

El papel de la Santa Sede en el desarrollo de la ética de la inteligencia artificial

O papel da Santa Sé no desenvolvimento da ética da inteligência artificial

The Role of the Holy See in the Development of Artificial Intelligence Ethics

Andrea Ciucci  *

A diferencia de otras innovaciones tecnológicas, el rápido desarrollo de los sistemas de inteligencia artificial ha generado un debate, que no puede ser definido como trivial, sobre la ética que debe caracterizar esta transformación. En los últimos cinco años, la Santa Sede ha contribuido de manera significativa al nacimiento de la “algorética”, tanto a través de varias intervenciones del Pontífice como gracias a iniciativas y documentos propuestos por algunos organismos del Vaticano directamente involucrados en este campo. A través de una metodología que ha privilegiado la participación de todos los actores de este proceso, la cuestión de la ética de la inteligencia artificial se ha basado en un juicio positivo, pero no ingenuo, de estas tecnologías. El resultado es una reflexión ética que no se limita a la reducción de posibles riesgos y que, más bien, plantea la pregunta de qué futuro para el hombre y la sociedad se quiere y se puede construir, gracias a la introducción de los sistemas de inteligencia artificial, y cuáles son las condiciones para que esto pueda suceder.

143

Palabras clave: Santa Sede; algorética; inteligencia artificial; ética; Papa Francisco

* Coordinador de la Secretaría de la Academia Pontificia para la Vida, Secretario General de la Fundación RenAissance para la Ética de la Inteligencia Artificial, Ciudad del Vaticano. Correo electrónico: andrea.ciucci@pav.va. ORCID: <https://orcid.org/0009-0006-5859-3023>.

Ao contrário de outras inovações tecnológicas, o rápido desenvolvimento dos sistemas de inteligência artificial gerou um debate, que não pode ser definido como trivial, sobre a ética que deve caracterizar essa transformação. Nos últimos cinco anos, a Santa Sé contribuiu significativamente para o nascimento da “algorética”, tanto por meio de várias intervenções do Pontífice quanto graças a iniciativas e documentos propostos por alguns organismos do Vaticano diretamente envolvidos neste campo. Através de uma metodologia que privilegiou a participação de todos os atores desse processo, a questão da ética da inteligência artificial baseou-se em um julgamento positivo, mas não ingênuo, dessas tecnologias. O resultado é uma reflexão ética que não se limita à redução de possíveis riscos e que, ao contrário, levanta a questão de qual futuro para o homem e a sociedade se deseja e se pode construir, graças à introdução dos sistemas de inteligência artificial, e quais são as condições para que isso possa acontecer.

Palavras-chave: Santa Sé; algorética; inteligência artificial; ética; Papa Francisco

144

Unlike other technological innovations, the rapid development of artificial intelligence systems has sparked a debate, which cannot be defined as trivial, on the ethics that should characterize this transformation. Over the past five years, the Holy See has contributed significantly to the birth of “algorithethics”, both through various interventions by the Pontiff and thanks to initiatives and documents proposed by some Vatican bodies directly involved in this field. Through a methodology that has privileged the participation of all actors in this process, the question of the ethics of artificial intelligence has been based on a positive, but not naive, assessment of these technologies. The result is an ethical reflection that is not limited to reducing potential risks but rather raises the question of what future for humanity and society we want and can build thanks to the introduction of artificial intelligence systems, and what conditions are necessary for this to happen.

Keywords: Holy See; algorithethics; artificial intelligence; ethics; Pope Francis

Introducción

A diferencia de otros grandes campos de la investigación científica y del desarrollo tecnológico, la irrupción de los sistemas de inteligencia artificial que se ha producido en los últimos años ha puesto de manifiesto una cuestión ética. Esta urgencia no solo ha involucrado a los especialistas, como sucedió en el pasado reciente con el desarrollo de la investigación genética o en los años 70 con el auge de la investigación biológica, sino que se ha convertido en un tema ampliamente debatido en la comunidad científica, en el debate político, incluso en el sentir común, comparable quizás a lo que sucedió con el tema nuclear.

A esta proliferación del tema ético relacionado con los sistemas de inteligencia artificial ha contribuido, de manera significativa, una intensa actividad de la Santa Sede. El Papa Francisco, personalmente, y varios organismos vaticanos han abordado y debatido el tema en los últimos años, proponiendo una agenda de temas y una metodología de trabajo. Este artículo ofrece una presentación razonada y crítica de la contribución de la Santa Sede en la construcción de la actual reflexión ética occidental sobre la inteligencia artificial.

Método

Los sujetos de la Santa Sede que participan de manera significativa y estructurada en el trabajo sobre la ética de la inteligencia artificial son, en primer lugar, el Pontífice, seguido de la Pontificia Academia para la Vida con la Fundación RenAIssance, el Dicasterio para la Cultura y la Educación y el Dicasterio para la Comunicación. Otras entidades vaticanas han abordado el tema de manera esporádica. El artículo examina las intervenciones e iniciativas de los sujetos relacionados con el tema de la ética de la inteligencia artificial, mostrando sus cuestiones fundamentales, los diferentes sujetos involucrados y la metodología de acción implementada.

