

Cortina Orts, Adela (2024). *¿Ética o ideología de la inteligencia artificial? El eclipse de la razón comunicativa en una sociedad tecnologizada*. Editorial Paidós.

Basilo A. Martínez-Villa
Centro de Estudios Universitarios Xochicalco, México



Recibido: 13/10/2025

Revisado: 11/12/2025

Aceptado: 15/01/2026

Cómo citar: Martínez-Villa, B. A. (2024). Reseña: ¿Ética o ideología de la inteligencia artificial? El eclipse de la razón comunicativa en una sociedad tecnologizada. *Revista Estudios de la Información*, 4(1), 88-93. <https://doi.org/10.54167/rei.v4i1.2118>

Resumen: En la era digital hoy en día, la tecnología no sólo es intuitiva y fluida; las interfaces gráficas, la inteligencia artificial (IA) y la automatización han hecho que la interacción con sistemas complejos sea más fácil que nunca. Pero ¿hacia dónde nos lleva este desarrollo tecnológico? ¿Cuál es la dimensión ética de la tecnología? Ante la posibilidad de crear una IA más capaz que la inteligencia del ser humano, ¿Se corre el riesgo de la instrumentalización de la vida humana? Conectarse no es comunicarse, es necesario cuidar afanosamente la palabra, para que ocurra la comunicación se tiene que ser veraz. En torno a estos y otros cuestionamientos radica la composición del libro que se reseña.

Palabras Clave: Ética, Ideología, Inteligencia Artificial, Racionalidad, Información.

Abstract: In today's digital age, technology is not only intuitive and fluid; graphical interfaces, artificial intelligence (AI), and automation have made interacting with complex systems easier than ever. But where is this technological development taking us? What is the ethical dimension of technology? Given the possibility of creating an AI more capable than human intelligence, is there a risk of instrumentalizing human life? Connecting is not communicating; it is necessary to carefully guard one's words; for communication to occur, it must be truthful. The book under review is structured around these and other questions.

Keywords: Ethics, Ideology, Artificial Intelligence, Rationality, Information.

A lo largo de sus más de 240 páginas, la Dra. Adela Cortina, Catedrática Emérita de Ética y Filosofía Política de la Universidad de Valencia, nos lleva a un ejercicio de análisis y reflexión crítica sobre la Inteligencia Artificial (IA) y su entorno en un contexto humano desde una perspectiva ética.



Este libro se estructura en 12 capítulos de la siguiente manera: 1. Lo que nos hace humanos. 2. Tres tipos de Inteligencia Artificial: Tres tipos de ética. 3. Transhumanismo y Posthumanismo ¿Pronóstico o Ideología? 4. Inteligencia Artificial confiable ¿Un cambio de época? 5. Ética Robótica. 6. ¿Qué ética para las máquinas inteligentes? 7. Las fronteras de la persona en un mundo tecnocientífico. 8. La libertad en la era digital. 9. Eclipse de la razón comunicativa: un reto radical para la democracia. 10. Un espacio público libre de dogmatismos. 11. Lenguaje Claro. 12. Educar en la era de la Inteligencia Artificial.

El Libro publicado por editorial Paidós en octubre de 2024, se apoya en un trabajo de investigación previo: el Proyecto de Investigación Científica y Desarrollo Tecnológico FFI2016-76753-C2-1-P, financiado por el Ministerio de Economía y Competitividad (ahora Ministerio de Ciencia, Innovación y Universidades) y en las actividades del grupo de investigación de excelencia PROMETEO/2018/121 de la Generalidad Valenciana. Ese trabajo *Ética de la inteligencia artificial* se integra por cuatro apartados de los cuales al menos tres forman parte de los capítulos de este libro. Parafraseando a la autora: El mundo de la IA es irreversible. La IA ha nacido en el espectro de las promesas y los riesgos; suscitando a la vez promotores y detractores. Pero, con la IA generativa, ¿Se busca que los seres humanos utilicen la IA o que la IA sustituya a los seres humanos?

