
Inteligencia artificial y fe. Niveles de interpelación crítica desde la perspectiva de la razón creyente

GONZALO AEMILIUS *

Pontificia Universidad Gregoriana – Instituto de Antropología y
Centro Hurtado

g.aemilius@unigre.it

Recibido 30.04.2025/ Aprobado 03.06.2025

ORCID: <https://orcid.org/0009-0005-9753-7000>

DOI: <https://doi.org/10.46553/teo.62.147.2025.p205-233>

ESTEBAN MADRID PÁEZ **

Pontificia Universidad Católica Argentina

estebanmadrid@uca.edu.ar

Recibido 30.04.2025/ Aprobado 03.06.2025

ORCID: <https://orcid.org/0009-0003-8704-921X>

DOI: <https://doi.org/10.46553/teo.62.147.2025.p205-233>

RESUMEN

El presente artículo propone una interpelación crítica sobre la Inteligencia Artificial (IA) desde la perspectiva de la razón creyente, delineando tres niveles progresivos de análisis. El primer nivel ofrece un *acercamiento empírico-descriptivo* de los «fru-

* El coautor de este artículo es Presbítero uruguayo, doctor en Teología por la Pontificia Universidad Gregoriana (PUG). Docente en el Instituto de Antropología y en el Centro Hurtado (PUG) desde 2019 y en la Facultad de Teología del Uruguay desde el 2025. Entre 2020 y 2023, ocupó el cargo de Secretario Personal del Papa Francisco. Actualmente está realizando un postdoctorado en Teología e Inteligencia Artificial. Sus publicaciones incluyen la coautoría de “La Inteligencia Artificial: Una mirada pastoral desde América Latina y el Caribe” en marzo de 2025.

** El coautor de este artículo es Presbítero de la Arquidiócesis de Paraná es delegado del Vicerrectorado de Integración del saber de la Universidad Católica Argentina sede Paraná. Doctor en Teología Moral por la Academia Alfonsiana (PUL-Roma). Profesor de Moral y compromiso social en dicha universidad y de Ética filosófica y Moral en el Seminario Arquidiocesano Nuestra Señora del Cenáculo Es miembro de la asociación de moralistas argentinos Morar y del grupo de Investigación teológica Hypsosis (Roma).

tos» del fenómeno, mapeando la omnipresencia de la IA en la vida contemporánea y sus impactos socioculturales, económicos y éticos. El segundo nivel, *epistemológico-hermenéutico*, indaga en el «cuerpo» del fenómeno de la IA como desafiante de nuestras categorías de conocimiento y verdad, explorando las lógicas subyacentes. El tercer nivel, aborda el origen y el impacto de la IA en el *paradigma cultural* (sus «raíces») que ejerce una influencia profunda en los valores y en la concepción del ser humano, y discute sus implicaciones para la dignidad humana, la capacidad creativa y el discernimiento moral. Finalmente, el artículo enfatiza la posibilidad y responsabilidad del *diseño ético* proactivo en la IA y la *esperanza cristiana* como motor para una participación consciente y activa, que transforme el *chronos* devorador en un tiempo nuevo de creación cultural al servicio de la dignidad humana y el bien común.

Palabras claves: Inteligencia artificial; Ética de la IA; Teología de la tecnología; Antropología cristiana; Paradigma cultural; Gobernanza algorítmica; Esperanza cristiana

Artificial Intelligence and Faith

Levels of Critical Interpellation from the Perspective of Believing Reason

ABSTRACT

This article proposes a critical interpellation on Artificial Intelligence (AI) from the perspective of believing reason, delineating three progressive levels of analysis. The first level offers an *empirical-descriptive approach* to the «fruits» of the phenomenon, mapping the omnipresence of AI in contemporary life and its sociocultural, economic and ethical impacts. The second level, *epistemological-hermeneutic*, investigates the «body» of the AI phenomenon as challenging our categories of knowledge and truth, exploring the underlying logics. The third level addresses the origin and impact of AI on the *cultural paradigm* (its «roots») that exerts a profound influence on values and the conception of the human being, and discusses its implications for human dignity, creative capacity and moral discernment. Finally, the article emphasizes the possibility and responsibility of proactive *ethical design* in AI and *christian hope* as an engine for conscious and active participation, transforming the devouring *chronos* into a new time of cultural creation in the service of human dignity and the common good.

Keywords: Artificial Intelligence; AI Ethics; Theology of Technology; Christian Anthropology; Cultural Paradigm; Algorithmic Governance; Christian Hope

Introducción: relevancia y pertinencia de un diálogo complejo

La irrupción de la inteligencia artificial (IA) en los diversos pliegues de la vida contemporánea ha desatado un inédito nivel de atención y debate público, suscitando las más variadas narrativas con sus variopintas imágenes, sentimientos

y preguntas. La bibliografía científica y ensayística sobre el tema ha experimentado un crecimiento exponencial en los últimos años, reflejando una creciente (omni)presencia cultural que es patente en nuestras sociedades.¹ En un mundo que parece vivir una verdadera «primavera de la IA», la aceleración tecnológica y su impacto social y antropológico no dejaron indiferente tampoco al pensamiento teológico ni al magisterio eclesial.

La intervención del Papa Francisco en el encuentro con los líderes del G7 en Apulia, la nota en conjunto de los dicasterios para la Doctrina de la Fe y para la Cultura y la Educación (*Antiqua et Nova*), el documento del CELAM (*IA: Una mirada pastoral desde América Latina y el Caribe*), así como el significativo mensaje inaugural del Papa León XIV (explicando su mismo nombre), dan testimonio de una creciente conciencia eclesial de la relevancia de la cuestión así como de la necesidad de participar activamente en la interpretación y el examen prudente de las consecuencias y del modo de gestionar y diseñar la revolución tecnológica en curso.

El interés por la IA no se reduce a una toma de posición externa o reactiva. La reflexión teológica no quiere actuar por oportunismo, ni para mantener presencia en la agenda mediática, sino porque ve en esta transformación epocal una interpelación a su vocación más profunda: anunciar el Evangelio en las coordenadas del mundo actual, acompañando al hombre y a la mujer en sus gozos, dolores y desafíos (cf. GS 1). Tampoco nace solo del deseo de ofrecer respuestas sino, ante todo, de cultivar una escucha activa y lúcida, capaz de acompañar las

¹ Actualmente la IA está presente en todos los dominios del quehacer social y ya se la considera fundamental para delinear el futuro. Tal es así que se prevé que el mercado señalado por esta tecnología crezca hasta los 190.000 millones de dólares en este 2025 con una tasa de crecimiento anual compuesto superior al 36% en lo que va del 2018 y 2025. Cf. Yuchen Jiang et al, «Quo vadis artificial intelligence?» en *Discover Artificial Intelligence* 2/4 (2022). Para visualizar la creciente presencia e impacto cultural, cf. Nestor Maslej et al., «The AI Index 2024 Annual Report» (Stanford: Stanford University, 2024), acceso el 10 de mayo de 2025, <https://arxiv.org/pdf/2405.19522>.

nuevas preguntas por el sentido, el futuro, el sufrimiento y el poder. Frente a esta novedad disruptiva, la tradición creyente tiene la oportunidad –y la responsabilidad– de ofrecer un pensamiento que, sin clausurarse ante los cambios y sin claudicar de su historia sapiencial, ayude a discernir y a seguir dando cuerpo (visibilidad) y cuerpo de sentido (credibilidad) a la Palabra de Dios en un mundo con profundas y rápidas transformaciones (cf. GS 4).

