



*Conferencia Episcopal  
de Colombia*



“INTELIGENCIA  
ARTIFICIAL:

**UNA GALAXIA  
DE REALIDADES  
DISTINTAS”**



*Conferencia Episcopal de Colombia*

**COMISIÓN EPISCOPAL DE DOCTRINA  
DEPARTAMENTO DE DOCTRINA  
COMITÉ TEOLÓGICO**

**“INTELIGENCIA ARTIFICIAL:  
UNA GALAXIA DE REALIDADES  
DISTINTAS”**

***SUBSIDIO DE REFLEXIÓN***

**CXVII ASAMBLEA PLENARIA  
Bogotá, D.C., 01 al 05 de julio de 2024**

## **Comisión Episcopal de Doctrina**

S.E. Mons. Óscar José Vélez Isaza, CMF  
Presidente  
*Obispo de Valledupar*

S.E. Mons. Germán Medina Acosta  
*Obispo Auxiliar de Bogotá*

S.E. Mons. Félix Ramírez Barajas  
*Obispo de Málaga – Soatá*

## **Departamento de Doctrina**

Raúl Ortiz Toro, Pbro.  
*Director*

## **Comité Teológico**

Olga Consuelo Vélez Caro  
Silvio Cajiao Pabón, S.J.  
Guillermo Zuleta Salas, Pbro.  
Gloria Liliana Franco Echeverri, O.D.N.  
María Isabel Gil Espinosa  
José Samuel Torres Tangua, C.Ss.R.  
Santiago Andrés Sierra González

# CONTEMPORANEO

Presentación	5
1. Introducción	11
2. Aspectos positivos de la Inteligencia Artificial	13
3. Algunos desafíos éticos de la Inteligencia Artificial	17
4. Regulación, ética y formación en el uso de la Inteligencia Artificial	19
Bibliografía	22
Anexo 1: Cronología del surgimiento y desarrollo de la Inteligencia Artificial	23
Anexo 2: Glosario de la Inteligencia Artificial	32



# PRESENTACIÓN

El creciente interés de distintos sectores de la sociedad por el tema de la Inteligencia Artificial no deja indiferente a la Iglesia. Particularmente, en este año 2024, el magisterio pontificio ha prestado especial atención al asunto, dedicándole sus mensajes para las Jornadas Mundiales de la Paz (1 de enero) y de las Comunicaciones Sociales (12 de mayo). Además, el Papa Francisco ha participado en la reciente cumbre del G7 (Apulia, Italia, 14 de junio) insistiendo en la necesidad de actuar preventivamente a través de modelos de regulación ética que conlleven a un tratado internacional vinculante que regule el desarrollo y el uso de la Inteligencia Artificial. El Pontífice, en esta cumbre, ha insistido en que todo esto es posible a partir de una sana política que “ponga de nuevo al centro la dignidad de la persona” y que, buscando el bien común, abra camino “a oportunidades diferentes, que no implican detener la creatividad humana y su sueño de progreso, sino orientar esa energía con cauces nuevos”.

Difundir esta propuesta ha sido la principal tarea del “Llamamiento de Roma para la ética de la Inteligencia Artificial” (*Rome call for AI ethics*) firmado en Roma el 28 de febrero de 2020 por la Pontificia Academia para la Vida, Microsoft, IBM, FAO, el Ministerio de Innovación del Gobierno de Italia y suscrito posteriormente por otras entidades religiosas, gubernamentales y empresariales. Hoy en día, esta convocatoria está liderada por la Fundación *RenAIssance* creada por la Pontificia Academia para la Vida por mandato del mismo pontífice en 2021.

El interés del Papa Francisco por este tema es de vanguardia ya que no son menores las preocupaciones acerca de los efectos adversos que puede llegar a traer una Inteligencia Artificial sometida a las leyes del poder despótico, el consumismo salvaje, la competencia desleal, la manipulación mediática, la discriminación y segregación por distintas causas y, en general, rendida a

aquellos ambientes lesivos de la dignidad humana. Él mismo ha rescatado las oportunidades que se presentan, siempre y cuando haya una “orientación del corazón” que ayude a superar la tentación de querer “ser como Dios sin Dios” y por ello empezó a usar, desde 2020, el término “*algorética*” que él mismo define como “moderación ética de los algoritmos y de los programas de inteligencia artificial” (Discurso en la sesión del G7. Apulia, 14 de junio de 2024).

Reconociendo esta realidad, que el Papa llama “ambivalencia” de la Inteligencia Artificial, entre potencialidades y riesgos, la Comisión Episcopal de Doctrina tiene el gusto de presentar este interesante y actual subsidio de reflexión, elaborado en conjunto entre el Departamento de Doctrina y el Comité Teológico, con el acompañamiento de Monseñor Félix Ramírez Barajas, que no pretende ser un producto acabado sobre el tema, sino que, sobre todo, busca presentar de manera propedéutica el asunto, particularmente para los menos entendidos. De manera particular, han sido muy útiles los aportes de la profesora Olga Consuelo Vélez, los padres Guillermo Zuleta y José Samuel Torres, y el profesor Santiago Andrés Sierra.

El documento lleva por título una frase usada por el Santo Padre en su mensaje para la celebración de la ya citada Jornada Mundial de la Paz 2024, cuando, a propósito de la Inteligencia Artificial, señalaba que “debe ser entendida como una galaxia de realidades distintas”. En lo que concierne la labor evangelizadora de la Iglesia, esta “galaxia” plantea numerosos desafíos en igual número de ámbitos; a estos se ha referido la primera sesión de la XVI Asamblea General Ordinaria del Sínodo de los Obispos, reunida en Roma en octubre de 2023 cuando en el Informe de Síntesis “Una Igle-

sia Sinodal en Misión” anotaba que en la actualidad existen “cuestiones que son controvertidas también al interior de la Iglesia, como los efectos antropológicos de las tecnologías y de la inteligencia artificial” (*Informe de síntesis, 15b*); o, también, “algunas cuestiones, como [...] los problemas éticos conectados a la inteligencia artificial, resultan controvertidas no sólo en la sociedad, sino también en la Iglesia, porque suscitan preguntas nuevas (I.S., 15g).

Para ello, la Asamblea Sinodal propuso animar “a los expertos en los diferentes campos del saber a madurar una sabiduría espiritual que haga de su competencia especializada un verdadero servicio eclesial” (I.S., 15i); “crear las condiciones que hagan posible una investigación teológica y cultural que sepa partir de la experiencia cotidiana del Pueblo Santo de Dios y se ponga a su servicio (I.S., 15j); y “promover iniciativas que permitan un discernimiento compartido sobre cuestiones doctrinales, pastorales y éticas controvertidas, a la luz de la Palabra de Dios, de la enseñanza de la Iglesia, de la reflexión teológica y valorando la experiencia sinodal” (I.S., 15k). Es en este sentido que se conformó recientemente en el CELAM el Grupo de Trabajo sobre Inteligencia Artificial, al cual auguramos buenos frutos de reflexión e incidencia.

Junto con estas motivaciones sinodales, reconocemos, además, que el tema de la Inteligencia Artificial está presente en nuestra vida cotidiana, a veces, incluso, tan sutil e imperceptible que podríamos tener la tentación de juzgarlo como irrelevante o solo competencia de las grandes empresas tecnológicas. Y porque nos confronta y exige respuestas en el campo de la evangelización, presentamos este Subsidio de Reflexión, destinado inicialmente a los Obispos con motivo de la próxima CXVII

Asamblea Plenaria, pero con la esperanza de que pueda llegar a muchas personas interesadas en el asunto. Se ha elaborado sin tono alarmista, no solo pensando en los riesgos que la Inteligencia Artificial presenta a la humanidad, sino considerando los aspectos positivos que se convierten en oportunidades y presentando la propuesta de un camino común que lleve, con base en principios éticos, a la regulación mundial y a la formación de las presentes y futuras generaciones.

Así, proponiendo este tema desde las bases, esperamos que continúe creciendo en la humanidad la conciencia de estar frente a unos desafíos teológicos, antropológicos y éticos que no pueden dejarnos indiferentes y que han de ser resueltos a partir de nuestra fe en Jesús que ha venido al mundo a dignificar a la persona humana. Precisamente, el Concilio Vaticano II, en la Constitución Pastoral “Gozos y Esperanzas”, entendiendo a la Iglesia como mediación, nos puso esta tarea: ayudar a que el género humano descubra el sentido de su propia existencia, su verdad más profunda, a partir de la respuesta dada por Dios a todos sus interrogantes, en la persona de su Hijo Jesucristo, el hombre perfecto.

Finalmente, deseo aprovechar esta ocasión para agradecer a la Asamblea Plenaria del Episcopado Colombiano, que en julio de 2021 nos confió a algunos Obispos la delicada tarea de conformar la Comisión Episcopal de Doctrina para el periodo 2021 - 2024. Junto con Mons. Germán Medina Acosta y Mons. Félix Ramírez Barajas, reconocemos los valiosos aportes que, en su momento, también como miembros, hicieron Mons. Hugo Alberto Torres Marín, hoy arzobispo de Santa Fe de Antioquia, quien por este motivo pasó a ser parte de la Comisión Permanente, y Mons. Froilán Casas Ortiz, ahora obispo emérito de Neiva.

