
Mathew J. GAUDET, Noreen HERZFELD, Paul SCHERTZ y Jordan J. WALES (eds.), *Encountering Artificial Intelligence: Ethical and Anthropological Investigations*, Oregon: Pickwick Publications, 2023, 262 pp., 15,5 x 23, ISBN 979-8-3852-1030-5.

La Iglesia reconoce que los avances en la Inteligencia Artificial (IA) exigen escrutar los signos de los tiempos e interpretarlos a la luz del Evangelio. En los últimos años la temática de la IA ha captado la atención de la Santa

Sede: la necesidad de una atenta consideración filosófica y teológica sobre su impacto multidimensional goza de consenso unánime. En esta línea destaca la colaboración de la Academia Pontificia para la Vida –con el patrocinio del *Rome Call for AI Ethics*–, y la Pontificia Academia de las Ciencias, así como la de los Dicasterios para la Promoción del Desarrollo Humano Integral, para la Comunicación y para la Cultura y Educación. Dentro de este último cabe mencionar la creación del *Centro de Cultura Digital* con vistas a fomentar el dialogo entre la fe y la tecnología con particular énfasis en las cuestiones éticas relacionadas con el uso de la IA en el ámbito social, económico, político y cultural, entre otros. El presente libro forma parte de la serie *Theological Investigations of Artificial Intelligence* la cual ofrece reflexiones sobre esta revolución tecnológica desde la tradición del pensamiento católico. La publicación de esta serie es fruto de la colaboración entre el *Journal of Moral Theology* y el *Grupo de Investigación en Inteligencia Artificial* del mencionado Centro de Cultura Digital del cual forman parte teólogos y filósofos norteamericanos. *Encountering Artificial Intelligence* reflexiona sobre algunas cuestiones antropológicas y éticas en juego en los debates sobre la IA a la luz de la fe y del Magisterio contemporáneo. El libro se dirige a un público amplio augurando contribuir a que toda persona de buena voluntad pueda tomar decisiones sabias sobre el uso de esta tecnología emergente.

La obra aborda dos conjuntos de preguntas –antropológicas y éticas– en dos partes respectivas precedidas por un primer capítulo que presenta el *status quaestionis* de la reflexión ética sobre la IA. Este capítulo introductorio tiene como fin aproximar al lector a los distintos interlocutores con los cuales se confronta la tradición católica. Entre ellos se cuenta la ética de los Principios –la cual tiene su referente en el Principialismo de Beauchamp y Childress–, la perspectiva de los Derechos Humanos, las posturas Crítica y Casuística, el Utilitarismo y la Ética de la Virtud. La presentación de cada interlocutor permite comprender sus luces y sombras, las cuales responden a las diversas concepciones antropológicas subyacentes. Cabe mencionar que la aproximación más prominente a la ética de la IA es la de los Principios. En efecto, muchas organizaciones y gobiernos elaboran un listado de principios y valores para regir el diseño y el uso de la IA. No obstante, el carácter abstracto y general de los principios y valores no sólo dificulta su aplicación a situaciones concretas, sino que tampoco eximen del riesgo de incurrir en un conflicto entre los mismos. El enfoque de la Ética de la Virtud, por su parte, tiene mucho que ofrecer a una perspectiva católica sobre la ética de la IA puesto que plantea preguntas a un nivel

más profundo que el meramente procedimental, esto es, acerca del florecimiento humano y de bienes humanos específicos. El desarrollo de la IA estimula a actualizar la comprensión tradicional de la ética de la virtud desplegándose hacia lo que S. Vallor denomina virtudes tecnomorales las cuales permiten el florecimiento humano dentro del contexto de la IA y otras tecnologías.

La Primera Parte titulada *Investigaciones Antropológicas* aborda en 4 capítulos cuestiones antropológicas fundamentales: el significado de persona, inteligencia, conciencia y relación. La reflexión sobre dichas cuestiones pone en evidencia las insuficiencias del paradigma tecnocrático, el cual, reduciendo la inteligencia y la personalidad a una dimensión cuantitativa, acaba perdiendo de vista aspectos fundamentales de la persona como son la relación y el don de sí.

