



HUMANIDADES DIGITALES: UNA REFLEXIÓN FILOSÓFICA

El estado de la cuestión: F. ARENAS DOLZ. **Reflexión y crítica:** A.L. TERRONES RODRÍGUEZ, J.M. MARTÍNEZ CASTELLÓ. **Ágora:** I. BELTRÁ VILLASEÑOR. **Didáctica:** F. GARCÍA MORIYÓN. **Informaciones.**

Diálogo Filosófico

**Revista cuatrimestral de reflexión, crítica e información
filosóficas editada por Diálogo Filosófico®.**

Diálogo Filosófico articula su contenido en artículos solicitados en torno a un tema o problema filosófico de actualidad en las secciones «Estado de la cuestión» y «Reflexión y crítica». Además, publica siempre artículos no solicitados en la sección «Agora» (filosofía en general) y ocasionalmente en la sección «Didáctica» (relacionada con la enseñanza de la filosofía y la filosofía de la educación). Privilegia los de contenido no meramente histórico y expositivo, sino que reflexionan de manera original sobre los problemas reales o dialogan creativamente con los pensadores y las corrientes filosóficas presentes y pasadas. Dichos artículos pasan por un proceso de evaluación ciega por pares. Asimismo, acepta el envío de reseñas que recojan una confrontación crítica con libros de reciente publicación.

Director: Antonio Jesús María Sánchez Orantos (Universidad Pontificia Comillas).

COMITÉ DE DIRECCIÓN

Juan Jesús Gutierrez Carrasco (Universidad Católica de Ávila. ESCUNI Centro Universitario de Educación), Alberto Lavín Fernández (IE University), Mario Ramos Vera (Universidad Pontificia Comillas).

COMITÉ CIENTÍFICO

Vittorio Possenti (Università degli Studi di Venezia), Erwin Schadel (Otto-Friedrich Universität Bamberg), Mauricio Beuchot (Universidad Nacional Autónoma de México), Adela Cortina (Universidad de Valencia), Jean Grondin (University of Montreal), Charles Taylor (McGill University), João J. Vila-Chã (Universidade Católica Portuguesa), Miguel García-Baró (Universidad Pontificia Comillas), Peter Colosi (The Council for Research in Values and Philosophy).

CONSEJO DE REDACCIÓN

José Luis Caballero Bono (Universidad Pontificia de Salamanca), Ildelfonso Murillo (Universidad Pontificia de Salamanca), José M.^a Vegas Mollá (Seminario Diocesano de San Petersburgo), Ignacio Verdú (Universidad Pontificia Comillas), Jesús Conill (Universidad de Valencia), Camino Cañón Loyes (Universidad Pontificia Comillas), Jorge M. Ayala (Universidad de Zaragoza), Félix García Moriyón (Universidad Autónoma de Madrid), Juan Antonio Nicolás (Universidad de Granada), Juan J. García Norro (Universidad Complutense de Madrid), Agustín Domingo Moratalla (Universidad de Valencia), Manuel Sánchez del Bosque, Leonardo Rodríguez Duplá (Universidad Complutense de Madrid).

Administración:

M.^a Jesús Ferrero

Dirección y Administración DIÁLOGO FILOSÓFICO
Corredera, 1 - Apartado de Correos 121 - 28770 COLMENAR VIEJO (Madrid)
Teléfono: 610 70 74 73
Información Electrónica: dialfilo@hotmail.com
www.dialogofilosofico.com

Esta revista está indexada en LATINDEX, RESH, CARHUS+,
ISOC, DICE, MIAR, FRANCIS, PASCAL, CIRC, DULCINEA,
The Philosopher's Index, Répertoire Bibliographique de la Philosophie,
International Directory of Philosophy.

Edita:

DIÁLOGO FILOSÓFICO / PUBLICACIONES CLARETIANAS

PRECIOS SUSCRIPCIÓN EN PAPEL (2023)

Número suelto: 16 euros (IVA incluido)

Suscripción anual: España: 34 euros (IVA incluido)
/ Extranjero: 42 euros (correo normal)

EN PORTADA: Robot (imagen tomada de Pixabay)

I.S.S.N.: 0213-1196 / Depósito Legal: M.259-1985

Diálogo Filosófico

Año 39

Enero/Abril

I/23

Presentación..... 3

El estado de la cuestión

ARENAS DOLZ, F.: *Las humanidades en la era digital. De las humanidades digitales al humanismo tecnológico*..... 4

Reflexión y crítica

TERRONES RODRÍGUEZ, A.L.: *Sabiduría práctica e innovación de la inteligencia artificial*..... 31

MARTÍNEZ CASTELLÓ, J.M.: *Justicia juvenil y digitalización de los aprendizajes*... 47

Ágora

BELTRÁ VILLASEÑOR, I.: «Atenas y Jerusalén» *La necesidad intrínseca del encuentro entre fe bíblica y filosofía en el pensamiento de Joseph Ratzinger*..... 65

Didáctica

GARCÍA MORIYÓN, F.: *Los dilemas morales en la educación*..... 87

Informaciones

Acontecimientos	105
Efemérides Filosóficas 2023	
V Simposio Internacional de Sofic –Interioridad Humana	
Crítica de libros	109
GORDILLO ÁLVAREZ-VALDÉS, Lourdes / HURTADO CONTRERAS, Ángel	
Luis: <i>Aprendiendo a ser libres</i> (Ildefonso Murillo).	
Noticias de libros	113

Reflexión y crítica

Sabiduría práctica e innovación de la inteligencia artificial

Practical wisdom and innovation of artificial intelligence

Antonio Luis Terrones Rodríguez

Es, por tanto, manifiesto que, aunque no estuviera orientada a la acción, tendríamos menester de ella por el hecho de ser virtud de una parte del alma; y porque no habrá recta elección sin el concurso de la Prudencia y de la virtud: la una nos proporciona el fin, la otra realizar los actos conducentes al fin¹.

