la DTP incluye decisiones respaldadas por algoritmos, incluso si no son completamente automatizadas, y refuerza el derecho a una explicación detallada *ex post*. En este sentido, se resuelven las dudas que cuestionan la existencia de este derecho en el RGPD. El artículo 22 del RGPD, que garantiza el derecho a la intervención humana en la toma de decisiones automatizada, está enfocado en «decisiones basadas únicamente» en el procesamiento automatizado. Esta redacción, como ha mostrado tanto la doctrina como cierta práctica administrativa, deja un vacío legal para las decisiones que eran apoyadas, pero no totalmente determinadas, por sistemas automatizados (Viollier & Fischer, 2022, pp. 151-169). La DTP soluciona este problema al abarcar todos los tipos de decisiones que son apoyadas o tomadas por sistemas automatizados de toma de decisiones (ADMS, por sus siglas en inglés), incluso cuando no afectan de forma significativa a las personas que realizan trabajo en plataformas (Rainone & Aloisi, 2024, pp. 5-6). Además, y considerando lo avanzado en el RGPD, la DTP otorga la posibilidad de solicitar revisiones formales de las decisiones y obliga a las plataformas a responder por escrito en un plazo de dos semanas. La redacción de los artículos 10 y 11 dejan en claro que el derecho a una explicación no se reduce a una explicación *ex ante* o de carácter abstracto; en efecto, lo que se busca es que las personas puedan comprender cómo se adoptaron las decisiones en el caso concreto, dejando fuera de toda duda la existencia de un derecho a la explicabilidad *ex post* y en concreto.

Pese a sus avances, la DTP enfrenta críticas. Su enfoque colectivo no es plenamente inclusivo, ya que algunos derechos están limitados a quienes tienen un contrato laboral formal. Por ejemplo, el artículo 15 estipula que solo los proveedores con la condición de trabajador tienen derecho a recibir asistencia de representantes en la supervisión del impacto de los sistemas de gestión algorítmica en las condiciones laborales (art. 10, num. 1). Esto da lugar a un marco fragmentado que afecta la coherencia del panorama normativo y genera el riesgo de duplicaciones onerosas e ineficaces en los sistemas de representación. Además, la falta de un mecanismo claro para que los trabajadores impugnen decisiones reduce su efectividad, aunque introduce mejoras respecto al RGPD, como prohibiciones específicas de procesamiento de datos y mayor transparencia.

INTERVENCIÓN,
SUPERVISIÓN Y
CONTROL HUMANO
EN LAS DECISIONES
AUTOMATIZADAS

HUMAN
INTERVENTION,
OVERSIGHT
AND CONTROL
IN AUTOMATED
DECISION-MAKING

IV.2. El artículo 152 quinquies D del Código del Trabajo

La Ley N.º 21.431, vigente desde el 11 de marzo de 2022, que incorporó al Código del Trabajo chileno un capítulo especial sobre la regulación del trabajo mediante plataformas digitales de servicios, incluyó una cláusula específica sobre la transparencia y el derecho a la información que es aplicable tanto a los trabajadores dependientes como independientes. El artículo 152 quinquies D establece derechos específicos para trabajadores de plataformas digitales en relación con sus datos personales, así como obligaciones de transparencia para las empresas. Tal como señala Contreras (2022):

El fundamento constitucional de estas normas radica en el derecho de acceso a la información que es parte del contenido de la autodeterminación informativa y en el derecho a la información que es parte del derecho a la negociación colectiva (p. 167).

En primer lugar, el artículo señalado parte afirmando que los datos de los trabajadores son estrictamente confidenciales y que solo pueden ser utilizados en el contexto de los servicios prestados por la empresa. Además, esta cláusula reconoce dos derechos específicos que constituyeron sendas innovaciones, sobre todo considerando la vigencia de una reforma a la legislación laboral que es anterior a la modificación de la normativa sobre protección de datos. En primer lugar, los trabajadores pueden solicitar en cualquier momento acceso a sus datos personales, especialmente aquellos relacionados con su calificación y desempeño laboral, cuestión que podría modificar la interpretación que actualmente existe del poder de disciplina y mando por parte del empleador. La plataforma debe responder dentro de un plazo de quince días hábiles. En segundo lugar, se incorporó por primera vez en el derecho chileno una norma que reconoce el derecho a la portabilidad, que permite a los trabajadores solicitar sus datos en un formato estructurado y común, compatible con diferentes sistemas informáticos. Por último, la normativa incorporó una obligación de las empresas de permitir el acceso de las autoridades a la programación del algoritmo, explicaciones completas sobre su funcionamiento, los datos de entrenamiento y otros factores relevantes para garantizar el cumplimiento de la ley.

La norma omite de manera explícita el derecho a solicitar una intervención humana en decisiones tomadas por algoritmos, lo que genera interrogantes sobre su alcance real en la protección de los trabajadores frente a sistemas automatizados. El acceso al funcionamiento del algoritmo y sus datos de entrenamiento permiten una mayor fiscalización por parte de las autoridades; sin embargo, esto está dirigido más a garantizar el cumplimiento legal que a ofrecer un recurso práctico a los trabajadores en situaciones individuales. En cuanto a los derechos de acceso y portabilidad de datos, estos permiten a los trabajadores

revisar y transferir información sobre su desempeño, lo que podría ser un paso preliminar hacia impugnar decisiones automatizadas, pero no establece un mecanismo directo para que intervenga una persona. Además, y aunque las autoridades pueden acceder a la programación y el funcionamiento del algoritmo, esto no asegura una protección inmediata y práctica para los trabajadores.

