



Via Inveniendi Et Iudicandi
ISSN: 1909-0528
revistaviei@usantotomas.edu.co
Universidad Santo Tomás
Colombia

Zabala Leal, Tatiana Dulima
La ética en inteligencia artificial desde la perspectiva del derecho*
Via Inveniendi Et Iudicandi, vol. 16, núm. 2, 2021, Julio-Diciembre
Universidad Santo Tomás
Bogotá, Colombia

Disponible en: <https://www.redalyc.org/articulo.oa?id=560269078009>

- ▶ Cómo citar el artículo
- ▶ Número completo
- ▶ Más información del artículo
- ▶ Página de la revista en redalyc.org

La ética en inteligencia artificial desde la perspectiva del derecho*

[Artículos]

Tatiana Dulima Zabala Leal**

Fecha de recepción: 4 de marzo de 2021

Fecha de aprobación: 18 de mayo de 2021

Citar como

Zabala Leal, T. D. (2021). La ética en inteligencia artificial desde la perspectiva del Derecho. *Via Inveniendi Et Iudicandi*, 16(2).

<https://doi.org/10.15332/19090528.6785>



Resumen

El presente artículo es producto de un estudio que tiene como objetivo demostrar la relevancia de la ética en el desarrollo de la inteligencia artificial, por lo que presentará una exploración de las posturas de expertos y diferentes organismos internacionales, entre ellos la Unión Europea, los cuales coinciden al establecer paradigmas similares sobre el alcance potencial de la autonomía de la IA como ser dotado de capacidad cognitiva o de inteligencia, pero no de conciencia o razón, por lo que el

* El presente artículo está adscrito al proyecto de investigación Marco Normativo de las Industrias Creativas, Tecnología, Contratos y Arbitraje en Colombia de la Institución Universitaria Politécnico Grancolombiano (Bogotá, Colombia).

** Magíster en Derecho Comercial y Contratos Internacionales. Candidata a doctora en Derecho empresarial, abogada, docente investigadora, Grupo de Investigación Derecho, Sociedad e Industrias Creativas, de la Escuela de Derecho y Gobierno, línea de investigación en Derecho de la Información y las Tecnologías, del Institución Universitaria Politécnico Grancolombiano (Bogotá, Colombia). Correo electrónico: tzabala@poligran.edu.co; ORCID <https://orcid.org/0000-0001-8938-7106>; CvLac: http://scienti.colciencias.gov.co:8081/cvlac/visualizador/generarCurriculoCv.do?cod_rh=0000155298

derecho a la libertad desempeña un papel trascendental en la definición de la categoría de la IA como sujeto.

Para acercar al lector al objeto de estudio, se expondrán los objetivos y las categorías de la conceptualización de los objetivos de la IA, la confrontación entre autonomía y automatización e inteligencia y razonamiento, de tal manera que este pueda hacer una análisis crítico que le permita tener una postura propia sobre la generación de responsabilidades jurídicas en cabeza de la IA como sujeto de derecho y sus programadores.

De igual forma, se abordarán temas como el transhumanismo, entendido como un mecanismo para mejorar la vida y desempeño del humano, la bioética, la cual debe partir del respeto a la dignidad humana y la protección de los derechos fundamentales para, finalmente, concretar las posturas de los expertos y de la Comisión Europea sobre la adaptación de la Carta de los Derechos Fundamentales para la consecución de una IA confiable y ética.

Palabras clave: inteligencia artificial, digitalización del derecho, ética, smart contracts, blockchain, token, botsourcing, heurística, algoritmo, autodeterminación, biohacking, transhumanismo.

Ethics in Artificial Intelligence from the perspective of Law

Abstract

This article is the product of a study that aims to demonstrate the relevance of ethics in the development of artificial intelligence, for which it will present an exploration of the positions of experts and different international organizations, including the European Union, which coincide in establishing similar paradigms on the potential scope of the autonomy of AI as a being endowed with cognitive capacity or intelligence, but not with conscience or reason, so the right to freedom

plays a transcendental role in defining the category of AI as an individual.

In order to bring the reader closer to the object of study, the objectives and categories of the conceptualization of the purposes of AI, the confrontation between autonomy and automation and intelligence and reasoning, will be presented, so that the reader can make a critical analysis that allows him to have his own position on the generation of legal responsibilities of the AI as a legal person and its programmers.

Similarly, issues such as transhumanism, understood as a mechanism to improve human life and performance, bioethics, which should be based on respect for human dignity and the protection of fundamental rights, will be addressed, to finally, specify the positions of experts and the European Commission on the adaptation of the Charter of Fundamental Rights to achieve a reliable and ethical AI.

Keywords: artificial intelligence, digitalization of law, ethics, smart contracts, blockchain, token, botsourcing, heuristics, algorithm, self-determination, biohacking, transhumanism.

Introducción

El eje central de este artículo es la ética vinculada a la inteligencia artificial (IA) y que será abordado a partir de la perspectiva de la legislación internacional en lo que respecta a tratados, acuerdos y convenios promulgados por la Unión Europea (UE). Estos han venido contribuyendo a la generación de diversos axiomas, que no han sido apropiados por la doctrina nacional e internacional con la importancia que debería asignárseles, razón por la que por medio de esta investigación se pretende mostrar el impacto y la funcionalidad de la aplicación de la ética en el proceso de digitalización del derecho.

En razón a que la IA, junto con sus aplicaciones y servicios conexos, ha sido objeto de múltiples estudios, de los que han resultado, en la última

década, investigaciones y publicaciones elaboradas tanto por la comunidad científica, como por el sector empresarial y la académica. Por esto, es oportuno aseverar que estos se han centrado en tratar aspectos técnicos propios del funcionamiento habitual que de ellos se hace en la cotidianidad de los contextos sociales, ya que se encuentran presentes en actos tan simples como el escuchar música, la apertura de una cuenta virtual, la suscripción de *smart contracts*, la implementación de *blockchain*, *token* o *botsourcing*, entre otros.

