

LOS SESGOS DE LA INTELIGENCIA ARTIFICIAL EN LA DOBLE DISCRIMINACIÓN LABORAL DE LA MUJER POR CAUSAS RELIGIOSAS¹

[ENG] *Artificial Intelligence Biases in the Double Employment Discrimination of Women on Religious Grounds*

Fecha de recepción: 29 septiembre 2025 / Fecha de aceptación: 28 noviembre 2025

RAFAEL GARCÍA-PURRIÑOS GARCÍA
Universidad Católica San Antonio de Murcia
(España)
ragarcia@ucam.edu
ORCID: 0000-0002-2495-674X
DOI: 10.12800/vg.21.3

Resumen: La utilización de sistemas de inteligencia artificial (IA) está transformando el mundo laboral, pero su implementación sin supervisión puede reforzar desigualdades estructurales. Entre los colectivos vulnerables se encuentran las mujeres con convicciones religiosas visibles, quienes sufren una doble marginación: género y adscripción religiosa. Esta discriminación se manifiesta en tres momentos clave: selección, desarrollo del contrato y despido. En cuanto a la selección, los algoritmos de reclutamiento aprenden de datos históricos con sesgos preexistentes que pueden perpetuar exclusiones. Además, los sistemas de cribado pueden penalizar nombres asociados a minorías religiosas o identificar vestimenta religiosa como un rasgo atípico. En culturas occidentales, el hiyab, la falda larga, la kipá o determinados símbolos cristianos visibles pueden generar prejuicios culturales, que se trasladan al algoritmo. A lo largo del desarrollo de la relación laboral, los sistemas de IA que evalúan desempeño, promociones, selecciones para formación o riesgos laborales pueden consolidar barreras discriminatorias, pudiendo llegar a recomendar la desvinculación de trabajadoras cuyos hábitos religiosos impliquen ajustes en la jornada o vestimenta, reforzando situaciones de exclusión. Es crucial la supervisión humana, las auditorías algorítmicas y la transparencia para construir entornos donde la tecnología actúe para garantizar la inclusión y no para perpetuar sesgos.

Palabras clave: IA, Sesgos, discriminación laboral, religión, mujer.

¹ La publicación es parte de la ayuda PID2023-147932OB-I00: HERRAMIENTAS PARA EL EJERCICIO DE LOS DERECHOS LABORALES DE MUJERES EN RIESGO DE EXCLUSIÓN, financiado por MCIU/AEI/10.13039/501100011033 y por FEDER, UE



Abstract: The use of artificial intelligence (AI) systems is transforming the labor market; however, their implementation without proper oversight can reinforce structural inequalities. Among vulnerable groups are women with visible religious convictions, who experience double marginalization based on both gender and religious affiliation. This discrimination manifests at three key stages: recruitment, the development of the employment relationship, and dismissal. In recruitment processes, hiring algorithms learn from historical data containing pre-existing biases, which can perpetuate exclusion. In addition, screening systems may penalize names associated with religious minorities or identify religious attire as an atypical trait. In Western cultures, the hijab, long skirts, the kippah, or certain visible Christian symbols may trigger cultural prejudices that are then transferred to the algorithm. Throughout the development of the employment relationship, AI systems used to assess performance, promotions, training selection, or occupational risks may entrench discriminatory barriers. These systems may even recommend the termination of employment for female workers whose religious practices require adjustments to working hours or dress codes, thereby reinforcing situations of exclusion. Human oversight, algorithmic audits, and transparency are crucial to building environments in which technology acts to guarantee inclusion rather than to perpetuate bias.

Keywords: Artificial Intelligence, Bias, Employment Discrimination, Religion, Women.

1. IMPLICACIONES ÉTICAS DE LA INCORPORACIÓN DE SISTEMAS DE INTELIGENCIA ARTIFICIAL EN EL ÁMBITO LABORAL. EFECTOS EN LA DISCRIMINACIÓN DE LOS TRABAJADORES

La inteligencia artificial (en adelante IA) marca una nueva fase en la relación de la humanidad con la tecnología, de manera que constituye el centro de lo que el Papa Francisco ha descrito como un “cambio de época”².

En la materia que nos ocupa, la incorporación progresiva de sistemas de IA en el ámbito laboral constituye uno de los fenómenos más relevantes de la actual transformación digital.

Las herramientas algorítmicas intervienen cada vez más en procesos tradicionalmente reservados a la valoración humana, como la selección y contratación de personal, las necesidades de formación, la evaluación del desempeño, la asignación de promociones, la gestión de la productividad o el estudio y decisión sobre despidos.

No podemos olvidar, en este aspecto, las diferencias que existen entre la IA y la inteligencia humana, puesto que, “en lo que respecta al ser humano, la inteligencia es una facultad relativa a la persona en su conjunto, mientras que, en el contexto de la IA, solo la podemos entender en un sentido funcional, sin asumir que las actividades características de la mente humana pueden descomponerse en pasos digitalizados, de modo que las máquinas puedan replicarlas. Por ello,

*“si bien, se trata de una extraordinaria conquista tecnológica capaz de imitar algunas acciones asociadas a la racionalidad, la IA obra solamente realizando tareas, alcanzando objetivos o tomando decisiones basadas sobre datos cuantitativos y sobre la lógica computacional”*³.

Pese a que estas tecnologías suelen presentarse como instrumentos objetivos, eficientes y neutrales, capaces de reducir la arbitrariedad y los sesgos humanos en la toma de decisiones empresariales, la realidad es que, lejos de erradicar las desigualdades existentes, los sistemas de IA pueden reproducir e incluso intensificar sesgos estructurales

² FRANCISCO, «Discurso a los participantes a la Asamblea Plenaria de la Pontificia Academia para la Vida (28 de febrero de 2020)», en *AAS* 112 (2020), p. 307. Cf. Id., «Discurso a la Curia Romana con motivo de las felicitaciones navideñas (21 de diciembre de 2019)», en *AAS* 112 (2020), p. 43

³ DICASTERIO PARA LA DOCTRINA DE LA FE, DICASTERIO PARA LA CULTURA Y LA EDUCACIÓN, *Antiqua et nova. Nota sobre la relación entre la inteligencia artificial y la inteligencia humana*, 28 de enero de 2025, p. 6.



presentes en los datos con los que son entrenados y en los contextos sociales en los que operan.

En el ámbito laboral, esto se puede traducir en la automatización de prácticas discriminatorias que, bajo una apariencia técnica y objetiva, consolidan patrones de exclusión preexistentes. La utilización acrítica de algoritmos en la gestión de recursos humanos plantea, por tanto, importantes interrogantes desde la perspectiva del principio de igualdad y no discriminación, así como del respeto a los derechos fundamentales en el trabajo.

Entre los colectivos que se pueden ver afectados por estos procesos se encuentran las mujeres con convicciones religiosas visibles, quienes se sitúan en una posición de especial vulnerabilidad al confluir en su persona dos motivos clásicos de discriminación: el sexo y la religión.

La discriminación laboral por razón de sexo continúa manifestándose en la actualidad en fenómenos como la segregación ocupacional, la brecha salarial, la penalización de la maternidad o las dificultades de acceso a puestos de responsabilidad. Por su parte, la discriminación por motivos religiosos adquiere una especial intensidad cuando las creencias se exteriorizan de forma visible, ya sea mediante la vestimenta, el uso de símbolos religiosos o la observancia de determinadas prácticas vinculadas a horarios, festividades o períodos de ayuno.

La confluencia de ambas causas da lugar a una discriminación agravada o interseccional que suele quedar invisibilizada tanto en las políticas empresariales de igualdad como en los propios sistemas automatizados de toma de decisiones. En efecto, en contextos laborales marcados por estándares culturales mayoritarios, determinados signos religiosos externos como el hiyab, la falda larga, la kipá o símbolos cristianos visibles pueden ser percibidos como elementos ajenos a la norma organizativa, generando prejuicios culturales que son posteriormente trasladados al diseño y funcionamiento de los algoritmos. Y así, la IA no solo refleja sesgos sociales preexistentes, sino que contribuye a su consolidación y normalización a través de decisiones automatizadas difíciles de detectar y, en consecuencia, prácticamente imposibles de cuestionar.

La relevancia del presente artículo se justifica, por tanto, en la necesidad de analizar críticamente el impacto de la IA en la reproducción de la doble discriminación

laboral de la mujer por causas religiosas, un fenómeno todavía escasamente abordado de manera específica en la literatura jurídica y sociolaboral.

Aunque existen numerosos trabajos sobre sesgos algorítmicos o sobre discriminación en el empleo, el análisis de la intersección entre sexo, religión e IA sigue siendo poco tratado, a pesar de su creciente importancia práctica. Y se trata de abordar el tema en el marco de la necesidad de una utilización ética de la IA, en el que va a ser necesario poner límites a su uso indiscriminado en la toma de decisiones laborales sobre, en este caso, las trabajadoras.

El objetivo principal es examinar cómo los sesgos algorítmicos pueden manifestarse a lo largo de las distintas fases de la relación laboral (selección y contratación, desarrollo del contrato y extinción), afectando a mujeres con convicciones religiosas visibles, con especial atención a las actividades relacionadas con el acceso al empleo. Asimismo, se pretende poner de relieve los límites de una concepción meramente técnica de la IA, llamando la atención sobre la necesidad de incorporar enfoques jurídicos, éticos y de derechos humanos en su diseño, implementación y supervisión.

Desde esta perspectiva, el trabajo aspira no solo a describir los riesgos, sino también a contribuir al diseño de marcos éticos, normativos y organizativos que garanticen un uso inclusivo y responsable de la IA.

Desde el punto de vista metodológico, el artículo se va a centrar en el análisis crítico de la literatura especializada sobre IA, sesgos algorítmicos y discriminación interseccional, así como en el estudio del marco jurídico aplicable en materia de utilización de IA, igualdad y no discriminación en el empleo. Este análisis se complementa con ejemplos ilustrativos que permiten visualizar cómo los sistemas de IA pueden operar como mecanismos de exclusión indirecta, incluso aunque no exista una intención discriminatoria explícita.

En cuanto a la estructura del artículo, en primer lugar, se abordan los marcos legal y conceptual relativos a la discriminación por razones de sexo o de religión, la inteligencia artificial y los sesgos algorítmicos en el ámbito laboral, con especial atención a los procesos de toma de decisiones automatizadas.



En segundo término, se analiza la discriminación laboral por razón de género y religión desde una perspectiva interseccional, identificando a las mujeres con convicciones religiosas visibles como un colectivo particularmente expuesto.

A continuación, se examina de forma detallada el impacto de la IA en las distintas fases de la relación laboral. Posteriormente, se estudian las principales respuestas jurídicas y éticas frente a la discriminación algorítmica, destacando el papel de la supervisión humana, las auditorías algorítmicas y la transparencia.

Finalmente, el artículo concluye con una serie de reflexiones y propuestas orientadas a garantizar que IA actúe como una herramienta al servicio de la inclusión y la igualdad efectiva en el empleo.

2. APROXIMACIÓN AL MARCO NORMATIVO

2.1. Principio de igualdad de trato y oportunidades y discriminación de la mujer

El principio de igualdad de trato y de oportunidades, así como la no discriminación en el acceso al empleo constituye una preocupación real para las instituciones relacionadas con los derechos de los trabajadores desde hace décadas.

Así, el Convenio número 111 de la Organización Internacional del Trabajo (en adelante OIT) sobre la discriminación (empleo y ocupación) de 1958 señalaba que se entendía como discriminación “*cualquier distinción, exclusión o preferencia, basada en motivos de raza, color, sexo, religión, opiniones políticas (...) que tenga por efecto anular o alterar la igualdad de oportunidades o de trato en el empleo y la ocupación*”. Entendiendo los términos empleo y ocupación también como referidos al acceso a los medios de formación profesional y la admisión en el empleo, así como las condiciones de trabajo y despido⁴. Hay que destacar que este concepto de discriminación no se refiere de forma exclusiva a la legislación y prácticas de los estados, sino que alcanzará a “*cualquier distinción, exclusión o preferencia*” sea cual sea su autor y el medio utilizado. Se extiende,

⁴ Ratificado por España con fecha 6 de noviembre de 1967 (BOE de 4 de diciembre de 1968).

en consecuencia, que afecta tanto a particulares e interlocutores sociales, y a normas, contratos (individuales y colectivos), órdenes y prácticas⁵.

En la Unión Europea esta preocupación se viene reflejando desde sus tratados fundacionales, con la aparición de numerosas directivas sobre la igualdad de trato, como es la Directiva 2000/43/CE, sobre igualdad por motivo racial o étnico, la Directiva 2000/78/CE sobre igualdad en el empleo y la ocupación, o la Directiva 2006/54/CE, del Parlamento Europeo y del Consejo, de 5 de julio de 2006, sobre aplicación del principio de igualdad de oportunidades e igualdad de trato entre hombres y mujeres en asuntos de empleo y ocupación, entre otras.