Fuerza crítica del fenómeno

Tenemos una nueva tecnología que irrumpió con la pretensión de replicar procesos «inteligentes» –por decirlo de alguna manera– con capacidad superlativa de acceder a datos, al análisis de estos, con los que toma decisiones, pero principalmente aprende a modificar autónomamente su «comportamiento» y generar nuevas acciones. Implementación, efecto y repercusión que se impone (¿deterministamente?) con agenda propia y que diseña y transforma nuestros sistemas cognitivos, educativos, laborales, legales, democráticos, políticos, sanitarios, bélicos, etc. Aún más, nuestra idea misma de individuo, de persona, la forma en que nos relacionamos y en la que configuramos nuestro mundo e interactuamos con él. Realidad que inexorablemente despierta viejos como nuevos interrogantes: «¿Nos harán estas tecnologías más poderosos y hábiles o, por el contrario, nos obligarán dentro de espacios físicos y conceptuales más limitados, obligándonos silenciosamente a adaptarnos a ellos, ya que ésta será la mejor manera, y a veces la única, de hacer que las cosas funcionen? ¿Nos ayudarán a resolver los problemas sociales y medioambientales más acuciantes o acabarán agravándolos?». ²

² Luciano Floridi, *The fourth revolution* (Oxford: Oxford University Press, 2014), 6. (La traducción es propia).

Riquezas y límites del aporte teológico

Tan importante como las preguntas es importante explicitar desde dónde estas quieren ser abordadas y la pretensión o no que tengamos en su tratamiento. De ahí que nos parezca pertinente explicitar que no hablamos desde una *expertise* técnica acerca de la IA. No somos ingenieros ni desarrolladores por lo que el lector podrá encontrar en esta reflexión un cierto tono de *outsiders*. Esta condición, sin embargo, no habilita a permanecer al margen. Es importante saber de lo que hablamos, cómo funciona y por qué funciona; su impacto y como afecta la cotidianeidad de la vida en sociedad. Realidad que, sin embargo, no se reduce a aspectos meramente técnicos. Las tecnologías y, especialmente las tecnologías 4.0 con su pretensión omniabarcativas reclaman miradas también omnicomprensivas capaz de dilucidar lo que está en juego. De ahí que realicemos esta reflexión confiados en que miradas externas pueden ser enriquecedoras para el debate contemporáneo como lo ilustran numerosas experiencias de encuentro y discusión entre diversos actores de la sociedad civil, política, religiosa: el diálogo en el G7 (Apulia), en la Alianza para la gobernanza de la IA (en el *World economics forum*), los principios éticos comunes del *Rome calls for ethics*, etc.

La irrupción de la IA nos interpela a una reflexión que trasciende la mera tecnificación o la eficiencia productiva. La tecnología, y por ende la IA, nunca es neutral³. Siempre conlleva una dimensión antropológica y ética inherente, pues es producto y reflejo de la acción humana y sus intenciones (cfr. *Gaudium et spes*, 33-34; *Caritas in veritate*, 69).

Nuestra perspectiva está marcada desde una reflexión ética y teológica fundamental. Este enfoque epistemológico posibilita asumir el fenómeno con una singularidad especial que autoriza a realizar un enriquecimiento circular: la profun-

³ Cf. Stephanie Hare, *Technology Is Not Neutral* (London: London Publishing Partnership, 2022).

dización de la lectura sobre el impacto antropológico y sus (eventuales) consecuencias para la teología y para la vida del creyente en nuestro tiempo. En efecto, como recordó el renombrado teólogo canadiense Bernard Lonergan una de las principales funciones de la teología radica en realizar el ejercicio creyente de mediación «entre una determinada matriz cultural y el significado y función de una religión dentro de dicha matriz».⁴ Lo cual implica, tal como ha sostenido el teólogo francés Joseph Moingt que el pensamiento de la fe no refiere o reclama solamente confrontarla con la fe de los orígenes, sino que invita a volverla inteligible para sus contemporáneos, es decir, confrontar la fe con la razón de la época en la cual se vive.⁵

La propia dinámica creyente nos impulsa a realizar este ejercicio no como algo externo o forzado sino como una instancia inmanente de sí misma porque le posibilitará al pensamiento creyente la oportunidad de ser «expuesto» en un nuevo escenario de sentido que reclama y pide interlocutores capaces de llevar adelante dicho ejercicio. Para poder pensar la fe, pero, a su vez, para ofrecer la *sapienza* que esta posee en el conjunto de saberes. El axioma-base que configura este abordaje nace de la conciencia y certeza que «ninguna época está privada de la gracia, al contrario, cada una es una puerta que se abre al misterio cristiano».⁶

Una propuesta en tres pasos

Asumiendo esta perspectiva ética-teológica fundamental, proponemos mapear tres niveles necesarios para abordar y reflexionar sobre el fenómeno disruptivo de la IA. En este sentido, la teolo-

4 Bernard Lonergan. *Método en teología* (Salamanca: Sígueme, 19882), 9.

5 A modo de ejemplo de lo anteriormente citado cf. Joseph Moingt, *Croire au Dieu qui vient I. De la croyance à la foi critique* (Paris: NRF-Gallimard, 2014.), 33-34, 176-177.

6 Elmar Salmann, *Contro Severino* (Casale Monferrato: Edizione Piemme, 1996), 316.

gía puede ayudar a ofrecer un marco de complejidad que distinga entre los *frutos* visibles, el *cuerpo* de su funcionamiento y las *raíces* epistémicas y culturales que la sustentan. Tres niveles de abstracción que parecen indispensables para evitar respuestas simplistas o moralizantes y, al mismo tiempo, abrir la posibilidad de una palabra antropológica y teológica que pueda resultar significativa. Así, intentaremos llegar con cierta mirada lúcida a los núcleos donde se gestan los nuevos relatos y paradigmas.⁷ Creemos importante no apelar solamente a las transformaciones que se están gestando sino al surgimiento de una nueva *episteme* que implica asumir una nueva forma de diseñar el *humanum*.

Por tanto, este artículo no pretende abordar de modo exhaustivo los múltiples enfoques e innumerables temáticas a desarrollar en torno a la IA. Tampoco busca publicitar (o denostar) lo que ya hoy está por sí mismo profusamente publicitado por los propios desarrolladores y medios de comunicación. En cambio, consideramos esencial poder comenzar a mapear y exponer (tímidamente) la complejidad y ambivalencia de la IA, especialmente desde una perspectiva que no eluda las preguntas fundamentales sobre el sentido, el destino humano y la trascendencia. Lejos de ofrecer narraciones simplificadoras, ya sean consoladoras o apocalípticas, este texto busca acoger el tiempo presente con sus paradojas, oportunidades y desvelos, para contribuir a la construcción de un mundo que reclama inventiva y honestidad de análisis.

El diálogo propuesto se estructura a partir de lo que denominamos *niveles de interpelación crítica*, entendidos como lentes de aumento progresivo a través de las cuales se analiza el fenómeno de la IA: desde la experiencia concreta hasta las implicaciones más trascendentales. Así, el *primer nivel* de análisis se centrará en una aproximación *empírico-descriptiva*, delineando las manifestaciones y el impacto tangible de la IA en la sociedad. La preocupación ética

7 Cf. Francisco, *Exhortación Apostólica Evangelii gaudium*, n. 74.

aquí se manifestará en la identificación de los desafíos y riesgos que emergen de su aplicación concreta.

