

INTELIGENCIA ARTIFICIAL GENERATIVA: ¿ES POSIBLE ESCAPAR A LA DICOTOMÍA PESIMISMO/OPTIMISMO?

Generative Artificial Intelligence: Is it Possible to Escape the Pessimism/Optimism Divide?

Dairon Alfonso Rodríguez Ramírez

Universidad Industrial de Santander, Bucaramanga, Colombia.



Resumen

El artículo aborda la dicotomía entre tecno-optimismo y tecno-pesimismo en el debate sobre la inteligencia artificial generativa. El objetivo principal es trascender esta polarización, mostrando cómo ambas posturas, a pesar de sus diferencias, comparten una visión simplificada de la tecnología, concibiéndola como totalmente controlable o completamente autónoma. La conclusión central del trabajo es que la inteligencia artificial generativa debería reconocerse como un posible socio cognitivo y no como un mero reemplazo o rival de la inteligencia humana. Adicionalmente, la atribución de la responsabilidad de sus resultados debería distribuirse entre diseñadores, curadores de datos, reguladores y usuarios, quienes interactúan de formas complejas con el propio diseño tecnológico. Al reconocer esta agencia distribuida y el funcionamiento algorítmico de la inteligencia artificial generativa, el artículo propone una relación donde los sistemas inteligentes artificiales potencien las capacidades humanas sin despojarnos de nuestra agencia.

Palabras clave: *tecno-optimismo; tecno-pesimismo; Inteligencia Artificial; agencia; interacción humano-máquina.*

¿Cómo citar?: Rodríguez Ramírez, D. A. (2025). Inteligencia artificial generativa: ¿es posible escapar a la dicotomía pesimismo/optimismo? *Praxis Filosófica*, (62S), e20615436. <https://doi.org/10.25100/pfilosofica.v0i62S.15436>

Recibido: 01 de julio de 2025. **Aprobado:** 04 septiembre de 2025.

Generative Artificial Intelligence: Is it Possible to Escape the Pessimism/Optimism Divide?

*Dairon Alfonso Rodríguez Ramírez*¹

Universidad Industrial de Santander, Bucaramanga, Colombia.

Abstract

The article addresses the dichotomy between techno-optimism and techno-pessimism in the debate on generative artificial intelligence. The main goal is to transcend this polarization, showing how both positions, despite their differences, share a simplified vision of technology, conceiving it as either totally controllable or completely autonomous. The central conclusion of the work is that generative artificial intelligence should be recognized as a possible cognitive partner, and not as a mere replacement or rival of human intelligence. Additionally, responsibility for its results should be distributed among designers, data curators, regulators, and users, who interact in complex ways with the technological design itself. By recognizing this distributed agency and the algorithmic functioning of generative artificial intelligence, the article proposes a relationship where artificial intelligent systems enhance human capabilities without stripping us of our agency; Human-machine interaction.

Keywords: *Techno-optimism; Techno-pessimism; Artificial Intelligence; Agency.*

¹ Profesor adscrito a la Escuela de Filosofía de la Universidad Industrial de Santander, de donde también se graduó como profesional en Filosofía. Obtuvo su doctorado en Humanidades (Filosofía) en la Universidad Autónoma Metropolitana y una maestría en Ciencias Cognitivas en la Universidad Autónoma del Estado de Morelos, en México. Sus áreas de especialidad incluyen la filosofía de la tecnología, la epistemología social y la filosofía del lenguaje.

INTELIGENCIA ARTIFICIAL GENERATIVA: ¿ES POSIBLE ESCAPAR A LA DICOTOMÍA PESIMISMO/OPTIMISMO?

Dairon Alfonso Rodríguez Ramírez

Universidad Industrial de Santander, Bucaramanga, Colombia.

I. Introducción

Nuestra existencia se desenvuelve, la mayor parte del tiempo, en el marco de un complejo entramado de objetos que han sido producidos por la actividad humana para satisfacer necesidades reales o simbólicas (Preston, 2013). Reflexionar sobre esta inmersión cotidiana en lo tecnológico, y sobre la naturaleza de los artefactos y las técnicas asociadas a su uso y producción (Sharff y Dusek, 2014, p. 241), debería ser un objetivo primordial de la filosofía de la tecnología. Sin embargo, una revisión de los referentes clásicos de la disciplina revela una tendencia predominante a inquirir más bien sobre el significado o el impacto general de la tecnología en la vida humana, priorizando las consecuencias sobre las condiciones que posibilitan nuestra relación con lo tecnológico.

Relacionado con este énfasis, una dicotomía recurrente y problemática se manifiesta en la filosofía de la tecnología, como señalan Mary Tiles y Hans Oberdiek (1995). Se trata de la polarización entre el tecno-optimismo y el tecno-pesimismo, discursos que, pese a su antagonismo, comparten una visión reduccionista. Conciben la tecnología, y en particular la inteligencia artificial generativa², de manera excluyente: o bien como un instrumento plenamente dócil y controlable, o bien como una fuerza incontrolable y autónoma. Esta visión binaria distorsiona la complejidad de los sistemas

² En el presente artículo, siempre que hablemos de inteligencia artificial (IA) nos estaremos refiriendo a su versión generativa, la cual se distingue por el uso de modelos de aprendizaje profundo para crear contenido novedoso.

técnicos y limita el desarrollo de un marco conceptual que propicie una interacción más responsable con lo tecnológico.

El presente trabajo busca cuestionar indirectamente dicha dicotomía, no para minimizar la importancia de las preguntas sobre las repercusiones de la tecnología, sino para explorar las posibilidades y consecuencias de trascender esta visión polarizada. Se analizará cómo el tecno-optimismo y el tecno-pesimismo interpretan los impactos de la tecnología a partir de una misma base conceptual. Posteriormente, se esbozará una perspectiva de la inteligencia artificial generativa que rompa con esta dualidad, concibiéndola no como una mera sustituta o competidora de la inteligencia humana, sino como un elemento susceptible de ser incorporado en diversos procesos de creación y cuyos resultados inmediatos son el producto de la intervención de distintos agentes. Esta redefinición es crucial para forjar una relación más madura, ética y sostenible con la tecnología, permitiendo que los sistemas inteligentes sirvan a los fines humanos sin comprometer nuestra propia agencia y responsabilidad.