Tal como se sostiene en un reciente informe preliminar de cumplimiento de la Ley N.º 21.431, existen muchos desafíos en relación con la transparencia y el acceso a la información, sobre todo considerando que casi la totalidad de los trabajadores están registrados bajo la modalidad de trabajadores independientes (Consejo Superior Laboral, 2024, p. 45). Además, si bien las aplicaciones «parecen cumplir las disposiciones que obligan a transparentar los criterios para determinar los ingresos a los que accederán las personas que trabajan en ellas» (p. 44), el gran problema se produce en la garantía de un derecho a la explicabilidad, en concreto, con respecto a las remuneraciones o ingresos. El informe del Consejo Superior Laboral, la entidad encargada de monitorear el cumplimiento de la normativa, sostuvo que

la gestión algorítmica del trabajo en plataformas y en particular, la complejidad para establecer las tarifas (tarifas dinámicas, en función de las evaluaciones, algunas en función de los pesos de los paquetes transportados, etc.) introduce una dificultad manifiesta para que las personas trabajadoras tengan claridad de la remuneración a la cual pueden acceder prestando sus servicios, incluso si todas las reglas están claras y son explicitadas en las plataformas la investigación cualitativa muestra que no resulta posible para los trabajadores reproducir los resultados de los cálculos de la remuneración (p. 44).

En otras palabras, si bien parece existir claridad acerca del cumplimiento de un derecho a la explicación en abstracto, no existen formas de analizar o evaluar el cumplimiento de la explicabilidad en concreto. Considerando esta dificultad, la ausencia de una norma que permita la intervención humana ante el resultado de mecanismos automatizados de decisión parece convertirse en una barrera particularmente difícil de superar, sobre todo cuando se contrasta con la reciente legislación europea en la materia.

Además, como destaca el mismo informe, los trabajadores señalan que la nueva normativa no ha mejorado el escaso conocimiento que tienen los trabajadores para ejercer sus derechos de reclamo ante problemas o contratiempos que sufran. El derecho a tener una «interfaz humana» ha sido considerado como uno de los más problemáticos en cuanto a su cumplimiento:

INTERVENCIÓN,
SUPERVISIÓN Y
CONTROL HUMANO
EN LAS DECISIONES
AUTOMATIZADAS

HUMAN
INTERVENTION,
OVERSIGHT
AND CONTROL
IN AUTOMATED
DECISION-MAKING

En todas las regiones donde se han establecido organizaciones, se ha levantado información sobre la presencia de las diferentes plataformas tanto de reparto como de transporte de pasajeros. En todas ellas no existen oficinas, ni espacios, ni representantes de las plataformas que puedan atender casos locales como desvinculaciones, problemas relacionados a accidentes laborales, reclamos por pérdida de ranking, incluso problemas relacionados al ejercicio del trabajo como por ejemplo el no pago de productos de clientes en los casos que las transacciones se realizan en efectivo entre repartidores - comercios - clientes (p. 50).

Como lo muestra el informe, el problema parece ir más allá de la automatización y de su impacto, y se relaciona con cuestiones básicas que las aplicaciones o plataformas digitales de trabajo aún no son capaces de cumplir.

Por otra parte, la ausencia de un derecho a solicitar la revisión, la intervención o el control humano de una decisión no puede ser suplido por el limitado rol de la autoridad pertinente y el ejercicio de potestades fiscalizadoras y sancionatorias. El mismo informe parte de la base de que el órgano fiscalizador, en este caso la Dirección del Trabajo, no tiene la capacidad de acceder a la información señalada en inciso final del artículo 152 quinquies D. Tal como lo reconoce el Consejo Superior Laboral, no existe información disponible acerca de casos en que funcionarios fiscalizadores hayan tenido «acceso a la programación del algoritmo de la plataforma digital de servicios» o que hayan podido evaluar el cumplimiento de la obligaciones de «entregar explicaciones completas y suficientes sobre la forma en que toma el algoritmo sus decisiones, los datos con los que fue entrenado y todo otro factor relevante para el cumplimiento íntegro de la ley» (p. 47).

Las potenciales lecciones de la implementación de la normativa europea nos darán algunas luces acerca del impacto de este nuevo derecho a solicitar una revisión por seres humanos en casos de decisiones automatizadas en el ámbito laboral. Lo que está claro es que los procesos de revisión no solo deben ser procedimentales, sino que deben incluir una evaluación sustantiva de la decisión algorítmica. Esto implica que los revisores humanos deben tener la capacidad de reinterpretar la decisión en términos humanos comprensibles y asegurarse de que esté respaldada por evidencia. Además, tal como se ha avanzado en diversos estudios sobre transparencia algorítmica, es imperativo distinguir las audiencias: para garantizar el derecho a una revisión efectiva, es necesario separar la lógica del algoritmo de la base empírica en la que se basa la decisión. Esto evita la circularidad en la que la decisión se considera aceptable simplemente porque se basa en el algoritmo, y exige a los empleadores crear las condiciones para que la revisión sea comprensible ante audiencias no expertas. Además, y quizás en un ámbito aún no discutido aquí, los

revisores humanos deben tener en cuenta la situación de trabajadores en situaciones fácticas similares para garantizar la consistencia y evitar la arbitrariedad en las decisiones, cuestión que está en el centro de la relación entre el derecho a la información y el derecho a la igualdad y no discriminación algorítmica. En efecto, los revisores humanos pueden tener en mente la dimensión normativa que ilustran las categorías protegidas por el derecho antidiscriminación y así evitar correlaciones espurias que pudieran quedar fuera del ámbito de protección (Wachter, 2022, p. 150).