El *segundo nivel*, de carácter *epistemológico-hermenéutico*, explorará cómo la IA desafía nuestras categorías de conocimiento y comprensión. En este punto, la reflexión se enriquecerá con aportes del pensamiento estructuralista y posestructuralista, lo que nos permitirá analizar cómo la IA, en tanto sistema de procesamiento de información, revela y a la vez oculta estructuras de sentido, poder y valores inherentes a los datos y algoritmos. La dimensión filosófica y teológica se manifestará en la interrogación sobre la naturaleza de la inteligencia y la verdad, y las limitaciones de una racionalidad puramente algorítmica.

Finalmente, el *tercer nivel*, de *teología y ética* estrictamente «*fundamental*» (reflexiones de segundo orden) abordará la irrupción de la IA a nivel de «paradigmas» como aquel «soberano oculto» que no solo moldea nuestras herramientas, sino también las mentes y relaciones y, en definitiva, nuestra visión del mundo, de Dios, de la salvación y del ser humano. Aquí, la reflexión se centrará en las implicaciones más abarcadoras de la IA para la antropología, la teología de la creación, la moral, la soteriología y la escatología, buscando discernir los caminos para orientar la IA hacia el florecimiento humano y el bien común, a la luz de la revelación y la esperanza cristiana.

1. Primer nivel: los frutos de la IA que nos movilizan

La bibliografía y el discurso público sobre ética de la IA se suele concentrar en lo que podríamos llamar sus «frutos» visibles. Son las novedades tecnológicas con sus implicaciones prácticas, las transformaciones concretas en economía (con sus poten-

ciales de descentralización y de control),⁸ en política,⁹ en ámbito de educación,¹⁰ salud,¹¹ comunicaciones y movilidad,¹² etc.

Surgen una infinidad de casos críticos a estudiar. Ha despertado así el interés ético con cuestionamientos que obligan a reformular viejos dilemas: ¿qué es lo específicamente «humano» de nuestra condición? ¿Dónde descansa la dignidad de *sapiens*? ¿qué constituye nuestra libertad? ¿es posible delegar el juicio moral? ¿han surgido nuevos «sujetos morales»? En definitiva, la IA toca el núcleo de nuestra humanidad, de sus dilemas (dolor, trabajo, sociedad) y de su redención: la posibilidad y el camino de la felicidad.

Algunos de los problemas pueden compararse con antiguas cuestiones y enfrentarse con «sabiduría práctica» (prudencialidad) según la dinámica de la analogía (como el cy-

8 En el *sector financiero*, por ejemplo, la IA se utiliza para el análisis de mercados, detección de fraudes, gestión de riesgos, optimización de carteras de inversión. Algoritmos predictivos asisten en la toma de decisiones, mientras los *chatbots* mejoran la atención al cliente. Distintas industrias han incorporado la robótica colaborativa y la IA para mejorar la eficiencia, la calidad y la seguridad en las fábricas inteligentes (Industria 4.0).

9 La IA se ha convertido en una herramienta crucial en el ámbito de *seguridad y de defensa*. Desde el reconocimiento facial y la videovigilancia inteligente para la seguridad pública, hasta la ciberseguridad para detectar amenazas informáticas. En el ámbito *militar*, el desarrollo de sistemas de armas autónomas (armas letales autónomas, LAWS) es un área de intenso debate ético y político, dado su potencial para tomar decisiones sin intervención humana directa. También en la *Administración Pública* y en la *Justicia* la IA se utiliza para optimizar la prestación de servicios, desde la gestión de trámites hasta la asignación de recursos. En el ámbito judicial, se experimenta con IA predictiva para asistir en decisiones de fianzas o sentencias, aunque con serias preocupaciones sobre la equidad y los sesgos.

10 La IA transforma la *educación*. Se emplea en plataformas de aprendizaje adaptativo, que personalizan el contenido y el ritmo de estudio según las necesidades de cada estudiante. Los tutores virtuales y los sistemas de evaluación automatizada liberan tiempo a los educadores, aunque también plantean preguntas sobre la calidad de la interacción pedagógica. Los análisis de datos impulsados por IA pueden identificar patrones de aprendizaje y riesgo de abandono escolar.

11 En el ámbito de la medicina, la IA está revolucionando también el diagnóstico, la investigación farmacéutica y la medicina personalizada. Algoritmos de visión artificial pueden detectar patrones en imágenes médicas (radiografías, resonancias, patologías) con una precisión comparable o incluso superior a la de los especialistas humanos, facilitando la detección temprana de enfermedades y su accesibilidad. En el descubrimiento de fármacos, la IA acelera la identificación de moléculas candidatas y la simulación de sus interacciones. La medicina de precisión utiliza la IA para analizar grandes volúmenes de datos genéticos, historial clínico y estilo de vida, permitiendo tratamientos adaptados a las características individuales de cada paciente.

12 El impacto en la movilidad es prometedor. Los vehículos autónomos son quizás el ejemplo más visible de la IA en el transporte, prometiendo una reducción drástica de accidentes y una mayor eficiencia del tráfico. Además, la IA optimiza las rutas de entrega, gestiona flotas de transporte público y contribuye al desarrollo de sistemas de control aéreo más seguros.

berbullying, cyberstalking, ciberadicción, que se equiparan a violencias y adicciones ya conocidas). Otros, en cambio, son nuevos, como los que surgen de la *agency* de las máquinas de los asistentes con IA¹³ o los sistemas de decisión autónomos que exponen la humanidad a cuestiones y a preguntas apenas formuladas.¹⁴

A modo de ejemplo, uno de los problemas más documentados es la amplificación de sesgos existentes en los datos de entrenamiento. Si un algoritmo se entrena con datos que reflejan desigualdades históricas (raciales, de género, socioeconómicas), el sistema de IA puede perpetuar y hasta exacerbar estas discriminaciones. Esto se ha observado en sistemas de reconocimiento facial con menor precisión para minorías étnicas, en herramientas de contratación que descartan currículums de mujeres o minorías, o en sistemas judiciales que recomiendan sentencias más duras para ciertos grupos demográficos.¹⁵ La discriminación algorítmica es un efecto palpable y preocupante de la IA que exige una intervención ética urgente.

Frente a desafíos tan acuciantes, la algorética se ha concentrado principalmente en las implicaciones prácticas (de responsabilidad ética) y en los desafíos de diseño que conlleven. Un enfoque principalmente orientado a quienes tienen la tarea de desarrollar tecnología (*AI developers*), a quienes de-

13 Los debates entorno a la moralidad de las *Agency* se han vuelto cruciales. Cuando un sistema de IA comete un error o causa un daño (ej. un coche autónomo en un accidente, un algoritmo de diagnóstico que falla), ¿quién es el responsable? ¿El desarrollador, el usuario, el propio algoritmo? La cadena de responsabilidad se vuelve difusa, lo que plantea un desafío legal y ético fundamental para la atribución de culpas y la compensación de víctimas.

14 La capacidad de la IA para procesar y analizar vastas cantidades de datos personales plantea, por ejemplo, serias interrogantes sobre la privacidad individual. Los sistemas de vigilancia masiva, el rastreo de comportamientos en línea y el perfilamiento automatizado pueden erosionar las libertades civiles y crear una sociedad de la vigilancia. El desarrollo y control de la IA están altamente concentrados en unas pocas grandes empresas tecnológicas; que inclusive logran sobrepasar a los gobiernos. Esta concentración de poder puede llevar a monopolios de la información y la capacidad de influir en la opinión pública a través de la manipulación algorítmica de contenidos (ej. noticias falsas, cámaras de eco). Esto representa un riesgo significativo para la pluralidad democrática y la participación ciudadana.