Resumen

La presencia de la inteligencia artificial se ha incrementado en los últimos años como consecuencia del soporte que ofrecen para la realización de actividades que destacan por un alto nivel de complejidad. Este incremento ha ocasionado el surgimiento de numerosos desafíos éticos que advierten la necesidad de configurar una investigación e innovación responsable de esta tecnología emergente. En este sentido, la virtud aristotélica de la *phronēsis* constituye un legado sapiencial que favorece el cultivo y la asunción de responsabilidad en el ámbito de los intelectos sintéticos como una alternativa a la racionalidad estrictamente técnica.

Abstract

The presence of artificial intelligence has increased in recent years as a result of the support they offer for carrying out activities that stand out for their high level of complexity. This increase has caused the emergence of numerous ethical challenges that warn of the need to set up responsible research and innovation of this emerging technology. In this sense, the Aristotelian virtue of *phronēsis* constitutes a sapiential legacy that favors the cultivation and assumption of responsibility in the realm of synthetic intellects as an alternative to strictly technical rationality.

¹ ARISTÓTELES: *Ética a Nicómaco*. Alianza, Madrid, 2005, pp. 200-201.

Palabras clave: Inteligencia artificial, ética, investigación e innovación responsable, *phronēsis*.

Keywords: Artificial intelligence, Ethics, Responsible research and innovation, *Phronēsis*.

1. Introducción

Al margen de los círculos de estudios dedicados a la figura de Aristóteles, hace unas cuantas décadas, el concepto de *phronēsis* del estagirita no despertaba el mismo interés que en la actualidad. Sin embargo, el reciente interés en este concepto clásico se ha originado debido a un incremento de la importancia reconocida a la ética de las virtudes. Este incremento coincide con el impulso de una tecnología emergente que ha consolidado su presencia en diversas esferas y actividades de la vida humana. La inteligencia artificial (IA) se ha convertido en uno de los pilares principales sobre los que pivota la cuarta revolución industrial y tras su desarrollo subyacen comprometedores desafíos morales que exigen ser abordados desde la óptica de una investigación e innovación responsable. En ese sentido, el estímulo para el planteamiento de este trabajo se encuentra en la creciente preocupación surgida a tenor de los valores instrumentales promovidos por un «tecnopensar» que condiciona el desarrollo de la IA de un modo preponderante. Para esta modalidad de pensamiento, la técnica es una suerte de antídoto frente a todos los problemas. Sin embargo, el surgimiento de numerosos desafíos exige una comprensión enraizada en una sapiencia que trascienda la simplicidad inspirada en la locura del solucionismo tecnológico², como señala el Papa Francisco:

«Buscar solo un remedio técnico a cada problema ambiental que surja es aislar cosas que en la realidad están entrelazadas y esconder los verdaderos y más profundos problemas del sistema mundial. Las soluciones meramente técnicas corren el riesgo de atender a síntomas que no responden a las problemáticas más profundas. Cualquier solución técnica que pretendan aportar las ciencias será impotente para resolver los graves problemas del mundo si la humanidad pierde su rumbo.»³

² Cf. MOROZOV, Evgeny: *La locura del solucionismo tecnológico*. Clave Intelectual, Madrid, 2015.

³ FRANCISCO, Papa: *Laudato si'*. Carta encíclica del 24 de mayo de 2015, pp. 88-153.

Así pues, en las siguientes páginas se reflexionará sobre el valor que representa la sabiduría práctica promovida por la virtud dianoética de la *phronēsis*, considerando, por un lado, la figura de Aristóteles y, por otro lado, las aportaciones de Pierre Aubenque, Kristján Kristjánsson y Bent Flyvbjerg. El legado sapiencial de este concepto puede fundamentar, cristalizar y supervisar cuidadosamente un enfoque ético para el desarrollo de una investigación e innovación responsable de la IA. Para este propósito el trabajo se dividirá en tres partes: en primer lugar, se definirá la IA, puntualizando algunos desafíos éticos subyacentes en su desarrollo, con el objetivo de plantear la necesidad de concebir esta tecnología como un sistema sociotécnico integrado por elementos técnicos y sociopolíticos; en segundo lugar, se enfatizará el potencial de una perspectiva de investigación e innovación que constituye un terreno fértil para el cultivo de la responsabilidad en el ámbito de los intelectos sintéticos a través de dinámicas participativas, inclusivas y democráticas; y, en tercer lugar, será en el último apartado cuando se argumentará la necesidad de fundamentar el desarrollo responsable de la IA a través de la superación de las lógicas tecnocráticas e instrumentales que, habitualmente, condicionan esta tecnología. La simultaneidad de esta superación y fundamentación se llevará a cabo poniendo en valor la tradición filosófica revelada en el concepto de *phronēsis*, entendida como «una disposición verdadera, acompañada de razón, relativa a la acción en las cosas buenas para el hombre»⁴.