Sin embargo, pese a la simpatía que generan las innumerables funciones de la IA, el desarrollo de esta también tiende a suscitar temor sobre los riesgos que el uso de esta implica; motivando investigaciones y estudios acerca del peligro que este tipo de tecnología supone para el ser humano; puesto que, de acuerdo con lo dicho por Marín (2019), algunas aplicaciones derivadas de la IA han venido generando conflictos éticos, como es el caso de los coches autónomos, en cuanto al hecho de tener en cuenta situaciones anormales como el encuentro con un peatón imprudente u otro vehículo invadiendo el mismo carril, acontecimientos en los que la IA implementada en el auto debe tener la capacidad de realizar proceso heurísticos para la toma de decisiones que den resolución al conflicto. Esta situación implica que el algoritmo base tenga un patrón de conducta ética, lo que, según Bonnefon, Sharif y Rahwan (2016), pondría a la máquina ante un dilema ético como, por ejemplo, ¿qué vida habrá de sacrificarse: la del conductor o la del peatón que cruza por donde no debe?; este problema vinculado a la autonomía no sería el único, pues otras tipologías de contrariedades éticas, como la privacidad, la seguridad de los datos personales o la creación de hábitos de consumo también preocupan a los consumidores de este tipo de servicios.

Y es que definitivamente, como lo expone Argandoña (2019), cualquier tecnología producto de la conciencia humana puede concluir en un uso beneficioso o nocivo para ellos mismos. Por esta razón, se puede decir que no todos los desarrollos tecnológicos son recomendables, afirmación respaldada por Martin y Freeman (2004), quienes aseveran que una de las perspectivas erradas de la comunidad en general es considerar la IA siempre como un fenómeno positivo, ya que no todas las veces esto constituye un avance, sino que para que esto se dé, se deben tener en cuenta los objetivos de la innovación tecnológica, el proceso de toma de decisiones, las personas que participan en el diseño y los riesgos que estos elementos podrían producir.

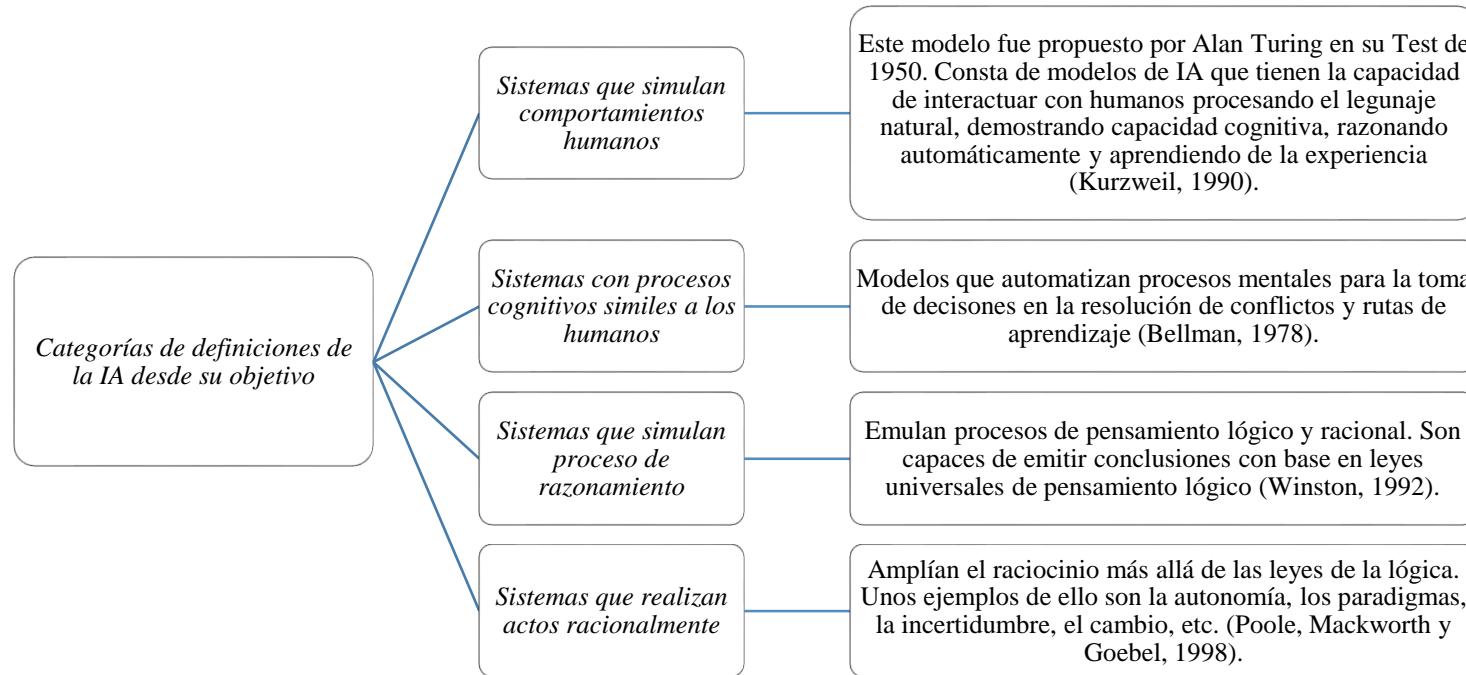
De esta manera, desde la perspectiva ética, la European Group on Ethics in Science and New Technologies (EGE) (2018), se ha planteado las siguientes cuestiones claves: el control humano, la autonomía artificial, la interacción emocional de seres humanos y productos de la IA, la responsabilidad, el rediseño institucional, la intersección entre el impulso y la manipulación, la transparencia, los límites a los sistemas de puntuación social, el perfilado humano sin consentimiento, la vigilancia masiva. A futuro, preocupan la existencia de conciencias artificiales con autodeterminación, los efectos de la responsabilidad contractual y extracontractual y los derechos a los no humanos como el reconocimiento de atributos de su personalidad.