En el contexto del desafío a las relaciones humanas que plantea la llamada IA sociable (robots cuidadores, ChatGPT, etc.), el capítulo 2 desarrolla dos interrogantes fundamentales –quién es la persona y qué es la inteligencia–, a fin de iluminar la naturaleza y alcance de este tipo específico de IA. El concepto de *encuentro*, cuyo fundamento es la teología trinitaria, resulta un punto de partida teológico adecuado para desplegar la reflexión antropológica puesto que permite comprender la relación como característica fundamental de la persona y, por tanto, de la imagen de Dios. Los teólogos sistemáticos del último siglo han destacado aspectos novedosos de la imagen como la capacidad para reflejar la vida de un Dios Trino que contiene tanto un “Yo” como un “Tú”. En efecto, esta relación Yo-Tú que existe en la naturaleza de Dios se refleja en nuestra naturaleza relacional. Conocer el recorrido histórico del concepto de razón hasta su actual connotación en el paradigma tecnocrático, arroja luz sobre las premisas teóricas del computacionalismo y del conductismo, posiciones que rigen modelos de IA. Ahora bien, ¿puede una *ratio* sin *intellectus* y sin intencionalidad ser causa de un comportamiento inteligente?

En la misma línea el capítulo 3 reflexiona sobre la dimensión experiencial de la conciencia cual componente necesario de las relaciones interpersonales. La conciencia conlleva la existencia de un sujeto y una experiencia de primera persona que excede el acceso funcional a los datos del ambiente y la producción de un comportamiento. Las explicaciones fenomenológicas de la conciencia evidencian cómo las explicaciones materialistas y funcionalistas de la misma diluyen su carácter eminentemente relacional e intersubjetivo. Desde esta reflexión sobre la conciencia se comprende la imprecisión de la llamada *machine consciousness*, y con ella, de la supuesta amistad con un dispositivo de IA. En efecto, dado que la IA no tiene experiencia consciente de sí misma, es in-

capaz de un verdadero encuentro intersubjetivo. No obstante, cabe preguntarse si puede desempeñar algún rol en el cuidado del vulnerable. El capítulo 4 profundiza en estas cuestiones evidenciando el carácter de simulación de la IA.

Esta Primera Parte se cierra con el capítulo 5 dedicado a las repercusiones espirituales de la IA y, más precisamente, cómo impacta en la relación con Dios. A diferencia de algunas posiciones extremas que consideran la IA como una especie de deidad, la teología cristiana rechaza esta posición por idolátrica. Como se pone de relieve, esta forma de idolatría termina reemplazando a Dios por el hombre puesto que, en última instancia, la IA sería el instrumento mediante el cual el hombre busca imponer control sobre todas las cosas.

Sentadas las bases antropológicas, la Segunda Parte afronta los desafíos éticos de la IA: desde una antropología centrada en la relación y en el encuentro es posible reflexionar sobre el potencial que puede desplegar la IA para proteger y ayudar al desarrollo de la dignidad humana o para su vulneración. Esta parte del libro, en sintonía con las enseñanzas magisteriales de Francisco, constituye un esbozo de una ética de la IA a la luz de la tradición moral católica.

El capítulo 6 pone en relación la IA con la enseñanza social católica. En este contexto, el énfasis de Francisco en la cultura del encuentro guía la reflexión. Con su compromiso de promover la justicia social, la solidaridad y la subsidiariedad, la enseñanza social católica ofrece una sólida propuesta para guiar el uso ético de la IA. En efecto, los avances en IA no deberían amenazar la libertad de las personas, ejercer formas de control o manipulación, aumentar las desigualdades o convertirse en una herramienta mediante la cual algunos poderosos puedan socavar las normas de la convivencia. Por el contrario, la IA debe ser una herramienta para el florecimiento compartido. Siguiendo los pasos del Buen Samaritano, el papa Francisco invita a considerar la IA a la luz de aquellos que están marginados y excluidos.

