



**VNiVERSiDAD
D SALAMANCA**

CAMPUS DE EXCELENCIA INTERNACIONAL

TRABAJO FIN DE GRADO

GRADO EN DERECHO

Departamento de Derecho del Trabajo y Trabajo Social

Área de Derecho del Trabajo y de la Seguridad Social

Curso 2018/2019

Inteligencia artificial y mercado de trabajo

Estudiante: Carmen Hernández González

Tutor: Dr. D. Rafael Sastre Ibarreche

Julio de 2019

Deliveratly left blank

TRABAJO FIN DE GRADO

GRADO EN DERECHO

Departamento de Derecho del Trabajo y Trabajo Social

Área de Derecho del Trabajo y de la Seguridad Social

Inteligencia artificial y mercado de trabajo

Artificial Intelligence and Labor Market

Carmen Hernández González

carhg@usal.es

Tutor: Dr. D. Rafael Sastre Ibarreche

Julio de 2019

Deliveratly left blank

RESUMEN:

La utilización de la inteligencia artificial (IA) puede modificar de manera profunda las relaciones laborales, tanto por su capacidad para destruir y crear puestos de trabajo, como por las consecuencias en la redistribución de la renta. Los cambios pueden presentarse de manera explosiva, lo que traería consigo la necesidad adaptar la actividad del sector público y las normas laborales a las nuevas situaciones. Se trata de cooperar a que los cambios inevitables sean compatible con el orden social o incluso con la democracia. Este trabajo se refiere a algunos de los avances de la IA, para poner en perspectiva la importancia de la materia. Se estudia la complejidad del concepto de IA, para estudiar luego su posible influencia sobre el número de empleos, los cambios pueden afectar a la salud o a la seguridad de los trabajadores, las alternativas para que se mantenga una distribución de la renta equitativa y las recomendaciones de la UE para que el desarrollo futuro de la inteligencia artificial responda a criterios éticos que la hagan fiable. En el contexto laboral, se hace referencia al RGPD y a la legislación española concordante sobre la materia.

PALABRAS CLAVE:

Inteligencia artificial, mercado de trabajo, protección de datos.

ABSTRACT:

The use of artificial intelligence (AI) can profoundly modify labor relations, both because of its capacity to destroy and create jobs, and because of its consequences in the redistribution of income. The changes can occur explosively, which would bring the need to adapt the activity of the public sector and labor standards to new situations. The inevitable changes must be made compatible with maintaining the social order or even with democracy. This work refers to some of the advances of the AI, to place in perspective the importance of the subject. The complexity of the concept of AI is studied, its possible influence on the number of jobs, how the changes can affect the health and safety of the workers and the alternatives to keep an equitable distribution of income. Also, this paper deals with the recommendations of the EU for the future development of artificial intelligence according to ethical criteria that make it trustworthy. In the labor market context, we refer to GDPR and to the Spanish legislation on the matter.

KEYWORDS:

Artificial intelligence, labor market, data protection.

Deliveratly left blank

Índice

Inteligencia artificial y mercado de trabajo

1. Introducción.....	1
2. Los avances de la inteligencia artificial	2
3. Qué es la inteligencia artificial	7
4. Avances tecnológicos y mercado de trabajo.....	14
4.1. Causas del cambio y antecedentes históricos	14
4.2. Cuántos y qué puestos de trabajo están en riesgo de desaparecer	19
4.3. Salud y seguridad del trabajador en el puesto de trabajo	20
5. Políticas públicas ante los cambios (im)previsibles.....	24
5.1. Educación para la transición y fomento de la inteligencia artificial	25
5.2. La Renta Básica Universal y otras formas de abordar la desigualdad	27
5.3. La financiación del sector público: reforma fiscal e IA	31
5.4. El capitalismo de la vigilancia y el Ciberleviatán.....	32
5.5. Los valores éticos promovidos desde la Unión Europea.....	35
5.6. Algunos desarrollos legislativos y jurisprudenciales	39
6. Conclusiones.....	42
7. Bibliografía.....	44

Deliveratly left blank

Inteligencia artificial y mercado de trabajo

CARMEN HERNÁNDEZ GONZÁLEZ

Trabajo de Fin de Grado

Universidad de Salamanca. Curso 2018/2019

1. Introducción

La utilización de la inteligencia artificial (IA) en un creciente número de actividades económicas es susceptible de modificar de manera profunda las relaciones laborales, tanto por su potencial capacidad para destruir y crear puestos de trabajo, como por las consecuencias en la redistribución de renta que estos cambios implican. Los cambios que ya han empezado a producirse pueden presentarse con un ritmo acelerado sin precedentes, lo que traerá consigo la necesidad adaptar la actividad del sector público a las nuevas situaciones. Se trata de cooperar a que los cambios inevitables no generen cargas insoportables para una parte de la sociedad, tanto por la pérdida de puestos de trabajo como por dar lugar a una distribución de la renta que no sea compatible con el orden social o incluso con la democracia. Este trabajo se refiere en primer lugar a algunos de los avances de la IA que de manera explosiva están implementándose desde hace apenas un lustro, para poner con ello en perspectiva la importancia de la materia. A continuación se estudia la complejidad del concepto de IA, basado en avances de las matemáticas que se han producido al amparo del aumento de la capacidad de los ordenadores para captar, transmitir y elaborar masivamente datos. La cuarta parte se refiere al mercado de trabajo, constatando la falta de certeza sobre el número de empleos que podrían destruirse y crearse al introducir IA en muchas actividades económicas. En cualquier caso, la IA va a cambiar el modo en que se ejecutan algunas tareas, introduciendo nuevas formas de organizar el trabajo susceptibles de afectar a la salud o a la seguridad de los trabajadores, lo que se analiza de manera general en los puntos siguientes.

El apartado quinto estudia las propuestas de políticas públicas para cooperar a una transición menos traumática ante cambios que pudieran ser drásticos, pero que son en buena parte imprevisibles. Por ello, se recomienda que el sector público cambie el papel del sistema educativo, para que coopere a la formación de los trabajadores para los nuevos

puestos de trabajo. Los cambios en la distribución de la renta se pueden abordar procurando la protección de los trabajadores (más que de los puestos de trabajo). La renta básica universal (RBU) es la propuesta más discutida, con puntos de vista diametralmente opuestos en cuanto a su viabilidad y conveniencia, aunque existen otras que siguen aproximaciones más tradicionales que adaptan modelos de subsidios ya conocidos. En cualquier caso, se pide con insistencia una reforma fiscal que recaude impuestos de las grandes empresas tecnológicas o de los robots.

Veremos también que no faltan autores que ven en los cambios ya acaecidos y en los que se anuncian, un peligro de socavamiento general del sistema democrático, contemplando la llegada de una sociedad distópica con distintos matices. Para evitar estos indeseables y profundos cambios, se proponen iniciativas públicas en el campo del derecho, por ejemplo, regulando los grandes oligopolios o redefiniendo los derechos a la privacidad. En esta línea nos referimos en este trabajo a las propuestas y desarrollos legislativos de la Unión Europea para implantar una IA *fiable*, basada en el respeto de los valores y derechos consagrados en el conjunto de la legislación europea, que no se deben poner en peligro. La Unión Europea ha avanzado recientemente en el diseño de una visión general de los principios que deben inspirar la aplicación de técnicas de IA, pero con la aprobación del Reglamento General de Protección de Datos (RGPD), también ha introducido ya cambios legales para garantizar derechos de los ciudadanos en general y de los trabajadores en concreto. A este RGPD ha seguido en España la Ley Orgánica 3/2018, a la que también haremos referencia también en el apartado final del trabajo.

Aunque la literatura sobre esta materia es abundante y no es posible abarcar todos los aspectos técnicos ni todas las implicaciones sociales del triunfo de la automatización y la robótica que van unidas a la IA, se ha procurado incorporar aquí alguna de la más reciente bibliografía nacional e internacional sobre los aspectos tratados.

2. Los avances de la inteligencia artificial

La inteligencia artificial (IA) representa un avance extraordinario en las capacidades de la tecnología para modificar un gran número de ámbitos de la actividad humana, desde el transporte, con la introducción de vehículos autónomos, hasta la configuración de buena parte de la actividad de cualquier empresa. En su aplicación práctica, la IA es un fenómeno que podría considerarse en ciernes, o en las primeras fases de implementación.

Tiene ya importantes consecuencias prácticas en la vida de los trabajadores y de los ciudadanos en general, pero las expectativas de su proyección futura son extraordinarias. Estamos ante una revolución que no tiene una fecha clara de comienzo, pero que ha tomado impulso partir del año 2017 y para la que el año 2015 podría considerarse todavía parte de *los viejos tiempos* en el calendario digital (Baldwin, 2019, p. 2 y 96).

El concepto de *inteligencia artificial* aparece por primera vez en el año 1956¹, pero el fenómeno de la IA que estamos viviendo ha despertado sólo recientemente, especialmente para los reguladores públicos y las autoridades económicas de los distintos gobiernos. Así vendría a corroborarlo que, por ejemplo, sólo desde 2018 la Comisión Europea se ha ocupado de establecer una estrategia común en relación con la IA (European Commission, 2018b), y sólo más recientemente en la Unión Europea un grupo de expertos de alto nivel sobre la materia se ha encargado de desarrollar directrices éticas para las aplicaciones de la IA y ha avanzado en la concreción práctica del mismo concepto de IA (European Commission. High-Level Expert Group on Artificial Intelligence. AI-HLEG, 2019a, 2019c).

Distintos países europeos y también China o Estados Unidos priorizan la planificación del desarrollo de la IA desde fechas igualmente recientes, en el entorno de los años 2017 y 2018². Los hitos más espectaculares de la IA se han alcanzado también en momentos próximos en el tiempo. Podemos citar los logros en el ámbito de los juegos. Los ordenadores pueden superar a los mejores ajedrecistas desde hace algunos años³, y

¹ Un grupo de profesores se reunió en la Universidad de Dartmouth, New Hampshire, Estados Unidos, con el propósito de encontrar “cómo hacer que las máquinas utilicen el lenguaje, formen abstracciones y conceptos, resuelvan tipos de problemas que ahora están reservados para los humanos y se mejoren a sí mismas”. Su intención era lograrlo en el transcurso del verano (McCarthy, Minsky, Rochester, & Shannon, 1955), pero la investigación sobre IA ha pasado desde entonces varios *inviernos* hasta su más reciente éxito.

² Lo mismo puede decirse de las iniciativas sobre IA de muchos países. En Francia, el Presidente Macron presentó la estrategia francesa sobre IA el 29 de marzo de 2018; en Alemania la fecha de referencia es noviembre de 2018, en el Reino Unido marzo de 2018. Sólo algunos países europeos tienen iniciativas sobre la materia en el año 2017 (por ejemplo, Suecia, Finlandia y Austria); en Estados Unidos, la llamada *IA Initiative* se aprobó por el Presidente Trump en 2019, fijándose por primera vez la materia como área prioritaria en el presupuesto federal; en China, país que para algunos es líder en los avances sobre la materia, los planes públicos de desarrollo de la IA datan también de los años 2017 y 2018 (Servoz, 2019, p. 139 y siguientes, anexo I). La Alianza Europea de la IA se creó en julio de 2018 (European Commission. High-Level Expert Group on Artificial Intelligence. AI-HLEG, 2019b).

³ El ordenador de IBM Deep Blue venció al campeón mundial Garry Kasparov en 1997. A partir de entonces, los programas de IA ayudaban a mejorar durante unos años a los campeones de ajedrez. Equipos *centauros*, formados por hombres y máquinas eran capaces de vencer a las máquinas durante algunos años. En años más recientes, los ordenadores han mejorado tanto que ha dejado de ser noticia el hecho de que superen a cualquier humano. Lo que fue noticia en 2017 fue que el programa AlphaZero de Google venció al programa de ajedrez más avanzado hasta entonces, Stockfish 8 campeón de 2016, capaz de calcular 70 millones de posiciones por segundo. AlphaZero ganó utilizando la última tecnología de aprendizaje

también a los mejores jugadores de *go*, un juego chino de una complejidad muy superior al ajedrez⁴. En diciembre de 2017 el programa de ordenador AlphaGo Master, propiedad de la empresa de IA Deep Blue, ganó al campeón mundial de *go*. Lo hizo a partir del estudio de 30 millones de posiciones en 160.000 juegos reales, y *entrenando* consigo mismo. Esta victoria supuso un logro impresionante que fue superado a los seis meses, cuando una nueva versión del programa llamada AlphaGo Zero fue capaz de vencer al que era entonces el mejor jugador del mundo, el programa AlphaGo Master. La diferencia fue que el nuevo programa no tuvo que analizar ningún juego real previo y aprendió a jugar solo y en solamente cuarenta días (Baldwin, 2019, p. 107).

Las aplicaciones de IA más recientes son cada vez más precisas en tareas como el reconocimiento facial, el entendimiento de cualquier idioma en forma oral o escrita, la lectura de escritura a mano, etc. Igualmente, ahora es posible identificar a las personas a través de características biométricas como su forma de andar o el modo singular como llevan o utilizan su teléfono móvil⁵, lo que plantea serias dudas en cuanto a la privacidad del individuo o el potencial uso de la información para usos no previstos, si bien aporta la capacidad de detectar y evitar el uso fraudulento de los teléfonos y otros aparatos electrónicos (The Economist, 2019b, 23 de mayo).

En la misma línea, los asistentes personales como *Siri* o *Alexa* son ya familiares a un número creciente de personas, y gracias a la IA alcanzan capacidades cada vez mayores en la selección de los contenidos o las respuestas precisas que el usuario demanda. La voz se está revelando como una forma eficaz de manejar esos asistentes, pero también como una fuente de datos susceptible de análisis mucho más allá de los que podríamos esperar hace poco tiempo.

automático (*machine-learning*), realizando sólo 80.000 cálculos por segundo y aprendiendo a jugar al ajedrez de forma automática en solo cuatro horas (Harari, 2018). Una descripción de cómo enseñó Apple a Siri el idioma Chino en su variante de Shanghai, puede verse en Baldwin, 2019, pp. 108-110.

⁴ Después de dos movimientos, un jugador de ajedrez puede elegir entre 400 movimientos, en tanto que el juego *go* las posibilidades son 130.000 en el mismo caso, y la complejidad luego crece. Hay más posiciones posibles en un tablero de *go* que átomos en el universo (Baldwin, 2019, p. 106).

⁵ El llamado *comportamiento biométrico* puede analizarse a partir del uso del teléfono móvil. En el pasado reciente, era posible identificar un teléfono concreto y vincularlo a su propietario a partir del conocimiento de datos como las aplicaciones que tuviera instaladas, el modelo de teléfono y el sistema operativo (lo que se conoce como *device fingerprinting*). Los desarrollos más recientes lo sustituyen por el *behavioural biometrics*, que se basa en los datos que revelan los sensores del teléfono sobre la forma en que cada individuo lo coge o lo maneja. La empresa UnifyID es capaz de identificar alrededor de 50.000 formas de andar individuales, que combinadas con la diferente presión de los dedos sobre la pantalla y otros datos permiten identificar al usuario de cada teléfono

Numerosos servicios de atención a los clientes pueden, mediante algoritmos que analizan la voz, detectar el estado de ánimo, las actitudes y las emociones de la persona con la que se comunican. Esto puede facilitar la atención más precisa de las necesidades de cada uno, incluso de ámbito médico⁶. En un avance que pudiera parecer extraordinario, se está experimentando con la posibilidad de reconstruir el rostro de una persona a partir sólo de una breve grabación de su voz, con unos resultados iniciales sorprendentes, como podemos juzgar a partir de la Imagen 1 siguiente.