En el capítulo 1, *Lo que nos hace humanos* la autora a partir de una referencia al novelista británico Ian McEwan, en su obra *Máquinas como yo*, inicia su reflexión sobre el proceso de comunicación ‘La verdad importa’. Precisamente a partir de la premisa que las máquinas no están hechas para entender una mentira, pero además nos dice que mentir no suele tener malas consecuencias para el mentiroso. Precisamente a partir de la ficción de McEwan sostiene que el eje de su reflexión inicial es ¿qué sucedería si las máquinas pudieran comprender una mentira? En tiempos de una ‘supuesta’ posverdad se dan por buenas las mentiras emotivas.

Continúa desarrollando su reflexión refiriendo *La isla del Doctor Moreau* y *las Leyes de Asimov*. De la primera obra recuerda el fallido intento de hominización, como los animales terminan por volver a la selva que es su hábitat después de haber incendiado el laboratorio y matar al doctor y ahí inserta su primer planteamiento ético, como una predisposición eminentemente humana, lo que muestra al hombre su deber en el caso concreto. Ahí entonces la cuestión, ¿se puede y debe introducirse en las ‘mentes’ de los robots normas éticas? De manera que coincidan con los intereses humanos. Desarrolla un análisis de las *Leyes de robótica de Asimov* y concluye que éstas a menudo entran en contradicción, ya que para aplicarlas es necesario entender el contexto e interpretarlo. Cierra este primer apartado con las interrogantes que desarrollará a lo largo del libro qué tipo de ética debe prevalecer para la IA y cómo es esencial cuidar la palabra y la comunicación.

El Capítulo 2, *Tres tipos de Inteligencia Artificial, tres tipos de ética* inicia narrando la anécdota de las elecciones en 2018 en el Distrito de Tokio, Tama New Town, donde un robot fue postulado a las elecciones y quedó en tercer lugar de la votación general y a partir de ahí desarrolla la idea de algunos que las cosas – en este caso – un mejor gobierno, se puede obtener resultado de la IA. Posteriormente aporta algunos datos históricos sobre IA y distingue tres tipos de IA, que a su juicio ameritan distintos tipos de ética cada una: (1) la inteligencia superior o superinteligencia, refiriéndose a una inteligencia que supera la inteligencia del hombre y que eventualmente lo sustituiría. Esta es la IA que da lugar a la perspectiva transhumanista y posthumanista con la idea de la singularidad; (2) la inteligencia general, la cual puede resolver problemas generales, cuya finalidad

es crear una IA, una máquina que tenga una inteligencia general similar a la humana; y (3) la inteligencia especial que es la que lleva a cabo trabajos específicos, realiza tareas concretas, pero de manera superior a la inteligencia humana. Concluye que el reto consiste en orientar el uso humano de estos sistemas de forma ética.

El capítulo 3, *Transhumanismo y Posthumanismo: ¿Pronóstico o ideología?* introduce la reflexión a través de la referencia literaria, ahora se apoya en los hermanos Huxley, particularmente en *Un mundo feliz* de Aldous y no tanto en *Julian*, pero sí menciona que este último quien en 1927 introduce una idea de transhumanismo. La creencia de que a través de la tecnología podemos ser mejores versiones del ser humano, la transformación a través de las tecnociencias, a través del acrónimo NBIC (nanotecnología, biotecnología, informática – big data, internet de las cosas – y ciencias cognitivas – IA, robótica, entre otras). ¿Se puede servir al hombre, trascendiendo al hombre, permaneciendo hombre al actualizar nuevas posibilidades de su naturaleza humana?

Se corre el riesgo de la instrumentalización de la vida humana. La autora plantea que ello no debe, ni puede ser permisible. Aquí realiza – a gusto de quien suscribe – uno de los análisis más interesantes del libro, apoyada en la crítica de Kant, sobre la responsabilidad de no mentir de la ciencia. Afirma: “es lícito aventurar bosquejos de totalidad y de futuro como idea regulativa, como orientación... desde los avances con que ya se cuenta, pero no es lícito darlos por realidad contrastada”. Muestra así tres desafíos para la ética en este entorno transhumanista: el primero, una ideología, ya liberar, ya progresista, ya indeterminada detrás de la técnica; el segundo, cuidar que no se intente seducir desde las pulsiones humanas, intereses económicos; y, tercero, esclarecer el tipo de ética al que se apela al mostrar una bondad moral en el entorno transhumano. Se debe distinguir entre las propuestas que gozan de un respaldo científico de aquellas que no lo tienen.

