

Mauricio Alejandro Ascencio Moreno* (Colombia)

La inteligencia artificial y el poder público: una mirada a su utilización como herramienta de eficiencia estatal, a partir de su impacto en los derechos humanos

RESUMEN

La inteligencia artificial (IA) ha sido utilizada por varios Estados para tomar decisiones sin intervención humana, con el objetivo de lograr mayores índices de eficiencia en el cumplimiento de sus fines esenciales; sin embargo, el uso de dicha tecnología en el ejercicio del poder público ha generado consecuencias negativas, especialmente frente a la garantía y el respeto de los derechos a la privacidad, la intimidad, la dignidad, la libertad, la igualdad, el debido proceso, etc., debido a la producción de situaciones de orden netamente técnico, que amenazan dichos derechos. Tal es el caso de los sesgos algorítmicos que discriminan a grupos poblacionales determinados por razones de raza, religión, género o condiciones de pobreza, los cuales pueden producirse desde la programación de los algoritmos o desde el mismo momento de procesar los datos necesarios para tomar una decisión automática.

En el presente trabajo se realiza un análisis reflexivo mediante el uso de la metodología cualitativa y el estudio de fuentes bibliográficas relevantes, sobre la utilización de la IA para la toma de decisiones en el ejercicio del poder público en Colombia, partiendo de las ventajas y los riesgos propios de dicha tecnología disruptiva, los cuales han afectado principalmente la garantía y protección de los derechos humanos, tal como lo han evidenciado las experiencias documentadas en otros Estados, las cuales llaman a cuestionarnos si esta herramienta técnica es la idónea para tomar decisiones administrativas o judiciales y si es el momento para implementarlas.

* Abogado de la Universidad Santo Tomás de Bogotá, magíster en Consultoría de la Universidad de Salamanca, magíster en Derecho Contractual Público y Privado de la Universidad Santo Tomás, especialista en Derecho Administrativo de la Universidad del Rosario y doctorando en Derecho de la Universidad Externado de Colombia. Docente universitario en el área de Derecho Público. Miembro del Observatorio Sector Público e Inteligencia Artificial de la Universidad de Cádiz, <https://www.ospia.org/>. mauricio.ascencio@est.uexternado.edu.co. Orcid: <https://orcid.org/0000-0002-9650-2368>.

Palabras clave: inteligencia artificial; derechos humanos; poder público.

Artificial intelligence and public power: An examination of its use as a tool for state efficiency, from the perspective of its impact on human rights

ABSTRACT

Artificial intelligence (AI) has been used by several states to make decisions with no human intervention to attain higher efficiency rates in the achievement of their fundamental objectives. However, the use of this technology in the exercise of public authority has led to negative consequences, especially in terms of guaranteeing and respecting the rights to privacy, intimacy, dignity, freedom, equality, due process, etc., as it gives rise to situations of a purely technical nature that jeopardize these rights. Such is the case of algorithmic biases that discriminate against certain population groups on the grounds of race, religion, gender, or poverty situations. Biases can occur from the moment the algorithms are programmed or at the time the data required to make an automatic decision is processed.

This research aims to conduct a reflective analysis of the use of AI for decision-making in the exercise of public power in Colombia, using qualitative methodology and the study of relevant bibliographic sources. The article examines the advantages and risks of this disruptive technology, which have mainly affected the guarantee and protection of human rights, as evidenced by experiences documented in other states, which cast doubt on whether this technical tool is suitable for administrative or judicial decision-making and whether this is the right time to implement it.

Keywords: Artificial intelligence; human rights; public authority.