15 Cf. Bernd Carsten Stahl, Doris Schroeder, Rowena Rodrigues, *Ethics of Artificial Intelligence. Case Studies and Options for Addressing Ethical Challenges* (Cham: Springer Open, 2023), 25-37.

tentan la responsabilidad de legislar o deben tomar decisiones políticas (*governance*) al respecto. Marcados por estos intereses, las propuestas privilegian, principalmente, enfoques deontológicos y consecuencialistas, soliéndose detener en la búsqueda de regulaciones prácticas o, al máximo, de una *ética soft* que sirva de base común para las discusiones que se llevan a cabo.

Desde una perspectiva empírica, la Doctrina Social de la Iglesia (DSI) ofrece una lente valiosa para identificar y describir estos impactos. Su enfoque en la *dignidad de la persona humana* como centro y fin de toda actividad social y económica (cfr. *Gaudium et spes*, 26) nos alerta sobre cualquier uso de la IA que instrumentalice al ser humano o lo reduzca a un mero productor o consumidor de datos. La DSI también enfatiza la *centralidad del trabajo* y su contribución a la realización personal y al bien común (cfr. *Laborem Exercens*, 6). El desplazamiento laboral masivo o la precarización del trabajo impulsada por la IA son, por tanto, preocupaciones éticas directas que se observan ya en este nivel descriptivo. Asimismo, el principio del *destino universal de los bienes* (cfr. *Populorum Progressio*, 22) y la preocupación por los *más vulnerables* nos llevan a cuestionar la distribución inequitativa de los beneficios y riesgos de la IA, identificando a aquellos que pueden quedar marginados o excluidos por su desarrollo. Cuestión que nos impulsa a dar un paso más y señalar que, inclusive, la accesibilidad en este nuevo ecosistema no genera –por sí misma– inclusión; ya que existe un efecto y un uso residual de la misma tanto en el proceso de desarrollo (entrenar el algoritmo) como en su aplicación. Hablar de marginación o de categorías como integrados/excluidos desde las características propias de la IA exige repensar estas realidades. La accesibilidad o el uso de determinada tecnología, a diferencia de los que se podía pensar en el pasado, no es más, por sí misma, un elemento integrador/inclusión o desintegrador/exclusión social.

En este primer nivel, el *status quo* se ve interpelado. La observación de los fenómenos descritos genera una inquietud, una «crisis» en el sentido de un punto de inflexión donde las viejas categorías de comprensión de la sociedad y la tecnología comienzan a tambalearse. No es aún un juicio profundo, sino una primera «revelación de la crisis» que nos exige una atención vigilante y una honestidad de análisis ante una realidad en constante mutación. Es un llamado a reconocer que, incluso en la descripción más objetiva, hay una dimensión humana que exige ser considerada. Este enfoque de ética aplicada, aunque necesario, corre el riesgo de quedar en la superficie si no se conecta con niveles más profundos de análisis. La teología –como la filosofía– no debería limitarse a exportar principios a problemas prácticos, sino dejarse interpelar también por los «*insights*» que esos problemas suscitan.¹⁶ La fe, en este escenario, no se presenta como un saber externo, sino como una palabra llamada a ofrecer lucidez y esperanza, que puede evitar tanto la mirada catastrofista paralizante como la utopía ingenua que alimenta viejos y nuevos mitos.

2. Segundo nivel: visibilizar el cuerpo del movimiento crítico

Reconocemos que el nivel de los frutos es el que despierta mayor interés por su efecto *shock* e implicancias concretas, y por lo tanto tiene más marketing y audiencia. Podríamos permanecer *in eterna* en dilemas éticos alucinantes. Sin embargo, más allá de los frutos, es necesario considerar el *cuerpo* de esta transformación: su lógica interna, su arquitectura técnica y su impacto ecosistémico y su pretensión totalizante en el *imprinting* cultural. Nos hallamos ante una reconfiguración de las estructuras cognitivas, afectivas, sociales y simbólicas. La IA –como cualquier otra tecnología en la historia de la humanidad– no produce solo resultados, sino que *mo-*

¹⁶ Cf. Vincent C. Müller, «The History of Digital Ethics», en Carissa Véliz (ed.), *The Oxford Handbook of Digital Ethics* (Oxford: Oxford University Press, 2024), 3-19, 15.

delata y moldea la forma en que percibimos, decidimos y nos vinculamos, con la salvaguarda o, mejor dicho, con la novedad de su carácter eminentemente intangible/“invisible” y su pretensión de omnipresencia. Como afirmaba el papa Francisco, si bien no es posible separar la historia del hombre y de la civilización de la historia de sus instrumentos, la condición tecno-humana nos exige superar la visión instrumentalista para asumir que la tecnología participa de nuestra formación moral y espiritual.¹⁷ En este sentido, resulta clave incorporar la categoría de infósfera (Floridi): aquel entorno donde el ser humano ya no interactúa desde fuera con la tecnología, sino que co-habita con sistemas informacionales.

Si el primer nivel se centró en la descripción empírica de la IA y sus efectos visibles, este segundo nivel se adentra en las implicaciones para nuestra comprensión del conocimiento, la racionalidad y, en última instancia, la propia condición humana. Aquí no solo nos preguntamos *qué es* la IA en su funcionalidad o cuáles son los beneficios explícitos o no que nos ofrece, sino *cómo la conocemos, cómo nos afecta* en nuestra forma de pensar y de interpretar el mundo, y *qué supuestos* epistemológicos subyacen a su diseño y operación. Es un nivel donde la filosofía, la teoría del conocimiento y la hermenéutica se vuelven herramientas indispensables para desentrañar la complejidad del fenómeno.

La IA, especialmente en sus formas más avanzadas como el aprendizaje profundo y los modelos generativos de lenguaje, nos confronta con la cuestión de la racionalidad y sus límites. ¿Puede una máquina «pensar» o «entender» en sentido estrictamente humano? ¿Qué implicaciones tiene para nuestra autocomprensión si las máquinas pueden replicar o incluso superar ciertas capacidades cognitivas que tradicionalmente considerábamos exclusivas del intelecto humano, como la generación de texto, la composición musical o el razonamiento lógico en dominios específicos? Esta discusión

¹⁷ Francisco, *Discurso en la sesión del g7 sobre inteligencia artificial* [14 de junio de 2024] Borgo Egnazia (Apulia - Italia).

nos lleva ineludiblemente a reflexionar sobre la especificidad de la inteligencia humana, que no se reduce a la capacidad de procesamiento de información (computar/predecir), sino que integra otras dimensiones.

2.1. *¿Inteligencia o inteligencias?*

La IA nos obliga, por lo tanto, a enfrentar una pregunta fundamental: ¿qué significa ser inteligente? Si una máquina puede jugar al ajedrez mejor que cualquier humano, o diagnosticar enfermedades con mayor precisión, ¿es esto sinónimo de inteligencia humana? La reflexión epistemológica nos lleva a diferenciar radicalmente entre la inteligencia artificial y la inteligencia humana como *apertura* a un *logos* que sólo puede ser dicho en plural. Un abordaje que a nivel científico declina como transdisciplinariedad e interdisciplinariedad del saber y que se abre metadisciplinarmente a la potencia evocativa de la poesía, el arte y hasta la locura (*demens*) de lo humano.

