

¿PUEDE LA INTELIGENCIA ARTIFICIAL SUSTITUIR A LA MENTE HUMANA? IMPLICACIONES DE LA IA EN LOS DERECHOS FUNDAMENTALES Y EN LA ÉTICA *

Can Artificial Intelligence Substitute the Human Mind? Implications of AI in Fundamental Rights and Ethics

MILAGROS OTERO PARGA **

Fecha de recepción: 25/7/2022
Fecha de aceptación: 01/09/2022

Anales de la Cátedra Francisco Suárez
ISSN: 0008-7750, núm. 57 (2023), 39-61
<http://dx.doi.org/10.30827/ACFS.v57i.24710>

RESUMEN La inteligencia artificial (IA) es una realidad cuyo desarrollo plantea múltiples problemas en la actualidad. El conflicto surge entre las posibilidades técnicas reales, y las consecuencias éticas de implementarla hasta sus últimas consecuencias. El presente trabajo analiza alguna de estas problemáticas y sostiene que la IA, es un instrumento al servicio del ser humano y sólo así debe ser utilizado, so pena de convertirse en un recurso ingobernable. Urge, por tanto, un ejercicio de responsabilidad que proteja el humanismo y los derechos fundamentales, defendiendo que, el ser humano es, y debe seguir siendo, el centro de la creación.

Palabras clave: Inteligencia Artificial, Humanismo, Ética, Derechos Humanos, Responsabilidad Social.

ABSTRACT Artificial intelligence (AI) is a reality whose development raises multiples problems nowadays. The conflict arises between the real technical possibilities and the ethical consequences of implementing them to their ultimate consequences. The present work analyzes some of these problems and maintains that AI is an instrument at the service of the human, being and only in this way should it be used under penalty of becoming an ungovernable resource. Therefore, an exercise of responsibility is urgently needed to protect humanism and fundamental rights, understanding that the human being is, and must continue to be, the center of creation.

Keywords: Artificial Intelligence, Humanism, Ethics, Human Rights, Social Responsibility.

* Para citar/citation: Otero Parga, M. (2023). ¿Puede la inteligencia artificial sustituir a la mente humana? Implicaciones de la IA en los derechos fundamentales y en la ética. *Anales de la Cátedra Francisco Suárez* 57, pp. 39-61.

** Universidad de Santiago de Compostela. Departamento de Derecho Público Especial y de la Empresa. Avda. Dr. Ángel Echeverri, s/n. 15782 Santiago de Compostela (España). Correo electrónico: milagros.otero@usc.es

1. INTRODUCCIÓN

Uno de los poemas de la poetisa gallega Rosalía de Castro que siempre me ha gustado especialmente se titula *El clavo*. En él, la famosa escritora escribe que tuvo un clavo clavado en el corazón. Ese clavo le producía mucho dolor y siempre pugnaba por sacárselo. Y un día lo consiguió, y se lo arrancó, pero después, ya no volvió a sentir nada; ni alegría ni dolor, y la soledad que experimentó por esa realidad, fue mucho más penosa que el dolor del clavo¹.

Traigo este poema a colación para adelantar mi posición sobre lo que es la Inteligencia Artificial, en adelante (IA). Creo que el contenido de este hermoso poema nos acerca a algunos de los problemas que plantea el desarrollo de la IA.

La llamada IA es, sin duda, un avance importantísimo de la ciencia, una mejora crucial diría yo, para propiciar algunos de los descubrimientos más sobresalientes de los últimos años de nuestra historia, y desde ese punto de vista debe ser estudiada, potenciada y valorada.

Pero la IA es artificial, no es natural y desde esa perspectiva entiendo que ni sustituye al ser humano ni debe hacerlo. Y por lo mismo debe ser utilizada dentro de los parámetros que corresponda, a fin de evitar el riesgo de situarla por encima de la propia humanidad.

Las máquinas entiendo, por muy perfeccionadas que puedan estar, nunca deben sustituir a las personas, pues si llegaran a hacerlo, asistiríamos al fin de la humanidad y de nuestro mundo concebido como hasta ahora lo hemos hecho.

Las máquinas sin duda pueden retener muchos más datos que las mentes humanas², pueden almacenar mucha más información, pueden combinarla, pueden retenerla y usarla simulando la memoria; pueden hacer cálculos basados en probabilidades sustituyendo la inteligencia y podrían, incluso, manifestar una forma de actuación constante suplantando a la

-
1. “Una vez tuve un clavo, clavado en el corazón, y yo no sé si aquel clavo era de oro, de hierro o de amor. Sólo sé que me hizo un mal tan hondo, que tanto me atormentó, que día y noche sin cesar lloraba, cual lloró Magdalena en la pasión. Señor que todo lo puedes, le pedí una vez a Dios, dame valor para arrancar de un golpe, clavo de tal condición. Y me lo dio Dios y lo arranqué, pero ¿quién pensara? después, ya no sentí más tormentos ni supe lo que era dolor; supe sólo que no sé qué me faltaba en donde el clavo faltó, y tal vez, tal vez, tuve soledades de aquella pena... ¡Buen Dios! Este barro mortal que envuelve el espíritu, ¡quién lo entenderá Señor!”. La traducción gallego-castellano es mía.
 2. De hecho, construimos computadores y programas por muchas razones; para servir a la sociedad y como herramientas para realizar las tareas económicas de la sociedad” (Newell y Simon, 1994, p. 123).

voluntad. De manera que memoria, inteligencia y voluntad, que son las tres potencias fundamentales del alma, podrían resultar similares en una máquina perfeccionada y en un ser humano. Pero siempre la primera sería una máquina y el segundo una persona. Confundirlas es peligroso desde el punto de vista humano, ético y jurídico.

Una máquina, aunque sea muy perfecta, no puede sentir la belleza de la poesía con la que iniciamos este trabajo, no puede emocionarse con su contenido, no puede incluso derramar una lágrima, producto del sobrecoimiento que produce estar ante algo verdaderamente hermoso.

Una máquina, seguramente podría escribir una poesía, pero nunca sentirla.

Por otro lado, las máquinas han sido pensadas, planeadas y corregidas por las personas, de manera que, de algún modo, y aunque lleguen a superarlas en prestaciones concretas, siempre dependerán de ellas. La hipótesis de que no fuera así es, a mi juicio, espeluznante. Por eso coincido con Truyol Serra cuando afirma: “hoy como siempre, estamos ante la tarea de conseguir que los avances tecnológicos, arrolladores, sean dominados por el hombre y puestos al servicio de su perfeccionamiento, en vez de implicar, como pudiera ocurrir y no pocos temen, una mayor o menor dependencia enajenante con respecto a los mismos” (Truyol, s/f, p. 118).

Es cierto que las máquinas han avanzado y siguen haciéndolo a un ritmo vertiginoso, de manera que resulta difícil saber hasta dónde pueden llegar³ (López Oneto, 2020, p. 229). Pero no es menos verdad que el avance de las máquinas debe ser controlado por las personas, de forma que sea acorde con los derechos humanos y respete los principios elementales de la ética. Ambos, derechos humanos y ética, están amenazados hoy en día⁴ (Barona, Vilar, 2021, p. 18).