Künstliche Intelligenz und Staatsgewalt: Überlegungen zu ihrer Verwendung als Instrument staatlicher Effizienz am Beispiel ihrer Auswirkungen auf die Menschenrechte

ZUSAMMENFASSUNG

Mehrere Staaten setzen künstliche Intelligenz (KI) ein, um ohne menschliche Mitwirkung Entscheidungen mit der Absicht zu treffen, bessere Leistungsindizes bei der Erfüllung ihrer wesentlichen Aufgaben zu erzielen. Der Einsatz dieser Technologie bei der Ausübung der staatlichen Gewalt hat jedoch negative Folgen mit sich gebracht, insbesondere hinsichtlich der Gewährleistung und der Achtung der Persönlichkeitsrechte sowie der Privatsphäre, menschlichen Würde, Freiheit, Gleichheit, Rechtsweg usw., die durch die Schaffung von Situationen ausschließlich technischer Natur bedingt sind, durch die die Sicherstellung der genannten Rechte bedroht wird. Dies betrifft zum Beispiel die algorithmischen Verzerrungen, durch die bestimmte Bevölkerungsgruppen aufgrund ihrer Rasse, ihrer Religion, ihres Geschlechts oder ihrer Armut diskriminiert werden. Solche Verzerrungen können durch die Programmierung der Algorithmen oder ab dem Moment auftreten, in dem die notwendigen Daten für automatische Entscheidungen verarbeitet werden.

Die vorliegende Arbeit beabsichtigt eine reflexive Analyse unter Verwendung qualitativer Methoden und der Auswertung relevanter Literaturhinweise über den Einsatz von KI zur Entscheidungsfindung im Rahmen der Ausübung der Staatsgewalt in Kolumbien. Dabei wird von den Vorteilen und Risiken dieser destabilisierenden Technologie ausgegangen, die vor allem die Garantie und den Schutz der Menschenrechte betreffen, wie aus den dokumentierten Erfahrungen aus anderen Staaten hervorgeht, aufgrund derer wir uns fragen müssen, ob dieses technische Instrument dazu geeignet ist, Verwaltungs- oder Gerichtsentscheidungen zu treffen, und ob schon der Zeitpunkt gekommen ist, solche Entscheidungen umzusetzen.

Schlüsselwörter: Künstliche Intelligenz; Menschenrechte; Staatsgewalt.

Introducción

En la actualidad, la vida del ser humano está estrechamente relacionada con algoritmos y sistemas de inteligencia artificial (IA) que facilitan el análisis de las grandes cantidades de datos que cada individuo produce. La influencia de esta tecnología en la vida cotidiana ha escalado tanto, que decisiones simples sobre cómo relacionarse socialmente, qué ruta tomar para ir a un destino, qué películas o series ver, qué tipo de bienes y servicios comprar, a quiénes seguir en redes sociales o qué tipo de contenido ver en estas se encuentran condicionadas por las valoraciones hechas por algoritmos, propiedad de las grandes corporaciones tecnológicas, los cuales influyen en la elección final de las personas, a partir de las sugerencias y recomendaciones hechas por canales digitales, situación que es uno de los efectos producidos por la así llamada “cuarta revolución industrial”.

El concepto de industria 4.0 o cuarta revolución industrial¹ surgió a partir de los avances científicos y tecnológicos desarrollados a finales del siglo XX y lo que va corrido del siglo XXI, cuando la sociedad entró en la era cibernética y se logró combinar la digitalización, la conectividad, la automatización, la robotización y la IA en los procesos industriales.

Esta transformación tecnológica ha cambiado prácticamente todos los aspectos de la vida humana, desde las nuevas formas de producción, las interacciones sociales, la economía y la política, y, sobre todo, se ha servido de los procesos de globalización para desvanecer las fronteras y las distancias; de esta manera, a partir de las tecnologías de la comunicación, las relaciones humanas se convirtieron en transfronterizas.

Pero esta incidencia va mucho más allá, al pasar del mundo analógico a una realidad digital, pues se integraron los sistemas técnicos y computacionales en las cadenas industriales de producción, lo cual condujo al auge de innovaciones técnicas

¹ Silvia Barona, *Algoritmización del derecho y de la justicia. De la inteligencia artificial a la Smart Justice* (Valencia: Tirant lo Blanch, 2021), 58.

y científicas que incidieron en la digitalización de dichas formas de producción, con la aparición del internet de las cosas,² el *big data* y las demás tecnologías disruptivas.³

Lo anterior minimizó los tiempos de producción, comunicación y procesamiento de información, lo que acentuó el papel de la eficiencia como protagonista del desarrollo y la administración de recursos, hecho que no fue ajeno a las relaciones entre el Estado y los ciudadanos, pues los problemas de eficiencia y transparencia en la prestación de los servicios públicos empezaron a ser tratados con soluciones tecnológicas a fin de obtener resultados a corto plazo.