La inteligencia humana es una inteligencia encarnada, imbricada en la experiencia corporal y emocional.¹⁸ Es consciente, autoconsciente, capaz de reflexionar sobre sí misma y sobre el sentido de la existencia. Se caracteriza por la creatividad genuina (no meramente recombinatoria), la intuición, la sabiduría, la capacidad de discernimiento moral, la empatía y la capacidad de establecer relaciones significativas. La racionalidad humana no admite un modelo único de inteligibilidad; es también narrativa, simbólica, afectiva y moral. Se mueve en el ámbito del *por qué* y del *para qué*, no solo del *cómo*. El ser humano es un ser-en-el-mundo, un ser-para-la-muerte, un ser que busca sentido y trascendencia. Estas dimensiones son experiencialmente inaccesibles para la IA.

¹⁸ En este sentido sugerimos dos obras que pueden orientar mucho: cf. Daniel Kahneman, *Pensar rápido, pensar despacio* (Buenos Aires: Penguin Random House, 2022); Antonio Damasio, *Lo strano ordine delle cose* (Milano: Adelphi, 2018).

Existe un riesgo epistemológico, una tentación reduccionista, de homogeneizar ambas inteligencias. Si la IA puede simular el lenguaje, podemos creer que «entiende». Si puede «crear», podemos pensar que tiene conciencia. Esta confusión, a menudo fomentada por el lenguaje antropomórfico que utilizamos para describir la IA (e.g., “pensar”, “aprender”, “crear”), es peligrosa porque puede llevar a una infravaloración de la especificidad y la dignidad de la inteligencia humana, reduciéndola a una mera máquina de procesamiento de información.

2.2. La IA y las estructuras ocultas: ¿generación o refuerzo?

Desde una óptica que dialoga con el estructuralismo, la IA puede ser entendida como un potente instrumento que, al procesar y analizar volúmenes masivos de datos, potencia las estructuras subyacentes que organizan el mundo, las relaciones sociales y el conocimiento mismo. De hecho, los algoritmos de aprendizaje automático no «inventan» las conexiones, sino que identifican patrones, correlaciones y regularidades que a menudo permanecen invisibles para el ojo humano debido a la escala y la complejidad de los datos.

En este sentido, la IA actúa como un espejo que nos devuelve, por un lado, la lógica implícita de nuestros propios sistemas; y, a su vez, lo procesa con la gramática en la que fue diseñada. Por ejemplo, un algoritmo de recomendación de contenidos, al identificar patrones en nuestras preferencias de consumo o información, no solo nos «sugiere» algo, sino que visibiliza las estructuras de nuestros gustos, intereses y sesgos que nosotros mismos hemos alimentado con nuestros datos codificándolos con su propio diseño algorítmico. Similarmente, en el ámbito científico, la IA es capaz de descubrir nuevas relaciones entre variables en vastos conjuntos de datos biológicos o físicos, revelando estructuras del mundo natural y/o de los estudios ya realizados que antes eran inaccesibles para la investigación humana.

Sin embargo, esta capacidad de «revelación estructural» tiene una doble arista. Por un lado, puede ofrecer nuevas perspectivas y conocimientos que amplían nuestra comprensión. Por otro, al operar con base en estructuras de datos existentes, la IA tiende a naturalizar y reforzar dichas estructuras. Si los datos reflejan desigualdades sociales, patrones de discriminación o ciertas ideologías implícitas (como señalamos anteriormente), la IA los codifica y reproduce, volviendo estas estructuras aún más opacas y resistentes al cambio. La «lógica» que la máquina aprende es, en última instancia, una lógica humana, con sus virtudes y sus falencias. La IA, en este nivel, no es un agente de cambio de estas estructuras, sino un amplificador (silencioso) de las mismas, lo que nos obliga a interrogarnos sobre la naturaleza de las estructuras sociales y cognitivas que estamos codificando y, a su vez, amplificando.

A este punto, podemos vislumbrar con mayor claridad cómo los frutos de la IA provienen y refuerzan una estructura antropológica y social que crece y nos influencia profundamente. Se abre así la necesidad de apelar a un horizonte epistemológico más complejo, con conciencia sistémica y atención a las configuraciones recursivas y a las interdependencias, que ofrece un marco fértil a la reflexión. Un nivel de análisis que permita sostener públicamente que cualquier discurso sobre este fenómeno no es (ni será) ni ingenuo ni neutro,¹⁹ como tampoco un plan diseñado de *governance* global.

2.3. *La ficción de neutralidad*

Si el estructuralismo nos ayuda a ver cómo la IA revela estructuras, el posestructuralismo nos permite acceder a la misma con una

¹⁹ Por ejemplo, Nvidia (GPU - hardware y software - pero sobre todo hw para IA) creció 2800% en 5 años - el valor accionario en 2022 era de USD 300 billones, hoy es de USD 3 trillones y monedas - se multiplicó por 10 en 2 años - llegando a ser la tercera empresa más valiosa a nivel mundial, sólo superada levemente por Apple y Microsoft). Esto nos recuerda que cuando hablamos de IA nos estamos refiriendo también y sobre todo a una industria, un modelo de negocios con millones en juego y no a fenómenos mágicos/ciencia ficción capaces de acallar todos nuestros mecanismos de análisis.

mirada capaz de criticar la supuesta neutralidad de estas estructuras y desvelar cómo la IA, al operar con ellas, puede desestabilizar todas nuestras configuraciones (de sentido, verdad, autoría, espacio-temporalidad, principio de realidad, entre otras). La IA no es un mero reflejo pasivo de datos; es un sistema activo que interpreta, clasifica, prioriza y, por tanto, jerarquiza e incluso crea contenidos que desafían las fronteras tradicionales del conocimiento.

Además, la concepción de la IA puede presuponer una racionalidad «sin espíritu», instituciones supuestamente neutras y resultados obtenidos «mágicamente» sin mediaciones de sentido. Esta crítica resuena fuertemente con la perspectiva posestructuralista, que cuestiona la existencia de una verdad unívoca o un sentido esencial e independiente del lenguaje, el poder y la interpretación. La IA, al generar textos, imágenes o voces de manera convincente, pero a veces con «alucinaciones» (información falsa o inventada que parece plausible), deconstruye la autoridad de la fuente y la estabilidad del significado. Las *deepfakes*, por ejemplo, no solo manipulan imágenes, sino que interpelan la noción misma de lo «real» y lo «auténtico», sembrando la duda sobre la veracidad de las evidencias.

En este contexto, la autoría (valor muy desarrollado en occidente) se vuelve difusa: ¿quién es el autor de un poema generado por una IA? ¿El programador, el algoritmo, los millones de textos con los que fue entrenado? Esta deconstrucción de la autoría también amplifica una deconstrucción que se viene ya realizando desde otras perspectivas del sujeto moderno y su pretensión de originalidad y agencia soberana. La IA, al operar en base a la recombinación de lo ya existente en sus vastos conjuntos de datos, nos obliga a repensar la creatividad, la originalidad y la relación entre el creador y la obra.

Uno de los puntos centrales de esta crítica es la *opacidad algorítmica* o el problema de la «caja negra». Muchos algoritmos de IA, especialmente los de aprendizaje profundo, son tan complejos que incluso sus propios desarrolladores no pueden explicar completa-

mente cómo llegan a sus resultados (trazabilidad).²⁰ Producen una «verdad» o una «respuesta» sin revelar el proceso/método de obtención de la misma (lo que clásicamente llamaríamos razonamiento o inteligibilidad de un proceso). Esta falta de transparencia mina la capacidad humana de comprender, auditar y, en última instancia, confiar en las decisiones de la IA, lo que impacta directamente en la noción de una verdad accesible y dialógica porque compartida y justificable (*accountability*). La asunción de determinadas respuestas sin la explicitación del método (camino) conduce inexorablemente al deterioro del diálogo y de todo lo que esto conlleva. Si no podemos entender *por qué* la IA llegó a una conclusión, ¿cómo podemos evaluar su validez o su equidad? Esto desafía la base misma de la autocomprensión y el lugar que la razón crítica está llamada a adquirir en este nuevo escenario. Desafío que resultará crucial para el pensamiento creyente y su método.