Agradezco al padre Raúl Ortiz Toro, quien durante este trienio dirigió el Departamento de Doctrina y que actualmente funge como Secretario Adjunto de la Secretaría General de la Conferencia Episcopal. Al Comité Teológico reiteramos nuestro agradecimiento por su ayuda constante; Asamblea tras Asamblea, con sus reflexiones, facilitan a los Obispos la mejor comprensión de los signos de los tiempos, desde una perspectiva teológica y realista, para poder responder pastoralmente a los desafíos de cada época. Algunos de ellos, como la profesora Olga Consuelo Vélez y los sacerdotes Silvio Cajiao, S.J. y Guillermo Zuleta, llevan varios trienios participando en este órgano consultivo. Desde hace un par de años la Comisión Permanente del Episcopado, además de confirmar para un nuevo periodo a los ya mencionados, eligió a otros igualmente destacados teólogos para completar las sillas de este Comité: la Hermana Gloria Liliana Franco, O.D.N., los profesores María Isabel Gil y Santiago Andrés Sierra, y el padre José Samuel Torres, C.Ss.R.

La Comisión Episcopal de Doctrina y su Departamento ejecutivo, junto con el Comité Teológico, hicieron posible durante este trienio 2021-2024, los siguientes Subsidios de Reflexión:

1. Por una Iglesia Sinodal: Comunión, Participación y Misión (CXII Asamblea Plenaria. Febrero de 2022).
2. Hacia un estilo sinodal de los organismos de participación eclesial (CXIV Asamblea Plenaria. Febrero de 2023).
3. Iglesia Misericordiosa: Reflexión teológica a partir de la realidad en torno a los temas de reconciliación y paz, víctimas, y promoción y defensa de la vida (CXV Asamblea Plenaria. Julio de 2023).

4. “Inteligencia Artificial: una galaxia de realidades distintas”. Oportunidades y desafíos en una Iglesia Sinodal en Misión (CXVII Asamblea Plenaria. Julio de 2024).

Las otras dos Asambleas de este periodo tuvieron como subsidios de reflexión dos documentos de especial trascendencia. La CXIII Asamblea Plenaria (julio de 2022) se dedicó al discernimiento de la Síntesis Nacional del Sínodo sobre la Sinodalidad, confeccionada con la participación del Departamento de Doctrina a partir de las síntesis locales de las 78 Jurisdicciones Eclesiásticas de Colombia; finalmente, la CXVI

Asamblea Plenaria (febrero de 2024) tuvo como principal tema el estudio del Informe de Síntesis “Por una Iglesia sinodal en misión”, elaborado en Roma por la Secretaría General del Sínodo de los Obispos, luego de la conclusión de la primera sesión de la Asamblea Sinodal.

Sirva entonces esta breve presentación como reconocimiento y gratitud, y sea, para todos, ocasión de alabanza a Dios, principio y fin de toda sabiduría.

**+ Óscar José Vélez Isaza, CMF**  
*Obispo de Valledupar*  
*Presidente de la Comisión Episcopal*  
*de Doctrina*

“INTELIGENCIA  
ARTIFICIAL:

**UNA GALAXIA DE**

**REALIDADES**

**DISTINTAS”**



# 1. INTRODUCCIÓN

La Inteligencia Artificial (IA) es un hecho actual que avanza rápidamente e influye en muchas áreas y contextos. Según la UNESCO, no existe una única definición de IA y, ciertamente, esa definición irá cambiando con el tiempo, por los continuos avances que ella representa. Sin embargo, una definición bastante difundida señala que la IA es un conjunto de algoritmos o de conexiones matemáticas que recopilan, reorganizan y reproducen información suministrada por las bases de datos previamente creadas por la inteligencia humana. Este desarrollo tecnológico es un medio poderoso que podría utilizarse para el bien o el mal, dependiendo de su propósito. Por ello, está en nuestras manos la evaluación de su potencial como ayuda para mejorar las condiciones de la vida humana y requiere de una profunda reflexión sobre los posibles peligros asociados al desarrollo de esta tecnología. El futuro de la humanidad dependerá de las decisiones que tomemos ahora, sobre cómo desarrollar y utilizar esta

nueva herramienta. La IA está cambiando rápidamente múltiples aspectos de la sociedad y esta transformación conlleva importantes desafíos éticos y sociales que requieren una atención cuidadosa.

Estos sistemas de procesamiento de la información integran modelos y algoritmos, que imitan con mayor celeridad y precisión nuestras capacidades para aprender y realizar tareas cognitivas, producen a gran escala resultados mecánicos como la predicción y la adopción de decisiones en entornos materiales y virtuales. Están diseñados para funcionar con diferentes grados de autonomía, mediante la modelización y representación del conocimiento, la explotación de datos y el cálculo de correlaciones. Pueden utilizarse en los sistemas ciberfísicos, robóticos y las interfaces entre seres humanos y ordenadores.

Su objetivo es alcanzar una inteligencia general que abarque toda la realidad a dife-

rencia de las inteligencias específicas de las personas concretas que solo abarcan un contexto definido. La IA es capaz de acumular una gran cantidad de datos y de procesarlos en instantes. Permite el desarrollo de sistemas especializados como el reconocimiento de la voz, procesamiento de lenguaje natural, visión por computadora, vehículos autónomos, asistentes virtuales, etc. Se está llegando incluso a emplear redes neuronales artificiales capaces de aprender comportamientos y particularidades de los datos de entrada, para ejecutar actividades de clasificación, detección de objetos, generación de textos, imágenes y audios. En otras palabras, estos sistemas son capaces de generar resultados muy semejantes a los de las personas, infiriendo incluso perspectivas u horizontes acordes con los intereses de quienes acuden a su uso. Es evidente que la IA está cambiando los procesos educativos, las economías, las disciplinas y al mismo ser humano.

Como se acaba de anotar, no existe una sola definición de IA e incluso el Papa Francisco ha afirmado recientemente que debería hablarse de “inteligencias artificiales”, pues el plural de esta proposición ayudaría a que hombres y mujeres de nuestro tiempo sean conscientes de la brecha infranqueable, entre la persona humana con inteligencia natural – aunque limitada y específica –, y los sistemas (ciencias, teorías y técnicas) que buscan reproducir o imitar sus capacidades cognitivas, siempre artificialmente, con una pretensión de universalidad.

Estos aspectos que hemos nombrado y otros que cada día se desarrollan, introducen debates éticos porque la IA, por una parte, puede contribuir decisivamente al desarrollo de la vida humana, solucionando problemas globales a nivel de tecnología,

educación, alimentación, conflictos bélicos o cambios climáticos, pero, por otra, puede empeorarlos (si solo se piensa en buscar los resultados inmediatos sin considerar las consecuencias a largo plazo) o generar nuevos problemas o más discriminaciones (al favorecer solo a algunos sectores y no a toda la población).



## 2. ASPECTOS POSITIVOS DE LA IA

En este horizonte nos preguntamos por los aspectos positivos de la IA que pueden ser acogidos y apoyados por la institución eclesial. Contamos con una clara, concisa y profunda reflexión ofrecida por el Papa Francisco en su Mensaje para la celebración de la 57ª Jornada Mundial de la Paz (1 de enero de 2024), en la que muestra las oportunidades, pero también llama a estar atentos a los desafíos éticos que esta nueva realidad conlleva. En primer lugar, queremos detenernos en aspectos u oportunidades que tiene la IA y que conviene destacar y apoyar.

Como lo reconoció la Conferencia de Aparecida (n. 44) en 2007, estamos ante un “cambio de época cuyo nivel más profundo es el cultural”. En ese momento no se habló de IA, pero hoy podríamos reconocer en ella, uno de esos cambios profundos. Lo primero que se puede señalar es la necesi-

dad de asumir estos cambios, de hacer el tránsito a ese nuevo paradigma para intentar comprenderlo, asumirlo, discernirlo y tomar decisiones éticas adecuadas. De lo contrario, se continuaría juzgando la nueva realidad desde presupuestos antiguos, invocando más el temor, la resistencia, el juicio negativo, en lugar de descubrir las ventajas y posibilidades actuales. La postura de rechazo o demonización del nuevo paradigma solo conlleva a quedarse atrás del caminar histórico, llegando muy tarde a asumir y a acompañar las nuevas realidades.

Ejemplo de cómo una realidad actual debe discernirse sin prejuicios es el Papa Francisco; precisamente, en otro Mensaje, esta vez dedicado a la 58ª Jornada Mundial de las Comunicaciones Sociales 2024, invita a esta actitud positiva: “conviene despejar el terreno de lecturas catastrofistas y de sus

efectos paralizantes. Hace un siglo, Romano Guardini, reflexionando sobre la tecnología y el hombre, instaba a no ponerse rígidos ante lo ‘nuevo’ [...] y concluía: Es cierto que se trata de problemas técnicos, científicos y políticos; pero es preciso resolverlos planteándolos desde el punto de vista humano. Es preciso que brote una nueva humanidad de profunda espiritualidad, de una libertad y una vida interior nuevas”.

