

DEFENDIENDO LOS DERECHOS FUNDAMENTALES FRENTE A LA INTELIGENCIA ARTIFICIAL

LECCIÓN INAUGURAL
DEL CURSO ACADÉMICO 2019-2020
PRONUNCIADA POR EL
PROFESOR DOCTOR JOSÉ LUIS GOÑI SEIN

CATEDRÁTICO DE DERECHO DEL TRABAJO Y DE LA SEGURIDAD SOCIAL
DE LA UNIVERSIDAD PÚBLICA DE NAVARRA

DEFENDIENDO LOS DERECHOS FUNDAMENTALES FRENTE
A LA INTELIGENCIA ARTIFICIAL

LECCIÓN INAUGURAL
DEL CURSO ACADÉMICO 2019-2020
PRONUNCIADA POR EL
PROFESOR DOCTOR JOSÉ LUIS GOÑI SEIN
CATEDRÁTICO DE DERECHO DEL TRABAJO Y DE LA SEGURIDAD SOCIAL
DE LA UNIVERSIDAD PÚBLICA DE NAVARRA



Pamplona, 13 de septiembre de 2019

upna

Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Edita: Universidad Pública de Navarra / Nafarroako Unibertsitate Publikoa
Coordinación: Servicio de Comunicación
Fotocomposición: Pretexto
Imprime: Imprenta Zubillaga
Depósito Legal: NA 2071-2019
Distribución: Sección de Publicaciones
Universidad Pública de Navarra
Campus de Arrosadia
31006 Pamplona
Fax: 948 169 300
Correo: publicaciones@unavarra.es

Sra. Presidenta de la Comunidad Foral de Navarra,
rector magnífico,
dignísimas autoridades,
miembros de la comunidad universitaria,
señoras y señores.

En el epílogo de su mandato, el anterior rector, don Alfonso Carlosena, tuvo a bien confiarme el encargo de pronunciar la lección inaugural en el acto de apertura del curso académico universitario 2019-2020, que solemnemente se celebra en el día de hoy. Quiero, ante todo, expresarle a él y al actual rector, con tanta firmeza como sinceridad, mi gratitud por concederme este privilegio. Me engañaría si pensara que ello se debe exclusivamente a mis méritos académicos y científicos. No se me oculta que es el inexorable paso del tiempo lo que, en buena medida, me trae hoy aquí ante todos ustedes. Como diría Mario Benedetti:

*los años se vinieron
y no se van
se quedan como troncos
pesan como desdichas*

*yo me hago el sordo/ignoro
sus truenos y mi pulso
miro hacia el horizonte
como si le tocara florecer*

Impulsado, pues, por los años alevosos, que van cayendo y que nos reservan, de vez en cuando, gratas sorpresas como la que tengo el honor de vivir con todos ustedes, me propongo disertar, sobre un tema de nuestro proteico presente como es *La inteligencia artificial y la defensa de los derechos fundamentales*. Convendrán conmigo en que

la inteligencia artificial (IA), que opera en un mundo sin ley, dando forma a nuestro futuro como personas, plantea importantes dilemas de tipo ético, político y jurídico. Me propongo, de forma muy modesta, observar de manera reflexiva su incidencia en el ámbito de la empresa y de las relaciones laborales e identificar los principios éticos que deben respetarse en su desarrollo y uso para preservar los derechos fundamentales.

I. La irrupción disruptiva de la tecnología digital

En un periodo de transición muy corto, en apenas tres décadas, se han fraguado unas innovaciones tecnológicas fantásticas, de enorme trascendencia, asociadas a los sistemas de gestión de la información. El desarrollo de Internet, las redes sociales, el mundo móvil, el teléfono inteligente, el uso del *big data*, la inteligencia artificial, la robótica o la impresión 3D se encuentran entre los ejemplos más destacados, pero ciertamente no son los únicos.

Todas estas innovaciones tecnológicas, que se engloban bajo el fenómeno de la «digitalización», al actuar de forma conjunta y combinada, están produciendo un repentino cambio en la realidad productiva, una auténtica revolución, que ha sido calificada como una verdadera «disrupción», en cuanto que vienen a alterar radialmente la forma en que la gente vive y trabaja¹.

Tales innovaciones tecnológicas tienen como base la obtención, gestión y procesamiento de la información. El elemento clave de esta revolución tecnológica es la información, la recopilación ingente de datos, el *big data*, y el análisis y tratamiento de los datos, basado en la inteligencia artificial. Y ello se proyecta en todos los órdenes de la vida social y económica, y, en particular, en lo que aquí más interesa destacar, sobre la actividad productiva, tanto en lo que se refiere a la forma de concebir el negocio, como a la gestión de los recursos humanos.

Las empresas centran sus inversiones y esfuerzos «en obtener más y mejor información, a mayor velocidad y en un tiempo reducido con el objetivo de incrementar su eficiencia, multiplicar la capacidad de respuesta y optimizar los recursos»². Ya no

1. Mercader Uguina, J.: *El futuro del trabajo en la era de la digitalización y la robótica*, Tirant lo Blanch, Valencia, 2017, pp. 29-30.

2. Gordon, A.: «Hacia nuevos horizontes corporativos: de la adaptación a la transformación», en Cuatrecasas, Instituto de Estrategia Legal en RRHH: *Nuevas tecnologías y gestión de recursos humanos. Proyecto Technos: impacto de las redes sociales y marco laboral neotecnológico*, Cuatrecasas, Wolters Kluwer, Madrid, 2017, p. 119.

toman decisión alguna importante sin llevar a cabo un análisis de los datos y de la información completa del mercado, de las preferencias de los potenciales clientes, usuarios o consumidores. Las empresas cada vez disponen de mayor información sobre su propio proceso productivo. Y ese conocimiento del proceso productivo les ayuda a predecir cómo se comportará el material, a saber cuándo tienen que renovar las máquinas, donde está cada producto en cada momento, etc.

Por otra parte, implementan nuevos modelos de gestión de personal basados en la información que, de forma mucho más metódica y rigurosa, proporciona el dispositivo digital que maneja el empleado. La digitalización ha traído consigo la pérdida de relevancia de la actividad física del trabajador. Hoy un fresador presta sus funciones manejando una *tablet* o agenda electrónica y no se mancha las manos. Muchos trabajadores solo están conectados a los dispositivos digitales que van generando una doble información: una propia y específica, conectada directamente con la actividad que desarrolla, y, otra indirecta o inespecífica, que describe su persona y que queda también registrada como huella digital en el dispositivo utilizado, creando nuevas áreas de conocimiento. Los empleadores se sirven de esa información para mejorar la gestión del trabajo, e incidir en los aspectos personales del individuo.