En el capítulo 4, *Inteligencia Artificial confiable ¿Un cambio de época?* se presenta el interés del marco regulatorio europeo para la IA, de manera que se pueda cuidar el desarrollo de la misma bajo supervisión humana y al servicio de las personas. Explica como el marco normativo en el espacio europeo es de vanguardia en ese sentido. Sin embargo, propone, precisamente la necesidad que ante los requerimientos del mercado y del entorno global, se genere una ética para este cambio de época, aquí es donde desarrolla su modelo normativo propuesto: (1) beneficencia; (2) autonomía y dignidad; (3) explicabilidad y rendición de cuentas; 4) no dañar; y (5) promover un mundo justo ¿Un cambio de época?

En el Capítulo 5, *Ética robótica la autora hace un recorrido por la llamada cuarta revolución industrial y la robótica*, la autora cuestiona en primer término ¿qué es un robot? Y ella misma aporta una variada serie de descripciones. Plantea el desafío ético para la robótica hablando entre otras cosas del riesgo de la deshumanización del trabajo de la salud. Sostiene la necesidad de generar una ética para las máquinas y cuestiona si se tratará de una ética para las máquinas o para agentes artificiales autónomos, reflexión que continúa en el capítulo 6.

El Capítulo 6; ¿Qué ética para las máquinas inteligentes? se centra en la reflexión continua sobre la necesidad ética para las máquinas, la autora sostiene que lo que es lo más popular o lo más solicitado, no necesariamente es lo éticamente correcto. Manifiesta cómo es un error común confundir alta incidencia de preferencias en una encuesta, con lo que se ‘debe’ hacer. Evidentemente refiere una lógica de mercado. Se debe – sostiene la autora – generar una ética que

vaya más allá de sólo prevenir daños o evite generarlos, precisamente ahí es donde introduce un punto de inflexión en su planteamiento: Se trata de sistemas funcionalmente autónomos, no ontológicamente autónomos. Lo que regresa el enfoque a si las máquinas pueden realmente ser consideradas agentes morales autónomos. Una reflexión que propone la autora es que una diferencia esencial entre la máquina y el humano – y ello es cualitativamente sustancial – es el ‘pensar despacio’ lo que permite identificar móviles de acción en el intercambio de argumentos.

Continúa y refiere cinco marcos éticos para el análisis: (1) el deber o deontologismo desde Kant hasta Asimov; (2) para pasar por el utilitarismo, donde la autora utiliza de manera equivalente los valores utilidad y placer; (3) luego aborda, brevemente, la ética desde la virtud; (4) el enfoque de las capacidades para concluir en su propuesta del ‘deontologismo dialógico’ o la noción de deber desde la acción comunicativa; y (5) una hermenéutica crítica.

Respecto al Capítulo 7, *Las fronteras de la persona en el mundo tecnocientífico*, se estudia el sintagma de la ‘singularidad de la humanidad’ donde se liga la idea de singularidad a la creación de máquinas inteligentes; de inteligencias superiores a la humana. Refiere también al *ciborg*, este ser de componentes híbridos, vinculando al hombre-máquina, menciona incluso a la ‘tecnopersona’ y precisamente aquí es donde amplía sobre aspectos regulatorios en el entorno europeo sobre la persona: Personas electrónicas. Además, se cuestiona si estamos ¿superando? el biocentrismo.

El Capítulo 8, *La libertad en la era digital*, ofrece un planteamiento de carácter filosófico que en este capítulo se efectúa consiste en señalar cómo en la actualidad existen los nuevos gurús que sostienen una variable de determinismo tecnológico, cómo la combinación de genes, neuronas y mundo social, que no hemos elegido, condicionan nuestro actuar y por ende nuestra falta de libertad. Describe de una manera muy clara, como es una tendencia histórica que en la medida en que una ciencia es reconocida en el ‘pódium’ de la excelencia, una parte significativa de sus defensores se tornan imperialistas y tratan de explicar la totalidad de los movimientos de la naturaleza y de la conducta de las personas, desde ese particular nicho de la ciencia.