2. Inteligencia artificial y tiempo de desafíos

Durante el verano de 1958 importantes investigadores, entre los que se encontraban John McCarthy, impulsor del evento, Ray Solomonoff, Herbert A. Simon, Trenchard More, Allen Newell, Oliver Selfridge y Arthur Samuel, entre otros, se dieron cita para celebrar el *Dartmouth Summer Research Project on Artificial Intelligence* en la Universidad Dartmouth College⁵. Este encuentro significó la génesis de la IA como campo de investigación. Concretar una definición de IA es una tarea que destaca por la complejidad, debido a la existencia de numerosas propuestas formuladas por los investigadores. No obstante, Ricardo Vinuesa *et al.* ofrecen una dilucidación que contribuye al esclarecimiento de este objeto de estudio:

⁴ ARISTÓTELES: *Ética a Nicómaco*. Alianza, Madrid, 2005, p. 188.

⁵ Cf. MCCORDUCK, Pamela: *Máquinas que piensan*. Tecnos, Madrid, 1999.

«[...] consideramos como IA cualquier tecnología de software con al menos una de las siguientes capacidades: percepción, que incluye audio, visual, textual y táctil (p. ej., reconocimiento facial), toma de decisiones (p. ej., sistemas de diagnóstico médico), predicción (p. ej., pronóstico del tiempo), extracción automática de conocimiento y reconocimiento de patrones a partir de datos (p. ej., descubrimiento de circulación de noticias falsas en las redes sociales), comunicación interactiva (p. ej., robots sociales o bots de chat) y razonamiento lógico (p. ej., desarrollo de teorías a partir de premisas). Este punto de vista abarca una gran variedad de subcampos, incluido el aprendizaje automático»⁶.

Han pasado sesenta y cuatro años y la presencia de la IA es cada vez más notoria en diversas esferas y actividades de la vida humana. Este incremento de la presencia de los intelectos sintéticos ha propiciado el surgimiento de numerosos desafíos éticos que reclaman una investigación e innovación responsable. En este sentido, estos desafíos advierten la necesidad de considerar a la IA como un sistema sociotécnico para facilitar un diagnóstico ético y político que constituya el exordio del cultivo de sabiduría práctica y la asunción de responsabilidad en el contexto de la cuarta revolución industrial. Klaus Schwab⁷ empleó el término «cuarta revolución industrial» para delimitar un momento histórico en el que el progreso de la humanidad está palmariamente condicionado por tecnologías emergentes como la IA, el Internet de las Cosas, el Big Data, los vehículos autónomos, la nanotecnología, etc., y la convergencia de soportes digitales, físicos y biológicos que contribuyen a una transformación del mundo⁸.

La IA se ha convertido en una innovación que estimula cambios a gran escala y pertenece al grueso de las denominadas tecnologías emergentes, de las que se han ofrecido varias definiciones. Henk ten Have y María do Céu Patrão Neves⁹ afirman que este tipo de tecnologías poseen cuatro elementos característicos: están motivadas, principalmente, por la innovación y no tanto por la producción de conocimiento, son el resultado de la convergencia de diversos campos de estudio y sirven como un detonante para el surgimiento de un

⁶ VINUESA, Ricardo et al.: «The role of artificial intelligence in achieving the Sustainable Development Goals», en *Nature Communications* 11 (2006), pp. 1-2.

⁷ Cf. SCHWAB, Klaus: *La cuarta revolución industrial*. Debate, Barcelona, 2016.

⁸ LACY, Peter et al.: *The Circular Economy Handbook*. Palgrave Macmillan, Londres, 2020, p. 45.

⁹ Cf. HAVE, Henk ten/ PATRÃO NEVES, María do Céu: *Dictionary of Global Bioethics*. Springer, Cham, 2021.

gran número de cuestiones éticas. Jason Hyon¹⁰ destacan su potencial habilitante de nuevas capacidades científicas y de exploración, que amplían el espectro de acción de los campos de estudio, ofrecen un alto rendimiento para el desempeño de determinadas actividades y proporcionan nuevas dinámicas operativas y parámetros de medición. Finalmente, Rotolo *et al.*¹¹, en un ejercicio de síntesis, puntualizan sus atributos principales: novedad radical, rapidez de crecimiento, coherencia, impacto prominente e incertidumbre y ambigüedad.

Los desafíos de la IA en el contexto de la cuarta revolución industrial se manifiestan de multitud de formas, entre las que pueden destacarse las siguientes. Stephen Cave y Kanta Dihal¹² han estudiado la preponderancia y los efectos de la blanquitud en los sistemas artificiales reflejados en robots, chatbots, asistentes virtuales y representaciones en la gran pantalla. Entre los efectos de esta «racialización» de la IA se encuentran la supresión en el imaginario colectivo de personas de otro color de piel, así como el incremento de prejuicios y la injusticia real. La falta de transparencia y explicabilidad ha suscitado nuevos reclamos de parte instituciones como la Comisión Europea (CE). La necesidad de transparencia y comprensión ciudadana sobre el funcionamiento de los intelectos sintéticos responde a la opacidad que caracteriza a los algoritmos¹³. La revolución de la automatización en el ámbito profesional es otro de los desafíos éticos y políticos de la IA, pues ha suscitado interesantes debates en torno a los beneficios y amenazas de este fenómeno¹⁴. Los vehículos autónomos como los drones y los coches sin conductor también han sido objeto de investigación, pues sitúan en un escenario atravesado por problemáticas y controversias. En el primer caso, algunos expertos señalan el inicio

¹⁰ Cf. HYON, Jason: «Emerging technologies», en NJOKU, Eni G. (ed.): *Encyclopedia of Remote Sensing*. Springer, Nueva York, 2014.