3. Tercer nivel: el paradigma de la complejidad y la IA

Tras haber explorado la manifestación empírica de la Inteligencia Artificial y sus desafíos epistemológicos, descendemos a un tercer nivel de interpelación. Aquí, la IA no es vista meramente como una herramienta o un objeto de estudio cognitivo, sino como un *paradigma* en el sentido de Edgar Morin: un “soberano oculto” que no solo moldea nuestras herramientas y tecnologías, sino que, de manera más profunda, configura nuestras mentes, nuestras relaciones, nuestros valores, nuestra visión del mundo y, en última instancia, nuestra comprensión de lo que significa ser humano y nuestro lugar en la creación. Un *paradigma* es un conjunto de principios fundamentales (conceptos, categorías, axiomas, tipos de racionalidad) que organizan, inconscientemente, el conocimiento y el pensamiento, y que determinan cómo se perciben, conciben y re-

20 Cf. Jenna Burrell, «How the machine ‘thinks’: Understanding opacity in machine learning algorithms», *Big Data & Society* 3/1 (2016): 1-12. Acceso el 26 junio de 2025, <http://dx.doi.org/10.2139/ssrn.2660674>.

suelven los problemas.²¹ La IA, al establecer nuevas lógicas de procesamiento, interacción y decisión, está instalando silenciosamente un nuevo paradigma de la existencia humana.

Cada cultura, cada sociedad posee intrínsecamente racionalidades que les son propias y sobre las que elabora un sinfín de afirmaciones que sustentan distintos mecanismos desde donde se legitima el saber, el poder, la moral, la autoridad, la credibilidad (en la más amplia acepción del término). Para poder dar razones de algo en un contexto determinado (de la fe, de la esperanza, o de lo que sea) es imprescindible acceder a los elementos que hacen posible que un discurso sea creíble (o no) al interior de una episteme determinada. Todo lo que atente o contradiga esa racionalidad, en el mejor de los casos, correrá con el riesgo del descrédito o la desvalorización o se volverá dialógicamente incomprensible.

Creemos imperante conocer y concentrarse en la transformación de los ejes estructurantes y estructurales de los individuos y de las sociedades como ser: la dimensión espacio-temporal y la comprensión que se tenga de estos; el ejercicio de la autoridad, el uso del poder y la justicia, así como los determinantes éticos para definir la valía de una acción.²² Esas racionalidades –casi siempre no evidentes o camufladas en determinadas promesas salvíficas– terminan por configurar las identidades de manera particular y terminan por gestar un modelo antropológico/socio-cultural determinado. No han sido pocas las veces que se ha corrido el riesgo de analizar y discutir sobre modelos antropológicos que configuran los entramados sociales y culturales ignorando las racionalidades y los mitos estructuradores que los posibilitan. Una palabra que quiera ser significativa no puede pasar solamente por la discusión fenome-

21 Cf. Edgar Morin, *Introducción al pensamiento complejo* (Barcelona: Gedisa, 2012³).

22 En este sentido, a modo de ejemplo, es interesante recordar como «les modèles sociaux et mondains auxquels nous restons encore agrippés par habitude ont un temps de retard sur nos façons de vivre et d'habiter un monde dans lequel distance et durée se transforment de plus en plus en proximité et instantanéité». Philippe Joron, «La prime à l'émotionnel» (Paris: François Bourin Editeur, 2012) 25. Se recomienda a su vez cf. Marc Augé, *Los no lugares. Espacios del anonimato* (Barcelona: Gedisa, 2000), 35-43.

nológica de los acontecimientos (principalmente el primer nivel de esta reflexión, pero también parte del segundo nivel) sino que sea capaz de pensar o repensar las formas que los constituyen.

Uno de los elementos determinantes en este nivel radica en si se mantiene o no una mirada capaz de albergar la riqueza infinita del ser humano, del mundo y de Dios, en su complejidad y variedad. Camino recorrido que nos lleva a preguntarnos si los presupuestos más hondos en las raíces del humanismo creyente se ven o no trastocados por un riesgo de homogeneización y simplificación oculta. Tenemos una llamada a repensar el presente y el futuro de nuestros mundos cada vez más tecnologizados/digitalizados donde las generaciones más jóvenes están naciendo. Hoy nos encontramos con una nueva y específica mediación de sentido (racionalidad) y, por tanto, de la realidad. La *reontologización* de la situación actual necesita también una necesaria *reepistemologización* adecuada a lo que acontece a fin de obtener una representación más clara de nuestro tiempo y, con ello, las condiciones adecuadas para darle la mejor forma y abordar con éxito sus problemas abiertos.²³ Nadar contra corriente no equivale a agitarse en el agua de forma confusa y caótica. Al contrario, el rigor se hace aún más imprescindible ya que una reflexión superficial no llevaría más que a agravar los problemas. La transformación epocal es profunda y el riesgo de moralizar el tiempo presente es alto. Actitud que no ayudaría a descubrir nuevas mediaciones para que la lectura de este tiempo sea en clave kairológica.²⁴

En la raíz de la transformación que introduce la IA hay un cambio de paradigma epistémico y antropológico que utiliza un

23 Cf. Luciano Floridi, *The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities* (Oxford: Oxford University Press, 2023), 243 pp.

24 En el fondo de esta perspectiva y de la presente reflexión se intenta responder a una de las invitaciones de la Constitución Apostólica *Veritatis Gaudium* cuando convoca a los estudios eclesásticos a que no se limiten: «a transmitir a los hombres y mujeres de nuestro tiempo, deseos de crecer en su conciencia cristiana, conocimientos, competencias, experiencias, sino que deben adquirir la tarea urgente de *elaborar herramientas intelectuales* que puedan proponerse como paradigmas de acción y de pensamiento, y que sean útiles para el anuncio en un mundo marcado por el pluralismo ético-religioso» VG 5 (el cursiva es propia).

lenguaje y una lógica determinada y que, por tanto, termina imponiendo su gramática en todos los dominios. Como hemos podido subrayar, la pretensión omnicomprendensiva de la IA no se desarrolla solamente en el nivel de los frutos y en su aplicabilidad en todos los sectores de la vida humana, sino que su irrupción se vuelve trascendental al configurar paradigmáticamente el corazón de la *episteme*. No se trata solo de algoritmos y máquinas, de potencialidades superlativas y escenarios futurísticos de ciencia ficción; sino de cosmovisiones y lenguajes que reconfiguran la manera en que entendemos la realidad, el conocimiento, el poder y la verdad. La fuerza radical de este tipo de tecnologías e inclusive, el menos visibilizado de sus poderes, a la hora de pensar el impacto masivo que tiene en el entramado social reside también en su performatividad. Creemos pertinente recordar que «el poder que opera a través de la costumbre es más eficiente y más estable que el poder que emite mandatos o ejerce coerciones».²⁵ La IA encarna –y a la vez propaga– una visión del mundo basada en la predicción estadística, en la eficiencia operacional y en la modelización de la complejidad a patrones legibles.