En España se recoge la igualdad como valor superior del ordenamiento jurídico, como derecho fundamental (art. 14, la llamada igualdad formal) y como mandato a los poderes públicos para que promuevan, a través de políticas positivas y concretas, las condiciones necesarias para su cumplimiento (la igualdad efectiva, del artículo 9.2).

Este mandato se refleja en un conjunto amplio de leyes, como el Estatuto de los Trabajadores y el Estatuto del Empleado Público y, sobre todo, a partir de la Ley Orgánica 3/2007, de 22 de marzo, de Igualdad Efectiva de mujeres y hombres.

Esta ley reconoce, en el artículo 5, la garantía del principio de igualdad de trato y de oportunidades entre mujeres y hombres en el ámbito del empleo público y privado, señalando que no constituirá discriminación en el acceso al empleo una diferencia de trato basada en una característica relacionada con el sexo cuando, debido a la naturaleza de las actividades profesionales concretas o al contexto en el que se lleven a cabo, dicha característica constituye un requisito profesional esencial determinante, siempre y cuando el objeto sea legítimo y el requisito proporcionado⁶.

Así, existirá discriminación cuando se dispense un trato desfavorable a las mujeres en el acceso al empleo, siendo directa cuando la persona recibe un trato menos favorable

⁵ MALDONADO MONTOYA, J. P., «Convenio sobre la discriminación (empleo y ocupación) 1958 (núm. 111)», en *Revista Internacional y Comparada de Relaciones Laborales y Derecho del Empleo* 7, número especial de conmemoración del Centenario de la OIT (2019).

⁶ GONZÁLEZ BUSTOS, M. A., «La discriminación por razón de género en el acceso al empleo en situaciones de exclusión social», en AA. VV., *La inserción laboral de las mujeres en riesgo de exclusión Social*, Tirant Lo Blanch, Valencia 2015, p. 16.



que otra en situación similar, e indirecta cuando al aplicar criterios aparentemente neutros se perjudica a un grupo protegido por la legislación antidiscriminatoria⁷.

Después de la Ley de Igualdad, numerosas normas posteriores han incluido disposiciones en la materia, en especial (para el tema que se va a tratar), el Real Decreto-ley 6/2019, de 1 de marzo, de medidas urgentes para garantía de la igualdad de trato y de oportunidades entre mujeres y hombres en el empleo y la ocupación. Que modifica varias normas para hacerlas compatibles con estos principios, así como la ley 15/2022, de 22 de julio, integral para la igualdad de trato y no discriminación, a la que más adelante se hará mención.

2.2. La libertad religiosa y la discriminación por razones religiosas

Actualmente, en la sociedad española conviven y se relacionan personas con diferentes creencias, ideas, identidades y culturas. Esta diversidad se traslada al mundo de la empresa, como reflejo que es de la sociedad en la que opera. En las empresas conviven hombres y mujeres de distintas religiones con otros que son agnósticos, ateos o indiferentes ante el hecho religioso.

Este aumento de la diversidad religiosa ha propiciado la aparición de diferentes conflictos en el marco laboral relacionados con el ejercicio del derecho de libertad religiosa por parte de las trabajadoras (días festivos, símbolos religiosos visibles, prendas de vestir, objeción de conciencia, etc.)

Podemos conceptualizar la libertad religiosa como una de las modalidades de libertad de conciencia que comprende derechos como: a) profesar una religión o abstenerse de hacerlo, b) expresar libremente una opinión con fundamento en una confesión religiosa, c) reunirse y asociarse, d) aprender y enseñar; y, e) practicar el culto⁸.

La tutela de la libertad religiosa y de la no discriminación por razón de religión en el ámbito de las relaciones laborales se encuentra recogida en varias normas jurídicas. Vamos a analizar las relevantes, sin un ánimo de ser exhaustivos, pero sí de presentar un panorama general del marco normativo de la libertad religiosa, en lo que pueda ser relevante en el plano laboral, que es el que nos ocupa.

⁷ OIT, *Lograr la igualdad de oportunidades en el empleo para las personas con discapacidades a través de la legislación: directrices*, Oficina Internacional del Trabajo, Ginebra 2007, p. 29.

⁸ PRIETO SÁNCHEZ, L., «Las Minorías Religiosas», en *Anuario De Derecho Eclesiástico Del Estado* (1993), p. 159.

La norma fundamental de referencia, como no puede ser de otra manera, es la Constitución Española de 1978 (en adelante CE). Además de las disposiciones que ya hemos abordado en referencia a la igualdad y no discriminación (artículos 9 y 14 fundamentalmente), hemos de referirnos, en especial, al artículo 16, que reconoce el derecho tanto de los individuos como de las comunidades a la libertad ideológica, religiosa y de culto, sin que nadie pueda ser obligado a declarar sobre su ideología, religión o creencias, así como la obligación de los poderes públicos, en el marco de un estado aconfesional, a tener en cuenta las creencias religiosas que están presentes en la sociedad. El artículo 18, por su parte, hace referencia, entre otros, al derecho a la propia imagen.

Otro artículo que puede ser importante en nuestro contexto es el 20 (libre difusión de pensamientos e ideas, así como la objeción de conciencia de los profesionales).

Por otra parte, La Carta de los Derechos Fundamentales de la Unión Europea reconoce el derecho de toda persona a la libertad de pensamiento, de conciencia y de religión (artículo 10), el derecho a la no discriminación por razón de religión o convicciones (artículo 21) y la diversidad cultural, religiosa y lingüística (artículo 22). Además, la Directiva 2000/78/CE, del Consejo, de 27 de noviembre de 2000, relativa al establecimiento de un marco general para la igualdad de trato en el empleo y la ocupación. Esta Directiva, que ha sido transpuesta al ordenamiento español por medio de la Ley 62/2003, de 30 de diciembre, de medidas fiscales, administrativas y del orden social, prohíbe la discriminación directa e indirecta por razón de religión en el ámbito laboral⁹.

En España, este derecho constitucional a la Libertad Religiosa fue desarrollado por la Ley Orgánica 7/1980, de 5 de julio, de Libertad Religiosa, que recoge los derechos protegidos, así como sus límites (protección de derechos de otras personas, orden público, seguridad y salud públicas, etc.)

En el ámbito laboral, el reconocimiento de este derecho a la libertad religiosa llega al Estatuto de los Trabajadores que, en su artículo 4.2 apartado c les reconoce el derecho a no ser discriminados, directa o indirectamente, por razón de religión o convicciones en el momento del acceso al empleo o una vez empleados. Asimismo, el artículo 17 proscribe

⁹ RODRÍGUEZ BLANCO, M., CANO RUIZ, I., «La libertad religiosa en el ámbito de las relaciones de trabajo», en *Observatorio del Pluralismo religioso en España* (2021), p. 9



cualquier clausula, norma o precepto que pueda dar lugar a situaciones de discriminación directa o indirecta desfavorables por razón de religión o convicciones. Será, asimismo nulo (53.4) todo despido que tenga como motivo una causa de discriminación (también la religiosa, claro). Por fin, el artículo 54.3 considerará causa de despido el acoso por causa de religión o convicciones.

Todo ello con independencia de las disposiciones que puedan incluir los diferentes convenios colectivos al abordar esta problemática y adaptarla a las características de sectores y de empresas (festividades, vestimenta, etc.)

2.3. Regulación de la utilización de algoritmos en el ámbito laboral

Si buscamos referencias legales a los algoritmos, nos encontramos la Ley 12/2021, de 28 de septiembre, por la que se modifica el texto refundido de la Ley del Estatuto de los Trabajadores, aprobado por el Real Decreto Legislativo 2/2015, de 23 de octubre, para garantizar los derechos laborales de las personas dedicadas al reparto en el ámbito de plataformas digitales, más conocida como Ley Riders, que introduce en el Estatuto de los Trabajadores una nueva letra d) en el artículo 64.4, referido a los derechos de los representantes legales de los trabajadores con la siguiente redacción:

“Ser informado por la empresa de los parámetros, reglas e instrucciones en los que se basan los algoritmos o sistemas de inteligencia artificial que afectan a la toma de decisiones que pueden incidir en las condiciones de trabajo, el acceso y mantenimiento del empleo, incluida la elaboración de perfiles”.

Aparte de este deber de información, la referencia a los algoritmos para la toma de decisiones no va más allá.

Sin ser una norma positiva, la carta de los derechos digitales de los trabajadores sí que hacía referencia a una serie de derechos ante la IA que, además de perseguir el bien común y centrarse en la dignidad de la persona, deberá garantizarse “el derecho a la no discriminación algorítmica, cualquiera que fuera su origen, causa o naturaleza del sesgo, en relación con las decisiones y procesos basados en algoritmos”, así como la “transparencia, auditabilidad, explicabilidad y trazabilidad” y la “accesibilidad, usabilidad y fiabilidad”.

Se reconoce, además, el derecho de la persona a que, sobre ella, no se tomen decisiones exclusivamente mediante algoritmos, por lo que podrán impugnar esas decisiones o solicitar intervención humana.

En este contexto, la ley 15/2022, de 22 de julio, integral para la igualdad de trato y no discriminación, introduce por primera vez una regulación positiva respecto de la utilización por las empresas y las administraciones públicas de algoritmos de IA. Ya en su artículo 3.1, al establecer el ámbito objetivo de aplicación de la ley, se refiere de forma expresa a la IA, recogiendo en su letra o) que la ley se aplicará a la “*Inteligencia Artificial, y gestión masiva de datos, así como otras esferas de análoga significación*”.

En la regulación de estos aspectos, trata de diseñar unas líneas de actuación que, a grandes rasgos, sirvan para promover buenas prácticas respecto de aquellos algoritmos que puedan servir para la toma de decisiones.

Lo hace en el artículo 23, cuyo tenor literal es el siguiente:

“(...) las administraciones públicas favorecerán la puesta en marcha de mecanismos para que los algoritmos involucrados en la toma de decisiones que se utilicen en las administraciones públicas tengan en cuenta criterios de minimización de sesgos, transparencia y rendición de cuentas, siempre que sea factible técnicamente. En estos mecanismos se incluirán su diseño y datos de entrenamiento, y abordarán su potencial impacto discriminatorio. Para lograr este fin, se promoverá la realización de evaluaciones de impacto que determinen el posible sesgo discriminatorio”.

Amplia esta regulación a las empresas en su apartado tercero, que dispone que

“Las administraciones públicas y las empresas promoverán el uso de una Inteligencia Artificial ética, confiable y respetuosa con los derechos fundamentales, siguiendo especialmente las recomendaciones de la Unión Europea en este sentido”

Se trata de una regulación poco concreta y de carácter eminentemente programático, pero tiene importancia porque con ella se pone de manifiesto, por primera vez, la preocupación del legislador por estos aspectos.

Por fin, la Ley 3/2023 de 28 de febrero, de empleo, dedica los artículos 16 y 17 al tratamiento de los datos, haciendo, en primer lugar, remisión expresa a la legislación vigente en la materia (Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, 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 y por el que se deroga la Directiva 95/46/CE, y la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.).

Contempla, en los apartados siguientes del artículo 16, una serie de disposiciones destinadas a delimitar los datos personales que son objeto de tratamiento en el ámbito de los planes y políticas de empleo, así como la delimitación de los responsables y las finalidades aceptadas para el tratamiento de los datos (*“ocupación y mejora de la empleabilidad de personas demandantes de los servicios de empleo”, “adecuada protección y activación de las mismas cuando se encuentran en situación de necesidad y de falta de empleo”, “prestación de servicios garantizados a las personas, empresas y demás entidades empleadoras”* y, por último, *“datos recabados con fines estadísticos o para la realización de estudios de investigación científica”*). Los datos se limitarán a los imprescindibles para el cumplimiento de estas finalidades.

Además, en el artículo 17 se hace referencia a la toma de decisiones (automatizadas o no) basadas en el análisis de datos, las evidencias estadísticas y el análisis del mercado de trabajo. En este artículo se nos adelantan algunos mandatos importantes, de manera que se actúe *“siempre en defensa de los derechos al empleo y al acceso a los servicios públicos de empleo de la persona interesada y se evitará la discriminación directa o indirecta”*, el acceso de los interesados *“al contenido íntegro de dichas decisiones y recomendaciones”* y, en el caso de la utilización de herramientas basadas en algoritmos de Inteligencia Artificial (en adelante IA), *“el conjunto de instrucciones que compondrán el algoritmo subyacente a las decisiones y recomendaciones basadas en el análisis de datos y las evidencias estadísticas.”*

Por fin, los datos de categoría especial de acuerdo con el artículo 9.1 del Reglamento (UE) 2016/679. No se podrán utilizar en estos procesos.