En el presente escrito se pretende hacer un análisis crítico sobre el uso de la IA para la toma de decisiones estatales, sea en la administración pública o de justicia, de cara al riesgo que dicha práctica implica para la garantía y el ejercicio de los derechos humanos, poniendo el acento en el derecho a la privacidad, la igualdad y la dignidad humana, e iniciando con la descripción de cómo se pretende resolver los problemas de legitimidad que sufre el sistema democrático colombiano mediante el uso de tecnologías disruptivas.

1. Problemas de legitimidad del sistema democrático y su legitimación *ex post* a través de la IA

Entre los rasgos característicos de nuestro actual Estado social de derecho se encuentra la primacía de un texto constitucional rígido, integrado por una parte dogmática y otra orgánica, las cuales parten de la separación de poderes, el reconocimiento de un catálogo de derechos, los mecanismos para hacerlos efectivos y un aparato estatal dedicado a garantizar el interés general, en el marco de la legalidad y la seguridad jurídica, principios democráticos que requieren de legitimidad al interior de la sociedad.

Desde un punto de vista sociológico, el sistema democrático goza de legitimidad en la medida en que los destinatarios de sus decisiones compartan la creencia social de que estas son legítimas.⁴

Siguiendo a Weber,⁵ este concepto de legitimidad, en las relaciones de dominación, está supeditado a la creencia de los ciudadanos en la legitimidad del Estado, y de dicha convicción social dependerá su grado de obediencia.

Por lo anterior, la legitimidad depende del tipo de relación de poder del que se trate, en función de lo cual el propio Weber estructuró tres tipos de dominación: i)

² “En su forma más simple, se puede describir como una relación entre las cosas (productos, servicios, lugares, etc.) y la gente, que resulta posible mediante tecnologías conectadas y plataformas varias”. Klaus Schwab, *La Cuarta Revolución Industrial* (Barcelona: Debate, 2018) 34.

³ Barona, *Algoritmización del derecho y de la justicia*, 59.

⁴ Vincenzo Ferrari, *Funciones del derecho* (Bogotá: Universidad Externado de Colombia, 2014), 291.

⁵ Max Weber, *Economía y sociedad* (México: Fondo de Cultura Económica, 2014), 327.

una dominación de carácter racional, que se centra en la legalidad del orden establecido y en el ejercicio del poder a partir del derecho; ii) una dominación tradicional, fundada en la creencia en la santidad de las tradiciones y en la legitimidad de las autoridades que ejercen el poder en razón de estas, y iii) una dominación carismática, que se funda en las condiciones de heroísmo o cualidades de un individuo particular y a las órdenes emanadas de este.⁶

En ese orden de ideas, con la expedición de la Constitución Política de Colombia se organizó el país como un Estado social de derecho, es decir que, en principio, la legitimidad de las autoridades parte del orden jurídico, el respeto por las garantías fundamentales y la creencia de los ciudadanos en dicha legitimidad, lo que constituye la legitimidad formal o positiva de la que habla Weber.

Sin embargo, el concepto de legitimidad expuesto por Weber ha sido ampliamente cuestionado, en cuanto estaría condicionarlo por una visión iuspositivista puramente formal. De acuerdo con una noción más cercana al pensamiento de Habermas, la legitimidad parte del consenso social obtenido gracias a los procesos de comunicación tendientes a concretar los contenidos materiales de justicia en un determinado grupo social.⁷

Siguiendo a Rúa,⁸ existe una legitimidad material, que se refiere al cumplimiento de los fines definidos por el modelo de Estado, es decir, “la satisfacción de los postulados contenidos en la parte dogmática de la Constitución por el aparato orgánico definido en el mismo texto superior”;⁹ para el caso colombiano, desde el texto superior, se creó una organización administrativa dotada de las competencias necesarias para cumplir los objetivos del Estado social de derecho.