En el trabajo que ahora inicio quiero hacer un pequeño estudio de cómo surgió y de lo que es la IA⁵ (Nilson, 2001, p. 1). A continuación,

-
3. “La humanidad está transitando hacia un nuevo tipo de civilización. Por supuesto es muy difícil, por no decir imposible, describir cómo será, en sus aspectos básicos, esta sociedad del futuro. Sin embargo, podría conjeturarse que esa imaginaria sociedad del porvenir sería una sociedad global organizada en red, impulsada por la sinergia de las ciencias y las tecnologías”.
 4. Advierte Barona que “la crisis de la humanidad está servida. Los sistemas algorítmicos controlados humanamente y los que carecen de supervisión humana han inundado nuestras vidas, adquiriendo un enorme poder en la sociedad del siglo XXI, un poder económico privado, dado que las capacidades avanzadas de la misma están acumulándose y concentrándose en manos del sector privado, y, por lo general, con derechos exclusivos”.
 5. Según Nilson, “la IA es una definición un tanto circular, que tiene por objeto el estudio del comportamiento inteligente de las máquinas. A su vez el comportamiento inteligente supone percibir, razonar, aprender, comunicarse y actuar en entornos complejos”.

veré algunas de sus posibles implicaciones en el ámbito de la ética y de los *humana iura*; y, por último, concluiré afirmando que la IA supone un gran avance al tiempo que facilita y mejora la actividad de los seres humanos en el mundo especialmente científico. Desde esa perspectiva debe ser estudiada y perfeccionada como coadyuvante de la labor humana. Pero todo eso tiene que hacerse sin olvidar que el ser humano es, y debe seguir siendo, el rey de la creación, y como tal, debe gobernarla y gobernar sus designios, aprovechando para ello todos los avances que ofrece la propia IA⁶ (Crespo García, 2016, p. 14).

De modo que la IA y su evolución deben supeditarse a la humanidad (Nilson, 2000, p. 2)⁷ porque no hacerlo así, transgrediría de una forma definitiva el orden natural, trayendo como consecuencia situaciones imposibles de prever pero que sin duda alterarían definitivamente nuestra manera de ser y de vivir⁸ (Nilson, 2000, p. 270). Por eso, coincido con Lagares cuando advierte de que “la evolución tecnológica y el desarrollo de las ciencias jurídicas deben ir de la mano, proyectándose a un ritmo prácticamente común y constante, no quedando en detrimento nunca una respecto de la otra” (Lagares, 2000, p. 98).

2. EL SURGIMIENTO DE LA IA. BREVE RESEÑA HISTÓRICA

Tal y como afirman Hintos, McClelland y Rumelhart, las personas pueden pensar y pueden recordar a voluntad aquello que han pensado en el momento en que precisen hacerlo, esto es, tienen memoria. Y sin embargo eso no es suficiente⁹ (Hintos, McClelland y Rumelhart, 1994, p. 280).

6. Dice Crespo García que son “numerosos y significativos los escenarios y aplicaciones en los que el Big Data conlleva la obtención de grandes logros y beneficios. Entre ellos pueden destacarse la prevención, detección e investigación de actividades terroristas y delincuencia (y también avances) en el campo de la salud y atención sanitaria.

7. Afirma Nilson que “la cuestión de si alguna vez seremos capaces, o no, de construir máquinas pensantes de nivel humano no admite aún una respuesta definitiva. El progreso de la IA hacia esta meta ha sido constante, aunque más lento de lo que algunos pioneros del tema habían predicho”.

8. Coincido con Nilson en que “el conocimiento de sentido común era el adecuado y aún lo es para muchas de las cosas que los hombres querían hacer. El pensamiento científico, poco a poco, se fue separando del sentido común, a medida que la gente buscaba descripciones más precisas de su mundo”.

9. “Las personas pueden tener acceso a su memoria de una manera muy flexible; pueden recordar elementos a partir de descripciones parciales de sus contenidos... Además, pueden hacerlo, aunque algunos aspectos de la descripción parcial estén equivocados”.

Poco a poco, con prisa y sin pausa se han ido introduciendo en nuestra vida sistemas computacionales que pretenden realizar procesos mentales y organizativos con mucha mayor rapidez¹⁰ que lo que podría hacer cualquier mente humana¹¹.

Estas máquinas, juegan con algoritmos y con bases de datos que pueden conseguir un sistema de almacenamiento, de infinidad casi incalculable de información, que se pueden recuperar con rapidez y exactitud en el momento en que se formule a la máquina la pregunta adecuada.

El proceso al que me refiero se inició a principios de la década de los años 50¹² y el resultado, que estamos experimentando hoy en día, presenta, a mi juicio, luces y sombras (Dreyfus & Dreyfus, 1994, p. 344).

Sus antecedentes son claros y se inician en la obra de Leibniz. Según narra Trillas, “Leibniz había estudiado ávidamente los escritos de Lull y los adaptó de forma crítica al proponer un método objetivo y mecánico para poder fundamentar las indagaciones racionales; de hecho, su libro *Dissertatio de arte combinatoria*, escrito en 1666 es la mejor crítica y homenaje que Lull haya recibido. A lo largo de cuatro siglos, Lull influyó en importantes pensadores como Montaigne, Pascal, Descartes, Giordano Bruno y Newton... Sin embargo, el paso decisivo para poder llegar a considerar los razonamientos como un tipo especial de cálculo, que permitiese pasar de las premisas o datos a las consecuencias o soluciones por medio de la aplicación de reglas de “cálculo” en número finito de pasos, tuvo que esperar hasta la mitad del siglo XIX” (Trillas, 1998, p. 29).

La expresión inteligencia artificial (IA) nació concretamente en 1956 en los Estados Unidos de Norteamérica durante una conferencia en Dartmouth. Su creación se debió a un conjunto de científicos cuyo interés se

10. Aunque en realidad, “el cerebro quizá ostente muchos más mapas topográficos de los que hasta ahora se han investigado o sospechado” (Churchland, 1994, p. 391).

11. “En los primeros días de la IA, su grito de guerra era: “las computadoras no mascan números, manipulan símbolos”. Con ello se pretendía inspirar a un público incrédulo, mostrándole que la computación era muy similar al pensamiento. Ahora la rueda ha completado su giro. La virtud de los sistemas conexionistas radica al parecer en que “no manipulan símbolos, mascan números”. Y en la actualidad todos sabemos (¿verdad?) que ¡el pensamiento no es una mera manipulación de símbolos! Y la rueda sigue girando (Clark, 1994, p. 342).

12. Dicen Dreyfus y Dreyfus, que “a principios de la década de los años 50 cuando las máquinas calculadoras lograron el éxito que merecían, algunos pensadores pioneros comenzaron a darse cuenta de que las computadoras digitales podían hacer más que masticar números. En ese punto surgieron dos visiones opuestas acerca de lo que podían ser las computadoras... Unos pretendían usarlas para ejemplificar concretamente una representación formal del mundo; los otros, (querían) simular las interacciones de las neuronas”.

centraba en escribir programas cuya finalidad era dotar de inteligencia a los ordenadores.

En 1944, y después de largos años de trabajo, algunos técnicos de Harvard junto con los de IBM culminaron la primera máquina que podía realizar cálculos electrónicos. Poco después en 1946 en Pensilvania se construyó la primera máquina de calcular totalmente electrónica¹³, que puede considerarse el primer ordenador¹⁴ de propósito general y digital electrónico (Trillas, 1998, p. 38).

Poco después, en 1950 surgió la teoría de Alan Turing que pronto se convirtió en el “sustento” de las teorías y de los trabajos en IA. La teoría de Turing trataba, en resumen, de averiguar si una máquina podía ser inteligente o no, y para ello planteaban una pregunta jurídica a un ordenador con el fin de averiguar si la máquina podía dar la misma respuesta que un agente humano (Bourcier, 2003, p. 57).

Turing pensaba, sin ningún tipo de duda, que llegaría a haber, según él en el año 2000, máquinas que podrían pensar de la misma manera que las personas.