Por lo tanto, ya son varios los organismos que han tratado la ética de la IA, como es el Instituto de Ingenieros Eléctricos y Electrónicos (IEEE) (2016), quienes construyeron la primera versión del documento de políticas del IEEE, que abordaba la ética en el diseño, para el cual fueron consultados más de 250 líderes de opinión; la International Telecommunication Union (ITU), que desde 2017 organiza cumbres internacionales llamadas “All for

Good”, en las cuales se producen documentos relevantes; la Association for Computing Machinery (ACM), que desde el mismo año realiza conferencias sobre IA, ética y sociedad, en la que nació la Alianza sobre a IA, que para el 2018 presentó siete principios y reunió más de ochenta entidades públicas y privadas; la Future of Life Institute, que ha llevado a cabo, desde 2015, movilizaciones para evitar el uso excesivo de la IA en la producción de armas.

Al abordar la definición de la IA, diferentes autores han tratado de acercarse a esta tesis, direccionada desde el objetivo de ella, razón por la que Russell y Norving (2016) clasifican las conceptualizaciones existentes en cuatro categorías y que se describirán en la figura 1.

Figura 1. Categorías de definiciones de la IA desde su objetivo

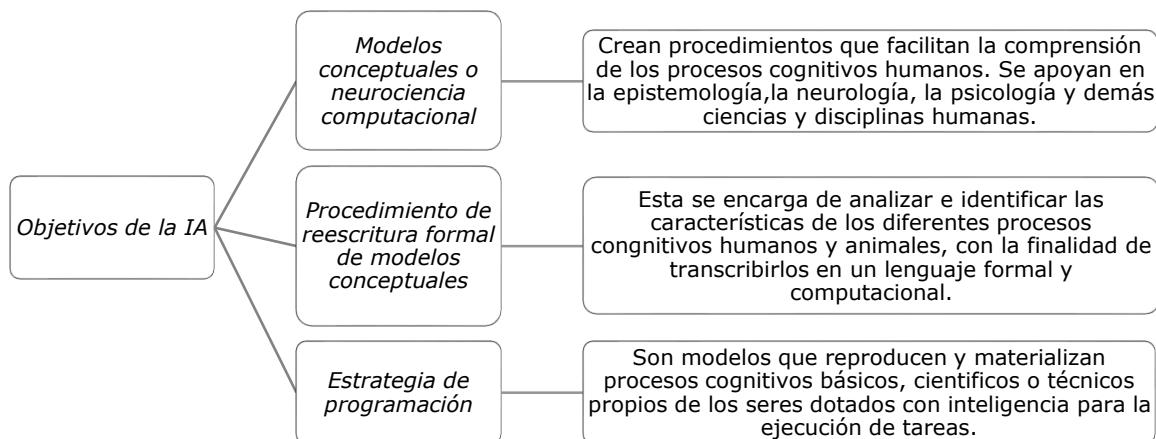


Fuente: elaboración propia, datos tomados de Russell y Norving (2016).

Del recorrido por las diferentes definiciones de la IA, se puede deducir que el principal objetivo de esta es la réplica de la inteligencia humana en sistemas informáticos y robots. Sin embargo, estos se diferencian en las facultades cognitivas que esta tecnología busca copiar y programar en sus productos.

En consonancia con las categorías mencionadas en la figura 1, también existen diversos objetivos a desarrollar por medio de la IA. Mira et ál. (1995), los clasifican como se observa en la figura 2.

Figura 2. Objetivos de la IA



Fuente: elaboración propia datos tomados de (Mira et ál., 1995).

De acuerdo con los objetivos mostrados en la figura 2, la IA tiene dos perspectivas, la primera como ciencia y la segunda como ingeniería del conocimiento Palma y Marín (2008); desde la primera perspectiva, se pretende una teoría computable del conocimiento humano y se encarga del objetivo expuesto en la figura 2: “desarrollar modelos conceptuales”, y en

pro de este, se analiza la estructura del conocer humano, estudiando los procesos y mecanismos a nivel neuronal y subcelular que soportan las funciones de percepción, memoria, lenguaje, decisión, emoción y acción. Para lo anterior, se dividen los procesos cognitivos en subtareas hasta alcanzar el nivel de razonamientos que no se descomponen, posteriormente, razón por la que pueden ser fácilmente expresados en un lenguaje formal.

Por otra parte, la perspectiva del conocimiento posee como fin la reescritura de las inferencias que resulten del análisis de los procesos cognitivos para, de esta manera, programar los operadores y los diversos sistemas. Se puede concluir, de esta primera parte, que el principal objetivo de la IA es replicar la inteligencia humana en máquinas y sistemas informáticos, y que la controversia se da en la posibilidad concreta de que eso se materialice y hasta qué punto se pueda dar.

Ahora bien, en este debate, ¿cuál es la relevancia de la ética en el desarrollo de la IA?; si se tiene en cuenta que la ética se extiende a todos aquellos seres dotados de autonomía e inteligencia, entendidos como aquellos que tienen la capacidad de tomar decisiones y actuar de forma racional, la *autonomía*, es entonces, un término utilizado para referirse a la capacidad de los seres racionales de escoger un curso de acción libremente, por lo que se puede asegurar, que es un rasgo distintivo de los seres humanos (EGE, 2018).

Es en este concepto de *autonomía* en el que se fundamenta la confusión que da paso a los diversos conflictos éticos que hay alrededor de la IA, ya que se afirma que muchas de las aplicaciones y productos de la IA tienen la capacidad de operar de forma autónoma (Quiroz, 2014). Sin embargo, en la actualidad, ninguno de ellos es capaz de dar cuenta de sus propios actos o de tomar decisiones con todos los elementos que los seres humanos lo

hacen, por lo que autores, como Marín (2019), afirman que es un error calificarlos *autónomos* desde la perspectiva ética del concepto y, además, que el uso de dicho término, para referirse simplemente a aplicaciones capaces de operar sin supervisión humana, pero, desde el punto de vista del análisis ético, el término adecuado para referirse a estos dispositivos sería *automático* y no *autónomo*.