Esto plantea una seria interpelación al *ethos* contemporáneo. El sujeto habitual de la ética, con su sabiduría práctica (la *phronesis* griega - la *prudentia* clásica) que la tradición ética y bíblica cultivó parece ceder su “lugar ideal” a un sujeto decisonal signado por sistemas algorítmicos que reposan en una racionalidad (aparentemente) acorpórea, ahistórica y acultural. De ahí el desafío de transformar la búsqueda de meros *datos históricos* a *datos híbridos* o al “ideal” de *datos* totalmente *sintéticos*, es decir, datos duraderos, reutilizables, rápidamente transportables, simultáneamente compartidos, sin finalidades y sin rivalidades.²⁶

La cuestión radica en que si presuponemos, amplificamos y autorizamos paradigmáticamente solo esta forma de racionalidad, ¿se podrá modelizar integralmente la experiencia humana? Bajo la

25 Byung Chul Han, *Sobre el poder* (Barcelona: Herder, 2016), 51.

26 Cf. Luciano Floridi, *The Ethics of Artificial Intelligence*, 32-37.

hegemonía de este “soberano oculto”, ¿cómo digitalizar el valor de una presencia amiga en medio del dolor? ¿cómo expresar la fuerza del silencio? ¿Cómo valorar la potencia de los gustos o la memoria de los olores? ¿Cómo traducir a datos (sin necesidad de excluir) las situaciones de tensión, las crisis de las personas y las comunidades, las alegrías del crecimiento y la necesidad de la espera paciente? En todas estas traducciones a datos habrá interpretaciones y reducciones. El problema, entonces, no es que existan grandes modelos de lenguaje (*Large Language Model*), sino si pretenden modelar “neutramente” la totalidad de la experiencia humana, incapacitando o dejando afuera otros lenguajes para decir, también, lo humano. ¿No será oportuno, entonces, desarrollar una teología capaz de reconocer y asumir la fuerza de este “soberano oculto”?

Se vuelve urgente, entonces, identificar y desmitificar los paradigmas subyacentes que configuran la IA haciéndola capaz de callar u ocultar otros lenguajes. La tecnología nace de necesidades, intereses, expresa cosmovisiones, distribuye poder y corre el riesgo de totalizar su pretensión. Desde la reflexión ética y creyente podemos elevar una palabra que ayude a discernir estos procesos y proponga una visión que habilite el lenguaje de la alteridad, que valore la finitud y la gratuidad y posibilite la irrupción de la novedad (frente a la repetición por reforzamiento de las dinámicas mayoritarias).

Desde esta perspectiva, la reflexión teológica y ética fundamental se vuelve indispensable, pues es aquí donde se abordan las grandes preguntas sobre el sentido de la vida, la trascendencia, el bien común y el destino último del ser humano. La IA, como un *signo de los tiempos*, nos exige un discernimiento que integre la luz de la fe para orientar su desarrollo de manera que promueva la dignidad inalienable de cada persona y el florecimiento de toda la creación.

La teología reafirma que el ser humano es único por su vocación a la relación con Dios y por su capacidad de trascendencia. La *Imago Dei* se manifiesta no en una inteligencia meramente computacional, sino en la capacidad de amar, de elegir, de crear con un

propósito y de buscar la verdad y el bien. Ignorar estas dimensiones sería ceder a una antropología de tipo reduccionista que, en última instancia, empobrece la comprensión de lo humano y justifica usos de la IA que atenten contra la persona (cfr. Papa Francisco, *Discurso a los participantes en la Asamblea Plenaria de la Pontificia Academia para la Vida*, 28 de febrero de 2020).

La provocación creyente nos invita a concebir al ser humano como *co-creador* con Dios. La capacidad de innovar, de construir, de transformar la creación es un don divino y una responsabilidad. La IA es, en este sentido, un producto de la inteligencia y la creatividad humanas, un fruto de la capacidad de “nombrar” y “cultivar” la tierra (Gn 2,15). Sin embargo, esta capacidad lleva consigo la responsabilidad de asegurar que lo creado sirva a los propósitos del Creador y al bien de la creación.

4. Balance y perspectivas: diseño y esperanza

En este camino, la reflexión creyente debe abrirse al horizonte de la esperanza. No basta con diagnosticar y realizar una crítica aguda de sus potencialidades y riesgos. Es necesario imaginar, proponer, co-crear: diseñar. El cristianismo no es un pensamiento nostálgico, sino una fuerza de renovación pascual. Por eso, la pregunta decisiva no es si estamos a favor o en contra de la IA, sino cómo podemos contribuir a que su desarrollo esté orientado al bien común, a la justicia y a la dignidad de todos y abra caminos de vida en nuestro tiempo.

El ámbito del diseño ético y técnico es hoy un terreno estratégico donde se juega buena parte del futuro. La participación de los creyentes en estos procesos –desde la investigación académica, la educación, el acompañamiento, la regulación legal hasta la creación de políticas públicas– es un imperativo ético y pertenece a su misión en el mundo de hoy. Ofrecer así una sabiduría que ayude (y nos ayude) a mantener viva la pregunta por el sentido.

En este sentido, la Doctrina Social de la Iglesia tiene mucho que aportar. Principios como la centralidad de la persona, con su condición frágil, corpórea y de unicidad; el destino universal de los bienes, la subsidiariedad y la participación pueden inspirar marcos de gobernanza algorítmica más humanos. Pero, más aún, es necesaria una espiritualidad del discernimiento que permita habitar este tiempo con una confianza lúcida: sin ingenuidad, sin miedo, con creatividad evangélica.

La IA no es el fin de la historia, pero sí un signo de los tiempos. Escuchar ese signo, a la luz de la fe, es parte de nuestra vocación. En este tiempo de la historia, los cristianos estamos llamados a ser testigos de una esperanza que no teme a la novedad, porque confía en que el Espíritu sigue haciendo nuevas todas las cosas (cf. Ap 21,5).

4.1. Hacia una ética y teología del diseño

La ética y la teología en la reflexión sobre la IA no pueden ser un apéndice o un “parche” que se añadan al final del proceso de desarrollo tecnológico. La complejidad de la IA y su capacidad de generar impactos a escala masiva exigen un enfoque “ético-teológico del diseño” (que sea capaz de incluir una reflexión *desde* el diseño -como la *ethics by design*- pero que la supera). Esto requiere una metodología que integre consideraciones ético-teológicas que posibiliten lucidez en los presupuestos paradigmáticos en cada fase del ciclo de vida de un sistema de IA: desde la concepción y la recopilación de datos, pasando por el algoritmo y su entrenamiento, hasta su despliegue, monitoreo y perpetuación de algunas de sus estructuras. Esto implica una colaboración interdisciplinaria y transdisciplinaria constante entre ingenieros, científicos de datos, filósofos, teólogos, sociólogos, juristas, profesionales en los distintos niveles de aplicación y ejecución y los propios usuarios o comunidades afectadas.

Este enfoque proactivo busca incorporar principios éticos como la justicia, la equidad, la transparencia, la privacidad, la responsabilidad y la seguridad desde el inicio. No se trata de imponer restricciones punitivas, sino de fomentar una cultura de desarrollo responsable que anticipe los posibles riesgos y maximice los beneficios. El objetivo es construir una IA centrada en el ser humano (*Human-Centered AI*), donde la tecnología sea concebida como un ecosistema (*enviroment*) capaz de potenciar las capacidades humanas sin esconder sus fragilidades y sombras. Que busque promover el florecimiento integral de las personas, sin potenciar la cultura de la sustitución y el sometimiento. Esto implica, por ejemplo, asumir la inevitabilidad de las mediaciones (por más intangibles que estas parezcan) y sus límites: la perpetuación de sesgos y estructuras, sistemas que no son totalmente explicables y, en consecuencia, difícilmente auditables. El diseño y la gobernanza algorítmica, por tanto, no es solo un tema de regulación, sino un conjunto de reflexiones, prácticas y principios que guíen la innovación misma.