Sabía, desde luego, que a su opinión se podían hacer una serie de objeciones importantes de distinto calado, como las llamadas teológicas, la de cabeza de arena, matemáticas etc., pero él fue respondiendo a cada una de ellas (Turing, 1994, pp. 63 ss.), buscando argumentos que pudieran avalar su teoría. Todo ello con la intención de probar que en las críticas a la IA en realidad “no se ofrecía ningún fundamento para estas afirmaciones. Pienso (decía) que en su mayoría se basan en el principio de la inducción científica” (Turing, 1994, p. 67). Y por eso, continuaba Turing, a pesar de todos estos inconvenientes, “con el tiempo podríamos esperar que las máquinas llegaran a competir con el hombre en todos los campos puramente intelectuales... el proceso podría ser similar al de enseñanza de un niño... hay mucho por hacer” (Turing, 1994, p. 80).

Para probar y reafirmar sus teorías elaboró una prueba llamada *test de Turing*. Esta prueba consistía en la interacción de un juez con una máquina. El juez debería formularle una serie de preguntas a la máquina durante 5 minutos y la prueba se consideraba un éxito siempre y cuando más del 30% de los jueces participantes (la prueba se realiza en la actualidad con varios jueces), después de 5 minutos de conversación, llegara a la conclusión de que la computadora era un ser humano¹⁵.

13. Su construcción se debió a J. Eckert y J. Mauchly.

14. Su nombre fue ENIAC, Electronic Numerical Integrator and Calculator.

15. En otro experimento de Turing, participan tres personas. Un hombre (A), una mujer (B) y un examinador (C), que puede ser de cualquier sexo. El examinador permanece en una

Conviene aclarar que esta prueba no ha sido pasada satisfactoriamente por ninguna máquina. Y de todas formas parece evidente que la IA presenta un grado de dependencia excesivamente alto de las capacidades de los ordenadores, que no todos tienen las mismas, y por lo mismo de todas las tecnologías electrónicas y fotónicas (Trillas, 1998, p. 64).

Pese a ello, y aún con estos inicios, los trabajos sobre IA siguieron avanzando, pues sus defensores estaban firmemente convencidos de poder conseguir mediante un aprendizaje, que equiparaban al de los animales o al de los niños, unas máquinas capaces de pensar, de sentir emociones, de tener conciencia y, en suma, de igualar o incluso superar la propia mente humana.

La implementación de la IA es la opción de la Unión Europea que lleva tiempo trabajando en un marco legal apropiado¹⁶ para el uso y aprovechamiento de la IA¹⁷. Su finalidad sería: 1) Permitir la adopción y desarrollo de la IA en Europa. 2) Conseguir su prosperidad en el mismo marco geográfico. 3) Garantizar que la IA no sólo funcione, sino que ayude a la sociedad y 4) facilitar la construcción de un liderazgo europeo en sectores estratégicos de impacto.

Se trata de conseguir un marco de IA europeo fiable que aborde la protección de los derechos fundamentales y evite los riesgos de seguridad, al tiempo que proponga un nuevo marco de responsabilidad civil adaptado a las nuevas necesidades.

Frente a los que así opinan se sitúan otros que afirman que “ningún modelo puramente formal será jamás suficiente por sí mismo para la intencionalidad¹⁸, porque las propiedades formales no son ellas mismas constitutivas de la intencionalidad y porque carecen por sí mismas de poder causal, salvo el poder de producir, cuando se ejemplifican correctamente, la siguiente etapa del formalismo cuando la máquina está funcionando” (Searle, 1994, p. 97).

Y otros simplemente dudan, y sus dudas se manifiestan en interrogantes como estos: “la pregunta importante no es ¿Cuándo entiende algo una

habitación apartado de los otros dos. El objeto del juego para el examinador consiste en determinar cuál de las otras dos personas es el hombre y cual la mujer (Turing, 1994, p. 53)

16. El último pronunciamiento europeo fue en mayo de 2022. Se trata de una Resolución del Parlamento Europeo sobre IA en la era digital.

17. <https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence>.

18. Pues según Searle “sea lo que fuere lo que el cerebro haga para producir intencionalidad, ello no puede consistir en ejemplificar concretamente un programa, ya que ningún programa por sí mismo es suficiente para lograr la intencionalidad” (Searle, 1994, p. 103).

máquina? Si no, ¿qué cosas necesita ser capaz de hacer una máquina a fin de poder comprender?” (Boden, 2017, p. 120).

3. ¿QUÉ ES LA INTELIGENCIA ARTIFICIAL? ¿CUÁLES SON SUS UTILIDADES?

La evolución de la IA fue imparable desde sus comienzos. Poco a poco fue extendiéndose su uso y ampliándose los estudios para intentar abarcar su concepto y conocimiento.

Pero ¿qué es en realidad la IA? Podemos decir junto con Bourcier que, “la inteligencia artificial es una rama de la informática que intenta reproducir las funciones cognitivas humanas como el razonamiento, la memoria, el juicio o la decisión, y después confiar una parte de estas facultades, que consideramos signos de inteligencia, a los ordenadores” (Bourcier, 2003, p. 54).

Las utilidades más importantes de este proceso son las siguientes:

- 1) Comunicación hombre-máquina mediante lenguaje natural.
- 2) Reconocimiento de las palabras para poder hablar a una máquina, pero también para hacer que la máquina hable.
- 3) Interpretación de imágenes y
- 4) Creación y utilización de sistemas expertos vinculados a campos de conocimiento muy especializados (Bourcier, 2003, p. 56).

El objetivo de la IA es “identificar problemas interesantes y resolubles del procesamiento de información para resolverlos” (Marr, 1994, p. 153).

En realidad, “la tarea que enfrenta el investigador en IA parece ser la de diseñar un sistema que pueda planificar, utilizando elementos bien seleccionados tomados de su almacén de conocimientos acerca del mundo en el que opera” (Dennett, 1994, p. 176).

Una vez identificado el objetivo, la IA podría “aislar un problema particular de procesamiento de información, formular una teoría computacional para él, construir un algoritmo que lo lleve a cabo y demostrar en la práctica que el algoritmo es adecuado” (Marr, 1994, p. 154).

Claro está, no siempre se consiguen buenos resultados aún identificado el objetivo. En realidad “la principal dificultad de la IA es que nunca se tiene la suficiente seguridad de que el problema tiene una teoría que lo resuelva... Pero no encontrarla no significa que no exista” (Marr, 1994, p. 155).

Sea como fuere, y aún con la posibilidad de los fallos que comentamos, resulta claro que las funciones o utilidades de la IA son muchas y muy

variadas¹⁹. Algunas de ellas, como la recopilación de datos, interpretación de imágenes, creación de sistemas expertos, reconocimiento de palabras, son muy claras y me parecen muy útiles. Son labores propias de máquinas que en realidad pueden ayudar mucho a las personas.

Por otro lado, creo que hay operaciones de inteligencia que precisan la utilización de características que sólo las personas tienen y nunca podrán tener las máquinas, por ejemplo, las emociones, la conciencia, la intuición etc. Una máquina tiene una capacidad de relación y almacenamiento muy grande²⁰, mayor que la de los seres humanos. Creo que también puede tenerla de reacción e incluso de respuesta en casos concretos, y siempre que se le hayan proporcionado previamente ciertos datos, pero no creo que vaya más allá. Muy al contrario, creo, incluso, que no debe hacerlo.

En el ámbito de lo jurídico, que es el que ahora nos interesa, creo que la IA buscó facilitar la interpretación, mejorar la localización de legislación y tratar de corregir la oscuridad del lenguaje jurídico excluyendo ambigüedades de los textos mediante la utilización de la lógica (Bourcier, 2003, p. 65).