Continúa el Papa valorando todo esfuerzo humano por seguir investigando, creando, haciendo avances y progresos: “La Sagrada Escritura atestigua que Dios ha dado a los hombres su Espíritu para que tengan ‘habilidad, talento y experiencia en la ejecución de toda clase de trabajos’ (Ex 35,31). La inteligencia es expresión de la dignidad que nos ha dado el Creador al hacernos a su imagen y semejanza y nos ha hecho capaces de responder a su amor a través de la libertad y del conocimiento. Ciencia y tecnología son producto extraordinario de su potencial creativo”. Por lo tanto, mantener una actitud positiva frente a esta nueva realidad puede vincular la creación de Dios con el progreso humano.

Esta mirada puede complementarse señalando algunos aspectos positivos de la IA:

- La IA presta innumerables beneficios a la sociedad al facilitar la vida de los seres humanos a través de herramientas y aplicaciones en los diversos contextos de la vida, que van desde la atención médica hasta la industria automotriz. Por ejemplo, la longevidad actual es fruto del avance médico capaz de prevenir, detectar y corregir problemas de salud desde el seno materno hasta el último día de existencia.

- La IA permite automatizar y agilizar tareas cuya ejecución resulta repetitiva o requiere de mucho tiempo a los seres humanos.
- Puede analizar datos en menos tiempo que los humanos, permitiendo tomar decisiones fundamentadas con mayor rapidez y asertividad.
- Al disminuir el tiempo que se dedica a las tareas rutinarias que no aportan gran valor ni interés, los empleados pueden centrarse en desarrollar su creatividad.
- La IA prácticamente no comete errores con respecto a los datos que procesa, lo cual es muy común en los seres humanos por cuidadosos que pretendan ser.

Por otra parte, los usos de la IA son innumerables en distintos ámbitos. Señalemos algunos de ellos:

- En el sector de la salud la IA puede analizar los síntomas y permitir un diagnóstico preliminar. Puede llegar a predecir la propensión a ciertas enfermedades. Reduce los riesgos asociados a procesos operatorios y es capaz de intervenir incluso al feto en el vientre materno para corregir malformaciones.
- En el ámbito educativo las ventajas son inmensas. Puede ayudar a modelos de aprendizajes más centrados en el análisis y la reflexión y a crear modelos educativos con necesidades específicas para los diferentes estudiantes.
- La IA optimiza el servicio de transporte tanto en tiempo como en consumo energético. También permite reducir los accidentes en carretera, anticiparse a posibles problemas de los vehículos, planear

rutas de transporte según demanda y eficiencia. Ya es una realidad en los vehículos eléctricos.

- En el sistema financiero la IA es capaz de evitar fraudes, de predecir comportamientos financieros y aconsejar eficientemente para las diversas inversiones.
- En marketing y publicidad prevé los intereses de los individuos y los contextos y es capaz de proponer estrategias eficientes para promover los productos y vaticinar ganancias específicas.
- En cuanto a protección de menores y personas vulnerables, sabiendo que las nuevas tecnologías se han convertido en un elemento de alto riesgo de abusos de todo tipo, la IA puede mejorar el sistema público de identidad digital (SPID) para protegerlos del acceso a las plataformas inadecuadas y también impedir el acoso digital (grooming).
- En el mundo de la evangelización, específicamente en lo que concierne la misión digital, la IA puede ayudar en la creación de *chatbot* para difundir no solo informaciones sino cursos de formación, catequesis de iniciación cristiana, y otros usos. En efecto, actualmente es muy difundido <http://catholic.chat> con respuestas automatizadas del Catecismo de la Iglesia Católica según la edad y conocimiento del interlocutor.

Podrían señalarse más aspectos positivos, pero nos interesa mostrar de qué manera la institución eclesial puede apoyar este nuevo paradigma. En este sentido los grandes desafíos éticos que presenta también pueden encontrar eco en la misión y evangelización de la Iglesia para que apoye,

oriente y promueva aquellos valores que se consideran indispensables para una implementación de la IA que no afecte negativamente a los seres humanos, ni a la creación. En este sentido, el Papa Francisco insistió en su Mensaje para la Jornada Mundial de la Paz 2024 que la dignidad de cada persona y la fraternidad que nos vincula a todos como miembros de una sola familia humana, deben estar en la base del desarrollo de las nuevas tecnologías y servir como criterios indiscutibles para valorarlas antes de su uso. El progreso digital podrá realizarse en el marco del respeto de la justicia y así contribuirá a la causa de la paz. En otras palabras, el mensaje cristiano tiene mucho que aportar al desarrollo de la IA si refuerza su capacidad de comunicar la buena noticia del Reino de Dios y promueve la fraternidad/sororidad como fundamento de la comunidad humana, de manera que todos los creadores o los usuarios de dichas tecnologías posean una formación cristiana suficiente para garantizar que el progreso contribuya al bien común y sea generadora de bien y de verdad.



## 3. ALGUNOS DESAFÍOS ÉTICOS DE LA IA

Si bien los aspectos positivos de la IA nos permiten acercarnos a esta realidad con la esperanza de ser administradores de la creación, sin la pretensión de ocupar el puesto de Dios, también debemos aceptar con sensatez que es constante la tentación de su uso despótico, aplicado en contra de los principios éticos que defienden la dignidad humana. El siglo XXI no deja de sorprendernos con avances en muchas áreas del conocimiento humano que generan, al mismo tiempo, grandes retos y desafíos para la sociedad, al punto de ir transformando a una velocidad inimaginable, como lo afirmaba Klaus Schwab en “La cuarta revolución industrial”.

La IA ubica al ser humano frente a amplias posibilidades, aunque, en algunos casos, su desarrollo y uso tiene consecuencias imprevistas generando riesgos para la vida humana y el medio ambiente. ¿Cómo conciliar, por ejemplo, la producción de los aparatos electrónicos, fundamentales para el

desarrollo de la IA, con la destrucción del medio ambiente a causa de las minas de coltán y otros minerales que son materia prima? ¿Y qué decir de la basura tecnológica que contamina diariamente en cantidades abismales? ¿Cómo olvidar a los campesinos de zonas vulnerables que no cuentan con IA para enfrentar el cambio climático y otros desafíos? La infrautilización de la IA es ciertamente una amenaza para grandes sectores de la población mundial que no cuentan con los medios mínimos para su utilización, lo que conlleva a la poca competitividad y a la marginación. Además, puede conducir a la manipulación de los comportamientos, al prejuicio y aumento de desigualdades, la discriminación en los mercados de trabajo, la desinformación, entre otros.

Hoy es muy común acudir a la asesoría de Siri o Alexa, o dialogar con Chat GPT, o utilizar las herramientas de copilot en Word o en Power Point. Los algoritmos de la IA, en

nuestros celulares, retoman la información que les brindamos, analizan los datos históricos, confirman nuestros deseos, predicen nuestros comportamientos y los ponen a merced de las cadenas comerciales, los sistemas políticos o las ideologías imperantes; y no es un secreto que la hiperinformación de las redes está sujeta a influencias culturales. “Inofensivos” juegos o aplicativos, con la excusa de entretener, recogen nuestros datos biométricos (huellas dactilares, iris ocular, voz, reconocimiento facial, fenotipo, y demás) y la información, que creíamos privada, se vuelve pública y vulnerable. Ya no somos dueños de nuestra propia intimidad. Los datos personales y sensibles entregados en instituciones de salud podrían ser utilizados para descartar a candidatos de un empleo. Los avances sofisticados de la IA han logrado que se difundan contenidos falsos y engañosos (deepfakes), haciendo que, en no pocas ocasiones, sea difícil distinguir entre un contenido real y uno tendencioso.

En fin, sin un “límite algorético” los sistemas de IA pueden ser utilizados para vigilar, controlar, predecir, publicitar. La sociedad actual se encuentra frente a un gran desafío en su uso: maximizar los beneficios y minimizar los riesgos relacionados con su desarrollo, adopción y despliegue. El trabajo debe ser mancomunado entre los desarrolladores, reguladores y usuarios de la IA, para establecer una cultura del uso responsable y ético de esta nueva tecnología. En últimas, el desafío mayor consiste en orientar todos estos cambios hacia el horizonte que beneficie a la sociedad, mantenga el desarrollo sostenible y se ejecute de forma ventajosa para la humanidad acogándose a las normas éticas y estándares mundiales, apoyando la paz y el desarrollo.

Para enfrentar estos desafíos, el Papa Francisco nos ha invitado a actuar de forma

sensata, respetando valores fundamentales como la inclusión, la transparencia, la seguridad, la equidad, la privacidad y la responsabilidad. Un aspecto muy interesante que señala el Pontífice en sus mensajes es asumir el “sentido del límite”, presente siempre en la realidad humana porque “por más prodigiosa que pueda ser nuestra capacidad de cálculo, habrá siempre un residuo inaccesible que escapa a cualquier intento de cuantificación”. Ahí el mensaje cristiano también tiene mucho que aportar para ayudar a reconocer y aceptar el límite de la criatura y acoger la plenitud del don que solo viene de lo alto. Desde ese horizonte podría ser posible que los seres humanos no se dejen encantar del paradigma tecnocrático, animado por una presunción de autosuficiencia, pensando en sobrepasar todo límite con la obsesión de querer controlarlo todo, de perder el control de sí mismo en la búsqueda de la libertad absoluta, cayendo en la dictadura tecnológica.