Por último, es importante reconocer que la supervisión humana no es una solución infalible. Si el derecho a una revisión humana busca superar los problemas derivados del sesgo de automatización —esto es, que los seres humanos tendemos a creer en los resultados de procesos automatizados de decisión—, es imperativo no caer en un sesgo equivalente de evaluación. En efecto, los revisores humanos pueden estar, a su vez, influenciados por sesgos o presiones.

215

INTERVENCIÓN,
SUPERVISIÓN Y
CONTROL HUMANO
EN LAS DECISIONES
AUTOMATIZADAS

HUMAN
INTERVENTION,
OVERSIGHT
AND CONTROL
IN AUTOMATED
DECISION-MAKING

V. RECAPITULACIÓN: ¿DE QUÉ HABLAMOS CUANDO HABLAMOS DE INTERVENCIÓN, SUPERVISIÓN O CONTROL HUMANO EN EL CONTEXTO DE LAS DECISIONES AUTOMATIZADAS?

Las normas analizadas en los epígrafes anteriores ofrecen mecanismos para concretar la intervención humana a través de medidas de supervisión y control de forma sectorial en los contextos de protección de datos personales, regulación de la IA y trabajo en plataformas, con los alcances que se resumen en las tablas N.º 1 y N.º 2.

Tabla N.º 1. Visión de conjunto sobre normas de supervisión y control humano (seleccionadas) en la Unión Europea

Nombre del instrumento	Reglamento General de Protección de Datos de la Unión Europea (RGPD)	Reglamento de IA (RIA)	Directiva sobre el Trabajo en Plataformas (DTP)
Principales artículos sobre supervisión/control humano	Artículo 22	Artículos 14, 16 y 26	Artículos 10 y 11
Responsable	Responsable del tratamiento de datos personales	Proveedores/responsables del despliegue	Plataformas digitales de trabajo

Tipo/forma de la supervisión	Obligaciones requeridas para habilitar la toma de decisiones basada en el tratamiento automatizado de datos personales, así como el deber de adoptar las medidas necesarias para asegurar los derechos y libertades del titular. Adicionalmente, se entregan derechos al titular: derecho a la intervención humana, a expresar su punto de vista y a solicitar la revisión de la decisión	«Supervisión efectiva, por todo el ciclo de vida»: para prevenir o reducir riesgos/desde el diseño/conforme a instrucciones de uso. Adicionalmente, se entregan vías de recurso a los particulares (artículos 85, 86 y 87)	Obligación de evaluación periódica del impacto de los sistemas automatizados en las condiciones laborales, con supervisión humana calificada y autoridad para anular decisiones (las decisiones críticas deben ser tomadas por un humano). Además, reconoce el derecho a explicaciones claras sobre decisiones automatizadas y permite su revisión. Si se vulneran derechos, las decisiones deben ser corregidas o compensadas
Momento de la supervisión (ex ante/ex post)	Ex post	Ex ante/ex post	Ex ante/ex post

Fuente: elaboración propia.

Tabla N.º 2. Visión de conjunto sobre normas de supervisión y control humano (seleccionadas) en Chile

Nombre del instrumento	Ley N.º 21.719, que regula la protección y el tratamiento de los datos personales y crea la Agencia de Protección de Datos Personales	Proyecto de Ley que Regula los Sistemas de Inteligencia Artificial (no en vigor)	Ley 21.431 (modificación al Código del Trabajo chileno)
Principales artículos sobre supervisión/control humano	Artículo 8 bis	Artículos 4 y 8	Artículo 152 quinquies D
Responsable	Responsable del tratamiento de datos personales	Operadores (proveedor, implementador, representante autorizado, importador y distribuidor)	Empresas de plataformas digitales de servicios

Tipo/forma de la supervisión	Obligaciones requeridas para habilitar la toma de decisiones basada en el tratamiento automatizado de datos personales, así como el deber de adoptar las medidas necesarias para asegurar los derechos y libertades del titular. Adicionalmente, se entregan derechos al titular: derecho a una explicación, a la intervención humana, a expresar su punto de vista y a solicitar la revisión de la decisión	Desde el diseño y desarrollo del sistema hasta su implementación	No menciona explícitamente la obligación de supervisión (garantía de acceso de las autoridades competentes al programa) ni tampoco explicaciones completas sobre decisiones, datos de entrenamiento y factores relevantes para la fiscalización legal
Momento de la supervisión (ex ante/ex post)	Ex post	Ex ante/ex post (indeterminadas)	Ex post

Fuente: elaboración propia.

Desde el punto de vista de la regulación de protección de datos personales, tanto el artículo 8 bis de la nueva Ley N.º 21.719 como su norma inspiradora (el artículo 22 del RGPD) incorporan criterios de control y supervisión humana en el contexto de la toma de decisiones automatizadas. Estas salvaguardas solo son aplicables en la medida que: a) la decisión haya sido tomada con base en el tratamiento automatizado de los datos personales del titular y b) dicha decisión haya afectado significativamente al titular.