Efectivamente, el lenguaje presenta un campo muy amplio e interesante para el trabajo de la IA. Por eso, “los lingüistas computacionales intentan encontrar una teoría computacional del lenguaje. Buscan respuestas a estas preguntas: ¿Cómo se identifican las estructuras de las frases? ¿Cómo puede ser modelado el conocimiento y el razonamiento? ¿Cómo puede ser usado el lenguaje para realizar tareas específicas?” (González Quevedo, 1999, p. 7).

Esta labor es necesaria porque lamentablemente asistimos a un momento en el que los operadores jurídicos se expresan muy mal (Otero Parga, 2011, pp. 139-157). Los legisladores hacen leyes ininteligibles, con repeticiones, frases larguísimas e incomprensibles, ausencia de signos de puntuación etc. En suma, las leyes son difíciles de leer y de interpretar y esa característica, como es lógico, no exime de su cumplimiento, pero sí lo dificulta sobremanera.

Los jueces lamentablemente no se pronuncian mejor. Las sentencias que contienen fallos cuyo contenido concede o deniega derechos son incomprensibles para la mayoría de los ciudadanos. Incluso lo son para los profesionales que deben, en muchas ocasiones, leer varias veces el contenido completo de las sentencias para poder entenderlas. Y eso se hace especialmente preocupante cuando es preciso recurrirlas.

19. *Vid.*: (Otero Parga, 2012, pp. 18-32).

20. “Muchos críticos de la IA, están convencidos de que cualquier sistema de IA no es, ni debe ser, nada más que engranajes de ruedas de conocimiento” (Dennett, 1994, p. 188).

Por lo que se refiere al poder ejecutivo el problema es similar. Es raro el mandato claro, coherente, completo y sencillo de entender.

Ante esta situación real resulta muy útil la participación de la IA en el mundo del lenguaje en general y del jurídico en particular. Un ejemplo es el sistema PROLOG que es “un lenguaje y un demostrador de teoremas. A partir de un problema planteado bajo la forma de predicado y una base de conocimientos organizada en forma de cláusulas, busca todas las soluciones en las que el predicado es verdadero” (Bourcier, 2003, p. 67).

Sin duda la utilidad es clara, pero hay que tener precaución en su utilización pues “el lenguaje de programación para el derecho aseguraría la misma función que la ley, por lo que sería un lenguaje de control de la sociedad. De aquí a programar directamente la sociedad mediante ordenadores habría solo un paso” (Bourcier, 2003, p. 67).

Y ¿qué sucede con los llamados sistemas de expertos? Estos sistemas datan de los años 70. Se trata de herramientas de IA muy conocidas en los campos efectivos de su utilización. Por ejemplo, en el ámbito de lo jurídico, los sistemas de expertos podrían diagnosticar, aconsejar o asistir en una decisión, siendo útiles, entre otras cosas para las siguientes:

- 1) Asistencia en la decisión.
- 2) Localización de una solución jurídica mediante “reglas válidas”.
- 3) Tratamiento de cuestiones de derecho que puedan surgir del problema en debate.
- 4) Ofrecimiento de un medio para proporcionar un resultado válido y justificable jurídicamente (Bourcier, 2003, p. 71).

Como se puede observar de nuevo, la IA aporta elementos que facilitan y hacen incluso más segura la actividad de los operadores jurídicos, pero coincido con Bourcier cuando afirma que “las máquinas de pensar no tratan más que de una parte del conocimiento de los juristas, pero lo que, a la interpretación, la íntima convicción o el sentimiento de justicia (se refiere), quedarán como operaciones axiológicas que nunca podrán ser generalizadas en un ordenador” (Bourcier, 2003, p. 91).

Y es que las emociones no pueden traducirse, preverse o contemplarse por una máquina. Lo mismo sucede con un estado de ánimo que “puede afectar a la manera en que percibimos las cosas, interpretamos los actos de los demás, predecimos las consecuencias de los actos, realizamos planes etc.” (Sloman, 1994, p. 271). Eso, no puede hacerlo ni sentirlo una máquina por muy avanzada o perfeccionada que esté.

En cualquier caso, no podemos negar las grandes ventajas de la IA, y, es más, debemos fomentar su investigación y su uso en todo aquello que

nos beneficie y haga más claro, rápido y seguro el resultado de la resolución de un problema.

Pero, nos preguntamos: ¿Puede un ordenador pensar? Creo con Copeland que “esto está lejos de ser cierto... sin duda (la IA) consigue programas asombrosos, pero considerémoslos más atentamente y veremos que sus espléndidas presentaciones son, en su mayor parte simuladas” (Copeland, 1993, p. 64).

Y aún continúa: Si puede o no pensar un ordenador es un asunto filosófico y no técnico, pues “es un supuesto hartamente común que sólo los seres conscientes pueden pensar... en realidad no es obvio ni mucho menos, que un artefacto no pueda ser consciente... la percepción no tiene por qué ser una actividad consciente” (Copeland, 1993, pp. 65-67).

Desde esta perspectiva, los ordenadores pueden hacer muchas y grandes cosas, pero no pensar, porque éstos aun cuando parecen que están pensando “en realidad emplean una técnica conocida como comparación de patrones” (Copeland, 1993, p. 75).

Es más “todavía nadie sabe hacer que los computadores actúen con nada que se asemeje al grado humano de pericia en dominios más amplios y menos ordenados, y ni siquiera si esto es posible” (Copeland, 1993, p. 148).

La IA tiene dos objetivos principales “uno es tecnológico; usar los ordenadores para hacer cosas útiles (a veces empleando métodos muy diferentes a los de la mente). El otro es científico, consiste en usar conceptos y modelos de IA que ayuden a resolver cuestiones sobre los seres humanos y demás seres vivos” (Boden, 2017, p. 12).

Si la IA y sus cultivadores se ocuparan de realizar estas dos funciones no nos estaríamos preguntando si su capacidad superará la humana, pero en la situación actual debemos hacerlo, ya que “la IA ha puesto en entredicho nuestro concepto de humanidad y su futuro. Algunos incluso dudan de si de hecho tendremos futuro, porque prevén que la IA superará la inteligencia humana en todos los ámbitos” (Boden, 2017, p. 13).

Efectivamente dudan, y no sólo dudan, algunos pensadores, como hemos visto están totalmente convencidos de que la IA tiene mucho más que ofrecer al ser humano que su pobre mente falible y limitada.

Yo no pertenezco a ese grupo. Es más, desconfío bastante de la IA. No tanto como quienes afirman que no existe²¹, pero sí lo suficiente para confiar más en la mente humana que, con sus fallos e imperfecciones es

21. Según Carabantes López, “lo cierto es que la IA es una técnica que todavía no existe. Hay máquinas que hacen cosas asombrosas, como jugar al ajedrez mejor que el mejor ajedrecista humano, pero no se puede decir, en rigor que sea auténtica IA. Su inferioridad respecto a nosotros se debe, principalmente, a que carecen de dos habilidades; en el mundo social,

humana, y por lo mismo está dotada de posibilidades como inteligencia, voluntad, sentido común, capacidad de improvisación, sentimientos, sensaciones etc., que le están vedados a las máquinas²².

Pues como nos recuerda González Ruiz, parafraseando a Dreyfus, “las computadoras son dispositivos analíticos, esto es, funcionan razonando a partir de reglas y descripciones de hechos; sin embargo, la pericia (la calidad de “experto”), depende básicamente de la intuición (inalcanzable para las computadoras) y no del razonamiento; por consiguiente, las computadoras no pueden aspirar a comportarse con la pericia propia de los expertos humanos” (González Ruiz, 1992b, p. 27).