II. Entre la utopía y el apocalipsis: las visiones dicotómicas de la tecnología

II.1 Tecno-optimismo y tecno-pesimismo

El tecno-optimismo, también denominado liberalismo tecnológico (Dotson, 2012), utopismo tecnológico (Spilioti, 2016) o progresismo tecnológico (Marwick, 2013), se inclina a juzgar favorablemente el impacto de la tecnología en nuestras vidas. Sostiene que su adopción y desarrollo pueden tener un efecto positivo sobre el bienestar humano (Danaher, 2022). Este bienestar se concibe de forma multidimensional, al incluir aspectos sociales, como la prosperidad económica y la mejora de nuestros estándares de vida (López-Portillo, 2018), pero también rasgos de nuestra cotidianidad, en particular, el tiempo libre disponible y la capacidad de autorrealización personal (Dotson, 2012). Es importante resaltar que el tecno-optimismo, al menos en sus versiones más robustas, no afirma que el impacto positivo de la tecnología sea una consecuencia inevitable, sino más bien una posibilidad alcanzable bajo ciertas condiciones. Veremos esto con más detalle más adelante.

El tecno-optimismo se manifiesta claramente en los discursos y actitudes de quienes, con diversos intereses, promueven la creación y adopción de nuevas tecnologías en las sociedades capitalistas contemporáneas, ya sean

desarrolladas o en vías de desarrollo. Por ejemplo, el tono de los gurús de las nuevas tecnologías digitales, como Elon Musk o Mark Zuckerberg, al presentar sus desarrollos más recientes, es casi sin excepción triunfalista, pues creen haber gestado las innovaciones que transformarán radical y positivamente alguna faceta de nuestra vida cotidiana.

Los defensores del tecno-optimismo han resonado con más fuerza últimamente, impulsados por los recientes avances en IA, cuyo potencial para generar transformaciones positivas en casi todos los campos de la actividad humana, desde la programación, pasando por la medicina (Sai *et al.*, 2024), y las finanzas, hasta la creación artística, ha sido enfatizado ampliamente (Bubeck *et al.*, 2023). Un claro ejemplo de esta visión optimista proviene de Sonya Huang, gestora de un importante fondo de inversiones en Silicon Valley, quien afirma:

The fields that generative AI addresses —knowledge work and creative work—comprise billions of workers. Generative AI can make these workers at least 10% more efficient and/or creative: they become not only faster and more efficient, but more capable than before. Therefore, Generative AI has the potential to generate trillions of dollars of economic value. (Huang, 2022, párr. 4)

La perspectiva tecno-optimista de corte capitalista subraya cómo la tecnología resolverá problemas globales, impulsará la prosperidad económica (Chui *et al.*, 2023), aumentará el tiempo libre o potenciará la autorrealización personal. Ejemplos claros de estas promesas en el ámbito de la IA incluyen predicciones sobre la automatización completa (vehículos autónomos), el desarrollo de la medicina personalizada impulsada por IA (Zhang *et al.*, 2025), o los grandes modelos del lenguaje como motores de la creatividad y la eficiencia (OpenAI, 2023). Todo esto, mientras se ignoran las posibles limitaciones y riesgos inherentes a tales desarrollos tecnológicos.

Sin embargo, el tecno-optimismo no es una visión exclusiva de la cultura tecnológica que se gesta en las sociedades capitalistas occidentales. Para Marx y Engels,

[l]a tecnología, siempre que pertenezca y sea gestionada por toda la comunidad para el bien común, se concibió como sustituta de los esclavos, liberando a todos de la necesidad de trabajar, poniendo así a su disposición la posibilidad de una vida humana plena. (Tiles y Oberdiek, 1995, p. 252)

Considerar la visión marxiana o marxista sobre lo tecnológico es valioso porque nos permite hacer explícito un aspecto muy importante del tecno-optimismo. La postura defendida por Marx, Engels y otros teóricos marxistas, no supone ingenuamente que la tecnología sea siempre positiva para la humanidad. De hecho, se argumenta que, particularmente durante la etapa de producción capitalista, la tecnología puede servir a fines que favorecen exclusivamente a la clase social dominante: la burguesía. En contraste, la vasta mayoría de la población, incluidos los operadores directos de las herramientas y máquinas, quedan excluidos del disfrute de los beneficios que estas producen, o para decirlo con el lenguaje de *El Capital* (Marx, 2017[1867]), los medios de producción (el término de Marx para las tecnologías agrícolas e industriales del siglo XIX) solo sirven al fin específico de la acumulación de capital. No obstante, Marx advierte en tono profético que, con un cambio político radical —la revolución proletaria—, los beneficios inherentes a los medios de producción podrán redistribuirse entre toda la población.

El futuro prometedor que Marx augura para los trabajadores del mundo sería impensable si no se concibiera lo tecnológico como una fuente inagotable de potenciales beneficios sociales y personales. Sin embargo, su vasto potencial para la transformación positiva de la sociedad se ve, según Marx, constreñido por el acaparamiento egoísta de sus frutos por parte de la burguesía, en tanto que clase social privilegiada. En este sentido, el problema no residiría en la tecnología en sí, pues esta solo produce prosperidad material, sino más bien en los usos que de ella se hacen en la etapa de producción capitalista. En un sistema donde los medios de producción pasen a manos del proletariado, esos mismos réditos podrán redistribuirse entre todos los trabajadores de manera justa, por lo que el potencial benéfico de los medios de producción quedaría liberado.

