

## NOTICIAS

**XXI Semana de Ética y Filosofía Política.  
Modelos humanos ante la falsa infalibilidad de la inteligencia  
artificial. Reflexiones sobre la polarización desde la filosofía\***  
XXI Week of Ethics and Political Philosophy. Human models in the face of the  
false infallibility of artificial intelligence. Reflections on polarisation from a  
philosophical perspective

**Rosana Sanahuja Sanahuja**

Universitat Jaume I  
sanahujr@uji.es

ORCID iD: <https://orcid.org/0000-0002-0697-1567>

**Martha Rodríguez Coronel**

Universitat Jaume I

ORCID iD: <https://orcid.org/0000-0002-3228-9019>

**Patrici Calvo**

Universitat Jaume I

ORCID iD: <https://orcid.org/0000-0002-8029-1877>

*Copyright:* © 2023 CSIC. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia de uso y distribución Creative Commons Reconocimiento 4.0 Internacional (CC BY 4.0).

---

\* Este congreso, coorganizado con la Asociación de Ética y Filosofía Política (AEEFP), se enmarca dentro de los objetivos de los proyectos de Investigación Científica y Desarrollo Tecnológico «Bioética cordial y Democracia algorítmica para una sociedad hiperdigitalizada» [PID2022-139000OB-C22] y «Ética aplicada y confiabilidad para una Inteligencia Artificial» [PID2019- 109078RB-C21], financiados por el Ministerio de Ciencia e Innovación de España, así como en las actividades del grupo de investigación de excelencia CIPROM/2021/072 de la Comunitat Valenciana.

La polarización y la tendencia al dogmatismo se presentan como riesgos acrecentados en la sociedad actual por el auge imparable de la inteligencia artificial. Frente a la imagen de fiabilidad y objetividad que muchas veces se atribuye erróneamente a la inteligencia artificial (IA), la falibilidad humana se presenta como una ventaja para afrontar la tendencia al dogmatismo, ya que obliga a las personas a compartir sus verdades y rediseñarse conjuntamente. Sin embargo, esta capacidad dialógica se ve amenazada por el refuerzo a través de la *algoritmización* de la vida pública de posiciones polarizadas y dogmáticas que acrecientan la espiral del silencio. Estas son algunas de las reflexiones lanzadas por Roberto R. Aramayo, Adela Cortina y Emilio Martínez durante la XXI Semana de Ética y Filosofía Política (Congreso internacional) de la Asociación Española de Ética y Filosofía política, presidida por Concha Roldán (FIS-CSIC), en la que más de 200 filósofos y filósofas debatieron a lo largo de tres días, del 2 al 4 de febrero de 2023, sobre *Ética, Democracia e Inteligencia Artificial* en la Universitat Jaume I de Castelló.

Aramayo sintetizaba sus reflexiones en la conferencia<sup>1</sup> poniendo en valor cómo la inteligencia artificial permite desgranar por contraste las mejores cualidades del ser humano entre las que destaca la falibilidad. «Si no pudiéramos equivocarnos, no podríamos aprender de nuestros errores y tendríamos a imponer nuestros dogmas por la fuerza, despreciando el pluralismo y a quienes no piensan igual», indica, añadiendo que «la ética no es cosa de dioses o robots omnipotentes e infalibles. Requiere compatibilizar la búsqueda de nuestro interés con el ajeno, porque nuestra libertad tiene como límites las libertades ajenas».

En la conferencia inaugural “Democracia y posverdad: la espiral del silencio”<sup>2</sup>, la filósofa Adela Cortina abordó asimismo el tema de la polarización, resaltando la importancia de dialogar desde la razón y el riesgo que para ello supone el dogmatismo y la «agresión moralista», una moral que cohesiona a los grupos, pero que logra esa cohesión provocando la vergüenza social en aquellos que están fuera del discurso hegemónico y, siguiendo la teoría de la «espiral del silencio» de la politóloga alemana

Elisabeth Noelle-Neumann (1982), silencian su opinión por no coincidir con la mayoritaria. Las redes sociales y las crecientes posibilidades de la inteligencia artificial aumentan el poder *viralizador* de los contenidos, tanto a la hora de reforzar las opiniones hegemónicas, como a la de «viralizar las afirmaciones no aceptadas por la opinión pública» acrecentando el riesgo de exclusión social y pérdida de reputación en la llamada «cultura de la cancelación», señalaba la filósofa. En este contexto, las personas «temen al aislamiento más que a la mentira sobre todo en tiempos de presunta postverdad».