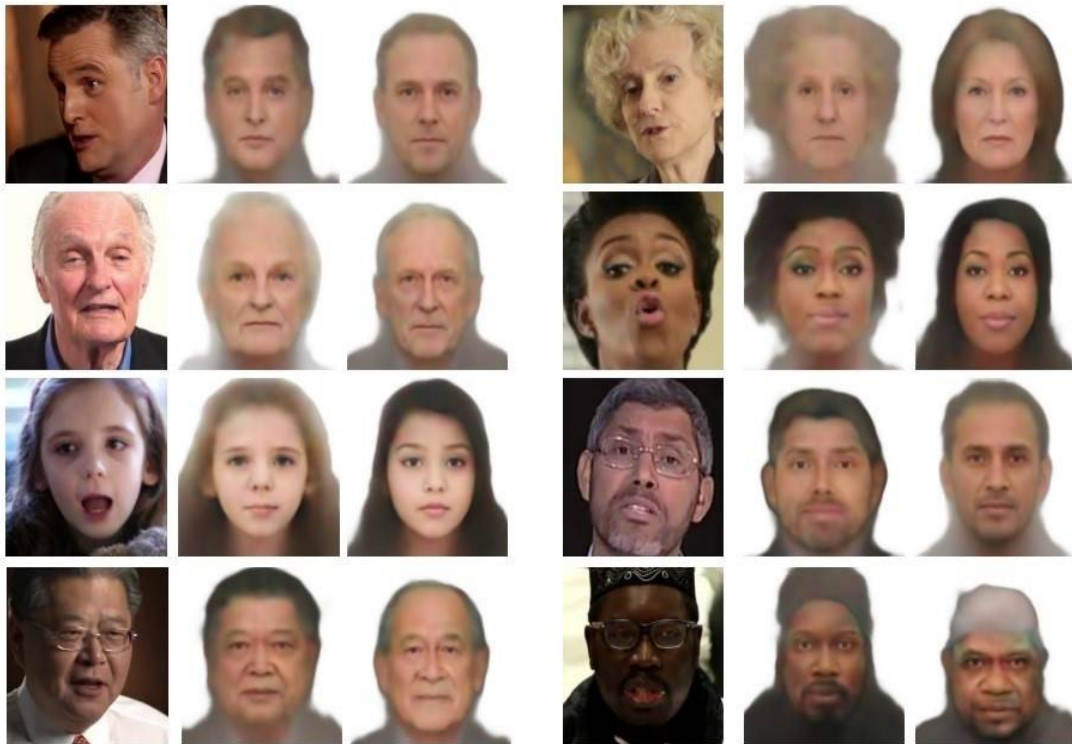


Imagen 1. Retratos de los individuos de las fotografías generados mediante inteligencia artificial a partir únicamente del análisis de un fragmento de seis segundos de su voz. El sistema tiene una precisión del 94 por ciento en la determinación del género y alcanza una buena correlación en la estructura de la nariz y la boca, principales determinantes del habla. Fuente: The Verge, 2019 y <https://speech2face.github.io/>

Los avances recientes en el reconocimiento y manipulación de imágenes son también importantes. Son esenciales para el perfeccionamiento de la conducción autónoma de vehículos, así como en el reconocimiento facial y la manipulación de imágenes. Es posible obtener así lo que se denominan *deep fakes*, que no son otra cosa que vídeos por completo falsos en los que se hace aparecer la imagen de cualquier persona haciendo o diciendo cosas que nunca ha hecho o dicho, pero que son difíciles de distinguir de registros de vídeo o audio verdaderos. En esta línea de avances espectaculares, la Imagen

⁶ La voz se puede usar para determinar qué pacientes tienen una más urgente necesidad de atención en un hospital. Las compañías de seguro analizan la voz para estimar las probabilidades de que un usuario esté mintiendo o diciendo la verdad sobre los hechos que fundamentan su reclamación.

2 nos muestra una serie de rostros con aspecto real pero que en realidad no corresponden a ninguna persona, sino que han sido elaborados con herramientas de IA⁷.

En el texto que sigue haremos referencia a otros avances y proyectos en el ámbito de la IA, especialmente por lo que se refiere a sus consecuencias en el ámbito general del mercado de trabajo y en concreto en el mundo sindical, en el que el uso de herramientas de IA afecta al desenvolvimiento individual del trabajador en su puesto de trabajo. Son dos aspectos relacionados, pero distintos de las consecuencias para el trabajador de las nuevas herramientas. Por un lado, es posible que su puesto de trabajo esté en peligro como



Imagen 2. Estas personas 'fotografiadas' en esta imagen no existen, sus rostros han sido generados mediante algoritmos de inteligencia artificial. Fuente: The Verge, 13 de junio de 2019.

consecuencia de la automatización total o parcial de las tareas que realiza, pero por otro lado, incluso si su puesto de trabajo no pelagra, es posible que tenga que enfrentarse a cambios en el desarrollo de su actividad derivados, por ejemplo, de la pérdida de privacidad, de la supervisión constante por parte del empresario o de la evaluación de su

⁷ La IA aplicada a los *deep fakes* ha entrado hace algún tiempo en la pornografía, pero llega a hacerse accesible a cualquiera. Por ejemplo, la aplicación *DeepNude* es capaz en solo unos minutos de reproducir el cuerpo desnudo de una mujer a partir de una fotografía cualquiera de la misma (The Verge, 27 de junio de 2019: <https://www.theverge.com/2019/6/27/18760896/deepfake-nude-ai-app-women-deepnude-non-consensual-pornography>). Además, los algoritmos para detectar las falsificaciones no remedian el daño, incluso si fueran perfectos (The Verge, 27 de Junio de 2019: <https://www.theverge.com/2019/6/27/18715235/deepfake-detection-ai-algorithms-accuracy-will-they-ever-work>).

desempeño por procedimientos automatizados con sesgos o no suficientemente transparentes. A todo ello nos referiremos en las páginas siguientes, pero antes vamos a ocuparnos de explicar al menos brevemente el concepto de *inteligencia artificial*.

3. Qué es la inteligencia artificial

Una definición de inteligencia artificial es probablemente necesaria si vamos a estudiar las consecuencias de su desarrollo presente y futuro para la sociedad en general y más concretamente en el mercado de trabajo. No es fácil dar una definición. En la literatura podemos encontrar un buen número de ellas, pero en general parece que pretenden *salir del paso* más que llegar a una definición cerrada y comprensible. Con frecuencia las definiciones son poco concretas, por ejemplo del Rey Guanter (2018, p. 32), después de no abordar la materia en un apartado dedicado a ello (*op. cit.* p. 31), afirma que “la inteligencia artificial es un marco que engloba diferentes especialidades que comparten un objetivo común: dotar a un sistema artificial de cierto grado de inteligencia” (del Rey Guanter, 2018, p. 32).

Lo cierto es que lo que se conoce como IA es una etiqueta genérica que abarca un variado conjunto de técnicas de predicción y toma de decisiones. El fundamento último de estas técnicas reside en complejos avances en distintos campos de las matemáticas en los que se basa el diseño de algoritmos de muy distinta naturaleza, capaces unas veces de análisis biométricos y otras del reconocimiento de imágenes, de la traducción de un texto o de jugar y ganar en un videojuego. Las distintas *familias* de algoritmos las explica de manera relativamente intuitiva Domingos (2015), que aspira a que en un distante futuro pueda existir un único todopoderoso *Master Algorithm*. Nosotros no vamos a entrar en estos aspectos técnicos, pero creemos deseable tener presente que el campo de la IA se fundamenta en avances en el campo de las matemáticas aplicadas que han sido posibles por la extraordinaria mejora de la capacidad de proceso de datos de los ordenadores actuales.

Tal vez la dificultad de definir lo que es la IA esté en el hecho mismo del nombre elegido para incluir comprensivamente el conjunto de campos que abarca. Baste recordar que el

concepto de *inteligencia* en su acepción tradicional está lejos de ser pacífico en el mundo de la psicología⁸.

Incluso el documento más reciente publicado por la Comisión Europea en relación con el concepto de IA, al que vamos a referirnos a continuación, aborda la definición de IA con un sonoro *'disclaimer'* advirtiéndole que su descripción y definición de las capacidades y áreas de investigación de la IA “es una muy cruda simplificación del estado de la técnica” (European Commission. High-Level Expert Group on Artificial Intelligence. AI-HLEG, 2019a).

La propuesta de definición de este Grupo de Expertos de la Comisión Europea es la siguiente:

“Los sistemas de inteligencia artificial (IA) son sistemas de software (y posiblemente también hardware) diseñados por humanos que, dado un objetivo complejo, actúan en la dimensión física o digital percibiendo su entorno a través de la adquisición de **datos**, interpretando los datos estructurados o no estructurados, razonando en el conocimiento, o el **procesamiento de la información**, derivada de estos datos y decidiendo la mejor alternativa o **alternativas de acción para lograr el objetivo dado**. La IA puede usar reglas simbólicas o elaborar un modelo numérico, y también puede **adaptar su comportamiento** analizando cómo el entorno se ve afectado por sus acciones previas” (2019a, p. 6).

El documento añade que como disciplina científica, la IA incluye varios enfoques y técnicas, como los siguientes:

1. *El aprendizaje automático [machine learning]*, que incluye por ejemplo el aprendizaje profundo [*deep learning*] y el aprendizaje por refuerzo [*reinforcement learning*].
2. *El razonamiento automático [machine reasoning]*, del que forman parte la planificación, la programación, la representación del conocimiento y del razonamiento, y la búsqueda y optimización.

⁸ Puede verse como ejemplo la página web <https://www.psicoadictiva.com/blog/concepto-inteligencia-ha-evolucionado/> donde se afirma que “los psicólogos llevan más de cien años intentando definir la inteligencia, y parece que todavía no hay dos que digan lo mismo en torno a este concepto”.

3. La robótica, que abarca campos como el control, la percepción, los sensores y actuadores, además de la integración en sistemas cibernéticos y físicos de todas las demás técnicas.

Siguiendo el texto citado, y a riesgo de esquematizar abusivamente lo que ya sabemos que es simplificado, vamos a intentar una descripción intuitiva de lo que es la IA. Cualquier sistema que utilice la IA necesita *datos* que pueden proceder de todo tipo de sensores adecuados para la tarea que se pretende realizar. Por ejemplo, si queremos que una aspiradora limpie el suelo de nuestra casa⁹ los datos pueden proceder de cámaras, micrófonos, teclados, sensores táctiles, de presión o de distancia, etc. (esto lo decidirá el diseñador del sistema). Los datos pueden estar organizados con algún criterio o no estarlo, lo que quiere decir que los datos pueden ser *estructurados* o *no estructurados*.

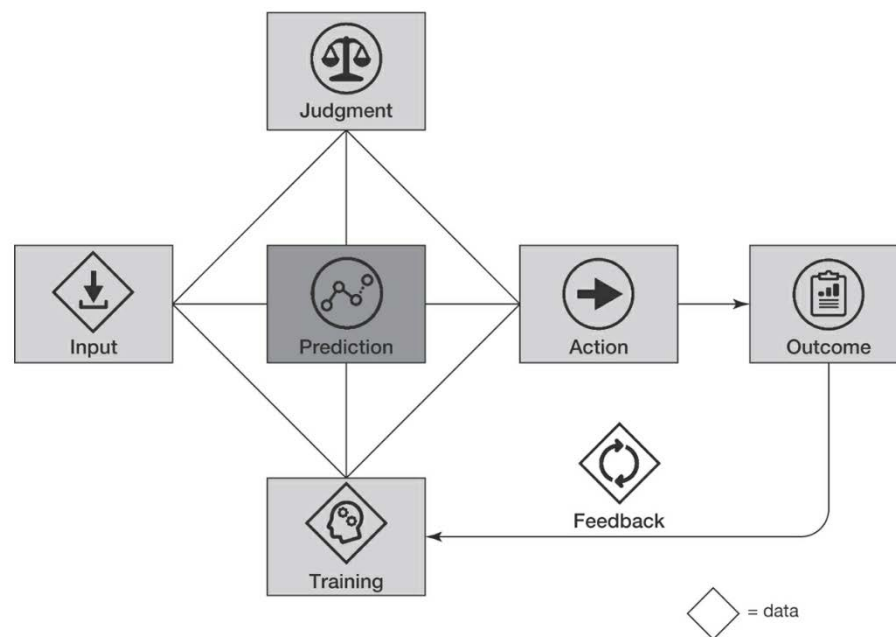
Con los datos correspondientes, el sistema de IA del que estuviera dotado nuestro aspirador doméstico sería capaz de procesar la información (*razonar* a partir de los datos) y decidir la mejor forma de realizar su función. Por ejemplo, a partir de una imagen del suelo de nuestra vivienda (que para la máquina es una secuencia de datos digitalizados, ceros y unos), tendría que (1) interpretar o *saber* si el suelo está o no limpio y (2) basándose en ello obtener un modelo numérico, una fórmula matemática, que establezca qué ‘decisión’ debe tomarse: limpiar o no, con qué intensidad, qué trayectoria seguir, etc.

En este ejemplo, una vez *tomada la decisión* el aparato pone en marcha los motores, los cepillos o los filtros que proceda para llevar a cabo la limpieza. Estos motores, cepillos y filtros serían los elementos *actuadores* o ejecutores de la tarea (si no se trata de un robot o aparato físico, la *actuación* podría ser la activación de algún tipo de *software*, por ejemplo, respondiendo a alguna pregunta si se trata de un sistema *asistente* que utiliza IA).

Si además nuestro aspirador es capaz de percibir el entorno en el que trabaja y de reconocer cómo cambia ese entorno después de su actuación, podría estar dotado de la capacidad de adaptar su actividad a cada nueva situación a fin de limpiar mejor (alcanzar mejor su objetivo). En ese caso se trataría de un sistema racional *capaz de aprender*.

⁹ Adaptamos el ejemplo del mismo documento (European Commission. High-Level Expert Group on Artificial Intelligence. AI-HLEG, 2019a)

Agrawal *et al.* (2018) destacan en este proceso que lo que permite hacer la IA es *predecir*. Consideran que la capacidad de predecir es la esencia de la IA, y que el hecho de que obtener predicciones fiables sea ahora mucho más barato es la clave del éxito de la IA. Según esto, la IA es una tecnología para hacer predicciones, y esas predicciones son el *input* para la toma de decisiones. Traslado esta aproximación a nuestro ejemplo anterior, nuestro aspirador doméstico parte de los datos que obtiene (a través de cámaras, sensores, etc.) y los usa para generar una predicción, en este caso la mejor forma de realizar la limpieza. A partir de ahí, la parte *robótica* de la máquina (lo que hemos llamado los actuadores) se programa para que de manera automática se ponga en funcionamiento de la manera *prevista*. El resultado final de las acciones vendría a aportar nuevos datos que ayudarían a mejorar el sistema basado en la IA, que *aprendería* así a partir de la información adicional derivada de su funcionamiento, la nueva información ayuda al *entrenamiento* del sistema (Agrawal *et al.*, 2018).



Esquema 1. El proceso de toma de decisiones. Los datos son el input del sistema, estos datos permiten hacer una predicción (comparando con los millones de otros datos que han 'entrenado' al sistema). La predicción debe ser valorada (judgement) por el humano decisor. La decisión da lugar a un resultado que, a su vez, ayuda a mejorar o entrenar el sistema. Fuente: Agrawal *et al.*, 2018.