La lección que se desprende del mensaje marxiano es que la tecnología, si bien posee un gran potencial benéfico (tecno-optimismo puro), también podría emplearse de tal forma que solamente favorezca a unos pocos. Pensemos, por ejemplo, en un arma, como un fusil automático, que posee un gran poder de fuego debido a su cadencia de cientos de disparos por minuto y al calibre mediano de sus balas. Aunque siempre es posible usarlo de manera egoísta, digamos para asaltar un banco, un uso adecuado de dicha arma traería beneficios a la mayor cantidad de personas posibles. En este caso, un uso «correcto» podría consistir en que solo sea portado por miembros autorizados de los cuerpos de seguridad oficiales con el propósito de disuadir o contrarrestar conductas delictivas, lo cual es posible gracias a su gran poder de fuego. En última instancia, para el tecno-optimismo, todo

se reduce al uso que se le dé a la tecnología y a los fines que dichos usos persigan.

Esto último abre la puerta a una consideración adicional implícita en el tecno-optimismo, a saber: la idea de que la tecnología es benéfica porque sus fines han sido elegidos de acuerdo con consideraciones y cálculos orientados al máximo beneficio social e individual (con la salvedad, por supuesto, de que capitalistas y marxistas entienden «beneficio social» de maneras muy diferentes). Así, las distintas versiones del tecno-optimismo asumen que el carácter benéfico de la tecnología no se explica por una suerte de bondad inmanente, sino porque las tecnologías disponibles conllevan un tipo de uso que tiene un impacto positivo.

En el recuento marxiano, esto es más que evidente. Pero ¿qué ocurre con las versiones más «capitalistas»? En ellas también es posible identificar el mismo supuesto, aunque de formas menos explícitas. Por ejemplo, al defender que la innovación impulsada por el libre mercado conduce, casi siempre, a productos y servicios que mejoran la calidad de vida y la eficiencia, se argumenta que los emprendedores y las empresas, al buscar su propio beneficio económico, se esfuerzan por diseñar productos que resuelvan necesidades o problemas reales de sus posibles compradores. En síntesis, la postura tecno-optimista asume que el propósito y la aplicación de la tecnología pueden alinearse con la generación de un beneficio universal, ya sea en virtud de una planificación consciente (véase: Mesthene, 1970) o debido a la dinámica inherente del mercado (por ejemplo, Shumpeter, 1942; Arrow, 1962). Analizaremos algunas implicaciones de esto más adelante.

En oposición a esta postura, encontramos el tecno-pesimismo, que también se manifiesta tanto en el discurso académico como en la cultura popular. El ejemplo clásico de pesimismo tecnológico en la literatura es la novela decimonónica *Frankenstein o el moderno Prometeo*, de Mary Shelley (1818). En ella, la creación se vuelve contra su creador y lo destruye, una narrativa interpretada como una metáfora de nuestra siempre conflictiva relación con lo tecnológico. Así, la visión tecno-pesimista reconoce lo que su opuesto parece soslayar: la introducción de nuevas tecnologías, por más inocuas que parezcan, conlleva riesgos y efectos insospechados que son tremendamente difíciles de controlar o prevenir. Así, lejos de ser predecible, la tecnología es incontrolable y, por ende, altamente riesgosa.

Basándose en esta premisa, el pesimista afirma que, en lugar de hacer realidad nuestros sueños de prosperidad y realización personal —como ingenuamente cree el optimista—, la tecnología plantea, más bien, una amenaza existencial para la humanidad, o al menos para aquello que nos hace verdaderamente humanos, como, por ejemplo, el pensamiento (Han,

2022) o nuestra libertad (Ellul, 1964). La razón para magnificar los riesgos de la tecnología parece hallarse en los miedos que despiertan ciertos episodios trágicos de la historia humana, donde las innovaciones tecnológicas jugaron un papel central. La bomba atómica, los desastres ecológicos o el calentamiento global, por mencionar solo algunos ejemplos, son usados como argumentos para reforzar la tesis central del tecno-pesimismo, a saber, que los males que puede llegar a producir la tecnología son mayores a sus beneficios iniciales.

Curiosamente, el tecno-pesimismo es la postura que más resuena en los clásicos de la filosofía de la tecnología. Autores como Heidegger, Ellul o Mumford, provenientes de tradiciones de pensamiento muy diversas entre sí, se alzan como profetas del desastre potencial que la tecnología, en particular las más novedosas, parece entrañar. Al respecto, Heidegger declara en el transcurso de una famosa entrevista lo siguiente:

Todo funciona [en el mundo tecnológico actual]. Esto es precisamente lo inhóspito, que todo funciona y que el funcionamiento lleva siempre a más funcionamiento y que la técnica arranca al hombre de la tierra cada vez más y lo desarraiga. No sé si usted estaba espantado, pero yo desde luego lo estaba cuando vi las fotos de la Tierra desde la Luna. No necesitamos bombas atómicas, el desarraigo del hombre es un hecho. Solo nos quedan puras relaciones técnicas. Donde el hombre vive ya no es la Tierra. (Heidegger, 1976, p. 206)

Avances técnicos, como el viaje a la Luna, refuerzan, según Heidegger, una cierta concepción de la “Tierra” como un mero objeto susceptible de ser explotado, un recurso más dentro de un complejo sistema técnico. Bajo esta perspectiva, el problema central no es que la tecnología no funcione, sino que funciona demasiado bien, sumergiéndonos en una existencia puramente técnica y despojándonos de una relación más auténtica con el ser y con nuestro entorno. Más adelante en la misma entrevista, Heidegger agrega que nada bueno puede salir de este desarraigo, por lo que cabe pensar que, para el autor alemán, nuestros tiempos están marcados por una decadencia sin precedentes (véase: Cárdenas Arenas, 2005).

Una vez más, esta visión general en torno a la tecnología encuentra su correlato en el discurso sobre la IA. En este ámbito, el tecno-pesimismo se manifiesta con fuerza en la preocupación por la autonomía creciente de los sistemas de inteligencia artificial (Bostrom, 2014), la pérdida de control humano sobre decisiones críticas, y la posibilidad de que estas tecnologías, lejos de ser herramientas al servicio de la humanidad, se conviertan en

agentes con agendas propias o con efectos secundarios impredecibles y devastadores (Russell, 2019). Se teme que la IA pueda socavar la verdad, manipular la opinión pública a una escala sin precedentes (O'Neill, 2016), o incluso alcanzar un punto en el que su capacidad de aprendizaje y adaptación haga incomprensible su funcionamiento y, por ende, los vuelva incontrolables para sus creadores (Bender *et al.*, 2021; Marcus y Davis, 2019).