II. El desarrollo y utilización de la inteligencia artificial en el ámbito laboral

El factor clave de todo este desarrollo tecnológico digital radica en la inteligencia artificial, que se aplica a un número creciente de nuevos productos y servicios³. Como señala la Comunicación de 2018 de la Comisión al Parlamento Europeo, al Consejo Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones, titulada *Inteligencia artificial para Europa*⁴: «lejos de ser ciencia-ficción, la inteligencia artificial (IA) forma parte de nuestras vidas».

La IA, además de facilitarnos la vida, nos está ayudando a resolver algunos de los principales retos a los que se enfrenta nuestro mundo: desde el tratamiento de enfermedades crónicas o la reducción de las tasas de mortalidad en los accidentes de tráfico hasta la lucha contra el cambio climático o la previsión de las amenazas a la ciberseguridad.

3. European Group on Ethics in Science and New Technologies: *Artificial Intelligence, Robotics and 'Autonomous' Systems*, European Commission, Brussels, 2018, p. 6.

4. Comisión Europea Bruselas, 25 de abril de 2018, COM (2018) 237 final.

La importancia de la inteligencia artificial se hace especialmente visible en el mundo de la empresa, aunque las pymes están muy lejos de alcanzar el avance tecnológico de las grandes empresas. Pueden apreciarse efectos positivos no solo en cuanto a analizar el flujo de datos y predecir comportamientos futuros que permitan decidir, sino en términos eficiencia y ahorro de costos en la gestión de recursos humanos. Las empresas recurren cada vez más a la inteligencia artificial en todas las fases de la relación laboral.

Así, ya es una realidad en los procesos de selección, donde las empresas están tomando decisiones sin intervención humana o con una intervención mínima con base en la experiencia digital. Son los *softwares* especializados los que realizan el primer filtro de selección en el reclutamiento de personal parametrizando una serie de palabras clave relacionadas con la empresa o el puesto vacante⁵ y en muchos casos la selección completa.

Los modelos matemáticos están presentes también en el ámbito de la planificación de los horarios, donde las decisiones de organización de la empresa las toma un programa de *software*. Los trabajadores tienen unos horarios cada vez menos predecibles, el *software* se encarga de identificar los días y las horas laborables, y de calcular el número de trabajadores que se precisa en cada momento, sin que el trabajador sepa cuando le van a llamar, así las empresas se ahorran los tiempos muertos y ganan en eficiencia.

A menudo el algoritmo juega también un papel importante en la evaluación y control de los trabajadores. Impone el ritmo de trabajo y mide automáticamente la velocidad y la eficiencia de cada trabajador individual. Por ejemplo, Amazon proporciona a los trabajadores un dispositivo digital, mediante el cual el trabajador es guiado y controlado en el tiempo que invierte en transportar paquetes de un sitio a otro, de manera que si emplea más de lo establecido se le envía un mensaje de forma automática.

Los algoritmos son, asimismo, capaces de decidir de forma automatizada el término de la relación contractual. Es posible vislumbrar esta aplicación de la inteligencia artificial en algunas plataformas digitales, donde con base en indicadores predeterminados, se toman decisiones sobre la permanencia del prestador de servicio; así, si un

5. Zemsania Group: «¿Y si tu próxima entrevista de trabajo la hiciera un robot?», <https://zemsania-globalgroup.com/futuro-rrhh-robots-automatizacion/>

conductor obtiene una calificación por debajo de un nivel, o una mala valoración por parte de los clientes, según los parámetros introducidos en el algoritmo, el trabajador es desconectado automáticamente de la plataforma⁶.

Además, gracias al *machine learning*, encontramos, en las grandes corporaciones, máquinas o robots inteligentes, que interactúan de una forma natural y precisa con las personas durante las 24 horas del día y los siete días de la semana. Estos robots tienen autonomía y capacidad de aprender a través de la experiencia y capacidad de adaptar su comportamiento a su entorno⁷. Su capacidad de autonomía surge de los algoritmos que desempeñan un papel importante, no solo en la identificación de patrones útiles en conjuntos de datos, sino también al tomar decisiones que dependen de estos patrones⁸.

III. El espejismo «Big dataísta»

Se ha dicho que la inteligencia humana corre hoy el riesgo de ser sepultada por la realidad virtual. Y en efecto, las grandes corporaciones han iniciado un culto al *big data* y a la inteligencia artificial convirtiéndolo en el centro de todas sus decisiones. Se pretende extraer de ello no solo información privilegiada sobre sus mercados y clientes, y patrones significativos de comportamiento de los trabajadores, sino el conocimiento propiamente dicho: es decir, profundas revelaciones que el ser humano no es capaz de obtener solo con las capacidades analíticas humanas.

Hay una tendencia creciente a reducir la realidad y las personas a los parámetros de la informática y a dejar libertad a una inteligencia artificial para que extraiga nuevas revelaciones que resultan inaccesibles con los métodos de interpretación humana. Aún más, se confía en ella para que incluso adopte todo tipo de decisiones en función de las circunstancias y al margen del ser humano.

6. Todoli Signes, A.: «La gobernanza colectiva de la protección de datos en las relaciones laborales: *Big data*, creación de perfiles, decisiones empresariales automatizadas y los derechos colectivos», *Revista de Derecho Social*, n° 84, 2018, p. 72.

7. De Stefano, V.: «Negotiating the algorithm: Automation, Artificial Intelligence and labour protection», *International Labour Organization, Employment Policy Department Employment Working Paper*, n° 246, Genève, Suiza, p. 3.

8. Kullmann, M.: «Platform Work, Algorithmic Decision-Making and EU Gender Equality Law», *Journal of Comparative Labour Law and Industrial Relations*, n° 1, vol. 34, 2018, p. 6.

Estamos viviendo una época de «fascinación»⁹ «Big dataísta», o de exaltación de la robótica y de la inteligencia artificial. Las decisiones que afectan a nuestras vidas no están hechas por humanos, sino por modelos matemáticos¹⁰, hasta el punto de que se han hecho dueños de las actividades humanas y casi del saber humano, dada la capacidad de generar conocimiento propio, y tomar decisiones por sí mismos, mediante aprendizaje automático (*machine learning*) y aprendizaje profundo (*deep learning*).

Los robots no solo interactúan de una forma natural y precisa con las personas, sino, y fundamentalmente, con las cosas: cada vez hay más cosas u objetos conectados a internet que personas¹¹. Los objetos se han convertido «en entidades cada vez más cognitivas, capaces de valorar opciones de decisión y actuar consecuentemente»¹².

Mediante la inteligencia artificial y la comprensión del lenguaje verbal, los robots realizan actividades físicas rutinarias de trabajo, y generan efectos físicos en el entorno; e incluso son capaces de realizar actividades cognitivas, como entrevistas de trabajo, resolver múltiples preguntas, hacer juicios tácitos o (hasta) sentir emociones. No sabemos cuántos ámbitos de actividad humana pueden llegar a colonizar.

