



Daniel Innerarity, (Coord.)

Una teoría crítica de la Inteligencia artificial

Ediciones Galaxia Gutenberg, 552 páginas, 2025



“La automatización generalizada plantea el problema de qué lugar corresponde a la decisión humana” (Innerarity, 2025, 23). Considerando la co-evolución de los procesos de digitalización y democráticos, en este libro el autor interpela las relaciones entre las tecnologías y las prácticas sociales. Por esto se sugiere como una lectura imprescindible para docentes.

Sus páginas exponen a la razón algorítmica, tecnocrática, desmitificando el oxímoron: *inteligencia artificial*. Innerarity explica claramente las aplicaciones y el diseño de las funciones predictivas, generativas, y agenticas de los algoritmos de aprendizaje automático. También, cómo las dificultades

humanas para comprender abstracciones complejas e interacciones entre las tecnologías y el mundo social provocan visiones reduccionistas

y antropomorfizadas de estas tecnologías; las cuales son categorizadas como 'herramientas' y caracterizadas como 'humanizadas'.

Así nos introduce en la noción de *democracia de las recomendaciones*, explicando cómo la vigilancia algorítmica infiere, perfila y puede condicionar nuestras decisiones micro y macro políticas. En esta obra, "la datificación no es solo un continuo crecimiento del volumen de los datos sociales sino un cambio cualitativo en la constitución de la realidad social" (Innerarity, 2025, 21). Y porque inciden en nuestros pensamientos y prácticas, "al hablar de la dimensión política de los algoritmos no solo debemos pensar en su utilización sino en la lógica específica con la que se inscriben en el mundo social" (Innerarity, 2025, 435). Si como agentes educativos comprendemos que las tecnologías reconfiguran las prácticas sociales en las organizaciones escolares y académicas que habitamos, seguramente, nos resonará su enfoque filosófico-político y cuestionaremos muchas prácticas e ideas irreflexivas circulantes que nos influncian.

Galardonado con el III Premio de Ensayo Eugenio Trías, el título del libro define su contenido: Una teoría crítica de la inteligencia artificial. Daniel Innerarity, titular de la Cátedra de Inteligencia Artificial y Democracia, construye y finaliza en Fráncfort una *teoría crítica* que propone itinerarios posibles para la gobernanza de los sistemas algorítmicos. Su trabajo se compone con una introducción, catorce capítulos organizados en tres secciones, una conclusión y una extensa bibliografía de consulta. Cada párrafo convoca a reflexionar problematizando los sesgos de la digitalización. Sus argumentos y ejemplos nos interpelan y, a la vez, demuestran *cómo* y *porqué* la presencia silente de estas tecnologías reconfigura el sistema democrático interviniendo en las instituciones.

En su extenso y profundo ensayo Innerarity se distancia de visiones apocalípticas y también del optimismo ingenuo; no obstante reconoce amenazas y explica que el fortalecimiento o la destrucción de la Democracia dependen de cómo se configure la gobernanza de la inteligencia artificial. Donde "la cuestión no es si las máquinas pueden pensar sino si son capaces de entender relaciones de sentido" (Innerarity, 2025, 61). En palabras del autor, es necesaria una inteligencia artificial más fenomenológica y hermenéutica que cartesiana. Dicha oposición, y sinergia, entre paradigmas atraviesa el texto del catedrático de filosofía quien realiza preguntas imprescindibles, por ejemplo: "¿Hasta qué punto un mecanismo que presume de no juzgar puede estar a cargo de nuestra existencia

personal y colectiva?” (Innerarity, 2025, 430)... “si intrínsecamente su campo de validez de lo calculable es limitado” (Innerarity, 2025, 75). El autor estimula nuestro pensamiento crítico y aprecia el ‘sentido común’, y junto a dichas formas de inteligencia humana destaca otras inexistentes en los sistemas algorítmicos: “la inteligencia es corporal, contextual y situacional; incluye el entorno, a los otros humanos y la cultura”, Larson (2021) en (Innerarity, 2025, 71).

Daniel Innerarity jamás desestima las ventajas de la digitalización pero advierte que la *gobernanza algorítmica*, tecnocrática, se exhibe como despolitizada aparentando una realidad desideologizada. Porque “la tecnología basada en el aprendizaje automático se ha extendido a tantos ámbitos y se ha convertido en una parte tan integral de la vida social que parece haberse vuelto invisible y suele considerarse como no problemática” (Innerarity, 2025, 435). Es decir, que los sistemas algorítmicos despolitizan simulando objetividad y presentando como neutrales e incuestionables los resultados generados con criterios opacos y sesgos de programación ocultando su interés a quienes no agencian el proceso de datificación. En dicho escenario cabrían deliberación y autodeterminación; porque no se puede desestimar que en la sociedad de los *big data* (optimizada mediante datos, con espacios monitorizados y automatizados por tecnologías, usando recursos naturales) la irreflexividad y la abundancia de datos influyen en la gobernanza de cada organización: educativa, económica, política o social.

El catedrático creó una teoría exenta de pensamiento binario que invita al lector a reflexionar. Propone que “deberían existir métodos de intervención humana en las diferentes fases de implementación de los procesos de automatización que, sin poner en peligro los beneficios de la automaticidad, permitan considerar este proceso como verdaderamente democrático” (Innerarity, 2025, 468). Más que regular la automatización, enfatiza en prever y gobernar democráticamente cada fase del diseño de las tecnologías de aprendizaje automático.

Además, su libro interpela al campo educativo analizando las transformaciones de forma y de sentido provocadas en dichas organizaciones que alteran tanto al aprendizaje, la enseñanza y la evaluación como la comunicación. En el capítulo sobre la *inteligencia de la inteligencia artificial*, el autor indaga en los procesos de aprendizaje preguntándose: “¿cómo distinguir entre una función subsidiaria y un horizonte de reemplazo de nuestra inteligencia?” (Innerarity, 2025, 82). Y aproxima

respuestas: “las herramientas cognitivas solo pueden ser utilizadas de forma significativa por las personas si estas tienen una idea de los procesos cognitivos de los que las libera” pero “deberíamos también ponderar hasta qué punto tiene sentido minusvalorar el aprendizaje de ciertos procesos cognitivos por el hecho de que la inteligencia artificial pueda hacerlos” (Innerarity, 2025, 81-84).

Finalmente disfruté leer este libro, en su versión física, porque enfoca dilemas y prácticas del escenario mundial actual. Admito que es muy interesante leerlo y que no es fácil reseñar una obra tan crítica. Recomendando este ensayo a docentes de distintos ámbitos educativos porque problematiza lo que parece obvio y porque plantea preguntas que atraviesan la formación permanente del profesorado y la educación ciudadana.

MARINA PATRICIA DE-LUCA
Doctora por la Universitat Rovira i Virgili
marinapatricia.deluca@alumni.urv.cat