En realidad, “una computadora sólo puede producir lo que el programador ha puesto en ella y no algo más” (Negrete, 1992, p. 241). Con esta afirmación el problema debería quedar resuelto, de modo que entendemos y defendemos que la inteligencia artificial es una forma nueva de conseguir, ordenar y recuperar datos y conocimientos que se somete a las reglas de unos sistemas de expertos previamente establecidos y metidos en un programa.

Por la gran amplitud de datos y algoritmos que puede manejar se convierte en un instrumento de extraordinaria utilidad para los seres humanos, que ya ha avanzado de forma imparable, pero que no sabemos aún hasta dónde puede llegar.

Los productos de la IA deben ser utilizados, desde luego, pero manejados con precaución porque pueden interferir, e incluso hacerlo grandemente, en el desarrollo de los derechos humanos y en el respeto por la ética en las actuaciones de los poderes del Estado.

Dedicaremos los próximos dos epígrafes a analizar esta realidad.

4. INTELIGENCIA ARTIFICIAL Y DERECHOS HUMANOS

El ámbito de lo jurídico, no debe, no puede, dar la espalda a las continuas modificaciones que se producen en el mundo. No puede hacerlo porque los operadores jurídicos, todos ellos, están al servicio de la justicia y de los ciudadanos y malamente podrían servirlos si no evolucionaran con los tiempos.

del lenguaje, y en el mundo físico, de nuestra versatilidad para hacer un intento pasable en casi cualquier cosa” (Carabantes López, 2016, p. 10).

22. Como dice González Ruiz, “los sistemas de expertos poseen “inteligencia resolutive” y no “inteligencia propositiva”, esto es, capacidad de resolver problemas, pero no de generarlos”, (González Ruiz, 1992a, p. 19).

Sin embargo, y aun admitiendo este principio, es necesario vigilar que la utilización de las nuevas tecnologías de la IA no ponga en riesgo el sistema, atacando o desprotegiendo los derechos fundamentales.

Este problema preocupa a los juristas hasta el punto de que “los estudiosos de los asuntos internacionales y de las más altas instituciones en las que se intenta establecer y mantener un orden, y una justicia en el mundo civilizado, tienen cada vez más difícil encontrar un equilibrio entre la libertad y la seguridad” (VV.AA., 1996, p. 16).

No me extrañan estas palabras, es más, las considero un poco escasas incluso. La IA ha puesto en desequilibrio claro los derechos humanos fundamentales, y no sólo el de la libertad versus seguridad, sino otros muchos.

En realidad, el uso de internet, por ejemplo, como forma de IA ha traído múltiples problemas de ámbito jurídico; problemas que encierran dentro de sí cuestiones previas y consecuencias que deben ser resueltas en la tirantez entre evolución y seguridad (VV. AA.b), 2000, pp. 25 ss.).

Todos esos problemas ponen en jaque muchos derechos humanos que de forma directa o indirecta se ven involucrados con la IA. Entre ellos, protección de datos, contenidos lícitos e ilícitos de internet, comercio electrónico, firma electrónica, certificación, pornografía infantil etc.

Llamo la atención de nuevo sobre el hecho de que todas estas aplicaciones de la IA, excepto la última claro está, son legítimas en sí mismas, e incluso plausibles. El problema no está en que surjan, ni tampoco en utilizarlas. El problema está en el abuso que se puede producir sobre ellas, en la mala utilización, y, en suma, en la falta de cuidado cuando se aproximan a la franja de los derechos humanos.

Desde luego, y aunque la IA surgió y se desarrolló fundamentalmente en el sector privado, coincido con Puentes Cociña y Quintiá Pastrana cuando afirman que “la transformación digital puede coadyuvar al cumplimiento de los principios de actuación del sector público, aumentando la transparencia, la imparcialidad, la objetividad, la eficiencia y la eficacia en la prestación de servicios” (Puentes Cociña, y Quintiá Pastrana, 2019, p. 20).

Pero todo esto no se puede hacer ignorando los riesgos, o dejándolos pasar confiando en que no produzcan daños excesivos. Los daños son muchos y además afectan a la parte más sensible del ordenamiento jurídico, esto es, a la protección de los derechos humanos.

En realidad, y junto al derecho a la salvaguarda de la privacidad de datos que parece que es el más evidente, podemos señalar otros derechos que se ven o pueden verse al menos, seriamente alterados por el mal manejo de la IA.

Pensemos, por ejemplo en la necesidad de “garantizar el respeto al principio de igualdad bien sea por restricciones de acceso para ciertos colectivos sociales o por la aparición de sesgos o discriminaciones en el manejo de

la información por parte de sistemas inteligentes; tensiones entre la falta de iniciativa pública y el alto grado de desarrollo de estas tecnologías en el sector privado; o, incluso, dificultad para determinar la responsabilidad patrimonial de la administración en algunos supuestos de uso de organismos autónomos o de decisiones tomadas por algoritmos” (Puentes Cociña, y Quintiá Pastrana, 2019, p. 21).

Todo ello sin mencionar todos los problemas que surgen en cuanto a qué información puede ser o no almacenada, utilizada o cedida sin consentimiento de las personas.

Capítulo aparte merece lo que se refiere a la posible vulneración de los derechos derivados de la personalidad. Pensemos por ejemplo en los modernos sistemas de reconocimiento de la voz, en el testamento digital, en el comercio on line y en todas sus derivaciones, o incluso en el control de los reos usando dispositivos electrónicos con control telemático.

Los problemas abarcan todo el ordenamiento y no sólo una de sus ramas. Por eso, “por una parte urge regular las incidencias de las tecnologías emergentes en la sociedad y en las instituciones públicas y sus relaciones con los ciudadanos... Y, por otra parte (atender) a las repercusiones de los avances tecnológicos en los derechos fundamentales (que) no están plenamente evaluadas. La innovación tecnológica puede afectar a los derechos fundamentales relativos a la intimidad, la privacidad, los datos personales, la seguridad incluida la jurídica, y a su vez a la libertad y a la igualdad” (Canals Ametller, 2019, p. 35).

Todo este trabajo está pendiente, pero mientras no se hace, conviene unir todos estos cabos y algunos más que quedan sueltos, para poder ver los riesgos jurídicos que puede entrañar el uso de estas tecnologías jurídicas.

Entre los más importantes señalaremos:

- 1) Aparición de sesgos y heurísticas derivados de la aplicación de la IA que podrían provocar discriminaciones en ciertos colectivos.
- 2) Dificultades para motivar las decisiones administrativas como consecuencia de la opacidad o falta de transparencia algorítmica.
- 3) Fragilidad de los derechos sociales para ejercer un control efectivo sobre las vulneraciones provocadas por las plataformas digitales.
- 4) Dificultades para determinar la responsabilidad por los actos imputables a órganos autónomos y
- 5) Amenaza para la privacidad y la protección de datos personales (Puentes Cociña y Quintiá Pastrana, 2019, p. 23).

Ante el surgimiento de todos estos problemas nos preguntamos ¿es de verdad la IA un avance? ¿Merece la pena conocerla y utilizarla? ¿Su utilización presenta más ventajas o *más* inconvenientes?

La respuesta a estas preguntas y a muchas más relacionadas con el tema resulta, a mi juicio, evidente. Sí, es necesario, útil, y hasta imprescindible diría yo, el uso de la IA en el ámbito jurídico. La experiencia jurídica avanza y con ella todo lo referente a la utilización e implementación de nuevas tecnologías. No tomar este tren sería fallar a la historia y apostar por el inmovilismo, y hasta me atrevería a decir por el involucionismo. En el ámbito del avance de las técnicas sociales, no aceptar la evolución y no adaptarse a ella no supone quedarse quieto, sino asumir el retroceso, y eso es algo que iría en contra de los derechos y, sobre todo, del derecho a la justicia que es, lo hemos dicho, el más importante de todos ellos.