Cuando las predicciones son suficientemente buenas y los riesgos de error no implican graves consecuencias, el sistema puede diseñarse para *decidir* actuar automáticamente (en nuestro ejemplo, mediante la parte robótica de la máquina, ejecutar la limpieza del suelo), pero no siempre es necesario ni conveniente que las acciones se lleven a cabo sin

intervención humana, que puede ser imprescindible para realizar una *evaluación o juicio* sobre la alternativa más deseable.

Cuando los riesgos son grandes y la necesidad de acción es inmediata, el sistema solo podría *decidir* por sí mismo cuando la predicción sea extremadamente precisa. Por ejemplo, para que los vehículos autónomos puedan ser una realidad deben ser capaces de *predecir* lo que va a pasar en millones de situaciones, ya que algunos errores pueden poner en peligro la vida humana. La tecnología no será viable hasta que se alcance una muy elevada capacidad de evaluar correctamente el entorno, para poder juzgar y decidir automáticamente sin intervención humana (por ejemplo, si es necesario frenar, girar, etc.).

En otros casos la IA puede realizar las predicciones, planteando un escenario al humano que debe juzgar y tomar la decisión. Por ejemplo, si se trata de seleccionar trabajadores para un puesto de trabajo a partir de sus currículums o de establecer quién debe salir de la cárcel en libertad condicional, las aplicaciones de IA *predicen* lo que consideran más conveniente a partir de los datos con los que se han desarrollado (que podrían tener sesgos, como veremos), pero son las personas las llamadas a juzgar cuál es la decisión más conveniente, siguiendo o no la información que la predicción proporciona.

Algunos problemas pueden plantearse al establecer las fronteras, al establecer cuando *fiarse* de un sistema inteligente o cómo usar la información que nos proporciona. En algunos casos habrá que ponderar la importancia de los errores humanos con la de los errores del sistema inteligente, o los sesgos humanos con los sesgos del sistema inteligente. Por ejemplo, en el reconocimiento de imágenes, ya hace algunos años que algunas aplicaciones superan a los humanos, es decir, cometen menos errores¹⁰. No querrá esto decir que pueda prescindirse de toda supervisión humana (por ejemplo, en el caso de la interpretación de imágenes médicas), pero las predicciones facilitadas por las aplicaciones basadas en la IA es probable que sean cada vez mejores.

¹⁰ En el *ImageNet Challenge*, celebrado hasta 2017, se trataba de predecir el nombre del objeto representado en una imagen. Los datos de partida contienen mil categorías de objetos, incluyendo por ejemplo distintas razas de perro y otras cosas similares. Los humanos cometen en torno a un cinco por ciento de errores. La IA se introdujo en el año 2012 mejorando con ello mucho la precisión de los sistemas informáticos, que superaron a los humanos a partir del año 2015 (Agrawal et al., 2018).

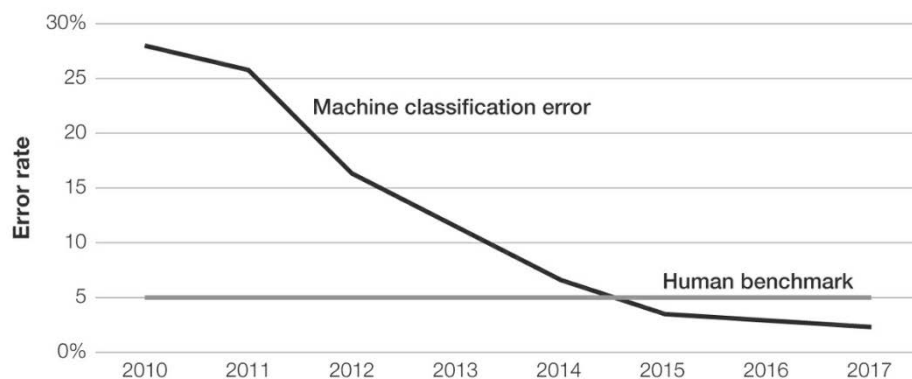


Gráfico 1. Resultado del ImageNet Challenge. Desde 2015 los sistemas de inteligencia artificial cometen menos del 5% de errores, que son los cometidos por los humanos. Fuente: Agrawal et al, 2018.

En el ámbito científico de la IA hay distintas técnicas que pueden considerarse campos de estudio independientes dentro del área general de la IA. Estas técnicas pueden utilizarse simultáneamente y a veces aparecen confundidas o separadas en la literatura. Los avances en el uso de estas técnicas son la base de las predicciones de la IA. Entre las más conocidas de estas técnicas están el *aprendizaje automático*, las *redes neurales* y el *aprendizaje profundo* (*machine learning*, *neural networks* y *deep learning*). Estas técnicas permiten a los sistemas de IA aprender a resolver problemas que no pueden ser especificados con precisión, o cuyo método de solución no puede ser descrito mediante reglas de razonamiento simbólico; éste es el caso del reconocimiento del lenguaje hablado y escrito o la predicción de comportamientos. En todos los casos el resultado de estas técnicas es un modelo numérico, una fórmula matemática, que permite llegar a una decisión a partir de los datos (European Commission. High-Level Expert Group on Artificial Intelligence. AI-HLEG, 2019a, p. 3)¹¹.

¹¹ Es conveniente comprender la diferencia entre las técnicas matemáticas utilizadas en la IA y las técnicas estadísticas o de programación 'tradicionales'. Los habituales modelos estadísticos de regresión multivariante son superados en su capacidad de predicción por las técnicas de aprendizaje automático que utilizan simultáneamente miles de variables relativas a millones de situaciones. También se distingue la IA de las técnicas de programación previas en que ahora no se intenta establecer dentro del programa todas las alternativas o situaciones posibles, sino que se alimenta el sistema con miles o millones de ejemplos con las que las técnicas de aprendizaje automático pueden realizar predicciones correctas (con mayor o menor intervención humana en la fase de *entrenamiento* del algoritmo correspondiente). En un ejemplo clásico, en vez de intentar programar simbólicamente todas las características de un gato en las distintas situaciones posibles, se proporcionan al sistema millones de imágenes etiquetadas correctamente y el sistema *aprende* a distinguir cuando una imagen que no está entre las utilizadas para entrenar el sistema tiene o no tiene un gato.

Como cabría esperar, la calidad del resultado al que se llegue depende crucialmente de la calidad de los datos iniciales. Si los datos no son buenos o contienen sesgos, los resultados no serán de la mejor calidad y estarán también sesgados. Por ejemplo, si se trata de seleccionar trabajadores y los datos que se han proporcionado al sistema incluyen pocas mujeres o pocas personas de un determinado grupo, es probable que el algoritmo esté sesgado en contra de esos grupos al presentar sus resultados.

Un problema conectado con el anterior es que los sistemas de IA son muy poco o nada transparentes. Es decir, que aunque sus resultados sean muy precisos y sus predicciones fiables, no es posible saber cómo se llega a ellos. Funcionan como una auténtica *caja negra*, que puede hacer imposible explicar a los afectados cómo o por qué se ha llegado a la predicción correspondiente, en la que pueden haberse basado decisiones importantes (por ejemplo, despedir a un trabajador, promocionarlo, aumentarle el sueldo, etc.).

Por último, debemos señalar que las aplicaciones de IA, incluso si pueden superar las capacidades humanas, están limitadas a un rango concreto de actividades que no tienen un tipo de *inteligencia general*, que es la humana, sino otra más estrecha o débil (*narrow* o *weak*)¹². La definición de uno u otro tipo de inteligencia no siempre es clara, se dice que la inteligencia general implica la conciencia de la propia existencia, sentimientos propios, lo que está lejos de ocurrir en las aplicaciones de IA, que hoy no son en ningún caso de *inteligencia general* (Harari, 2018)¹³.

¹² Las aplicaciones IA no pasan todavía el test de Turing. En 1950 Alan Turing se planteó si las máquinas podrían pensar. A su juicio, la pregunta no puede contestarse por las dificultades para definir lo que es 'pensar', pero propuso el siguiente test práctico: Un ordenador debería considerarse inteligente si es capaz de conversar con alguien a través de una conexión electrónica de modo que el interlocutor no se dé cuenta de que no está conversando con un humano. La conversación se supone que tendría lugar a través de teclados y pantallas de ordenador, con el humano situado en una habitación diferente. Ver, por ejemplo, Ramge (2019, pp. 31–32).

¹³ Los robots y otras herramientas de IA no pueden desarrollar sentimientos, aunque puedan ser capaces de interpretar los sentimientos humanos a partir de datos biométricos o de otro tipo. No obstante, los humanos pueden tener sentimientos hacia los robots. Por ejemplo, sentir pena al ver el mal trato de un robot con forma humana (<https://www.theverge.com/tldr/2019/6/17/18681682/boston-dynamics-robot-uprising-parody-video-cgi-fake>) o lamentar su fallecimiento (<https://www.theverge.com/2019/6/19/18682780/jibo-death-server-update-social-robot-mourning>).

4. Avances tecnológicos y mercado de trabajo

4.1. Causas del cambio y antecedentes históricos

Los inciertos cambios introducidos por la IA en la economía en general y en los mercados de trabajo en particular se tratan de anticipar con una revisión de la historia. El problema es que no todos interpretan la historia de la misma manera, y las conclusiones a las que llegan son considerablemente diferentes. Para algunos, la IA es un fenómeno temible que dará lugar a la desaparición de numerosísimos empleos, especialmente los de baja cualificación, más automatizables. Para otros, la IA creará tantos o más empleos como los que desaparezcan y el volumen de empleo no es un aspecto preocupante, aunque haya que prever algunos cambios, como los sistemas educativos o los mecanismos de redistribución de la renta. Con carácter general, las propuestas en cuanto a las iniciativas públicas relevantes están influidas en cada caso por planteamientos ideológicos generales en cuanto al papel que debe desempeñar el sector público y la mayor o menor preocupación por la desigualdad.

Los argumentos *tranquilizadores* sobre el futuro inmediato proceden al menos de dos frentes que pueden aparecer combinados. Por una parte, la observación de la actualidad de los mercados de trabajo, y por otro la observación del hecho histórico de que anteriores cambios tecnológicos no destruyeron empleo en términos agregados. En esta línea, se destaca que en este momento, salvo en algunos países como España, se están alcanzando muy bajos porcentajes de desempleo, y además también altos niveles de empleo, especialmente por la incorporación de la mujer al mercado de trabajo. Un reciente informe de la revista *The Economist* se refiere al aumento sin precedentes de la disponibilidad de puestos de trabajo (The Economist, 2019a). Las cifras que recoge este informe son contundentes: en el conjunto de la OCDE, la tasa de desempleo no había sido tan baja en varias décadas, con algunos países como Estados Unidos alcanzando cifras de desempleo por debajo de las vistas en más de cincuenta años (3,6%). En 2018 la tasa de actividad era la más alta alcanzada nunca en países como Gran Bretaña, Canadá, Alemania, Australia y otros 22 países de la OCDE. Sólo en Italia, España y Grecia la tasa de desempleo es superior a las cifras previas a la crisis económica de 2008¹⁴, y el informe al

¹⁴ A pesar de ello, el número de cotizantes a la Seguridad Social en España está en este momento muy próximo al alcanzado antes de la crisis, algo menos de 20 millones de personas. Otra cosa pueden ser las bases de cotización, puesto que los salarios de muchos trabajadores son menores que antes.

que nos estamos refiriendo lo atribuye a la rigidez de las normas laborales que impiden el ajuste de los mercados. La conclusión es que, si bien la expansión de los servicios ha hecho que haya menos empleos para personas con una formación media, hay que poner en duda las predicciones pesimistas sobre las consecuencias de cualquier nueva tecnología en el futuro del trabajo. Más bien la lección que se obtiene de la experiencia de los últimos quinientos años es que el cambio tecnológico crea empleos más que destruirlos y, en consecuencia, las ya buenas cifras actuales es probable que mejoren en el actual periodo de expansión económica.

Un informe reciente sobre el futuro del mercado de trabajo publicado por la Comisión Europea y cuyo autor es Michel Servoz (2019) se alinea con los más optimistas. La segunda página del largo informe sobre la materia está exclusivamente dedicada a recoger el lema que preside el informe: “*Don’t panic*”¹⁵. Desde luego, habría poco que temer si se cumple, como considera el informe, que la revolución digital no sólo creará más empleos que los que destruirá, sino que también mejorará la calidad de los empleos y los servicios, si bien esto sucederá *a largo plazo*, con la aparición de empleos que ahora no nos podemos imaginar. Parece más una llamada a tener fe en el futuro que un sólido argumento, especialmente si se considera que el propio autor acepta que habrá trastornos significativos en el mercado de trabajo a corto plazo. Afirma que los avances de la IA y la robótica tienen el potencial de producir cambios nunca vistos desde la revolución industrial y que para gestionar la transición y mitigar los riesgos potenciales es de importancia capital la intervención de los gobiernos europeos (Servoz, 2019, p. 13). De hecho, el informe plantea luego distintos escenarios, el segundo de los cuales considera creíble que a medio plazo haya alteraciones importantes que lleven a la completa reorganización de algunos trabajos, como los vinculados al sistema judicial, la selección de recursos humanos o concesión de préstamos bancarios. El informe además propone distintas reformas y actuaciones públicas de todo orden para hacer viable sin sobresaltos la transición digital.

La mayoría de los autores son cautos. Las publicaciones más recientes avisan al menos de la posibilidad de una reducción drástica de puestos de trabajo para personas con baja y media cualificación. Si no se puede predecir el futuro, menos se pueden conocer los detalles de los cambios que nos esperan, pero la dirección del cambio sí es previsible y la

¹⁵ Palabras que constan también en grandes caracteres en la cubierta de ‘*the hitchhiker’s guide to the galaxy*’, de Douglas Adam, 1979 (Servoz, 2019, p. ii).

posibilidad de que sea drástico entraña un gran peligro para muchos. Prepararse para lo que pudiera pasar tiene todo el sentido.

Pero si las revoluciones anteriores han traído más empleo y crecimiento, ¿por qué esta vez podría ser diferente? Ya hemos vivido una Revolución Industrial que hizo desaparecer la mayoría de los puestos de trabajo en la agricultura, y una transformación de la industria que ha sustituido los puestos de trabajo en ese sector para crearlos en el ámbito de los servicios. Sin embargo, el análisis histórico pone de manifiesto diferencias con el fenómeno actual, además de destacar que aquellos cambios tampoco se produjeron sin sacrificio. Para Baldwin (2019) existen dos diferencias con el pasado: la velocidad con la que se van a producir los cambios y la enorme sensación de injusticia que pueden introducir.

Los cambios se están ya produciendo *a un ritmo explosivo*, es decir, que supera la capacidad de adaptación del sistema. No se podrán reemplazar los puestos de trabajo perdidos a la velocidad necesaria, y esto puede dar lugar a reacciones violentas y a la ruptura de la estructura social, especialmente si la frustración de muchos trabajadores es recogida por partidos políticos populistas *à la* Trump o los movimientos pro Brexit¹⁶.

La transformación que en el pasado se produjo de una sociedad agraria a una industrial, y la transformación posterior que nos llevó a la actual sociedad de servicios se produjeron en intervalos de tiempo muy diferentes. Para Baldwin, la primera *Gran Transformación* llevó a la población a las ciudades para trabajar en la industria abandonando la agricultura. El cambio, impulsado por el motor de vapor y la mecanización subsiguiente se produjo a lo largo de un siglo aproximadamente, desde mediados del siglo XVIII. Sus consecuencias se han sentido mucho más allá, hasta la aparición de las distintas formas de Estado de bienestar ya en el siglo XX¹⁷. La segunda Gran Transformación es la que

¹⁶ Garicano (2019, p. 118) también atribuye la elección de Trump y el Brexit en parte a los cambios producidos a partir de la IA. Pare este autor, la automatización y el avance de la IA llevan al regreso del nacionalismo, el tribalismo y las políticas extremistas de izquierdas y de derechas, azuzadas por líderes populistas (p. 129).

