

COECKELBERGH, M. (2021). *Ética de la inteligencia artificial*. Madrid: Cátedra (183 pp.)

La admiración y el miedo, respuestas humanas muy comunes ante los avances tecnológicos y presentes en diferentes contextos culturales, residen en el corazón de los futuristas y los catastrofistas. Por ello, el ejercicio que demanda el análisis de estos comportamientos debe ponerse en manos de un constante cuestionamiento ético. Y es que no son tanto las ideas futuristas las que se escapan de esta conveniente reflexión, sino que son aquellas que se encuentran en los lugares más comunes, carentes de todo atisbo de duda, las que resultan más difíciles de cuestionar con detenimiento, mirada austera y amplitud de miras. De este modo, nos encontramos en dominios en donde se asumen ciertos presupuestos sin someterse a juicio, lo que resulta ser un terreno muy fructífero para la mirada inquisidora de quien pretende hacer filosofía. Y es este papel de vocación interrogadora el que desempeña Mark Coeckelbergh en las páginas de *Ética de la inteligencia artificial*, para conducirnos hacia la duda y constante interrogación, es decir, para enseñarnos «*que lo que pensamos los seres humanos tiene también varias caras*» (Coeckelbergh, 2021).

El presente ensayo dedica sus esfuerzos al propósito de presentar de forma completa, y sistemática, el estado de la cuestión de las problemáticas éticas que se derivan de las tecnologías de inteligencia artificial (IA) en diferentes disciplinas, integrando y agregando claridad al trabajo en este campo. La evaluación de los posibles avances en su desarrollo ha producido una cierta des-

confianza que ha llevado al autor a la necesidad de repensar y discutir los problemas éticos y sociales que suscitan estos monstruos del apocalipsis. El autor los define de este modo haciendo referencia al monstruo de Frankenstein, una figura que aporta una valiosa mirada acerca de la ambivalente posición de estas nuevas tecnologías.

Dos de los temas más recurrentes del ensayo y que producen mucha inquietud académica y social son la superinteligencia y el transhumanismo. Quizás, resuenen mucho en la actualidad, pero para sorpresa de muchos desde bien temprano el ser humano sueña con alcanzar el poder de un dios y, hacerse con atributos divinos tales como la inmortalidad. Esta aspiración se ha trasladado a la medicina actual en forma de medidas integrales contra el envejecimiento. Su origen no es otro que el miedo producido, por la fragilidad humana y su capacidad de cometer errores, lo que nos ha llevado a mejorar al ser humano hasta llegar a comprenderlo como un *Homo deus*: humanos que han ascendido a la categoría de dioses (Harari, 2016). Muchos piensan que la biología humana necesita mejorarse para no arriesgarnos a quedar como «la parte lenta y cada vez más ineficiente de la IA (Armstrong, 2014)».

Actualmente, tal como menciona Coeckelbergh, existen dos caminos hacia la superinteligencia. Uno consiste en que la IA se automejore continuamente, diseñando y mejorando versiones de sí misma, que, a su

vez, diseñarían versiones mejoradas exponencialmente: un acontecimiento llamado La Singularidad Tecnológica, que consiste, en el escenario en el que los seres humanos serán superados por máquinas inteligentes o inteligencias, o inteligencias biológicas cognitivamente mejoradas o ambas.

Si bien tales avances pueden resultar escalofrantes, por suerte para algunos o para desgracia de otros, los avances en neurociencia todavía nos sitúan lejos de estas posibles aplicaciones. Y, aunque pueda relajarnos, no significa que debamos permanecer exentos de preocupaciones. Existen numerosas aplicaciones de software inteligente que nos acompañan en nuestra vida cotidiana. Y son el origen de múltiples problemáticas que afectan a la forma en la que estamos y convivimos en el mundo. Una de las áreas que más gravemente se ha visto afectada recientemente y que hemos dejado pasar por alto, además de que se haya instaurado como modelo económico, han sido los datos, datos que forman parte de nuestra privacidad y que, como bien define Carissa Véliz (2022) «la realidad es tan colectiva como es personal». Y que como define Shoshana Zuboff (2021), el capitalismo tiene el poder de impulsar la difusión de la conexión y la vigilancia en línea, de modo que los espacios sociales que antes permanecían dentro de un círculo social cerrado, ahora se están abriendo a corporaciones dirigidas a la obtención de beneficios y o la regulación de conductas. Este último fenómeno conlleva el riesgo de que los usuarios puedan ser manipulados y explotados. La IA tiene múltiples aplicaciones y, puede ser utilizada para manipular nuestras necesidades y recoger datos muy personales acerca de nosotros, es decir, creencias religiosas, ideologías políticas u, orientación sexual o datos referentes a la salud. Recabar este tipo de datos, en muchos países se considera un