El sacerdote jesuita Antonio Spadaro, en su libro “Ciberteología”, nos invita a “pensar la fe en tiempos de la red” y, a la vez, a reflexionar “sobre la pensabilidad de la fe a la luz de la lógica de la red”. Son tiempos de renovación misionera y pastoral, de buscar nuevos caminos para anunciar la buena noticia de Jesucristo por las mediaciones actuales, aprovechándolas al máximo y comprendiendo que ellas no son un obstáculo para la vida de la fe sino una oportunidad para mostrar su profundo significado para los tiempos actuales. No son simplemente nuevas mediaciones para anunciar la fe – aunque también lo son –, sino, sobre todo, una nueva manera de considerar la creación, el progreso, la educación, la ciencia y, por supuesto, el ser humano creyente, no en el sentido de “a pesar del progreso”, sino precisamente “contando con el progreso y los grandes avances de la humanidad”.

## 4. REGULACIÓN, ÉTICA Y FORMACIÓN EN EL USO DE LA IA

La *Fundación RenAIssance*, creada por mandato del Papa Francisco en 2021 ([www.romecall.org](http://www.romecall.org)) y la UNESCO, como lo ha manifestado en la “Recomendación sobre la ética de la Inteligencia Artificial” documento publicado el mismo año (<https://www.unesco.org/es/articles/recomendacion-sobre-la-etica-de-la-inteligencia-artificial>), proponen tres áreas de impacto para el manejo de la IA: regulación, ética y formación, considerando seis principios basales: transparencia, inclusión, responsabilidad humana, imparcialidad, confiabilidad, y seguridad y privacidad.

**Estrategias de Regulación:** estas iniciativas de regulación permiten mantener un equilibrio en una variedad de contextos para garantizar que los sistemas de IA se desarrollen y utilicen de manera responsable y favorable. Un aspecto fundamental es la implementación de marcos de evaluación del impacto social en los cuales se deben identificar, analizar y potenciar los

beneficios asociados con los sistemas de IA, así como implementar medidas para prevenir, atenuar y monitorear las posibles amenazas, problemas y riesgos para la civilización humana. Se debe prestar especial atención a los derechos humanos y a las libertades fundamentales, especialmente de las personas marginadas y vulnerables, así como a los impactos en el medio ambiente y a las consecuencias sociales y éticas.

La responsabilidad y rendición de cuentas deben permanecer en manos humanas. Esto significa que, aunque los sistemas de IA pueden asistir en la toma de decisiones y ejecución de tareas, los seres humanos deben mantener el control y la responsabilidad final. Los marcos regulatorios también deben asegurar que los actores de la IA desarrollen algoritmos explicables y entendibles, facilitando así la comprensión y la confianza del público en estas tecnologías. Los mecanismos de gobernanza de la IA deben ser inclusivos, transparentes, multi-

disciplinarios y multilaterales. Esto incluye asegurar que los derechos humanos, las libertades fundamentales y el estado de derecho sean respetados tanto en el ámbito digital como físico.

Ahora bien, como lo ha dicho el Papa Francisco, “la sola reglamentación no es suficiente. Estamos llamados a crecer juntos, en humanidad y como humanidad [...] El reto que tenemos ante nosotros es dar un salto cualitativo para estar a la altura de la sociedad compleja, multiétnica, pluralista, multirreligiosa y multicultural” (Mensaje para la 58ª Jornada Mundial de las Comunicaciones Sociales 2024). Por ello, adquiere sentido la siguiente área de impacto.

**La ética** en la IA implica no solo evitar daños, sino también promover activamente el bien social. La transparencia es un componente clave de la ética en la IA. Las personas tienen derecho a conocer los motivos detrás de las decisiones que afectan sus derechos y libertades. Por lo tanto, los actores de la IA deben comprometerse a desarrollar algoritmos explicables que permitan comprender cómo y por qué se han tomado ciertas decisiones. Por otro lado, se debe promover la auditabilidad y la trazabilidad de los sistemas de IA, ya que permitirá una supervisión efectiva y asegurará que los sistemas operen de manera justa y transparente durante todo su ciclo de vida.

Los sistemas de IA deben probarse exhaustivamente, incluso en condiciones reales, antes de implementarse comercialmente, de manera particular aquellos considerados como riesgos potenciales para los derechos humanos. La ética también requiere revisar los efectos laborales, sociológicos y psicológicos de las decisiones tomadas apoyándose en la IA. Es crucial que estos sistemas no disminuyan la capacidad y au-

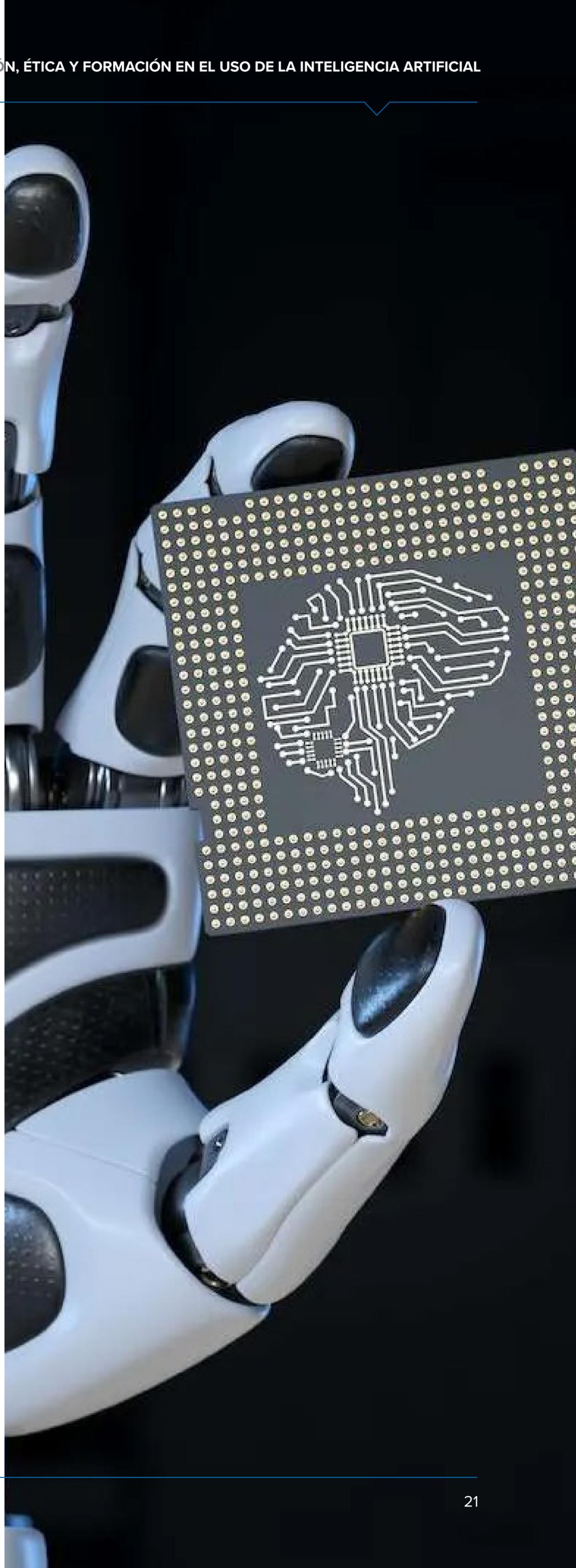
tonomía de las personas para tomar decisiones acertadas, y cualquier riesgo en este sentido debe ser ampliamente investigado y eliminado.

**La formación**, educación en ética de la IA y la sensibilización de la persona es vital para una adopción responsable y beneficiosa de estas tecnologías. El acceso a estas competencias digitales y alfabetización mediática debe ser abierta a todos e impulsada conjuntamente por los gobiernos, la Iglesia, la sociedad civil, las universidades, medios de comunicación, líderes comunitarios y el sector privado. Además, es prioritario que se promueva la diversidad y la inclusión, en los equipos de desarrollo de IA y en los conjuntos de datos de entrenamiento. Estos deben reflejar la diversidad de la población para evitar la perpetuación de desigualdades y prejuicios. En la misma línea, se debe incluir un enfoque particular en las regiones donde los datos son escasos para asegurar que la brecha digital no se amplíe con la adopción de la IA. En últimas, se pretende asegurar que la IA se desarrolle y utilice de manera responsable y ética, en beneficio de toda la humanidad con algunas recomendaciones puntuales:

- Desarrollar e implementar marcos regulatorios completos para la IA basados en los principios de los derechos humanos.
- Establecer principios éticos para guiar el desarrollo y la aplicación de la IA.
- Invertir en la formación de profesionales en ética de la IA, en todos los ámbitos del conocimiento.
- Realizar evaluaciones periódicas y rigurosas del impacto de los sistemas de IA.

- Promover el desarrollo y la implementación de recursos inclusivos y diversos de la IA.
- Establecer un esquema claro de responsabilidad y rendición de cuentas para los sistemas de IA.