Para Cortina el mecanismo más sutil para silenciar propuestas es a través del repudio desde la opinión pública porque, siguiendo a Tocqueville, «se teme más al aislamiento que al error». El arma de la vergüenza social refuerza la polarización y atenta contra la democracia: «La moralización de las posiciones polarizadas imposibilita la deliberación y la negociación, imprescindibles en cualquier sociedad democrática, en la que deben entrar en diálogo la diversidad de opiniones», apunta. El problema para Cortina no reside en la moral, sino que la «responsabilidad» es de los que la manipulan para reforzar su estatus y reputación incrementando la polarización: «No se consigue absolutizar una opinión moralizándola, sino dogmatizándola, es decir, sustrayéndola a la crítica racional. Los contenidos de las propuestas no son absolutos, las absolutiza el modo de defenderlos». Y es precisamente este uso absolutista de la moral el que hace patente que es una manipulación porque una norma es moral porque está dispuesta a dar razones de su validez. «Este uso absolutista de lo moral es una manipulación espuria», sentencia Cortina. Frente a ello, el diálogo desde la disposición a dar razones se presenta como un arma fundamental para evitar la polarización, desde la posición de falibilidad contrapuesta al dogmatismo.

En su intervención, Cortina lamentó la vuelta que supone la polarización a las reminiscencias humanas de las sociedades tribales y al instinto de defender lo propio en confrontación con lo externo, cuando la evolución fue mostrando «que lo inteligente es reciprocarse, estar dispuesto a entrar en relación con cualesquiera personas traspasando la barrera de los nuestros siempre que se pueda recibir algo a cambio de aquel a quien se da o de otro que asume esa responsabilidad», señala. «Por eso fueron ganando derecho las sociedades contractuales frente a las tribales». Sin embargo, lamenta, las tendencias tribales quedan latentes y se ven reforzadas por la polarización, que utiliza herramientas lo más cercanas posible a la emoción para evitar razona-

<sup>1</sup> Artículo de opinión escrito por Roberto R. Aramayo a partir de su participación en el congreso y la presentación de Patrici Calvo: <https://elobrero.es/opinion/104409-cual-es-el-papel-de-las-cautelaseticas-frente-a-la-inteligencia-artificial-la-dialectica-entre-tecno-fobia-y-la-tecno-latria.html>

<sup>2</sup> Grabación de la conferencia de Adela Cortina presentada por Domingo García Marzá: <https://www.youtube.com/watch?v=YDbTBY8oykM>

mientos serenos, señala, advirtiendo del riesgo de la polarización como un «fenómeno letal para la sociedad democrática».

La huida de estos razonamientos serenos es uno de los riesgos que Aramayo atribuye también al uso de la inteligencia artificial. Sin entrar en una «necia tecno-fobia» que obvie las ventajas de la inteligencia artificial, pero sin caer tampoco en «una tecno-latría que pudiera servir para delegar irresponsablemente nuestras decisiones», el filósofo advirtió de los riesgos de delegar en los algoritmos nuestros pensamientos y decisiones. «Paulatinamente vamos dejando de utilizar nuestras facultades porque nos resulta muy fácil delegarlas en ciertos dispositivos digitales», afirmó.

Durante la conferencia de clausura del congreso celebrado en la Universitat Jaume I de Castellón, Aramayo advirtió de la engañosa tendencia a creer que las máquinas pueden actuar «con una mayor objetividad e incluso las imaginamos como infalibles, al poder analizar un ingente número de datos para sus complejos cálculos. Pero en realidad los algoritmos fijan y amplifican el sesgo que se haya introducido en su programación». Asimismo, señala el riesgo de delegar en la inteligencia artificial acciones claves del ser humano como es la capacidad de creación. «El ser humano se ha caracterizado por configurar su entorno e interactuar con sus propias creaciones culturales, pero ahora podría perder su condición de artífice y oficiar como un mero espectador de las modificaciones introducidas por la inteligencia artificial», señala.

Ante estos riesgos, la filosofía considera que tiene que actuar como una piedra de toque para estudiarla con una mirada crítica que, sin dejar de reconocer sus virtudes, haga inventario de los eventuales abusos en que se pueda incurrir. Para Aramayo uno de los principales riesgos radica en primar las dimensiones económicas a cualquier precio, de forma que la misión de la inteligencia artificial deje de lado «los aspectos que no sirvan a la consecución de tal objetivo, cuyo logro se podría ver lastrado por otro tipo de consideraciones como las que preocupan a la reflexión ética, empeñada en diseñar criterios morales que causen el menor daño posible a los demás al ejercitar nuestra libertad». Los algoritmos que permiten aumentar la productividad a la vez que ponen en riesgo muchos puestos de trabajo son para el filósofo «el sueño dorado del ultra-neoliberalismo». Como alternativa, propone avanzar hacia el *homo empathicus* como modelo diferencial al de *homo oeconomicus* y, de esta manera, «recuperar la empatía que la mentalidad neoliberal no recoge entre sus principios reguladores». «En ese caso no

se trataría de imponer a toda costa y sin reparar en sus dañinos efectos los propios caprichos» de forma que la solidaridad module la persecución del propio beneficio en detrimento del ajeno. La inteligencia artificial tendrá las presentaciones que se establezcan en su diseño, concluye Aramayo, destacando la importancia de que en esta fase se deje a un lado la pura rentabilidad mercantilista y se apueste por poner el acento «en facilitar una cohesión social donde no proliferen las desigualdades más extremas y se intenten preservar los recursos naturales, con el fin de no robar su futuro a las nuevas generaciones. El artefacto, por sofisticado que sea, se regirá por las pautas de quienes lo diseñen y con arreglo al contexto socio-político del que absorba su arsenal informativo». También en el caso de la inteligencia artificial, como en otros ámbitos «somos la ética que adoptemos».

