



**Título:** Una teoría crítica de la inteligencia artificial

**Autor:** Daniel Innerarity

**Editorial:** Galaxia Gutenberg, S.L

**País:** España

**Fecha:** 19 marzo 2025

**Páginas:** 552

**Idioma:** español

**Formato:** Impreso

## Recensión realizada por:

**Dilia Monasterio**

Universidad Central de Venezuela

ailidadm@gmail.com

ORCID: <https://orcid.org/0000-0002-4341-5850>

Caracas-Venezuela

**Fecha de recepción:** 12/11/2024

**Fecha de aprobación:** 22/11/2024

## Descripción

En *Una teoría crítica de la inteligencia artificial*, Daniel Innerarity presenta un análisis profundo y multifacético sobre las implicaciones éticas, sociales y políticas de la inteligencia artificial (IA) en el contexto contemporáneo. A través de su trabajo, el autor, es titular de la Cátedra de Inteligencia Artificial y Democracia en el Instituto Universitario Europeo de Florencia, cuestiona la naturaleza de la tecnología algorítmica y su influencia en la vida democrática y el autogobierno.

Innerarity establece que la crítica efectiva a la IA debe ir más allá de ser una simple advertencia moral, destacando la importancia de entender la complejidad intrínseca de la tecnología. La IA no es meramente una cuestión técnica, también involucra dimensiones ontológicas y socioculturales. Esto implica repensar categorías fundamentales como sujeto, acción, responsabilidad, conocimiento y trabajo en un entorno donde la algoritmización se ha vuelto omnipresente.

## Introducción

En la era digital, la inteligencia artificial (IA) ha emergido no solo como una herramienta de eficiencia, sino como un actor crucial en la reconfiguración de nuestras estructuras sociales, políticas y económicas. Daniel Innerarity, en su obra *Una teoría crítica de la inteligencia artificial*, se sumerge en esta complejidad, ofreciendo un análisis crítico y profundo, el filósofo en esta obra desafía las narrativas simplistas vinculadas al tema de la IA. Publicado en un contexto de creciente preocupación por el impacto de la tecnología en la democracia y el sujeto humano, el libro nos invita a cuestionar qué significa ser humano en un mundo donde las máquinas toman decisiones las cuales alguna vez consideramos exclusivas de nuestro juicio.

El libro *Una teoría crítica de la inteligencia artificial* se destaca como una obra esencial para aquellos interesados en entender el impacto de la IA en la democracia y las dinámicas sociales. Innerarity ofrece herramientas valiosas



para cuestionar las implicaciones éticas y políticas de la automatización.

El libro está estructurado en varias secciones que abordan diferentes dimensiones de la IA y sus repercusiones. Innerarity comienza por la Introducción titulada *Crítica de la razón algorítmica*, Innerarity aborda cómo la tecnología algorítmica ha transformado la manera de percibir y evaluar el mundo. Su eficacia plantea un desafío a la crítica reflexiva, limitando la capacidad de cuestionar las construcciones sociales detrás de nuestras percepciones. La introducción menciona la necesidad de un diálogo continuo entre el desarrollo tecnológico y sus protagonistas, enfatizando como la teoría crítica no solo busca una ética responsable de la IA, sino que investiga las condiciones estructurales que influyen en su regulación y legitimidad democrática.

La primera parte, *Teoría de la razón algorítmica*, explora la naturaleza de la inteligencia artificial y su intersección con el conocimiento humano. Se examinan dos tipos de inteligencias: la artificial y la humana, y se discute la encrucijada actual de la inteligencia artificial. Además, se investiga la especificidad del conocimiento humano, abordando aspectos como el sentido común, la reflexividad y el conocimiento implícito, además de las limitaciones inherentes a la inexactitud y el aprendizaje. Finalmente, se analizan las diferencias entre los techies (tecnólogos) y los fuzzies (humanistas), destacando sus diferentes enfoques y contribuciones en el contexto algorítmico.

La segunda parte, *Pragmática de la razón algorítmica*, se centra en la infraestructura tecnológica de la sociedad digital y su impacto en la toma de decisiones. Se discuten tres enfoques fundamentales: el tecnosolucionismo, el cual aboga por soluciones tecnológicas a problemas sociales; el tecnoneutralismo, este enfoque sostiene que la tecnología es inherentemente neutra; y el tecnodeterminismo, esta perspectiva argumenta que la tecnología determina el progreso social. Cada uno ofrece una visión única sobre la relación entre humanos y tecnología, planteando interrogantes sobre cómo las decisiones automatizadas influ-

yen en nuestra soberanía y en la construcción de la democratización en un contexto mediado por algoritmo.