Con posterioridad, se analizará el Reglamento Europeo de Inteligencia Artificial (Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024) que puntualiza y añade una nueva dimensión a la regulación de la IA, de la que hace, por fin, una definición legal.

De igual modo, garantiza asimismo la vigilancia y la evaluación humanas de las decisiones automatizadas con inclusión del derecho a recibir explicaciones sobre dichas

decisiones y a que éstas sean revisadas la Directiva relativa a la mejora de las condiciones laborales en el trabajo en plataformas digitales aprobada en marzo de 2024 que, entre otras previsiones, garantiza que *“toda decisión de restringir, suspender o poner fin a la relación contractual o a la cuenta de una persona que realice trabajo en plataformas, o cualquier otra decisión que cause un perjuicio equivalente, será adoptada por un ser humano”*.

Además, la Directiva, prohíbe el empleo de sistemas automatizados de supervisión o de toma de decisiones que requieran el tratamiento de ciertos tipos de datos personales de estos trabajadores, como datos biométricos o datos sobre el estado emocional o psicológico de las personas.

3. INTELIGENCIA ARTIFICIAL Y SU UTILIZACIÓN EN EL ÁMBITO LABORAL: EL PELIGRO DE LOS SESGOS ALGORÍTMICOS

Al hablar de discriminación por razón de sexo en el acceso al empleo nos referimos tanto a los procesos de selección como a las condiciones ofertadas para la colocación. De esta manera, al analizar si ha existido discriminación no hemos de fijarnos exclusivamente en los requisitos exigidos para el acceso al empleo y las condiciones ofrecidas, sino que hemos de analizar también los resultados posteriores, abarcando los criterios de selección, la oferta de empleo, el propio proceso de selección, y todas las acciones previas al acto de la contratación y entrada en la empresa¹⁰.

En caso de existencia de duda (demostrada con indicios suficientes) respecto del respeto a la igualdad en el proceso de contratación, será el empresario quien habrá de demostrar de manera objetiva y razonable la selección del candidato y las razones que llevaron a la misma. Lo mismo cabrá decir respecto de la utilización de la IA una vez iniciada la relación laboral, en aspectos como la selección de personas para la formación, promoción y desarrollo, la evaluación del desempeño y la selección de personas trabajadoras para procesos de suspensión de contrato, reducciones de jornada o despidos por razones técnicas, económicas, organizativas y de producción. Y no digamos a la hora

¹⁰ GONZÁLEZ BUSTOS, M. A., «La discriminación por razón de género en el acceso al empleo en situaciones de exclusión social», cit., p. 18.



de realizar actividades de control y la imposición de posibles sanciones e, incluso, despidos de carácter disciplinario.

En todos estos momentos, la discriminación en el ámbito laboral puede acentuarse cuando el hecho de ser mujer se suma a cualquier otro de los condicionantes de la discriminación y desigualdad, y así ocurre en el caso de las mujeres con convicciones religiosas visibles. Lo que se conoce como discriminación interseccional, entendida como aquella que se refiere a situaciones en las que varios motivos discriminatorios actúan e interactúan entre sí al mismo tiempo, de tal manera que son inseparables y dan lugar a tipos particulares de discriminación¹¹.

Esta intersección es difícil de identificar y da lugar a un motivo de discriminación nuevo y, como tal, más difícil de detectar, denunciar y juzgar¹².

En todo caso, y antes de abordar la problemática específica planteada por la utilización de sistemas de IA en el ámbito laboral y sus posibles sesgos discriminatorios, hemos de aclarar los conceptos básicos que nos permitan entender con claridad el funcionamiento de estas tecnologías, así como los riesgos que puede acarrear a su aplicación en el ámbito laboral y, en concreto, en los procesos de toma de decisiones, cuando éstas pueden afectar directamente a los derechos de las personas trabajadoras.

En particular, resulta imprescindible analizar el modo en que los sesgos algorítmicos pueden emplearse y operar, así como su posible producción o perpetuación de acciones discriminatorias.

3.1. Aproximación al concepto de Inteligencia Artificial

Antes de abordar el estudio del impacto de la IA en la selección de personal y la prevención de los sesgos, de manera particular en los colectivos que pueden ser objeto de discriminaciones, hemos de tratar de aproximarnos al concepto de IA.

En este sentido, y aún sin intención de profundizar en exceso en nuestro análisis, nos podemos encontrar con multitud de definiciones, puesto que los autores abordan el concepto de acuerdo con el sector y el enfoque hacia el que estén orientadas sus

¹¹ Agencia de los Derechos Fundamentales de la Unión Europea y Consejo de Europa, *Manual de legislación europea contra la discriminación* (2018) (https://www.echr.coe.int/Documents/Handbook_non_discrim_law_SPA.pdf)

¹² FRÍAS, M., «La interseccionalidad de la discriminación por razones de raza, etnia y género», en *Proyecto “Convivir sin discriminación: un enfoque basado en los derechos humanos y la perspectiva de género”* (2022), p. 16.

investigaciones. Por ello, dentro del término, nos podemos encontrar con técnicas, instrumentos o herramientas muy diferentes y heterogéneas.

En el lenguaje común, se introduce este término cuando una máquina es capaz de imitar las funciones cognitivas propias de la mente humana, tales como: creatividad, sensibilidad, aprendizaje, entendimiento, percepción del entorno y utilización del lenguaje¹³.

El primer gran hito en el desarrollo de la IA se remonta al año 1956, cuando el informático norteamericano John McCarthy organizó un congreso veraniego en la Universidad de Dartmouth para afrontar el problema de la IA, que él definía como “hacer una máquina capaz de mostrar un comportamiento que se calificaría de inteligente si fuera un ser humano quien lo produjera”¹⁴.

En esa misma línea, para Nieva Fenoll, supone “la capacidad de que las máquinas puedan pensar o imitar el pensamiento humano a base de aprender y utilizar las generalizaciones que usamos para tomar nuestras decisiones habituales”¹⁵.

En efecto, con carácter general, podemos definir a la IA como aquella habilidad de las máquinas y dispositivos para aprender información y usarla como medio en la toma de decisiones. Se trata, efectivamente, de atribuir a una máquina una habilidad puramente humana¹⁶.

Entonces, la IA sería “*aquella inteligencia que exhiben ciertos sistemas o artefactos contruidos por el hombre. Un sistema artificial poseería inteligencia cuando es capaz de llevar a cabo tareas que, si fuesen realizadas por un humano, se diría de este que es inteligente*”¹⁷.

En el propio término encontramos la palabra “artificial” acompañada de una cualidad propiamente humana, la inteligencia, en contraposición, pues, con la “natural”, la que es propia de los seres humanos, lo que, ya de entrada, nos sugiere que el origen de

¹³ GRIGORE, A. E., «Derechos humanos e inteligencia artificial», en *Ius Et Scientia: Revista electrónica de Derecho y Ciencia* 8 (2022), p. 165.

¹⁴ MCCARTHY, J., MINSKY, M. L., ROCHESTER, N., SHANNON, C. E., «A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence» (31 de Agosto de 1955) (<http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>) (acceso: 12 de enero de 2026).

¹⁵ NIEVA FENOLL, J., *Inteligencia artificial y proceso judicial*, Marcial Pons, Madrid 2018, p. 20.

¹⁶ ROUHIAINEN, L., *Inteligencia artificial*, Alienta Editorial, 2018, p. 17.

¹⁷ LÓPEZ ONETO, M., *Fundamentos para un derecho de la Inteligencia Artificial. ¿Queremos seguir siendo humanos?*, Tirant Lo Blanch, Valencia 2020, p. 39.



la IA ha de ser el resultado de un esfuerzo informático intencionado, y no de la inteligencia de una persona¹⁸.

La mayor parte de los sistemas actuales, en particular aquellos que usan el aprendizaje automático, están basados en inferencias estadísticas más que en deducciones lógicas. Se trata de analizar enormes conjuntos de datos con el objetivo de identificar patrones, de manera que se pueda predecir efectos, imitando así procesos cognitivos típicos de la capacidad humana de resolución de problemas¹⁹.

Se trata, en definitiva, de “*emular las diversas capacidades del cerebro humano para presentar comportamientos inteligentes sintetizando y automatizando tareas intelectuales*”²⁰.

Se da la circunstancia de que, además, las máquinas desempeñan estas habilidades con mayor facilidad y rapidez que un ser humano, sin sentir fatiga y cometiendo muchos menos errores²¹.

Un sistema de IA necesita de una secuencia de instrucciones que especifique las diferentes acciones que debe ejecutar el ordenador para resolver un determinado problema. Esta secuencia de instrucciones es lo que se conoce como estructura algorítmica del Sistema de IA²².

De acuerdo con el glosario de la Carta Ética Europea sobre el uso de la IA en los sistemas judiciales y su entorno de 4 de diciembre de 2018, podemos definir el algoritmo como

“una secuencia finita de reglas formales (operaciones e instrucciones lógicas) que permiten obtener un resultado de la entrada inicial de información.

¹⁸ BARRIO ANDRÉS, M., *El reglamento Europeo de Inteligencia Artificial*, Tirant Lo Blanch, Valencia 2024, p. 22.

¹⁹ DICASTERIO PARA LA DOCTRINA DE LA FE, DICASTERIO PARA LA CULTURA Y LA EDUCACIÓN, *Antiqua et nova. Nota sobre la relación entre la inteligencia artificial y la inteligencia humana*, cit., p. 2.

²⁰ NAVAS NAVARRO, S., «Derecho e inteligencia artificial desde el diseño. Aproximaciones», en Navas Navarro, S. (dir.), *Inteligencia artificial. Tecnología. Derecho*, Tirant lo Blanch, Valencia 2017, p. 24.

²¹ ROUHIAINEN, L., *Inteligencia artificial: 101 cosas que debes saber hoy sobre nuestro futuro*, Alienta Editorial, Barcelona 2018, p. 17.

²² NAVAS NAVARRO, S., «Derecho e inteligencia artificial desde el diseño. Aproximaciones», cit., p. 25.

Esta secuencia puede ser parte de un proceso de ejecución automatizado y aprovechar modelos diseñados a través del aprendizaje automático²³”

Nótese como, en este caso, estamos utilizando un verbo propio de humanos, como es el verbo aprender, para referirnos a operaciones que realizan las máquinas. Realmente aquí, “aprender”, significa que “*el algoritmo encuentra patrones por sí solo y después, con el paso del tiempo, los conecta con resultados*²⁴”. Para ello, será necesario manejar enormes cantidades de datos y unos sistemas enormemente potentes y capaces de hacerlo²⁵. No se trata de un aprendizaje similar al humano, no es un proceso cognitivo, es un proceso estadístico en el que un algoritmo tiene la capacidad de encontrar patrones y crear sus propios modelos que se relacionan con los datos. Esto es, las predicciones realizadas por los algoritmos de IA no se basan en una teoría, sino en los propios datos. Datos que, por otra parte, determinan la predicción²⁶.

El citado glosario de la Carta Ética Europea sobre el uso de la Inteligencia Artificial en los sistemas judiciales y su entorno se refiere a la IA como “conjunto de métodos, teorías y técnicas científicas cuyo objetivo es reproducir, mediante una máquina, las habilidades cognitivas de los seres humanos²⁷”

La IA se basa en algoritmos con capacidad autónoma, que, una vez sistematizados y programados para la consecución de unos fines determinados fijados por la empresa, llevan a cabo tareas de filtración de información, predicción de situaciones, identificación de patrones, y evolucionan y aprenden de su propia experiencia para la consecución de los objetivos diseñados²⁸.

Esta capacidad para aprender y para tomar decisiones permite a la IA realizar multitud de tareas, sea de manera autónoma, sea como asistencia, que antes correspondía

²³ Carta ética europea sobre el uso de la inteligencia artificial en los sistemas judiciales y su entorno (<https://protecciondata.es/wp-content/uploads/2021/12/Carta-Etica-Europea-sobre-el-uso-de-la-Inteligencia-Artificial-en-los-sistemas-judiciales-y-su-entorno.pdf>) (acceso 12 de enero de 2026), p. 49.

²⁴ O’NEIL, C., *Armas de destrucción matemática. Cómo el Big Data aumenta la desigualdad y amenaza la democracia*, Capitán Swing, Madrid 2017, p. 96.

²⁵ RODRÍGUEZ FERNÁNDEZ, M. L., «Inteligencia Artificial, Género y Trabajo», en *Temas Laborales* 171 (2004), p. 14.