El capítulo 7 examina las promesas y los riesgos asociados a la IA en los diversos ámbitos de la vida. El análisis sigue el esquema de círculos concéntricos comenzando por el ámbito más íntimo de encuentro que es la familia. No es posible valorar el impacto de la IA en las relaciones familiares sin una adecuada comprensión de la naturaleza de la familia y de su finalidad como institución primaria. Desde esta perspectiva es posible valorar cómo las relaciones de amor propias de esta institución pueden verse afectadas por los imperativos que gobiernan la lógica de la IA como es la optimización y la eficiencia de los resultados. No se trata de un tema menor puesto que la IA está adquiriendo un protagonismo nada despreciable en las relaciones que modelan la dinámica familiar,

sea desde los momentos fundacionales de la familia –por ejemplo, algoritmos que calculan la compatibilidad de las personas, o algoritmos asociados a una medicina perinatal no del todo ajena a una mentalidad eugenésica o a las técnicas de reproducción asistida–, hasta los momentos de mayor vulnerabilidad como son aquellos relacionados con el cuidado de los enfermos. En efecto, uno de los mayores riesgos es incurrir en la tentación de reemplazar el cuidado por la tecnología, la cual es incapaz de empatía y compasión. Estos desafíos conviven con los numerosos beneficios que la IA reporta a la familia como es facilitar la organización de algunas tareas y contribuir a preservar la historia familiar.

Uno de los objetivos fundamentales de la educación es el cultivo de las virtudes. Ahora bien, dado que la mayoría de los softwares educativos tienen un enfoque centrado en adquirir habilidades comerciales, es necesario reflexionar en qué medida y cómo puede la IA ayudar al crecimiento personal en orden a alcanzar los hábitos conducentes a la sabiduría. Es preciso reflexionar también sobre el impacto de la IA para el cultivo de las relaciones interpersonales. En este contexto, especial atención merece el cuidado de la salud. Aunque ciertamente hay beneficios en la integración de la IA en este ámbito, la IA desafía el paradigma humanista del cuidado de la salud restando importancia a la relación personal y reduciendo al paciente a un conjunto de datos cuantificables. El resto del capítulo ilustra cómo impacta la IA en las relaciones humanas que tejen ámbitos más globales como son la política, el mundo del trabajo y la economía, la esfera de la comunicación y de la cultura y, con creciente actualidad, el ámbito militar. De especial interés resulta la reflexión sobre el impacto de la tecnología en la casa común. Es sabido que el discurso público sobre el impacto ambiental de la IA es pobre. Aunque es frecuente pensar en la informática como algo que ocurre en alguna dimensión no material, lo cierto es que se trata de un proceso físico que requiere bastante energía y produce residuos electrónicos no-biodegradables. No menor es la advertencia de Francisco contra la declaración de “independencia de la realidad” la cual se aplica a los intentos de vivir en el ciberespacio: una especie de mundo paralelo que separa de la naturaleza y, por tanto, también del Dios de la naturaleza al comprometer la capacidad contemplativa. Un tal aislamiento se sitúa en la huella del rechazo de la relación con Dios que comenzó con la caída, con la diferencia que ahora no nos escondemos detrás de hojas de higuera, sino en los espacios virtuales, el metaverso y otros entornos artificiales.

Es responsabilidad de la sociedad discernir cómo utilizar la IA de modo que fortalezca el florecimiento humano. A este fin el libro se cierra con un ca-

pítulo de índole más práctica con algunas propuestas particulares para un futuro ético de la IA la cual, sin lugar a duda, necesitará formar parte de la transformación hacia una sociedad más justa que el pensamiento social católico ha demandado desde la Revolución Industrial.

M. Soledad PALADINO
Universidad Austral
DOI 10.15581/006.56.2.512