¹⁷ La tesis de Baldwin es que toda gran transformación en la economía da lugar a una reacción en la sociedad y la correspondiente reacción en contra de la que surge la resolución de los conflictos. Por ejemplo, en el caso de la Revolución Industrial, pasaron cien años desde 1769, año en que James Watt introduce el motor de vapor, hasta el primer barco de vapor comercial. El mismo tiempo se necesitó para establecer los cerramientos en el ámbito rural. La nueva tecnología, el vapor, permitió los subsiguientes cambios en el sector textil, minero y siderúrgico. La *Segunda Revolución Industrial* (desde principios del siglo XIX hasta 1970) se caracteriza por el desarrollo del ferrocarril, el alumbrado, el motor de explosión, etc. Los cambios trajeron progreso, pero también desigualdad y el sufrimiento de muchos, como atestiguan por ejemplo las Leyes de Pobres en Inglaterra (1834) o las novelas de Dickens. La Revolución Francesa, el fascismo y el

introduce la sociedad de servicios y no se intensifica sino a partir de 1973, con el desarrollo de los ordenadores y los avances en el ámbito de las telecomunicaciones. La automatización y la robótica que trajeron consigo son responsables de la creación de muchos puestos de trabajo en el ámbito de los servicios y de la destrucción de otros en la industria. Las consecuencias se han dejado notar ya en el éxito de movimientos populistas cuyas consecuencias todavía desconocemos pero que van a solaparse con el ritmo volcánico del progreso tecnológico, una revolución que no debe ignorarse (Baldwin, 2019, p. 83)¹⁸.

La globalización del trabajo y la IA son una nueva frontera que para Baldwin cobra impulso a partir del año 2017 al ritmo explosivo que hemos señalado. Este autor destaca que junto a la IA existe otro desarrollo que es clave para determinar la evolución de los mercados de trabajo, la *inteligencia remota (IR)*, terminología con la que hace referencia a la posibilidad de realizar muchos trabajos desde cualquier parte del mundo con acceso a Internet. Ese acceso y un mediano dominio del idioma inglés permite (junto a las herramientas de traducción automática) que personas con cualificaciones en un amplio rango de materias puedan competir con los trabajadores de países más desarrollados, cobrando salarios mucho más bajos. Formarían una legión de *telemigrantes* situados en China o en cualquier otra parte del mundo, completando una globalización que hasta el momento había afectado al mercado de bienes y servicios, pero no tanto al trabajo¹⁹. Para Baldwin puede hablarse de *globoización*, terminología con la que une la nueva globalización de los mercados de trabajo y la introducción por la IA de un buen número de robots que automatizan tareas y destruyen empleo.

La clave de la situación está en el avance exponencial en tres campos que contribuyen a la digitalización de la economía, y como cualquier crecimiento exponencial *sorprende* al cabo de unas cuantas iteraciones (lo que señala Baldwin como el *'holy cow moment'*). Por un lado, desde 1956 se duplica aproximadamente cada 18 meses la velocidad de proceso de los ordenadores (Ley de Moore), por otro lado, desde 1989 crece la velocidad

comunismo serían para Baldwin la *reacción* inevitable que conduce como *resolución* a la *New Deal* y el Estado de bienestar.

¹⁸ Para este autor la velocidad del cambio es claramente difícil de comprender y muchos no son conscientes del mismo o prefieren ignorarlo (Baldwin, 2019, p. 87)

¹⁹ El salario de un contable, un programador informático o un ingeniero en China es aproximadamente el 4% de lo que sería en Estados Unidos o en Europa (op. cit., p. 117).

de transmisión de datos a un ritmo igualmente extraordinario, (Ley de Gilder)²⁰ y también crece el número de individuos conectados a la Internet, lo que hace a la red más valiosa y expande las posibilidades de desarrollo de grandes empresas en un modelo en el que el ganador desplaza a todos sus competidores (*winer takes-all*)²¹. El resultado conjunto no es otro que el abaratamiento de la recogida y gestión de datos (el desarrollo de los *big data*) que ahora se pueden gestionar con ordenadores cuya capacidad de proceso y precio era difícil de imaginar hasta hace poco.

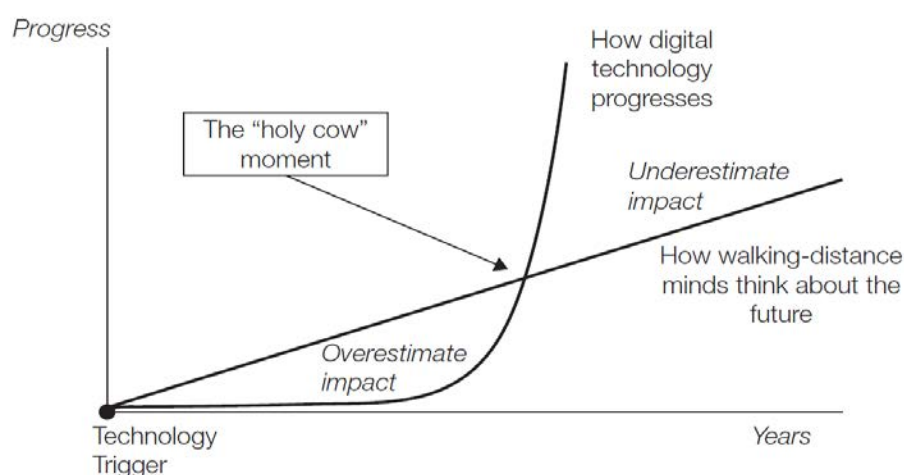


Gráfico 2. The "Holy Cow" diagram. El crecimiento exponencial de la capacidad de procesamiento de datos y su transmisión, además del crecimiento de las redes hace que se subestime la importancia de los cambios más recientes y del inmediato futuro (Baldwin, p. 91)

La interpretación del informe de Servoz (2019) no es muy diferente. Se refiere en este caso a tres cambios sísmicos: la globalización, la digitalización y la atomización económica. En conjunto, se trata del mismo fenómeno: muchos servicios se refieren a flujos digitales intangibles, lo que permite que las correspondientes actividades se localicen en cualquier parte y que la producción de bienes pueda descentralizarse en redes

²⁰ George Gilder anticipó en 1989 un crecimiento de los datos transmitidos tres veces por encima del crecimiento de la velocidad de los ordenadores. En 2017, cada minuto se envían medio millón de Tweets, se ven cuatro millones de videos en YouTube, 47 millones de mensajes de Instagram, 4 millones de *likes* en Facebook y 15 millones de mensajes de texto.

²¹ Hace referencia a que una red es más valiosa cuanto más grande es. Por ejemplo, la aplicación WhatsApp puede que no sea estrictamente mejor que otras, pero su generalización produce un efecto de bola de nieve que invita a sumarse al uso de tal aplicación a quien nunca había pensado hacerlo. En general, el resultado es que el mercado está dominado por grandes empresas. Varios ejemplos pueden verse en Garicano (2019, p. 85 y ss.).

horizontales de empresas que colaboran²² de muchas actividades (agencias de viajes, servicios financieros, etc. y la atomización económica (p. 16-19).

4.2. Cuántos y qué puestos de trabajo están en riesgo de desaparecer

La IA se considera una tecnología de aplicación general y muy poderosa (Agrawal, Gans, & Goldfarb, 2019b). Que sea de aplicación general quiere decir que es susceptible de ser utilizada en un amplio número de sectores y que puede tener consecuencias de todo tipo a largo plazo, como el descubrimiento de la bombilla o el motor de vapor o el microprocesador. Este hecho, y su previsible implantación acelerada hacen que sea muy difícil prever qué puestos de trabajo van a desaparecer o se van a modificar y cuántos y qué nuevos empleos pueden aparecer.

Distintos trabajos se han ocupado de esta materia, e incluso no falta quien ha puesto la información en Internet, accesible para que todos puedan tener noticia de si sus puestos de trabajo van a ser ocupados por robots (<https://willrobotstakemyjob.com/>). Servoz (2019, p. 35 y ss) recoge las conclusiones de los trabajos más recientes y conocidos sobre la materia que podemos sintetizar en los siguientes puntos:

1. En cuanto al volumen de empleos que pueden desaparecer, las previsiones varían enormemente, desde el 14% hasta el 47%²³. Sin entrar en detalles, cabe destacar que los informes a los que se refieren las cifras se refieren a distintas áreas geográficas (Estados Unidos, la OCDE, etc.) y tienen diferente horizonte temporal.
2. Las pérdidas de empleos no se distribuirán geográficamente de manera uniforme. Puede afectar más a países con menor desarrollo tecnológico o con mercados de trabajo menos flexibles, lo que puede poner en riesgo la cohesión social e incrementar las desigualdades.
3. Los trabajos más amenazados requieren baja o media cualificación, que están vinculados a actividades repetitivas más fácilmente automatizables (servicios de reparto, atención telefónica a clientes, taxistas y otros conductores, contables, etc.).

²² Como ejemplo, puede citarse la fabricación de vehículos, que se hace ahora ensamblando piezas que pueden proceder de muy distintas empresas auxiliares no dependientes del fabricante principal.

²³ La cifra del 47% corresponde a un muy citado análisis de Carl Frey y Michael Osborne publicado en 2013. Otro análisis de la empresa Forrester para Estados Unidos estima que en 2015 habrá una disminución neta del 7% de puestos de trabajo en Estados Unidos (citado por Baldwin, 2019, p. 161). Otros análisis muy citados son los del *McKinsey Global Institute*, el último de 2017 (<https://www.mckinsey.com/>).

Los trabajos menos amenazados son los que requieran creatividad, como el cuidado de personas mayores o niños²⁴.

4. Algunos trabajos se transformarán más que desaparecer. En algunos casos la IA permitirá redefinir las *tareas* dentro de un puesto de trabajo, y aunque algunas sean asumidas por los ordenadores o los robots, estos pueden ser auxiliares de tareas efectuadas por los trabajadores (por ejemplo, un robot puede auxiliar a un trabajador en el control de calidad de un montaje, pero el trabajador no desaparece)²⁵.
5. Las mujeres y los jóvenes pueden ser más afectados por la automatización, ya que son los que ocupan trabajos más probablemente llamados a desaparecer (por ejemplo, cajeros) y además están menos representados en sectores en los que el empleo es más probable que crezca, como los de TIC o STEM²⁶.
6. Si la productividad general no crece, o no lo hace de manera general, es probable que los salarios se mantengan bajos para los trabajadores con baja cualificación, en tanto puede crecer para los trabajadores cualificados, incrementando significativamente la desigualdad.

La conclusión de Servoz, como la de otros, es relativamente pobre. El efecto neto de la automatización depende de muchos factores, como la velocidad de adopción de la IA, la capacidad de recualificar a los trabajadores desplazados o el número de empleos que se creen en actividades ahora conocidas o en otras nuevas que no podemos ni siquiera imaginar. En todo caso, los riesgos son grandes y la necesidad de que el sector público intente prever y suavizar la transición también parece ineludible.

4.3. Salud y seguridad del trabajador en el puesto de trabajo

Las perspectivas de crecimiento económico y la necesidad de competir en el mundo globalizado hacen que no pueda prescindirse de incorporar los avances que trae consigo el uso de la IA. Por eso como hemos visto, los gobiernos se ocupan de potenciar y coordinar el uso efectivo de los más modernos algoritmos. Sin embargo, también es necesario estudiar cómo afectan al entorno de trabajo los inevitables cambios, a fin de

²⁴ Nada es seguro. La IA se utiliza ya para componer música, cuadros e incluso literatura, imitando el estilo del autor de referencia (o los diálogos de *Friends*, como citamos con anterioridad). Sobre la IA y el arte, puede verse Du Sautoy (2019).

²⁵ Por lo que se refiere al trabajo de los médicos radiólogos Agrawal *et al* (2019b, p. 43 y ss) dividen su trabajo en 29 tareas y encuentran que es difícil prever si los radiólogos serán más o menos demandados en el futuro o cómo evolucionará su remuneración.

²⁶ Science, technology, engineering y mathematics.

hacerlos al ritmo correcto o siguiendo alternativas que no pongan en riesgos derechos de los trabajadores o su misma salud y seguridad. Moore (2019) se ocupa de estas cuestiones con algún detalle para estudiar las ventajas y los inconvenientes de la IA. La introducción de herramientas de IA lo que hace es facilitar una intensificación del seguimiento y la monitorización del trabajador, lo que, dependiendo de la forma en se haga, puede aumentar el estrés en el puesto de trabajo, la discriminación, intensificar el ritmo de trabajo, etc., todo ello unido a los riesgos psicológicos correspondientes. Resumimos a continuación algunos de los problemas que diagnostica Moore, referidos a distintos ámbitos siguientes de las relaciones laborales:

4.3.1. Gestión de recursos humanos

Cada vez es más habitual recurrir en la gestión de recursos humanos a lo que se conoce como “herramientas analíticas de personas” (*people analytics, human resource analytics o talent analytics*) que utilizan *big data* para medir, seguir e interpretar la actividad del empleado, planificar su actividad o gestionar en general los recursos humanos de la empresa. Los algoritmos pueden funcionar aquí como mecanismos de predicción a los que ya nos hemos referido, con sus características de falta de transparencia en la formación de las predicciones (dada la *caja negra* de su interior). Las predicciones asisten la toma de decisiones de la empresa, agilizando la toma de decisiones. La fuente de datos de los algoritmos pueden ser cosas tales como el número de letras que escriben los empleados (*keyboard clicks*), las páginas web visitadas, el número o el contenido de llamadas telefónicas, la presencia física del trabajador, datos de geolocalización fuera de la empresa o dentro de la misma, el contenido de los correos electrónicos enviados o recibidos, su tono de voz o sus mismos movimientos corporales. Existen aplicaciones que, a partir de la captación en video de las entrevistas de trabajo, permiten obtener recomendaciones a partir del lenguaje verbal y no verbal de los candidatos²⁷.

La IA se utiliza para decidir a quién contratar, despedir o promocionar, para detectar que trabajadores es más probable que dejen la empresa, para detectar actitudes de liderazgo,

²⁷ Estas aplicaciones se usan también en el sistema judicial americano, para obtener recomendaciones sobre los períodos de encarcelamiento a los que se va condenar a los procesados o si deben o no ser favorecidos por un régimen de libertad condicional. Como sucede en el ámbito laboral, los algoritmos, que pueden contener sesgos que discriminen a algún grupo, se introducen precisamente para evitar los sesgos de los individuos llamados a tomar las decisiones. Por ejemplo, en las aplicaciones para decidir sobre la libertad condicional de un individuo, puede no tenerse en cuenta en los algoritmos correspondientes la raza del sujeto o su lugar de residencia (que pudiera estar correlacionado).

el nivel de la moral de los empleados o su estado general de salud. No es ciencia ficción, estas herramientas se utilizan ya en el 40% de los departamentos de recursos humanos de las grandes empresas internacionales, y la tendencia es creciente.

Los algoritmos se utilizan con el propósito de tomar decisiones más acertadas en la gestión de la empresa, y pueden ser tan buenos como los datos con los que se hayan construido. Es decir, que pueden tener sesgos, si estaban sesgados los datos utilizados, pero su propósito general es precisamente eliminar los sesgos que pudieran introducir los decisores *humanos* tradicionales. Se busca tener *datos objetivos* para la toma de decisiones.