El recelo hacia los desarrollos más recientes en IA se ve alimentado porque su núcleo técnico, las redes neuronales artificiales, a menudo da la impresión de funcionar como «cajas negras» (Pasquale, 2015). Esto se debe a que están conformadas por numerosos módulos, lo que se traduce en una cantidad ingente de líneas de código difícilmente rastreables para cualquier programador. Así, cada día es más difícil, primero, identificar cómo los modelos de IA más avanzados llegan a sus resultados y, segundo, prever cómo podrían comportarse en situaciones imprevistas, reforzando la idea de que este tipo de desarrollo «es lo que es», inescrutable y potencialmente peligroso, tal como lo anunciaban algunos filósofos en el siglo XX (Ellul, 1960).

Adicionalmente, la capacidad de los modelos generativos para producir contenido aparentemente indistinguible del creado por humanos, combinada con la falta de transparencia en algunas de sus aplicaciones, suscita inquietudes sobre la responsabilidad algorítmica y la posibilidad de sesgos incorporados que podrían perpetuar o amplificar desigualdades existentes. Más allá de la manipulación o la desinformación, la preocupación central radica en la posibilidad de que estos sistemas desarrollen capacidades emergentes no programadas, llevando a resultados que desafíen nuestras expectativas. El temor a una autonomía total, en este contexto, no se limita a la independencia operativa, sino que se extiende a una posible independencia cognitiva, donde la IA, a través de su aprendizaje continuo, podría evolucionar de maneras que escapan a nuestra comprensión y, por lo tanto, a nuestra capacidad de intervención efectiva, lo cual plantea un dilema existencial sobre el verdadero papel del ser humano en un mundo cada vez más mediado por inteligencias artificiales complejas.

II.2 Bases conceptuales de la dicotomía

A pesar de ser miradas antagónicas sobre lo tecnológico, tanto el pesimismo como el optimismo se basan en la misma dicotomía conceptual que distorsiona nuestra comprensión de la naturaleza de los sistemas tecnológicos (Tiles y Oberdiek, 2014). Me refiero a la polaridad que concibe la tecnología, o bien como una entidad completamente sujeta

al control humano, o bien como una fuerza inherentemente autónoma e incontrolable. A continuación, veremos algunas implicaciones de esta dicotomía y exploraremos sus posibles causas.

En el caso del tecno-optimismo, encontramos una postura ontológica particular que conviene, al menos, esbozar. Lo primero es que parece asumir una cierta pasividad en los componentes de nuestro universo tecnológico, en particular, cuando se afirma que las funciones que pueden desempeñar dependen enteramente de nuestras elecciones y acciones. Al respecto, podríamos pensar que son varios los momentos en que es posible determinar el tipo de funcionalidad que cumplirá una herramienta. El más obvio resulta ser la etapa de diseño, donde se conceptualizan sus características, se eligen los materiales y se definen los principios operativos de su funcionamiento. Adicionalmente, y como cualquier persona puede atestiguar, cuando manipulamos una herramienta nos es dado decidir el propósito inmediato de su aplicación, la intensidad con la que la usamos y el contexto específico en el que se inserta (Rodríguez y Maldonado, 2021). Todo ello nos da una impresión de control absoluto que, como hemos visto, el tecno-optimista no teme generalizar.

La idea de la pasividad también se manifiesta en la confianza que solemos depositar en el progreso tecnológico, sobre todo en nuestras visiones utópicas del futuro. En ellas, se asume el cambio tecnológico como una serie ininterrumpida de innovaciones cada vez más avanzadas y superiores, diseñadas para resolver todos nuestros problemas y satisfacer nuestras necesidades (Kurzweil, 2005). Desde la erradicación de enfermedades hasta la superación de las limitaciones biológicas humanas, el tecno-optimismo proyecta un futuro donde la tecnología nos conduce a una utopía de abundancia y bienestar, tanto social como personal. Si surgen problemas, como averías o efectos secundarios inesperados (véase, por ejemplo, Tenner, 1996), la solución es invariablemente buscar más y mejor tecnología, bien sea corrigiendo los diseños de las tecnologías existentes, o bien proponiendo nuevos tipos de tecnologías que corrijan los defectos de las generaciones anteriores (Brynjolfsson y McAfee, 2014). Por supuesto, todo ello solo es posible a condición de que la tecnología pueda ser siempre mejorada en cada etapa de su desarrollo, por lo que, una vez más, ha de quedar sometida a la racionalidad humana, la cual eventualmente discernirá los mejores diseños y las mejores aplicaciones.

La confianza en la racionalidad humana como guía del desarrollo tecnológico es un pilar fundamental del tecno-optimismo. Sin esta convicción, resultaría difícil sostener la idea de que la tecnología, percibida

como maleable y pasiva, pueda desarrollarse en una dirección específica o que, como señalase Aristóteles en su momento, pueda tener siquiera propósito (*Metafísica*, Libro VII, 1044a8; *Física*, Libro II, 199b28-30). En este contexto, la racionalidad suele entenderse como la capacidad para elegir los mejores medios disponibles para alcanzar ciertos fines específicos. Es a esto justamente a lo que aspiran estrategias como la planificación centralizada (Mesthene, 1970; Winner, 1977), o el llamado diseño intencional en el ámbito del desarrollo y la innovación tecnológica (Verbeek, 2005).

A pesar de sus diferencias, estas estrategias comparten la confianza en que el ser humano, a través de su intelecto y capacidad de cálculo, podrá controlar, eventualmente, el devenir tecnológico. Esto implica una visión lineal y predecible del progreso, donde cada innovación se concibe como un paso hacia un futuro mejor, y donde los artefactos tecnológicos estarán depurados de contingencias y efectos no deseados. Se asume que cualquier desafío tecnológico puede ser superado con más ingenio y un enfoque más refinado, consolidando así la idea de que la tecnología es meramente una herramienta al servicio de nuestras aspiraciones más elevadas.