La tercera parte, *Filosofía política de la razón algorítmica*, aborda cómo la IA transforma la gobernanza y las dinámicas democráticas. Se plantea la delegación del control a máquinas y el impacto de esta tendencia en la toma de decisiones. La gobernanza algorítmica promete objetividad y eficiencia, pero también limita la democracia al hacer opacas las cadenas de responsabilidad. La obra critica tanto la superficialidad de enfoques éticos como la simple moratoria en el desarrollo tecnológico. Se cuestiona qué significa el autogobierno democrático en un contexto donde las decisiones se delegan a algoritmos, destacando la necesidad de un examen crítico y político de estas tecnologías.

El autor cierra con un apartado denominado conclusión que la bautiza bajo el nombre *El futuro de la democracia en la era digital*, donde el autor hace un llamado a la reflexión colectiva sobre el papel de la IA en nuestras vidas. Innerarity destaca la urgencia de establecer diálogos continuos que aborden las intersecciones entre tecnología, humanidad y democracia, enfatizando la responsabilidad de cada uno en esta conversación crucial. Se centra en la crítica a la automatización en la toma de decisiones humanas, resaltando un dilema clave: ¿hasta qué punto puede la inteligencia artificial reemplazar procesos humanos? Se analizan tres enfoques para abordar este reto: la moratoria, la ética y la crítica política. Se destaca la necesidad de una regulación proactiva en lugar de detener el desarrollo tecnológico. La obra propone que, en lugar de temer a la IA, es crucial fomentar un enfoque que integre reflexión y supervisión ética, para garantizar no solo la efectividad, sino también la justicia en su implementación.

## Conclusión

El trabajo de Innerarity es una invitación a la reflexión colectiva sobre cómo queremos que el desarrollo tecnológico influya en nuestra vida cotidiana y nuestras instituciones democráticas. Al plantear preguntas fundamentales sobre el significado de ser humano en el contexto de la IA,



el autor nos empuja a ser actores críticos en el proceso de definir un futuro en el que la tecnología sirva a la humanidad, en lugar de dominarla.

En este sentido, el aporte a la obra de Innerarity, se traduce en presentar las dimensiones ontológica-epistemológica y axiológica que se captan en la obra. Se percibe que la naturaleza de la realidad se presenta como un constructo ontológico y sociocultural influido por la tecnología y la IA. En este contexto, la digitalización no es meramente un fenómeno técnico, sino que reconfigura conceptos fundamentales como el sujeto, la acción y la responsabilidad. Innerarity argumenta que la automatización plantea desafíos a la autodeterminación democrática, sugiriendo una complejidad inherente en la que la crítica se vuelve esencial para desvelar la articulación del poder en la sociedad digital y promover un autogobierno que no sea sacrificado por la racionalidad algorítmica.

La obra de Daniel Innerarity se caracteriza por un enfoque crítico hacia la IA, examinando la interacción entre esta y la democracia. La naturaleza de su obra se basa en cuestionar la forma en que la automatización impacta en la autogobernanza y el control social, planteando la necesidad de una reflexión ética y política que supere las meras regulaciones técnicas. Asimismo, se distingue que la relación entre el investigador y su obra es profundamente introspectiva, en la que Innerarity no solo analiza la tecnología, sino que también desafía las concepciones preestablecidas sobre la misma y su integración en la vida humana.

Además, desde un plano axiológico, el autor señala que el desafío del *ethics washing*, por tanto, critica el uso superficial de la ética como una herramienta para evitar cambios estructurales significativos, resaltando la necesidad de una reflexión conceptual más profunda. La tendencia a utilizar la ética como una especie de “lavado de cara”, donde se presenta una imagen de responsabilidad sin abordar los problemas estructurales importantes. Esto sugiere que, en lugar de simplemente aplicar principios éticos de manera superficial, debemos profundizar en la complejidad de la tecnología y su impacto en la sociedad. (Innerarity, 2025; p. 31).

En esta línea, Innerarity resalta que el desafío en esta materia es conceptual. Desde estas palabras el filósofo como en sus obras anteriores, nos convoca a profundizar sobre el conocimiento. Efectivamente, la digitalización ha permeado todos los aspectos de nuestras vidas, estamos entrando en lo que diversos autores han denominado la ‘Era inteligente’ en este siglo.