145

El magisterio del Papa Francisco

El magisterio del Papa Francisco sobre el tema de la inteligencia artificial consta, en primer lugar, de tres documentos principales, todos ellos de 2024: “Mensaje para la Jornada de la Paz”, “Mensaje para la Jornada de las Comunicaciones Sociales” y “Discurso ante el G7 en Borgo Ignazia”. Además de estos textos, especialmente estructurados, hay que reseñar una decena de discursos y mensajes, relacionados con iniciativas promovidas por otras realidades de la Santa Sede. En detalle y en orden cronológico:

- Discurso ante el Congreso “Child Dignity in the Digital World” (2019)
- Discurso ante la Pontificia Academia para la Vida sobre la algorética (2020)
- Discurso ante los líderes religiosos firmantes de la *Rome Call* (2023)

- Discurso ante la Pontificia Academia para la Vida sobre tecnología y vida (2023)
- Discurso ante la Sociedad Max Planck (2023)
- Discurso ante la Minerva Dialogues (2023)
- Discurso ante la Fundación Centesimus Annus (2024)
- Mensaje para el encuentro “AI Ethics for Peace” (2024)
- Discurso ante la Academia Pontificia de las Ciencias (2024)
- Mensaje ante la cumbre de París “para la acción sobre la inteligencia artificial” (2025).

Llama la atención, en primer lugar, el número y la importancia de las intervenciones del Papa Francisco, teniendo en cuenta que la tecnología y la inteligencia artificial no son temas tradicionalmente religiosos. De especial relevancia simbólica, más que de contenido, son el discurso ante el G7 (el Papa podría haber elegido hablar a los grandes de la Tierra sobre otros temas quizás más afines, como, por ejemplo, la libertad religiosa o la paz) y la acuñación de un neologismo, «algorética», que ha recibido un reconocimiento oficial y ha gozado de cierta difusión. El motivo por el que la Iglesia católica interviene en un tema tecnológico lo explica el Papa Francisco en el discurso ante la Pontificia Academia para la Vida de 2020:

146

“La llamada ‘inteligencia artificial’ está en el corazón mismo del cambio de época que estamos atravesando. La innovación digital, efectivamente, alcanza a todos los aspectos de la vida, tanto personales como sociales. Afecta a la forma en que entendemos el mundo y a nosotros mismos. Está cada vez más presente en las actividades e incluso en las decisiones humanas, y está cambiando nuestra forma de pensar y actuar [...] En efecto, la profundidad y la aceleración de las transformaciones de la era digital plantean problemas inesperados que imponen nuevas condiciones al ethos individual y colectivo” (Francisco, 2020).

Lejos de caer en juicios alarmistas o incluso apocalípticos, el texto reconoce la capacidad transformadora de estas tecnologías, definidas poco después como “don de Dios”, que impone una reflexión amplia y renovada, con el fin de salvaguardar la vida humana y los vínculos sociales, gracias a un nuevo compromiso en el ámbito ético y educativo. El Pontífice lleva a cabo esta reflexión con un triple marco de referencia: en primer lugar, la Doctrina Social de la Iglesia (dignidad de la persona, justicia, solidaridad y subsidiariedad), luego el díptico de las dos encíclicas *Laudato Si'* (sobre el cuidado de la casa común) y *Fratelli tutti* (sobre la fraternidad universal), citadas en varias ocasiones en muchos discursos, y, por último, los derechos humanos que “en el encuentro entre diferentes visiones del mundo, constituyen un importante punto de convergencia para la búsqueda de un terreno común” (Francisco, 2024b).

En el discurso ante el G7, el Papa Francisco (2024b) argumenta la necesidad de una inspiración ética de la inteligencia artificial a partir de dos consideraciones fundamentales: la peculiaridad de este instrumento definido como *sui generis* por su complejidad y autonomía, la no neutralidad de cualquier innovación tecnológica y el consiguiente impacto en la vida y en la sociedad humanas.

“No debemos olvidar que ninguna innovación es neutral. La tecnología nace con un propósito y, en su impacto en la sociedad humana, representa siempre una forma de orden en las relaciones sociales y una disposición de poder, que habilita a alguien a realizar determinadas acciones impidiéndoselo a otros. Esta dimensión de poder que es constitutiva de la tecnología incluye siempre, de una manera más o menos explícita, la visión del mundo de quien la ha realizado o desarrollado. Esto vale también para los programas de inteligencia artificial. Con el fin de que estos instrumentos sean para la construcción del bien y de un futuro mejor, deben estar siempre ordenados al bien de todo ser humano. Deben contener una inspiración ética. La decisión ética, de hecho, es aquella que tiene en cuenta no sólo los resultados de una acción, sino también los valores en juego y los deberes que se derivan de esos valores” (Francisco, 2024b).

147

El bien del ser humano evocado por el Pontífice tiene, ante todo, en la custodia y centralidad de la persona humana, que debe ser custodiada y no reducida o marginada, uno de los pilares de la reflexión ética sobre la inteligencia artificial, como se desprende de la repetición de esta intuición en muchos de sus discursos. “Nuestro último desafío es y seguirá siendo siempre el hombre”, escribe el Papa Francisco (2025) a Macron, y un tiempo antes:

“Desde esta perspectiva, creo que el desarrollo de la inteligencia artificial y del aprendizaje automático tiene el potencial de aportar una contribución beneficiosa al futuro de la humanidad, no podemos descartarlo. Sin embargo, estoy seguro de que este potencial sólo se hará realidad si existe una voluntad coherente por parte de quienes desarrollan las tecnologías para actuar de forma ética y responsable. Me anima el compromiso de tantas personas que trabajan en estos campos para garantizar que la tecnología se centre en el ser humano, se funde en bases éticas durante el diseño del proyecto y tenga por finalidad el bien [...] El concepto de dignidad humana -este es el núcleo- exige que reconozcamos y respetemos el hecho de que el valor fundamental de una persona no puede medirse con un conjunto de datos [...] No podemos permitir que los algoritmos limiten o condicionen el respeto a la dignidad humana, ni que excluyan la

compasión, la misericordia, el perdón y, sobre todo, la apertura a la esperanza de cambio en el individuo” (Francisco, 2023d).