La IA llegó para quedarse y su integración a la sociedad no es algo evitable, aunque sí requiere un enfoque equilibrado que contemple estrategias de regulación, principios éticos claros y una formación adecuada. La creatividad, el pensamiento crítico, la responsabilidad y la rendición de cuentas deben recaer en los seres humanos, garantizando así que la IA se utilice para el bien común de la humanidad. La transparencia, la auditabilidad y la trazabilidad son esenciales para construir confianza en estas herramientas informáticas, mientras que la educación y la inclusión aseguran que todas las personas puedan participar y beneficiarse del avance de las nuevas tecnologías. Solo a través de un enfoque integral y multidisciplinario se podrá asegurar que la IA se desarrolle y utilice de manera ética y responsable.



# BIBLIOGRAFÍA

- Benanti, Paolo. *Human in the loop. Decisioni umane e intelligenze artificiali*. Mondadori, 2022.
- Burgos, Jesús Alonso. *Teoría e historia del hombre artificial*. Akal, 2017.
- Cave Peter. *¿Puede ser humano un robot? 33 rompecabezas filosóficos desconcertantes*. Ariel, 2007.
- Colegio de Escritores de la Civiltà Cattolica. IA. Serie Accénti, 13. Editrice Civiltà Cattolica, 2020.
- Dávila Moran, Roberto Carlos, Agüero Corzo, Eucaris del Carmen (2023). *Desafíos éticos de la inteligencia artificial: implicaciones para la sociedad y la economía*. Revista Conrado, 19 (94) 137-144.
- De Martin, Juan Carlos. *Contro lo smartpone. Per una tecnologia più democratica*. ADD Editore, 2023.
- Ferraris, Maurizio – Saracco, Guido. *Tecnosofia. Tecnologia e umanesimo per una scienza nuova*. Editori La Terza, 2023.
- Floridi, Luciano. *Etica dell'Intelligenza artificiale*. Editoriale Raffaello Cortina, 2022.
- Francisco, *Mensaje para la 58 Jornada Mundial de las Comunicaciones Sociales, "Inteligencia artificial y sabiduría del corazón para una comunicación plenamente humana"*, 24 de enero de 2024.
- Francisco, *Mensaje para la celebración de la 57 Jornada Mundial de la Paz, "Inteligencia artificial y paz"*, 8 de diciembre de 2023.
- Francisco, *Discurso en la sesión del G7 sobre Inteligencia Artificial, "Un instrumento fascinante y tremendo"*, 14 de junio de 2024.
- *Intelligenza artificiale (e dintorni) a la prova di filosofia e teología*. En: Archivio Teologico Torinese. Facoltà Teologica dell'Italia Settentrionale. Anno XXX, 2024 No. 1.
- López Baroni, Manuel Jesús. *Bioética y tecnologías disruptivas*. Herder, 2021.
- Peyron, Luca. *Incarneazione digitale. Custodire l'uomo nell'infosfera*. Editrice Elledici, 2019.
- Rapetti, Yuri Berio. *La società senza sguardo. Divinizzazione della tecnica nell'area della teocrazia*. Mimesis Edizioni, 2021.
- Schwab, Klaus. *La cuarta revolución industrial*. Editorial Debate, 2016.
- Sierra, Santiago. *Bien Común: desafío para una sociedad excluyente*. Bogotá: Javeriana, 2020.
- Sierra, Santiago. *La ecología integral es total, es humana*. En: Diez años del Pensamiento Social del papa Francisco: aportes a la Doctrina Social de la Iglesia y perspectivas. Bogotá: Javeriana, 2023.
- Sigman, Mariano – Bilinkis, Santiago. *Artificial. La nueva inteligencia y el contorno de lo humano*. Debate, 2023
- Spadaro, Antonio (2014). *Ciberteología. Pensar el cristianismo en tiempos de la red*. Herder, 2014.
- Tham Joseph – Losito, Massimo. *Bioetica al futuro. Tecnicizzare l'uomo o Umanizzare la tecnica?* Librería Editrice Vaticana, 2010.
- UNESCO. *Recomendación sobre la ética de la inteligencia artificial*, 23 de noviembre de 2021.

## ANEXO 1

# CRONOLOGÍA DEL SURGIMIENTO Y DESARROLLO DE LA INTELIGENCIA ARTIFICIAL

Para ayudar a entender la realidad fenoménica de la IA, se ofrece a continuación una pequeña reseña con los hitos y personajes más relevantes en su diseño, consolidación y proyección, ordenados en forma cronológica por décadas. El anexo 2 ofrece un Glosario de términos útiles para entender y consolidar la apreciación de esta realidad y cuyas consecuencias e implicaciones se han abordado en este documento.

## LOS INICIOS

### 1950: Alan Turing y el Test de Turing

En 1950, el matemático británico Alan Turing publicó un artículo titulado *Computing Machinery and Intelligence* ("Maquinaria informática e inteligencia"), donde planteaba la pregunta fundamental: "¿Pueden pensar las máquinas?". En este trabajo, Turing propuso un experimento mental, ahora conocido como el Test de Turing, para determinar si una máquina puede exhibir un comportamiento inteligente indistinguible del de un humano. El test implica que un humano juzga una conversación entre él mismo y una máquina, sin saber cuál de los interlocutores es cual. Si el juez no puede consistentemente identificar correctamente al humano y a la máquina, la máquina se considera "inteligente".

### 1956: La Conferencia de Dartmouth

En el verano de 1956, se llevó a cabo un taller de dos meses en Dartmouth College, en Hanover, New Hampshire, que reunió a varios jóvenes investigadores interesados en las máquinas inteligentes. Esta conferencia fue organizada por John McCarthy, Marvin Minsky, Nathaniel Rochester y Claude Shannon. Fue en esta conferencia donde John McCarthy acuñó el término IA para describir este nuevo campo de investigación. El objetivo declarado del taller era "explorar cómo hacer que las máquinas usen el lenguaje, formen abstracciones y conceptos, resuelvan tipos de problemas reservados para humanos y se mejoren a sí mismas". Aunque los resultados inmediatos del taller fueron modestos, estableció las bases y la dirección del futuro desarrollo de la IA.

### 1958: John McCarthy inventa el lenguaje de programación Lisp

En 1958, John McCarthy inventó Lisp, un lenguaje de programación que se convertiría en uno de los más utilizados en la investigación de IA durante décadas. Lisp se destacó por su capacidad de manipular símbolos y listas, lo que lo hizo particularmente adecuado para las aplicaciones de IA.

## PRIMEROS DESARROLLOS

### **1960: Desarrollo de programas de ajedrez**

En 1960, se desarrollaron los primeros programas de ajedrez por computadora en el MIT (*Massachusetts Institute of Technology. Instituto de Tecnología de Massachusetts*). Estos programas sentaron las bases para la investigación en IA aplicada a juegos de estrategia.

### **1961: Marvin Minsky publica Steps Toward Artificial Intelligence**

En 1961, Marvin Minsky publicó *Steps Toward Artificial Intelligence* ("Pasos hacia la Inteligencia artificial"), un documento que recogía los primeros pasos y desafíos en el campo de la IA. Este trabajo ayudó a establecer una dirección clara para su investigación futura.

### **1964: ELIZA – Un primer paso hacia la comunicación en lenguaje natural**

En 1964, Joseph Weizenbaum en el MIT desarrolló ELIZA, uno de los primeros programas de procesamiento de lenguaje natural. ELIZA fue diseñada para simular una conversación mediante el uso de un script (guion), siendo el más famoso de ellos el "psicoterapeuta". Utilizaba métodos de patrón de reconocimiento y sustitución de palabras para dar la impresión de entender el lenguaje humano y responder de manera coherente. Aunque su capacidad era superficial y no entendía realmente el contenido de las conversaciones, ELIZA marcó un avance importante en el desarrollo de interfaces de comunicación huma-

no-computadora y demostró las posibilidades del procesamiento del lenguaje natural.

### **1967: DENDRAL**

En 1967, se desarrolló DENDRAL, el primer sistema experto que interpretaba estructuras moleculares. DENDRAL fue un hito importante en la aplicación de la IA a problemas científicos específicos.

### **1969: Shakey el robot – Integración de la percepción y la acción**

En 1969, Shakey el robot, desarrollado por el Stanford Research Institute (SRI), fue el primer robot móvil capaz de razonar sobre su propio comportamiento. Shakey podía moverse de manera autónoma a través de su entorno utilizando cámaras y sensores para percibir obstáculos, y un conjunto de motores para moverse. Lo innovador de Shakey fue su capacidad para hacer mapas de su entorno y usar estos mapas para planificar rutas y resolver problemas simples de navegación y manipulación de objetos. El desarrollo de Shakey combinó la percepción, el razonamiento y la acción física, y fue un gran avance en la robótica y en la IA.

## AVANCES Y APLICACIONES

### **1972: Primer robot humanoide**

En 1972, WABOT-1, el primer robot humanoide del mundo, fue presentado en la Universidad Waseda de Japón. Este robot podía realizar tareas básicas de comunicación y movilidad, y fue un importante paso adelante en la robótica humanoide.

### **1974: Uso de Unimate en General Motors**

En 1974, General Motors utilizó el robot industrial Unimate en una línea de ensamblaje; fue uno de los primeros robots industriales y marcó el comienzo de la automatización en la manufactura.