Como era de esperar, el tecno-pesimismo se situará, en lo que respecta a su visión del control y la previsibilidad de la tecnología, en las antípodas del tecno-optimismo. Así, asume que la tecnología, y particularmente los sistemas complejos como la IA, son inherentemente opacos e imprevisibles. En este sentido, su fundamento conceptual reside en nuestra aparente impotencia frente a la tecnología, ya sea porque los sistemas se descontrolan por sí mismos o porque su gran complejidad hace imposible anticipar todos sus efectos.

Ahora bien, una razón poderosa para ver la tecnología de esta manera es el reconocimiento de la complejidad de los sistemas tecnológicos, en particular de aquellos posteriores al siglo XVII, momento en el que en Europa comienza la primera revolución industrial. Como lo describe Mumford (1934) las creaciones tecnológicas a partir de ese momento histórico, sobre todo en Occidente, sufren un proceso de complejización en el que pasaron de estar constituidas por un número reducido de partes a conformar ensamblajes de decenas o incluso cientos de piezas. Todo ello fue posible gracias a la introducción de nuevas fuentes de energía, como el carbón, con su abundancia de energía potencial química que permitió el funcionamiento de máquinas a una escala y complejidad nunca vistas. Estas innovaciones, hicieron posible soñar en su momento con autómatas que realizaran las más complejas y tediosas tareas, lo cual, se suponía, nos liberaría definitivamente de la horrible carga del trabajo manual industrial (Sennett, 2008).

Sin embargo, la contracara de todo esto fue la renovación de un antiguo temor hacia la tecnología, el cual, paradójicamente, creció a la par de las promesas de liberación. La sociedad se enfrentó a máquinas cuya intrincada naturaleza las volvía incomprensibles para la mayoría y, en consecuencia, amenazantes. Las personas comenzaron a temer la pérdida de sus empleos, que fuesen reemplazadas por autómatas o la liberación de fuerzas incontrolables. Este miedo se manifestaba de diversas maneras, por ejemplo, en movimientos como el de los luditas, que, en la Inglaterra de principios del siglo XIX, provocó la destrucción de innumerables telares mecánicos, o en la literatura, con obras que exploraban los peligros de una tecnología desbocada, como *Frankenstein* o *La isla misteriosa* de Julio Verne (1875). Este temor se actualiza hoy en día en el miedo que despierta la IA, reeditando las mismas aprehensiones que surgieron con las tecnologías de la revolución industrial.

Por supuesto, en el tecno-pesimismo la confianza en la racionalidad humana ha desaparecido, ya que, según se argumenta, la creciente complejidad de los sistemas tecnológicos actuales a menudo plantea un desafío a nuestra comprensión. En ese sentido, no solo es cuestión de que puedan volverse ajenos a nuestros designios o preferencias morales, sino que de entrada su operación escapa, en algún grado, a nuestro entendimiento. Esto obedecería a varios factores. En primer lugar, y como cualquiera pueda constatar, la ya señalada complejidad hace difícil para un usuario promedio seguir el funcionamiento de innovaciones como computadoras, celulares y en general casi cualquier tipo de tecnología digital o electrónica. Esta asimetría de información limita la capacidad de los individuos para ejercer un control informado de las nuevas tecnologías. Pero incluso la falta de comprensión y, en consecuencia, la falta de control puede aquejar a usuarios expertos y con la formación adecuada. En particular, en el caso de la inteligencia artificial generativa, donde por tratarse de redes neuronales profundas el procesamiento de la información se ve distribuido entre miles de millones de conexiones y ponderaciones, es imposible para cualquier ser humano entender su funcionamiento en cada etapa.

Se teme que la misma complejidad algorítmica termine por generar los efectos nocivos tan anunciados por los tecno-pesimistas. Se trataría de comportamientos o propiedades de los sistemas, que, si bien no fueron programados ni previstos, surgirían como un resultado emergente a partir de la interacción dinámica y no lineal de sus diferentes módulos o partes. De hecho, los modelos de inteligencia artificial generativa, en general, funcionan así: no fueron programados para «razonar», «escribir poesía» o «resumir textos», de la misma manera en que un programador escribiría un

código específico para cada una de esas tareas, sino que más bien todas esas capacidades emergen a partir de su entrenamiento con grandes volúmenes de texto y del aprendizaje automático de patrones y relaciones.

III. La IA más allá de la dicotomía pesimismo/optimismo

Habiendo explorado las bases conceptuales del tecno-optimismo y del tecno-pesimismo, resulta evidente que ambas posturas, a pesar de su antagonismo, se anclan en una misma raíz: la concepción de la tecnología como una entidad con una agencia polarizada, o bien totalmente controlable, o bien completamente autónoma. Esta dicotomía simplifica una realidad mucho más compleja, limitando nuestra capacidad para interactuar de manera efectiva y ética con los avances tecnológicos, especialmente con la IA. Para trascender esta polaridad, necesitamos adoptar una perspectiva que reconozca el carácter co-constitutivo de la relación entre humanidad y tecnología (Latour, 1993; Malafouris, 2013), y donde ni la una ni la otra sean vistas como agentes pasivos o dominantes por completo. En la presente sección intentaré esbozar los contornos de esta perspectiva.

En lugar de preguntar si la IA es intrínsecamente buena o mala, o si está bajo nuestro control absoluto o totalmente fuera de él, deberíamos centrarnos en las interacciones dinámicas que la IA posibilita, las condiciones específicas de su desarrollo y uso, así como las maneras en que se integra en la sociedad. Esto supone reconocer una serie de aspectos que suelen soslayarse en el debate entre tecno-pesimistas y tecno-optimistas.

El primero, y quizás el más importante, es que las IA generativas, dada sus arquitecturas específicas y limitaciones técnicas inherentes, solo pueden ejecutar una serie de operaciones predefinidas, como transformar datos para encontrar patrones y predecir palabras. En consecuencia, todos sus resultados pueden ser explicados en términos de estas funciones. Veamos, a guisa de ejemplo, el caso del “razonamiento” en una IA generativa.