Además de la centralidad de la persona humana, el Papa Francisco (2020) plantea como segundo tema ético central el de las repercusiones sociales de la implementación de los sistemas de inteligencia artificial, mostrando en varios pasajes cómo este fenómeno está aumentando la injusticia y las desigualdades en el mundo en lugar de reducirlas, reduciendo los espacios de democracia y alimentando una cultura de la eficiencia que aumenta la brecha de los más débiles y frágiles.

Los diferentes discursos están plagados de ejemplos concretos de estas repercusiones sociales. Dependiendo de los contextos, el Pontífice recuerda los riesgos que se vislumbran en el mundo laboral, en la medicina, en los fenómenos migratorios, en la práctica de la justicia, en la educación, en la elaboración de perfiles de las personas, en el control y la calificación social, en la seguridad. El Pontífice dedica especial atención a los temas de la paz y de la comunicación en la era de la inteligencia artificial, que condensa en dos documentos específicos.

En el largo mensaje para el Día de la Paz 2024 (Francisco, 2023e), el Papa Francisco dedica un capítulo a los “temas candentes para la ética”, abordando brevemente cuestiones como la fiabilidad de los sistemas, un sistema generalizado de manipulación y control social, un cambio en el mundo laboral que aumenta la injusticia social. En este contexto, el Pontífice afirma que:

“La búsqueda de las tecnologías emergentes en el sector de los denominados ‘sistemas de armas autónomos letales’, incluido el uso bélico de la inteligencia artificial, es un gran motivo de preocupación ética. Los sistemas de armas autónomos¹ no podrán ser nunca sujetos moralmente responsables. La exclusiva capacidad humana de juicio moral y de decisión ética es más que un complejo conjunto de algoritmos, y dicha capacidad no puede reducirse a la programación de una máquina que, aun siendo ‘inteligente’, no deja de ser siempre una máquina. Por este motivo, es imperioso garantizar una supervisión humana adecuada, significativa y coherente de los sistemas de armas” (Francisco, 2023e).

En clave positiva, el Pontífice declara que:

¹ El tema de las armas autónomas también se retoma, con la misma preocupación, en el discurso ante el G7 (Francisco, 2024b) y en “Mensaje para el encuentro de Hiroshima” (Francisco, 2024d).

“Realmente lo último que el mundo necesita es que las nuevas tecnologías contribuyan al injusto desarrollo del mercado y del comercio de las armas, promoviendo la locura de la guerra. [...] Las aplicaciones técnicas más avanzadas no deben usarse para facilitar la resolución violenta de los conflictos, sino para pavimentar los caminos de la paz” (Francisco, 2023e).

El uso de la inteligencia artificial para construir un mundo más justo y pacífico —concluye el mensaje— requiere un esfuerzo educativo y comunicativo adecuado, así como un compromiso renovado de los actores políticos involucrados en las grandes decisiones globales.

“Mensaje para la Jornada de las Comunicaciones Sociales 2024” (Francisco, 2024a) explora con mayor precisión los desafíos que la inteligencia artificial plantea al gran mundo de la comunicación.² El mensaje comienza con un llamamiento a una visión sapiencial y espiritual de esta tecnología que, de alguna manera, invade el espacio propio del lenguaje humano, con el fin de preservar “una comunicación plenamente humana” que no se reduzca a un flujo de datos desencarnado. Aquí la lista de cuestiones sobre las que reflexionar es más detallada: explosión de noticias falsas, preponderancia de las redes sociales, riesgo de *echo chamber*, construcción de un pensamiento único, campañas de desinformación, peligros para los trabajadores del sector, reducción de los espacios de participación, aumento de las desigualdades sociales también debido a la gestión de las noticias. Ante estos peligros, el documento destaca:

149

“El reto que tenemos ante nosotros es dar un salto cualitativo para estar a la altura de una sociedad compleja, multiétnica, pluralista, multirreligiosa y multicultural. Nos corresponde cuestionarnos sobre el desarrollo teórico y el uso práctico de estos nuevos instrumentos de comunicación y conocimiento” (Francisco, 2024a).

Esta dificultad, continúa el Pontífice, solo puede afrontarse manteniendo relaciones existenciales que impliquen los cuerpos, el arraigo en la realidad, compartir experiencias y no solo datos. En este sentido, parece decisiva la “alianza entre generaciones, entre quienes tienen memoria del pasado y quienes tienen visión de futuro”. La algorética se convierte en la clave de síntesis para comprender el mensaje del Papa Francisco. El neologismo aparece por primera vez en los discursos del Papa en 2019 y se retoma varias veces en sus intervenciones. La algorética se define como una nueva frontera de la reflexión ética:

² Sobre el mismo tema, véase también el Documento del Dicasterio para la Comunicación (2023), dedicado específicamente al tema de las redes sociales, en el que la IA emerge como un factor de multiplicación de noticias peligrosamente problemático.

“Su objetivo es asegurar una verificación competente y compartida de los procesos con los que se integran en nuestra era las relaciones entre los seres humanos y las máquinas [...] La ‘algor-ética’ podrá ser un puente para que los principios [de la Doctrina Social de la Iglesia] se inscriban concretamente en las tecnologías digitales, mediante un diálogo transdisciplinario eficaz” (Francisco, 2020).