Pero, no debemos caer en el espejismo «Big dataísta». Como escribe el filósofo Jordi Pigem: «ver en los datos y en la inteligencia artificial, la esencia del verdadero conocimiento es un espejismo»¹³.

Las máquinas pueden llegar a interactuar con los humanos, pero, en realidad,

no piensan ni entienden, solo ejecutan mecánicamente. Pueden calcular prodigiosamente a base de aplicar simples reglas mecánicas, pero eso no es inteligencia. Nada entienden de lo que hacen. [...] La inteligencia artificial no

9. Ponce Solé, J.: «Inteligencia artificial, derecho Administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico», *Revista General de Derecho Administrativo, Iustel*, nº 50, 2019, p. 13.

10. O'Neil, C.: *Armas de destrucción matemática. Cómo el big data aumenta la desigualdad y amenaza la democracia*, Capitán Swing, Madrid, 2017.

11. Cuatrecasas, Instituto de Estrategia Legal en RRHH: *Internet of Things y su impacto en los Recursos Humanos y en el Marco Regulatorio de las Relaciones laborales*, Cuatrecasas, Wolters Kluwer, Madrid, 2017, pp. 28-29.

12. *Ibidem*, p. 33.

13. PIGEM, J.: *Ángeles o robots, La interioridad humana en la sociedad hiperconectada*, Fragmenta Editorial, 2018, p. 58.

es inteligencia, del mismo modo que una flor artificial no es ninguna flor, por más que lo parezca hasta que prestamos atención. Inteligencia artificial es un oxímoron¹⁴ que ofende a la inteligencia. No hay inteligencia sin vida y sin sensibilidad¹⁵.

La robótica seguirá siendo un subproducto humano, porque la creatividad es algo esencialmente humana. Por más que los robots actúen con un alto grado de autonomía, y sean capaces de generar sus propias instrucciones, no hay que dejar de ver detrás de cada robot a la acción humana. Los robots son mecanismos tecnológicos concebidos por el hombre y guiados en su diseño y configuración finalista por la inteligencia humana, bien que en ocasiones los humanos pierdan el control de las decisiones tomadas, y ni siquiera entiendan por qué una decisión errónea ha sido tomada. Las máquinas adoptan la decisión por delegación humana. Y, en esa medida, el hombre es el responsable último de la decisión.

Pero más allá de las disquisiciones filosóficas sobre la inteligencia artificial, aquí importa especialmente valorar el impacto que la adopción de estas nuevas tecnologías disruptivas genera en los trabajadores y, en particular, sobre los derechos fundamentales, pues es evidente que los sistemas de inteligencia artificial presentan riesgos y pueden tener, junto a múltiples efectos beneficiosos, un impacto negativo.

En este sentido, la Comunicación de 2018 de la Comisión Europea sobre «Inteligencia artificial sobre Europa», observa que «como ocurre con cualquier tecnología transformadora, algunas aplicaciones de la IA pueden plantear nuevos problemas de carácter ético y jurídico, por ejemplo, en relación con la responsabilidad o la adopción de decisiones parcialmente sesgadas».

IV. Robótica y tensiones sociales: la destrucción de empleo y la deshumanización

En efecto, hay que reconocer que los desarrollos tecnológicos no son neutros o inocuos, exhiben siempre un lado positivo y un lado negativo. Es lo paradójico de las innovaciones tecnológicas. Estas nuevas expresiones de la tecnología digital, en especial la robótica y la inteligencia artificial, pueden ser tanto un medio de progreso y

14. Que significa: «usar dos conceptos de significado opuesto en una sola expresión».

15. Pigem, J.: *Ángeles o robots...*, cit., pp. 72-73.

de bienestar del ser humano, como la expresión de una nueva forma de utopía negativa. «En el fondo lo decidimos nosotros»¹⁶ los seres humanos.

Los robots (entendiendo por tales, los «robots humanoides» o «colaborativos» que conviven e interactúan con los trabajadores, como los «robots convencionales» o estáticos protegidos en barreras) y los sistemas IA presentan innumerables ventajas en los ámbitos de la asistencia sanitaria, la seguridad en el transporte y la energía, la lucha contra el cambio climático y la anticipación de las amenazas en materia de ciberseguridad¹⁷.

En lo que respecta al ámbito laboral, una de las posibles ventajas de la robótica está relacionada con esa «utopía positiva» de la «sociedad liberada del trabajo» de la que habla Vincenzo Ferrari en su ensayo sobre «El incierto futuro de los derechos en el trabajo humano»¹⁸; las máquinas o robots tienen esa aptitud de liberar a los humanos de las formas más odiosas de trabajo.

Los robots desempeñan un papel importante en la minoración del nivel de exposición de los trabajadores a los riesgos laborales. Pueden contribuir a liberar a los trabajadores de las actividades pesadas y peligrosas, principalmente en el sector manufacturero y de logística, así como de todas aquellas actividades repetitivas, o trabajos peligrosos, sucios, u opacos, que pueden representar un riesgo para la salud de los trabajadores. Los robots se revelan, en suma, claves para aliviar la fatiga física y la penosidad que presentan muchos trabajos no solo de la industria, sino del sector servicios y para mejorar el bienestar de los trabajadores.

Por otro lado, el desarrollo de ciertos robots o extensiones robóticas aporta también beneficios positivos a las personas con movilidad reducida, personas mayores, y, en particular, a las personas con discapacidad, en tanto que puede ser un elemento relevante para mitigar los problemas de desigualdad laboral. Un exoesqueleto o las extensiones robóticas «se presentan como una gran oportunidad de integrar a traba-

16. Del Molino, S.: «Nosotros en el centro de todo, sin viejos mapas, ni puntos cardinales», *El País*, 14 de abril de 2019; *Ideas*, p. 5.

17. Como destaca el Consejo Económico y Social Europeo: «Inteligencia artificial: anticipar su impacto en el trabajo para garantizar la transición justa», dictamen oficial de 20 de septiembre de 2018. Disponible en <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52018IE1473&from=ES>.

18. Ferrari, V.: «El incierto futuro de los derechos en el trabajo humano», *Trabajo y Derecho*, n° 52, abril de 2019, p. 16.

adores con discapacidad en el ámbito laboral respecto a tareas o funciones que, por limitaciones físicas, no les sería posible desarrollar sin ese suplemento»¹⁹.

Esos complementos robóticos pueden ayudar a las personas físicas con algún tipo de deficiencia física o sensorial a conseguir un puesto de trabajo o, tratándose de un trabajador con una discapacidad sobrevenida, a que pueda desarrollar una actividad laboral con un rendimiento normal, haciendo que la participación de tales personas en la sociedad sea plena y efectiva. Desde esta perspectiva, los complementos robóticos muestran una consistente potencialidad como medida antidiscriminatoria.