De acuerdo con lo anterior, se genera otro interrogante, ¿solo los seres humanos son capaces de actuar de forma autónoma?, pregunta de suma relevancia, ya que los postulados éticos se diferencian dependiendo de quién sea el responsable de los razonamientos y las decisiones, y en este caso, el análisis se dirige a los productos generados mediante la IA como sujeto. En razón a este cuestionamiento, autores, como López de Mantaras (2015), hablan de una IA débil y una IA fuerte; la primera es aquella con la que se diseñan y programan ordenadores capaces de realizar tareas de forma inteligente, mientras que la segunda es aquella que permitiría replicar en máquinas la inteligencia humana, siendo esta última la que posibilitaría que estos dispositivos actuasen de forma autónoma.

En razón a la relevancia ética de la IA, desde el cuestionamiento que da la confrontación “autonomía vs. automático”; es que se da la importancia jurídica de la responsabilidad ética de los productos de la IA, así como los principales retos éticos que implica el desarrollo de la IA, y como tal este es el eje fundamental que justifica el planteamiento de esta investigación y la generación de conocimiento en esta temática, que le dará a los diferentes lectores, la información suficiente para establecer una posición ética desde la construcción de normativas hasta el establecimiento de una postura personal.

Método

Tipo de estudio

Se diseñó un estudio de tipo cualitativo (Hernández et ál., 2010), especialmente útil cuando se necesita desarrollar un instrumento de recolección que permita recoger la percepción de los participantes. Por lo que dio como resultado un estudio cualitativo exploratorio, donde se utilizó como instrumento una entrevista semiestructurada que se aplicó a la muestra (Ato, López y Benavente, 2013).

Participantes

Se obtuvo una muestra de diez expertos, escogidos por medio de un muestreo no probabilístico e intencional.

Técnicas e instrumentos

En la recolección de datos, se desarrollaron entrevistas semiestructuradas a expertos, siguiendo una guía como instrumento en el que se construyeron preguntas relacionadas con las categorías que orientaron la investigación. Dicha guía fue evaluada por jueces en aspectos como pertinencia, redacción y coherencia.

Procedimiento

Las entrevistas se desarrollaron en el lugar de domicilio de los participantes, donde se solicitó inicialmente la firma de consentimiento informado y se abordaron posteriormente las preguntas de la guía de entrevista. Las entrevistas fueron grabadas en audio para su posterior transcripción, y a partir de las transcripciones se realizó el proceso de codificación abierta, axial y selectiva en el programa Atlas. Ti 8.

Consideraciones éticas

La investigación adoptó todos los principios legales y éticos formulados por la Resolución 8430 de 1993 y la Ley 1090 de 2006; muestra de esto fue que se les informó a los entrevistados sobre los objetivos de la investigación, la destinación de los resultados obtenidos, los riesgos de su participación en la investigación y, finalmente, se verificó, de forma detallada, el cumplimiento de las normas éticas en la realización del estudio.

Resultados

Al realizar el análisis sociodemográfico, se pudo observar que el 100 % de los participantes fueron hombres, el rango de edad más frecuente está entre los 30 y 35 años, siendo la profesión más común la ingeniería electrónica y las áreas de experiencia con mayor incidencia *big data*, automatización, modelos de predicción y biomecánica, inteligencia artificial, modelos de predicción, biomecánica y telemática, de igual manera, la empresa que más utiliza este tipo de expertos es B-Yond, con el 30 %, como se puede observar en la tabla 1.

Tabla 1. Datos sociodemográficos

| Datos sociodemográficos | |
|---------------------------|-----|
| Población de expertos: 10 | |
| Variable | % |
| Sexo | |
| Masculino | 100 |
| Edad | |
| 25-30 años | 10 |
| 30-35 años | 65 |
| 35-40 años | 25 |
| Profesión | |

| Datos sociodemográficos | |
|--|----|
| Población de expertos: 10 | |
| Variable | % |
| Ingeniería Electrónica | 50 |
| Ingeniería de Sistemas | 30 |
| Mecatrónica | 10 |
| Ingeniería de Datos | 10 |
| <i>Área de especialidad de expertos o campo profesional</i> | |
| <i>Big data, automatización, modelos de predicción y biomecánica</i> | 30 |
| Automatización, telemática y desarrollo de telecomunicaciones | 20 |
| Inteligencia artificial, modelos de predicción, Biomecánica y telemática | 30 |
| Desarrollo de telecomunicaciones y programación | 10 |
| Programación | 10 |
| <i>Empresa donde desempeñan sus especialidades</i> | |
| B-yond | 30 |
| Oracle | 20 |
| Huawei | 20 |
| Microsoft | 10 |
| IBM | 10 |
| Independiente | 10 |

Fuente: elaboración propia.

En lo que respecta a la aplicación de los principios éticos a la IA, los expertos en su mayoría reafirman la concepción de que esta debe girar alrededor de aquellas situaciones habituales que se pueden presentar entre los robots y las máquinas inteligentes y su interacción con el ser humano, la sociedad y el medio ambiente, como lo expresa E1 (2021), al referirse al *biohacking*, “la aplicación de la ciencia en introducir la tecnología al cuerpo humano, con fines de diagnóstico, curar o mejorar los atributos humanos, siendo uno de los aspectos, en los que se debe regular desde la ética la IA”(p.2).