Cumpliendo estas dos condiciones, ambas regulaciones impiden este tipo de toma de decisiones automatizadas, a menos que se cumplan ciertos requisitos habilitadores (en general, que el tratamiento se realice con base en el consentimiento del titular, que la decisión sea necesaria para la ejecución de un contrato o que así lo señale la ley). Incluso en dichos casos, es necesario establecer salvaguardas para los derechos y libertades del titular, y adicionalmente se le otorgan ciertos derechos al titular que este puede ejercer. Entre ellos, los más vinculados a la supervisión y el control humano de los algoritmos y sistemas de inteligencia artificial son: a) el derecho a obtener una explicación (solo mencionado explícitamente en la ley chilena), b) a la intervención humana, c) a expresar su punto de vista y d) a solicitar la revisión de la decisión.

Es importante tener en consideración que ambos artículos no regulan el tratamiento de datos personales realizado por sistemas algorítmicos

y automatizados (el que debe cumplir con todos los requisitos de la legislación de datos personales), sino que regulan en qué condiciones se pueden tomar decisiones basadas en el tratamiento automatizado de los datos personales del titular en aquellos casos en que dicha decisión lo afecte significativamente. Esto implica que cualquier decisión que no cumpla con el nivel de automatización o el nivel de afectación exigido por la regulación de datos personales (ambos criterios son distintos en la regulación chilena y europea) puede ser realizada libremente sin cumplir con los estándares de supervisión y control humano contenidos en los artículos 8 bis de la nueva LPD chilena o el artículo 22 del RGPD de la Unión Europea.

En el caso de la regulación general de la IA, tanto el RIA como el proyecto de ley chileno enfatizan la importancia de la supervisión humana, especialmente en el caso de los sistemas de alto riesgo, buscando proteger la salud, la seguridad, los derechos fundamentales y el medio ambiente de los riesgos asociados con la IA. El artículo 14 del RIA establece la supervisión humana como un requisito obligatorio para los sistemas de alto riesgo, y el artículo 8 del proyecto de ley chileno también incluye la supervisión humana como un principio general y un requisito para los sistemas de alto riesgo. El artículo 14 del RIA busca imponer una serie de evaluaciones, obligaciones y prohibiciones aplicables, la mayoría de ellas, por los proveedores *ex ante* (antes de su comercialización) para evitar que se produzcan los eventuales daños o afectaciones a los derechos fundamentales (Botero Arcila, 2024). No obstante, también supone obligaciones y medidas de supervisión humana para los responsables del despliegue e implementación de los sistemas aplicables tanto antes como después de la toma de decisiones, en este caso para corregirlas (Lazcoz, 2024, p. 576). En ese sentido, el RIA es más específico que el proyecto chileno en cuanto a cómo se debe implementar la supervisión humana en los sistemas de IA de alto riesgo.

En el ámbito laboral, los artículos 10 y 11 de la DTP de la UE establecen un marco normativo para proteger los derechos de los trabajadores ante decisiones automatizadas, abordando limitaciones del RGPD en el ámbito laboral. Además, este instrumento ha sido considerado como uno de los más detallados y específicos a la hora de abordar el estándar de la intervención humana. El artículo 10 impone a las plataformas la obligación de evaluar regularmente el impacto de los sistemas automatizados, garantizar la supervisión independiente y permitir la intervención humana en decisiones que afecten significativamente la relación laboral. Por su parte, el artículo 11 refuerza garantías al derecho a revisiones humanas y explicaciones claras, incluso para decisiones respaldadas por algoritmos, pero no totalmente automatizadas. Este marco enfatiza la explicabilidad concreta y el acceso a revisiones formales, ofreciendo mayor transparencia que el RGPD. Tal como

señalamos más arriba, antes que un enfoque basado en el ejercicio voluntario de derechos por parte de sus titulares, la DTP parte de la base de obligaciones o prohibiciones estrictas que aspiran a mejorar la condición laboral de este tipo de trabajadores.

En Chile, por su parte, el artículo 152 quinquies D del Código del Trabajo reconoce derechos de transparencia y acceso a datos personales para trabajadores de plataformas digitales de trabajo, como el acceso a sus calificaciones y el derecho a la portabilidad de datos. Sin embargo, omite explícitamente el derecho a solicitar intervención humana en decisiones algorítmicas, limitando su capacidad para abordar el sesgo de automatización. Aunque se requiere que las plataformas entreguen explicaciones completas y permitan acceso a la programación algorítmica, estas medidas están diseñadas para permitir una fiscalización administrativa antes que para la protección inmediata de los trabajadores. De todos modos, como se ha señalado por diversos análisis preliminares, el problema sigue siendo la capacitación y el entrenamiento de los funcionarios públicos que podrán revisar o cuestionar el proceso o el producto del tratamiento automatizado de datos personales en el ámbito laboral.

VI. CONCLUSIONES

El objeto principal de este trabajo se centró en describir y analizar diferentes intentos regulatorios que consagran la supervisión, el control o la evaluación humana en los SDA para realizar una reconstrucción o un análisis crítico sobre sus premisas o supuestos principales. Las conclusiones se centran en la forma en la que se implementa la intervención humana en las decisiones automatizadas, destacando tanto los avances como los desafíos persistentes en los ámbitos seleccionados para el estudio.

Los intentos regulatorios examinados buscan ir más allá de una vaga referencia a la presencia humana, garantizando —*mutatis mutandis*— una supervisión efectiva con la autoridad y capacidad necesarias para evaluar, corregir y, si es necesario, anular las decisiones tomadas o sugeridas por algoritmos.