Recientemente, la interfaz de *Gemini* de Google incorporó una opción que muestra su hilo de razonamiento para generar respuestas. Así, ante la pregunta, ¿Cuál es la raíz cuadrada de un número extenso como 359830495830495? el modelo muestra que siguió los siguientes pasos: (1) trazar un plan para encontrar la cifra, el cual consiste en buscar en línea un motor de búsqueda con capacidad de cálculo, introducir el número y solicitar la operación; (2) refinar el plan; (3) realizar el plan trazado; y (4) verificar el resultado. Esto es revelador por dos motivos. En primer lugar, porque deja claro que lejos de razonar como nosotros para resolver, cuando usamos papel y lápiz, la raíz cuadrada de un número, Gemini sigue

un camino diferente. De hecho, muchas de las formas en que un modelo resuelve problemas o responde preguntas guardan poca relación con los procesos mentales involucrados cuando afrontamos esas mismas tareas. Y lo hace así porque su “razonamiento” no es proceso cognitivo orgánico, sino el resultado de algoritmos complejos y el procesamiento, de acuerdo con esos algoritmos, de una vasta cantidad de datos. La «inteligencia» de un gran modelo del lenguaje como *Gemini* radica en su capacidad para procesar y organizar información (léase datos) a una escala y velocidad inalcanzables para los humanos. Comprender esto nos ayuda a ver la IA como lo que es: un instrumento potente cuyas capacidades son explicables en términos de sus principios operativos y su entrenamiento, sin necesidad de atribuirle conciencia o inteligencia humana.

Esta última consideración nos impulsa a concebir la IA no como un mero sustituto (tecno-optimismo) o competidor (tecno-pesimismo) de la inteligencia humana, sino, más bien, como un reconfigurador de nuestras capacidades cognitivas (Englebart, 1962; Brooks, 2019; Rivera Novoa y Duarte Arias, 2025). Al permitirnos extender nuestra habilidad para procesar y organizar información a escalas y velocidades sin precedente, la IA emerge como un socio cognitivo (Hutchins, 1995). En este sentido, algunas tareas que antes eran exclusivas del intelecto humano, relacionadas con la sistematización y organización de información, pueden ahora ser asistidas o delegadas parcialmente (Suchman, 2007), posibilitando que nuestra atención y energía se redirijan hacia los aspectos más complejos, creativos o relacionales de una tarea. Un ejemplo paradigmático de esta dinámica se observa en el ámbito de la escritura creativa: el empleo de un modelo como Gemini por parte de un autor no suprime la necesidad de la creatividad humana, sino que puede liberar al escritor de labores tediosas como la investigación exhaustiva o la superación de bloqueos creativos iniciales, permitiéndole así concentrarse en la narrativa, la profundidad de los personajes o la transmisión de emociones, es decir, en la escritura misma (Floridi, 2025). De este modo, la IA se consolida como un socio cognitivo cuyas capacidades se entrelazan con las nuestras, configurando nuevas prácticas y modalidades de interacción con lo tecnológico.

Por último, pero no menos importante, encontramos el asunto de la agencia. Al respecto, es importante advertir que en el caso de la IA resulta más realista entender la agencia en términos distribuidos y emergentes³. Como ya se mencionó, la IA no es un agente autónomo con decisiones

³ Si bien el término de “agencia distribuida” lo tomo de Latour (1993), en el presente texto he buscado desarrollarlo en una dirección un tanto diferente. De hecho, el argumento que presento se encuentra más alineado con Suchman (2007).

propias, producto de una superinteligencia o una conciencia propia. Su «comportamiento» es el resultado de una compleja red de interacciones entre sus algoritmos, los vastos conjuntos de datos con los que fue entrenada (que a menudo contienen sesgos humanos), las decisiones de diseño de sus programadores y el contexto social y técnico en el que se implementa. Los posibles «errores» o «sesgos» de un modelo no son actos de una voluntad maliciosa propia, sino propiedades que emergen como producto de la complejidad de sus redes neuronales profundas y de las características inherentes a los datos de entrenamiento que le han sido suministrados.

Esta agencia distribuida implica que la responsabilidad por los resultados de la IA también debería estarlo. Ni el creador puede ejercer un control absoluto sobre todas las posibles ramificaciones de un sistema, ni el usuario puede ser el único responsable de sus consecuencias (Nissenbaum, 1996; Coecklbergh, 2012). Más bien, los efectos de la IA serían el resultado de un sistema sociotécnico donde diseñadores, reguladores, usuarios, desarrolladores de datos y la propia tecnología, con sus propias características, interactúan de formas no-lineales y dinámicas.

Así los diseñadores y desarrolladores tienen la responsabilidad inicial de cómo se concibe el algoritmo, lo cual incluye la selección de arquitecturas y los objetivos de entrenamiento. En cuanto a los datos de entrenamiento, su calidad, volumen y los posibles sesgos presentes en ellos, dependerá de quiénes los etiquetan y preparan, por lo que la responsabilidad también recaerá sobre estas personas. Si un modelo perpetúa estereotipos, la raíz puede estar en los datos con los que fue alimentado, no en una «intención» del algoritmo o de sus creadores.

Este recorrido rápido por la agencia distribuida no puede terminar sin mencionar a los usuarios finales y, por supuesto, a la propia tecnología. Esta última, si bien no tiene una agencia propia, sí incorporará en su diseño y materialidad *affordances*, o posibilidades para la acción, que influyen en cómo los usuarios interactúan con ella y en los resultados que se pueden obtener (Norman, 1988; Gaver, 1991). Por ejemplo, un software de edición de video no solo permite editar, sino que su interfaz, las herramientas que ofrece y sus flujos de trabajo preestablecidos guían al usuario hacia ciertas formas de creatividad y producción. En el caso de la IA, el diseño de un algoritmo, la elección de su arquitectura (como una red neuronal transformadora para los llamados grandes modelos de lenguaje como *Gemini*) y los parámetros iniciales de su entrenamiento predefinen de entrada un espacio de posibilidades y limitaciones.