Sin embargo y siguiendo a Ferrari,¹⁰ legitimar el sistema democrático (entendido como el sistema electoral y la acción del Estado) a partir del derecho implica justificar su existencia en razón a la adhesión de los ciudadanos al modelo normativo reconocible, es decir, un ordenamiento jurídico claro que identifique derechos, deberes y los mecanismos idóneos para hacerlos efectivos, lo cual significa que dicho proceso de legitimación debe realizarse *ex ante* y no *ex post*, puesto que una máxima del Estado de derecho es que las normas rigen hacia el futuro, como paradigma del principio de legalidad en el cual se sustenta.

Pero la creencia de la ciudadanía en la legitimidad estatal de la que habla Weber se ve disminuida cuando, a título enunciativo, esa barrera de irretroactividad de las normas se ve desconocida en el actuar estatal: así por ejemplo, en los eventos en que

⁶ Weber, *Economía y sociedad*, 374-375.

⁷ Rafael del Águila, “La política: el poder y la legitimidad”, en *Manual de ciencia política*, coord. por Rafael del Águila (Madrid: Trotta, 2000), 29-34.

⁸ Carlos Rúa, “La legitimidad en el ejercicio del poder político en el Estado social de derecho. Una revisión desde el caso colombiano”, *Ius et Praxis* 19, n.º 2 (2013), <http://dx.doi.org/10.4067/S0718-00122013000200004>.

⁹ Rúa, “La legitimidad en el ejercicio del poder político en el Estado social de derecho”, 97.

¹⁰ Ferrari, *Funciones del derecho*, 294.

los jueces, al momento de resolver casos particulares iniciados por *actores sociales comunes*, deben interpretar el ordenamiento jurídico, adaptándolo al caso concreto, y hacen incansables esfuerzos argumentativos para desconocer dicho principio en favor de intereses privados. Situación que también ocurre en el proceso legislativo, donde se expiden normas ambiguas de forma que se puedan modelar sus efectos a las necesidades del momento o en función de la satisfacción de intereses políticos.¹¹

Por otra parte, existe una acentuada y constante tensión entre el sistema capitalista y la democracia, lo que crea un riesgo para los mecanismos democráticos, tales como el modelo representativo y el interés general como un fin del Estado, los cuales se ven progresivamente disminuidos como producto del crecimiento de los intereses económicos.¹²

La globalización también ha contribuido a esta problemática, pues, con la creación de instituciones multilaterales y normas supranacionales, la vida de los ciudadanos y las políticas públicas han venido a ser, cada vez con mayor frecuencia, impuestas por organismos técnicos extranjeros carentes de legitimidad democrática, la cual reside exclusivamente en los órganos de elección popular, los cuales se han visto gravemente deteriorados debido a su desconexión con los electores, la corrupción y la omisión en el cumplimiento de su obligación de expedir normas encaminadas a hacer realidad los postulados del Estado social de derecho, dejando así en manos de los jueces la efectividad de los derechos constitucionalmente reconocidos.¹³

Tales prácticas, sumadas a fenómenos sociales como la corrupción, la captura del Estado por parte de grupos de interés, el clientelismo y demás prácticas similares, han minado la creencia de los ciudadanos en que las autoridades toman sus decisiones a partir de intereses altruistas, con la intención de lograr el bien común y en aras de materializar los fines esenciales del Estado, previstos en el artículo 2.º de la Constitución Política de Colombia.

Estos fenómenos sociales, que a nuestro modo de ver tienen una relación de causalidad con la pérdida de legitimidad del sistema democrático y del Estado, han sido estudiados desde la economía por la escuela de pensamiento conocida como *public choice* o elección pública. Según esta, los políticos solo obedecen a sus intereses particulares y sus decisiones son el producto de las presiones propias del mercado político que invaden todas las ramas del poder público,¹⁴ lo cual hace que los electores partan de la premisa de que su voto no tiene injerencia alguna en la elección de un determinado candidato y pierdan, por tanto, el interés en el proceso democrático.

