



Filosofía de la Inteligencia Artificial

Carlos M. Madrid Casado

Pentalfa Ediciones
2024

El libro sobre Filosofía de la Inteligencia Artificial (IA) de Carlos M. Madrid Casado, ofrece una visión general de los temas asociados con el progreso y los retos de la IA. Ello, bajo la perspectiva de:

- un enfoque que privilegia la crítica a las visiones optimista de la IA y que precisa los riesgos y desafíos que se deben enfrentar en su desarrollo.
- las coordenadas del Materialismo Filosófico de Gustavo Bueno¹ desde las cuales se considera que la IA es un producto humano y social. Lo cual implica que no se puede concebir como un ente con una existencia independiente.
- la defensa y centralidad de los valores humanos en la utilización, avances y mitigación de riesgos como consecuencia de la IA.

Teniendo como prefacio un texto intitulado: «El Mito de la IA», el autor desarrolla el contenido del libro en tres partes (figura 1), las cuales están referidas a:

- 1) Gnoseología de la IA: centrada en el análisis y fundamentación de si la IA es una ciencia, una técnica o una tecnología.
- 2) Ontología de la IA: exploración focalizada en delimitar qué es la «inteligencia» y qué es lo «artificial», la crítica a la concepción algorítmica de la inteligencia y en el debate sobre las distintas modulaciones (fuerte/débil, general/restringida) previstas para el progreso de la IA.

¹ Destacado filósofo español autor del enfoque denominado Materialismo Filosófico.

- 3) Ética de la IA: se concentra en la reflexión y debates, entre otros temas, sobre si los algoritmos o los robots son sujetos éticos; en los dilemas éticos, morales y políticos de la IA; en las indefiniciones de los marcos normativos vigentes; en la geopolítica de la IA.

En otros términos, la propuesta filosófica del autor al realizarse desde las coordenadas del materialismo filosófico articula las tres partes señaladas: gnoseología, ontología y ética de la IA.

Figura 1

Contenido



Fuente: elaboración propia.

En la primera parte, la clave es la revisión de los asuntos que plantean los saberes sobre la IA a partir de la teoría del cierre categorial.² En la segunda parte, el interés se concentra en las realidades de la IA, y en la última parte se analizan las relaciones entre ética, moral y política.

Las ideas fuerza de estas partes se han sistematizado en una tabla, a los efectos de tener una visión de conjunto sobre lo que considera el autor una filosofía de la inteligencia artificial, y a los fines de precisar el cómo se desarrolla el hilo conductor de las reflexiones y argumentaciones.

² Constituye otra forma de concebir la ciencia como un sistema de operaciones que construye objetos de estudios. Donde las categorías u objetos de estudios son las construcciones conceptuales.

Cuadro 1

Ideas fuerza

Parte	Capítulo	Idea Fuerza
I. Gnoseología de la Inteligencia Artificial	¿Cuál es el estatuto gnoseológico de la Inteligencia Artificial?	» La IA es una herramienta, y por ende no debe ser considerada como un sujeto cognoscente.
		» La capacidad de la IA de convertir datos en información útil depende de los algoritmos y modelos que sean utilizados, en ningún caso del sentido común, la creatividad y la comprensión del contexto. No posee conciencia propia, ni capacidad de comprender.
		» La IA no puede comprender el significado profundo de los patrones, ni establecer mecanismos causales entre la correlación de datos.
		» La IA no puede reproducir de manera completa la experiencia humana en la construcción del conocimiento, ni la comprensión integral que se requiere un determinado contexto.
		» Es un mito que la IA pueda reemplazar a la inteligencia humana y no requiera de control humano.
II. Ontología de la Inteligencia Artificial	¿De qué hablamos cuando hablamos de «inteligencia artificial»? Argumentos en contra de la IA. La IA no es inteligencia ni es artificial.	» La ontología de la IA debe considerar además de los factores propiamente técnicos, a la historia , la cultura, la política, la tecnología, etc.
		» La IA es un conjunto de algoritmos y datos que funcionan en un hardware específico. Es decir, no posee una existencia independiente del hardware.
		» La IA no posee conciencia ni la intencionalidad de un humano. Tampoco dispone de la capacidad de crear realidades propias.
		» La IA puede ser una herramienta útil para el desarrollo de la humanidad.
		» La IA es parte de la realidad material y depende de los componentes a partir de los cuales fue construida.
III. Ética de la Inteligencia Artificial	Ética, moral y política de la IA	» Bajo la perspectiva ética, se enfatiza la determinación de la responsabilidad en caso de daños, los riesgos en los sesgos de los algoritmos, la privacidad de los datos personales y el hecho que la herramienta debe estar al servicio del ser humano.
		» Es un peligro dar la idea que la IA tiene cualidades humanas como la conciencia y la intencionalidad.
		» La IA requiere de marcos regulatorios para asegurar su uso en democracia y evitar que contribuya a preservar las desigualdades sociales.