Los usuarios finales estaremos inmersos en esta maraña de condicionamientos, elecciones previas y *affordances*, intencionados o no,

de los sistemas tecnológicos. Pero aun y con todo eso, cada acción, cada entrada de datos y cada juicio que los usuarios hagamos sobre la utilidad o los sesgos de sus resultados añaden una capa a esta compleja urdimbre de agencia.

IV. Conclusión

En este recorrido, hemos desentrañado cómo las posturas del tecno-optimismo y el tecno-pesimismo, a pesar de su aparente oposición, comparten una base conceptual común: la dicotomía de la tecnología como un ente o bien completamente controlable o bien inherentemente autónomo. Hemos visto que el optimismo se nutre, por una parte, de una fe profunda en la racionalidad humana y, por otra parte, en la creencia de que la tecnología es totalmente pasiva, concibiéndola como una herramienta maleable y siempre perfectible, cuyo propósito y bondad residen en la intención de sus creadores y usuarios. Por el contrario, el pesimismo, arraigado en la experiencia de la complejidad y la opacidad de los sistemas tecnológicos modernos, especialmente la IA, nos advierte sobre la incontrollabilidad, los efectos emergentes y la posible amenaza existencial que estas innovaciones conllevan.

Sin embargo, al examinar la inteligencia artificial generativa hemos argumentado que estas visiones dicotómicas son insuficientes. Si bien es cierto que su complejidad, manifestada en su comportamiento de “caja negra” y sus capacidades emergentes no programadas, desafía la noción de un control humano absoluto, al mismo tiempo, su dependencia de diseños algorítmicos y de vastos conjuntos de datos generados por humanos nos impide considerarla como una fuerza completamente autónoma.

La clave para una comprensión más matizada reside en reconocer que la tecnología, y en particular la IA, posee un carácter que podríamos llamar sociotécnico, donde la responsabilidad de sus resultados es compartida. No es simplemente un objeto pasivo ni un agente completamente independiente. En cambio, opera en una zona intermedia, donde sus capacidades y límites están inextricablemente ligados a las prácticas humanas, a los contextos sociales en los que se inserta y a las decisiones de diseño que, paradójicamente, le otorgan un grado de agencia, cuya manifestación más tangible son las posibilidades para la acción que permite. Comprender que el “razonamiento” de un gran modelo del lenguaje, por ejemplo, no es una réplica de la cognición humana, sino un proceso algorítmico de escala masiva, nos libera de la falsa dicotomía que aquí hemos reseñado y nos permite enfocar nuestra atención en las condiciones reales de su empleo. Solo al reconocer

y gestionar esta agencia distribuida podremos desarrollar una relación más madura, ética y sostenible con la tecnología, forjando un futuro en el que los sistemas inteligentes pueda contribuir a tareas complejas sin que ello signifique que hemos abandonado nuestra propia agencia y responsabilidad.

Referencias bibliográficas

- Aristóteles. (1994). *Metafísica* (T. Calvo Martínez, Trad.). Gredos.
- Aristóteles. (1995). *Física* (G. R. de Echandía, Trad.) Gredos.
- Arrow, K. J. (1962). Economic Welfare and the Allocation of Resources for Invention. En National Bureau of Economic Research, *The Rate and Direction of Inventive Activity: Economic and Social Factors* (pp. 609–626). Princeton University Press. <https://doi.org/10.1515/9781400879762-024>
- Bender, E. M., Gebru, T., McMillan-Major, A., y Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *FAccT'21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610-623. <https://doi.org/10.1145/3442188.3445922>
- Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.
- Brooks, R. A. (2019). *Robots and the Human Future*. MIT Press.
- Brynjolfsson, E. y McAfee, A. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. W. W. Norton and Company.
- Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, G., Horvitz, E., Kamar, E., Lee, P., Lee, Y. T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M. T. y Zhang, Y. (2023). *Sparks of Artificial General Intelligence: Early experiments with GPT-4*. Arxiv. <https://doi.org/10.48550/arXiv.2303.12712>
- Cárdenas Arenas, J. C. (2005). Filosofía de la tecnología en Martín Heidegger. *Praxis Filosófica, Nueva serie*, (21), 97-110.
- Chui, M., Hazan, E., Roberts, R., Singla, A., Smaje, K., Sukharevsky, A., Yee, L. y Zimmel, R. (2023). *The Economic Potential of Generative AI: The Next Productivity Frontier*. McKinsey & Company.
- Coeckelbergh, M. (2012). *Human Being @ Risk: Enhancement, Technology, and the Evaluation of Vulnerability and Agency*. Springer. <https://doi.org/10.1007/978-94-007-6025-7>
- Danaher, J. (2022). Techno-optimism: an Analysis, an Evaluation and a Modest Defense. *Philosophy and Technology*, 35(54). <https://doi.org/10.1007/s13347-022-00550-2>
- Dotson, T. (2012). Technology, choice and the good life: Questioning technological liberalism. *Technology in Society*, 34(4), 326–336. <https://doi.org/10.1016/j.techsoc.2012.10.004>
- Ellul, J. (1960). *La Technique ou l'Enjeu du siècle*. Armand Colin.
- Ellul, J. (1964). *The Technological Society* (J. Wilkinson, Trad.). Alfred A. Knopf.