¹¹ Ferrari, *Funciones del derecho*, 302-303.

¹² Floralba Padrón, "Presentación", a ¿El Estado constitucional en jaque? Tomo I. *Los retos del componente democrático* (Bogotá: Universidad Externado de Colombia, 2018), 17-25.

¹³ Padrón, "Presentación", 19-20.

¹⁴ Mario Pinzón, *Aproximaciones al Análisis Económico del Derecho* (Bogotá: Universidad Externado de Colombia, 2010), 110.

Para Lafont,¹⁵ actualmente los ciudadanos se sienten abandonados y no reconocen a las instituciones políticas como órganos de representación popular, hecho que no es otra cosa que la pérdida de legitimidad del sistema democrático y no corresponde a una situación desconocida para los gobernantes, quienes han utilizado múltiples estrategias a fin de lograr, entonces, una legitimación *ex post*.

En efecto, al resquebrajarse la creencia en la legitimidad estatal por parte de los ciudadanos, como consecuencia de fenómenos sociales como los mencionados, el Estado ha intentado superar tales desafíos legitimando el sistema *ex post*, entre otras cosas, a través de reacciones exacerbadas en el ejercicio del *ius puniendi*. Tal es el caso de la creación de múltiples organismos dirigidos a combatir la corrupción y mantener la imagen de control al interior de la administración pública, así como de la adopción de múltiples regímenes de responsabilidad de los servidores públicos y particulares que ejercen funciones o administran recursos públicos.

Otra de las herramientas utilizadas para legitimar el sistema *ex post* es la excesiva garantía del debido proceso, en virtud de la cual se generan actuaciones judiciales y administrativas en las que, con la bandera de la protección del debido proceso, se disponen instancias y oportunidades procesales tendientes a implantar en el imaginario colectivo la creencia de que existe pleno respeto de sus garantías fundamentales afectadas por los problemas de base del sistema democrático y del Estado en general.

Por otra parte, los gobiernos se han servido de los avances técnicos para facilitar ese proceso de legitimación del Estado, intentando dar soluciones tecnológicas a los problemas de base de la sociedad, con herramientas disruptivas como la IA y el *big data*, así como con estrategias de gobierno abierto y canales de transparencia a través de redes sociales, a fin de mejorar la percepción ciudadana sobre las instituciones públicas.

Respecto a la utilización de la ciencia y los avances tecnológicos como una forma de legitimar el sistema democrático, Habermas¹⁶ ha sostenido que la racionalización de la sociedad está supeditada a la instrumentalización de las nuevas tecnologías, pues en la medida en que la ciencia y la técnica avanzan, sus logros penetran en todos los ámbitos institucionales, cambiando las creencias de legitimación; cambio de creencias que también se produce con la secularización y el desencadenamiento de las cosmovisiones.

El hecho es que la aplicación de la razón técnica a un sistema social significa la institucionalización de un dominio político irreconocible,¹⁷ y que, en las relaciones de dominio, cuando se presentan situaciones que generan la pérdida de legitimidad del sistema democrático o del Estado mismo, las soluciones tecnológicas y la racionalidad de las acciones ciudadanas buscan alcanzar la legitimación *ex post*.

¹⁵ Cristina Lafont, *Democracia sin atajos* (Madrid: Trotta, 2021), 18.

¹⁶ Jürgen Habermas, *Ciencia y técnica como "ideología"* (Madrid: Tecnos, 2007), 54.

¹⁷ Habermas, *Ciencia y técnica como "ideología"*, 55.