Las conferencias plenarias se completaron con la intervención de Emilio Martínez sobre “Ética y datos de salud”<sup>3</sup>, quien apelaba por su parte al *homo commiserans*. En su intervención, el filósofo señalaba que hay tres maneras de entender al ser humano a efectos de ética y salud: el *homo oeconomicus*, que persigue la maximización de sus intereses particulares y no debe nada a la sociedad, un modelo que entiende que está «desfasado» y que ha ido dando paso al *homo reciprocans* propio de las sociedades contractuales basado en la idea de «yo ayudo a otros con la condición que me ayuden a mí» en el marco del pacto social. Frente a estos dos modelos, entiende que «deberíamos avanzar éticamente hacia el *homo commiserans*, es decir, yo reconozco a las personas como iguales y libres, como seres valiosos en sí mismos, a quienes debo en gran medida lo que he llegado a ser. Y libremente me comprometo a respetar a los demás, a cuidar de ellos, más allá del contrato social de la reciprocidad». «La reciprocidad está bien, pero deberíamos avanzar e ir un poquito más allá», señala.

En su intervención, Martínez relacionó este modelo de *homo commiserans* con los principios éticos que han de regular el uso de los datos de salud y que, a su entender, pasarían en primer lugar por tratar de reducir la vulnerabilidad, aumentando el empoderamiento de las personas afectadas, y en segundo, por procurar la mayor confidencialidad posible de los datos personales, incluidos los datos sanitarios y genéticos, con límites razonables. De esta forma reivindica el valor de la intimidad que lleva a que los datos personales no se puedan

<sup>3</sup> Grabación de la conferencia de Emilio Martínez presentada por Elsa González Esteban: <https://www.youtube.com/watch?v=pctwIZvPf98>

hacer públicos salvo en casos muy excepcionales por razones de interés común. En tercer lugar, sitúa como principio ético el control público sobre los procedimientos, permisos y prohibiciones de acceso y manejo de datos personales de salud, resaltando en este sentido la necesidad de una regulación y de una vigilancia que abarque «desde el diseño mismo de las tecnologías que tiene que tener participación de expertos, porque normalmente son tecnologías muy complejas, y del público al que van dirigidas, es decir, de todos nosotros».

En suma, el Congreso Internacional de la AEEFP sobre *Ética, Democracia e Inteligencia Artificial* constituyó un momento disruptivo para el desarrollo de una filosofía moral de la inteligencia artificial a la altura de los tiempos. Tanto las tres conferencias plenarias como las comunicaciones y debates generados alrededor de estas giraron en torno a una idea de fondo: la sociedad mundial se halla inmersa en los albores de una nueva era, sin precedentes en la historia de la humanidad, cuyo correcto desarrollo exige salvar retos tecnológicos, pero también éticos y democráticos. Se trata de una transformación de la sociedad y sus esferas funcionales mucho más global, compleja y profunda que cualquier otra anterior, cuyos impactos se están dejando sentir con fuerza en los procesos relacionales, asistenciales,

productivos, comunicativos, educativos e, incluso, decisivos en clave de sostenibilidad, predictibilidad, velocidad, exhaustividad, extensibilidad, capacidad, completitud, consistencia, eficiencia, ratificación, precisión o detección.

Sin embargo, tras esta transformación subyacen problemas vinculados con la inmadurez tecnológica y, sobre todo, el aumento de la vulnerabilidad de los afectados y afectadas por el diseño, aplicación y uso de la IA, especialmente aquellos y aquellas pertenecientes a los colectivos más vulnerables de la sociedad. Cabe, por tanto, trabajar en el desarrollo científico-técnico de la IA para poder extraer de ella su máxima potencialidad. Pero también en su desarrollo moral y democrático para erradicar aquellos impactos éticos y políticos que, como la brecha digital, los sesgos (aporófobos, xenófobos, homófobos, misóginos, racistas o edadistas), la irresponsabilidad, la opacidad, la aceleración o la anomia limitan su progreso y castigan a la sociedad, especialmente a los grupos más vulnerables.

El congreso resultó clave para mostrar y orientar el compromiso de la filosofía moral con el diseño, aplicación y uso de una inteligencia artificial confiable y responsable, capaz de promover una sociedad más justa y *felicitante* y no el mero beneficio particular de las grandes tecnológicas.