²⁶ SERRA-NAVARRO, D., «Inteligencia Artificial, vampirismo y la caja negra: aproximación especulativa en el marco del aprendizaje automático», en *Arte, Individuo y sociedad* 36 (2024), p. 969.

²⁷ Carta ética europea sobre el uso de la inteligencia artificial en los sistemas judiciales y su entorno (<https://protecciondata.es/wp-content/uploads/2021/12/Carta-Etica-Europea-sobre-el-uso-de-la-Inteligencia-Artificial-en-los-sistemas-judiciales-y-su-entorno.pdf>) (acceso 12 de enero de 2026), p. 49.

²⁸ PÉREZ LÓPEZ, J. I., «Inteligencia Artificial y contratación laboral», en *Revista de Estudios Jurídico-Laborales y de Seguridad Social* 7 (2023), p. 189.



realizar tan sólo a seres humanos, que eran los únicos con capacidad para poder llevarlas a cabo. También podrá servir para realizar predicciones sobre comportamientos. Además, la investigación sobre IA está haciendo que ésta evolucione y experimente enormes mejoras a un ritmo vertiginoso.

Todas estas capacidades, evidentemente, comportan también importantes riesgos. Como ya se ha comenzado, y respecto de la discriminación, resultaría ciertamente cínico decir ahora que existe una realidad analógica libre de discriminaciones o vulneraciones de derechos fundamentales y una realidad algorítmica plagada de ellos. Lo que ocurre es que un algoritmo no tiene vida propia, es sencillamente una fórmula matemática que no hace sino reproducir estereotipos, prejuicios y discriminaciones que ya existen en la sociedad, de donde estos algoritmos extraen sus datos²⁹.

Por eso, desde hace años se viene trabajando en normas que definan y acoten la utilización de esta IA a aquellos campos en los que realmente constituye un avance y una mejora en eficacia y eficiencia, proscribiendo y limitando aquellos en los que puede constituir un peligro, y favoreciendo un desarrollo ético de estos sistemas.

Evidentemente, que una decisión adoptada por medio de IA sea posible tecnológicamente no la hace deseable, de manera automática y acrítica, desde el punto de vista social. Es la mano humana quien, en última instancia, habrá de determinar las decisiones que pueden y no pueden adoptarse mediante tecnologías de IA³⁰. Es una grave responsabilidad, en clave ética e incluso legal, de la persona que lo maneja y lo utiliza en la toma de decisiones, el hecho de conocer el razonamiento que está detrás de cada decisión autónoma tomada por un algoritmo, es decir, es imprescindible la llamada “*explicabilidad*” de las decisiones³¹.

Por fin, llegamos al Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de IA y por el que se modifican los Reglamentos (CE) n° 300/2008, (UE) n° 167/2013, (UE) n° 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas

²⁹ ARÁNGUEZ SÁNCHEZ, T., «Sesgos sexistas de los algoritmos e inteligencia artificial», en Aránguez Sánchez, T., Olariu, O. (coord.), *Algoritmos, teletrabajo y otros grandes temas del feminismo digital*, Dykinson, Madrid 2022, p. 79.

³⁰ RODRÍGUEZ FERNÁNDEZ, M. L., «Inteligencia Artificial, Género y Trabajo», cit., p. 14.

³¹ SERRA-NAVARRO, D., «Inteligencia Artificial, vampirismo y la caja negra: aproximación especulativa en el marco del aprendizaje automático», cit., p. 969.

2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (el llamado Reglamento de IA, y que, en adelante, denominaremos como RIA), que define, en su artículo 3.1., el término “Sistema de IA” como

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

Esta definición hace referencia especialmente a la capacidad de adaptación y aprendizaje, así como a las capacidades predictivas, de toma de decisiones de manera auxiliar o incluso autónoma (donde es el propio sistema quien, de acuerdo con un algoritmo, toma todas las decisiones), y de influencia en su entorno inmediato.

La revolución tecnológica que actualmente estamos experimentando ha determinado que exista una enorme disponibilidad de datos y una enorme capacidad para procesarlos rápidamente, lo que ha desencadenado la expansión de la IA y su utilización en numerosos ámbitos, también en la gestión de los recursos humanos de la empresa³².

Más adelante, en este trabajo, se hará referencia a las cautelas y limitaciones que introduce este Reglamento a la utilización en la empresa de Sistemas basados en algoritmos de IA, y las obligaciones respecto de éstos y, en particular, en lo que puede afectar a los trabajadores, y que va mucho más allá de la igualdad y no discriminación, pudiendo afectar a otros derechos fundamentales de los trabajadores.

3.2. Inteligencia artificial y toma de decisiones automatizadas en el ámbito laboral

Adaptando las definiciones que hemos venido estudiando al contexto laboral, podríamos considerar a la IA como un conjunto de sistemas capaces de procesar grandes volúmenes de datos, identificar patrones y realizar predicciones o recomendaciones que influyen en decisiones relativas a la gestión de personas.

A diferencia de los sistemas informáticos tradicionales, estos modelos no se limitan a ejecutar instrucciones preprogramadas, sino que aprenden a partir de datos

³² RODRÍGUEZ FERNÁNDEZ, M. L., «Inteligencia Artificial, Género y Trabajo», cit., p. 14.



históricos, ajustando su comportamiento mediante técnicas de aprendizaje automático (el proceso conocido como *machine learning*). En este sentido, el sometimiento exclusivo de las decisiones laborales a los algoritmos se traduciría en posibles efectos negativos sobre el ejercicio de los derechos fundamentales de los trabajadores. Por eso, la IA puede crear oportunidades de mejora de las condiciones laborales, así como incentivar el empleo decente, pero también plantea importantes retos y dilemas éticos, genera una responsabilidad de las empresas que va a necesitar de una regulación que encuentre un equilibrio entre la necesidad de optimizar los procesos y la protección y respeto de los derechos fundamentales de los trabajadores³³.

Así, por una parte, la IA se utiliza en una gran variedad de aplicaciones y herramientas para el trabajo asistido y el análisis de datos, lo que permite la automatización de tareas cada vez más complejas, así como la toma de decisiones y la gestión automatizadas o semiautomatizadas del lugar de trabajo³⁴.

En la relación laboral se puede utilizar la IA en la contratación laboral, bien como mecanismo de selección de personal más eficaz, precisa y rápida (análisis y recopilación de curriculum, etc), bien como método de descarte automatizado de candidatos sin necesidad de pasar por un proceso personal, cara a cara, con el responsable de reclutamiento correspondiente³⁵.

El tratamiento de datos y la IA pueden contribuir a una mayor agilidad y eficiencia en los procesos de colocación, selección y contratación de personal, acortando la duración de estos procesos y la toma de decisiones³⁶.

La IA, además, acompaña todo el proceso de vigencia de la relación laboral, se puede utilizar, así, para el control de tiempos y desempeño, el examen de los niveles de absentismo, para la prevención de riesgos laborales, para los procesos de vigilancia y

³³ TOYAMA MIYAGUSUKU, J., «La Inteligencia Artificial y las Relaciones Laborales», en *Revista Ius Et Veritas* 69 (2024), p. 71.

³⁴ MACÍAS GARCÍA, M. C., «La inteligencia artificial para el entorno laboral. Un enfoque en la predicción de accidentes», en *e-Revista Internacional de la Protección Social* (2022), p. 88.

³⁵ PÉREZ LÓPEZ, J. I., «Inteligencia Artificial y contratación laboral», cit., p. 188.

³⁶ OLARTE ENCABO, S., «La aplicación de inteligencia artificial a los procesos de selección de personal y ofertas de empleo: impacto sobre el derecho a la no discriminación», en *Doctrina Laboral* 119 (2020), p. 83.

control de los trabajadores, e incluso para la prueba la hora de justificar un despido, etc. También para la gestión administrativa ordinaria de todos los procesos³⁷.

Durante el desarrollo de la relación laboral, estos sistemas pueden intervenir en la evaluación del desempeño, la selección y detección de necesidades de formación, la asignación de promociones, la identificación de personas “de alto potencial”, la selección para programas de formación o la medición de la productividad³⁸. En este ámbito, hemos de tener muy en cuenta que la utilización de datos personales en el control y vigilancia de los trabajadores puede dar lugar a graves afectaciones a derechos fundamentales de los trabajadores como los derechos a la intimidad, privacidad y no discriminación.

Finalmente, en la fase de extinción, algunos modelos se utilizan para detectar riesgos de rotación, recomendar despidos o identificar perfiles considerados “menos adaptables” a las necesidades empresariales.

En este sentido, y dada su capacidad predictiva, pueden determinar, en función de patrones observados en el trabajador (o del candidato), cuál será su respuesta futura ante determinadas situaciones y órdenes empresariales, lo que -puede condicionar las decisiones de la empresa tanto en el momento de su contratación, como para conceder incentivos o incluso sanciones que pueden llegar hasta el despido³⁹.

Podemos ver aspectos positivos en la aplicación de la IA, como pueden ser el aumento de la seguridad laboral, el estudio de la flexibilidad de horarios; e incluso la supervisión del cumplimiento de derechos laborales, así como la supresión de actividades repetitivas y monótonas que no agregan valor⁴⁰.

Sin embargo, La creciente automatización de estas decisiones plantea un desplazamiento progresivo de la valoración humana hacia sistemas técnicos cuya lógica interna no siempre resulta comprensible ni transparente. Este fenómeno se ve agravado cuando las decisiones adoptadas por la IA se perciben como objetivas o neutrales por el mero hecho de estar basadas en cálculos matemáticos, obviando que dichos cálculos se

³⁷ Vid. ASQUERINO LAMPARERO, M. J., «Algoritmos y discriminación», en Morales Ortega, J. M. (ed.), *Realidad social y discriminación: estudios sobre diversidad e inclusión laboral*, Laborum, Murcia 2022, p. 336.

³⁸ Vid. ÁLVAREZ FERNÁNDEZ, E., «Sobre las decisiones automatizadas en el ámbito laboral y el derecho a la privacidad de los trabajadores», en *Diario La Ley* 10587 (2024).

³⁹ ASQUERINO LAMPARERO, M. J., «Algoritmos y discriminación», cit., p. 340.

⁴⁰ TOYAMA MIYAGUSUKU, J., «La Inteligencia Artificial y las Relaciones Laborales», cit., p. 70.



sustentan en elecciones humanas previas relativas a los datos utilizados, las variables consideradas y los objetivos perseguidos.

En este sentido, las Directrices Éticas para una IA Fiable, presentadas por el Grupo de Expertos de Alto Nivel sobre la IA de la Comisión Europea el 8 de abril de 2019 exigían que los sistemas de IA fueran transparentes y explicables, de manera que las personas trabajadoras puedan conocer el proceso de toma de decisiones que puedan afectar a sus condiciones laborales, al tiempo que enfatizaban la necesidad de intervención y supervisión humanas en los sistemas de IA que se utilizan por las empresas en el ámbito laboral. Además, se insiste en el derecho de los trabajadores a conocer cómo se utilizan sus datos en los sistemas de IA, y a que se asegure que su privacidad sea respetada y que los datos se manejen de manera ética y segura.⁴¹

3.3. Sesgos en la gestión algorítmica del trabajo

Los sesgos algorítmicos constituyen uno de los principales riesgos asociados al uso de la IA en el empleo. Lejos de ser anomalías excepcionales, estos sesgos son inherentes a sistemas que aprenden a partir de datos sociales, los cuales reflejan desigualdades históricas y estructuras de poder preexistentes.

En este sentido, es importante no dejar de lado el hecho de que un algoritmo es una creación humana, por lo tanto, debe ser entendida como una secuencia de instrucciones, desarrollada por humanos, que se encargan de realizar tareas y encontrar soluciones de forma automatizada. Por esto mismo, puede reproducir y arrastrar los estereotipos que estas personas, las que elaboran las instrucciones, pueden tener, dada la sociedad o la cultura a la que pertenecen⁴².

La gestión algorítmica del trabajo está basada en una serie de aplicaciones de IA que pueden, en primer lugar, generar datos sobre la prestación laboral de un servicio o la elaboración de uno o varios productos; también para captar esos datos, ya sean tareas, imágenes, patrones, sonidos, desplazamientos o valores de cualquier clase; o para procesarlos y elaborar de informes de perfiles o de cualquier otro tipo; y al fin, para tomar

⁴¹ TERUEL FERNÁNDEZ, C., «Derecho de Información de los Trabajadores ante la toma de decisiones automatizadas», en *Revista Internacional y Comparada de Relaciones Laborales Y Derecho Del Empleo* 4 (2024), p. 190.