Pero la robótica no solo trae efectos positivos. Desde la perspectiva del trabajador se perciben también importantes efectos globales que generan inquietud. En general preocupa la amenaza que representa la automatización que los robots van a llevar a cabo en los próximos años para el nivel de empleo existente y las condiciones de trabajo. La inclusión masiva de los robots en los procesos productivos provocará la destrucción de puestos de trabajo sin que se sepa si se van a crear otros puestos de trabajo de sustitución en la misma proporción, y una presión a la baja de los salarios. Todo lo cual podrá desencadenar, además, problemas de desigualdad social.

Los sectores que ya cuentan con tecnología robótica son fundamentalmente los de la industria manufacturera (fábricas de montaje de vehículos), el de almacenamiento y la logística (v. gr. Amazon) y el del transporte (v. gr. ferrocarril, como el metro sin conductor)²⁰. También se está imponiendo en el sector de la sanidad: Japón ha apostado con fuerza por la robótica para suplir la falta de personal y atender a una población envejecida. Así, en muchas residencias y casas particulares se utilizan artilugios como «como la máquina Hug, que ayuda a los ancianos a incorporarse; sensores que miden su grado acumulación de orina o que monitorean su estado de sueño; o el robot Sota, que interactúa con los ancianos y tiene capacidad de activar, a través de una serie de comandos orales, funciones de otros dispositivos domésticos o monitorear constantes vitales»²¹.

19. Vid. Cuatrecasas, Instituto de Estrategia Legal en RRHH: *Robótica y su impacto en los Recursos Humanos y en el Marco Regulatorio de las Relaciones Laborales*, Cuatrecasas, Wolters Kluwer, Madrid, 2018, p. 344.

20. *Ibidem*, p. 49.

21. Rizzi, A: «En busca del equilibrio entre mayores, migrantes y robots», *El País*, 25 de marzo de 2019, pp. 10-11.

Como destaca el Consejo Económico y Social Europeo en el dictamen de 2018 sobre inteligencia artificial, la repercusión que tendrá la introducción de la robótica en los procesos de producción es controvertida. La diversidad de resultados (del 9% al 54%) de empleos amenazados²², refleja la complejidad de las opciones metodológicas y su diversa repercusión en los resultados de la investigación.

De momento, el avance de la automatización está siendo limitado y no se han perdido tantos puestos de trabajo como se ha venido vaticinando. Tal vez sea pronto para evaluar el impacto de esta automatización. Cabe notar, no obstante, que autorizadas voces de la economía ponen en cuestión ese incremento de destrucción de empleo por cuanto creen firmemente que la incorporación masiva de la robótica derivará en una reubicación y reconversión de empleos²³. En todo caso parece innegable, que se producirá una transformación de los modelos de producción y que ello tendrá repercusiones laborales y sociales para el colectivo de trabajadores.

Al margen de esta amenaza, la automatización por medio robots humanoides o *co-bots* con los que trabajadores interactúan en el mismo espacio que ellos emitiendo o recibiendo instrucciones de ellos y ejecutando con ellos, puede tener, también, como ha señalado De Stefano,²⁴ «implicaciones imprevistas para la dignidad humana de las personas físicas involucradas en estos procesos».

Observa el autor que los procesos de automatización aumentan los sentimientos de alienación del trabajo, y las personas que trabajan codo con codo con entidades físicas no-humanas que disfrutan de personalidad jurídica corren el riesgo, aún más, de ser consideradas meros engranajes en el proceso de producción, algo que podría conducir a una mercantilización severa de su trabajo con consecuencias deshumanizantes no deseadas.

22. Un estudio realizado por McKinsey Global Institute: «A Future that Works: Automation, Employment, and Productivity», enero de 2017, estima que aproximadamente la mitad de todas las actividades actuales podrían automatizarse mediante la adaptación de tecnologías actualmente demostradas. Tales actividades serían las que los trabajadores dedican el 49% de su tiempo de trabajo a actividades mecánicas o automatizables.

23. Tesis mantenida en un informe de 2013 por los estudiosos de la Universidad de Oxford, Osborne, M. y C. Frey: «The Future of Employment: How Susceptible Are Jobs to Computerisation?». Disponible en: https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf

24. De Stefano, V.: *Negotiating the algorithm: Automation, Artificial Intelligence...*, cit., p. 5.

Al margen de lo expuesto, no cabe pasar por alto los efectos psicosociales que pueden generarse por la convivencia con los robots colaborativos en la persona del trabajador. Colocar a un trabajador junto a un robot es someter al trabajador a una mayor presión psicológica. El trabajador habrá de desarrollar sus tareas al ritmo determinado por el robot.

Ese tecnoestrés implica un «sobre esfuerzo realizado para adaptarse a las nuevas exigencias»²⁵, que puede desembocar en problemas de dependencia y psicológicos graves como la fatiga o la tecnofobia. Supone una compresión de la esfera de la dignidad del trabajador en favor de las exigencias productivas, en tanto que el ejercicio automático e inhumano del robot influye en el trabajador con la consiguiente pérdida de autonomía en la ejecución de su actividad laboral.

V. La digitalización del trabajo y los conflictos individuales: desafíos para la privacidad y la igualdad

La introducción de la tecnología digital y el uso del *big data* en la organización del trabajo suscita también otra serie de dilemas y tensiones individuales importantes en relación a los derechos fundamentales y el uso de los dispositivos digitales facilitados por la empresa como herramientas de trabajo. Estos dispositivos proporcionan, como se ha señalado anteriormente, a las empresas una mayor información sobre sus empleados, y, sin perjuicio de que eso les permite adoptar decisiones más inteligentes y efectivas, puede, en ocasiones, afectar a los derechos de las personas afectadas.

La idoneidad de los dispositivos informáticos para registrar muchos datos ajenos respecto de lo verdaderamente útil para el cumplimiento de la prestación, supone un riesgo para la privacidad del trabajador. Los registros digitales del comportamiento humano en los dispositivos digitales proporcionados por el empresario en el lugar de trabajo, como el ordenador (correo electrónico, la navegación por Internet, la mensajería instantánea), la *tablet*, o el móvil, permiten al empresario hacer un seguimiento exacto del trabajador durante el tiempo de trabajo, tanto de la ejecución de la prestación, como de su persona.

En el mercado se encuentran disponibles técnicas muy sofisticadas para el control del uso que el trabajador hace de las aplicaciones y dispositivos instalados en el

25. Cuatrecasas, Instituto de Estrategia Legal en RRHH: *Internet of Things y su impacto en los Recursos Humanos...*, cit., p. 323.

ordenador, como algunos conjuntos de paquetes de seguridad, con capacidad para monitorear todo el uso del ordenador. Mediante estas técnicas de inteligencia artificial altamente invasivas y omnipresentes de monitoreo, resulta muy fácil obtener una visión muy cercana de los usos de los dispositivos (profesional o personal) y patrones de comportamiento del titular, y llegar a saber las preferencias de todo tipo (su orientación sexual, su punto de vista político o sindical).