### **1979: Ordenador vence en backgammon**

En 1979, un ordenador venció a un campeón mundial de backgammon (un juego de mesa antiguo y popular), marcando un avance en la capacidad de las máquinas para competir en juegos más complejos y estratégicos.

## EXPANSIÓN Y CONSOLIDACIÓN

### **1980: Auge de los sistemas expertos**

La década de 1980 es conocida como la "era de oro" de los sistemas expertos, una clase especializada de programas de IA diseñados para emular la toma de decisiones de un experto humano en campos específicos. Estos sistemas utilizaban bases de conocimiento llenas de información sobre un área particular, junto con un motor de inferencia que aplicaba reglas lógicas para resolver problemas y hacer recomendaciones. Ejemplos destacados incluyen MYCIN, desarrollado en la Universidad de Stanford para diagnosticar enfermedades infecciosas, y XCON (también conocido como R1), utilizado por Digital Equipment Corporation para configurar sistemas informáticos. Estos sistemas demostraron el potencial comercial de la IA y llevaron a una amplia adopción en diversas industrias.

### **1980: WABOT-2**

En 1980, WABOT-2, el sucesor del WABOT-1, fue capaz de leer partituras musicales y tocar un teclado, marcando un avance significativo en la robótica humanoide.

### **1981: Sistemas de visión artificial**

En 1981, se desarrollaron los primeros sistemas de visión artificial conocidos como "sombreado y borde", que mejoraron la capacidad de las máquinas para interpretar y entender imágenes visuales.

### **1981: Sistemas de visión artificial**

En 1981, se desarrollaron los primeros sistemas de visión artificial conocidos como

"sombreado y borde", que mejoraron la capacidad de las máquinas para interpretar y entender imágenes visuales.

### **1985: NETtalk**

En 1985, Terrence Sejnowski y Charles Rosenberg desarrollaron NETtalk, un sistema temprano que aprendió a pronunciar palabras de manera similar a como lo hacen los humanos. Este fue un importante avance en el procesamiento del habla y mostró el potencial de las redes neuronales para el aprendizaje de tareas complejas.

### **1986: Renacimiento de las redes neuronales**

A mediados de los 80, las redes neuronales, que habían caído en desgracia en los años 70 debido a limitaciones técnicas y teóricas, experimentaron un renacimiento significativo. En 1986, la publicación del algoritmo de retropropagación por parte de David Rumelhart, Geoffrey Hinton y Ronald Williams permitió el entrenamiento eficaz de redes neuronales de múltiples capas, abriendo nuevas vías para el aprendizaje automático. Este avance revitalizó el interés en las técnicas de aprendizaje profundo, preparando el escenario para los desarrollos futuros que eventualmente llevarían a los modernos sistemas de IA.

### **1987: Primer vehículo autónomo**

En 1987, Ernst Dickmanns convirtió una camioneta Mercedes-Benz en el primer vehículo autónomo dirigido por visión artificial, marcando un avance crucial en la conducción autónoma.

## **INNOVACIONES Y LOGROS**

### **1996: Agentes inteligentes**

En 1996, nacieron los agentes inteligentes que pueden percibir su entorno, aprender de él y tomar decisiones basadas en esa información, marcando un avance significativo en la autonomía de las máquinas.

### **1997: Deep Blue de IBM derrota a Garry Kasparov**

En 1997, Deep Blue, una supercomputadora de ajedrez de IBM, derrotó al campeón mundial Garry Kasparov en un partido de ajedrez. Fue la primera vez que una computadora derrotaba a un campeón mundial en un match bajo condiciones de torneo regulares. Este evento fue crucial para mostrar el potencial de la IA en juegos de estrategia compleja y fue ampliamente cubierto por los medios globales.

### **1998: Introducción de Kismet**

En 1998, Cynthia Breazeal en el MIT introdujo Kismet, un robot con inteligencia emocional capaz de detectar y responder a las emociones humanas, abriendo nuevos caminos en la interacción humano-robot.

### **1999: Lanzamiento de Aibo**

En 1999, Sony lanzó Aibo, un perro robot con habilidades y personalidad que se desarrollan con el tiempo; fue uno de los primeros robots de consumo que mostró el potencial de la robótica en el entretenimiento y la compañía.

## AUGE DE LA INTELIGENCIA ARTIFICIAL

### **2002: Lanzamiento de Roomba**

En 2002, iRobot lanzó Roomba, el primer robot aspiradora autónomo que revolucionó la industria de los electrodomésticos al introducir la robótica en el hogar, permitiendo que las tareas de limpieza se realizaran de manera automática y eficiente.

### **2004: Asimo de Honda**

En 2004, Honda presentó a Asimo, el primer robot humanoide en caminar de forma autónoma. Podía realizar movimientos complejos, como subir y bajar escaleras, lo que marcó un hito importante en la robótica humanoide y la movilidad autónoma.

### **2005: Stanley gana el DARPA Grand Challenge**

En 2005, Stanley, un coche autónomo desarrollado por la Universidad de Stanford, ganó el DARPA Grand Challenge. Este logro demostró el potencial de los vehículos autónomos para navegar y tomar decisiones de manera independiente en entornos reales y complejos.

### **El Amanecer del Aprendizaje Profundo**

En 2006, Geoffrey Hinton y sus colegas introdujeron una nueva técnica para entrenar redes neuronales profundas, conocida como "entrenamiento previo no supervisado". Este método permitió a las redes neuronales aprender características útiles sin necesidad de supervisión directa en las pri-

meras etapas del entrenamiento. Este avance resolvió el problema de la desaparición del gradiente que anteriormente limitaba la efectividad de las redes profundas y marcó el comienzo de la "era del aprendizaje profundo" en la IA.

## INTELIGENCIA ARTIFICIAL EN LA VIDA COTIDIANA

### **2011: IBM Watson y la Competencia en Jeopardy! e introducción de Siri**

En 2011, el mundo presenció un hito significativo en el campo de la IA cuando IBM Watson compitió en el juego de televisión "Jeopardy!" contra dos de los campeones más exitosos del juego. Watson, un sistema de IA diseñado para procesar y entender lenguaje natural, no solo compitió, sino que ganó, demostrando un entendimiento y una capacidad de respuesta a preguntas formuladas en lenguaje natural que superaban todo lo visto hasta entonces. Este evento mostró las capacidades prácticas de la IA e impulsó el interés y la inversión en tecnologías de procesamiento del lenguaje natural y análisis de datos. Ese mismo año, Apple lanzó Siri, un asistente personal de reconocimiento de voz que utilizaba IA para interactuar con los usuarios mediante comandos de voz, marcando un hito en la accesibilidad y uso cotidiano de la IA.

### **2012: AlexNet**

En 2012, Alex Krizhevsky, Ilya Sutskever y Geoffrey Hinton desarrollaron AlexNet, un modelo de red neuronal convolucional. AlexNet ganó por un amplio margen la competencia ImageNet Large Scale Visual Recognition Challenge en 2012, y su éxito

fue fundamental para el resurgimiento del interés en el aprendizaje profundo dentro de la comunidad científica y tecnológica.

### **2013: Creación del robot Atlas**

En 2013, Boston Dynamics creó Atlas, un robot bípedo humanoide capaz de realizar una amplia variedad de tareas que requerían destreza y movimientos precisos, avanzando significativamente en la robótica humanoide.

### **2014: Los Coches Autónomos de Google, Eugene Goostman pasa la prueba de Turing y Lanzamiento de Alexa**

En 2014, Google (ahora parte de Alphabet Inc.) realizó pruebas significativas de sus coches autónomos en vías públicas. Estos coches utilizaban avanzadas tecnologías de IA, como el aprendizaje profundo para el procesamiento de imágenes y la percepción sensorial, y algoritmos de decisión para la navegación y el manejo de situaciones de tráfico. Este avance no solo destacó la capacidad de la IA para realizar tareas complejas y críticas, sino que también planteó importantes cuestiones sobre la seguridad, la regulación y la ética en la robótica autónoma. Ese mismo año, Eugene Goostman, un chatbot, pasó la prueba de Turing, demostrando que podía engañar a los humanos haciéndoles creer que estaban conversando con otra persona. Además, Amazon lanzó Alexa, un asistente virtual inteligente que utiliza IA para responder a preguntas, reproducir música, controlar dispositivos domésticos inteligentes y mucho más.

### **2015: Asistente IA de Facebook y AlphaGo de DeepMind derrota al Campeón Europeo de Go**

En 2015, Facebook anunció su asistente personal inteligente llamado "M". Este asistente combinaba el procesamiento de lenguaje natural y la IA para responder preguntas y realizar tareas. M podía manejar una variedad de solicitudes, desde hacer reservas en restaurantes hasta enviar mensajes. Sin embargo, el proyecto completo fue eventualmente discontinuado en 2018 debido a su incapacidad para escalar eficazmente y a los altos costos asociados con su mantenimiento. A pesar de su discontinuación, marcó un importante desarrollo en el uso de la IA para asistentes virtuales personalizados.

En octubre de 2015, AlphaGo, desarrollado por DeepMind (una empresa adquirida por Google), derrotó a Fan Hui, el campeón europeo de Go. Esta victoria fue un prelude importante para el posterior enfrentamiento de AlphaGo con Lee Sedol.