En algún caso, estos nuevos entornos pueden, en potencia, dar lugar a una mejora en el cumplimiento de los trabajadores, en la medida que comprendan el sistema y pueda servirles para identificar áreas de mejora, sin embargo, si falta transparencia, es posible que el trabajador sufra un aumento de los riesgos unidos al estrés y la ansiedad, haga que los trabajadores se sientan injustamente espiados o aumenten también los riesgos físicos en el puesto de trabajo.

4.3.2. 'Cobots' en los puestos de trabajo

Los robots sustituyen a los trabajadores en algunos puestos de trabajo (por ejemplo, evitando que realicen actividades peligrosas), pero también es frecuente que los robots sean solo colaboradores del trabajador, que se ocupa de algunas tareas auxiliado por un robot colaborador (*cobot*). Por ejemplo, la máquina puede detectar errores que el trabajador corrige, evitando algunas tareas penosas y repetitivas anteriormente realizadas por el trabajador.

Sin embargo, dependiendo de cómo se introduzcan, no faltan los riesgos. Por ejemplo, si el trabajador es forzado a trabajar al ritmo del cobot (y no al contrario), o si asume riesgos físicos en casos en que, por fallos o por errores y la programación de la máquina, ésta pueda golpear al trabajador²⁸.

²⁸ Se ha publicado cómo Amazon utiliza algoritmos para despedir automáticamente a trabajadores basándose en su productividad medida a través de algoritmos. El ritmo de trabajo impuesto se considera abusivo y las condiciones de trabajo extenuantes. Los trabajadores son supervisados por robots, que envían avisos a los trabajadores ante lo que aprecien como una reducción del ritmo de trabajo., lo que a llevado a huelgas y protestas por parte de los empleados de esta empresa.
<https://www.theverge.com/2019/4/25/18516004/amazon-warehouse-fulfillment-centers-productivity-firing-terminations>.

Los *chatbots* utilizados en los centros de atención telefónica pueden ser un ejemplo. La mejora en las capacidades de reconocimiento y análisis de voz hacen que los chatbots puedan ocuparse ya de atender las preguntas más frecuentes, siendo en otros casos necesaria la intervención de humanos. Estos, sin embargo, pueden verse sometidos a un elevadísimo nivel de monitorización y evaluación de su actividad, estudiándose no sólo la voz del que llama, sino también la voz y las expresiones faciales del empleado, para revelar su nivel de cansancio o hacerle inmediatas recomendaciones sobre cómo desenvolver su actividad. El robot se puede convertir en el auténtico jefe del trabajador, recomendándole modificar su tono de voz, hablar más despacio, etc. (Roose, 2019). Estas herramientas, a falta de la formación adecuada, pueden suponer riesgos psicológicos vinculados al estrés o al riesgo de ser despedido o discriminado en la parte del salario que pueda ser dependiente de la medición del desempeño.

4.3.3. 'Wearables' en el puesto de trabajo y fabricación 'a medida'

La monitorización de los trabajadores es susceptible de incrementarse en gran medida a través de aparatos industriales o de 'relojes' capaces de obtener distintos parámetros físicos del trabajador, como su ritmo cardiaco, las horas de sueño, etc. Los aparatos industriales también pueden tener incorporados GPS, identificadores de radio-frecuencia, etc. Algunos de esos pueden jugar un papel similar a los asistentes de centros de atención telefónica a los que nos hemos referido, dando instrucciones al trabajador para asistirle en su cometido. Por ejemplo, en el caso de la fabricación de productos personalizados en una cadena de montaje, el trabajador puede tener que realizar tareas repetitivas junto a cambios menores para 'personalizar' la fabricación. Esas órdenes le llegarían a través de instrucciones en pantallas o a través de la voz, lo que puede tener la ventaja de evitar errores o proporcionar formación al trabajador a través de la introducción de cambios en su actividad.

Sin embargo, una vez más, estos nuevos sistemas pueden crear inconvenientes al trabajador, forzando un incremento del ritmo de trabajo, o incluso haciéndole perder habilidades, en la medida que la nueva forma de trabajar le exige únicamente seguir instrucciones precisas que le llegan a su lugar de trabajo sin margen para su autonomía o desarrollo personal.

Algunos problemas de esta índole podrían identificarse en las plataformas que permiten la prestación de servicios en los que se ha llamado la *gig economy*, que suele identificarse

con empresas como Uber o Amazon Mechanical Turk (AMT)²⁹. El trabajo puede realizarse *online* (desde casa o cualquier otro lugar) y *offline* (por ejemplo, conduciendo un vehículo, como en Uber o en Deliveroo). Las actividades del trabajador al relacionarse con sus clientes a través de las plataformas proporcionan datos que evalúan y sitúan al trabajador en un rango que se concreta en su reputación en la plataforma y es susceptible de tener consecuencias para no ser expulsado de la misma o asignársele más o menos tareas. Cada plataforma tiene su propia forma de funcionamiento, y todas ofrecen alguna ventaja a sus clientes (menor precio que otros servicios, más seguridad al identificar al trabajador, etc.), pero también se ha denunciado que la enorme monitorización de las actividades fuerza a los trabajadores a realizar muchas horas o seguir un ritmo de trabajo que pone en riesgo su seguridad. Por ejemplo, dada la forma de remuneración de la relación laboral, el trabajador solo recibe dinero por los minutos de trabajo (no por su disponibilidad, los tiempos de espera, la especial dificultad de las tareas -por ejemplo en días con mal tiempo). Las evaluaciones de los clientes se convierten en cruciales, y no consideran la salud física del trabajador o ninguna circunstancia personal que pueda afectarle. Como es conocido, la legislación laboral por ahora no se adapta bien a este tipo de tareas realizadas por trabajadores ‘autónomos’.

5. Políticas públicas ante los cambios (*im*)previsibles

Como venimos señalando, los cambios esperables son inciertos, pero es probable que sean muy importantes para algún grupo social, o incluso para toda la sociedad. Los riesgos invitan a estudiar iniciativas novedosas, como programas de redistribución de renta, más necesarios si nos encontrásemos ante un escenario con desempleo masivo. También, para facilitar la transición existe consenso en la necesidad de modificar los sistemas educativos de modo que los trabajadores tengan la oportunidad de *reinventarse* y afrontar los cambios. Finalmente, una tarea imprescindible y que sólo puede realizar el sector público es dirigir la intensidad y la dirección de los cambios, de manera que no se pongan en peligro valores democráticos o derechos esenciales de los trabajadores y de la población

²⁹ Sobre las malas condiciones de trabajo de esta empresa de Amazon en Estados Unidos, puede verse: <https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona>. La actividad de la empresa consiste en parte en hacer que los trabajadores ‘entrenen’ algoritmos de IA introduciendo datos. A esto hace referencia el nombre de la empresa, que recuerda la historia de un jugador de ajedrez con turbante que fingía en el siglo XVIII que una máquina era la que jugaba, y ganaba a sus oponentes jugando al ajedrez. La máquina, obviamente, era manipulada secretamente por su dueño pretendidamente turco (https://en.wikipedia.org/wiki/The_Turk).

en general. Nos estamos refiriendo a la necesidad de abordar nuevas regulaciones que garanticen el desarrollo ético de las nuevas aplicaciones, así como la privacidad y los demás derechos individuales reconocidos en el ordenamiento jurídico. Nos referiremos a continuación a cada uno de estos aspectos.

5.1. Educación para la transición y fomento de la inteligencia artificial

Una forma de hacer frente a la transición digital debe ser introducir cambios en el sistema educativo e intentar formar parte del grupo de países líderes en investigación y desarrollo de las aplicaciones de IA. Como afirma Servoz (2019, p. 57), los sistemas educativos jugarán un papel fundamental, especialmente cuando se trata de aprender materias nuevas por parte de muchísimos trabajadores que ya no son jóvenes y necesitan aprender urgentemente habilidades no específicas, sino transversales, válidas para ser adaptadas a los nuevos puestos de trabajo. La tarea no es fácil, menos cuando lo primero que se exige es afrontar el reto de decidir qué cualificaciones serán necesarias³⁰. Las propuestas principales de Servoz pueden resumirse en los siguientes puntos:

1. Mejorar la formación de estudiantes y profesores en materias relacionadas con la IA. (*STEM, science, technology, engineering and mathematics*). Debe hacerse desde edades tempranas y sin restar importancia a los estudios de arte y humanidades que fomentan habilidades creativas y analíticas que es probable que sean más importantes en el futuro³¹.
2. Flexibilizar los sistemas educativos, para adaptar los curriculums a las habilidades individuales de cada estudiante, incluso en el nivel universitario³². Deben adquirirse capacidades de comunicación y trabajo en grupo, desarrollando actitudes que no están vinculadas al puro *conocimiento* de materias concretas.
3. Vincular más el sistema educativo general con la formación profesional, haciendo desaparecer la mala imagen social que a veces tiene esta última y vinculándola con la

³⁰ Servoz (p. 59) da como referencia la información procedente de LinkedIn, red profesional con 560 millones de usuarios y la que pueden verse (LinkedIn's Economic Graph) tendencias del mercado de trabajo como las cualificaciones más demandadas o los salarios.

³¹ Los problemas en la educación de adultos tienen su origen en la formación secundaria. Según el informe PISA de 2015, un amplio porcentaje de estudiantes de 15 años mostraba graves deficiencias en ciencias (20,6%), lectura (19,7%) y matemáticas (22,2%), citado por Servoz (2019, p. 70).

³² Este modelo se sigue ya en el Swiss Federal Institute of Technology (Zurich). Otra iniciativa innovadora fue anunciada en octubre de 2018 en el MIT (Estados Unidos).

práctica en empresas³³. La formación profesional debe estar mejor financiada y accesible para los más necesitados.

4. La formación debe extenderse a toda la vida del trabajador, rompiendo la frontera entre el sistema educativo y el mercado laboral. En muchas áreas de conocimiento, lo que se aprende en el primer curso universitario está obsoleto en el cuarto año de la carrera³⁴. Las empresas deben adaptarse para cumplir parcialmente una función de instituciones educativas. Las carreras profesionales lineales son algo del pasado y se trata de evitar que los desempleados pasen a ser *inempleables*.
5. La formación y la protección social, tradicionalmente vinculadas a un contrato laboral *tradicional*, tienen que extenderse a cualquier nueva forma de empleo, como las de las plataformas digitales o la economía colaborativa (trabajadores autónomos). En esa línea, para financiar la formación se está a favor de introducir el modelo de *mochila austriaca*, según el cual una parte del salario iría destinado a financiar un fondo libre de impuestos que el trabajador podría gastar cuando quisiera en actividades de formación, que lo preparasen mejor para los cambios de actividad, en la misma o en una empresa diferente.

La Comisión Europea (2018b) es consciente de esta necesidad de adaptar la formación de los trabajadores, incluso con anterioridad a la elaboración de una estrategia europea sobre IA³⁵. Se ha establecido un Plan de Acción para la Educación Digital que prevé planes piloto encaminados a los grupos cuyos trabajos puedan desaparecer más fácilmente, de modo que se evite exacerbar la desigualdades entre grupos sociales, regiones o sectores económicos (European Commission, 2018b).

Por otra parte, y puesto que la implantación de la IA es imparable, es preferible liderar el proceso que esperar pasivamente las consecuencias. Con ese fin, se busca reforzar lo que se conoce como *el ecosistema europeo de IA*, que incluye tanto el aumento de la investigación en el sector como la rápida implantación de los avances, particularmente en las empresas pequeñas y medianas. En esta materia podemos resumir las propuestas principales, de nuevo a partir del informe de Servoz (2019, pp. 76–92):

³³ Así se hace en la enseñanza secundaria en Suiza.

³⁴ World Economic Forum, 2016, citado por Servoz (2019, p. 65).

³⁵ Puede verse el documento COM/2016/0381 final, “Working together to strengthen human capital, employability and competitiveness.

1. Se necesita aumentar el número de especialistas profesionales en IA, vinculando las investigaciones académicas con los inversores tan rápidamente como sea posible.
2. Aumentar el número de estudiantes en los campos que comprende la IA.
3. Facilitar la entrada de especialistas extranjeros mediante visados especiales, como los que existen en Canadá, que se tramitan en dos semanas o menos.
4. Reforzar la política industrial europea, apoyando a empresas líderes (*campeones europeos*) que puedan competir en campos como la ciberseguridad, los semiconductores, etc. Para ello reforzar el mercado único europeo es necesario. En esta línea se ha aprobado ya, por ejemplo, la libre circulación de datos no personales dentro de la UE³⁶

La Unión Europea, aunque no es líder en el ámbito de la investigación o la puesta en práctica de la IA, ha articulado una política común que toma conciencia de la importancia de la materia. No obstante, es probable que las iniciativas europeas sean insuficientes y falta el esfuerzo individual de los países en el desarrollo y la financiación de los programas necesarios.

5.2. La Renta Básica Universal y otras formas de abordar la desigualdad

Toda la literatura sobre IA coincide en que, incluso si el número de puestos de trabajo fuese estrictamente el mismo después de introducir la IA, por crearse tantos puestos como los que se destruyen, es posible que algunos grupos de trabajadores no pudiesen adaptarse con la rapidez exigida. Se crearía así un problema de redistribución de renta que habría que abordar. El problema será más grave cuantos más puestos de trabajo pudieran destruirse y cuanto más rápido fuese el fenómeno. Para algunos autores, el riesgo es suficientemente grave para estar preparados, incluso si pudiera no suceder lo peor. Este es el caso de Bruun y Duka (2018), para los que si no se hace nada las consecuencias sociales podrían ser catastróficas. Incluso si sólo hubiera un cinco por ciento de probabilidades de que se produjese un cataclismo, merecería la pena tomárselo seriamente³⁷. Para estos autores, eso significa introducir una Renta Básica Universal

³⁶ Reglamento UE 2018/1807, de 14 de noviembre, en vigor desde el 18 de diciembre de 2018, la CE ha publicado una guía para su aplicación: COM(2019) 250 final, de 29 de mayo de 2019.

³⁷ En la misma línea puede verse Harari (2018 cap. 2), para quien el peligro es *la irrelevancia* de mucha parte de la población de modo que “la potencial alteración de la vida social y política es tan alarmante que aunque fuese baja la probabilidad de un desempleo masivo, debe considerarse seriamente”.

Incondicionada. El desempleo masivo daría lugar a un incremento de la desigualdad que vendría a cuestionar en muchos países desarrollados la legitimidad del sistema democrático, que ahora se sostiene en la posibilidad de *ganarse la vida* en un puesto de trabajo del que obtienen los ingresos necesarios para financiar el consumo de bienes y servicios, ahorrar o financiar los placeres que cada cual desee. Ante una situación de desempleo masivo, el sector público no podría financiar los programas del Estado de bienestar tradicional y se haría insostenible.