En lo que se refiere a la protección de datos personales, tanto el artículo 8 bis de la Ley N.º 21.719 de Chile como el artículo 22 del RGPD de la Unión Europea incorporan criterios de control y supervisión humana en la toma de decisiones automatizadas, ofreciendo salvaguardas que se aplican cuando la decisión se basa en el tratamiento automatizado de datos personales y afecta significativamente al titular. En ese contexto, se otorgan derechos al titular, incluyendo el derecho a obtener una explicación (más explícito en la ley chilena), a la intervención humana,

219

INTERVENCIÓN,
SUPERVISIÓN Y
CONTROL HUMANO
EN LAS DECISIONES
AUTOMATIZADASHUMAN
INTERVENTION,
OVERSIGHT
AND CONTROL
IN AUTOMATED
DECISION-MAKING

a expresar su punto de vista y a solicitar la revisión de la decisión. En la UE, el RGPD materializó uno de los principales impulsos regulatorios en materia de supervisión y control humano en el contexto de las decisiones automatizadas.

No obstante, las particularidades de los SDA han hecho necesarias regulaciones que van más allá del ámbito de la protección de datos personales. Con todo, el desarrollo e implementación de los sistemas de IA implicarán el tratamiento de datos personales en muchos casos y esta preocupación ha sido recogida por diversos apartados del RIA y la DTP.

Tanto el RIA de la Unión Europea como el proyecto de ley chileno sobre IA clasifican los sistemas según su riesgo y enfatizan la importancia de la supervisión humana en los sistemas de alto riesgo. El artículo 14 del RIA establece la supervisión humana como un requisito obligatorio para ellos, pero se complementa con otros tipos de obligaciones a proveedores y responsables del despliegue que se concretan tanto *ex ante* como *ex post*. Aunque contiene medidas más específicas y detalladas que la propuesta chilena, no está exenta de desafíos.

En el ámbito laboral, los artículos 10 y 11 de la DTP de la UE buscan proteger los derechos de los trabajadores frente a decisiones automatizadas, abordando limitaciones del RGPD. La DTP obliga a las plataformas a evaluar el impacto de los sistemas automatizados, así como garantizar la supervisión independiente y la intervención humana en decisiones que afecten la relación laboral. Asimismo, refuerza el derecho a revisiones humanas y explicaciones claras, incluso en decisiones apoyadas por algoritmos, estableciendo un marco más detallado y específico que el RGPD. Por su parte, el artículo 152 quinquies D del Código del Trabajo chileno reconoce los derechos de transparencia y acceso a datos personales, pero omite el derecho a la intervención humana en decisiones algorítmicas, limitando la protección de los trabajadores frente al sesgo de automatización.

A pesar de estos intentos y de algunos avances, los desafíos persisten porque la supervisión humana no es una solución infalible, ya que los revisores humanos pueden estar influenciados por sesgos u otros tipos de limitaciones, como diferentes concepciones de justicia. A ello se suma la necesidad de implementar medidas adicionales (por ejemplo, sobre responsabilidad) para que la intervención humana pueda efectivamente coadyuvar o garantizar decisiones justas y transparentes en los sistemas automatizados.

En definitiva, aunque los diferentes esfuerzos regulatorios analizados aquí constituyen innovaciones para asegurar un «derecho a la intervención humana» en la toma de decisiones automatizadas, su efectividad práctica sigue siendo incierta. Si bien aún tenemos que analizar y evaluar el modo

en que se interpretan e implementan las normativas pertinentes, el riesgo es que la supervisión humana se convierta en un ritual simbólico más que en un verdadero contrapeso al poder de los sistemas algorítmicos, reproduciendo la lógica de que basta con señalar la presencia de un humano para validar o legitimar el proceso. Si las obligaciones o salvaguardas no van acompañadas de criterios claros de competencia, autoridad y recursos para quienes ejercen esa supervisión, la promesa regulatoria corre el peligro de quedar reducida a un remedio superficial. Más aún, la proliferación de regímenes sectoriales fragmentados puede reforzar la ilusión de control sin enfrentar el principal dilema sustantivo: ¿qué procesos o decisiones deben permanecer en manos humanas por razones de dignidad, legitimidad democrática o justicia procedimental, más allá de la eficiencia técnica? Las soluciones regulatorias actuales constituyen un paso necesario, pero insuficiente. El verdadero desafío es transformar la intervención humana de un requisito formal en una garantía sustantiva de agencia, responsabilidad y rendición de cuentas en la era algorítmica.

REFERENCIAS

Abraha, H. (2023). Regulating algorithmic employment decisions through data protection law. *European Labour Law Journal*, 14(2), 172-191. <https://doi.org/10.1177/20319525231167317>

Adams-Prassl, J., Abraha, H., Kelly-Lyth, A., Silberman, M. 'Six', & Rakshita, S. (2023). Regulating algorithmic management: A blueprint. *European Labour Law Journal*, 14(2), 124-151. <https://doi.org/10.1177/20319525231167299>

Acemoglu, D., & Johnson, S. (2023). *Power and progress: Our thousand-year struggle over technology and prosperity*. Hachette UK.