- Englebart, D. C. (1962). *Augmenting Human Intellect: A Conceptual Framework*. Stanford Research Institute. <https://doi.org/10.21236/AD0289565>
- Floridi, L. (2025). Distant Writing: Literary Production in the Age of Artificial Intelligence. *Minds and Machines*, 35(30). <https://doi.org/10.1007/s11023-025-09732-1>
- Gaver, W. W. (1991). Technology Affordances. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Reaching Through Technology*, 79–84. <https://doi.org/10.1145/108844.108856>
- Han, B-Ch. (2022). *La expulsión de lo distinto*. Herder.
- Heidegger, M. (1976). Nur noch ein Gott kann uns retten. *Der Spiegel*, 30(23), 193–219.
- Huang, S. (19 de septiembre de 2022). *Generative AI: A Creative New World*. Sequoia. <https://sequoiacap.com/article/generative-ai-a-creative-new-world/>
- Hutchins, E. (1995). *Cognition in the Wild*. MIT Press. <https://doi.org/10.7551/mitpress/1881.001.0001>
- Kurzweil, R. (2005). *The Singularity Is Near: When Humans Transcend Biology*. Viking.
- Latour, B. (1993). *We Have Never Been Modern*. Harvard University Press.
- López-Portillo Romano, J. R. (2018). *La gran transición. Retos y oportunidades del cambio tecnológico exponencial*. Fondo de Cultura Económica.
- Malafouris, L. (2013). *How Things Shape the Mind: A Theory of Material Engagement*. MIT Press. <https://doi.org/10.7551/mitpress/9476.001.0001>
- Marcus, G. y Davis, E. (2019). *Rebooting AI: Building Artificial Intelligence We Can Trust*. Pantheon.
- Marwick, A. E. (2013). *Status Update: Celebrity, Publicity, and Branding in the Social Media Age*. Yale University Press.
- Marx, K. (2017[1867]). *El Capital, Libro I: El proceso de producción del capital*. Siglo XXI.
- Mesthene, E. G. (1970). *Technological Change: Its Impact on Man and Society*. Harvard University Press.
- Mumford, L. (1934). *Technics and Civilization*. Harcourt, Brace and Company.
- Nissenbaum, H. (1996). Accountability in a Computerized Society. *Science and Engineering Ethics*, 2(1), 25–42. <https://doi.org/10.1007/BF02639315>
- Norman, D. A. (1988). *The Design of Everyday Things*. Basic Books.
- Nova-Rivera, A. y Duarte Arias, D. A. (2025). Generative Artificial Intelligence and Extended Cognition in Science Learning Contexts. *Science & Education*. <https://doi.org/10.1007/s11191-025-00660-1>
- O’Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.
- OpenAI. (2023). *GPT-4 Technical Report*. ArXiv. <https://doi.org/10.48550/arXiv.2303.08774>
- Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press. <https://doi.org/10.4159/harvard.9780674736061>

- Preston, B. (2013). *A Philosophy of Material Culture: Action, Function and Mind*. Routledge. <https://doi.org/10.4324/9780203069844>
- Rodríguez, D. A., y Maldonado, C. E. (2021). Causal Cognition and Skillful Tool Use. *Techne: Research in Philosophy and Technology*, 25(3), 479-492. <https://doi.org/10.5840/techne2021118148>
- Russell, S. (2019). *Human Compatible: Artificial Intelligence and the Problem of Control*. Viking.
- Sai, S., Gaur, A., Sai, R., Chamola, V., Guizani, M., y Rodrigues, J. J. P. C. (2024). Generative AI for Transformative Healthcare: A Comprehensive Study of Emerging Models, Applications, Case Studies and Limitations. *IEEE Access*, 12, 31078-31106. <https://doi.org/10.1109/ACCESS.2024.3367715>
- Scharff, R. C., y Dusek, V. (Eds.). (2014). *Philosophy of Technology: The Technological Condition: An Anthology* (2nd ed.). Wiley-Blackwell.
- Schumpeter, J. A. (1942). *Capitalism, Socialism, and Democracy*. Harper & Brothers.
- Sennett, R. (2008). *The Craftsman*. Yale University Press.
- Shelley, M. (1818). *Frankenstein; or, The Modern Prometheus*. Lackington, Hughes, Harding, Mavor & Jones.
- Spilioti, T. (2016). Technological Utopianism: A Critical Introduction. En G. Claeys (Ed.), *The Palgrave Handbook of Utopianism* (pp. 1-15). Palgrave Macmillan.
- Suchman, L. A. (2007). *Human-Machine Reconfigurations: Plans and Situated Actions* (2nd ed.). Cambridge University Press.
- Tenner, E. (1996). *Why Things Bite Back: Technology and the Revenge of Unintended Consequences*. Vintage Books.
- Tiles, M. y Oberdiek, H. (1995). *Living in a Technological Culture: Human Tools and Human Values*. Routledge.
- Tiles, M., y Oberdiek, H. (2014). Conflicting Visions of Technology. En: Scharff, R. C. y Dusek, V. (Eds.), *Philosophy of Technology: The Technological Condition: An Anthology* (2nd ed.) (pp. 249-259). Wiley-Blackwell.
- Verbeek, P-P. (2005). *What Things Do: Philosophical Reflections on Technology, Agency, and Design*. Pennsylvania State University Press. <https://doi.org/10.1515/9780271033228>
- Verne, J. (1875). *L'Île mystérieuse*. Hetzel.
- Winner, L. (1977). *Autonomous Technology: Technics-out-of-Control as a Theme in Political Thought*. MIT Press.
- Zhang K., Meng X., Yan X., Ji J., Liu J., Xu H., Zhang H., Liu D., Wang J., Wang X., Gao J., Wang Y. G. S., Shao C., Wang W., Li J., Zheng M. Q., Yang Y. y Tang Y. D. (2025). Revolutionizing Health Care: The Transformative Impact of Large Language Models in Medicine. *Journal of Medical Internet Research*, 27(1), e59069. <https://doi.org/10.2196/59069>

Datos de financiación del artículo

El autor declara que no recibió financiación para este artículo.

Implicaciones éticas

El autor no tiene ningún tipo de implicación ética que se deba declarar en la escritura y publicación de este artículo.

Declaración de conflicto de interés

El autor declara que no tiene ningún conflicto de interés en la escritura o publicación de este artículo.

Contribuciones del autor

Escritura (borrador original), escritura (revisión del borrador y revisión/corrección).

20 Autor de correspondencia

Dairon Alfonso Rodríguez Ramírez. darodri@uis.edu.co. Carrera 27 Calle 9 Bucaramanga, Santander.

Declaración de uso de inteligencia artificial

El autor declara que no utilizó ningún programa o aplicación de inteligencia artificial.