⁴² CASTILLO CASTRO, K. P., «El sesgo algorítmico generado por la inteligencia artificial como acto de discriminación en las relaciones de trabajo», en *Revista de Derecho Procesal del Trabajo* 8 (2025), p. 79.

decisiones sobre trabajo, asignación de lugar y puesto, cambios, despidos, y dar instrucciones y órdenes sobre las tareas laborales, controlarlas o seleccionar personal o descartarlo, entre otros usos⁴³.

Este análisis de datos basados en patrones de comportamiento (movimientos, gestos, hábitos, compras, búsquedas en internet, actividad en redes sociales, datos de localización, etc.) de los individuos pueden ayudar a los empleadores a obtener información, en apariencia objetiva, sobre las personas que van a desempeñar un puesto, antes incluso de contratarlas.

La aplicación de estas técnicas tiene indudables implicaciones éticas y riesgo de intromisión en los derechos fundamentales del individuo, pudiendo introducir sesgos que den lugar a decisiones injustas y a desigualdades, incluso a discriminaciones por razón de sexo, raza u otras condiciones personales o sociales, por lo que han de ser utilizadas con las debidas cautelas y, especialmente, con la máxima transparencia⁴⁴.

Así lo refleja el Parlamento Europeo, al establecer que la baja calidad o el origen dudoso de los datos

“podrían dar lugar a algoritmos sesgados, correlaciones falsas, errores, una subestimación de las repercusiones éticas, sociales y legales, el riesgo de utilización de los datos con fines discriminatorios o fraudulentos y la marginación del papel de los seres humanos en esos procesos, lo que puede traducirse en procedimientos deficientes de toma de decisiones con repercusiones negativas en las vidas y oportunidades de los ciudadanos”⁴⁵.

⁴³ OTTAVIANO, J. M., «La amenaza fantasma: Inteligencia Artificial y derechos laborales», en *Nueva Sociedad* 294 (2021), p. 88.

⁴⁴ Según la Real Academia Española, el sesgo se define como «Error sistemático en el que se puede incurrir cuando al hacer muestreos o ensayos se seleccionan o favorecen unas respuestas frente a otras» (sesgo, sesga | Definición | Diccionario de la lengua española | RAE - ASALE). Por su parte, Sánchez Caparrós lo define como «aquella inclinación o prejuicio de una decisión realizada por un sistema inteligente que afecta a una persona o a un grupo determinado de una forma que se considere injusta y razonable» (Sánchez Caparrós, M., «Los riesgos de la inteligencia artificial para el principio de igualdad y no discriminación. Planteo de la problemática y algunas aclaraciones conceptuales necesarias bajo el prisma del Sistema Interamericano de Derechos Humanos», en *Revista elDial.com, Contenidos Jurídicos*, Editorial Albreumatica S. A., 2022).

⁴⁵ Parlamento Europeo: Considerando m). «[L]a información revelada por los análisis de los macrodatos no ofrece una visión general objetiva e imparcial de ninguna materia y que es tan fiable como lo permitan los datos subyacentes» (Resolución del Parlamento Europeo, de 14 de marzo de 2017, sobre las implicaciones de los macrodatos en los derechos fundamentales: privacidad, protección de datos, no discriminación, seguridad y aplicación de la ley (2016/2225(INI))).



Desde un punto de vista conceptual, los sesgos algorítmicos pueden surgir en distintas fases del ciclo de vida de un sistema de IA.

En primer lugar, pueden derivarse de los propios datos de entrenamiento. Si los conjuntos de datos utilizados reflejan una infrarrepresentación de determinados colectivos, como pueden ser las mujeres con convicciones religiosas visibles, el sistema aprenderá que dichos perfiles son atípicos o poco deseables, reproduciendo patrones de exclusión previos.

En segundo lugar, los sesgos pueden introducirse en la fase de diseño del algoritmo, cuando se seleccionan determinadas variables como indicativas de éxito, productividad o adecuación cultural, sin tener en cuenta su impacto discriminatorio indirecto.

Finalmente, los sesgos también pueden aparecer en la fase de interpretación y aplicación de los resultados, cuando las recomendaciones algorítmicas se aceptan acríticamente como decisiones finales.

Los datos son el elemento central y básico para la IA, pues conforman la materia prima que se utiliza para automatizar los procesos de aprendizaje en el que los sistemas de IA se entrenan para realizar sus predicciones⁴⁶.

Los modelos, pues, aprenden a partir de los datos. Existen a partir de aquí diferentes procesos, el más común es el llamado “aprendizaje supervisado”, que implica analizar los datos para buscar patrones que los distingan y los permitan agrupar por categorías. Durante el aprendizaje, el sistema tomará los datos y hará predicciones partiendo de ellos, asociando la información con las diferentes etiquetas. Estas asociaciones serán mejores a medida que avance el proceso de aprendizaje. Es esencial, en el proceso, la capacidad de generalización para poder predecir. Para ello, al algoritmo se le han de añadir una serie de datos de prueba. Estos datos serán similares a los que el sistema analizará, pero no los mismos. Estas variaciones determinarán la calidad de las predicciones.

En el ámbito laboral, estos sesgos adquieren una especial gravedad, ya que las decisiones adoptadas pueden afectar al acceso al empleo, a la estabilidad laboral y al

⁴⁶ FERRANTE, E., «Inteligencia Artificial y sesgos algorítmicos: ¿Por qué deberían importarnos?», en *Nueva Sociedad* 294 (2021), p. 29.

desarrollo profesional. Además, la automatización contribuye a ocultar el carácter discriminatorio de dichas decisiones, al diluir la responsabilidad individual y dificultar la identificación de los factores concretos que han conducido a un resultado desfavorable.

En resumen, los datos son un reflejo de la realidad, del ahora. Por ello, cabe preguntarse: ¿realmente queremos que el modelo perpetúe (o agrave, incluso) las desigualdades, por el solo hecho de que vivimos en una sociedad desigual o queremos modificar esta realidad?

La población utilizada para construir las muestras, las variables que se miden son decisiones humanas, decisiones que no son neutrales. No es necesario ni siquiera que este proceso sea voluntario, la mayoría de las veces ni siquiera hay consciencia o ideología detrás. Simplemente, los algoritmos son creados por personas que tienen sus prejuicios adquiridos a lo largo de su experiencia, que pueden filtrarse en el diseño y la definición de criterios de evaluación para estos modelos⁴⁷.

Cuanta mayor sea la diversidad en los equipos de diseño e implementación de la IA, más justos y equitativos serán sus resultados.

El almacenamiento de datos a largo plazo supone también riesgos adicionales, ya que en el futuro esa información podría aprovecharse de forma inadecuada⁴⁸.

Normalmente un algoritmo dotará de más valor a las decisiones tomadas con la mayor información disponible. Por lo que, y dado que en el caso de las minorías –raza, género, orientación sexual, etc.- hay menos datos disponibles el algoritmo deducirá que tomar una decisión favorable a alguien que pertenece a un colectivo minoritario tendrá más probabilidad de error que si se toma la decisión a favor de alguien perteneciente al colectivo mayoritario, por lo que exigirá más a aquellos trabajadores procedentes de minorías⁴⁹.

Los sesgos generados pueden provenir de distinto tipo de datos, tales como textos, redes sociales, ubicaciones de GPS e incluso imágenes. La IA parte de dichos datos

⁴⁷ *Ivi*, p. 34.

⁴⁸ GRIGORE, A. E., «Derechos humanos e inteligencia artificial», cit., p. 169.

⁴⁹ TODOLÍ SIGNES, A., «En cumplimiento de la primera Ley de la robótica: Análisis de los riesgos laborales asociados a un algoritmo/inteligencia artificial dirigiendo el trabajo», en *Labour & Law Issues* 5 (2019), p. 16.



desarrollando un aprendizaje automático que no tiene en cuenta estereotipos, ética, prejuicios, o valores⁵⁰.

Tampoco tiene en cuenta, en consecuencia, el respeto a los derechos fundamentales de los trabajadores. Los sistemas de IA pueden arrastrar valores sociales “incrustados” en su diseño y que éstos sean contrarios a la igualdad, o a otros principios constitucionales y derechos fundamentales de las personas⁵¹.

Esto es lo que lleva a autores como O’Neil a hablar de los algoritmos como “*opiniones integradas en matemáticas*”⁵² dado que arrastran una enorme cantidad de prejuicios ideológicos.

Los algoritmos, aunque a veces les atribuimos características humanas, como el aprendizaje, en realidad no tiene consciencia ni experiencia subjetiva de su comportamiento, ni se les puede atribuir intencionalidad por sí mismos. Sin embargo, en muchas ocasiones y si no tienen supervisión humana, nos pueden conducir a decisiones injustas⁵³.

Un elemento particularmente problemático es el uso de algoritmos de “caja negra”, es decir, sistemas cuyo funcionamiento interno resulta opaco incluso para sus propios desarrolladores. La falta de explicabilidad impide comprender qué variables han sido determinantes en una decisión concreta, obstaculizando tanto la detección de sesgos como el ejercicio efectivo de mecanismos de impugnación por parte de las personas afectadas.

Por eso la discriminación que se produce mediante este tipo de algoritmo genera en la persona trabajadora situaciones de especial indefensión, pues es totalmente indetectable e inevitable, y puede ocurrir incluso antes del comienzo de la relación laboral, dificultando la determinación de la legislación aplicable⁵⁴.

⁵⁰ Vid. FERRANTE, E., «Inteligencia Artificial y sesgos algorítmicos: ¿Por qué deberían importarnos?», cit., pp. 27-36.

⁵¹ COTINO HUESO, L., «Big data e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales», en *Dilemata* 24, p. 134.

⁵² O’NEIL, C. H., *Armas de Destrucción Matemática. Cómo el Big Data aumenta la desigualdad y amenaza la democracia*, Capitán Swing, Madrid 2017.

⁵³ MORANDÍN-AHUERMA, F., «Conciencia e inteligencia artificial Heidegger, Searle y Bostron», en *STOA* 14 (2023), p. 204.

⁵⁴ MORALES ARAGÓN, A., «Prácticas de IA prohibidas en el ámbito laboral. Análisis y posibles soluciones desde la normativa laboral española», en *Labos* 6 (2025), p. 192.

Esto es así porque la decisión del algoritmo queda oculta no solo para la persona trabajadora que se ve afectada por la decisión, sino hasta para el implementador del sistema y para el propio desarrollador, y para el empresario usuario, que solamente se limita a realizar una serie de preguntas y el algoritmo procesa la información que tiene y toma decisiones. Por lo tanto, en estos casos concretos, ¿quién se hará responsable de una mala decisión adoptada por un algoritmo que utiliza un modelo de autoaprendizaje opaco o cuyos parámetros ya nadie recuerda?⁵⁵

Esta pregunta es particularmente pertinente cuando se están afectando derechos fundamentales.

Por eso, con relación a este tipo de algoritmos, el Consejo de Europa, a través del Comité de expertos sobre Internet MSI-NET, en el documento “*Algorithms and human rights. Study on the human rights dimensions of automated data processing techniques (in particular algorithms) and possible regulatory implications*”, deja escrito que:

“El uso de un sistema de IA en cualquier proceso de toma de decisiones con un impacto significativo en los derechos humanos de las personas necesita ser identificable. El uso de un sistema de IA no solo deber ser dado a conocer públicamente en términos claros y accesibles, sino que las personas individuales deben ser capaces de comprender cómo se adoptan las decisiones y cómo se han verificado esas decisiones⁵⁶”.

Es necesario, en consecuencia, auditar y supervisar los modelos algorítmicos y los datos que los alimentan, con la idea de corregir sesgos de género y otras causas de discriminación, de manera que se evalúen de manera correcta y rigurosa los resultados obtenidos y la manera en que estos afectan a determinados grupos sociales⁵⁷. En el caso que nos ocupa, a la doble discriminación por ser mujer y por profesar y practicar una confesión religiosa lo que, además, manifiesta de manera visible.

⁵⁵ GALDON CLAVELL, G., ORTEGA SÁNCHEZ, H., «Abrir la caja negra. Digitalización y participación en el Trabajo», en Moreno Izquierdo, J. A., Estrada López, B. (coord.), *¿Una empresa de todos? La participación del trabajo en el gobierno corporativo*, Plataforma por la Democracia Económica y Libros de la Catarata, Madrid 2022.

⁵⁶ Consejo de Europa, *Algorithms and human rights. Study on the human rights dimensions of automated data processing techniques (in particular algorithms) and possible regulatory implications*, Committee of Experts on Internet Intermediaries (MSI-NET), DGI(2017)12 (2018) (https://ec.europa.eu/futurium/en/system/files/ged/algorithms_and_human_rights_en_final.pdf_1.pdf), (consultado el 14 de enero de 2026).

⁵⁷ BLASCO JOVER, C., «IA: dentro de la flexibilidad laboral y de un balance trabajo-vida», en *Revista Justicia & Trabajo* 4 (2024), p. 85.