4.2 Interpelaciones de la esperanza cristiana en un tiempo de transformación

En un mundo donde la IA suscita tanto fascinación como temor, la esperanza cristiana –como hemos ya tímidamente anunciado– emerge como una virtud teologal fundamental. Esta esperanza no es una forma de optimismo ingenuo que ignora los desafíos, ni tampoco una pasividad que se resigna a un destino tecnológico inevitable. Al contrario, es una fuerza dinámica que impulsa a la acción, al discernimiento y a la creatividad. Si partimos del presupuesto que «ninguna época está privada de la gracia, al contrario, cada una es una puerta que se abre al misterio cristiano»,²⁷ la esperanza cristiana transforma al presente y al futuro en un espacio ac-

27 E. Salmann, *Contro Severino* (Casale Monferrato: Edizione Piemme, 1996), 316. (La traducción es propia).

tivo de creación cultural. Si bien podemos coincidir en lo inevitable e irrefutable de la irrupción de esta nueva tecnología, el diseño y su desarrollo puede ser visitado y reconfigurado por otros actores valiosos. De ahí que uno de los rostros de la esperanza cristiana en el dominio de la IA está declinado –como acabamos de mencionar– por el diseño estructural y paradigmático de la misma.

No se trata de inventar más historias exitistas –abundan y demasiado– tampoco construir más cuentos apocalípticos –que también abundan y demasiado– sino acoger el tiempo presente con sus complejidades, paradojas, oportunidades y desvelos de construir un mundo que reclama o exige inventiva y honestidad de análisis sobre todo en lo que respecta a tratar de preservar la vida y dignidad de las personas.

Frente a la lógica del “chronos devorador” –un tiempo lineal que avanza sin sentido hacia un progreso sin alma, a menudo impulsado por la eficiencia algorítmica sin consideración humana–, la esperanza cristiana moviliza una reflexión teológica y ética de la responsabilidad que no anticipa solo los posibles daños de la IA, sino que también busca potenciar las oportunidades que esta realidad presenta para el diseño de un mundo más justo, solidario y humano. Esta esperanza nos llama a una participación consciente y activa en la configuración de la IA. Nos capacita para ver en la IA no solo un desafío, sino también una oportunidad para profundizar nuestra comprensión de la creación, de nuestra vocación a ser co-creadores con Dios y de nuestra responsabilidad de custodiar y cultivar el jardín del mundo. La esperanza nos impulsa a la acción concreta, a la creatividad evangélica, para que la IA sea una herramienta que abone el terreno para la vida plena y digna para todos, y no para su degradación o alienación.

La IA es una «fuerza de renovación pascual» en la medida en que nos desafía a una conversión teológica y pastoral, a repensar cómo la fe puede iluminar y guiar la construcción de un futuro donde la tecnología esté orientada al bien, a la justicia y a la dignidad de todos. Esto impulsa al pensamiento teológico a resituar

y reconsiderar, por ejemplo, una antropología teológica capaz de asumir el nuevo vínculo que la tecnología exige a la hora de pensar el *humanum* y la sociedad; a reposicionar el valor de la ciencia y la tecnología como parte del proyecto creador; a confrontar la realidad soteriológica y escatológica con el fuerte nihilismo (silencioso) de la gramática digital o con la ya impuesta cultura transhumanista.

No es menor asumir qué significa la realidad sacramental y sacramentaria de la Iglesia en su virtualidad expuesta hoy también a la digitalización de la misma y pide a la liturgia y a la eclesiología fundamental una toma de posición capaz de custodiar la riqueza de su existencia en momento determinado y determinante de la *forma mentis* del ser humano. Conceptos como ecología integral, formación de la conciencia, libre albedrío, responsabilidad moral - por nombrar brevemente algunos temas que exigen un tratamiento serio y responsable si quieren no abandonar al creyente a la merced de su buena voluntad.

La esperanza cristiana nos permite habitar la incertidumbre del presente con una confianza lúcida: sin ingenuidad, sin miedo, pero con creatividad evangélica. Esta confianza no es una negación de los problemas, sino la convicción de que, aun en medio de las crisis, Dios sigue actuando en la historia y llamándonos a colaborar en la construcción de su Reino.

Bibliografía

Augé, Marc. *Los no lugares. Espacios del anonimato*. Barcelona: Gedisa, 2000.

Burrell, Jenna. «How the machine ‘thinks’: Understanding opacity in machine learning algorithms». *Big Data & Society* 3/1 (2016): 1-12. Acceso el 26 junio de 2025, <http://dx.doi.org/10.2139/ssrn.2660674>.

Damasio, Antonio. *Lo strano ordine delle cose*. Milano: Adelphi, 2018.

Han, Byung Chul. *Sobre el poder*. Barcelona: Herder, 2016.

Hare, Stephanie. *Technology Is Not Neutral*. London: London Publishing Partnership, 2022.

Jiang, Yuchen; Li, Xiang; Luo, Hao; Yin Shen; Kaynak Okyay. «Quo vadis artificial intelligence?» en *Discover Artificial Intelligence* 2/4 (2022). Acceso el 11 junio de 2025, <http://dx.doi.org/10.2139/ssrn.2660674>.

Maslej, Nestor; Fattorini, Loredana; Perrault, Raymond; Parli, Vanessa; Reuel, Anka; Brynjolfsson, Erik et al. «The AI Index 2024 Annual Report». Stanford: Stanford University, 2024. Acceso el 15 junio de 2025, <http://dx.doi.org/10.2139/ssrn.2660674>. en <https://aiindex.stanford.edu/report/>

Floridi, Luciano. *The fourth revolution*. Oxford: Oxford University Press, 2014.

Floridi, Luciano. *The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities*. Oxford: Oxford University Press, 2023.

Francisco, *Evangelii gaudium* [= EG], Exhortación Apostólica, in AAS 105/12 (2013) 1019-1137.

- Francisco, *Discurso en la sesión del g7 sobre inteligencia artificial* [14 de junio de 2024] Borgo Egnazia (Apulia – Italia). Acceso el 10 junio de 2025, <https://www.vatican.va/content/francesco/es/speeches/2024/june/documents/20240614-g7-intelligenza-artificiale.html>.
- Kahneman, Daniel. *Pensar rápido, pensar despacio*. Buenos Aires: Penguin Random House, 2022.
- Joron, Philippe. «La prime à l'émotionnel» en *L'homme postmoderne*. Paris: François Bourin Editeur, 2012.
- Lonergan, Bernard. *Método en teología*. Salamanca: Sígueme, 1988².
- Moingt, Joseph. *Croire au Dieu qui vient I. De la croyance à la foi critique*. Paris: NRF-Gallimard, 2014.
- Morin, Edgar. *Introducción al pensamiento complejo*. Barcelona: Gedisa, 2012².
- Müller V.C., «The History of Digital Ethics», en C. VÉLIZ (ed.), *The Oxford Handbook of Digital Ethics*, Oxford. Oxford University Press, 2024.
- Salmann E., *Contro Severino*. Edizione Piemme, Casale Monferrato, 1996.
- Stahl B.C. – Schroeder D. – Rodrigues R., *Ethics of Artificial Intelligence. Case Studies and Options for Addressing Ethical Challenges*, Springer Open, Cham 2023, 25-37.