Según Francisco (2023e), debe ofrecerse un desarrollo ético de los algoritmos, un marco de valores que pueda orientar de manera eficaz y concreta el desarrollo tecnológico, una “moderación ética” (Francisco, 2024b) que debe construirse ya en la fase de diseño (*by design*) con los ingenieros y las realidades industriales implicadas. Resulta especialmente interesante la mención del Pontífice sobre esta colaboración en un tema ético con las realidades científicas implicadas, que aparece en el discurso ante el G7:

“En un contexto plural y global, en el que también se muestran las distintas sensibilidades y plurales jerarquías en las escalas de valores, parecería difícil encontrar una única jerarquía de valores. Pero en el análisis ético podemos recurrir además a otros tipos de instrumentos. Si nos cuesta definir un solo conjunto de valores globales, podemos encontrar principios compartidos con los cuales afrontar y disminuir eventuales dilemas y conflictos de la vida. Por esta razón ha nacido la *Rome Call*. En el término ‘algorética’ se condensa una serie de principios que se revelan como una plataforma global y plural capaz de encontrar el apoyo de las culturas, las religiones, las organizaciones internacionales y las grandes empresas protagonistas de este desarrollo” (Francisco, 2024b).

150

La Pontificia Academia para la Vida y la Fundación RenAissance

La contribución ofrecida por la Pontificia Academia para la Vida³ se inscribe en un proceso de reflexión científica más amplio, iniciado en 2016 con la reforma de la entidad deseada por el Papa Francisco. La academia destacó desde el principio la especial urgencia de la relación entre las nuevas tecnologías y la vida humana, iniciando un camino de investigación que la llevó a analizar, a través de una serie de congresos y publicaciones, el influjo de las tecnologías biomédicas en los momentos del nacimiento y la muerte, la redefinición del concepto de humano en relación con el avance de la robótica, el influjo poderoso de las llamadas tecnologías emergentes y convergentes.⁴

³ Los objetivos, la estructura y las actividades de la Academia pueden verse en: www.academyforlife.va.

⁴ Entre los frutos más prestigiosos de este trabajo se encuentra el artículo publicado en *Nature Machine Intelligence*: Sinibaldi *et al.*, 2020.

En este marco se inserta el congreso de 2020 (con las actas publicadas), *The Good Algorithm? Artificial Intelligence: Ethics, Law, Health* (Paglia, 2021b), y el lanzamiento asociado de la *Rome Call for AI Ethics*.

La sesión del congreso dedicada a la ética en el campo de la IA sirvió para sentar las bases del problema, que en aquel momento era objeto de estudio solo por parte de especialistas. A través de cuatro ponencias se puso de manifiesto la necesidad de abordar el tema en el marco más amplio de la transformación tecnológica y digital, la dimensión social y multicultural en la que debía situarse esta reflexión y las repercusiones educativas conexas. El acto de clausura del congreso fue la ceremonia de firma de la *Rome Call for AI Ethics*, inaugurada por el mensaje del Papa (Francisco, 2020), en la que participaron como firmantes: el director general de la FAO, Qu Dongyu; el presidente de Microsoft, Brad Smith; el vicepresidente de IBM, John Kelly III; la Ministra de Transición Tecnológica del Gobierno italiano, Paola Pisano; y el presidente de la Academia, monseñor Vincenzo Paglia. También asistió, sin firmar, el presidente del Parlamento Europeo, David Sassoli.

La *Rome Call*⁵ es un documento breve, compuesto por un marco antropológico de fondo articulado en tres breves capítulos (ética, educación, derecho) y seis principios:

- *Transparencia*: en general, los sistemas de inteligencia artificial deben ser explicables;
- *Inclusión*: las necesidades de todos los seres humanos deben tenerse en cuenta para que todos puedan beneficiarse y para que se ofrezcan a todas las personas las mejores condiciones posibles para expresarse y desarrollarse;
- *Responsabilidad*: quienes diseñan y ponen en práctica estas tecnologías deben proceder con responsabilidad y transparencia;
- *Imparcialidad*: evitar crear o actuar según prejuicios, salvaguardando así la equidad y la dignidad humana;
- *Fiabilidad*: los sistemas de inteligencia artificial deben ser capaces de funcionar de manera fiable;
- *Seguridad y privacidad*: los sistemas de inteligencia artificial deben funcionar de forma segura y respetar la privacidad de los usuarios.

151

El texto, fruto del trabajo de una comisión de expertos de diversas disciplinas y culturas dirigida por el profesor Paolo Benanti, no es deliberadamente un texto vaticano y, por lo tanto, no forma parte en modo alguno del corpus del magisterio católico. Más bien es un texto propuesto por un organismo de la Iglesia católica a todos aquellos interesados en

⁵ Para conocer el texto, todos los firmantes y las iniciativas relacionadas con la difusión del documento, se puede consultar la página oficial: www.romecall.org. El sitio está gestionado por la Fundación vaticana RenAIssance, creada por el Papa Francisco específicamente para custodiar, profundizar y difundir la *Rome Call* y, en general, un enfoque ético de la tecnología.

construir juntos una visión ética de la inteligencia artificial y dispuestos a hacer suya la visión del documento y suscribirlo. El único texto fundamental citado e inspirador es la Declaración Universal de los Derechos Humanos. Esta ubicación responde al método y al propósito del documento. De hecho, adopta el método indicado por el Papa Francisco de dialogar sobre estos temas con todos los sujetos involucrados y está formulado de manera que todos los sujetos que lo deseen y se comprometan en este sentido puedan suscribirlo, sin tener que declarar necesariamente una opción religiosa.

Desde el punto de vista de la difusión del documento, esta formulación ha sido especialmente eficaz. Hasta la fecha, lo han firmado varias otras industrias (entre ellas CISCO en 2024, y Qualcomm y Salesforce en 2025), decenas de universidades de todo el mundo y los líderes de todas las principales religiones del mundo.⁶ El Papa Francisco ha definido la *Rome Call* como “una plataforma global y plural capaz de encontrar el apoyo de las culturas, las religiones, las organizaciones internacionales y las grandes empresas protagonistas de este desarrollo” (Francisco, 2024b). El *AI Index Report 2021* publicado por el Institute for Human-Centered AI de la Universidad de Stanford lo ha definido como uno de los cinco documentos más importantes sobre el tema en 2020. Amandeep Singh Gill, enviado para la tecnología del secretario general de las Naciones Unidas, declaró en su discurso en el evento de Hiroshima para la firma de la *Rome Call* por parte de los líderes religiosos del mundo: “*The Rome Call for AI Ethics embodies the spirit needed for global AI governance*”.