Debe tenerse en cuenta, además, que la empresa tecnológicamente hiperconectada tiende a externalizar y a transferir su actividad a los trabajadores mediante fórmulas de trabajo a domicilio o trabajo a control remoto. Y es cada vez más común que el empresario entregue a sus trabajadores equipos de trabajo o de *software* de TIC, que los empleados instalan, bien en los ordenadores de su casa, o bien en sus propios dispositivos móviles (BYOD), permitiéndoles acceder a la red, los sistemas y recursos del empleador. La implementación de estas nuevas tecnologías representa un riesgo superior para la vida privada de los trabajadores porque los empleadores tienen un mayor acceso a datos personales del trabajador.

Junto a todas estas herramientas tecnológicas proporcionadas por el empresario para la prestación de trabajo es bastante común que las empresas cuenten también con tecnologías digitales con el único objeto del control directo del trabajador. Me refiero a la videovigilancia, o el uso de los dispositivos geolocalizadores (GPS), y, en general, cuantos dispositivos integran el Internet de las Cosas o IoT (*Internet of Things*) que se aplican para ejercer algún tipo de supervisión del trabajador.

Dentro de éstos últimos, particularmente inquietantes resultan, por ejemplo, los sistemas biométricos (por ejemplo, reconocimiento facial), los rastreadores de *fitness* que recopilan datos de salud de los empleadores (ritmo cardiaco) proponiendo medidas correctivas para mejorar su salud y también la productividad del negocio; o los *wearables*, como la «ropa conectada» de Google, donde unos hilos de metal conductor, tejidos de forma imperceptible en un pantalón o en una camisa, sirven para interactuar con dispositivos externos y ejercer un control sobre el trabajador²⁶, o la pulsera electrónica de Amazon²⁷ que geolocaliza al trabajador y mide el tiempo que tarda en realizar su actividad, advirtiéndole cuando se retrasa, mediante una luz roja.

26. Cuatrecasas, Instituto de Estrategia Legal en RRHH: *Internet of Things y su impacto en los Recursos Humanos...*, cit., pp. 59-62.

27. Sobre el tema vid. Di Meo, R.: «Technologie e poteri datoriali: commento a margine del c.d. braccialetto Amazon», *Labour & Law Issues*, vol. 4, n° 1, 2018.

En todos estos dispositivos digitales tiene un peso importante la inteligencia artificial, que analiza el entorno y pasa a la acción con cierto grado de autonomía. Trabajando con datos, las decisiones automatizadas se toman con base en algoritmos, que se confeccionan con arreglo a la métrica que ha elegido libremente el programador. Lo cual no deja de suscitar algunos problemas importantes, en la medida en que el algoritmo mide aquello que el programador quiere medir.

En este sentido, observa el Grupo Europeo de Expertos de Alto Nivel en Inteligencia Artificial en *Directrices éticas para una inteligencia artificial fiable* (documento elaborado en abril de este año 2019) que la inteligencia artificial plantea desafíos éticos importantes en relación con el impacto en las personas y en la sociedad, las capacidades de toma de decisiones y la seguridad.

Y ello porque los algoritmos no son neutros. Son, como ha escrito la matemática y activista Cathy O’Neil²⁸ «opiniones encerradas en matemáticas». «En función de los datos que se utilizan para entrenar los sistemas de inteligencia artificial, sus resultados pueden ser sesgados»²⁹. Si se parte de prejuicios, es evidente que se corre el riesgo de que los resultados acaben reflejando los sesgos que incorpora la fuente de datos.

Este riesgo deviene aún más grave cuando estas prácticas se basan en el auto-aprendizaje de la inteligencia artificial, con el *software* capaz de reprogramar sus propios criterios y métricas para alcanzar un resultado predefinido muy general, como mejorar la productividad del trabajo, porque los sistemas de IA aprenden de los datos de formación³⁰.

Ocurre, además, que los sistemas de IA adolecen de falta de transparencia. Normalmente, se ignora cómo las grandes compañías recolectan nuestros datos, cómo los etiquetan, y qué procesos técnicos determinan la decisión del sistema de inteligencia artificial. Los sistemas de inteligencia artificial se basan sencillamente en modelos «opacos, no regulados e incontestables». Rara vez se hacen públicos, por lo que se hace difícil controlar el resultado de las decisiones tomadas por el sistema de inteligencia artificial; esto es, si contienen inexactitudes, errores, sesgos o discriminaciones directas.

28. O’Neil, C.: *Armas de destrucción matemática...*, cit.

29. Como advierte la Comisión Europea en el mencionado documento de 2018.

30. De Stefano, V.: *Negotiating the algorithm: Automation, Artificial Intelligence...*, cit., p. 10.

En algún caso en que se ha hecho público el algoritmo, se ha podido comprobar la falacia que, a veces, esconden. Sirva de muestra el algoritmo que determina el repunte inusual del precio de la electricidad; el algoritmo está publicado en el BOE, pero la Comisión Nacional de la Competencia y de los Mercados ha debido de llamar la atención a las compañías eléctricas porque ha detectado que el algoritmo utilizado descarta ofertas más baratas que el precio que resulta aplicado³¹.

El uso de la IA en el ámbito de las relaciones laborales presenta también problemas de impacto, porque, si bien en teoría los modelos matemáticos debieran conducir a la equidad en las relaciones laborales, en realidad no ocurre así. Ya solo por el hecho de trabajar con probabilidades y no con certezas, no es descartable que se produzcan errores, pero, si además se trabaja con claros prejuicios sobre las características del candidato, o se utilizan datos no inclusivos, las probabilidades de sesgo o discriminación son inevitables.

Existe, así, el riesgo de que los sistemas de IA reflejen, incluso de forma involuntaria, los sesgos de sus programadores humanos, ya que, si se parte exclusivamente de la idea la productividad y el rendimiento laboral del programador, es probable que se favorezca a los más capaces o inteligentes, «descartando o penalizando a candidatos o trabajadores con discapacidad o con características que difieren de las expectativas de los programadores»³².

En este sentido, el algoritmo puede afectar a las oportunidades de empleo o de promoción laboral o de estabilidad en el empleo si se utilizan criterios sesgados. Por lo tanto, la gestión por algoritmo e inteligencia artificial en el lugar de trabajo, lejos de deparar resultados neutrales y reducir la discriminación, podría aumentar las prácticas discriminatorias³³. Un modelo de algoritmo mal diseñado puede propagar sesgos a escala masiva y acelerada.

El uso de los algoritmos genera, además, algún otro impacto humano que conviene también tener en cuenta, especialmente en el ámbito de funcionamiento de las pla-

31. Comisión Nacional de la Competencia de los Mercados: «Informe de supervisión del mercado peninsular de producción de energía eléctrica». Año 2015, IS/DE/025/16, p. 27: «... se siguen detectando episodios en los que ofertas más baratas que el precio que resulta finalmente casado, resultan descartadas por las limitaciones impuestas al algoritmo utilizado en este proceso». Disponible en https://www.cnmc.es/sites/default/files/1403915_9.pdf. Vid. Ponce Solé, J.: «Inteligencia artificial...», cit., p.17.