¹¹ Cf. ROTOLO, Daniele et al.: «What is an emerging technology?», en *Research Policy* 44 (2015), pp. 1827-1842.

¹² Cf. CAVE, Stephen / DIHAL, Kanta: «The Whiteness of AI», en *Philosophy & Technology* 33 (2020).

¹³ Cf. RIBEIRO, Marco Tulio et al.: «“Why should I trust you?»: Explaining the predictions of any classifier», en *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (2016), pp. 1135-1144; MILLER, Tim: «Explanation in artificial intelligence: Insights from the social sciences», en *Artificial Intelligence* 267 (2019), pp. 1-38.

¹⁴ Cf. FLORIDI, Luciano: «Robots, Jobs, Taxes, and Responsibilities», en *Philosophy & Technology* 30 (2017), pp. 1-4; SMIDS, Jilles et al.: «Robots in the Workplace: a Threat to—or Opportunity for—Meaningful Work?», en *Philosophy & Technology* 33 (2020), pp. 503-522.

de una era posheroica¹⁵, mientras que otros advierten de las consecuencias que conlleva el uso de drones y una, supuesta, deshumanización de la guerra¹⁶. En el segundo caso, los coches sin conductor invitan a pensar las implicaciones de esta tecnología y su presencia en las vías públicas¹⁷. Finalmente, el grave proceso de degradación ambiental experimentado por el planeta Tierra en las últimas décadas ha estimulado la investigación sobre alternativas sostenibles en el ámbito de la IA¹⁸.

Estos desafíos de la IA representan un llamado para considerar a esta tecnología emergente como un sistema sociotécnico, abriendo paso a una nueva mirada de carácter ético y político. La reducción de los intelectos sintéticos a simples medios o fines tecnológicos responde a una agenda tecnocrática reduccionista que obstaculiza la reflexión ética como consecuencia de la desvinculación de los contextos sociales y políticos en los que esta tecnología se despliega. Como alternativa a esta desvinculación, Laura Sartori y Andreas Theodorou¹⁹ señalan la importancia de concebir la IA como un sistema sociotécnico caracterizado por la combinación de componentes técnicos y elementos sociopolíticos. Este enfoque sociotécnico de los sistemas artificiales enfatiza que los artefactos tecnológicos no deben pensarse al margen del contexto donde se insertan y los valores a los que afectan. Una comprensión adecuada de las tecnologías emergentes demanda dirigir la mirada hacia los impactos sociales y políticos que ocasiona. En este sentido, el enfoque sociotécnico invita a pensar la tecnología en un entramado de relaciones desplegadas en los contextos sociales y políticos donde subyacen desafíos²⁰. Arrojar una mirada sociotécnica sobre la IA contribuye a un tratamiento diferenciado, informando de las oportunidades que reporta la sabiduría práctica para un desarrollo tecnológico alternativo a las lógicas estrictas

¹⁵ Cf. COKER, Christopher: *Future War*. Polity, Cambridge, 2015.

¹⁶ Cf. SINGER, Peter Warren: *Wired for War: The Robotics Revolution and 21st Century Conflict*. Penguin, Nueva York, 2009.

¹⁷ Cf. DINESH KUMAR, Amara et al.: «A Brief Survey on Autonomous Vehicle Possible Attacks, Exploits and Vulnerabilities», 2018.

¹⁸ Cf. WYNSBERGHE, Aimee van: «Sustainable AI: AI for sustainability and the sustainability of AI», en *AI and Ethics* 1 (2021), pp. 213-218.

¹⁹ Cf. SARTORI, Laura / THEODOROU, Andreas: «A sociotechnical perspective for the future of AI: narratives, inequalities, and human control», en *Ethics and Information Technology* 24 (2022), pp. 1-11.

²⁰ Cf. GREEN, Ben: The Contestation of Tech Ethics: A Sociotechnical Approach to Technology Ethics in Practice, en *Journal of Social Computing* 2 (2021), pp. 218-219.

tamente técnicas y reduccionistas. Finalmente, implica reconocer que la ética no se circunscribe únicamente al entorno del agente moral, sino a un conjunto de relaciones, metodologías, infraestructuras y herramientas que sientan las bases de un escenario determinante para la innovación²¹.

3. Innovación y necesidad de responsabilidad

La innovación posee un valor instrumental y representa una herramienta para dar forma a una vida con mejores condiciones. La persecución de la innovación traducida en avances tecnológicos siempre ha estado presente como una fuerza impulsora del progreso de la humanidad desde sus orígenes. Los avances en materia de innovación proceden: por un lado, del plano teórico, como resultado de la investigación; y, por otro, son fruto del desarrollo práctico o incluso de descubrimientos imprevistos. Al margen de su naturaleza, la innovación inicia nuevos caminos, ofrece la posibilidad de acercamiento a inéditas parcelas de la realidad y provoca cambios a gran escala.