La idea de una renta básica universal (RBU) tiene antecedentes en el siglo XVIII, pero con más seriedad se ha planteado en el siglo XX como un medio de combatir la pobreza unida a la situación de desempleo (Milton Friedman, por ejemplo, propuso un impuesto negativo sobre la renta). En su forma más estricta, una RBU es un ingreso que el sector público proporcionaría a todos los ciudadanos por el hecho de serlo, con independencia de su nivel previo de renta, su estado civil, o incluso su edad (también podría pagarse a los niños). El pago se haría de forma incondicionada, es decir que no está sujeto a seguir ninguna conducta o circunstancia predeterminada, como trabajar o estar buscando trabajo, ser rico, etc. Bruun y Duka discuten únicamente si debería pagarse a los no nacionales del país, o si debería seguirse pagando a los nacionales que fijasen su residencia fuera del mismo³⁸. La RBU se propone asegurar a todos unos ingresos mínimos, de modo que la seguridad de obtenerlos se convierte en un derecho humano fundamental que separa la atención de las necesidades de la necesidad de trabajar (Bruun & Duka, 2018, p. 8). En este caso, los autores aprecian dificultades políticas para introducir la RBU, pero llegan incluso a hacer cálculos aproximados de cuál podría ser su cuantía en Estados Unidos y cómo financiarla.

Sin entrar en los cálculos presupuestarios, uno de los autores que más defiende introducir una renta básica es el historiador Rutger Bregman (2016). No lo hace conectándolo estrictamente a los avances de la IA, sino a una idea de justicia. No se trata de hacer favores, sino de reconocer derechos, a partir del hecho de que *somos ricos*. Efectivamente lo somos si nos comparamos con sociedades pasadas, incluso hace sólo doscientos años se vivía mucho menos y mucho peor, el problema es que hoy vivimos en función de unos valores que deberían cambiar, moviéndonos hacia una sociedad donde los ingresos

³⁸ Para estos autores, los menores también deberían recibir la RBU, aunque introducen algunos matices para aminorar la cuantía de algunos adultos que habiendo trabajado tengan derecho a pensiones u otros ingresos públicos (Bruun & Duka, 2018, p. 8).

mínimos estuvieran garantizados y se trabajase solamente quince horas a la semana. La propuesta de Bregman es utópica, pero, como señala en el título de su libro, es también viable, aunque seguramente no a corto plazo.

Bregman recorre con cierto detalle los antecedentes históricos y las experiencias concretas en las que se ha introducido una RBU o programas muy parecidos. Hay distintas experiencias en Canadá y en distintos países africanos. A su juicio, los resultados son buenos, y descarta todos los inconvenientes con los que los detractores suelen oponerse a la implementación de una RBU, como que haría que la gente no trabajase o incurriese en conducta antisociales, o que sería imposible financiarla. Por el contrario, la gente viviría con menos estrés, habría menos accidentes, se reducirían las emisiones de CO₂ (al trabajar menos horas), se contribuiría a la emancipación de la mujer, a reducir el desempleo y, especialmente, a reducir la desigualdad (Bregman, 2016)³⁹.

En todo caso, la propuesta de RBU no es mayoritaria. Lo más frecuente es descartarla como inviable, bien por la dificultad de financiarla ya mencionada o por los incentivos perversos que se supone que introduciría. En esta línea podemos citar a Garicano (2019, p. 207), que considera este programa imposible de financiar, piensa que los menores se pasarían las horas muertas jugando a los videojuegos en vez de formándose, lo que les descolgaría del mercado de trabajo y además cree que trabajar es lo que da sentido a la vida. Curiosamente, Harari (2018), que también se ocupa del tema extensamente y considera la RBU algo ineludible en la edad de la IA y la robótica, pone como ejemplo a los rabinos judíos ortodoxos en Israel, que no trabajan en absoluto, son sostenidos por el Estado y viven felices.

En su visión distópica de la sociedad gestionada a todos los niveles por la IA, Lassalle (2019, p. 120 y ss) también se refiere a la RBU como un mecanismo para evitar una reacción ludita y poner a los ricos a salvo de la ira popular⁴⁰, un ingreso que nos

³⁹ Resumimos aquí la propuesta de Bregman de una manera que la simplifica en exceso. Su libro se refiere con detalle a una propuesta de la administración Nixon en Estados Unidos que estuvo a punto de aprobarse en los años 70 y a otras experiencias en Canadá, Estados Unidos, Uganda, Kenia, etc. Como puede suponerse, su propuesta implica un posicionamiento claro a favor de mayores niveles de igualdad y el abandono del PIB como medida del bienestar, sustituido por algún otro índice que mida mejor el progreso, como un índice de bienestar sostenible o incluso un índice de felicidad del planeta. Su propuesta de utopía incluye también la desaparición de cualquier frontera ya que hoy la desigualdad es en buena medida dependiente del país de nacimiento. Desde una posición ética, está a favor de la desaparición de las fronteras.

⁴⁰ Para Lassalle (2019, p. 125), es sintomático que los principales valedores de la RBU sean las grandes corporaciones tecnológicas cuyos portavoces son, entre otros, Mark Zuckerberg, Bill Gates, Jeff Bezos o

preservaría del miedo a ser prescindibles, del miedo a la llegada de los robots. La RBU no se fundaría en la equidad, sino “en pacificar la transformación digital y evitar que la paulatina erradicación del trabajo humano se convierta en un conflicto laboral y social... la clave es impedir que el legislador introduzca objeciones éticas o laborales a la IA y a la robótica... Se conseguiría la adhesión acrítica y pasiva del conjunto de la sociedad al proceso de automatización planteado por las plataformas” (p. 126-127). En su planteamiento pesimista del futuro de la democracia, Lassalle concluye que la RBU no es un programa deseable, ya que su fin no es tanto la equidad como instaurar un Leviatán tecnológico que neutralizaría la política y la desvestiría de ropajes democráticos, actualizando la reflexión de Juvenal sobre el *panem et circenses*, pan y circo (Lassalle, 2019, p. 130)⁴¹.

La RBU no es probable que se establezca a corto plazo, aunque de alguna manera sea necesario procurar una mejor distribución de la renta (ineludible para muchos desde hace tiempo, con independencia de la IA y la robótica). Una alternativa razonable creemos que puede ser la que propone Baldwin (2019, p. 271). Este autor es optimista en cuanto a los posibles resultados finales de la introducción masiva de IA, incluso pudiera ser un Nirvana en el que nadie tuviera que trabajar. Sin embargo, considera que la educación adicional no es probable que sirva de mucho, y tampoco es partidario de la RBU (por los motivos habituales). Su propuesta está dentro de una cierta ortodoxia económica en la que debe protegerse a los trabajadores, y no los puestos de trabajo. Cambios radicales y rápidos pueden dar lugar a reacciones radicales o a violencia que deben ser abordados por los gobiernos ayudando a los trabajadores desplazados, incluso introduciendo regulaciones que hagan más lento el proceso⁴² con el fin de reducir la fricción producida por el cambio inevitable y repartir más equitativamente los beneficios. Un ejemplo sería el modelo de *flexiseguridad* que se aplica en Dinamarca, que incluye flexibilidad para despedir y contratar trabajadores junto a generosos subsidios de desempleo para rentas moderadas de hasta 2000 euros al mes. El sistema se complementa con ayudas eficaces

Elon Musk. Todos participan del interés de lograr que la transformación digital no se interrumpa ni reduzca su ritmo de crecimiento exponencial.

⁴¹ Juvenal fue un poeta romano de finales del siglo I y comienzos del II, autor de dieciséis sátiras en las que critica la sociedad de su época. Es conocido por algunas de sus expresiones, como la citada de *panem et circenses* refiriéndose a las costumbres romanas del Imperio. También tienen origen en sus escritos expresiones como ‘*rara avis*’, ‘*mens sana in corpore sano*’ y ‘*sed quis ipsos custodes*’ (quién vigilará a los vigilantes). (Wikipedia, 2019).

⁴² Recuerda a lo que ha pasado en España en relación al dominio de Uber y Cabify. Las normas autonómicas o municipales han restringido, al menos temporalmente, la actividad de estas empresas.

en la búsqueda de nuevos empleos o la formación adicional del trabajador. La fórmula no es nueva, pero sería válida en un contexto en el que el libre mercado y a globalización no van a desaparecer. Desde una posición optimista, la fórmula de política pública que propone, proteger a los trabajadores y no los puestos de trabajo, confía en que pueda dar frutos aceptables para la sociedad, aunque advierte de la posibilidad del éxito de movimientos fascistas o totalitarios de los que la presidencia de Trump podría ser un augurio.

5.3. *La financiación del sector público: reforma fiscal e IA*

Michel Servoz (2019, pp. 103–107) dedica un interesante apartado a la necesidad de cambios en el sistema fiscal antes los cambios que nos encontramos en la economía digital. Algunas de las propuestas de cambios son las siguientes:

1. Reducir los impuestos sobre las rentas del trabajo a fin de promover con ello la inversión en educación, capital humano, que debe tener un tratamiento fiscal al menos idéntico al de las inversiones en capital físico o en investigación y desarrollo. La pérdida de recaudación en las rentas de trabajo se compensaría gravando más la riqueza, las ganancias de capital y las actividades que dañen el medioambiente.
2. Estudiar la posibilidad de que los robots paguen impuestos⁴³. Aunque no parece fácil, podría tener la ventaja de frenar la pérdida de puestos de trabajo, al hacerlos menos atractivos, favoreciendo así la transición hacia su uso. En todo caso, establecer un impuesto de este tipo también sería difícil sin coordinación internacional, pues las empresas se situarían donde no existiese el impuesto. Este tipo de impuesto es además susceptible de frenar los incentivos a la innovación, lo que también se considera indeseable.
3. Aumentar la cooperación internacional para evitar la elusión fiscal. Muchas empresas actúan en el mercado sin apenas base física. Nueve de las veinte empresas más grandes por capitalización son hoy digitales. Esto les permite con mayor facilidad pagar impuestos en las jurisdicciones donde estos son más bajos, lo que les da una ventaja competitiva. La Comisión Europea ha propuesto introducir lo que se ha llamado *tasa Google* sobre los ingresos de las empresas

⁴³ Esta propuesta se vincula a Bill Gates y otros líderes de empresas tecnológicas.

digitales (ya que, si bien es difícil controlar los beneficios de estas empresas, es más fácil controlar sus ingresos en cada país). El proyecto se ha abandonado a nivel europeo, pero subsiste en algunos países como Francia y el Reino Unido.

La elusión fiscal es un problema que preocupa en el ámbito internacional mas allá de los cambios promovidos por la IA. La OCDE ha promovido iniciativas para evitar lo que se conoce como prácticas BEPS (*base erosion and profit shifting*) que reducen la carga fiscal de las grandes empresas que pueden, con artificios contables, llevar a cabo lo que se conoce como *planificación fiscal agresiva*, utilizando las lagunas legales que dejan sistemas fiscales complejos para reducir el importe de los impuestos pagados⁴⁴.

5.4. *El capitalismo de la vigilancia y el Ciberleviatán*

Como hemos visto, Baldwin hace una llamada a la intervención pública con paliativos a las disrupciones sociales nacidas de la inequidad en la distribución de los beneficios y costes de la revolución digital. Para otros, lo que se anuncia o ya está aquí es un cambio completo del modelo capitalista en el que el capital físico deja de ser relevante y es sustituido por el control sobre el conocimiento o *los datos*. En esta visión, el dominio de fuerzas que ponen en peligro la libertad y la organización social tal como la conocemos está a punto de llegar, o tal vez ya ha llegado y es ahora irreversible.

Para Shoshana Zuboff (2019) vivimos ya en una sociedad organizada como *capitalismo de la vigilancia* (*'surveillance capitalism'*), que es profundamente antidemocrática. La concentración de poder y de conocimiento da lugar a una tiranía que se alimenta de los individuos, aunque estos le son indiferentes. Se ha producido un golpe de estado que, al amparo del caballo de Troya de la tecnología, ha eliminado la importancia del individuo que fundamenta la democracia. El nuevo tirano no necesita campos de concentración o ejercer la violencia, le basta con las herramientas de manipulación que le son accesibles mediante el manejo de toda la información que tiene del conjunto de los individuos, lo que le permite reemplazar el predominio del derecho o la política. Para Zuboff, estamos en una nueva era que llama 'civilización de la información' en la que ya no se intenta

⁴⁴ En la Unión Europea, la propuesta más importante busca establecer una base imponible común y consolidada en el impuesto sobre sociedades (BICIS). Se trataría de que las empresas multinacionales paguen impuestos en los países donde se generan sus beneficios, impidiendo la elusión de impuestos unida a la posibilidad de hacer aparecer contablemente los beneficios en las jurisdicciones de más baja fiscalidad. Se ha publicado recientemente, por ejemplo, los casi nulos impuestos pagados en España por empresas como Amazon o Netflix. Puede verse: <http://www.oecd.org/tax/beps/>.

dominar *la naturaleza*, como en la civilización industrial anterior, sino dominar *la naturaleza humana*. No se buscan máquinas que superen los límites del cuerpo humano, sino máquinas capaces de modificar el comportamiento de la población, los individuos y los grupos, haciendo así radicalmente imposible la democracia.

La doctrina de la *indiferencia radical* caracteriza al capitalismo de la vigilancia, lo que importa es crecer, aumentar el número de usuarios conectados a una red controlada por unos pocos interesados (como Facebook), sin importar si los contenidos que circulan en la red son desinformación, son falsos o promueven la violencia. Esto sólo es un problema en el caso de que, en algún momento, ponga en peligro a la propia empresa, por ejemplo, atrayendo el interés de los reguladores o suscitando la animadversión de los usuarios. Por eso, cualquier actividad de gestión de los contenidos que circulan en la red es más una política defensiva que un ejercicio de responsabilidad pública. La actividad de los ‘moderadores de contenidos’ se rige por principios que buscan eliminar de la red tan pocos materiales como sea posible, con el objetivo de minimizar la pérdida de usuarios, no de valorar la validez o la veracidad de los contenidos. La ‘indiferencia radical’ sólo aspira a que la red se use, no importa para qué. Se espera de todos “que cedamos nuestra autoridad, relajemos nuestras preocupaciones, bajemos el tono de voz, nos acomodemos al flujo general y nos sometamos a los visionarios de la tecnología de cuyo mejor juicio da testimonio su riqueza y su poder... Nos han educado en que la tecnología es ‘inevitable’, pero en realidad es un narcótico existencial que nos han recetado para inducirnos a la resignación” (Zuboff, 2019)⁴⁵. Pese a todo, esta autora termina rechazando la inevitabilidad de los acontecimientos, podemos rebelarnos contra los cambios que no hemos elegido, concienciarnos de que cada generación puede deshacer los errores de la anterior, y en esa línea, nos llama a indignarnos y actuar sin cobardía. El muro de Berlín cayó por muchas razones, pero sobre todo, porque los berlineses dijeron “¡basta ya!”