32. De Stefano, V.: *Negotiating the algorithm: Automation, Artificial Intelligence...*, cit., p. 9.

33. *Ibídem* p. 10.

taformas digitales como Deliveroo, Glovo, Take Eat Easy, etc., que, como se sabe, se dedican a poner en contacto a grandes grupos de prestadores de servicios con los clientes finales, a través de un colaborador (trabajador autónomo) que se encarga de ejecutar el servicio *on line* o de transportar el producto pactado entre el prestador y el cliente, acudiendo al establecimiento del prestador a recoger el producto y haciéndoselo llegar al consumidor, utilizando, normalmente, su propio medio de transporte.

En estas plataformas digitales, el algoritmo determina varios aspectos del trabajador colaborador, entre ellos, la reputación del trabajador, que depende de la valoración del cliente, la puntuación del *partner*, las horas de alta demanda y eficiencia, y las posibilidades de acceso o continuidad en la plataforma. Los riesgos de aplicación del algoritmo son aquí, no tanto de sesgos o de discriminación, cuanto de deshumanización de la persona del colaborador o trabajador que realiza el reparto de su producto. La inteligencia artificial puede socavar la dignidad de la persona, porque el autómata toma el sitio del empleador y muchas veces las operaciones del sistema generan resultados contrarios a la honorabilidad y estabilidad en el empleo del trabajador, sin que este pueda participar en el proceso de toma de decisión, ni impugnar tal decisión.

El efecto perverso de la inteligencia artificial en el ámbito laboral es constatable, por ejemplo, en el caso descrito por O'Neil en su libro *Armas de destrucción matemática*. En 2011, Sarah Wysocki trabajaba como maestra de primaria en un centro educativo de Washington D. C. A finales de curso fue calificada por un algoritmo de evaluación docente y sacó una penosa puntuación en enseñanza de lengua y matemáticas. El algoritmo, en vez de valorar su forma de enseñar, corregir, tutorizar, etc., medía tan solo un indicador indirecto, en concreto: las puntuaciones de sus alumnos en matemáticas y lectura.

Como esa puntuación representaba la mitad de la valoración global, el distrito escolar no tuvo más remedio que despedirla, a pesar de que gozaba de la estima del director del colegio y de los padres de sus alumnos, que la consideraban una excelente profesora. La cosa no acabó ahí; en los años siguientes, una inspección detectó numerosas tachaduras en las pruebas: los maestros, temiendo ser echados por el algoritmo, corrigieron al alza los ejercicios de sus estudiantes³⁴. No obstante, Sarah Wysocki no tuvo ningún problema porque estuvo tan solo un par de días en paro y enseguida la contrataron en un colegio de un próspero distrito de Virginia.

34. O'Neil, C.: *Armas de destrucción matemática...*, cit., pp.11 y ss.

VI. La protección de los derechos humanos en el trabajo digital

Advertidas las tensiones y preocupaciones críticas que las tecnologías digitales basadas en la inteligencia artificial plantean para los derechos fundamentales de los trabajadores, es el momento de reflexionar sobre el marco en el que se han de desarrollar y utilizar los sistemas de inteligencia artificial, a fin de mitigar el posible impacto negativo sobre el estado de derecho y los derechos fundamentales de las personas, en particular, en el ámbito de la relación laboral.

La Comisión Mundial de la OIT en el documento titulado *Trabajar para un futuro más prometedor*, elaborado con motivo del centenario de la OIT, que se conmemora este año 2019, aboga por un «enfoque de la inteligencia artificial ‘bajo control humano’ que garantice que las decisiones finales que afectan al trabajo sean tomadas por seres humanos y no por algoritmos».

Por su parte, el Consejo Económico y Social Europeo, en el dictamen oficial de finales de 2018 sobre inteligencia artificial, advierte asimismo de la necesidad de establecer «un límite claro en la interacción entre los trabajadores y las máquinas inteligentes, de modo que el ser humano no se convierta nunca en ejecutor de la máquina».

No parece razonable pensar que la inteligencia artificial vaya a traer por sí misma un punto de equilibrio entre los intereses de las partes en conflicto. Los algoritmos no los diseñan los trabajadores, sino las corporaciones, y lógicamente en su propio beneficio, por lo que no es de extrañar que, junto a grandes posibilidades de mejora del trabajo e incluso beneficios sustanciales a los individuos y a la sociedad, presenten también algunos impactos y efectos negativos.

Ya lo estamos viendo en la vigilancia injustificada, la posibilidad subrepticia de elaboración de perfiles sobre la salud y vida privada del empleado, la adopción de decisiones sesgadas y discriminatorias, la falta de justificación de las decisiones automatizadas y la incapacidad de reacción, amén de los abusos de orden laboral que se están produciendo, en particular en las plataformas digitales, al franquear las condiciones mínimas de calidad de trabajo (trabajo autónomo, sin protección, salarios micro, etc.) y donde se constata la existencia de una correlación entre la precariedad y estas plataformas.

La IA es una herramienta que debe ser utilizada con sentido ético por las empresas. Y ese sentido ético debe quedar claramente establecido, como señala el documento mencionado de la OIT, por un «sistema de gobernanza internacional», o subsidiaria-

mente, por las leyes de los distintos países o mediante convenios colectivos. Hay que volver al control social de los cambios tecnológicos donde se exija el respeto a ciertos derechos y protecciones mínimas.

La forma correcta de construir un futuro con inteligencia artificial es, como ha destacado la Comisión Europea, garantizando el establecimiento de un marco ético y jurídico de la IA, basado en los derechos fundamentales consagrados en la UE, la Carta de la UE y la legislación internacional sobre derechos humanos.

En este sentido, siguiendo las indicaciones de la Comisión, se ha constituido un Grupo de Expertos de Alto Nivel de la UE, que ha alumbrado el 8 de abril de 2019 un importante documento sobre *Directrices éticas para una inteligencia artificial fiable (Ethics Guidelines for Trustworthy AI)*³⁵. Estas directrices no tienen carácter vinculante, ni pretenden sustituir a ninguna formulación política ni legal, pero son cruciales para garantizar la confianza en la inteligencia artificial. Diseñan un marco para lograr una IA de confianza, que se estructura en tres capas de abstracción.

En dicho documento, se identifican cuatro principios éticos, que van más allá del cumplimiento formal de las leyes y que pueden inspirar los instrumentos regulatorios nuevos, y guiar la justificación de desarrollo, despliegue y uso de los sistemas de inteligencia artificial. Estos principios son: 1) el respeto a la autonomía humana; 2) la prevención del daño; 3) la equidad; 4) y la explicabilidad.