En el apartado anterior, se advirtió la necesidad de concebir la IA como un sistema sociotécnico para impulsar un diagnóstico ético y político. En ese sentido, en el ámbito de la innovación es importante destacar la aplicación de la metáfora de los ecosistemas, surgida en la década de 1990 a raíz de una publicación de James F. Moore²², como una perspectiva útil para comprender la dimensión política y social de la innovación. Bernd Carsten Stahl²³, señala que el atractivo de los ecosistemas se encuentra en el concepto sencillo y ampliamente aceptado que ofrecen para entender la dinámica de funcionamiento del vínculo existente entre los sistemas económicos y los mecanismos de innovación. Al mismo tiempo que brindan coordenadas para la actividad económica, permiten analizar el rol de los actores desde una perspectiva sociopolítica. Esta perspectiva invita a una interpretación novedosa de la innovación, pues no la reduce, exclusivamente, a

²¹ Cf. CHOPRA, Amit K. / SINGH, Munindar P.: «Sociotechnical Systems and Ethics in the Large», en *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society (AIES '18)*. Association for Computing Machinery, Nueva York, 2018, pp. 48-53.

²² Cf. MOORE, James F.: «Predators and prey: A new ecology of competition», en *Harvard Business Review* 71 (1993), pp. 75-86.

²³ Cf. STAHL, Bernd Carsten: «Responsible innovation ecosystems: Ethical implications of the application of the ecosystem concept to artificial intelligence», en *International Journal of Information Management* 62 (2022), pp. 1-13.

aspectos económicos o técnicos. La proyección de esta mirada sobre los ecosistemas de innovación contribuye a una puesta en valor del papel de los actores individuales y colectivos de una sociedad para someter las innovaciones a un escrutinio sobre su aceptabilidad, deseabilidad y sostenibilidad. Beneficia el surgimiento de preocupaciones y despierta la atención como resultado de un impacto que debe estar sujeto a la reflexividad ética. Así pues, es razonable defender que todo estudio actual sobre la innovación en el ámbito de la IA debe integrar una preocupación a nivel teórico y práctico para prestar atención a sus desafíos morales y políticos en el contexto de la cuarta revolución industrial.

El grueso de la literatura existente sobre la relevancia de la innovación en la dinámica del crecimiento económico y el progreso de las sociedades, no deja lugar a dudas de que estamos frente a un concepto que destaca por su complejidad²⁴. No obstante, dada la consideración de la IA como un sistema sociotécnico, es preciso centrar la atención en una perspectiva de la innovación vinculada a la investigación y la responsabilidad.

El modelo Investigación e Innovación Responsable (RRI, siglas en inglés de *Responsible Research and Innovation*) representa una metacategoría de innovación surgida y adoptada en Europa durante la última década. La CE introdujo la RRI como un elemento transversal del programa *Horizonte 2020*, iniciando un debate sobre la necesidad de cultivar la responsabilidad colectiva en los procesos de generación de conocimiento científico y tecnológico, frente a los desafíos de la cuarta revolución industrial y los imperativos éticos recogidos en la Agenda 2030. El interés de la CE por la transversalidad de la RRI ha quedado reflejado, recientemente, en *Horizonte Europa* (HE), un programa de investigación e innovación que se desarrollará durante el periodo 2021-2027. El objetivo principal de HE consiste en «alcanzar un impacto científico, tecnológico, económico y social de las inversiones de la UE en I+D+i, fortaleciendo de esta manera sus bases científicas y tecnológicas, y fomentando la competitividad de todos los Estados Miembros» (CE 2021).

Existen diversas narrativas de la RRI en el contexto europeo, sin embargo, la CE ha adoptado las tesis de René von Schomberg, entendiendo este modelo como «un proceso transparente e interactivo mediante el cual los actores sociales y los innovadores se responden

²⁴ EDWARDS-SCHACHTER, Mónica: «The nature and variety of innovation», en *International Journal of Innovation Studies* 2 (2018), p. 65.

mutuamente con miras a la aceptabilidad (ética), la sostenibilidad y la deseabilidad social del proceso de innovación y sus productos comercializables (con el fin de permitir una adecuada incorporación de los avances científicos y tecnológicos en nuestra sociedad)»²⁵. Este modelo promueve la corresponsabilidad de los diversos actores de la sociedad que se encuentran involucrados en un proceso que debe ser inclusivo, participativo y transparente. Impulsa la integración social a través del cultivo y fortalecimiento de habilidades y prácticas democráticas en el seno de los procesos de investigación e innovación, procurando un equilibrio entre la aceptabilidad, deseabilidad y sostenibilidad de las teorías y prácticas científicas y tecnológicas.

El valor de la responsabilidad de este modelo se encuentra en la adquisición de un compromiso con la praxis, entendido como un ejercicio de sabiduría práctica, enraizada en la prudencia, la deliberación y el juicio sosegado. En este sentido, es necesario recoger el testigo de la tradición filosófica clásica, pues ofrece coordenadas para el fortalecimiento sapiencial en un contexto impregnado por una epistemología estrictamente técnica que resulta insuficiente para afrontar los desafíos éticos actuales de la IA. Por lo tanto, si la responsabilidad constituye un elemento esencial para el desarrollo de la IA en el marco de la RRI, será preciso trascender los límites del ensimismamiento del «tecnopensar» y caminar por la senda de la sabiduría práctica que nos ofrece el concepto aristotélico de *phronēsis*.