### **2016: Microsoft lanza Tay y AlphaGo de DeepMind derrota al Campeón Mundial de Go**

En marzo de 2016, Microsoft lanzó Tay, un chatbot diseñado para interactuar con los usuarios en redes sociales; fue desarrollado para aprender de las conversaciones con los usuarios de Twitter. Sin embargo, rápidamente se volvió polémico debido a que hizo comentarios racistas y ofensivos después de ser influenciado por usuarios malintencionados. Este incidente resaltó los desafíos éticos y técnicos en el desarrollo de IA social, y Microsoft retiró a Tay poco después de su lanzamiento para

hacer ajustes y evitar futuros problemas similares.

En marzo de 2016, AlphaGo, desarrollado por DeepMind, derrotó a Lee Sedol, uno de los mejores jugadores de Go del mundo, en una serie de cinco partidas, ganando 4-1. Este evento fue ampliamente cubierto por los medios y fue considerado un hito significativo en la IA, demostrando su capacidad para manejar y ganar juegos de estrategia altamente complejos. La victoria subrayó el potencial de la IA para superar las capacidades humanas en tareas que requieren pensamiento estratégico profundo y adaptativo.

### **2017: Sophia, la ciudadana robot y Primer servicio de taxi autónomo**

En 2017, Sophia, un robot humanoide diseñado por Hanson Robotics, fue reconocida como ciudadana saudí, marcando un cambio significativo en la percepción social de los robots y planteando nuevas preguntas sobre los derechos y roles de las máquinas inteligentes. También en 2017, Waymo, una empresa pionera en vehículos autónomos, lanzó el primer servicio de taxi sin conductor en Arizona, marcando un avance importante en la movilidad autónoma y el transporte inteligente.

### **2018: Atlas aprende a hacer parkour**

En 2018, el robot Atlas de Boston Dynamics aprendió a hacer parkour, demostrando avances impresionantes en la agilidad y capacidad física de los robots humanoides. El parkour (del francés *parcours*, significando recorrido) es una actividad física basada en la capacidad motriz del individuo, desarrollada a partir del método natural. Tiene como objetivo trasladarse de un punto a otro del entorno de la manera más sencilla

y eficiente posible, adaptándose a las exigencias del mismo con la sola ayuda de su cuerpo.

## **MOMENTOS CLAVES EN LA INTELIGENCIA ARTIFICIAL**

### **GPT-1 (junio de 2018)**

El primer modelo Generative Pre-trained Transformer (GPT) fue lanzado por OpenAI en 2018. GPT-1 tenía 117 millones de parámetros y fue entrenado en un corpus masivo de texto para predecir la siguiente palabra en una secuencia. Esta versión demostró que los modelos preentrenados en grandes conjuntos de datos podían ser ajustados (fine-tuned) para tareas específicas, lo que estableció la base para futuros desarrollos en modelos de lenguaje natural.

### **GPT-2 (febrero de 2019)**

GPT-2 amplió significativamente el tamaño del modelo a 1.5 mil millones de parámetros. Este modelo mostró capacidades impresionantes para generar texto coherente y relevante a partir de un prompt dado, lo que aumentó el interés y la preocupación sobre los posibles usos indebidos de la tecnología. GPT-2 se destacó por su capacidad para realizar tareas de generación de texto con una fluidez notable.

### **GPT-3 (febrero de 2020)**

GPT-3, con 175 mil millones de parámetros, representó un salto enorme en términos de capacidad y rendimiento. GPT-3 fue entrenado con un corpus aún más grande y demostró habilidades avanzadas en generación de texto, comprensión y respuesta en lenguaje natural, programación asistida por

IA, y más. GPT-3 ha sido utilizado en una amplia gama de aplicaciones comerciales y académicas, mostrando la versatilidad y el poder de los grandes modelos de lenguaje.

### **2021: Presentación de DALL-E**

En 2021, OpenAI presentó DALL-E, un programa de IA generador de imágenes a partir de descripciones textuales, revolucionando la capacidad de la IA para crear arte y gráficos de alta calidad basados en instrucciones textuales.

### **2022: Creación de Robotaxi y Lanzamiento de ChatGPT**

En 2022, la compañía Cruise, filial de General Motors, creó Robotaxi, permitiendo pedir un coche sin conductor en San Francisco, lo que marcó un avance significativo en la movilidad autónoma. Ese mismo año, OpenAI lanzó ChatGPT, un chatbot capaz de mantener conversaciones naturales, mejorando la interacción humano-computadora.

### **ChatGPT (noviembre de 2022)**

Es una versión ajustada específicamente de GPT-3 para mejorar la interacción conversacional. Este modelo fue entrenado para manejar conversaciones de manera más coherente y contextualmente relevante, mejorando su capacidad para mantener la coherencia a lo largo de interacciones más largas. ChatGPT ha sido implementado en diversas aplicaciones que requieren interacción humana, proporcionando respuestas naturales y útiles.

### **2023: Inversión de Microsoft en OpenAI**

Microsoft anunció una inversión multimillonaria en OpenAI y la integración de ChatGPT en sus plataformas, incluyendo Bing y Windows. Esta inversión ha permitido una mayor accesibilidad y uso de tecnologías avanzadas de IA en aplicaciones cotidianas.

### **GPT-4 (marzo de 2023)**

GPT-4, la última versión hasta la fecha, introdujo mejoras adicionales en términos de tamaño, con más parámetros que GPT-3, y refinamientos en sus capacidades de razonamiento y comprensión. GPT-4 aborda algunas de las limitaciones de sus predecesores, reduciendo aún más la generación de respuestas incoherentes o irrelevantes, y mejorando la fiabilidad y la ética del modelo. Este modelo ha ampliado aún más las aplicaciones de la IA en áreas complejas que requieren un entendimiento profundo del contexto.

### **2023: Inversiones de Google, Meta y otras empresas y Acuerdo en la Casa Blanca para regular la IA**

En 2023, Google, Meta y muchas otras empresas anunciaron inversiones y desarrollos propios de IA, reflejando la creciente competencia y avances en el campo. Además, líderes mundiales firmaron un acuerdo en la Casa Blanca para regular la IA, subrayando la importancia de las consideraciones éticas y de seguridad en su desarrollo.

**2024: Lanzamiento de ChatGPT-4o**

OpenAI lanzó ChatGPT-4o, una versión avanzada con mejoras significativas en la comprensión del contexto, precisión de respuestas, capacidades multimodales y personalización.

**2024: Nace Sora**

En 2024, OpenAI lanzó Sora, una herramienta capaz de generar videos a partir de descripciones textuales, ampliando las capacidades de generación de contenido multimedia de la IA.



## ANEXO 2

# GLOSARIO DE INTELIGENCIA ARTIFICIAL CON ÉNFASIS EN ANÁLISIS ÉTICO

**Agentes Inteligentes:** Sistemas de software que perciben su entorno, toman decisiones autónomas y realizan acciones para alcanzar objetivos específicos.

**Algorética:** La ética de los algoritmos. Se enfoca en las implicaciones éticas del diseño, desarrollo y uso de algoritmos, especialmente aquellos utilizados en IA y aprendizaje automático. Los temas clave incluyen transparencia, justicia, privacidad, responsabilidad y seguridad.

**Algoritmo:** Un conjunto de instrucciones paso a paso diseñadas para realizar una tarea o resolver un problema.

**Algoritmo de Retropropagación:** Un método de entrenamiento para redes neuronales que ajusta los pesos de las conexiones minimizando el error de salida mediante la propagación de ese error desde la salida hasta las capas internas.

**Análisis Predictivo:** El uso de datos históricos, algoritmos de aprendizaje automático y técnicas estadísticas para predecir resultados futuros.

**Aprendizaje Automático (*Machine Learning*):** Un subcampo de la IA que se enfoca en el desarrollo de algoritmos que permiten a las computadoras aprender a partir de datos y hacer predicciones o decisiones

sin ser programadas explícitamente para realizar la tarea.

**Aprendizaje por Refuerzo (*Reinforcement Learning*):** Un tipo de aprendizaje automático donde un agente aprende a tomar decisiones optimizando una recompensa acumulada a través de interacciones con el entorno.

**Aprendizaje por Transferencia (*Transfer Learning*):** Una técnica en la que un modelo entrenado en una tarea se reutiliza como punto de partida para un modelo en una tarea relacionada.

**Aprendizaje Profundo (*Deep Learning*):** Una rama del aprendizaje automático basada en redes neuronales artificiales con múltiples capas (redes neuronales profundas) que pueden aprender representaciones jerárquicas de datos. Es especialmente efectiva en tareas como el reconocimiento de imágenes y el procesamiento del lenguaje natural.

**Autonomía:** El grado en que un sistema de IA puede operar y tomar decisiones de manera independiente. Los niveles de autonomía varían desde sistemas completamente controlados por humanos hasta sistemas completamente autónomos.

**Bien Común:** Un principio ético que se enfoca en el beneficio colectivo y el bienestar de toda la sociedad en lugar de los intereses individuales, crucial en la aplicación de la IA en políticas públicas y servicios sociales.