Por eso a la pregunta de si es necesario el uso de la IA, nuestra respuesta es que no solo es necesario, sino que es imprescindible. Pero inmediatamente después de dar esta respuesta debemos matizar, que solo será conducente el uso de la IA si la usamos como una técnica, es decir si la dominamos en vez de permitir que ella nos domine.

Pues coincido en que “aunque no hay duda de que la IA puede ayudar a hacer eficaz el derecho a una buena administración de los ciudadanos, mediante la personalización, y mediante la prevención de la mala administración y de la corrupción administrativa, tampoco hay dudas en cuanto al peligro que, a la vez, y paradójicamente, la IA puede presentar para la buena administración” (Ponce Sole, 2019, p. 57).

Es así, porque la IA es una técnica que funciona con algoritmos y éstos, “tienen el potencial de promover la eficacia, la eficiencia, la consistencia y la justicia, pero también pueden reforzar discriminaciones históricas que oscurecen comportamientos indeseables” (Ponce Sole, 2019, p. 57).

Por eso hay que actuar con precaución, tomando de la IA todo aquello que ésta nos ofrece para mejorar la vida de la sociedad, y rechazando, por el contrario, lo que entorpece el adecuado desarrollo de los derechos humanos.

Y es que, como afirma López Oneto, “en el contexto transicional convulso y caótico, la irrupción de la IA en la vida cotidiana de la humanidad primero ha tensionado y luego modificado fragmentariamente los sistemas jurídicos del mundo, diseñados históricamente para regular la convivencia intersubjetiva en un mundo construido con tecnologías analógicas” (López Oneto, 2020, p. 231).

En este estado de cosas, sobre todo por las consecuencias negativas que produce, es necesario proteger las bases fundamentales que rigen el Estado de Derecho, entendiendo que éste proporciona la cobertura de una vida en la que se pueden desarrollar los derechos y cumplir con los deberes en paz, justicia y orden. Esa es la finalidad de la sociedad y no debe ser olvidada metiéndonos en una especie de carrera que muchas veces conduce a nin-

guna parte y en la que, lastimosamente, nos olvidamos de lo que tenemos y de lo que sacrificamos.

Es así porque “la cultura digital de masas ha barrido derechos fundamentales como el derecho a la intimidad, a la privacidad, al honor, y un largo etc., y poco a poco está cercenando también el derecho a la libertad, alimentándose desde una perspectiva difícil de rebatir, pero no imposible, de que la IA, el mundo digital, mejora al ser humano, implica progreso y avances que sólo pueden redundar en el ser humano” (Barona Vilar, 2021, p. 19).

El riesgo es real y muy peligroso. Lo es porque la IA se nos presenta como un rasgo de modernidad que sólo puede traer beneficios, de manera tal que rechazarlo es convertirse casi en un paria de la sociedad. En una persona que se niega a crecer y evolucionar con los tiempos y que, por lo mismo, irremisiblemente está condenada al fracaso.

La defensa de los derechos humanos pasa a ser así, casi sin darnos cuenta²³, una necesidad menos que pierde brillo frente a la omnipotente IA, a la que todos debemos “adorar” sin mirar atrás, ya que no hacerlo sería de tontos.

Lejos de coincidir con esta idea creo que hoy más que nunca es preciso “aplicar el principio de precaución en el contexto de la IA (pues ello permitiría equilibrar los términos de la discusión en el ámbito político global), otorgando un mayor peso a quienes están en las posiciones más débiles en los debates, que actualmente están monopolizados por los intereses de las empresas privadas” (Velasco Rico, 2019, p. 85).

El equilibrio resulta fundamental y suele ser una buena receta en todo caso, pero mucho más cuando hablamos de derechos fundamentales y de Estado de Derecho, y en una situación que quizá sea un poco adversa, habida cuenta de la que podemos llamar “carrera casi sin retorno” en la que nos han metido los avances de la IA. Por eso ahora más que nunca, suenan muy reconfortantes las palabras de Truyol Serra, cuando recordaba que aún en estos momentos de evolución acelerada, es preciso “respetar la necesidad imperiosa de hallar normas para salvaguardar los derechos fundamentales de la persona, y en particular el derecho a la intimidad” (Truyol Serra, s/a, p. 120).

Coincido con Lagares cuando afirma que “internet se ha convertido en una esfera a la que todos tenemos un acceso tan libre y falto de esquemas

23. Al final como dice Crespo García, “todos somos objeto de investigación exhaustiva y detallada en relación con nuestros movimientos, opiniones, deseos e intereses. Esta intromisión indiscriminada en nuestra esfera privada, supone una vulneración, en la mayor parte de las ocasiones de nuestra intimidad” (Crespo García, 2016, p. 16).

correctores que comienza a cuestionarse la conveniencia de una tan grande villa sin vallado” (Lagares, 2000, p. 12).

Esta realidad nos sume en un nuevo problema que me gustaría analizar siquiera fuera brevemente. Se refiere a la ética en las relaciones seres humanos e IA.

5. ÉTICA E INTELIGENCIA ARTIFICIAL

Como hemos indicado, la IA puede, y de hecho quizá ya ha llegado, a convertirse en una técnica de trabajo peligrosamente parecida a la propia mente humana en su funcionamiento. Como afirma Boden, “los sistemas de IA ya pueden conocer emociones humanas de varias formas, por ejemplo, fisiológicas; observando la frecuencia respiratoria y la respuesta galvánica de la piel de la persona; verbales, observando la velocidad de habla y la entonación, así como el vocabulario; y visuales, analizando las expresiones del rostro” (Boden, 2017, p. 78).

Estas realidades, muy modernas e interesantes para algunos, para otros, quizá la mayoría, se convierten en inquietantes pues les ponen delante seres que son muy parecidos a ellos, pero no idénticos. Seres que en algunos casos pueden llegar incluso a resolver problemas mejor que los propios humanos, pero que carecen de sentimientos, empatía, etc.

Así las cosas, nos formulamos esta pregunta ¿podría la IA aprender a ser moral? ¿Puede llegar a enseñarse lo bueno o lo malo a través de algoritmos?

¿Es tan prometedora la IA como parece o como algunos nos quieren hacer creer? Aquí la doctrina duda. O más que dudar divide sus respuestas en dos grandes bloques.

Algunos apuestan por la IA afirmando que en realidad tiene más posibilidades que la propia inteligencia humana, ya que es más objetiva, puede almacenar más datos y por lo mismo sus posibilidades de respuesta, no solo son mucho mayores, sino mucho más objetivas, y por lo mismo más “inteligentes”. Tiene en realidad, muchas menores posibilidades de producir error o fallo. Se trata únicamente de meter en las bases de datos cuantas más posibilidades de respuesta mejor, a fin de que la máquina artificial con su IA pueda aportar en el menor lapso la respuesta adecuada. En ese ámbito de cosas la IA sería o podría llegar a ser superior a la humana²⁴.

24. De esta opinión tan favorable a la IA es el Dr. Senén Barro Ameneiro, ex rector de la Universidad de Santiago de Compostela y actualmente catedrático de la Facultad de Física de la

Otros autores como Boden afirman, por el contrario, que “la IA es menos prometedora de lo que mucha gente cree²⁵... hay infinidad de cosas que la IA no puede hacer... además, la IA se ha centrado en la racionalidad intelectual y ha ignorado la inteligencia social/emocional y ni hablamos de la sabiduría” (Boden, 2017, p. 151).