152

En resumen, se puede afirmar que la Pontificia Academia ha desempeñado cuatro funciones en el campo del desarrollo de la ética de la IA. En primer lugar, ha llamado la atención, antes que muchos otros, sobre un tema que luego se ha convertido en central en el debate académico y público. En segundo lugar, situó esta cuestión en una reflexión antropológica y social más amplia, sin limitarse a ofrecer, como muchos otros que le siguieron, una mera lista de principios reguladores. Benanti sitúa la algorética en el marco de una condición que él define como tecno-humana:

“The way humans have always understood and lived their existence: an interaction with the environment mediated through tools, technological artifacts. It seems clear, then, that it is not possible to separate the history of human beings and civilization from the history of the tools they made” (Benanti, 2023).

En tercer lugar, ha intentado aplicar un método transdisciplinario en el que el resultado final es fruto del trabajo conjunto de expertos pertenecientes a ámbitos científicos muy diferentes. De hecho, ha conseguido dar forma a lo que en varias ocasiones

⁶ En 2023, líderes religiosos judíos y musulmanes firmaron juntos la *Rome Call* en el Vaticano. En 2024, en Hiroshima, líderes budistas, sijs, confucianos, hindúes, bahaíes y zoroastrianos firmaron el documento. Para un análisis detallado de la participación de las religiones, véase: Ciucci (2025).

había sido solicitado por el mundo tecnológico, que denunciaba la problemática falta, en sus entornos, de un punto de vista humanista. Por último, la *Rome Call* ha sido un poderoso medio de difusión del tema y su visión en mundos y contextos muy diferentes. Paglia escribe sucintamente:

“Possiamo cogliere, anche se fra molte contraddizioni e tensioni, la graduale maturazione di un comune impegno per un autentico sviluppo, che intende il progresso in termini non solo economici o quantitativi. Il pur indispensabile contributo del sapere tecnico-scientifico viene infatti articolato con altre prospettive, incluse quelle della teologia e dell'etica, in modo da orientarci, con la collaborazione di tutte le persone di buona volontà, verso uno sviluppo umano effettivamente integrale” (Paglia, 2021a).

El Dicasterio para la Cultura y la Educación

El otro gran actor vaticano que se ha movido en el tema de la ética de la inteligencia artificial es sin duda el Dicasterio para la Cultura y la Educación. Este compromiso se ha desarrollado en cuatro direcciones fundamentales: la participación en congresos científicos y debates de algunas figuras de alto nivel del dicasterio, la promoción de los *Minerva Dialogues*,⁷ la constitución de un grupo de trabajo dedicado al tema que produjo en 2024 un volumen sobre el tema (Gaudet *et al.*, 2024) y la publicación en 2025, en colaboración con el Dicasterio para la Doctrina de la Fe, de la *Nota Antiqua et Nova*.

153

Si los dos primeros elementos muestran la adquisición efectiva del método que consiste en escuchar y dialogar, como pidió el Papa Francisco en su discurso a los participantes en los *Minerva Dialogues* de 2023, los dos documentos proponen un análisis del fenómeno de la inteligencia artificial y esbozan algunas perspectivas de carácter antropológico y ético. El volumen *Encountering Artificial Intelligence* dedica la segunda parte al tema de la ética:

“In this part of the volume, we respond to Pope Francis's call with an outline of an AI ethics that can assist individuals, institutions, and civil society in engaging personal, social, and political concerns related to AI in light of the Catholic theological and moral tradition” (Gaudet *et al.*, 2024, p. 147).

⁷ Los *Minerva Dialogues* son encuentros anuales, iniciados en 2016 por iniciativa del fraile dominico Éric Salobir. Como explica el secretario del Dicasterio, Paul Tighe, en un artículo publicado en el *Osservatore Romano* (30 de marzo de 2023), constituyen una oportunidad para que expertos del mundo de la tecnología —entre los que se encuentran científicos, ingenieros, directivos de empresas, juristas y filósofos— se reúnan con representantes de la iglesia, como oficiales de la curia, teólogos y moralistas. Nacidos para promover una mayor conciencia y reflexión sobre el impacto social y cultural de las tecnologías digitales, según Tighe, ofrecen la oportunidad de “*indagare insieme per capirci meglio, per assicurarci che la tecnologia sia applicata per il bene dell'umanità, parlando anche con schiettezza in certi momenti*”.

En el capítulo seis, gracias a un análisis general del magisterio del Papa Francisco (definido como positivo pero no ingenuo), el volumen radica la cuestión de la ética de la inteligencia artificial en el ámbito más amplio de la doctrina social de la Iglesia, mostrando cómo las categorías fundamentales de esta reflexión pueden abordar positivamente este nuevo fenómeno: bien común, solidaridad, subsidiariedad, libertad, educación y familia son las palabras que permiten identificar y asegurar un impacto positivo de la inteligencia artificial en la vida personal y social de las personas.

“The goal of chapter 7 is to identify how AI systems can make the realization of the ideal even more difficult. To serve this purpose, the chapter articulates the vision and the challenges at the same time. We will begin with how AI is shaping more intimate spheres of encounter, such as the family, before expanding to look at AI's effects on our broader social systems, like government and the economy” (Gaudet et al., 2024, p. 161).