Para ello, y citando el mencionado documento del Consejo de Europa, existe una obligación de transparencia que puede consistir en *“la divulgación pública de información sobre el sistema en cuestión, sus procesos, los efectos directos e indirectos sobre los derechos humanos, y las medidas tomadas para identificar y mitigar los impactos adversos del sistema en los derechos humanos”*⁵⁸ en particular si son utilizados por la administración pública.

En este contexto, el art. 35 Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo de 27 de abril de 2016 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 dispone que

cuando sea probable que un tipo de tratamiento, en particular si utiliza nuevas tecnologías, por su naturaleza, alcance, contexto o fines, entrañe un alto riesgo para los derechos y libertades de las personas físicas, el responsable del tratamiento realizará, antes del tratamiento, una evaluación del impacto de las operaciones de tratamiento en la protección de datos personales”.

Más concretamente cuando se trate de una

“evaluación sistemática y exhaustiva de aspectos personales de personas físicas que se base en un tratamiento automatizado, como la elaboración de perfiles, y sobre cuya base se tomen decisiones que produzcan efectos jurídicos para las personas físicas o que les afecten significativamente de modo similar”,

incluirá, como mínimo,

“una descripción sistemática de las operaciones de tratamiento previstas y de los fines del tratamiento, inclusive, cuando proceda, el interés legítimo perseguido por el responsable del tratamiento; una evaluación de la necesidad y la proporcionalidad de las operaciones de tratamiento con respecto a su finalidad; una evaluación de los riesgos para los derechos y libertades de los interesados; y las medidas previstas para afrontar los riesgos, incluidas garantías, medidas de seguridad y mecanismos que garanticen la protección de datos personales, teniendo en cuenta los

⁵⁸ Consejo de Europa, *Algorithms and human rights. Study on the human rights dimensions of automated data processing techniques (in particular algorithms) and possible regulatory implications*, cit.

derechos e intereses legítimos de los interesados y de otras personas afectadas”.

4. LA DOBLE DISCRIMINACIÓN DE LA MUJER POR MOTIVOS RELIGIOSOS EN EL ÁMBITO LABORAL

4.1. Inteligencia artificial y la Perspectiva de discriminación interseccional

Para comprender adecuadamente el impacto de los sesgos algorítmicos sobre las mujeres con convicciones religiosas visibles, resulta imprescindible, como ya se ha adelantado, adoptar una perspectiva de discriminación interseccional, en los términos de la Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación (que, en su artículo 6.3 apartado b la define como aquella que tiene lugar “*cuando concurren o interactúan diversas causas de las previstas en esta ley, generando una forma específica de discriminación*”).

La Ley 3/2023, de 28 de febrero, de Empleo alude a que los posibles programas de acción positiva se reforzarán cuando se produzcan situaciones de interseccionalidad (artículo 50.1). También se tendrá en cuenta dicho enfoque cuando se elabore el perfil de la persona usuaria.

Para que se produzca este tipo de discriminación no basta con que tenga origen en dos o más elementos, sino que es preciso, además, que esas identidades sociales discriminadas se interrelacionen intrínsecamente en un individuo o grupo de individuos, como es el caso de mujer y persona con signos religiosos visibles⁵⁹.

Obviamente, entendemos que esa combinación de dos o más factores, al interactuar, generan o aumentan la situación de discriminación. La discriminación en estos casos se hace más aguda y las consecuencias sufridas, por tanto, tienen un riesgo mayor⁶⁰.

⁵⁹ CAVALCANTE CARVALHO, A. M., «Discriminación interseccional: concepto y consecuencias en la incidencia de violencia sexual contra mujeres con discapacidad», en *Journal of Feminist, Gender and Women Studies* 7 (2018), p. 16.

⁶⁰ LÓPEZ SÁNCHEZ, C., VILASECA GARCÍA, C., SERRANO JAPA, J. M., «Interseccionalidad: la discriminación múltiple desde una perspectiva de género», en *Revista Crítica de Historia de las Relaciones Laborales y de la Política Social* 14 (2022), p. 77.



Este enfoque parte de la premisa de que las personas no experimentan la discriminación de manera aislada en función de un único motivo, sino a través de la interacción simultánea de múltiples ejes de desigualdad, como el sexo, la religión, el origen étnico o la clase social.

La consideración, conjunta y por separado, de estos factores de desigualdad, nos permitirá identificar con mayor claridad las dificultades que sufre cada mujer en su caso específico, lo que permite buscar recursos y ofrecer soluciones mucho más específicas y concretas⁶¹.

Y así, ciñéndonos al ámbito laboral, la interseccionalidad nos va a hacer notar cómo determinadas experiencias de discriminación quedan fuera de los marcos jurídicos tradicionales, que tienden a analizar los motivos de forma fragmentada. Así, una mujer puede no ser discriminada únicamente por ser mujer o únicamente por su religión, sino por la combinación específica de ambas circunstancias, lo que tenderá a generar sus propias dinámicas de exclusión.

La inteligencia artificial, lejos de neutralizar estas interacciones, puede amplificarlas. Los algoritmos no operan sobre categorías jurídicas abstractas, sino sobre correlaciones estadísticas entre múltiples variables. De este modo, características aparentemente neutras como el nombre, el lugar de residencia, la participación en determinadas actividades comunitarias o la vestimenta visible pueden dar lugar a discriminaciones indirectas difíciles de detectar. En el caso de las mujeres con convicciones religiosas visibles, esta lógica algorítmica puede traducirse en una penalización sistemática de perfiles que no se ajustan a los modelos dominantes de trabajador “ideal”, caracterizados por una disponibilidad horaria plena, una apariencia neutralizada y una desvinculación de compromisos identitarios visibles. La interseccionalidad permite, así, comprender por qué los sistemas de IA pueden generar efectos particularmente excluyentes sobre este colectivo, incluso cuando no incorporan explícitamente variables relacionadas con el género o la religión.

La conjunción entre inteligencia artificial, sesgos algorítmicos y discriminación interseccional plantea importantes desafíos para el Derecho del Trabajo y para los marcos normativos de igualdad. En particular, cuestiona la suficiencia de los enfoques

⁶¹ *Ivi*, p. 78.

tradicionales de no discriminación, basados en la identificación de un motivo único y en la prueba de una intención discriminatoria directa.

La automatización de decisiones laborales exige repensar conceptos como la discriminación indirecta, la carga de la prueba y la responsabilidad empresarial, así como reforzar mecanismos de transparencia y control sobre los sistemas utilizados.

Como primera conclusión, podríamos afirmar que no debe tomarse decisiones laborales automatizadas basadas solo en algoritmos como promociones, bonos, renovaciones de contratos o despidos. Es necesaria la intervención “humana” para evitar los sesgos y las discriminaciones⁶².

Asimismo, la empresa deberá establecer sistemas capaces de detectar y combatir los sesgos o desviaciones de la implementación de IA que generen vulneraciones a los derechos de los trabajadores. Un cauce adecuado para negociar y establecer estos sistemas dado participación a los trabajadores lo constituye la negociación colectiva.

Así, en la negociación de convenios colectivos, al margen de poder pactar diferentes limitaciones en el uso de la IA, se puede pactar la transparencia de información referida a la implementación de esta en las empresas.

4.2. El sexo y las convicciones religiosas como fuentes de discriminación y su reflejo en los sesgos algorítmicos

La distribución de espacios y tareas laborales entre hombres y mujeres ha venido resultando históricamente en que el trabajo asociado al espacio privado doméstico sea asignado a las mujeres y el trabajo en el espacio público y productivo sea asociado a los hombres⁶³.

En líneas generales, podemos decir que las diferencias biológicas (sexo) y las diferentes condiciones laborales y el diferente trato ofrecido a hombres y mujeres por la sociedad pueden afectar a las relaciones laborales.

La presencia de signos visibles de práctica de alguna religión, y los derechos asociados a la libertad religiosa y su ejercicio, añade una nueva fuente de sesgos discriminatorios que se pueden arrastrar en los algoritmos de IA.

⁶² TOYAMA MIYAGUSUKU, J., «La Inteligencia Artificial y las Relaciones Laborales», cit., p. 69.

⁶³ VEGA CASADO, M. T., «Análisis de la prevención de riesgos laborales desde perspectiva de género en España», en *Revista de la Escuela Jacobea de Posgrado* 26 (2024), p. 62.



Como ya hemos abordado en el presente trabajo, en muchas ocasiones, la utilización de estos algoritmos tanto para ayudar en la toma de decisiones como para tomar decisiones por sí mismos, pueden arrastrar sesgos que perpetúen las desigualdades y constituyan, por sí mismas, barreras en el acceso al empleo. En consecuencia, es necesario que exista una intervención humana que garantice que el proceso ha estado libre de errores y de sesgos, y se realiza en condiciones de igualdad y de no discriminación.

Es importante entender que, sin perjuicio de la delimitación del alcance de la responsabilidad, habrá discriminación donde haya diferencia de trato basada en una causa protegida, aun cuando esta conducta o acto no sean intencionados, esto es, no se busque la discriminación. Por ello, abarcará todos los efectos derivados de sesgos algorítmicos, se trate de sesgos buscados o de sesgos accidentales derivados de la inferencia de datos⁶⁴.

Estos algoritmos se utilizan en selección de personal en tareas como entrevistas previas, contacto con aspirantes, realización de pruebas, emisión de informes y, especialmente, cribado de currículos. Por lo tanto, una persona define el perfil requerido para un puesto de trabajo (qué cualidades se precisan, qué tipo de aptitudes profesionales y hasta personales se requieren), y los algoritmos, en función de esos parámetros, serán capaces de predecir cuál es el trabajador idóneo. A partir de ahí, podrán servir como apoyo o incluso tomar las decisiones por sí mismos.

También se utilizan, ya en el transcurso de la relación laboral en la toma de numerosas decisiones que afectan a las personas trabajadoras, como detección de necesidades de formación, selección de personas para movilidad geográfica, movilidad funcional o ascensos, e incluso suspensiones de contrato o despidos.

Otros procesos en los que las empresas pueden utilizar estos algoritmos son la prevención de riesgos laborales y las actividades de control, seguimiento y evaluación del desempeño de los trabajadores.

En todos estos ámbitos la utilización de la IA proporciona una apariencia fría y objetiva, sin el riesgo aparente de que se toman las decisiones por sensaciones, intuiciones o ideas preconcebidas de las personas. En este sentido, el sistema de IA, ciertamente, puede evitar decisiones basadas en corazonadas, primeras impresiones o prejuicios. El

⁶⁴ RIVAS VALLEJO, P., «Sesgos de género en el uso de inteligencia artificial para la gestión de las relaciones laborales: análisis desde el derecho antidiscriminatorio», en *e-revista Internacional de la Protección Social* VII (2022), p. 58.

problema, como sabemos a estas alturas, es que los algoritmos de IA también arrastran sesgos e ideas preconcebidas, sea en su diseño, sea en los datos con que han sido entrenados, sea en su propia aplicación. Muchas veces, como ya se ha visto, ni siquiera se puede saber exactamente los motivos o el mecanismo por el que el algoritmo ha tomado determinadas decisiones.

Esto es particularmente relevante en los procesos anteriores a la contratación, además, están (en la empresa privada) muy poco regulados, son espacios no contemplados expresamente por la normativa, ni por la negociación colectiva, existe escasa intervención de Inspección de Trabajo y prácticamente, en consecuencia y salvo excepciones, resoluciones judiciales dictados en esta fase⁶⁵.

Ocurre, sin embargo, que pueden aparecer sesgos discriminatorios desde el mismo diseño de la oferta de empleo, y del propio diseño del algoritmo utilizado para gestionar estos y otros procesos de tomas de decisiones en el ámbito laboral. Por ejemplo, el sistema puede dejar fuera candidatos por determinadas características físicas y psicológicas o de la mera apariencia, tales como el origen étnico, emociones y bienestar, símbolos visibles, que el algoritmo interprete como señales de no idoneidad de los candidatos. Como también podría ocurrir con otras características, como gestos, tono de voz y otros que hagan suponer al sistema de IA que está mintiendo o que oculta algo, o que, simplemente, no se adapta a la cultura o valores existentes en la empresa.

Evidentemente, en los procesos de selección y otros procesos de toma de decisiones en la empresa tradicionales también se valoran elementos que informan sobre características ajenas a la dimensión laboral y que también son susceptibles de ocasionar un trato discriminatorio por parte del humano-seleccionador, y las personas que toman las decisiones o las aconsejan y, en fin, el proceso, arrastran sus propios prejuicios y sesgos. En estos casos, al menos, deberían intervenir, como en el diseño de los algoritmos, varias personas, que constituyan una muestra diversa y que tengan formación específica.