En cuanto al interrogante, ¿hasta qué punto los humanos entrarían a depender de las máquinas desarrolladas bajo la tecnología de IA?, la mayoría de los participantes expresó que las aplicaciones derivadas de la IA aún son muy limitadas a nivel intelectual, ya que todavía no se han desarrollado en ellos procesos de aprendizaje automático, sino que requiere de un entrenamiento previo para el desarrollo de cada una de las tareas diseñadas a ejecutar, lo cual está reafirmado por E5 (2021) en su expresión:

el principal miedo o conflicto de los humanos con la IA es aquel que ha sido inculcado por las películas de ciencia ficción, ya que los muestran como máquinas capaces de tomar decisiones en contra de los seres humanos, lo cual no es cierto, ya que el aprendizaje automático no es una de las fortalezas de la IA. (p.4)

Por otra parte, otro de los grandes temores de la sociedad, es el hecho de la sustitución de las personas por máquinas en el sector productivo (Castrillón, Rodríguez y Leyton, 2008), lo que, de acuerdo con los resultados de las entrevistas, se debe a los grandes márgenes de reducción de costos en materia salarial y de seguridad social al lograr que las máquinas producto de la IA puedan realizar, de forma rápida y eficiente, las labores del humano, temor que fue ratificado por E10 (2021) al afirmar que “el gran temor de las personas es que la IA genere desempleo y un temor constante a la sustitución en la población trabajadora más vulnerable”.

También se extrajo de las entrevistas, el hecho del temor a los procesos de transformación en la sociedad como algo natural en el ser humano, sino que también cuestionan el tratamiento responsable que se le dé a las diferentes aplicaciones fruto del desarrollo de la IA, como se puede observar en la percepción de E3 (2021): “Considero que siempre una IA va

a poder llegar hasta donde su creador haya puesto los límites, entonces desde el principio se debe de acotar la ‘libertad’ que tenga la IA”.

Otro de los aspectos que emerge de las entrevistas es la responsabilidad por las decisiones, actos y omisiones de los robots, ya que definitivamente, se percibe una falta de seguridad jurídica enmarcada en el cuestionamiento sobre si la máquina producto de la IA puede ser parcial o totalmente responsable de sus actos y si pueden considerarse personas físicas, personas jurídicas, animales u objetos; situación que genera verdadera confusión desde el punto de vista de los expertos, lo que se puede observar desde la respuesta de E8 (2021): “es importante fijar parámetros de actuación en lo que respecta a límite, condiciones y responsabilidades a la conducta de los productos de la IA”.

Además, está la responsabilidad de los programadores de los productos derivados de la IA, en lo que los entrevistados fueron enfáticos, como lo muestra E6 (2021), al aseverar que “los diseñadores deben conservar principios éticos antes, en el transcurso y posterior al proceso de concepción, desarrollo y aplicación de la IA”. Esto mostraría una serie de axiomas éticos que el gremio utiliza, y la principal regla es no perjudicar, no herir, no incurrir en engaños que perjudiquen a los usuarios y a las personas que interactúan con estos productos.

Por otra parte, a través de las entrevistas, se indagó sobre los escenarios científicos que han tenido como objetivo el cuestionamiento de la relación de la IA con los seres humanos, como es el transhumanismo, definido por Russell y Cohn (2012) como aquel que “cuestiona la posibilidad y conveniencia en la transformación de la condición humana, creada y desarrollada mediante tecnologías que disminuyan el envejecimiento de la persona y aumenten su capacidad intelectual, física, y psicológica” (p. 58). Esta definición se encuentra totalmente aceptada entre los expertos, ya

que la mayoría estuvo de acuerdo con la explicación dada por E4 (2021), cuando se le preguntó por su conocimiento de este escenario: “Es el pensamiento de mejorar la raza humana a través de las bondades de la tecnología, ya sea mejorando física o intelectualmente las capacidades del hombre”.

A través de las entrevistas, emergieron categorías de investigación como la relación de la IA con la bioética, desde la perspectiva del aumento o el incremento de tiempo de las capacidades humanas, lo cual ha causado gran polémica, en razón a las implicaciones que trae el uso de la IA para manipularlas, a lo que también se presentó una tendencia mayoritaria de los participantes al estar de acuerdo con los principios transhumanistas, como la afirmación de E7 (2021), que identifica, entre las funciones de la IA, el hecho de “mejorar la vida del ser humano al integrarse plenamente a la tecnología, no solo en capacidades lógicas, sino físicas y psicológicas”.

Entre las categorías planteadas, desde el inicio de esta investigación, está la autonomía, la cual se percibió desde el punto de vista de los expertos participantes de esta investigación como la capacidad que tiene el ser humano para legislarse a sí mismo (Vargas, 2018), así como para diseñar y elegir los principios, axiomas, normas y leyes que deben cumplir para garantizar una buena convivencia, las cuales relacionaron con la dignidad humana y la actividad netamente humana, rechazando que los productos de la IA puedan ser autónomos, como se observa en la afirmación de E2 (2021): “la autonomía es un aspecto importante y único del ser humano”.

Discusión

Por medio del análisis expuesto en el apartado de resultados y de acuerdo con el propósito de esta investigación, que es el análisis de la ética vinculada a la IA, abordado desde la perspectiva de la legislación

internacional en lo que respecta a tratados, acuerdos y convenios promulgados por la Unión Europea (UE), se confrontó lo encontrado con la literatura y los preceptos surgidos desde la experiencia de la UE.

Se abre este análisis, aclarando por qué la UE se encuentra a la vanguardia de la ética de la IA, lo que, según Cotino (2019), se debe, por una parte, al alto nivel de desarrollo tecnológico de esta comunidad y, por otra, en razón a los altos estándares de derecho fundamentales establecidos en los países que la integran; razones que han hecho que la UE ofrezca al mundo los principios de una ética específicamente aplicada a la IA.

Definitivamente, a través de la revisión documental y las entrevistas a expertos, se puede extraer que el punto de partida y axiomas principales de la ética aplicada a la IA, son por una parte los derechos fundamentales y la dignidad, desde esa perspectiva se pueden extraer axiomas éticos que, según Cotino (2019), “parten de la dignidad como presupuesto de los derechos fundamentales y los derechos sean derechos subjetivos sean principios objetivos del ordenamiento jurídico” (p. 36).

Este fue uno de los factores que se corroboró en esta investigación, ya que los expertos participantes enfatizaron en los diferentes miedos que tienen los seres humanos a que sus derechos fundamentales sean vulnerados al predominar productos derivados de la IA en sus contextos, pues verían vulneradas las garantías principalmente relacionadas con el derecho al trabajo, al mínimo vital, a la satisfacción de necesidades básicas, a escoger libremente profesión u oficio, a la privacidad, entre otros, generando con ello, riesgos a la dignidad humana.