154

A través de las categorías identificadas en el capítulo anterior, el volumen analiza los beneficios y riesgos —introducidos por los sistemas de inteligencia artificial— sobre la familia (cuyas relaciones están moldeadas por las nuevas tecnologías), sobre la educación y el paradigma tecnocrático, sobre la salud y la medicina, sobre la protección de los más débiles (arriesgadamente expuestos a la eficiencia de los algoritmos), sobre el mundo laboral con la pérdida y la creación de figuras profesionales, sobre el ámbito militar con especial atención a las armas autónomas, sobre los procesos culturales y sobre el impacto medioambiental de estos sistemas.

Por último, el capítulo ocho ofrece una visión sapiencial en la que es posible habitar un tiempo digital, destacando toda una serie de ejercicios de humanidad y atenciones que deben preservarse en la vida cotidiana (que no deben digitalizarse y autonomizarse indiscriminadamente), en el cuidado de las personas y situaciones más vulnerables (que deben encontrar en las nuevas tecnologías un beneficio y no un enésimo factor de exclusión), en el debate público y jurídico (preservando la libertad de las personas y valores como la equidad, la privacidad, la seguridad y la responsabilidad), en la economía (llamada a custodiar un rostro humano) y en las relaciones interpersonales (que no pueden reducirse a formas exclusivamente digitales). En cada uno de estos temas se presentan también puntos de trabajo para quienes diseñan y crean sistemas de inteligencia artificial.

La nota *Antiqua et Nova*

El 14 de enero de 2025, el Dicasterio para la Cultura y la Educación, junto con el Dicasterio para la Doctrina de la Fe, publicó *Antiqua et Nova - Nota sobre la relación entre la inteligencia artificial y la inteligencia humana*. El extenso documento ofrece

una reflexión articulada sobre el tema de la inteligencia artificial en diálogo con la tradición cristiana y recoge todo el trabajo realizado por la Santa Sede sobre el tema.

Si en las tres primeras partes el documento enmarca el tema de la inteligencia artificial en comparación con la inteligencia humana, la cuarta parte está dedicada específicamente a la cuestión ética. Gracias a este marco, al final, la larga quinta parte retoma varias cuestiones específicas que destacan, retomando el esquema y el índice del volumen *Encountering Artificial Intelligence*, las posibilidades y los riesgos del uso de la inteligencia artificial en los diferentes campos de la experiencia humana. Los primeros números del cuarto volumen definen el marco en el que se debe plantear la cuestión ética en este tema. La inteligencia artificial es una empresa humana y, como toda actividad humana, pide ser comprendida dentro del designio de Dios: el intelecto es un don suyo para ser usado para el bien (Dicasterio Doctrina, 2025, pp. 36-38).

El documento destaca dos características fundamentales. En primer lugar, el resultado de la comparación entre la inteligencia humana y la artificial muestra que la ética de la inteligencia artificial es, en realidad, siempre la ética de las personas:

“La dimensión ética es primordial, ya que son las personas las que diseñan los sistemas y determinan para qué se utilizan. Entre una máquina y un ser humano, sólo este último es verdaderamente un agente moral, es decir, un sujeto moralmente responsable que ejerce su libertad en sus decisiones y acepta las consecuencias de las mismas; sólo el ser humano está en relación con la verdad y el bien, guiado por la conciencia moral que le llama a “amar y practicar el bien y que debe evitar el mal”, certificando “la autoridad de la verdad con referencia al Bien supremo por el cual la persona humana se siente atraída”; sólo el ser humano puede ser lo suficientemente consciente de sí mismo como para escuchar y seguir la voz de la conciencia, discerniendo con prudencia y buscando el bien posible en cada situación. En realidad, esto también pertenece al ejercicio de la inteligencia por parte de la persona” (Dicasterio Doctrina, 2025, p. 39).

155

En segundo lugar, la custodia de la dignidad de la persona humana y del bien común (los fines de la acción ética también en este campo, como se señala en los números 42 y 48) deben ser perseguidos tanto por quienes utilizan estas tecnologías (custodiando sus fines, como se indica en el número 40) como por quienes diseñan y construyen estos sistemas (los medios no son moralmente indiferentes, como se indica en el número 41). En esencia, una ética *by design*, como ha señalado en varias ocasiones el Papa Francisco y se destaca en la *Rome Call*.

La segunda mitad de la parte dedicada a la ética subraya, por último, una cuestión que requiere especial atención:

“Dado que una causalidad moral en sentido pleno sólo pertenece a los agentes personales, no a los artificiales, es de suma importancia poder identificar y definir quién es responsable de los procesos de IA, en particular de aquellos que incluyen posibilidades de aprendizaje, corrección y reprogramación. Si bien, por un lado, los métodos empíricos (*bottom-up*) y las redes neuronales muy profundas permiten a la IA resolver problemas complejos, por otro lado, dificultan la comprensión de los procesos que condujeron a tales soluciones. Esto complica la determinación de responsabilidades ya que, si una aplicación de IA produjera resultados no deseados, sería difícil determinar a qué persona atribuirlos” (Dicasterio Doctrina, 2025, 44).

Aquí, en el documento, aparece el tema de la *accountability* y de la necesidad de instrumentos jurídicos reguladores adecuados, capaces de salvaguardar y reconocer al mismo tiempo las responsabilidades humanas, incluso en procesos tecnológicos y estructuras sociales extremadamente complejos.

Conclusiones

156 El largo análisis dedicado a las intervenciones e iniciativas promovidas por la Santa Sede en los últimos cinco años sobre el tema de la ética de la inteligencia artificial ofrece algunas conclusiones.