Los que así opinan entienden que la IA puede hacer mucho, pero no todo lo que puede hacer la inteligencia humana porque ésta dispone de un elemento emotivo, sensorial, capaz de repentizar y de sobreponerse a las dificultades, que hace que sea superior siempre a cualquier máquina.

Pero no es la supuesta superioridad el problema ético más importante que plantea la IA en su desarrollo conceptual.

Su generalización alteraría mucho el mercado de trabajo dejando sin ocupación a muchas personas que serían sustituidas por máquinas. Además, y “al margen de los efectos que tendría en el desempleo, el uso de la IA sin empatía en contextos tan esencialmente humanos es muy arriesgado en la práctica, y éticamente dudoso” (Boden, 2017, p. 157).

En efecto, los robots que no hace mucho nos parecían de ciencia ficción, amenazan con convertirse en parte integrante de nuestras vidas, y si ello llegase a suceder no se trataría únicamente de unas máquinas destinadas a hacernos la vida más fácil, a proporcionarnos compañía, o a trabajar por o para nosotros.

Los robots, como en nuestras peores pesadillas, cobrarán casi vida propia y se meterán en nuestra privacidad, y en nuestros derechos, tomando decisiones por nosotros y apoderándose de nuestras mentes. Todo ello por no hablar de las implicaciones negativas de usar la IA en contextos moralmente comprometidos.

Quizá la visión apocalíptica de la IA sea exagerada, pero lo cierto es que en la actualidad existen ya problemas reales de una toma de poder de las máquinas sobre las personas, que deben ser previstos, a fin de poder ser evitados.

Coincido con Sadín cuando advierte de que “la humanidad se está dotando a grandes pasos de un órgano de prescindencia de ella misma, de

misma Universidad. En su discurso de toma de posesión como académico de la Academia de Ciencias desarrolló esta idea.

25. En realidad, opina Boden, “considerada como vía para comprender la inteligencia humana, es probable que la WBE (Whole brain emulation) ascendente fracase. Nos puede enseñar mucho sobre el cerebro y puede llegar a que los científicos de la IA desarrollen más aplicaciones prácticas, pero la idea de que a mitad del siglo la WBE habrá explicado la inteligencia, parece excesiva.

su derecho a decidir con plena conciencia y responsabilidad las elecciones que le involucran” (Sadín, 2020, p. 21).

No importa que la IA ofrezca más soluciones, en un período más corto de tiempo, y más seguras, e incluso objetivas. No importa que la IA sea más rápida, que no se deje sorprender por emoción alguna, y que no tenga problemas de conciencia.

No importa, porque el ser humano es y debe seguir siendo el rey de la creación y, por lo mismo, debe dominar todas las cosas creadas, y no dejarse dominar por alguna de ellas.

Es necesario “defender nuestra facultad de juicio. Arendt la consideraba como la cuestión política principal en la medida en la que determinaba la posibilidad de acciones individuales y colectivas que niegan a toda normatividad infundada el juego de poderes ilegítimos” (Sadín, 2020, p. 43).

Por eso “depende de nuestra responsabilidad frente a nuestra herencia humana utilizar nuestro derecho a bloquear, donde sea que estemos, los mecanismos que trabajan para imponer, en todas las escalas de nuestra existencia, un orden unilateral e infundado de las cosas. A eso lo podemos denominar una ética en acto de nuestras convicciones o una saludable puesta en práctica de legítima defensa” (Sadín, 2020, p. 124).

La IA es un mecanismo muy útil que no debe ser desaprovechado. Debe ser entendido como un instrumento, como una técnica al servicio de las personas. No debemos permitir otro uso ni olvidar el lugar que cada cosa ocupa dentro de nuestro orden existencial.

El ser humano está primero, después las máquinas que le ayudan en su vida. Es posible, más que posible, es una realidad, que las máquinas hayan alcanzado un nivel de desarrollo muy elevado, pero es tarea de los seres humanos regular el proceso hasta donde su utilización pueda dejar de ser una necesidad para convertirse en un peligro.

Es preciso averiguar si es posible que los ordenadores y las demás máquinas que usan IA puedan llegar a tener inteligencia, conciencia y creatividad, pues de la respuesta que se dé a estos interrogantes dependerán, en gran medida, las acciones a implementar para controlar su adecuado uso.

La respuesta que se puede dar a estos interrogantes no es pacífica. En realidad “hay distintas posturas filosóficas en torno a estas cuestiones. Algunos dicen que las máquinas jamás podrán ser agentes morales, las máquinas argumentan, no tienen las capacidades necesarias para la agencia moral, tales como estados mentales, emociones, o libre albedrío. De ahí que sea peligroso suponer que puedan tomar decisiones morales correctas o delegar en ellas totalmente decisiones” (Coeckelberg, 2021, p. 52).

Los que así opinan entienden incluso que el debate, o la preocupación que se ha generado en torno a la posible utilización de la IA hasta el punto

de que sus técnicas puedan llegar a dominar a la propia mente humana, se exceden en sus preocupaciones imaginando un mundo futuro más propio de películas que de la realidad.

En el otro extremo están “quienes piensan que las máquinas pueden ser agentes morales completos de la misma forma en que los seres humanos” (Coeckelberg, 2021, p. 52).

Quienes defienden esta posición piensan que las máquinas pueden incluso llegar a ser mejores que los propios seres humanos pues sus razonamientos son más completos, más objetivos y se dejan influir menos por las emociones.

La polémica está servida. Hay argumentos a favor y en contra de casi cualquier respuesta, y desde luego en relación con las dos opciones generales que hemos expuesto.

Llegados a este punto no conviene a mi juicio, seguir moviéndonos en círculo. Es preciso defender alguna posición.

Yo creo que la IA no es algo que haya existido por generación espontánea, sino que es fruto de la mente humana. Sus aplicaciones hoy en día son muy importantes y afectan en gran medida al desarrollo de los derechos humanos. La recogida de información personal, la vigilancia de nuestras preferencias, la posible influencia en la toma de decisiones, son problemas que se deben evitar. Pero la ayuda que nos presta, la flexibilización del trabajo, los avances en la tecnología, medicina, ciencia en general, son, por el contrario, ventajas que se deben utilizar.

Hoy en día nos hemos vuelto en gran medida dependientes de la IA de manera que ésta ha dejado de ser una simple herramienta de trabajo para convertirse en parte de nuestra propia naturaleza. Este es a mi juicio el principal asunto que suscita un problema ético. Me refiero al de la responsabilidad. ¿Se puede atribuir responsabilidad moral a una máquina? ¿Se puede hacer basándonos simplemente en los resultados que acredita y en las posibilidades que maneja? ¿Es consciente una máquina de los pros y de los contras de las decisiones que toma, o decide únicamente en función de posibilidades algorítmicas?

Mi respuesta a estas preguntas es que las máquinas, aun las más perfectas, son máquinas y, por lo mismo, no son sujetos de deberes ni de derechos y no pueden ejercitar la responsabilidad.

Pero la respuesta, aunque parece fácil no lo es en realidad porque “los debates sobre la ética en la IA tocan de lleno delicados problemas políticos y sociales, que, a su vez, están ligados a problemas filosóficos de índole normativa, como los relativos a la justicia y la equidad, así como a problemas filosóficos y científicos relacionados con los seres humanos y sus sociedades. Uno de estos problemas es el futuro del trabajo” (Coeckelberg, 2021, p. 115).

Por lo que a los problemas jurídicos se refiere la situación dialéctica del uso de la IA es similar. “Los derechos a la vida, la libertad y la seguridad de las personas se ven afectados positivamente si son los sistemas inteligentes los que moderan los contenidos en la red. Pueden tener un impacto positivo y negativo en cuestiones de discriminación”. Y “los impactos negativos se ven en los derechos a la privacidad... y en los derechos de libertad de opinión, expresión e información” (Valls Prieto, 2021, p. 75).