Agencia Española de Protección de Datos. (2024). *Evaluación de la intervención humana en las decisiones automatizadas*. <https://www.aepd.es/prensa-y-comunicacion/blog/evaluacion-de-la-intervencion-humana-en-las-decisiones-automatizadas>

Agudo, U., Liberal, K. G., Arrese, M., & Matute, H. (2024). The impact of AI errors in a human-in-the-loop process. *Cognitive Research*, 9. <https://doi.org/10.1186/s41235-023-00529-3>

Aloisi, A., Rainone, S., & Countouris, N. (2023). An unfinished task?: Matching the Platform Work Directive with the EU and international. *ILO Working paper*, 101. <https://doi.org/10.54394/ZSAX6857>

Ashwini, K. P. (2024). *Informe de la Relatora Especial sobre las formas contemporáneas de racismo, discriminación racial, xenofobia y formas conexas de intolerancia*. Asamblea de las Naciones Unidas. <https://documents.un.org/doc/undoc/gen/g24/084/23/pdf/g2408423.pdf>

Atria, F. (2016). *La forma del derecho*. Madrid: Marcial Pons, Ediciones Jurídicas y Sociales.

Barona Vilar, S. (2023). Dataización de la justicia (Algoritmos, Inteligencia Artificial y Justicia, ¿el comienzo de una gran amistad?). *Revista Boliviana de Derecho*, (36), 14-45. <https://ssrn.com/abstract=4582084>

Barrio Andrés, M. (2024). Capítulo I. Objeto, ámbito de aplicación y sentido del Reglamento Europeo de Inteligencia Artificial. En Moisés Barrio Andrés (dir.), *El Reglamento Europeo de Inteligencia Artificial* (pp. 21-45). Valencia: Tirant lo Blanch.

Beck, J., & Burri, T. (2024). From 'Human Control' in International Law to 'Human Oversight' in the New EU Act on Artificial Intelligence. En Julio Mecacci, Daniele Amoroso, Luciano Cavalcante Siebert, David Abbink, Jeroen van den Hoven y Filippo Santoni de Sio (eds.), *Research Handbook on Meaningful Human Control of Artificial Intelligence Systems* (pp. 105-132). <http://dx.doi.org/10.2139/ssrn.4236554>

Botero Arcila, B. (2024). AI liability in Europe: How does it complement risk regulation and deal with the problem of human oversight? *Computer Law & Security Review*, 54. <https://doi.org/10.1016/j.clsr.2024.106012>

Bradford, A. (2019). *The Brussels Effect: How the European Union Rules the World*. Nueva York: Oxford Academic. <https://doi.org/10.1093/oso/9780190088583.001.0001>

Breidbach, C. (2024). Responsible algorithmic decision-making. *Organizational Dynamics*, 53(2). <https://doi.org/10.1016/j.orgdyn.2024.101031>

Citron, D. K., & Pasquale, F. (2014). *The scored society: Due process for automated predictions*. *Washington Law Review*, 89(1). <https://digitalcommons.law.uw.edu/wlr/vol89/iss1/2>

Coddou Mc Manus, A., & Padilla Parga, R. (2024). Discriminación algorítmica en los procesos automatizados de reclutamiento y selección de personal. *Revista Chilena De Derecho Y Tecnología*, 13. <https://doi.org/10.5354/0719-2584.2024.71312>

Comisión Europea, Dirección General de Redes de Comunicación, Contenido y Tecnologías. (2019). *Directrices éticas para una IA fiable*. Oficina de Publicaciones. <https://data.europa.eu/doi/10.2759/14078>

Connor, K. (2004). Functionalism and Artificial Intelligence. *Episteme*, 15(1).

Consejo Superior Laboral. (2024). *Evaluación implementación de la Ley 21.431, que modifica el Código del Trabajo, regulando el contrato de trabajadores de empresas de plataformas digitales de servicios*. https://www.mintrab.gob.cl/wp-content/uploads/2024/08/Informe_Implementacion_Ley21.431_CSL2024.pdf

Contreras, P. (2022). La protección de datos personales de trabajadores de empresas de plataformas digitales de servicios y la regulación de la gestión algorítmica del trabajo bajo la Ley N° 21.431. En R. Palomo (ed.), *El trabajo a través de plataformas digitales. Problemas y desafíos en Chile* (pp. 157-172). Santiago de Chile: Tirant Lo Blanch.

Dowding, K., & Taylor, B. R. (2024). Algorithmic Decision-Making, Agency Costs, and Institution-Based Trust. *Philosophy & Technology*, 37(68). <https://doi.org/10.1007/s13347-024-00757-5>

European Data Protection Board. (2024). *Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models*. https://www.edpb.europa.eu/our-work-tools/our-documents/opinion-board-art-64/opinion-282024-certain-data-protection-aspects_en

Faúndez-Ugalde, A., & Mellado-Silva, R. (2023). Uso de automatización de procesos robóticos por administraciones tributarias e impacto en los derechos humanos. *Revista Chilena De Derecho Y Tecnología*, 12. <https://doi.org/10.5354/0719-2584.2023.65457>

Gaudeul, A., Arrigoni, O., Charisi, V., Escobar-Planas, M., & Hupont Torres, I. (2024). *Understanding the Impact of Human Oversight on Discriminatory Outcomes in AI-Supported Decision-Making*. <https://doi.org/10.3233/FAIA240598>

Google. (2024). *Perspectives on Issues in AI Governance*. <https://ai.google/static/documents/perspectives-on-issues-in-ai-governance.pdf>

Grupo de trabajo sobre protección de datos del artículo 19 (WP19). (2017). *Directrices sobre decisiones individualizadas automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679, adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018 (WP215rev.01)*. <https://www.aepd.es/documento/wp251rev01-es.pdf>

Hacker, P. (2023). The European AI liability directives—Critique of a half-hearted approach and lessons for the future. *Computer Law & Security Review*, 51.