La primera y fundamental, porque ha hecho posibles las demás, es de orden metodológico. La dirección fue claramente dada por el Papa Francisco, quien ya en 2020 afirmó:

“En efecto, como creyentes, no tenemos nociones preestablecidas con las que responder a las preguntas sin precedentes que la historia hoy nos plantea. Nuestra tarea es, más bien, caminar junto con los demás, escuchando atentamente y poniendo en contacto la experiencia y la reflexión. Debemos dejarnos interpelar como creyentes, para que la Palabra y la Tradición de la fe nos ayuden a interpretar los fenómenos de nuestro mundo, identificando caminos de humanización, y por tanto de evangelización amorosa, para recorrerlos juntos. Así podremos dialogar provechosamente con todos aquellos que buscan el desarrollo humano, manteniendo a la persona en todas sus dimensiones, incluidas las espirituales, en el centro del conocimiento y las prácticas sociales. Nos enfrentamos a una tarea que involucra a la familia humana en su totalidad” (Francisco, 2020).

Todas las intervenciones del Vaticano han seguido esta modalidad de diálogo y apertura con todos los sujetos involucrados. En este sentido, por ejemplo, la Iglesia Católica ha optado por no designar la *Rome Call* como un documento vaticano.

Al seguir este método, la Santa Sede ha ocupado el espacio público no para defender una visión del mundo, algunas de sus prerrogativas o, peor aún, algunas posiciones privilegiadas, sino para poner a disposición del debate público una visión articulada del hombre y de la sociedad. En cierto modo, sucedió exactamente lo contrario de lo que ocurrió en el episodio más llamativo de enfrentamiento entre la iglesia y los científicos de la época moderna: el caso Galileo. En aquella ocasión, los cardenales que trataron el asunto se opusieron a Galileo no tanto por razones científicas, sino más bien porque intuían que la aceptación de lo propuesto por el científico toscano socavaría gravemente la visión teológica de la época, que situaba en el centro del universo el acontecimiento cristiano y la consiguiente función de la Iglesia. En el caso de Galileo, la Iglesia se atrincheró en un intento apologético; hoy, ante la explosión de la inteligencia artificial, el Santo Padre ha salido al ágora pública para contribuir a la reflexión común, sin “nociones preconcebidas con las que responder a las preguntas sin precedentes que la historia hoy nos plantea”.

La ampliación de los sujetos implicados ha supuesto asumir algún riesgo: ha surgido cierta perplejidad en torno a la participación de algunas de las mayores empresas mundiales del sector tecnológico y de sus líderes. La decisión de trabajar con ellos, tomada después de largos períodos de reflexión y discernimiento, se deriva de la percepción de que una reflexión ética que no involucrara a los actores y autores de esta revolución digital privaría al proyecto de consistencia y aplicabilidad reales. Las divergencias, incluso profundas, sobre aspectos de carácter principalmente social no impidieron encontrar convergencias y esperanzas comunes sobre las que trabajar juntos, por el bien de todos.

157

La segunda conclusión se refiere al juicio general que la Santa Sede da al fenómeno de la inteligencia artificial. Como se desprende del análisis realizado en las páginas anteriores, en ningún texto aparece un tono apocalíptico y de grave juicio condenatorio. También son muy escasas las referencias al poshumanismo y al transhumanismo, tradiciones de pensamiento que a menudo se asocian a juicios extremadamente negativos sobre la transición digital en curso. Ciertamente, la Santa Sede denuncia el paradigma tecnocrático “que promete un progreso incontrolado e ilimitado se impondrá y quizás incluso eliminará otros factores de desarrollo con enormes peligros para toda la humanidad” (Francisco, 2019). Pero el juicio sobre la inteligencia artificial es positivo y, al mismo tiempo, no ingenuo; reconoce esta tecnología como uno de los mejores frutos del hombre y, al mismo tiempo, destaca sus riesgos y oportunidades.

La algorética no pretende simplemente enumerar una serie de principios reglamentarios destinados a reducir los riesgos del impacto de esta tecnología; más bien, inserta la inteligencia artificial en un marco prospectivo positivo. La pregunta que surge es: ¿cómo podemos, juntos, construir y utilizar estos sistemas para construir

el bien de las personas, de la sociedad, del planeta?⁸ La idea de una tecnología no neutral y la atención a sus repercusiones sociales, y no solo antropológicas, se derivan de este enfoque y caracterizan de manera peculiar la reflexión ética propuesta por la Santa Sede sobre el tema. La pregunta que permanece en los corazones y mentes de muchos es la relativa a la eficacia real que todo este camino puede tener, en un contexto global caracterizado por inmensos intereses económicos y políticos. La discusión de los últimos años ha mostrado un primer resultado: ciertamente, la Santa Sede ha contribuido de manera significativa a la construcción y promoción del tema de la ética de la inteligencia artificial, y ha sido esencial para su posicionamiento entre las grandes cuestiones de la discusión pública mundial. No es poca cosa y no ocurre con otros sectores, igualmente decisivos, de la investigación científica y tecnológica. En un futuro próximo, será decisivo ampliar la conciencia de la urgencia de la cuestión entre los diferentes sujetos implicados y la capacidad de los políticos y constructores de encontrar reglas y soluciones capaces de apoyar al mismo tiempo la justicia y el desarrollo, la protección y la innovación.

Por último, para la Iglesia católica otra prueba de fuego será la gestión de la transformación digital en la práctica pastoral de las distintas iglesias locales. Reconocer el don de Dios de la inteligencia artificial y no pensar que esto también afecta a las condiciones concretas de la práctica eclesial podría ser el primer fracaso.

158

Bibliografía

Benanti, P. (2020). Algor-Ethics: Artificial Intelligence and Ethical Reflection. *Revue d'éthique et de théologie morale*, 307(3), 93-110.

Benanti, P. (2023). The urgency of an algorethics. *Discovering Artificial Intelligence*, 3(11). DOI: <https://doi.org/10.1007/s44163-023-00056-6>.