4. El valor del legado sapiencial

Durante el proceso de investigación e innovación en el campo de la IA afloran cuestiones que son moralmente relevantes y reclaman ser abordadas con agilidad sapiencial. Para reflexionar sobre estas cuestiones, la historia de la filosofía moral y sus figuras más representativas ofrecen un conjunto de elementos del pensamiento que pueden resultar de utilidad. Entre estas figuras destaca el filósofo clásico Aristóteles, uno de los pensadores más influyentes en la historia del pensamiento de la civilización occidental, con una extensa obra que abarca diversas temáticas, entre las que se encuentran la ética y la política. En vista de los aspectos moral y políticamente subyacentes en la RRI y la IA, la temática de estudio sobre la que se profundizará

²⁵ SCHOMBERG, René von: *Towards responsible research and innovation in the information and communication technologies and security technologies fields*. European Commission-DG Research and Innovation, Luxemburgo, 2011, p. 84.

corresponderá a la ética. Se llevará a cabo un análisis introductorio del concepto aristotélico de *phronēsis*, con el objetivo de extraer aquellas ideas que puedan facilitar la construcción de un vínculo de compromiso práctico entre el saber filosófico y el terreno de la IA.

Aristóteles empleó el término *phronēsis* para representar la excelencia de la inteligencia práctica, equivalente al concepto de *sophia* en el ámbito teórico, sin embargo, no existe una dicotomía absoluta entre ambas formas de racionalidad, sino una gradación de carácter dialéctico²⁶. La clasificación de la epistemología aristotélica se explica en función de un criterio dualista que es más complejo de lo que parece: si la racionalidad tiene como objeto la acción humana (*technē, phronēsis*), se la denomina racionalidad práctica, a causa de su vinculación con la técnica, la moral y la política; en cambio, si se ocupa de conceptos o esencias universales y necesarias (*nous, epistēmē, sophia*), esta racionalidad será de tipo teórico, pues muestra interés por cuestiones científicas y filosóficas.

La concepción aristotélica de la virtud dianoética de la *phronēsis* se erige sobre un equilibrio donde la excelencia práctica consiste en actuar de manera adaptativa a la contingencia. Este equilibrio le atribuye a la prudencia un lugar privilegiado entre las demás virtudes, a tenor de una condición factual que la hace dependiente de un horizonte contingente en el que dibuja su significado. En este contexto, lo contingente se entiende como aquello que puede ser o no ser. De este modo, el estagirita cuestiona aquella perspectiva científica para la que nada puede ser de otra manera. En un mundo así, opuesto a lo contingente, no habría cabida para una acción humana comprometida y responsable. Por el contrario, el reconocimiento de lo accidental y circunstancial supone una invitación constante a la acción del ser humano. En este sentido, Aristóteles argumenta el carácter extensible y extendido de la *phronēsis* a la particularidad de la praxis. Por lo tanto, legado sapiencial que contiene esta virtud se confirma en su aplicación completa e inmediata a cada particularidad, como una reflexión comprensiva que empuja al juicio a una imperecedera adaptación²⁷.

²⁶ VEGA, Jesús: «Aristotle on Practical Rules, Universality, and the Law», en CONTRERAS, Francisco José (ed.): *The Threads of Natural Law*. Springer, Dordrecht, 2013, p. 3.

²⁷ GADAMER, Hans-Georg: *Verdad y método*. Sígueme, Salamanca, 1977, pp. 393-394.

Pierre Aubenque²⁸ lleva a cabo un análisis sobre la prudencia aristotélica, delimitando su estructura antropológica a partir de tres pilares que pueden evidenciar la pertinencia del legado sapiencial en el contexto de la RRI. En primer lugar, destaca que, para el estagirita, es prudente aquella persona que delibera bien. En este sentido, la deliberación acredita la primera dimensión antropológica de la prudencia, constituyendo una especie de investigación o búsqueda encargada de las cosas humanas contingentes. Aubenque aprecia en la deliberación la condición *sine qua non* para que la acción humana pueda ser buena, es decir, virtuosa, fundamentada en la rectitud del entendimiento. En segundo lugar, si la buena deliberación es lo que posibilita a la persona ser prudente, entonces, la elección configura el carácter moral, comprometiendo la libertad, la responsabilidad y el mérito. El filósofo francés recurre a la *Ética* a Eudemo para resaltar el valor de la *proaíresis*, entendida como un momento posterior a la deliberación donde es manifestada la voluntad deseante de carácter intelectual que recoge el testigo de la deliberación y también le pone fin. Así pues, la segunda dimensión antropológica de la *phronēsis* consiste en la elección del mejor medio. En tercer lugar, se encuentra el juicio, considerado por Aristóteles como una cualidad favorecedora del reconocimiento de los límites de la ciencia. Vincula la justicia a la indulgencia, considerando, que el juicio que decide sobre lo equitativo destacará por su rectitud y, por lo tanto, por circunscribirse a lo verdadero. En este sentido, el buen juicio constituye la reunión entre la determinación intelectual y el predicado moral²⁹. Esta concepción del juicio promueve la identificación de los límites equilibrados entre lo particular y lo universal, lo sensible y lo inteligible, sin pretensiones de imponer la rigidez normativa a una realidad contingente, situando lo razonable en una posición preeminente respecto a lo racional, o como apuntaría Hans-Georg Gadamer: «lo bueno no puede determinar nuestro actuar según algo racionalmente dado, sino que es únicamente en esta unidad indisoluble de *ethos* y *dianoia* donde se concreta»³⁰.

Por su parte, Kristján Kristjánsson³¹ también analiza el concepto de *phronēsis*, atribuyéndole dos funciones que enriquecen la argu-

²⁸ AUBENQUE, Pierre: *La prudencia en Aristóteles*. Crítica, Barcelona, 1999.