1) Bajo el principio de respeto a la autonomía humana se propugna que los seres humanos que interactúan con los sistemas de inteligencia artificial sean capaces de mantener la autodeterminación plena y efectiva sobre sí mismos. Se considera que los sistemas de IA no deben subordinar, coaccionar, engañar, manipular, o condicionar injustificadamente a los seres humanos. Al contrario, deben diseñarse para aumentar, complementar y potenciar las aptitudes cognitivas, sociales y culturales humanas.

En el caso de la profesora Sarah Wysocki, antes comentado, se descubrió que los algoritmos de evaluación docente son una herramienta poderosa de modificación del comportamiento. Los maestros, ante la posibilidad de que sus empleos se vieran amenazados si los alumnos suspendían, terminaron aprobando a sus alumnos. Las

35. Sitio Web: <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.

numerosas rectificaciones en los exámenes y puntuaciones anormalmente altas hacían sospechar que los maestros habían cedido ante el miedo³⁶.

En un contexto de IA fiable no deberían tener cabida modelos matemáticos que, como la aplicada en el referido centro de enseñanza, coartan la libertad individual de las personas mediante amenazas más o menos veladas y coerciones indirectas.

La aplicación de este primer principio de respeto a la autonomía humana incluye diversos aspectos sistémicos:

a) Ante todo, la realización de una evaluación del impacto en los derechos fundamentales para minimizar los riesgos. No es nada nuevo, y viene exigida por la normativa de protección de datos. El artículo 35 del Reglamento (UE) 2016/679 ha previsto tal evaluación ante decisiones automatizadas que produzcan efectos jurídicos sobre personas físicas o que les afecten significativamente, y el artículo 28 de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y Garantías de los Derechos Digitales, lo concreta, en particular, entre otros supuestos, cuando el tratamiento pudiera generar situaciones de discriminación, o vulneración de derechos fundamentales, o cuando se trata de crear perfiles personales mediante análisis o la predicción de aspectos referidos al rendimiento en el trabajo, su situación económica, su salud, sus preferencias, su comportamiento, su localización o sus movimientos.

b) Asimismo, el respeto a la autonomía humana comporta que los usuarios puedan tomar decisiones informadas y autónomas con respecto a los sistemas de IA. Lo cual solo es posible si se proporciona el conocimiento y herramientas para comprender e interactuar satisfactoriamente con los sistemas de inteligencia artificial.

En el caso comentado anteriormente, cuando Sarah Wysocki fue despedida del centro por su mala puntuación en su evaluación de IMPACT pensó que los números eran injustos y quiso saber de dónde venían y qué medía exactamente el modelo matemático aplicado. No encontró a nadie que le pudiera explicar su mala puntuación.

Como comenta Cathy O’Neil al respecto, «el modelo en sí es una caja negra, su contenido, un secreto corporativo fieramente guardado». «El análisis se subcontrata a programadores y estadísticos y, por lo general, dejan que sean las máquinas las que

36. O’Neil, C.: *Armas de destrucción matemática...*, cit., pp. 18-19.

hablen». Y sus veredictos, incluso cuando están equivocados o son perjudiciales, son indiscutibles e inapelables³⁷.

Sin embargo, es clave, como señala el Grupo de Expertos, reconocer el derecho a no estar sujeto a una decisión basada únicamente en el procesamiento automatizado cuando esto produzca efectos jurídicos en los usuarios o los afecte de manera similar, de manera significativa.

c) Por lo demás, el respeto a la autonomía humana requiere capacidad de supervisión humana sobre los procesos de trabajo en los sistemas de inteligencia artificial. Significa que los seres humanos deben tener la capacidad de supervisar la actividad general del sistema (incluidos su impacto económico, social, legal y ético más amplio) y la capacidad de decidir cuándo y cómo usar el sistema en cualquier situación particular.

2) Bajo el segundo principio ético de prevención de daños se entiende que la inteligencia artificial no debería causar ni exacerbar el daño ni afectar negativamente a los seres humanos. Los sistemas de inteligencia y los entornos en los que operan deben ser seguros, impidiendo usos maliciosos.

Este segundo principio ético implica, según el Grupo de Expertos de Alto Nivel, varios requerimientos:

a) Que los sistemas de IA se desarrollen con un enfoque preventivo de los riesgos, comportándose de manera fiable según lo previsto, y, al mismo tiempo, minimizando el daño no intencional e inesperado y previniendo el daño inaceptable.

b) Que los sistemas de IA, como todos los sistemas de *software*, estén protegidos contra vulnerabilidades, esto es, frente a ataques que pueden llevar al sistema a tomar diferentes decisiones o al cierre del mismo.

c) Que exista un plan de repliegue y de seguridad, en función de la magnitud del riesgo planteado por el sistema de IA, que a su vez depende de las capacidades del sistema.

37. *Ibíd.*, p. 17.

d) Que pueda hacer juicios, predicciones, recomendaciones o decisiones correctas basadas en datos o modelos, indicando, en caso de no poder evitar la inexactitud, la probabilidad de error.

Todo ello, aplicado al ámbito laboral, significa, en relación con los robots, que debe analizarse el impacto que para la propia seguridad de las personas presenta el uso de los robots, particularmente cuando están diseñados para funcionar en un espacio compartido con los trabajadores, sin necesidad de las protecciones convencionales. Pueden surgir nuevos riesgos para los trabajadores –derivados tanto de las características de los sistemas tecnológicos utilizados, como de la mayor presión psíquica derivada del ritmo de trabajo impuesto por el robot o de la reducción de los tiempos muertos³⁸–, que deberán ser contemplados en las evaluaciones de riesgos y en los correspondientes planes de prevención de riesgos laborales. No existe a nivel internacional normativa que recoja las actuaciones preventivas que, en relación con el uso de los robots, deben seguirse para proteger la salud física y psicológica de los empleados, aunque, es sabido que una de las tres leyes de la robótica, que Asimov desarrolló en su libro *Yo, robot*, es impedir que las máquinas dañen a las máquinas.

Por otra parte, y en relación con los sistemas de *software*, significa que los datos que se recaben deben ser exactos y los equipos informáticos deben estar protegidos contra posibles vulnerabilidades con medidas de seguridad. A este tipo de exactitud y de seguridad se refieren el art. 5.1 d) cuando señala que los datos serán «exactos y si fuera necesario actualizados», y el art. 5.1 f) del RGPD (UE), al disponer que los datos personales «deben ser tratados de tal manera que se garantice una seguridad adecuada de los datos personales, incluida la protección contra el tratamiento no autorizado o ilícito y contra su pérdida, destrucción o daño accidental, mediante la aplicación de medidas técnicas u organizativas apropiadas».