Ciucci, A. (2024). Le religioni e l'etica dell'intelligenza artificiale. Il caso della Rome Call for AI Ethics. *La Società*, 5-6, 100-109.

Dicasterio para la Comunicación (2023). Hacia una plena presencia - Reflexión pastoral sobre la interacción en las Redes Sociales. Recuperado de: https://www.vatican.va/roman_curia/dpc/documents/20230528_dpc-verso-piena-presenza_es.html.

Dicasterio para la Doctrina de la Fe y Dicasterio para la Cultura y la Educación (2025). *Antiqua et Nova*. Nota sobre la relación entre la inteligencia artificial y la inteligencia

⁸ En este sentido, se puede pensar que el marco ético propuesto por la Santa Sede puede atribuirse, al menos en cierto modo, al sociotecnicismo pragmático defendido por Watson, Mokander y Floridi (2024), ya que no se corresponde ni con el optimismo ingenuo del sociotecnicismo dogmático ni con el pesimismo riguroso del escepticismo, considerados insuficientes por los autores.

humana. Recuperado de: https://www.vatican.va/roman_curia/congregations/cfaith/documents/rc_ddf_doc_20250128_antiqua-et-nova_sp.html.

Gaudet, M., Herzfeld, N., Scherz, P. & Wales, J. (2024). *Encountering Artificial Intelligence: Ethical and Anthropological Explorations*. Eugene: Pickwick Publications.

Francisco (2019). Discurso a los participantes en el congreso “Promoting Digital Child Dignity”. Recuperado de: https://www.vatican.va/content/francesco/es/speeches/2019/november/documents/papa-francesco_20191114_convegno-child%20dignity.html.

Francisco (2020). Discurso a los participantes en la plenaria de la Pontificia Academia para la Vida. Recuperado de: https://www.vatican.va/content/francesco/es/speeches/2020/february/documents/papa-francesco_20200228_accademia-perlavita.html.

Francisco (2023a). Discurso a los participantes en el encuentro “Rome Call” organizado por la Fundación RenAIssance. Recuperado de: <https://www.vatican.va/content/francesco/es/speeches/2023/january/documents/20230110-incontro-romecall.html>.

Francisco (2023b). Discurso a los Miembros de la Pontificia Academia para la Vida, recuperado de <https://www.vatican.va/content/francesco/es/speeches/2023/february/documents/20230220-pav.html>

Francisco (2023c). Discorso a una delegazione della Società Max Planck. Recuperado de: <https://www.vatican.va/content/francesco/it/speeches/2023/february/documents/20230223-delegazione-mplanck.html>.

159

Francisco (2023d). Discurso a los participantes en los “Minerva Dialogues” organizado por el Dicasterio para la Cultura y la Educación. Recuperado de: <https://www.vatican.va/content/francesco/es/speeches/2023/march/documents/20230327-minerva-dialogues.html>.

Francisco (2023e). Mensaje para la celebración de la 57 Jornada mundial de la Paz: Inteligencia artificial y paz. Recuperado de: <https://www.vatican.va/content/francesco/es/messages/peace/documents/20231208-messaggio-57giornatamondialepace2024.html>.

Francisco (2024a). Mensaje para la 58 Jornada mundial de las Comunicaciones Sociales: Inteligencia artificial y sabiduría del corazón para una comunicación plenamente humana. Recuperado de: <https://www.vatican.va/content/francesco/es/messages/communications/documents/20240124-messaggio-comunicazioni-sociali.html>.

Francisco (2024b). Discurso en la Sesión del G7 sobre inteligencia artificial. Recuperado de: <https://www.vatican.va/content/francesco/es/speeches/2024/june/documents/20240614-g7-intelligenza-artificiale.html>.

Francisco (2024c). Discurso ai partecipanti al convegno internazionale promosso dalla Fondazione Centesimus Annus Pro Pontifice. Recuperado de: <https://www.vatican.va/content/francesco/it/speeches/2024/june/documents/20240622-centesimus-annus-propontifice.html>.

Francisco (2024d). Mensaje para la reunión “AI Ethics for Peace”. Recuperado de <https://www.vatican.va/content/francesco/es/messages/pont-messages/2024/documents/20240710-messaggio-ai-ethics-forpeace.html>.

Francisco (2024e). Discurso a los participantes en la plenaria de la Academia de las Ciencias, recuperado de: <https://www.vatican.va/content/francesco/es/speeches/2024/september/documents/20240923-plenaria-accademia-scienze.html>.

Francisco (2025). Mensaje al Presidente de la República francesa con ocasión del “Sommet pour l’action sur l’intelligence artificielle”. Recuperado de: <https://www.vatican.va/content/francesco/es/messages/pont-messages/2025/documents/20250207-messaggio-summit-parigi-ia.html>.

Paglia, V. (2021a). Tecnologie digitali ed etica. *Rivista Di Scienze dell'Educazione*, 59(1), 68–80.

160

Paglia V. Pegoraro R. (2021b). The Good Algorithm? Artificial Intelligence: Ethics, Law, Health. Recuperado de: [https://www.academyforlife.va/content/dam/pav/documenti%20pdf/2020/Assemblea/Abstract/Abstract%20book_GOLD%20\(1\).pdf](https://www.academyforlife.va/content/dam/pav/documenti%20pdf/2020/Assemblea/Abstract/Abstract%20book_GOLD%20(1).pdf).

Sinibaldi, E. *et al.* (2020). Contributions from the Catholic Church to ethical reflections in the digital era. *Nature Machine Intelligence*, 2, 242-244.

Watson, D. S., Mökander, J. & Floridi, L. (2024). Competing narratives in AI ethics: a defense of sociotechnical pragmatism. *AI & Society*.