Hoff, K., & Bashir, M. (2015). Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 57(3), 407-434. <https://doi.org/10.1177/0018720814547570>

Jiménez López, J. (2024). El Reglamento de Inteligencia Artificial y el Reglamento de protección de datos. En L. Cotino Hueso y P. Simón Castellanos (dirs.), *Tratado sobre el Reglamento de Inteligencia Artificial de la Unión Europea* (pp. 106-134). Navarra: Aranzadi.

Kaminski, M. (2019). The Right to Explanation, Explained. *Berkeley Technology Law Journal*, 34(1). <https://dx.doi.org/10.2139/ssrn.3196985>

Kellogg, K., Valentine, M., & Christin, A. (2020). Algorithms at Work: The New Contested Terrain of Control. *The Academy of Management Annals*, 14(1). <https://doi.org/10.5465/annals.2018.0174>

Klimas, T., & Vaiciukaite, J. (2008). The Law of Recitals in European Community Legislation. *ILSA Journal of International & Comparative Law*, 15.

Laux, J. (2024). Institutionalised distrust and human oversight of artificial intelligence: towards a democratic design of AI governance under the European Union AI Act. *AI & Society*, 39, 2853-2866. <https://doi.org/10.1007/s00146-023-01777-z>

Lazcoz Moratinos, G., & Castillo Parrilla, J. A. (2020). Valoración algorítmica ante los derechos humanos y el Reglamento General de Protección de Datos: el caso SyRI. *Revista chilena de derecho y tecnología*, 9(1), 207-225. <https://dx.doi.org/10.5354/0719-2584.2020.56843>

INTERVENCIÓN,
SUPERVISIÓN Y
CONTROL HUMANO
EN LAS DECISIONES
AUTOMATIZADAS

HUMAN
INTERVENTION,
OVERSIGHT
AND CONTROL
IN AUTOMATED
DECISION-MAKING

Lazcoz Moratinos, G. (2024). La vigilancia o supervisión humana en el artículo 14 del Reglamento de inteligencia artificial: ¿Un mero requisito obligatorio para los sistemas de alto riesgo? En L. Cotino Hueso y P. Simón Castellanos (dirs.), *Tratado sobre el Reglamento de Inteligencia Artificial de la Unión Europea* (pp. 570-586). Navarra: Aranzadi.

Ledford, H. (2019). Millions Affected by Racial Bias In Health-Care Algorithm. *Nature*, 574(7780), 608-609.

Le Moli, G. (2022). AI vs Human Dignity: When Human Underperformance is Legally Required. *Rethinking Capitalism*, (4). <https://geopolitique.eu/en/articles/ai-vs-human-dignity-when-human-underperformance-is-legally-required/>

López-Tarruella Martínez, A. (2024a). El futuro Reglamento de Inteligencia Artificial y las relaciones con terceros Estados. *Revista Electrónica De Estudios Internacionales*, (45). <https://reei.tirant.com/reei/article/view/2257>

López-Tarruella Martínez, A. (2024b). Derecho a presentar una reclamación y derecho a una explicación. Vías de recurso para los particulares en el reglamento de inteligencia artificial. En L. Cotino Hueso y P. Simón Castellanos (dirs.), *Tratado sobre el Reglamento de Inteligencia Artificial de la Unión Europea* (pp. 757-780). Navarra: Aranzadi.

López, C., Davidoff, A., Luco, F., Humeres, M., & Correa, T. (2024). Users' Experiences of Algorithm-Mediated Public Services: Folk Theories, Trust, and Strategies in the Global South. *International Journal of Human-Computer Interaction*, 41 (8), 5039-5056. <https://doi.org/10.1080/10447318.2024.2356910>

Matsumi, H., & Solove, D. J. (2023). The Prediction Society: Algorithms and the Problems of Forecasting the Future. *Illinois Law Review*, (1). https://scholarship.law.gwu.edu/faculty_publications/1708

Maxwell, W. (2023). *Meaningful Human Control to Detect Algorithmic Errors. Artificial Intelligence Law: Between Sectoral Rules and Comprehensive Regime - Comparative Law Perspectives*, Forthcoming. <https://ssrn.com/abstract=4551063>

Medina, M. (2022). El derecho a conocer los algoritmos utilizados en la toma de decisiones. Aproximación desde la perspectiva del derecho fundamental a la protección de datos personales. *Teoría y realidad constitucional*, (49), 141-171.

Míguez Macho, L., & Torres Carlos, M. (2024). Capítulo II. Sistemas prohibidos y sistemas de alto riesgo. En Moisés Barrio Andrés (dir.), *El Reglamento Europeo de Inteligencia Artificial* (pp. 49-85). Valencia: Tirant lo Blanch.

Obregón Fernández, A., & Guillermo Lazcoz Moratinos, G. (2021). La supervisión humana de los sistemas de inteligencia artificial de alto riesgo. Aportaciones desde el Derecho Internacional Humanitario y el Derecho de la Unión Europea. *Revista electrónica de estudios internacionales (REEI)*, (42). <https://doi.org/10.17103/reei.42.08>

Ossandón Cerda, F. (2021). Inteligencia Artificial en las Administraciones Tributarias: Oportunidades y Desafíos. *Revista De Estudios Tributarios*, 1 (24), 123-156. <https://revistaestudiostributarios.uchile.cl/index.php/RET/article/view/60703>

Palma Ortigosa, A. (2019). *Decisiones automatizadas en el RGPD. El uso de algoritmos en el contexto de la protección de datos*. <https://laadministraciondia.inap.es/noticia.asp?id=1509629>

Ponce del Castillo, A., & Naranjo, D. (2022). *Regulating algorithmic management An assessment of the EC's draft Directive on improving working conditions in platform work* [ETUI Research Paper]. Policy Brief 2022.06.