²⁹ *Ibid.* p. 173.

³⁰ GADAMER, Hans-Georg: *El giro hermenéutico*. Cátedra, Madrid, 1998, p. 215.

³¹ KRISTJÁNSSON, Kristján: «Phronesis as an ideal in professional medical ethics: some preliminary positionings and problematics», en *Theoretical Medicine and Bioethics* 36 (2015), p. 303.

mentación en torno a la pertinencia del legado sapiencial para una investigación e innovación responsable de la IA: en primer lugar, el filósofo inglés le atribuye una función constitutiva, consolidada tras un proceso deliberativo en el que las personas consideran todos los matices y opciones posibles y determinan qué opción es la más adecuada en favor del bien; y, en segundo lugar, afirma que la *phronēsis* no es, únicamente, sensible al bien de una determinada situación o actividad particular, sino también a la totalidad, procurando promover el bien en general³². Estos apuntes de Kristjánsson ayudan a percibir una concepción integradora de esta virtud, admitiendo su aproximación como fundamento de la RRI para la IA.

Finalmente, Bent Flyvberg impulsa una interpretación contemporánea de la virtud aristotélica. El filósofo danés pone en tela de juicio el empleo exacerbado del modelo epistémico enraizado en las ciencias naturales, indicando, que ha servido para expandir una lectura de la realidad en términos de teorías y leyes que resulta insuficiente y subestima el reconocimiento de aspectos políticos subyacentes³³. Este modelo ha impregnado un «tecnopensar» que fundamenta el desarrollo de tecnologías emergentes como la IA, esquivando y rechazando aquellos aspectos de la realidad que se manifiestan problemáticos y no caen bajo el paraguas de la racionalidad instrumental. Como alternativa, el danés promueve una concepción sapiencial de la investigación a partir del reconocimiento de la necesidad del análisis reflexivo de valores e intereses y su impacto en la sociedad. Este análisis conduce a un tratamiento responsable, promoviendo la reflexión en torno a aquellas notas características que dotan a la realidad de complejidad y solicitan ser afrontadas con sabiduría práctica³⁴.

Así pues, los diversos planteamientos sobre esta virtud dianoética dotan al conocimiento de poder, nutriéndolo de un valor político que accede al cultivo del deber cívico de usarlo sabiamente. En este punto se sitúa la importancia del legado sapiencial de la *phronēsis* para un desarrollo responsable de la IA. La puesta en valor de este legado brinda un conjunto de coordenadas esenciales para promover la RRI, fortaleciendo habilidades cívicas y democráticas en el entorno de los intelectos sintéticos. Además, estimula un modo de pensar sosegado

³² *Ibid.*

³³ FLYVBERG, Bent: «Making Organization Research Matter: Power, Values, and Phronesis», en CLEGG, Stewart R. et al. (eds.): *The Sage Handbook of Organization Studies*. SAGE Publications, Thousand Oaks, California, 2006, p. 39.

³⁴ Cf. FLYVBERG, Bent: «Phronetic planning research: theoretical and methodological reflections», en *Planning Theory & Practice* 5 (2004).

en un mundo caracterizado por la velocidad amenazante para el control de las vulnerabilidades de los sistemas artificiales y el diagnóstico de sus desafíos éticos.

5. Conclusión

En estas páginas se ha argumentado la importancia del cultivo de la sabiduría práctica como un elemento fundamental para promover un desarrollo ético de la IA en el contexto de una investigación e innovación responsable. Al comienzo se hizo hincapié en la importancia de reconocer la IA como un sistema sociotécnico, debido a la necesidad de poner en valor aquellos aspectos morales y políticos que subyacen tras su desarrollo y que, habitualmente, se encuentran ocultos como resultado de la preponderancia de un modelo lineal epistémico que fundamenta el pensar técnico. El interés del fortalecimiento sapiencial del ámbito de la IA posee un espíritu que trasciende las divisiones convencionales entre los saberes técnicos y las humanidades. Además, supone una fructífera manera de promover la ética en el seno de las tecnologías emergentes y una herramienta para la adquisición de compromiso con la praxis en aras de una mejor comprensión de los problemas de la realidad originados tras el desarrollo de los intelectos sintéticos.

Así pues, se inicia un camino crítico que debe ser abordado y profundizado por las humanidades como una repuesta de compromiso con el fenómeno digital, proporcionando coordenadas para pensar el mundo actual, facilitando la observación de las dimensiones morales y políticas de los artefactos que proyecta nuestro pensamiento y emprende nuestra acción. Además, la sapiencia de los saberes humanísticos motiva el florecimiento de preguntas como un mecanismo de asunción de responsabilidad con nuestro destino.

La racionalidad de la praxis humana y la racionalidad de la filosofía práctica no se encuentran con lo contingente como si se tratase de algo diferente, pero a su vez inherente a ellas mismas. Se fundamentan en lo fáctico de la praxis que constituye nuestra realidad vital, y no en la derivación a partir de un principio, como correspondería al ideal demostrativo lógico de la ciencia³⁵.

³⁵ GADAMER, Hans-Georg: *El giro hermenéutico*. Cátedra, Madrid, 1998, p. 218.