delito, ya que permanecen altamente protegidos jurídicamente. Y dentro del contexto capitalista, la herramienta más utilizada para recabar datos es a través del uso de las redes sociales. «Este tipo de tecnologías de aprendizaje automático y ciencias de datos puede conducir a nuevas formas de manipulación, vigilancia y totalitarismo, no necesariamente bajo la apariencia de regímenes autoritarios, sino de manera subrepticia y altamente efectiva» (Coeckelbergh, 2021, p.88). Las múltiples aplicaciones de la IA introducen numerosas problemáticas que implican graves injusticias, como es el caso de los sesgos algorítmicos, a la hora de conceder una ayuda financiera, calcular la factura de la luz, predecir la reincidencia de un preso, etc. Nuestro éxito para enfrentarnos con los grandes problemas de este periodo como predice Coeckelbergh, viene determinado por «una combinación de inteligencia abstracta (humana y artificial) y sabiduría pragmática concreta desarrollada sobre la base de experiencia humana situacional y práctica (incluyendo nuestra experiencia con la tecnología)» (2021).

Esta obra combina de forma satisfactoria un lenguaje accesible para el público general que pretenda conocer las inquietudes y problemáticas éticas que derivan de la IA con un tratamiento detallado de los diferentes debates que conforman este campo de investigación. En este sentido, profundiza de forma resumida en cada una de estas temáticas a través de doce capítulos breves, cada uno de los cuales se ocupa de desglosar minuciosamente distintas problemáticas. La presentación y exposición de las posiciones y argumentos es fluida y favorece una lectura ágil y clara. La forma en la que trata los temas que se discuten, y su pretensión de abordar reflexivamente una variedad de debates y discusiones de especial relevancia filosófica y social, dota al lector de una serie

de capacidades y herramientas para comprender de mejor forma los desafíos éticos planteados por la IA.

Por todo lo anterior, las páginas que conforman *Ética de la inteligencia artificial* se erigen, en definitiva, como ejemplo de cómo la ética y la ciencia poseen espacios de reflexión y pensamiento común y por qué es importante ahondar en ellos. Algo de agradecer habida cuenta de que necesitamos medios conceptuales para afrontar los nuevos retos de las sociedades tecnológicas.

Referencias

Armstrong, S. (2014). *Smarter Than Us: The Rise of Machine Intelligence*. Machine Intelligence Research Institute.

Coeckelbergh, M. (2021). *Ética de la inteligencia artificial*. Ediciones Cátedra.

Harari, Y. N. (2016). *Homo Deus: A Brief History of Tomorrow*. Random House.

Véliz, C. (2022). *Privacidad es poder: Datos, vigilancia y libertad en la era digital*. Penguin Random House Grupo Editorial.

Zuboff, S. (2021). *La era del capitalismo de la vigilancia: La lucha por un futuro humano frente a las nuevas fronteras del poder*. Ediciones Culturales Paidós.

Jorge Couceiro Monteagudo
(Universidad Complutense de Madrid. jorgecouceiromonteagudo@gmail.com)

GONZÁLEZ-ESTEBAN, ELSA y SIURANA, JUAN CARLOS (eds.) (2023). *Inteligencia Artificial: concepto, alcance y retos*. Valencia: Tirant Blanch.

El libro *Inteligencia Artificial: Concepto, Alcance, Retos*, editado por Elsa González y Juan Carlos Siurana y dividido en tres grandes bloques, es un trabajo focalizado en la creciente penetración de la Inteligencia Artificial (IA) en todos los aspectos de la vida humana y sus impactos sobre la sociedad *hiperdigitalizada e hiperconectada* y sus diferentes esferas funcionales, como la economía, la política, la comunicación o la investigación científica, entre otras.

Entre las múltiples aportaciones del libro, destaca su propuesta de un diseño institucional y una infraestructura ética que permita orientar la gestión y toma de decisiones en las organizaciones en un sentido responsable, justo y equitativo y garantizar

la acción y supervisión humanas, entendidas éstas desde la ética del discurso como participación y búsqueda deliberativa de acuerdos por parte de todos los afectados.

El **Bloque I. Inteligencia Artificial: concepto** consta de cuatro capítulos escritos por Juan Arana, Jesús Conill, Dieter Sturma y Pedro Jesús Teruel respectivamente. Este bloque explora diferentes aspectos del concepto de la IA y su relación con la mente humana y la tecnología. En él se discuten los enfoques naturalistas de la mente humana, se analizan las críticas de los neurocientíficos al paradigma computacional y se exploran las posibilidades de la IA en diversos campos.

El primer capítulo, “Algunas críticas de los neurocientíficos al paradigma computa-