**Big Data:** Conjuntos de datos extremadamente grandes y complejos que requieren tecnologías avanzadas de almacenamiento, procesamiento y análisis para extraer información valiosa.

**Chatbot:** Un programa de computadora que simula una conversación con usuarios humanos, utilizando procesamiento del lenguaje natural para responder a preguntas y mantener diálogos.

**Computación Cuántica:** Un paradigma emergente de computación que utiliza principios de la mecánica cuántica para realizar cálculos a velocidades mucho mayores que las computadoras clásicas, con potenciales aplicaciones significativas en IA.

**Consentimiento Informado:** El proceso de asegurar que las personas comprendan y acepten de manera explícita cómo se utilizarán sus datos por los sistemas de IA. Es un componente esencial de la ética de la IA y la privacidad de los datos.

**Datos de Entrenamiento:** Conjunto de datos utilizados para entrenar un modelo de aprendizaje automático. Estos datos permiten que el modelo aprenda patrones y haga predicciones.

**Derechos Humanos:** Un enfoque ético que asegura que el desarrollo y uso de la IA debe respetar y promover los derechos humanos fundamentales, como la dignidad, la privacidad y la libertad.

**Desigualdad Tecnológica:** El fenómeno donde el acceso desigual a tecnologías avanzadas, como la IA, exacerba las disparidades socioeconómicas y de oportunidades entre diferentes grupos sociales y regiones.

**Ética de la IA:** Un campo de estudio que aborda las implicaciones morales y éticas del desarrollo y uso de la IA, incluyendo temas como la privacidad, el sesgo algorítmico y la responsabilidad de las decisiones automáticas.

**Explicabilidad:** La capacidad de un sistema de IA para proporcionar explicaciones claras y comprensibles de sus decisiones y acciones. La explicabilidad es importante para la confianza del usuario y la supervisión regulatoria.

**Generative Pre-trained Transformer (GPT):** Una serie de modelos de lenguaje desarrollados por OpenAI que utilizan la arquitectura de transformadores y se entrenan en grandes corpus de texto para generar lenguaje natural coherente y relevante.

**Gobernanza de la IA:** El conjunto de políticas, normas y regulaciones que guían el desarrollo y uso de la IA, asegurando que se adhiera a principios éticos y beneficie a la sociedad en general.

**Hiperparámetros:** Parámetros que se configuran antes de iniciar el proceso de aprendizaje y que no se ajustan durante el entrenamiento del modelo. Ejemplos incluyen la tasa de aprendizaje y el número de capas en una red neuronal.

**IA Fuerte (Strong AI) vs. IA Débil (Weak AI):** La primera es la hipótesis de que una máquina puede poseer una mente, conciencia y entendimiento igual o superior al

humano, mientras la segunda afirma que los sistemas de IA son diseñados para realizar tareas específicas sin tener conciencia o entendimiento genuino.

**Imagen de Entrenamiento:** La imagen utilizada como entrada para entrenar modelos de reconocimiento de imágenes y visión por computadora.

**Impacto Social:** Las consecuencias y efectos de la implementación de sistemas de IA en la sociedad, incluyendo consideraciones sobre el empleo, la educación y la interacción social.

**Interpretabilidad:** La capacidad de un modelo de IA para ser entendido y explicado por los humanos, especialmente en términos de cómo se toman las decisiones.

**Justicia (Fairness):** El principio de que los sistemas de IA deben tratar a todas las personas de manera justa, sin discriminación ni parcialidad. La equidad es fundamental para garantizar que la IA beneficie a todos los segmentos de la sociedad por igual.

**Posthumanismo:** Es una corriente de pensamiento que aspira a la superación del humanismo. Algunos consideran que es la última etapa del transhumanismo donde el ser humano se integrará profundamente con la tecnología, como, por ejemplo, que un computador sustituya la mente de un individuo humano, o que las mejoras genéticas aporte características nuevas al actual ser humano.

**Privacidad de Datos:** El derecho de los individuos a controlar cómo se recopilan, utilizan, comparten y almacenan sus datos personales. Es un aspecto crucial en el desarrollo y uso de sistemas de IA que manejan información sensible.

**Procesamiento del Lenguaje Natural (NLP):** Un campo de la IA que se enfoca en la interacción entre computadoras y humanos mediante el lenguaje natural. Incluye tareas como el análisis de sentimientos, la traducción automática y la generación de texto.

**Proporcionalidad:** El principio de que los beneficios y riesgos de la IA deben ser equilibrados de manera justa, asegurando que las tecnologías no causen daño desproporcionado a ningún grupo social.

**Prueba de Turing:** Un experimento propuesto por Alan Turing para determinar si una máquina puede exhibir inteligencia. Si un juez humano no puede distinguir consistentemente entre las respuestas de una máquina y las de un humano, la máquina se considera "inteligente".

**Red Neuronal Convolutiva (CNN):** Un tipo de red neuronal profunda diseñada para procesar datos estructurados en forma de una cuadrícula, como imágenes. Las CNN utilizan convoluciones para extraer características locales y combinarlas en representaciones más abstractas.

**Red Neuronal Recurrente (RNN):** Un tipo de red neuronal que se especializa en procesar datos secuenciales, como series temporales o texto. Las RNN tienen conexiones que forman bucles, permitiendo que la información persista y se use en pasos futuros.

**Redes Neuronales Artificiales:** Modelos computacionales inspirados en el cerebro humano que consisten en capas de nodos (neuronas) interconectados. Cada conexión tiene un peso ajustable que se aprende durante el entrenamiento del modelo.

**Regularización:** Una técnica utilizada para prevenir el sobreajuste (overfitting), añadiendo un término de penalización a la función de pérdida del modelo durante el entrenamiento.

**Resiliencia Tecnológica:** La capacidad de los sistemas de IA para resistir y recuperarse de fallos, ataques y errores, asegurando su fiabilidad y seguridad a largo plazo.

**Responsabilidad (*Accountability*):** El principio de que los desarrolladores y operadores de sistemas de IA deben ser responsables de las acciones y decisiones de esos sistemas. Incluye la necesidad de tener mecanismos para auditar y revisar el desempeño y las decisiones de la IA. De allí que la Inteligencia Artificial Responsable sea un enfoque de desarrollo de IA que prioriza la seguridad, la ética y la equidad, asegurando que las tecnologías sean beneficiosas para todos los segmentos de la sociedad.

**Roboética:** El campo de estudio que aborda las implicaciones éticas del diseño, desarrollo y uso de robots y sistemas autónomos, incluyendo su impacto en la sociedad y la moralidad humana.

**Salvaguardias:** Medidas y mecanismos implementados para proteger contra los riesgos y efectos adversos potenciales de los sistemas de IA, incluyendo la seguridad, la privacidad y la equidad.

**Sesgo Algorítmico:** La presencia de prejuicios o favoritismos sistemáticos en los algoritmos de IA, que pueden resultar en decisiones injustas o discriminatorias. Esto puede ocurrir debido a datos de entrenamiento no representativos o enfoques de modelado inadecuados.

**Sesgo de Confirmación:** La tendencia de los sistemas de IA a reforzar prejuicios existentes debido a datos de entrenamiento sesgados, lo que puede perpetuar injusticias y estereotipos.

**Sistema Experto:** Un programa de computadora que emula la toma de decisiones de un humano experto en un campo específico, utilizando una base de conocimiento y un motor de inferencia para resolver problemas o hacer recomendaciones.

**Sobreajuste (*Overfitting*):** Cuando un modelo de aprendizaje automático se ajusta demasiado a los datos de entrenamiento, capturando el ruido en lugar de los patrones generales, lo que lleva a un mal rendimiento en datos no vistos.

**Subajuste (*Underfitting*):** Cuando un modelo es demasiado simple para capturar la estructura subyacente en los datos de entrenamiento, resultando en un mal rendimiento tanto en los datos de entrenamiento como en los no vistos.

**Tecnofobia:** El miedo o aversión a las nuevas tecnologías, incluida la IA, debido a preocupaciones sobre su impacto en el empleo, la privacidad y la seguridad.

**Tokenización:** El proceso de dividir texto en unidades más pequeñas (tokens), como palabras o subpalabras, que se utilizan como entrada para modelos de procesamiento del lenguaje natural.

**Transformadores (*Transformers*):** Una arquitectura de modelos de aprendizaje profundo que utiliza mecanismos de atención para procesar datos secuenciales y paralelizarlos, mejorando la eficiencia en tareas de procesamiento del lenguaje natural.

**Transhumanismo:** Movimiento que propugna la superación de las limitaciones actuales del ser humano, tanto en sus capacidades físicas como psíquicas, mediante el desarrollo de la ciencia y la aplicación de los avances tecnológicos.

**Transparencia:** La práctica de hacer que los procesos, decisiones y algoritmos de los sistemas de IA sean comprensibles y accesibles para los usuarios y reguladores. La transparencia es clave para la confianza y la rendición de cuentas en la IA.

**Validación Cruzada:** Una técnica utilizada para evaluar el rendimiento de un modelo de aprendizaje automático, dividiendo los datos en partes y utilizando diferentes partes para entrenamiento y validación.



*Conferencia Episcopal  
de Colombia*