En el Capítulo 9, *Eclipse de la razón comunicativa: Un reto radical para la democracia*, parte aquí de una pregunta esencial: ¿Hacia dónde queremos ir? señala como con el progreso de las tecnociencias la aceleración ‘parece ser la clave vital’ lo que hoy es una solución, queda obsoleta mañana y hace un planteamiento lógico inverso: ‘las respuestas nos cambiaron las preguntas’. Ella se identifica como partidaria del ‘cosmopolitismo democrático’ postula en términos de Kant la idea de ir democratizando los países uno a uno hasta el entorno mundial y afirma que no es una utopía sino una idea regulativa. Defiende la razón comunicativa más allá de la razón instrumental, considerando que la primera, permite la identificación y logro de fines a diferencia de la segunda.

Menciona la simulación democrática en entornos de dictadura y reflexiona precisamente sobre el papel de la verdad en tiempos de posverdad y la necesidad de cuidar la palabra. Detalla como la connotación de posverdad es contradictoria, ya que –en una interpretación semántica del término–, quien la utiliza distorsiona deliberadamente la realidad porque la conoce ‘y no le interesa comunicarla’. Además, ofrece un recorrido por diferentes modelos normativos que refieren el valor del lenguaje y la palabra, menciona la teoría de la justicia social crítica y cierra su reflexión con el siguiente análisis: ‘la profusión de prácticas que defienden la legitimidad de utilizar en el debate público términos con significantes ambiguos o vacíos, pero con una connotación positiva para la

ciudadanía...' vacían a las palabras de significado. Anulando, en términos de Habermas y su Teoría de la Acción Comunicativa, la pretensión de verdad y de validez del enunciado.

En relación con el Capítulo 10, *Un espacio público libre de dogmatismos*, el libro enriquece su análisis y reflexión al dedicarlo al espacio público y a la libertad de pensamiento y expresión, exaltando el valor de la acción comunicativa. Aborda dos concepciones de opinión pública una de carácter normativo y otra de carácter descriptivo. Establece cómo el espacio público debe propiciar las condiciones para acabar con el dogma a través de la crítica abierta. Desarrolla también una reflexión en torno a la moral y como la 'vergüenza moral' refiriéndola como 'agresión moralista' ha sido históricamente un instrumento para moldear la opinión o el espacio público. Para cerrar el apartado postulando la necesidad de crear un entorno donde sea permisible hablar con verdad sin vergüenza de hacerlo.

El enfoque del Capítulo 11, lenguaje claro, ofrece un análisis que gira entorno a la palabra o el lenguaje; cómo palabras como 'democracia', 'progreso' y 'progresista' son términos manipulados, por lo que es necesario cuidar la palabra en estos casos. Evitar la desmesura. Y como es necesario regresar a un elemento esencial de occidente el reconocimiento al 'logos' combatir la 'logofobia'. Ya que el 'logos' permite la interpretación del mundo social y natural y no implica, como actualmente se hace en muchos casos una interpretación de individuos aislados.

Aborda también de una manera más explícita la atención a la teoría de la acción comunicativa de Habermas vinculada con la teoría de los actos de habla de Austin y Searle y como el hablante debe aceptar las pretensiones de validez para que la comunicación sea efectiva entre ellas la inteligibilidad, el mensaje debe ser claro, evitar la ambigüedad. La autora identifica la ambigüedad con la oscuridad como arma ideológica. Amplia además la reflexión al establecer que la virtud o cortesía del filósofo es un punto hacia la igualdad como derecho de las personas.

Como cierre, el libro presenta el Capítulo 12, Educar en la era de la inteligencia artificial, acude al ejemplo de China y si bien reconoce que no se puede saber cuáles serán las necesidades concretas de la educación de los niños en un futuro; afirma: 'lo que sí podemos saber es qué valores morales queremos transmitirles para que los acepten o los desechen' la concepción de que la educación es esencial para la vida humana, trasciende las fronteras y las civilizaciones. Regresa su reflexión a la IA como instrumento, pero señala sus luces y sombras. Cierra su análisis indicando que la complejidad del siglo XXI podría ser mejor comprendida con inteligencia aumentada y distribuida como es el caso de la IA, pero la educación presenta limitantes que se deben atender: la brecha digital; los sesgos de los algoritmos; inequidad educativa, ausencia de parámetros ético en el uso de la tecnología, entre otras. Lo cual puede poner en peligro la capacidad crítica y la autonomía de las personas.