Estrechamente vinculado al principio de prevención de daños se halla la privacidad, que, como se ha observado, resulta especialmente afectada por los sistemas de IA. El uso de los sistemas de IA permite inferir, a partir de los registros digitales del comportamiento humano, una nueva información personal sobre el usuario. Por ello resulta primordial para la fiabilidad del sistema de IA garantizar la privacidad tanto de la información facilitada por el usuario, como de la generada sobre el usuario a lo largo de su interacción con el sistema.

38. Cuatrecasas, Instituto de Estrategia Legal en RRHH: *Robótica y su impacto en los Recursos Humanos...*, cit., pp. 320 y ss.

En el ámbito laboral, la prevención de daños a la privacidad de los empleados requiere asegurar que el ejercicio de la gestión, vigilancia y control a través de tales dispositivos digitales se realiza dentro de unos límites de respeto a la dignidad de la persona humana (art. 20.3 ET). Lo cual significa que el procesamiento de datos personales del trabajador debe estar limitado a lo estrictamente necesario para la satisfacción de un fin legítimo, y debe cumplir con los principios de proporcionalidad y subsidiariedad. Por ello, el Grupo de Trabajo del art. 29 (Dictamen 2/2017)³⁹, considera que, antes de la utilización de cualquier herramienta de monitoreo, debería realizarse una prueba de proporcionalidad para considerar si los datos son necesarios, si este procesamiento prevalece sobre los derechos generales de privacidad que los empleados tienen también en el lugar de trabajo, y qué medidas deben adoptarse para garantizar que la injerencia en el derecho a la intimidad y al secreto de las comunicaciones se limite al mínimo necesario.

3) Bajo el tercer principio ético de la equidad se prescribe que los sistemas de IA deben ser justos. Esta dimensión implica un compromiso negativo y otro positivo: uno negativo, por cuanto se trata de que las personas y los grupos estén libres de sesgos injustos, discriminación y estigmatización y, por tanto, el sistema sea capaz de garantizar una distribución equitativa y justa de los beneficios y los costos. Y otro positivo, porque se trata de aprovechar la potencialidad de la IA para fomentar la igualdad de oportunidades. Se considera que los sistemas de IA deben ser diseñados siguiendo los estándares de accesibilidad y de manera inclusiva, de forma que todas las personas puedan acceder a los dispositivos de IA.

La tecnología digital puede crear, como señalábamos, situaciones discriminatorias, producto de las decisiones automatizadas, adoptadas con programas que tienen sesgos, intencionados o no. El daño o perjuicio resulta a veces de prejuicios y sesgos históricos inadvertidos, pero otras muchas veces de la explotación intencional contra ciertos grupos o personas. La forma en que se desarrolla un programa informático puede llevar a clasificar a las personas en el grupo que no les corresponda.

Y esto ocurre señaladamente en el ámbito laboral. Es evidente que, si el algoritmo se construye pensando en un perfil determinado de trabajadores, la aplicación del mismo

39. Grupo de Trabajo sobre Protección de Datos del Artículo 29: Dictamen 2/2017 sobre el tratamiento de datos en el trabajo. Sitio web: http://ec.europa.eu/justice/data-protection/index_en.htm17/ESWP249.

reflejará el sesgo programado, y a algunas personas se les negará el acceso al empleo o se les considerará, como en el caso de Sarah Wysocki, trabajadores prescindibles, y, todo ello, además, bajo la pátina de objetividad y legalidad que exhiben los algoritmos.

Sin embargo, no es éticamente aceptable que una persona se vea discriminada por un algoritmo que encierra una lógica discriminatoria. La discriminación artificial es tan reprochable como la realizada mediante la sola intervención humana. El sesgo identificable y discriminatorio debe eliminarse tanto en la fase de acopio de datos, como de entrenamiento y programación de algoritmos.

A este respecto, es preciso promover, como recomienda el Consejo Económico y Social de la UE, en su dictamen de 2018, sobre «Inteligencia artificial», una sensibilización de los expertos y diseñadores de máquinas inteligentes en la posible mala utilización de los resultados de sus investigaciones. Hay que incluir en los programas y planes de estudio de las universidades, directrices éticas sobre la IA para garantizar la protección de derechos como la igualdad y no discriminación.

4) Bajo el cuarto y último principio de explicabilidad se apuesta por que los sistemas de inteligencia artificial sean transparentes y, en la medida de lo posible, explicables a los afectados directamente.

Este principio comporta el cumplimiento de tres aspectos básicos. En primer lugar, la trazabilidad: se debe dejar constancia de los datos y los procesos que determinan la decisión del sistema de IA; lo cual es fundamental para identificar en su caso, los posibles errores y para ayudar a prevenir futuros errores. En segundo lugar, la explicabilidad: en las decisiones que tengan incidencia significativa en la vida de los seres humanos, debería ser posible exigir una explicación adecuada del proceso de toma de decisiones del sistema IA. Y, en tercer lugar, la comunicación: los seres humanos tienen derecho a ser informados de que están interactuando con un sistema de IA y a oponerse a esa interacción con la máquina.

Se ha de notar que el RGPD (UE) no impide la posibilidad de una decisión automatizada sin intervención humana (esto es, que sea tomada por algoritmos y no por seres humanos) cuando sea necesario para la celebración o ejecución de un contrato (art. 22.2 a). Por tanto, en principio, y con carácter general, el trabajador no tendría derecho a oponerse a un procesamiento enteramente automatizado decidido por el empleador cuando está relacionado con el contrato.

Ahora bien, el responsable de tratamiento debe cumplir con una serie de obligaciones, para salvaguardar los derechos e intereses legítimos del mismo. En concreto,

debe informar al trabajador de que va a ser objeto de una decisión automatizada; de la lógica aplicada; y debe explicarle la importancia y las consecuencias previstas de dicho tratamiento [arts. 13.2 f) y 15.1 h)].

Además, una vez tomada por la empresa la decisión automatizada, el interesado tiene derecho a: a) expresar su punto de vista (efectuar alegaciones); b) recibir una explicación de la decisión adoptada después de la evaluación; y c) ejercer su derecho a impugnar la decisión, si considera lesiva de sus derechos (art. 22.3).

Se trata de evitar que la decisión quede exclusivamente en las manos del algoritmo o de la inteligencia artificial y que el trabajador pierda completamente el control de las decisiones que le conciernen. Si queremos una tecnología de trabajo decente y bajo control humano, como pretende la OIT, debemos asegurarnos de que estos sistemas tengan un impacto justo en la vida de las personas y que estén en línea con los valores de la democracia, el estado de derecho y los derechos fundamentales.

Si no lo hacemos así pronto llegarán los desencantos y pasaremos de la utopía digital a la distopía. El espejismo tecnoutópico, sin unos valores y principios éticos que protejan a los individuos y a los grupos potencialmente vulnerables, nos puede llevar a una sociedad indeseable en sí misma.

Esto es todo. Muchas gracias.