Sería bueno tratar de potenciar los efectos positivos del uso adecuado de la IA²⁶ y, por el contrario, intentar minimizar los negativos. Debe hacerse así, pues “el impacto de la IA en el Estado de Derecho puede ser positivo y negativo siempre que se refleje en principios como la justificación, la proporcionalidad y la igualdad, puede mejorar la eficacia de las instituciones y la confianza en el sistema y la legitimidad de las instituciones democráticas” (Valls Prieto, 2021, p. 79).

Las posibilidades son muchas, de modo que no hay que aparcarse los problemas de IA entendiendo que son en sí mismos simples. No lo son. Se trata de un instrumento cuya utilización plantea en el presente problemas reales que se debaten entre la ética de su adecuada utilización y la necesidad de utilizarla.

Pero a mi juicio y coincido en ello plenamente con Coeckelberg, “la IA es buena a la hora de reconocer patrones, pero la sabiduría no se puede delegar en las máquinas” (Coeckelberg, 2021, p. 164).

REFERENCIAS BIBLIOGRÁFICAS

- Barona Vilar, S. (2021). *Algoritmización del derecho y la justicia. De la IA a la Smart justice*. Valencia: Tirant lo Blanch.
- Boden, M. A. (1994). Escape de la habitación china, en Boden, M. A. (compiladora), *Filosofía de la IA*. México: Fondo de Cultura Económica.
- Boden, M. A. (2017). *Inteligencia artificial*. Madrid: Editorial Turner.
- Bourcier, D. (2003). *Inteligencia artificial y derecho*, Barcelona: Editorial UOC.
- Bueno de Mata, F. (coord.) (2016). *Fodertics, estudio sobre nuevas tecnologías y justicia, 5.0*. Granada: Comares.
- Canals Ametller, D. (2019). Incidencias del avance tecnológico en el Derecho Público, en Puentes Cociña, B y Quintiá Pastrana, A. (directores). *El derecho ante la transformación digital*. Barcelona: Atelier.

26. En este sentido serían sin duda de gran ayuda las Directrices éticas para una IA fiable publicadas en <https://op.europa.eu/es/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1>

- Carabantes López, M. (2016). *Inteligencia artificial. Una perspectiva filosófica*. Madrid: Escolar y Mayo editores.
- Churchland, P. M. (1994). Algunas estrategias reductivas en la neurobiología cognoscitiva, en Boden, M. A. (compiladora). *Filosofía de la IA*. México: Fondo de Cultura Económica.
- Clark, A. (1994), El conexionismo, su competencia y su explicación, en Boden, M. A. (compiladora). *Filosofía de la IA*. México: Fondo de Cultura Económica.
- Coeckelberg, M. (2021). *Ética de la Inteligencia Artificial*. Madrid: Cátedra.
- Copeland, J. (1993). *Inteligencia artificial*. Madrid: Alianza Universidad.
- Crespo García, P. (2016). El impacto del Big Data en los derechos fundamentales de las personas, en Bueno Mata, F. (coord.) *Fodertics, estudios sobre nuevas tecnologías y justicia 5.0*. Granada: Comares.
- Dennett, D. C. (1994). Las ruedas de conocimiento: el problema estructural de la IA, en Boden, M. A. (compiladora). *Filosofía de la IA*. México: Fondo de Cultura Económica.
- Dreyfus, H. L. & Dreyfus, S. E. (1994). La construcción de una mente versus el modelaje del cerebro: la IA regresa a un punto de ramificación, en Boden, M. A. (compiladora). *Filosofía de la IA*. México: Fondo de Cultura Económica.
- González Quevedo, M. A. y otros (1999). *Técnicas de inteligencia artificial*. Salamanca: Secretariado de la Universidad de Alicante.
- González Ruiz, E. (1992a). Algunas tendencias actuales en filosofía de la inteligencia artificial, en Negrete J., *De la filosofía a la Inteligencia Artificial*. México: Grupo Noriega editores.
- González Ruiz, E. (1992b). Sobre Dreyfus y Dreyfus: mente sobre máquinas, en Negrete, J. *De la Filosofía a la Inteligencia artificial*. México: Grupo Noriega editores.
- Hintos, G. E., McClelland, J. L. & Rumelhart, D. E. (1994). Las representaciones distribuidas, en Boden, M. A. (compiladora). *Filosofía de la IA*. México: Fondo de Cultura Económica.
- Lagares, D. (2000), *Internet y el Derecho*, Barcelona: Ediciones Carena.
- López Oneto, M. (2020). *Fundamentos para un derecho de la inteligencia artificial ¿Queremos seguir siendo humanos?* Valencia: Tirant lo Blanch.
- Marr, D. (1994). La inteligencia artificial: un punto de vista personal, en Boden, M. A. (compiladora). *Filosofía de la Inteligencia Artificial*. México: Fondo de Cultura Económica.
- Negrete, J. y otros (1992). La vulneración del dictum de la no-inteligencia de las máquinas, en Negrete J. *De la Filosofía a la inteligencia artificial*. México: Noriega editores.
- Newell, A. y Simon, H.A., (1994). La ciencia y la computación como investigación empírica: símbolos y búsqueda, en Boden, M. A. (compiladora). *Filosofía de la Inteligencia Artificial*. México: Fondo de Cultura Económica.
- Nilson, N. I. (2000). *Inteligencia Artificial. Una nueva síntesis*. Madrid: McGraw Hill.

- Ponce Sole, J. (2019). El derecho a una buena administración y la personalización de los servicios públicos. Sesgos nudging e IA, en Puentes Cociña, B & Quintiá Pastrana, A. (directores). *El derecho ante la transformación digital*. Barcelona: Atelier.
- Puentes Cociña, B. y Quintiá Pastrana, A. (directores) (2019). *El derecho ante la transformación digital*. Barcelona: Atelier.
- Sadín, E. (2020). *La IA o el desafío del siglo. Anatomía de un anti-humanismo radical*. Buenos Aires: Caja Negra Editora.
- Searle, J. R. (1994). Cerebros y programas, en Boden, M. A. (compiladora). *Filosofía de la inteligencia artificial*. México: Fondo de Cultura Económica.
- Slovan, A. (1994). Los motivos, los mecanismos y las emociones, en Boden M. A. (compiladora). *Filosofía de la IA*. México: Fondo de Cultura Económica.
- Trillas, E. (1998). *La inteligencia artificial. Máquinas y personas*. Madrid: Editorial Debate.
- Truyol Serra, A. (s/f). Bases filosóficas y metodológicas para un derecho de la sociedad de la información, en VV. AA., *Implicaciones socio-jurídicas de las tecnologías de la información*, Madrid: CITEMA, Centro de informática, telemática y medios afines.
- Turing, A. M. (1994). La maquinaria de computación y la inteligencia, en Boden, M. A. (compiladora). *Filosofía de la inteligencia artificial*. México: Fondo de Cultura Económica.
- Velasco Rico, C. I. (2019). Vigilando al algoritmo. Propuestas organizativas para garantizar la transparencia, en Puentes Cocina, B. & Quintiá Pastrana, A. (directores). *El derecho ante la transparencia digital*. Barcelona: Atelier.
- Valls Prieto, J. (2021). *Inteligencia Artificial, derechos humanos y bienes jurídicos*. Cizur Menor: Thomson Reuters.
- VV.AA. (1996). *Ámbito jurídico de las tecnologías de la información*. Madrid: Consejo General del Poder Judicial.
- VV.AA. (2000). *Problemática jurídica en torno al fenómeno de internet*. Madrid: Consejo General del Poder Judicial.