Se deduce entonces que la afirmación del Parlamento Europeo (2017), de que la ética aplicada a la IA debe partir de la dignidad humana, infiere además que sus principios deben tener como base los derechos fundamentales contenidos en la Carta de los Derechos Fundamentales, lo

que está reforzado por la solicitud que la Comisión Europea (2018), quien estructuró un grupo de expertos, que especificaron puntualmente que este documento debe ser tenido en cuenta para lograr el desarrollo de productos derivados de la IA, confiables y desarrollados desde una perspectiva ética.

Desde la perspectiva del derecho a la dignidad humana, la UE ha establecido cuatro principios básicos que en la literatura han sido considerados universales y que a través de la investigación fueron corroborados por los expertos, el primero es el de justicia, el cual se acerca al principio de la bioética con una postura de beneficencia y no maleficencia al humano, este fue identificado en esta investigación como aquel que exige que la IA sea desarrollada para el bien común y en beneficio de toda la humanidad, generándole prosperidad, valor agregado y aportando a la sostenibilidad del mundo; por lo tanto, su interacción con la sociedad debe contribuir a la justicia, la paz y la autonomía de los ciudadanos, así como al logro de la equidad y la paz en el ámbito social, económico y político.

Lo anterior está apoyado en los objetivos propuestos por el Grupo de Expertos de Alto Nivel en Inteligencia Artificial (AI-HLEG) (2018), que son la protección al proceso democrático y el Estado de Derecho, el suministro de bienes y servicios a bajos costos y altos estándares de calidad, mayor representatividad de los datos, mitigación de daños y mayor confianza de los usuarios, así como la contribución al logro de los objetivos de desarrollo sostenible de la ONU.

Todo lo anterior es coherente con lo planteado en la bibliografía y en los análisis de los expertos entrevistados, ya que coinciden en que la IA y sus productos son diseñados sin causar daños físicos, psicológicos, financieros o sociales, lo que implica que su aplicación debe hacerse bajo los mismos

parámetros de prevención de perjuicios evitando el excesivo o indebido uso de los productos de la IA; situación que, según Cotino (2019), inicia con los desarrolladores y se hace extensiva a los usuarios y los productos de la IA.

Esta perspectiva del principio de justicia, de acuerdo con la AI-HLEG (2018), apunta a evitar el uso de la IA en la carrera armamentista, protegiendo la democracia de cada país, evitando la polarización ideológica y el determinismo algorítmico, permitiendo la inclusión en el proceso de diseño de los productos de la IA, evitando daños sobre el medio ambiente y teniendo en cuenta el consumo de quienes requieren cómputos masivos, almacenamiento de datos y los minerales necesarios para las baterías.

El segundo principio que se encontró, según la opinión de los expertos, está relacionado con la experiencia que al respecto tiene la UE sobre la autonomía o acción humana, que implica que, a partir de la IA, el ser humano cede voluntariamente parte de su poder de decisión a las máquinas, sin dejar de promover la independencia humana y generando restricciones a la autonomía de los productos de la IA, así como procurar que esta sea reversible, lo cual es respaldado por Floridi et ál. (2018), al afirmar categóricamente que “Los humanos siempre deben conservar el poder de decidir qué decisiones tomar” (pp. 697-698).

Por su parte, el tercer principio, conocido como la supervisión humana o gobernanza de la autonomía, propuesto por AI-HLEG (2018), se deriva del anterior, al involucrar una relación directamente proporcional entre la autonomía y el impacto social de la IA, así como también que, al aumentar el control humano, existirán mayores registros que viabilicen la auditoría de los productos de la IA y, a la vez, la posibilidad de que el ser humano

tenga la posibilidad de desviar las decisiones tomadas por máquinas inteligentes.

Con respecto a lo anterior, el EGE (2018) ha establecido una posición muy clara cuando afirma que la autonomía se vincula al derecho de ser libre, por lo tanto, es una característica específica y única del ser humano, lo que quiere decir que a los productos de la IA no se les puede atribuir autonomía por más inteligentes que estos sean, ya que la autonomía está supeditada al principio de libertad del ser humano. Esto, según Floridi et ál. (2018), implica que el desarrollo de la IA es una responsabilidad humana, por lo que se necesita siempre del acompañamiento de control humano en el diseño y desarrollo de los productos de la IA, así como la facilidad para prevenir diversas acciones futuras de estos.

Floridi et ál. (2018) es radical en afirmar que el mejoramiento social que pueda traer el desarrollo de la IA, no puede sacrificar el control humano ni el límite a la prevención de daños; sin embargo, Thaler y Sunstein (2009), alertan sobre el riesgo de que tecnologías como la IA dañen la autodeterminación humana a cambio de rutinas que faciliten el trabajo y la vida del ser humano.

El último principio propuesto por la UE es el de la explicabilidad y transparencia que, según Cotino (2019), implica la necesidad de entender y pedir cuentas, lo que implica directamente identificar de quién es el funcionamiento de la IA y la responsabilidad que surja de ese funcionamiento, lo que supone proponer procesos transparentes para que el humano pueda entender cómo funcionan los productos desarrollados por la IA, sobre todo en los casos en los que la autonomía sale a relucir para así poder determinar la justicia, la responsabilidad, la explicabilidad, la auditabilidad y la transparencia, aspectos que pueden variar desde la perspectiva en la que son analizados, por lo que el círculo de personas que

puedan acceder al conocimiento del algoritmo sean también restringido por muchos motivos, como la seguridad o propiedad intelectual (AI-HLEG, 2018).

Este último principio ha generado muchas recomendaciones de autores, como Floridi et ál., (2018), que exponen involucrar al sector productivo y a los profesionales; desarrollar normativas apropiadas, mejorando la infraestructura de los sistemas de justicia que le permita a los tribunales discutir las decisiones algorítmicas; el desarrollo de mecanismos de auditoria, para identificar consecuencias no deseadas, posibilitar la trazabilidad y la auditabilidad, entre otros.