Ponce Solé, J. (2019). Inteligencia artificial, Derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico. *Revista general de Derecho administrativo*, (50), 58-67.

Ponce Solé, J. (2024). Límites jurídicos de la toma de decisiones discrecionales automatizadas mediante inteligencia artificial: Racionalidad, sabiduría y necesaria reserva jurídica de humanidad en el ámbito digital. *Revista General de Derecho Administrativo*, (66).

Radtke, T. (2025). Human Oversight under the AI Act and its interplay with Art. 22 GDPR. En B. Raue, A. von Ungern-Sternberg, L. K. Kumkar y T. Rűfner (eds.), *Artificial Intelligence and Fundamental Rights: The AI Act of the European Union and its implications for global technology regulation* (pp. 91-109). IRDT.

Ranchordas, S. (2021). Empathy in the digital administrative state. *Duke Law Journal*, 71, 1341-1389. <https://scholarship.law.duke.edu/dlj/vol71/iss6/4>

Rainone, S., & Aloisi, A. (2024) *The EU Platform Work Directive: what's new, what's missing, what's next?* [ETUI Research Paper]. Policy Brief 2024.6. <https://dx.doi.org/10.2139/ssrn.4957467>

Robbins, S. (2023). The many meanings of meaningful human control. *AI Ethics*, 4, 1377-1388. <https://doi.org/10.1007/s43681-023-00320-6>

Todolí Sígenes, A. (2024). La regulación de la inteligencia artificial en la Directiva de trabajo en plataformas digitales. *LABOS Revista De Derecho Del Trabajo Y Protección Social*, 5, 26-46. <https://doi.org/10.20318/labos.2024.9030>

Veale, M., & Edwards, L. (2018). Clarity, Surprises, and Further Questions in the Article 29 Working Party Draft Guidance on Automated Decision-Making and Profiling. *Computer Law & Security Review*, 34(2), 398-404. <https://doi.org/10.1016/j.clsr.2017.12.002>

Viollier, P., & Fischer, E. (2023). La intervención humana como resguardo ante la toma automatizada de decisiones: implicancias éticas y jurídicas. En Michelle Azuaje Pirela (ed.), *Introducción a la ética y el derecho de la inteligencia artificial* (pp. 151-169). Madrid: La Ley.

Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. *International Data Privacy Law*. <https://dx.doi.org/10.2139/ssrn.2903469>

Wachter, S. (2022). The theory of artificial immutability: Protecting algorithmic groups under anti-discrimination law. *Tulane Law Review*, 97(2). <https://dx.doi.org/10.2139/ssrn.4099100>

Wieczorek, K. (2021). The Conscience of a Machine? Artificial Intelligence and the Problem of Moral Responsibility. *ER(R)GO. Teoria-Literatura-Kultura*, (42), 15-34.

Jurisprudencia, normas y otros documentos legales

Directiva (UE) 2024/2831 del Parlamento Europeo y del Consejo de 23 de octubre de 2024 relativa a la mejora de las condiciones laborales en el trabajo en plataformas (Parlamento Europeo y Consejo Europeo, 2024). *Diario Oficial de la Unión Europea*, 11 de noviembre de 2024.

Directrices Éticas para una IA fiable (Grupo independiente de expertos de alto nivel sobre IA, 8 de abril de 2019). <https://digital-strategy.ec.europa.eu/es/library/ethics-guidelines-trustworthy-ai>

Dun & Bradstreet, Sentencia C-203/22 (TJUE, 27 de febrero de 2025).

Ley N.º 21.719, que regula la protección y el tratamiento de datos personales y crea la Agencia de Protección de Datos Personales (Ministerio Secretaría General de la Presidencia [Chile], 13 de diciembre de 2024). <https://www.bcn.cl/leychile/navegar?idNorma=1209272>

Ley N.º 21.431, que modifica el Código del Trabajo regulando el contrato de trabajadores de empresas de plataformas digitales de servicios (Ministerio del Trabajo y Previsión Social [Chile], 11 de marzo de 2022). <https://www.bcn.cl/leychile/navegar?idNorma=1173544>

Proyecto de Ley que Regula los Sistemas de Inteligencia Artificial (Boletín N.º 16821-19, Refundido con: 15869-19) (Cámara de Diputados [Chile], 7 de mayo de 2024).

Reglamento (UE) 2016/679, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos (Parlamento Europeo y Consejo Europeo, 2016). *Diario Oficial de la Unión Europea*, 27 de abril de 2016.

Reglamento (UE) 2024/1689, por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) 300/2008, (UE) 167/2013, (UE) 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de Inteligencia Artificial) (Parlamento Europeo y Consejo, 13 de junio de 2024). *Diario Oficial de la Unión Europea*, 13 de junio de 2024.

Schufa, Sentencia C-634/21 (TJUE, 7 de diciembre de 2023).

Sentencia C/09/550982 / HA ZA 18-388 (Corte del Distrito de La Haya [Países Bajos], 5 de febrero de 2020).

Recibido: 27/01/2025
Aprobado: 02/09/2025