En todo caso, la capacidad de tomar decisiones discriminatorias o basadas en datos obtenidos con violación del derecho a la intimidad son mucho mayores si se cuenta con sistemas de IA, capaces de recabar, manejar y procesar enormes cantidades de datos, y tomar decisiones a partir de los mismos, conforme con los parámetros introducidos.

⁶⁵ ASQUERINO LAMPARERO, M. J., «Algoritmos y discriminación», cit., p. 362.



Además, las decisiones adoptadas mediante algoritmos se presentan como una verdad científica y, por ello mismo, objetiva, impersonal⁶⁶, lo que dificulta la asunción de responsabilidades respecto de las decisiones sesgadas a los sujetos que las cometen, dado que los sesgos y los responsables de haber adoptado las decisiones correspondientes quedan ocultos detrás de una fórmula matemática⁶⁷.

Por ello, es especialmente preocupante la falta de transparencia, puesto que esa opacidad propia de los mecanismos automatizados incrementa las habituales dificultades en la reclamación contra las decisiones empresariales en vía judicial⁶⁸.

4.3. Cautelas necesarias para la utilización de la IA en la toma de decisiones empresariales que puedan afectar a las personas trabajadoras

La gestión algorítmica de la empresa puede generar importantes dilemas de carácter ético e incluso afectar a los derechos fundamentales de los trabajadores.

Más allá de ese modelo ideal, ese que presenta una IA puesta al servicio de la mayor eficacia, rapidez y objetividad en los procesos de empresariales de toma de decisiones sobre las personas trabajadoras, que garantice la decisión adecuada sin entrar en impresiones o valoraciones subjetivas, es muy frecuente la utilización de la IA para detectar no solo situaciones de valoración positiva de las trabajadoras, sino sobre todo conductas indebidas, y ello utilizando datos y realizando un tratamiento algorítmico de los mismos.

Tales datos podrían ser datos biométricos, análisis de gestos capaces de recoger estados de ánimo, lugares visitados, problemas de carácter personal o familiar, bajas o ausencias al trabajo, entre otros datos que pueden ser invasivos para la intimidad.

En estos casos, cuando se realice un tratamiento de datos de carácter personal de los trabajadores se tendrá en cuenta la Ley de Protección de Datos y sus obligaciones respecto de datos que vayan más allá de los imprescindibles para el cumplimiento del contrato y desempeño de la actividad, o sean datos de especial protección⁶⁹.

⁶⁶ O'NEIL, C. H., *Armas de Destrucción Matemática. Cómo el Big Data aumenta la desigualdad y amenaza la democracia*, cit.

⁶⁷ ARÁNGUEZ SÁNCHEZ, T., «Sesgos sexistas de los algoritmos e inteligencia artificial», cit., pp. 71-88.

⁶⁸ RIVAS VALLEJO, P., «Sesgos de género en el uso de inteligencia artificial para la gestión de las relaciones laborales: análisis desde el derecho antidiscriminatorio», cit., p. 54.

⁶⁹ RODRÍGUEZ SANZ DE GALDEANO, B., «Los sistemas de inteligencia artificial en el ámbito laboral y el marco regulador europeo de seguridad del producto», en Rodríguez Sanz de Galdeano, B., Egúsqiza

En líneas generales, y pese a sus indudables ventajas, los métodos de gestión digitalizados, incluida la utilización de algoritmos informáticos para coordinar y supervisar el trabajo puede dar lugar a que la dirección de los trabajadores dependa, en gran medida, de la IA, de forma que la empresa se esté, de alguna manera, refugiando detrás de simples algoritmos de caja negra, sin siquiera tener control inmediato sobre las decisiones que adopta mediante los mismos.

No podemos olvidar que, para alimentar de datos a la IA, se requiere de dispositivos que recaben los datos de cada persona, y algoritmos que procesen esa información y, eventualmente, decidan o aconsejen decisiones con base en la misma. Estos dispositivos representan un riesgo, de entrada, para la privacidad de quienes se encuentren sometidos a ese control laboral. Además, las instrucciones generadas por los algoritmos son capaces de asignar tareas a los más productivos, proponer premios y castigos y tomar otras decisiones para atraer o favorecer a algunos perfiles de trabajadores, animar a otros a que abandonen el proceso o seleccionar a quienes el modelo presupone como más adecuados, en función de los datos extraídos y el procesamiento de estos⁷⁰.

Estos aspectos constituyen riesgos de potenciales discriminaciones (errores de cálculo, ausencia de perspectiva de género, evaluación arbitraria o discriminación racial o por otras circunstancias, entre otras), si el sistema persigue exclusivamente la eficiencia, la contratación del trabajador más productivo y que menos problemas pueda suponer para la mera gestión empresarial considerada fríamente⁷¹.

Para evitar este riesgo, los criterios han de ser transparentes y conocidos. En particular si la empresa se puede basar en ellos para tomar decisiones que impliquen la toma de decisiones que determinen favorecer o elegir a unas personas y el perjuicio o rechazo de otras.

Con frecuencia la enorme capacidad de vigilancia, control y dirección de la IA va más allá de los límites de la privacidad, puede introducir sesgos que la hacen incurrir en

Balmaseda, M. A. (coord.), *Inteligencia Artificial y Prevención de Riesgos Laborales: Obligaciones y responsabilidades*, Tirant lo Blanch, Valencia 2023, p. 63.

⁷⁰ Vid. DE STEFANO, V., *Algorithmic Management and Collective Bargaining*, European Trade Union Institute, Bruselas 2021.

⁷¹ OTTAVIANO, J. M., «La amenaza fantasma: Inteligencia Artificial y derechos laborales», cit., p. 91.



graves discriminaciones, tiene efectos sobre el derecho a la intimidad y puede incidir en un debilitamiento de la protección social de las personas trabajadoras⁷².

La OIT ha mostrado su preocupación sobre este tema y propugna que la IA será usada “bajo control humano”, es decir, que las últimas decisiones que puedan afectar al trabajo no deben ser tomadas por algoritmos o programas sino por humanos encargados de ello⁷³.

Veremos un poco más adelante como el Reglamento Europeo que regula la IA establece, entre otras cautelas, la de la supervisión humana en determinados casos⁷⁴.

Por lo que hace referencia a la necesidad de transparencia, tras la actualización del Estatuto de los Trabajadores por el artículo único.1 del Real Decreto-ley 9/2021, de 11 de mayo, en vigor desde el 12/08/2021, respecto de los derechos de información y consulta y competencias de la representación de los trabajadores se añade la letra d) al apartado 4 del artículo 64 del citado Estatuto, que, textualmente, le otorga al Comité de Empresa el derecho a

“Ser informado por la empresa de los parámetros, reglas e instrucciones en los que se basan los algoritmos o sistemas de inteligencia artificial que afectan a la toma de decisiones que pueden incidir en las condiciones de trabajo, el acceso y mantenimiento del empleo, incluida la elaboración de perfiles”.

Como se ha adelantado, el empresario se va a encontrar también con obligaciones en materia de protección de datos derivadas de la utilización de IA. En este sentido, el Grupo de Trabajo del art. 29, actual Comité Europeo de Protección de Datos, destacaba en las Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679, de 2017, que:

“la elaboración de perfiles y las decisiones automatizadas pueden plantear riesgos importantes para los derechos y libertades de las personas que requieren unas garantías adecuadas. (...) La elaboración de perfiles puede perpetuar los estereotipos existentes y la segregación social. Asimismo, puede encasillar a una persona en una categoría específica y limitarla a las

⁷² Ivi, p. 88.

⁷³ ROMERO BURILLO, A. M., MORENO GENÉ, J., *Los nuevos escenarios laborales de innovación tecnológica*, Tirant Lo Blanch, Valencia 2023, p. 171.

⁷⁴ *Declaración del Centenario de la OIT para el futuro del trabajo*, adoptada en la 108ª Reunión de la Conferencia Internacional del Trabajo, celebrada en 2019.

preferencias que se le sugieren (...) lo que puede llevar a (...) una discriminación injustificada⁷⁵”.

En respuesta a esta preocupación, el art. 22 RPD se refiere a las “*decisiones individuales automatizadas, incluida la elaboración de perfiles*”, imponiendo ciertas exigencias y cautelas. En su apartado primero, el art. 22 RPD indica que “*todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado cuando pueda producir efectos jurídicos sobre su persona*”, como pudiera ser, en paradigmático ejemplo, decisiones que denieguen a una persona una oportunidad laboral o la coloquen en gran desventaja. Se exceptúa esta prohibición cuando el tratamiento sea necesario para la celebración o la ejecución de un contrato entre el interesado y un responsable del tratamiento.

Hemos de acotar en este punto la expresión “necesario”. El Grupo de Trabajo del art. 29 ha considerado que para entender que la toma de una decisión automatizada es necesaria, el responsable deberá demostrar que es la forma más apropiada para alcanzar el objetivo final, es decir, que no hay maneras menos intrusivas para alcanzar el resultado final, de manera que la intervención humana es imposible dado el elevadísimo número de datos procesados.

De todas formas, no cabe deducir sin más la nulidad de una decisión automatizada sin intervención humana, incluida la elaboración de perfiles, si se ha producido en el ámbito de una relación de trabajo o si ha sido consentida, a no ser que se trate de una opción totalmente descartada, de conformidad con el art. 9.1 LOPDyGDD. Esto es, cuando el tratamiento con inteligencia artificial vaya referido a datos relativos origen racial, étnico, opinión política, confesión religiosa o filosófica, afiliación sindical, datos genéticos, biométricos, de salud, la vida sexual o la orientación sexual de una persona física (los llamados datos especialmente protegidos que, en el ámbito de este trabajo, referido a discriminación por razones religiosas, es bastante claro que las decisiones automatizadas y sin supervisión humana van a resultar nulas en un alto porcentaje, cuando no todas).

⁷⁵ Grupo de trabajo sobre protección de datos del artículo 29, *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*, wp251rev_en (aepd.es).



El art. 22 RPD recuerda también que si se toman decisiones automatizadas fundadas únicamente en el tratamiento de datos (sin participación humana), el responsable ha de informar al afectado, aportando detalles sobre la lógica y explicando la importancia y consecuencias del proceso, con el fin de que este último pueda formular alegaciones o impugnar una decisión final eventualmente marcada por un sesgo o desviación de carácter discriminatorio.

Este derecho, según interpretación del Grupo de Trabajo del art. 29, implica la facultad de obtener información clara y simple (suficientemente comprensible) sobre el funcionamiento del proceso de elaboración de perfiles o adopción de decisiones de forma automatizada capaz de permitir formular impugnaciones.

Por fin, y dado que, a efectos del RIA, el empresario es “responsable del despliegue” del sistema de IA (conforme con la definición del artículo 3.4 de este reglamento) queda sujeto a las obligaciones que el RIA establece para éste, por lo que deberá introducir diferentes cautelas en la utilización de IA en la gestión laboral.

La norma opta por un enfoque basado en el riesgo, exigiendo, en función del nivel de riesgo, requisitos más o menos rigurosos para cada uno de los actores participantes, o prohibiendo la utilización de determinados Sistemas, con el fin de proteger a las personas resultando la existencia de sistemas prohibidos, sistemas de alto riesgo y el resto de los sistemas⁷⁶.

En lo que respecta a las relaciones laborales, los principales riesgos contemplados en la norma son, entre otros, la generación de patrones históricos de discriminación durante el proceso de contratación y en la evaluación, promoción y retención de personas en las relaciones contractuales hasta la extinción de la relación laboral, así como una eventual vulneración de los derechos fundamentales a la intimidad, dignidad o datos personales de la persona trabajadora⁷⁷.

El RIA contempla un marco legal que tiene como principal objetivo el de garantizar una utilización responsable de la IA y a todos los niveles, desde el desarrollador

⁷⁶ RODRÍGUEZ SANZ DE GALDEANO, B., «Obligaciones del empresario en materia de prevención de riesgos laborales derivadas de la utilización de sistemas de IA», en *Revista Galega de Dereito Social* -2ª ET- 18 (2023), p. 48.

⁷⁷ ROJAS ROSCO, R., «El impacto de la regulación de la IA en las relaciones laborales», en *Elderecho.com, Tribuna Abierta*, 3/3/2025 (<https://elderecho.com/el-impacto-de-la-regulacion-de-la-ia-en-las-relaciones-laborales>).

hasta la empresa encargada de su implementación, estableciendo un catálogo de obligaciones concretas aplicables a los Sistemas de Inteligencia Artificial (SIA) considerados de alto riesgo.