La autora – como ya se ha detallado – sostiene que los avances técnicos históricamente han precedido a las orientaciones morales sobre cómo hacer uso de ellos y que, en el caso de la IA, no es la excepción. Explica cómo la IA incluye varios enfoques y técnicas, tales como *Machine Learning* (*Deep Learning* y *Reinforcement Learning*), *Machine Reasoning* (planeación, programación, representación de conocimiento y razonamiento, así como búsqueda y optimización), y robótica (control, percepción, sensores y actuadores (*actuators*), y los sistemas ciberfísicos que incluyen todas las anteriores variables). Reitera la autora: el paso es irreversible.

De la lectura del libro surgen múltiples temas y cuestionamientos, uno de ellos sobresale y se distingue, en términos de Horkheimer en su *Crítica de la razón instrumental*, como asistimos a una reflexión sobre la maximización de los medios, sin tener claro el porqué de esa potencialización y, por ende – al paso de un tiempo – cómo éstos se convierten en el fin y se pierde la intención ética o moral de la utilización del instrumento. La verdadera esencia de la razón, consiste en hallar medios para lograr los objetivos, pero ello no implica que el medio – que es un instrumento - se convierta en el fin en sí mismo.

Si se pierde el humanismo, en aras de un transhumanismo o posthumanismo, sostiene la Doctora Cortina, se pierde la dimensión ética de la persona. La IA – al menos en este momento – carece de intencionalidad, son los humanos quienes postulan – arriesgadamente – un mejor futuro con el *Ciborg* o con el traslado de nuestras mentes a un soporte digital. La máquina no tiene intencionalidad, esto al margen de si estamos hablando de IA fuerte o débil. Pero vinculando el desarrollo de la tecnología a la ética, hace una referencia científica entre algunas otras de carácter literario: la conciencia o sentido moral no sería, pues, según Darwin, una dimensión intranscendente de la vida humana, (...) marca la diferencia más relevante entre el hombre y el animal y tiene supremacía sobre cualquier otro principio humano; precisamente en ese entono de desarrollo tecnológico es dable afirmar que el auge de la automatización ha profundizado aún más el desapego de la sociedad de la alfabetización tecnológica y de manera necesaria el desapego a la normatividad.

Las herramientas impulsadas por IA sustituyen los roles humanos, haciendo que los procesos sean más eficientes, pero a la vez, eliminan la necesidad de que una persona comprendan los mecanismos subyacentes en la herramienta. La automatización elimina la interacción humana, fortaleciendo un ciclo en el que los individuos se convierten en participantes pasivos en lugar de agentes activos para solución de problemas. En la medida en que una tecnología se torna amigable, menos la entiende la gente. La IA, las redes sociales, la automatización han hecho la vida más cómoda, pero también han creado un mundo en el que las personas están cada vez más ciegas a las fuerzas que dan forma a su realidad digital. Es la llamada ‘paradoja de la interfaz’. Ahí está el reto de la ética.

Sostiene la autora que conectarse no es comunicarse, para que ello ocurra es necesario cuidar afanosamente la palabra, para que ocurra la comunicación se tiene que ser veraz. Agregaría: se tiene que ser activo. La autora, además, sostiene que desgraciadamente son malos tiempos para la veracidad, no tanto para la verdad, que continúa conviviendo con la presunta ‘posverdad’. Lo que peligra sobre todo es la veracidad, y conviene recordar que lo contrario de la verdad es el error, lo contrario de la veracidad es la mentira. De ahí la necesidad de criterios normativos. En un entorno de tecnología, algoritmos, oferta y demanda, empresas, la ética es un imperativo. El Libro desarrolla cinco principios éticos que deben prevalecer en la era digital: (1) beneficencia; (2) autonomía y dignidad; (3) explicabilidad y rendición de cuentas; (4) no dañar; y (5) promover un mundo justo.

El libro es muy interesante y está escrito con una prosa dinámica, lo que permite profundizar en ideas complejas a través de ejemplos literarios y anécdotas de la técnica. En tiempos donde prevalece la orientación científica y mediática hacia la superficie, profundizar más allá de preferencias y algoritmos, es una bocanada de aire fresco. Los invito a leerlo.