De estos principios, se puede concluir que, a partir de la dignidad y la propia libertad del ser humano, relacionada directamente con la autonomía, se configuran como principios éticos de la IA los derechos fundamentales a la no discriminación, a evitar el riesgo y a mantener la privacidad. A la vez, se reafirman principios como el de responsabilidad del Estado de derecho, rendición de cuentas y no estigmatización; además de reforzar el concepto de regulación de la IA, el cual debe ser un proceso construido desde un debate democrático y de compromiso público, que implica transformaciones y actualizaciones en la educación de cada país que aseguren el conocimiento y los derechos de las futuras generaciones, así como la sostenibilidad del medio ambiente.

Las diversas propuestas mencionadas están apoyadas en los lineamientos del Parlamento Europeo (2017), que propone la creación de un Código de Comité de Ética que incluya los siguientes axiomas: evitar conflictos de intereses, acreditar la competencia y la experiencia de los desarrolladores, la independencia, transparencia y rendición de cuentas de los modelos de IA y sus productos y tener un carácter multidisciplinario que permita el

acompañamiento de mínimo un miembro con formación filosófica, ética o jurídica en el diseño de estos.

La segunda conclusión a la que se llega, refiriéndose específicamente a la perspectiva de los expertos colombianos, es que, a diferencia de la UE, en Colombia no se ha tenido en cuenta el hecho de reforzar los principios éticos desde la educación de las futuras generaciones, que en realidad son quienes llevarán la mayor parte de las interacciones con los productos desarrollados por la IA. A esta situación se debe prestar especial atención, ya que se está proponiendo siempre un desarrollo de esta tecnología de la mano de la supervisión del ser humano, por lo que los lineamientos educativos deben ser no solo a partir del conocimiento tecnológico, sino también desde el desarrollo de la ética y la moral.

Desde la perspectiva jurídica, en Colombia, tanto el derecho como la ética han dejado de lado el desarrollo de las nuevas tecnologías, como es el caso de la IA, a diferencia de la UE, que se ha concentrado en el interés social, institucional y doctrinal, desde el punto de vista de la prevención frente a los efectos, impactos y desafíos a los que llevan los productos de la IA.

Hoy es válido decir que la existencia de estas tecnologías han traspasado la ciencia ficción de las pantallas y han cautivado la atención de diversos especialistas por lo complejo de las implicaciones de sus efectos, generando así, la necesidad relevante de construir principios éticos ante los diversos riesgos, no solo de convivencia, sino jurídicos, en razón a los diversos vacíos que existen en muchos países, como es el caso de Colombia, territorio en el que no se cuentan con bases normativas que sirvan como fuente para la solución de los conflictos derivados del uso de los productos desarrollados desde la IA.

Aún se está a tiempo para diseñar propuestas regulatorias concretas en cada país, así como la construcción de una ética aplicada exclusivamente a

la IA, que sea inclusiva y alimentada por los principios jurídicos del sistema político de cada país, con una pretensión del cumplimiento normativo en el diseño y en las afectaciones futuras de los productos desarrollados por la IA.

Para el desarrollo de propuestas regulatorias en cada país, se debe tener en cuenta la evolución de la ética en la IA, pues es la única forma de establecer políticas y normas eficaces y eficientes en materia de cumplimiento de las fórmulas propuestas; desarrollando escenarios de prevención y cumplimiento normativo, propiciando una adecuada respuesta del derecho, ya sea por la vía regular o por nuevas vías que permitan la elaboración de parámetros contextualizados desde los axiomas éticos aplicados a la IA.

Es así que para la construcción de normas que regulen el funcionamiento de la IA y sus productos es menester que se defina si estos son o no sujetos, si han de contar con atributos de la personalidad o no, si están en nivel superior, igual o inferior con el humano, cuáles serán los límites de su funcionamiento, si serán catalogados como seres con cognición y razonamiento, si se les atribuirá autonomía o serán automatizados, si gozarán del derecho a la libertad y garantías constitucionales, si son sujetos de responsabilidad civil, disciplinaria, penal, entre otras, y cuál será el alcance de los efectos de sus actos.

De igual forma, la definición de los cuerpos normativos que regulen la interacción de la IA y sus productos con el ser humano deben tener como fundamento y fuente, el conjunto de acuerdos, tratados y convenios que, en materia de derechos fundamentales y humanos, se han ratificado por el ejecutivo del territorio nacional.

Finalmente, el transhumanismo es un tema en el que requiere que se adelanten debates contundentes por parte de los órganos de las ramas del

poder público, pues la implementación de esta, en la práctica ideológica, podría llevar a borrar la línea que diferencia a los humanos de los objetos o cosas no sintientes.

Conclusiones

Partiendo de la discusión planteada, se hace necesario, por parte de los Estados, definir códigos de ética de la IA que prevengan la mala praxis, el monopolio de las tecnologías en materia comercial y laboral, que prohíban su incursión en el uso armamentista, que limiten su alcance en el uso de datos personales, que prevenga cláusulas abusivas en los contratos electrónicos y que eviten el abuso de la posición dominante en las interacciones que con ellas se tengan.

De igual manera, es imperativo que se deje claro que la autonomía es una figura derivada de la libertad, la cual es exclusiva del ser humano, toda vez que es el único sujeto que goza del ejercicio de la autodeterminación, gracias a su capacidad de raciocinio, mientras que la IA, si bien cuenta con automatización, es decir, tiene capacidad cognitiva, carece de conciencia o razón.

Ahora bien, partiendo del hecho de que se hace menester la definición de un código de ética de la IA, este debe constituirse bajo el precepto de la protección de la dignidad humana, los derechos fundamentales y los principios de justicia, paz, autonomía exclusiva en el ser humano como especie libre, la supervisión del comportamiento de la IA y la transparencia y aplicabilidad de sus actos.