Con un sesgo también pesimista, se refiere a la IA el libro de José María Lassalle. Ya el título de su libro lo dice casi todo: *Ciberleviatán. El colapso de la democracia liberal frente a la revolución digital* (Lassalle, 2019). Ya nos hemos referido a que este autor ve la RBU como un programa que legitimaría los cambios hacia una sociedad no democrática. Como el Leviatán de Hobbes, el Ciberleviatán tendría como fin garantizar

⁴⁵ Señala Zulloff que se extiende un sentimiento de que la próxima generación no vivirá como la actual y se valora cada vez menos la democracia tanto en Estados Unidos como en Europa y no faltan los que proponen formas políticas autoritarias que emulen al régimen político chino actual.

la paz y la estabilidad, intercambiar protección por obediencia, pero esta vez con un poder irresistible que anula toda ciudadanía, que lo sabe todo de cada uno de sus miembros y es capaz de influir sobre cada uno de ellos con algoritmos personalizados que desactivarían sus malestares y sus miedos⁴⁶, con la IA como medida de todas las cosas (Lassalle, 2019, p. 133). Para este autor, sorprende también que el avance hacia la concentración del poder se produzca sin necesitar la violencia o la fuerza para imponerse, percibida por la ciudadanía como algo positivo pese a estar en juego la democracia y los valores que la sustentan. Ya ahora los datos muestran quienes somos, cuáles son nuestras preferencias y si somos o no conscientes de las mismas, nos revelan en qué gastar nuestro dinero, cómo diseñar nuestros modos de vida, qué amigos elegir, etc., “se vuelcan masivamente datos sobre nuestra capacidad de elegir y, al mismo tiempo, se nos libera de esa responsabilidad para ayudarnos a decidir cómo ajustarnos mejor a nuestras preferencias... la dependencia del entramado algorítmico es tan intensa que si desapareciera repentinamente de nuestras vidas, estaríamos desamparados y nos sentiríamos perdidos sin remisión” (Lassalle, 2019, p. 43 y 98). Para Lassalle, la solución sólo puede venir de la mano de una ‘sublevación liberal’ que todavía es posible si se establece una nueva generación de derechos fundamentales que protejan a la persona y den sentido a la relación con la IA y la robótica, recuperando el liberalismo su capacidad crítica. Algunas de las ideas que propone Lassalle son generales y abundan en el replanteamiento de valores abstractos que no es evidente cómo trasladar a la práctica, sin embargo también contiene su propuesta algunos elementos concretos que llaman al protagonismo de la ley, que hasta ahora se ha inhibido de legislar sobre las grandes corporaciones tecnológicas. Reclama Lassalle reiteradamente la intervención reguladora del Estado para restablecer y garantizar la propiedad privada de los datos, para fijar límites remuneratorios, para intervenir en el diseño de los algoritmos, para desmantelar los oligopolios controlados por las empresas tecnológicas y, en general para establecer una nueva generación de derechos que fortalezcan la nueva ciudadanía digital. Para este autor, Europa tiene un papel fundamental en esta materia, ofrece la posibilidad de ser una alternativa a la resistencia y cambio frente a la distopía tecnológica que Estados Unidos y China proyectan sobre el

⁴⁶ Para Harari (2018), el Leviatán tecnológico, el régimen no solo sabrá exactamente como sentimos, además podrá hacer que sintamos lo que quiera. Para Harari, además, lo que cuenta es lo que la gente ‘siente’ no lo que pensamos racionalmente, de modo que la democracia no puede sobrevivir en el futuro de la revolución biotecnológica e informática. Puede verse también O’Neil (2016).

futuro⁴⁷. Tal vez este liderazgo europeo se produzca efectivamente. Por ahora Europa carece de la legislación necesaria, la realidad ha evolucionado al ritmo explosivo de los avances en el procesamiento y la transmisión de datos, pero como veremos, se puede apreciar un cierto liderazgo europeo en la protección de datos personales con el Reglamento sobre esta materia.

Algunos autores no aprecian en los avances de los algoritmos creados por la IA un mundo distópico al estilo de Lassalle o Zuboff, pero si importantes riesgos, especialmente para los más desfavorecidos de la sociedad. El libro de Cathy O'Neil *Weapons of Math Destruction* (2016) es uno de los que pone de manifiesto cómo la utilización de algoritmos a gran escala, con sesgos ocultos y falta de transparencia es susceptible de dar lugar a graves injusticias, especialmente para los más desfavorecidos de la sociedad⁴⁸.

En cualquiera de los supuestos, sea por el peligro de una sociedad distópica o por los perjuicios más limitados que aprecia O'Neil, parece necesario considerar en futuros desarrollos las propuestas éticas para el desarrollo de la IA que se han elaborado recientemente en la Unión Europea, a las que nos referiremos a continuación.

5.5. Los valores éticos promovidos desde la Unión Europea

La Unión Europea tiene una preocupación mayor que Estados Unidos, China y otros países por la protección de datos y la protección de la privacidad. Así se pone de manifiesto con la regulación del llamado *derecho al olvido* y especialmente con la aprobación del Reglamento General de Protección de Datos de la Unión Europea (RGPD) y, por lo que se refiere a España, la aprobación de la Ley Orgánica 3/2018, de Protección de Datos Personales y Garantía de los Derechos Digitales, de 5 de diciembre (LOPD-

⁴⁷ Aunque las empresas tecnológicas que dominan el mercado europeo tienen su origen en Estados Unidos, China es una potencia mundial en IA y lidera algunos aspectos negativos de la misma, tanto por el régimen político del país como por la carencia de un sistema de protección de datos, lo que facilita que los mismos estén disponibles para la elaboración de algoritmos de todo tipo. En cuanto al control ciudadano, es sabido que China está introduciendo sistemáticamente el control biométrico y un perfilado de sus ciudadanos, a los que asigna una puntuación que le puede dar más o menos derechos. El control es más intenso sobre las minorías: <https://www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html>, y <https://www.nytimes.com/2019/07/02/technology/china-xinjiang-app.html>.

⁴⁸ Las “destructivas armas matemáticas” de O'Neil (2016), pueden tener sesgos que se refuerzan en círculos viciosos. Por ejemplo, si un algoritmo condena a la cárcel a alguien que es negro y pobre, la condena a la cárcel puede condicionar sus futuras dificultades para encontrar empleo y, en consecuencia, tal vez aumenta las posibilidades de reincidencia y cárcel. Igualmente, algoritmos para hacer llegar publicidad de créditos con altos intereses se centran en personas con dificultades económicas, baja autoestima o debilidad mental; el aceptar créditos que en realidad no pueden pagar puede empeorar su ya difícil situación de partida.

GDD). Estas normas están lejos de regular todos los aspectos de la IA, pero son un avance para la protección de los trabajadores y de los ciudadanos en general.

No obstante, los aspectos legales son solamente uno de los tres componentes que un reciente documento sobre *directrices éticas para una IA fiable* elaborado a impulso de la Comisión Europea considera relevantes para desarrollar una IA *fiable* o de confianza, *'trustworthy'* (European Commission. High-Level Expert Group on Artificial Intelligence. AI-HLEG, 2019c). A lo largo de todo su ciclo vital las aplicaciones de IA deben *cumplir las leyes* vigentes, pero también deben ser *éticas*, atenerse a principios y valores éticos, y ser de calidad, *robustas* o resistentes desde el punto de vista técnico o social, porque no bastan las buenas intenciones para no causar daños. Resumimos a continuación algunas de las ideas principales de estos principios éticos para el desarrollo de la IA, que la UE plantea como guía, es decir, de seguimiento voluntario:

1. Los sistemas de IA deben centrarse en el ser humano. Se trata de servir al bien común promoviendo el bienestar y la libertad. Se trata de aprovechar las oportunidades que la IA proporciona, pero evitando y minimizando los riesgos que comporta. La IA no es un fin en sí misma, sino una herramienta que debe estar al servicio de valores como los promovidos por los Objetivos de Desarrollo Sostenible (ODS) de la Organización de las Naciones Unidas, entre otros, la igualdad de género, la racionalización de uso de los recursos naturales, la cohesión social, etc.⁴⁹
2. Con el propósito anterior, la confianza en la IA debe ser la base de su desarrollo, el requisito esencial para que pueda servir de base a los cambios que se anuncian. Si no se acredita que la IA es fiable, sus inconvenientes pueden impedir que se alcancen los grandes beneficios que puede reportar. Se necesita una confianza del tipo que ha permitido el desarrollo de la aviación, la energía nuclear o la cadena de distribución de alimentos, una confianza que solo se puede lograr con una aproximación general u holística que exige la colaboración de todas las partes interesadas, desde los desarrolladores de algoritmos de IA hasta los usuarios finales.
3. La confianza en la IA debe basarse en los tres componentes que hemos señalado con anterioridad: cumplir las leyes, atenerse a principios y valores éticos, y tener calidad o robustez desde es punto de vista técnico y social para no producir daños incluso con buenas intenciones. Estos componentes pueden entrar en conflicto y deben ser

⁴⁹ Puede verse: <https://www.un.org/sustainabledevelopment/es/objetivos-de-desarrollo-sostenible/>.

armonizados (por ejemplo, principios éticos y legales), por lo que la reflexión sobre cada uno de los aspectos se hace todavía más importante, para identificar las oportunidades y los desafíos en cada caso.

4. La guía ética elaborada por la UE da por sentado el cumplimiento del derecho vigente, sin entrar explícitamente en este aspecto. Da por supuesto que debe la IA atenerse a los derechos fundamentales consagrados en los Tratados de la UE y también la Carta Europea de Derechos Humanos (CEDH) y otros textos concordantes que articulan el Estado de Derecho y la democracia. En definitiva esos textos legales consagran principios éticos de carácter general, como la dignidad del ser humano, la libertad individual, la igualdad, la no discriminación y la solidaridad, el derecho a la justicia, etc. Son todos ellos valores que son necesarios para que la IA esté centrada en el ser humano, lo que contribuirá crucialmente a que sus desarrollos inspiren confianza.
5. A partir de los derechos fundamentales anteriores tal como se expresan en la CEDH, los desarrollos de la IA deben seguir cuatro principios éticos cuyo contenido explicaremos brevemente: respeto a la autonomía personal, prevención de daños, justicia y explicabilidad.
 - i) *Principio de respeto a la autonomía personal.* Implica que los sistemas de IA no deben subordinar injustificadamente al individuo, coaccionarlo, engañarlo, condicionarlo o manipularlo. Por el contrario, deben estar diseñados para aumentar, complementar o dar mayores facultades a los individuos, promoviendo sus capacidades. Esto significa que debe existir siempre supervisión humana de los resultados de la IA. En el ámbito laboral, la IA debe apoyar a los humanos, procurando la creación de trabajo digno.
 - ii) *Principio de prevención de daños.* Es un principio ético que implica que los sistemas de IA no deben ni producir ni aumentar daños, ni individuales ni colectivos, incluidos los daños intangibles de carácter social, cultural o político. Debe procurarse la seguridad física y mental, lo que significa que deben cuidarse especialmente los efectos de la IA sobre personas vulnerables o en entornos con asimetrías de poder o de información, como puede ser el caso de las relaciones entre los empleadores y los empleados, entre las empresas y los consumidores o entre los gobiernos y los ciudadanos.
 - iii) *Principio de justicia ('fairness').* Implica una distribución justa e igualitaria de los beneficios y de los costes de la aplicación de IA, de modo que no haya sesgos que discriminen o estigmaticen a individuos o grupos concretos. La IA

debe aumentar la igualdad de oportunidades para el acceso a la educación, a los bienes y servicios y a la tecnología. Además, la justicia implica seguir un principio de *proporcionalidad* entre medios y fines para valorar el equilibrio entre los intereses afectados y los objetivos perseguidos. Implica también que las partes afectadas deben tener la oportunidad de oponerse y pedir la revocación de las decisiones tomadas por sistemas de IA o por los humanos que las supervisen.

- iv) *Principio de explicabilidad*. Es importante que la IA sea explicable para inspirar confianza. Es necesario para poder exigir responsabilidades a los decisores, que deben ser identificables. Los procesos implicados por la IA deben ser transparentes, sus capacidades y sus fines deben ser públicos y sus decisiones, hasta donde sea posible, deben explicarse a los afectados por ellas directa o indirectamente. Los casos de algoritmo de *caja negra* requieren especial atención en cuanto a su trazabilidad, la posibilidad de auditoría y transparencia de sus capacidades. El grado de explicabilidad necesario dependerá mucho del contexto y de la gravedad de las consecuencias de los errores o de la falta de precisión de los resultados. No es lo mismo equivocarse en la recomendación de compra de un producto que en la decisión de dejar libre a un presunto delincuente.

En la práctica, un problema será siempre las tensiones entre los principios a los que nos hemos referido, para las que no hay soluciones únicas. Por ejemplo, si se utiliza la IA en el ámbito de la prevención de delitos, la reducción de la criminalidad puede afectar al derecho a la libertad individual y la privacidad, e incluso ser los perjuicios causados menores que los beneficios.

En la práctica, y esta es la aportación principal del documento al que venimos refiriéndonos, los dilemas deben abordarse a partir de reflexiones basadas en la evidencia, y no la simple intuición o la decisión aleatoria de los responsables. Puede haber casos en los que los dilemas éticos no tengan solución, como en situaciones en las que la dignidad humana esté en juego. El grupo de expertos de la UE proporciona una guía detallada de los aspectos del desarrollo de la IA que deben evaluarse para que el sistema sea de confianza, tanto en la fase inicial de diseño, como en las posteriores de despliegue o puesta a disposición y efectivo uso por los destinatarios finales. La guía concreta los numerosos aspectos a evaluar para entender, por ejemplo, que un sistema de IA es

transparente, respeta la privacidad, es técnicamente robusto, etc. Se trata, en definitiva, de dar contenido a principios cuyo cumplimiento es importante y que no debe confiarse a valoraciones subjetivas de nadie, incluidos los interesados en el diseño y despliegue de las herramientas concretas de IA⁵⁰.

La importancia concedida al desarrollo de aplicaciones de IA fiables o de confianza ha dado lugar también a 33 recomendaciones del grupo independiente de expertos dentro de la Comisión Europea dirigidas al papel que deben tener las instituciones de la UE y los Estados Miembros (European Commission. High-Level Expert Group on Artificial Intelligence. AI-HLEG, 2019d). Sin entrar aquí en ellas, podemos destacar que muchas exigen iniciativas regulatorias por parte de la UE y de los Estados Miembros, y específicamente se recomienda evaluar y revisar si fuera necesario las leyes de la UE (recomendación 28). En conjunto, el conjunto de recomendaciones que recoge el documento se extiende a un amplísimo ámbito de sectores que inevitablemente afectan a muchos aspectos de las relaciones laborales y que podrían entenderse como susceptibles de constituir una verdadera modernización del sector público, afectado transversalmente por la innovaciones de la IA.

5.6. Algunos desarrollos legislativos y jurisprudenciales

Un largo y reciente artículo del profesor Jesús Baz (2019) hace un detallado recorrido sobre la garantía de los derechos digitales laborales en España, obligadamente en el marco del RGPD que es obligatorio en el conjunto de la Unión Europea desde mayo de 2018. En España la cuestión se regula en la Ley Orgánica 3/2018 citada anteriormente y además por la jurisprudencia del Tribunal Supremo, la del Tribunal de Justicia de la UE (TJUE) y la del Tribunal de Europeo de Derechos Humanos (TEDH). No se regula directamente la IA o los algoritmos, pero sí materias relacionadas que se refieren al derecho a la privacidad del trabajador dentro y fuera del puesto de trabajo. Este derecho nace del

⁵⁰ La guía del grupo de expertos de la UE hace un estudio considerablemente detallado de los requisitos necesario para desarrollar aplicaciones de IA en las que se pueda confiar, pone de manifiesto los requisitos para cumplir los principios enumerados en el texto. Por ejemplo, la transparencia, unida al principio de *explicabilidad*, exige la *trazabilidad* de los datos utilizados, de la forma en que se etiquetaron, del algoritmo utilizado para elaborar los datos, de los errores que se hubieran detectado, etc. Se trata de aumentar la capacidad de auditoría. Además, la *comunicación* del sistema con los humanos no debe dar lugar a confusión o hacer creer que el usuario se comunica con un humano cuando no es así, etc. Finalmente, el documento presenta una versión piloto de las cuestiones que hay que analizar y responder para valorar en la práctica el nivel de confianza o fiabilidad de las aplicaciones de IA, haciendo operativos los principios enunciados (European Commission. High-Level Expert Group on Artificial Intelligence. AI-HLEG. 2019c.pp. 26-31).

derecho a la protección de datos y jurídicamente se fundamenta en el artículo 18.4 de la Constitución Española (CE), la ley limitará el uso de la informática para garantizar el honor y la intimidad personal y familiar de los ciudadanos y el pleno ejercicio de sus derechos, junto a las normas antes citadas, el art. 8 CEDH y la importante STEDH de 5 de septiembre de 2017 (Barbulescu II), que reformula los límites de los poderes empresariales utilizando como criterio fundamental de referencia los principios y la normativa de protección de datos personales. Todas las referencias anteriores constituyen para Baz la normativa de protección de datos (NPD).