6. Bibliografía

- ARISTÓTELES: *Ética a Nicómaco*. Alianza, Madrid, 2005.
- CAVE, Stephen / DIHAL, Kanta: «The Whiteness of AI», en *Philosophy & Technology* 33 (2020), pp. 685-703.
- CHOPRA, Amit K. / SINGH, Munindar P.: «Sociotechnical Systems and Ethics in the Large», en *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society Association for Computing Machinery* (2018), pp. 48-53.
- COMISIÓN EUROPEA: *Horizonte Europa: nuevo programa marco de la UE*. 2021. Disponible en: <https://www.horizonteeuropa.es/que-es>
- DINESH KUMAR, Amara / RENU CHEBROLU, Koti Naga / SOMAN KP, Vinayakumar R.: «A Brief Survey on Autonomous Vehicle Possible Attacks, Exploits and Vulnerabilities», 2018. Disponible en: <https://bit.ly/3vRWegX>
- EDWARDS-SCHACHTER, Mónica: «The nature and variety of innovation», en *International Journal of Innovation Studies* 2 (2018), pp. 65-79.
- FLORIDI, Luciano: «Robots, Jobs, Taxes, and Responsibilities», en *Philosophy & Technology* 30 (2017), pp. 1-4.
- FLYVBJERG, Bent: «Phronetic planning research: theoretical and methodological reflections», en *Planning Theory & Practice* 5 (2004), pp. 283-306.
- FLYVBJERG, Bent: «Making Organization Research Matter: Power, Values, and Phronesis», en CLEGG, Stewart R. et al. (eds.): *The Sage Handbook of Organization Studies*. SAGE Publications, Thousand Oaks, California, 2006, pp. 370-387.
- GADAMER, Hans-Georg: *Verdad y método*. Sígueme, Salamanca, 1977.
- GADAMER, Hans-Georg: *El giro hermenéutico*. Cátedra, Madrid, 1998.
- GREEN, Ben: «The Contestation of Tech Ethics: A Sociotechnical Approach to Technology Ethics in Practice», en *Journal of Social Computing* 2 (2021), pp. 209-225.
- HAVE, Henk ten / PATRÃO NEVES, Maria: *Dictionary of Global Bioethics*. Springer, Cham, 2021.
- HYON, Jason: «Emerging technologies», en NJOKU, Eni G. (eds.): *Encyclopedia of Remote Sensing*. Springer, Nueva York, 2014.
- KRISTJÁNSSON, Kristján: «Phronesis as an ideal in professional medical ethics: some preliminary positionings and problematics», en *Theoretical Medicine and Bioethics* 36 (2015), pp. 299-320.
- LACY, Peter / LONG, Jessica / SPINDLER, Wesley: *The Circular Economy Handbook*. Palgrave Macmillan, Londres, 2020.
- MCCORDUCK, Pamela: *Máquinas que piensan*. Tecnos, Madrid, 1991.
- MILLER, Tim: «Explanation in artificial intelligence: Insights from the social sciences», en *Artificial Intelligence* 267 (2019), pp. 1-38

MOORE, James F.: «Predators and prey: A new ecology of competition», en *Harvard Business Review* 71 (1993), pp. 75-86

MOROZOV, Evgeny: *La locura del solucionismo tecnológico*. Clave Intelectual, Madrid, 2015.

RIBEIRO, Marco Tulio et al.: «“Why should I trust you?”: Explaining the predictions of any classifier», en *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (2016), pp. 1135-1144.

ROTOLO, Daniele / HICKS, Diana / MARTIN, Ben R.: «What is an emerging technology?», en *Research Policy* 44 (2015), pp. 1827-1842.

SCHOMBERG, René von: *Towards responsible research and innovation in the information and communication technologies and security technologies fields*. European Commission-DG Research and Innovation, Luxemburgo, 2011.

SCHRAG, Calvin O.: «Praxis», en AUDI, Roberto (ed.): *The Cambridge dictionary of philosophy*. Cambridge University Press, Nueva York, 1999, p. 731.

SCHWAB, Klaus: *La cuarta revolución industrial*. Debate, Barcelona, 2016.

SMIDS, Jilles / NYHOLM, Sven / BERKERS, Hannah: «Robots in the Workplace: a Threat to—or Opportunity for—Meaningful Work?», en *Philosophy & Technology* 33 (2020), pp. 503-522.

STAHL, Bernd Carsten: «Responsible innovation ecosystems: Ethical implications of the application of the ecosystem concept to artificial intelligence», en *International Journal of Information Management* 62 (2022), pp. 1-13.

VEGA, Jesús: «Aristotle on Practical Rules, Universality, and the Law», en CONTRERAS, Francisco José (ed.): *The Threads of Natural Law*. Springer, Dordrecht, 2013.

VINUESA, Ricardo / AZIZPOUR, Hossein / LEITE, Iolanda / BALAAM, Madeline / DIGNUM, Virginia / DOMISCH, Sami / FELLÄNDER, Anna / LANGHANS, Simone Daniela / TEGMARK, Max / NERINI, Francesco Fusco: «The role of artificial intelligence in achieving the Sustainable Development Goals», en *Nature Communications* 11 (2020), pp. 1-10.

WYNSBERGHE, Aimee van: «Sustainable AI: AI for sustainability and the sustainability of AI», en *AI and Ethics* 1 (2021), pp. 213-218.

Recibido el 4 de mayo de 2022
Aceptado el 15 de enero de 2023

Antonio Luis Terrones Rodríguez
Universitat de València / Instituto de Filosofía-CSIC
Antonio.Terrones@uv.es