Por lo tanto, los axiomas a seguir deben tener como pilar la mitigación de conflictos en todo evento, la acreditación de la idoneidad de los desarrolladores, la transparencia y honestidad del funcionamiento de la IA y sus productos, además de darle un contexto interdisciplinar a sus

modelos de creación y que abarquen aspectos morales, éticos, sociológicos, filosóficos, psicológicos y legales (Aristizábal, 2019).

Es así que para poder construir las propuestas antes mencionadas, desde la academia, se debe iniciar una transformación hacia el profesional digital, que cuente con una serie de habilidades no tradicionales, como el alfabetismo digital, el conocimiento especializado de la industria, la inteligencia emocional, la mentalidad emprendedora, el conocimiento especializado de un área de la ciencia o disciplina con competencias interdisciplinares, habilidades de gestión y conocimiento estratégico de los mercados.

En general, los profesionales del derecho han sido entrenados para pensar, escribir y reducir riesgos, pero el diseño futuro es todo lo contrario, es hacer, construir y asumir riesgos. La innovación avanza por un camino de prueba y error. La experiencia del uso de los sistemas digitales en materia legal ha sido creada con la mentalidad de un abogado del siglo XIX, que debe enfrentarse a las necesidades de los diseñadores del siglo XXI. Por lo que los nuevos productos de servicios legales pueden hacer una enorme contribución para la inclusión en la justicia y la mitigación de ellos en futuros riesgos que vendrán con el reconocimiento de la IA como sujeto de actos jurídicos.

Referencias

Argandoña, A. (2019). Ética e inteligencia artificial (I). *IESE Blog Network: Economía, Ética y RSE*. <https://blog.iese.edu/antonioargandona/2019/03/25/etica-e-inteligencia-artificial-i/>

Aristizábal, J. F. (2019). La teoría pura del derecho y la exclusión de la sicología. *Revista IUSTA*, 1(50), 121-143. <https://doi.org/10.15332/1900-0448.2019.0050.05>

- Ato, M., López, J. y Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología*, 29(3), 1038-1059.
<https://dx.doi.org/10.6018/analesps.29.3.178511>
- Bonnefon, J., Sharif, A. y Rahwan, I. (2016). The social dilemma of autonomous vehicles. *Science*, 352(6293), 1573-1576. <https://doi.org/10.1126/science.aaf2654>
- Castrillón, O., Rodríguez, M. y Leyton, J. (2008). Ética e inteligencia artificial, ¿necesidad o urgencia?
<http://www.iiis.org/CDs2008/CD2008CSC/CiSCi2008/PapersPdf/C054TM.pdf>
- Cotino, L. (2019). Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y *big data* confiables y su utilidad desde el derecho. *Revista Catalana de Dret Públic*, (58), 29-48. <https://doi.org/10.2436/rdcp.i58.2019.3303>
- European Group on Ethics in Science and New Technologies (EGE) (2018). Statement on Artificial Intelligence, Robotics and “Autonomous” Systems. Comisión Europea, Dirección General de Investigación e Innovación.
https://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf
- Floridi, L. et ál (Luciano Floridi, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, Robert Madelin, Ugo Pagallo, Francesca Rossi, Burkhard Schafer, Peggy Valcke yEffy Vayena). (2018). AI4People —An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines*, 28(4), 689-707.
- Grupo de Expertos de Alto Nivel en Inteligencia Artificial (AI-HLEG) (2018). Draft ethics guidelines for trustworthy AI: Working document for stakeholders' consultation.
<https://digital-strategy.ec.europa.eu/en/library/draft-ethics-guidelines-trustworthy-ai>
- Hernández, R., Fernández, C. y Baptista, M. (2010). *Metodología de la investigación*. McGraw-Hill.
- Instituto de Ingenieros Eléctricos y Electrónicos (IEEE) (2016-2017). Ethically aligned design: A vision for prioritizing human well-being with autonomous and intelligent systems. <https://ethicsinaction.ieee.org/>

López de Mantaras, R. (2015). Algunas reflexiones sobre el presente y futuro de la inteligencia artificial. *Novática*, 234(4), 97-101.
<https://digital.csic.es/handle/10261/136978>.

Marín Palma, R. (2008) Inteligencia artificial. Técnicas, métodos y aplicaciones.
<https://www.iberlibro.com/Inteligencia-artificial-T%C3%A9cnicas-m%C3%A9todos-aplicaciones-Palma/21841884913/bd>

Martin, K. y Freeman, R. (2004). The Separation of Technology and Ethics in Business Ethics. *Journal of Business Ethics*, 53(4), 353-364.
<https://ssrn.com/abstract=1410846>

Mira, J., Delgado, A., Boticario, J. y Díez, F. (1995). *Aspectos básicos de la inteligencia artificial*. Sanz y Torres.

Palma Méndez, J. T. (2008). Inteligencia artificial: métodos, técnicas y aplicaciones. McGraw-Hill España, 524-535. <https://elibro.net/es/lc/poligran/titulos/50116>

Parlamento Europeo. Resolución del Parlamento Europeo. Recomendaciones destinadas a la Comisión sobre normas de derecho civil sobre robótica (2015/2103[INL]). Febrero 16 de 2017.

Quiroz, M (2014). Acercamiento a las “oposiciones paradigmáticas” entre neoconstitucionalismo y positivismo jurídico. *Revista IUSTA*, 41(2), 77-97.

Rusell, S. y Norvig, P. (2016). *Artificial Intelligence: A Modern Approach* (3.^a ed.). Pearson Education Limited.

Rusell, J. y Cohn, R. (2012). *YAML*. Bookvika Publishing.

Thaler, R. y Sunstein, C. (2009). *Un pequeño empujón. El impulso que necesitas para tomar mejores decisiones sobre salud, dinero y felicidad*. Taurus.

Vargas, A. M. (2018). Los límites y las consecuencias reales del discurso jurídico de la autonomía privada de la voluntad en el individuo. *Revista IUSTA*, 2(49), 91-114.
<https://doi.org/10.15332/1900-0448.2018.0049.04>