Cuando ese riesgo es intolerable, entrará dentro de las prácticas prohibidas, que el RIA enumera y describe en su artículo 5. De acuerdo con el mismo, serán prácticas prohibidas, a grandes rasgos, cuando provoquen, o sean susceptibles de provocar prejuicios considerables a personas o grupos de personas, las relacionadas con técnicas subliminales o engañosas, las que exploten vulnerabilidades de personas, las que se utilicen para clasificar personas o grupos de personas de acuerdo con su comportamiento y características, o para predecir riesgo de delitos, así como las que se utilicen, básicamente, para crear datos o categorizaciones biométricas.

Se trata, en todo caso, de sistemas que contravienen los valores de la Unión de respeto a la dignidad humana, libertad, igualdad, democracia y Estado de Derecho, así como los derechos fundamentales, que reconoce la UE, como el derecho a la no discriminación, la protección de datos y la privacidad⁷⁸.

En el particular caso de la clasificación de personas o grupos de personas por sus características por razones ajenas a las laborales podría entrar la clasificación de las personas en función de si son hombre o mujer, o bien en función de si realizan o no prácticas religiosas, siempre que este encasillamiento de lugar a trato perjudicial o desfavorable.

Cabe advertir, en este punto, que el Reglamento es muy severo a la hora de castigar las prácticas prohibidas, pues implican, entre otros efectos, y además de su propia prohibición, esto es, la del uso, comercialización o puesta en servicio del sistema –como indica el propio artículo 5.1 del Reglamento–; sanciones especialmente graves, como “*multas administrativas de hasta 35.000.000 de euros o, si el infractor es una empresa, de hasta el 7 % de su volumen de negocios mundial total correspondiente al ejercicio financiero anterior, si esta cuantía fuese superior*”, así como “*multas administrativas de hasta 1.500.000 de euros*” cuando el sujeto infractor sea una institución, órgano u organismo de la Unión –según prevén, respectivamente los artículos 99.3 y 100.2.

⁷⁸ RODRÍGUEZ SANZ DE GALDEANO, B., «Obligaciones del empresario en materia de prevención de riesgos laborales derivadas de la utilización de sistemas de IA», cit., p. 49.



Cuando el riesgo no es intolerable, pero es alto, se podrá declarar el SIA como de alto riesgo, siempre que entrañe un riesgo significativo de causar perjuicios para la salud, la seguridad o los derechos fundamentales de las personas físicas, de acuerdo con los criterios que el RIA aprueba en su artículo 6. Estos SIA de alto riesgo son aquellos que, como vemos, pueden afectar derechos fundamentales, al empleo, seguridad y privacidad, sujetos a normas más estrictas en su desarrollo y supervisión.

En este sentido, de acuerdo con el Anexo III del RIA, son sistemas de alto riesgo, entre otros, en el ámbito del empleo, gestión de los trabajadores y acceso al autoempleo los

“Sistemas de IA destinados a ser utilizados para la contratación o la selección de personas físicas, en particular para publicar anuncios de empleo específicos, analizar y filtrar las solicitudes de empleo y evaluar a los candidatos» y a los «sistemas de IA destinados a ser utilizados para tomar decisiones que afecten a las condiciones de las relaciones de índole laboral o a la promoción o rescisión de relaciones contractuales de índole laboral, para la asignación de tareas a partir de comportamientos individuales o rasgos o características personales o para supervisar y evaluar el rendimiento y el comportamiento de las personas en el marco de dichas relaciones”.

Estos sistemas requerirán, en quienes los adquieran y utilicen y, en general, en todos los operadores, el respeto a unos principios que han de guiar su actuación y que son los de: intervención y vigilancia humana, solidez y seguridad técnica, privacidad y gobernanza de datos, transparencia, respeto a la diversidad y no discriminación y bienestar social y medioambiental⁷⁹. Estos requisitos y principios se van desgranando a lo largo de todo el RIA, en la Sección II del Capítulo III.

Las empresas encargadas de la implementación de estos sistemas, con la finalidad de reducir los riesgos, deberán seguir una serie de actuaciones, como

“la aplicación de las instrucciones del proveedor para el uso del sistema; garantía de la supervisión humana por personas que tengan la competencia, la formación y la autoridad necesarias; la validación de los datos de entrada para garantizar su idoneidad y suficiente representatividad para la finalidad

⁷⁹ Vid. GOÑI SEIN, J. L., «El Reglamento UE de Inteligencia Artificial y su interrelación con la normativa de seguridad y salud en el trabajo», en Rodríguez Sanz de Galdeano, B., Egúskiza Balmaseda, M. A. (coord.), *Inteligencia Artificial y Prevención de Riesgos Laborales: Obligaciones y responsabilidades*, Tirant lo Blanch, Valencia 2023, pp. 73-122.

prevista; la supervisión de la actividad del sistema de IA; información al proveedor o distribuidor del sistema de IA de cualquier fallo de funcionamiento, incidente o riesgo; guardar los registros; realización de una evaluación de impacto; así como facilitar la información necesaria a los representantes de los trabajadores y a las personas trabajadoras afectadas de que estarán expuestos a la utilización del sistema de IA de alto riesgo, de conformidad con la normativa nacional⁸⁰”.

Además, se impone un deber de transparencia para sistemas de IA, sean o no de alto riesgo, destinados a interactuar con personas físicas. En estos casos se deberá informar por el sistema de IA o por el propio proveedor o usuario a las personas expuestas de que están interactuando con un sistema de IA, salvo cuando resulte evidente. Además, se informará cuando proceda, de las funciones habilitadas por la IA, de si existe vigilancia humana y de quién es el responsable de la toma de decisiones.

Del mismo modo, en caso de que el empresario ponga a disposición sistemas de IA considerados de riesgo limitado, deberá realizar una evaluación del impacto de los riesgos residuales sobre la protección de datos y la seguridad y salud del trabajador desde la perspectiva de prevención de riesgos laborales.

Para facilitar el cumplimiento de estas obligaciones tanto éticas como legales, y además de la negociación colectiva⁸¹, podría ser un buen instrumento la negociación en la empresa de políticas de uso responsable de la IA. La introducción en la empresa de estas políticas sería, en principio, de carácter voluntario, si bien si la IA se va a utilizar para implementar medidas de control laboral de las personas trabajadoras, el artículo 87 LOPDGDD, exigiría la elaboración de unos criterios de utilización de la IA en el ámbito laboral, previo trámite de participación con la representación legal de los trabajadores en caso de que exista⁸².

⁸⁰ ROJAS ROSCO, R., «El impacto de la regulación de la IA en las relaciones laborales», cit.

⁸¹ Vid. ÁLVAREZ CUESTA, H., «Experiencias convencionales de regulación del impacto laboral de la inteligencia artificial y su uso con fines de control», en *Trabajo y derecho* 19 (2024).

⁸² En este sentido, nuestro Tribunal Supremo dictó una sentencia con fecha 6 de febrero de 2024 (Recurso 263/2022) donde señala la necesidad legal de conceder un trámite de participación a la representación legal de los trabajadores, no solo de forma previa a la implementación de la política de uso de medios tecnológicos, sino también ante cualquier modificación de dichos criterios de uso, de acuerdo con lo establecido en el artículo 87 LOPDGDD, en relación con el art. 20.3 ET.



5. CONCLUSIONES

I. La incorporación de sistemas de inteligencia artificial en el ámbito laboral constituye un cambio estructural con profundas implicaciones jurídicas, sociales y éticas. La IA no actúa sobre una sociedad sin leyes ni costumbres, sino que se inserta en contextos sociales que presentan desigualdades históricas, así como prejuicios culturales preexistentes.

En esta línea, el Reglamento (UE) 2024/1689 subraya expresamente que el desarrollo y uso de los sistemas de IA debe respetar los valores de la Unión y los derechos fundamentales reconocidos en la Carta de Derechos Fundamentales de la UE.

El documento *Antiqua et nova* insiste en que el progreso tecnológico no puede desligarse de la responsabilidad moral, advirtiendo que la técnica debe permanecer al servicio de la persona humana y del bien común, y no convertirse en un fin en sí misma.

II. En sociedades caracterizadas por una creciente diversidad religiosa y cultural, como es el caso de los Estados miembros de la Unión Europea y, en particular, de España, la visibilización de las convicciones religiosas en el espacio público y laboral constituye un elemento central de la identidad personal. El Reglamento Europeo de IA, al advertir del riesgo de reproducción de patrones históricos de discriminación en ámbitos como el empleo, presenta un marco normativo idóneo para abordar situaciones de discriminación en este ámbito desde una perspectiva preventiva.

En este contexto, la detección y corrección de sesgos algorítmicos que penalicen a mujeres con prácticas religiosas visibles se configura no solo como una exigencia jurídica, sino como un imperativo ético, particularmente si se han utilizado prácticas de las catalogadas como prohibidas.

III. La discriminación interseccional constituye un tipo nuevo y diferente de discriminación, distinta y agravada respecto de los rasgos particulares que la producen en cada caso. Además, puede intensificarse en entornos laborales en los que se toman decisiones a través de sistemas algorítmicos.

Hemos de estar alerta sobre la reproducción de desigualdades estructurales en el ámbito del empleo a través de la utilización de la IA, particularmente este tipo agravado de discriminaciones.

IV. La automatización de decisiones laborales mediante sistemas de IA introduce un riesgo añadido de invisibilización de las prácticas discriminatorias. La presentación de las decisiones algorítmicas como objetivas, neutras o científicas dificulta la identificación de los sesgos subyacentes y diluye las responsabilidades jurídicas. Por este motivo, el Reglamento de IA, insiste en la necesidad de garantizar la trazabilidad, la transparencia y la explicabilidad de los sistemas, especialmente cuando afectan a derechos fundamentales.

En el marco de la Doctrina Social de la Iglesia, el documento *Antiqua et nova* alerta contra la ilusión de neutralidad tecnológica y subraya que toda tecnología incorpora opciones humanas y valoraciones implícitas, por lo que resulta éticamente inaceptable ocultar estas decisiones tras una aparente objetividad técnica.

V. Por este motivo, la supervisión humana se configura como una garantía esencial e imprescindible para la protección de los derechos fundamentales en el ámbito laboral. El Reglamento Europeo de IA impone de forma expresa la obligación de supervisión humana efectiva en los sistemas de alto riesgo, exigiendo que las personas encargadas puedan comprender, supervisar y, en su caso, intervenir o anular las decisiones automatizadas.

En concreto, se deberá vigilar expresamente que estos sesgos no afecten a la discriminación por razón de género, y que la pertenencia a colectivos en riesgo de exclusión social, fuente de discriminación en sí misma, no se traslade también a estos sesgos adquiridos por los algoritmos. Para ello es necesario contar con personal, tanto en la selección de personal como en el diseño de los sistemas de IA, que represente diferentes colectivos y sensibilidades y que, además, tenga formación específica.

En todo caso, es esencial la información y la transparencia en la toma de decisiones mediante algoritmos, y en los parámetros que estos siguen en sus procesos.

VI. La atribución de responsabilidad en materia de discriminación algorítmica no puede diluirse en la complejidad técnica de los sistemas de IA. El Reglamento de IA establece un sistema de obligaciones graduadas que alcanza tanto a los proveedores como a los responsables del despliegue de los sistemas, imponiendo deberes de evaluación de riesgos, gobernanza de datos y control de resultados.



VII. La lucha contra la discriminación algorítmica no puede limitarse a mecanismos reactivos de tutela judicial. El Reglamento Europeo de IA apuesta por un enfoque preventivo basado en la evaluación de riesgos, la documentación técnica y las auditorías periódicas, especialmente en el caso de sistemas de alto riesgo utilizados en el ámbito laboral.

En este contexto, la negociación colectiva y la participación de la representación de las personas trabajadoras se presentan como instrumentos especialmente adecuados para concretar estas exigencias.

Sería, además, muy deseable que las empresas elaborasen políticas de uso responsable de la IA, así como la introducción en la negociación colectiva de cláusulas sobre utilización y limitación de la utilización de algoritmos de IA.

VIII. El presente artículo dedica una especial atención a la identificación de puntos de convergencia sustanciales entre la regulación civil de la IA y la reflexión ética contenida en *Antiqua et nova* y, en general, en la Doctrina Social de la Iglesia. Ambos enfoques parten de una visión antropocéntrica que sitúa a la persona en el centro del desarrollo tecnológico. Mientras el Reglamento de IA proclama como objetivo garantizar una IA fiable que respete los derechos fundamentales, la democracia y el Estado de Derecho, la Doctrina Social insiste en la primacía de la persona, el bien común y la justicia social como criterios rectores del desarrollo tecnológico.

En este campo, como en muchos otros relacionados con los derechos laborales, el Derecho positivo y la doctrina social de la Iglesia pueden mantener un diálogo fecundo en defensa de las personas trabajadoras, poniendo el acento en la dignidad del ser humano.