Fuente: elaboración propia.

Entre las conclusiones del texto podemos destacar:

- la IA es un conjunto integrado de algoritmos y datos que operan en un hardware.
- la IA es una pluralidad de tecnología cuyas operaciones se centran en fines prácticos. Por ejemplo, el diseño de máquinas y sistemas inteligentes que puedan realizar actividades de una manera más eficiente que los humanos.
- la inteligencia humana no se puede concebir sólo desde una visión algorítmica.
- «la singularidad tecnológica y las visiones apocalípticas son mitos y escatologías» que promueven la ciencia ficción y los transhumanistas.³
- el progreso de la IA como tecnología debe tener en cuenta sus consecuencias y riesgos para la sociedad.
- las aplicaciones de la IA requieren considerar sus implicaciones políticas y el establecer regulaciones que aseguren su uso responsable.
- en lugar de ética de la IA resulta más conveniente hablar de moral y, en especial, de políticas de IA.
- la investigación en IA debe fundamentarse en la búsqueda de una comprensión de los mecanismos causales de los fenómenos.

Ciertamente el evaluar lo que subyace en el concepto de la IA, su alcance y desafíos se requiere de un enfoque filosófico riguroso y sistemático. En este caso el enfoque utilizado por el autor fue el materialismo filosófico de Gustavo Bueno, el cual le ha facilitado una aproximación profunda y estructurada al tema en cuestión:

- En lo gnoseológico al rechazar las concepciones idealistas de la IA, colocar el énfasis en la IA como producto material-social, y al aplicar la teoría del conocimiento para analizar el cómo la IA logra y procesa la información, a partir de datos, algoritmos y hardware.

³ Movimiento que aspira a transformar la condición humana (limitaciones físicas y cognitivas) a partir del desarrollo y la utilización de tecnologías avanzadas.

- En lo ontológico al considerar los objetos de estudio de la IA como entidades reales, al utilizar la teoría del cierre categorial para clasificar los sistemas involucrados en la IA y al reconocer la emergencia de nuevas propiedades en los sistemas complejos.
- En lo ético al considerar la responsabilidad humana sobre los avances y la utilización de la IA, y en la reflexión sobre los valores y normas que deben orientar y regular los avances en esta materia.

En los anexos del texto se encuentran varios documentos, muy significativos, sobre proyectos de investigación, principios de la IA y una declaración en materia de IA. Tales como: a) Una propuesta para el proyecto de investigación de verano en la Universidad de Dartmouth sobre inteligencia artificial; b) Propuesta de investigación de M. L. Minsky; c) Propuesta de investigación de N. Rochester; d) Propuesta de investigación de John McCarthy; e) Pausar los experimentos gigantes de IA: una carta abierta. 22 de marzo de 2023; f) Los 23 Principios de la Inteligencia Artificial coordinados por el Instituto para el Futuro de la Vida y desarrollados en la conferencia celebrada en Asilomar en 2017; g) Declaración Transhumanista adoptada por la organización Humanity+, anteriormente conocida como la Asociación Transhumanista Mundial, en marzo de 2009.