Los cuatro derechos digitales laborales que la ley protege y que Baz (2019, p. 51) estudia con algún detalle son los siguientes:

- a) El derecho a la intimidad en el uso de dispositivos digitales en el ámbito laboral (art. 87 LOPDP-GDD).
- b) El derecho a la intimidad frente al uso de dispositivos de videovigilancia y de grabación de sonidos en el lugar de trabajo (art. 89 LOPDP-GDD).
- c) El derecho a la intimidad ante la utilización de sistemas de geolocalización en el ámbito laboral (art. 90 LOPDP-GDD).
- d) El derecho a la desconexión digital en el ámbito laboral (art. 88 LOPDP-GDD), que busca garantizar la efectividad de los tiempos de descanso, lo que se relaciona más con la protección de la intimidad personal y familiar que directamente con la protección de datos personales.

La configuración de estos derechos en la práctica tendrá que hacerse compatible con el art. 20.3 del Estatuto de los Trabajadores, que permite al empresario establecer medidas de vigilancia y control de la actividad laboral⁵¹. Esta compatibilidad puede alcanzarse a través de la negociación colectiva, sin que ésta pueda ser el cauce para renunciar colectivamente a derechos fundamentales que son individuales, sino para reforzarlos mediante ‘garantías adicionales’. Baz glosa la normativa vigente y la jurisprudencia con

⁵¹ Mercader Urquina (2017, pp. 121–181) estudia estas cuestiones refiriéndose a “los límites de la empresa panóptica” (en arquitectura, el término se aplica al edificio que está construido de manera que todo su interior se ve desde un solo punto, dando la capacidad de vigilancia permanente). Su análisis cubre con detalle la jurisprudencia sobre cada uno de los temas estudiados, aunque dada la fecha de publicación, no se refiere a la Ley Orgánica 3/2018, contiene referencias a un abundante número de obras doctrinales sobre la materia.

un acercamiento a la misma que parece inclinarse por entender que deben reforzarse los derechos del trabajador cuando la interpretación permita la duda⁵².

El trabajo del profesor Baz complementa bien nuestra aproximación a la IA en este trabajo, aunque nuestro estudio sigue una aproximación más general, centrada en comprender la magnitud del fenómeno y sus consecuencias en el mercado de trabajo y en el conjunto de la sociedad. El trabajo de Baz, junto a los de Mercader Urguina (2017, 2018) creemos que sientan las bases para la definición jurídica de algunos de los problemas concretos a los que se enfrentan los trabajadores y los empresarios. Los primeros reclamarán legítimamente el cumplimiento de sus derechos laborales, y los segundos deben atender la pretensión de los trabajadores sin perder por ello la capacidad de incorporar a su gestión empresarial las modernas técnicas de IA, que entre otras cosas, pueden ser vitales para mantener la competitividad en el sector económico correspondiente a su actividad.

Las cuestiones prácticas tienen que hacerse compatibles con principios generales, y eso no es inmediato en el mundo digitalizado en el que nos encontramos. Pongamos solo un ejemplo de las dificultades a las que nos referimos. En el derecho a la privacidad de los trabajadores en relación con el uso de dispositivos digitales en el ámbito laboral, la norma de referencia es el art. 87 LOPDP-GDD. Es un derecho modulable pero no eliminable, pero creemos que es difícil establecer cómo ha de ejercerse si queremos utilizar los medios menos invasivos posibles, como se deriva del principio de proporcionalidad. Por ejemplo, parece casi imposible renunciar al registro permanente de la actividad de un ordenador, que deja de manera casi inevitable un rastro de datos en la propia máquina y en las redes a las que se conecta. Tal vez pueda limitarse la capacidad sancionadora a partir de datos de los que es casi imposible no disponer, y que además pueden ser interesantes si se analizan globalmente, de forma anonimizada. No vemos fácil en algunos casos la manera de *recurrir a medios menos invasivos* que propone Baz (p. 55), ni tampoco es probable que deban prohibirse los usos secundarios de la información de la manera restrictiva, cuando esos datos pueden anonimizarse y además ser de gran importancia para la gestión de la empresa, debidamente tratados. Es más, la *vigilancia permanente* puede ser ejercida ya sobre el trabajador por terceros ajenos a la empresa,

⁵² De acuerdo con Mercader Urguina (2017, p. 123) la cuestión debe resolverse de acuerdo con el principio de proporcionalidad, para hacer compatibles los derechos fundamentales de alcance laboral con el derecho a la libertad de empresa.

como los proveedores de servicios digitales o los titulares de muchas de las aplicaciones instaladas en los dispositivos móviles o los ordenadores. No significa que esos terceros y la empresa estén interesados en la misma información o la usen con los mismos fines, pero sí sería susceptible de amenazar los mismos derechos a la privacidad, por ejemplo mediante técnicas de *triangulación* o cruce de información de diferentes procedencias⁵³.

Tal vez fuera preferible una interpretación más amplia, en el entendimiento de que el mejor interés del trabajador también se sirve favoreciendo que el empresario pueda ser competitivo a partir del más moderno acceso a todos los datos disponibles que sean compatibles con el respeto a los derechos del trabajador. La transparencia de los criterios y los deberes de información del empresario al trabajador y sus representantes deben servir para hacer compatibles los intereses de todos⁵⁴. Tal vez se podría conseguir poniendo el énfasis más en el uso de la información disponible, la necesidad de agregarla o de anonimizarla para su uso, o la prohibición completa de utilizarla con determinados propósitos. La existencia de nuevas herramientas de IA y la accesibilidad de los datos y las posibilidades de su tratamiento que el legislador nunca pudo haber considerado, es probable que planeen pronto una nueva frontera a la protección del derecho a la intimidad del trabajador.

6. Conclusiones

La IA es una fuerza transformadora que con carácter explosivo está modificando las relaciones sociales y laborales. Se trata de una tecnología de uso general, que, como la electricidad o el motor de explosión, dejará sentir sus efectos a lo largo de los próximos años de maneras que ahora no podemos ni siquiera imaginar.

Como hemos visto, los autores estudian el fenómeno con aproximaciones muy distintas. Algunos ven en el futuro la repetición de ciclos pasados como el de la primera industrialización que llevó a la población del campo a la ciudad, o el de la segunda, que vio aparecer muchos puestos de trabajo nuevos ligados a los servicios. Otros creen que

⁵³ Puede verse *The Economist* (2019b) y también el trabajo de Chen publicado en *The New York Times* (Chen, 2019) (<https://www.nytimes.com/2019/07/03/technology/personaltech/fingerprinting-track-devices-what-to-do.html>). Ambos trabajos ponen de manifiesto la enorme cantidad de información disponible para los proveedores de servicios o los diseñadores de algunas aplicaciones, y la dificultad de impedirlo.

⁵⁴ Sobre regulaciones específicas de la AI, o de aspectos concretos, puede verse (Przegalinska, 2019), (Navas Navarro, 2017) (Kim, 2019) (Holder, Iglesias Portela, Triaille, & Van Gyseghem, 2019) (European Commission, 2018a) (Agrawal, Gans, & Goldfarb, 2019a) (Agrawal et al., 2019b).

esta vez es diferente y creen que no hay que asumir riesgos, pues incluso si fuera relativamente poco probable que nos encontremos en el peor escenario posible, la catástrofe que iría unida a tal escenario hace que no pueda eludirse la introducción de cambios de calado, como la RBU. Tampoco faltan los que invitan a reconsiderar el futuro a partir de la interpretación de la realidad como un fracaso del liberalismo que está poniendo o ha puesto ya en riesgo el mismo sistema democrático.

Creemos, como reflexión final, que urge la reevaluación desde el ámbito jurídico de los problemas que plantea la IA en línea con los documentos del grupo de expertos sobre IA constituido dentro de la UE (European Commission. High-Level Expert Group on Artificial Intelligence. AI-HLEG, 2019c y 2019d).

No se pueden frenar, pero tampoco conviene ser espectadores de los cambios, como si vinieran impuestos indefectiblemente por los avances del aprendizaje automático o los algoritmos. Al contrario, en la línea de las iniciativas de la Comisión Europea, es deseable incorporar pronto cambios legislativos que sirvan para liderar un desarrollo de la IA de acuerdo con principios éticos que respeten los derechos fundamentales, sin frenar el cambio tecnológico que pueda exigir la competitividad europea en un mundo en el que se está globalizando el acceso a algunos datos y también el mismo mercado de trabajo, con el desarrollo de los migrantes digitales a los que nos hemos referido.

7. Bibliografía

- Agrawal, A., Gans, J., & Goldfarb, A. (2018). Prediction machines: the simple economics of artificial intelligence. In *Harvard Business Review Press*.
[https://doi.org/10.1016/0166-3615\(86\)90070-9](https://doi.org/10.1016/0166-3615(86)90070-9)
- Agrawal, A., Gans, J., & Goldfarb, A. (2019a). Economic Policy for Artificial Intelligence. *Innovation Policy and the Economy*, 19, 139–159.
<https://doi.org/10.1086/699935>
- Agrawal, A., Gans, J. S., & Goldfarb, A. (2019b). Artificial Intelligence: The Ambiguous Labor Market Impact of Automating Prediction. *Journal of Economic Perspectives*, 33(2), 31–50. <https://doi.org/10.1257/jep.33.2.31>
- Baldwin, R. (2019). *The Globotics Upheaval: Globalization, Robotics, and the Future of Work*. New York: Oxford University Press.
- Baz Rodríguez, J. (2019). La Ley Orgánica 3/2018 como marco embrionario de garantía de los derechos digitales laborales. Claves para un análisis sistemático. *Trabajo y Derecho*, (54), 48–78.
- Bregman, R. (2016). *Utopia for Realists. The Case for a Universal Basic Income, Open Borders and a 15-Hour Workweek*. The Correspondent.
- Bruun, E. P. G., & Duka, A. (2018). Artificial Intelligence, Jobs and the Future of Work: Racing with the Machines. *Basic Income Studies*.
<https://doi.org/10.1515/bis-2018-0018>
- Chen, B. X. (2019, July 3). ‘Fingerprinting’ to Track Us Online Is on the Rise. Here’s What to Do. *The New York Times*.
- del Rey Guanter, S. (Dir). (2018). *Inteligencia artificial y su impacto en los recursos humanos y en el marco regulatorio de las relaciones laborales [Proyecto Technos]*. Las Rozas Madrid: Cuatrecasas y Wolters Kluwer.
- Domingos, P. (2015). *The Master Algorithm. How the Quest for the Ultimate Learning Machine Will Remake Our World*. Retrieved from
<https://folk.idi.ntnu.no/agnar/Documents/Domingos-SVM-NN-CBR.pdf>
- Du Sautoy, M. (2019). *The Creativity Code. Art and Innovation in the Age of AI*. Cambridge, Massachusetts: The Belknap Press of Harvard University Press.

- European Commission. High-Level Expert Group on Artificial Intelligence. AI-HLEG. (2019a). *A Definition of AI: Main Capabilities and Disciplines*. Retrieved from <https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines>
- European Commission. High-Level Expert Group on Artificial Intelligence. AI-HLEG. (2019b). AI4EU – Advancing Europe through collaboration in AI. Retrieved June 21, 2019, from <https://www.ai4eu.eu/#about>
- European Commission. High-Level Expert Group on Artificial Intelligence. AI-HLEG. (2019c). *Ethics guidelines for trustworthy AI*. Retrieved from <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>
- European Commission. High-Level Expert Group on Artificial Intelligence. AI-HLEG. (2019d). *Policy and Investment Recommendations for Trustworthy AI*. Retrieved from <https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence>
- European Commission. (2018a). *Artificial Intelligence for Europe. COM(2018) 237 final. C*. Retrieved from <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN>
- European Commission. (2018b). Artificial Intelligence for Europe COM(2018) 273 final. *Digital Single Market*, 2. <https://doi.org/10.1628/978-3-16-157046-9>
- Garicano, L. (2019). *El contrataque liberal. Entre el vértigo tecnológico y el caos populista*. Barcelona: Ediciones Península.
- Harari, Y. N. (2018). *21 Lessons for the 21st Century*. London: Johathan Cape.
- Holder, C., Iglesias Portela, M., Triaille, J. P., & Van Gyseghem, J.-M. (2019). Legal and regulatory implications of Artificial Intelligence. The case of autonomous vehicles, m-health and data mining. In *European Commission, Luxembourg*. <https://doi.org/10.2760/527052> (online)
- Kim, P. T. (2019). Bid Data and Artificial Intelligence: New Challenges for Workplace Equality. *University of Louisville Law Review*, 57(June), 1–16.
- Lassalle, J. M. (2019). *Ciberleviatán. El colapso de la democracia liberal frente a la revolución digital*. Barcelona: Arpa.

- McCarthy, J., Minsky, M., Rochester, N., & Shannon, C. (1955). *A Proposal for the Summer Research Project on Artificial Intelligence*. Retrieved from <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>
- Mercader Uguina, J. R. (2017). *El futuro del trabajo en la era de la digitalización y la robótica*. Tirant lo Blanch.
- Mercader Uguina, J. R. (2018). El impacto de la robótica y el futuro del trabajo. *Revista de La Facultad de Derecho de México*.
<https://doi.org/10.22201/fder.24488933e.2017.269.62438>
- Moore, P. V. (2019). *OSH and the Future of Work: Benefits and Risks of Artificial Intelligence Tools in Workplaces*. Retrieved from <https://osha.europa.eu/en/tools-and-publications/publications/osh-and-future-work-benefits-and-risks-artificial-intelligence/view>
- Navas Navarro, S. (et al. . (2017). *Inteligencia artificial. Tecnología-a. Derecho*. Valencia: Tirant lo Blanch.
- O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown Publishing Group.
- Przegalinska, A. (2019). *Wearable Technologies in Organizations: Privacy, Efficiency and Autonomy in Work*. Palgrave Macmillan.
- Ramge, T. (2019). *Who's Afraid AI? Fear and Promise in the Age of Thinking Machines*. New York: The Experiment.
- Roose, K. (2019, June 24). A Machine May Not Take Your Job, but One Could Become Your Boss. *New York Times*. Retrieved from <https://www.nytimes.com/2019/06/23/technology/artificial-intelligence-ai-workplace.html>
- Servoz, M. (2019). The Future of Work? Work of the Future! On How Artificial Intelligence, Robotics and Automation are Transforming Jobs and the Economy in Europe. In *AI Report, European Commission*. Retrieved from <https://ec.europa.eu/digital-single-market/en/news/future-work-work-future>

The Economist. (2019a, May 23). *Across the rich world, an extraordinary jobs boom is under way*. Retrieved from

<https://www.economist.com/briefing/2019/05/23/across-the-rich-world-an-extraordinary-jobs-boom-is-under-way>

The Economist. (2019b, May 23). *Online identification is getting more and more intrusive - Behavioural biometrics*. pp. 05–23. Retrieved from

<https://www.economist.com/science-and-technology/2019/05/23/online-identification-is-getting-more-and-more-intrusive>

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. In *Profile Books*. New York: PublicAffairs Books.