En ese contexto y según informe de la Organización para la Cooperación y el Desarrollo Económicos (OCDE),¹⁸ Colombia es el principal referente en la región en materia de implementación de la IA, con la expedición del documento Conpes 3975 de 8 de noviembre de 2019,¹⁹ el cual recoge la política pública de transformación digital de los sectores público y privado a partir de la utilización de herramientas tecnológicas, en particular la IA. En el referido documento se resaltan los avances del país en materia de aplicación de esta tecnología en áreas como la administración de justicia, resaltándose a la Corte Constitucional y su sistema de digitalización denominado *PretoriA*, el cual lee y analiza datos de más de dos mil procesos de tutela en un máximo de treinta y seis minutos, siendo el primer tribunal supremo en utilizar la IA en el mundo.²⁰

Asimismo, la OCDE resaltó la aplicación de la IA en la digitalización de los procesos que se adelantan en la Superintendencia de Industria y Comercio, con el algoritmo *Justicia Digital*,²¹ el cual permite digitalizar en expedientes las audiencias adelantadas dentro de los procesos judiciales a cargo de dicha entidad, con la finalidad de buscar mayores índices de eficiencia en la toma de decisiones.

El informe continúa resaltando los casos de utilización de la IA en el sector público en Colombia, particularmente el caso del Departamento Nacional de Planeación, entidad responsable del Sistema de Identificación de Potenciales Beneficiarios de Programas Sociales (Sisben),²² que utiliza una tecnología de aprendizaje automático denominada Quantile Gradient Boosting, que hace perfiles automáticos de las personas a partir de un índice de prosperidad diseñado por el Estado, y determina los beneficiarios de programas sociales.

La OCDE finaliza su estudio del caso colombiano con la experiencia de la Agencia Nacional para la Defensa Jurídica del Estado,²³ la cual desarrolló, junto con el sector privado, un modelo matemático que utiliza *deep learning*, para estimar la probabilidad de éxito procesal de las acciones judiciales en contra del Estado colombiano y establecer los posibles rangos de las indemnizaciones, así como para determinar la posibilidad de conciliar en un caso particular, a partir del aprendizaje automático

¹⁸ Organización para la Cooperación y el Desarrollo Económicos (OCDE), *The strategic and responsible use of artificial intelligence in the public sector of Latin America and the Caribbean*, 22 de marzo de 2022, https://www.oecd-ilibrary.org/governance/the-strategic-and-responsible-use-of-artificial-intelligence-in-the-public-sector-of-latin-america-and-the-caribbean_1f334543-en.

¹⁹ Consejo Nacional de Política Económica y Social, República de Colombia, Departamento Nacional de Planeación, Documento Conpes 3975 de 8 de noviembre de 2019, “Política Nacional para la Transformación Digital e Inteligencia Artificial”, <https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/3975.pdf>.

²⁰ OCDE, *The strategic and responsible use of artificial intelligence*, 36-38.

²¹ OCDE, *The strategic and responsible use of artificial intelligence*, 36.

²² OCDE, *The strategic and responsible use of artificial intelligence*, 37.

²³ OCDE, *The strategic and responsible use of artificial intelligence*, 38.

generado como resultado de la valoración de indicadores financieros, datos sobre las condiciones actuales del país y teoría de juegos.

A la luz de los anteriores avances y reconocimientos, es preciso preguntarse: ¿por qué se considera todavía necesario realizar la transformación digital del Estado en Colombia? Una posible respuesta sería la que se expone en el Conpes 3975 de 8 de noviembre de 2019, donde se resalta que el objetivo de la implementación de tecnologías como la IA en la actividad estatal busca “contribuir en la construcción de un Estado más eficiente, transparente, participativo, y que preste mejores servicios a los ciudadanos, [para lo cual] el Gobierno nacional [ya] ha realizado esfuerzos orientados a la transformación digital en el sector público”.²⁴ No obstante lo anterior, la respuesta al señalado interrogante es mucho más compleja, y no se limita ciertamente al objetivo fijado en esta política pública.

Partiendo de la finalidad perseguida por Colombia con la implementación de la IA en el Estado, es claro el objetivo de mejorar factores de la administración en términos de eficiencia, transparencia, participación y la prestación de servicios a los ciudadanos. Las deficiencias en relación con estos aspectos obedecen, según Ramió,²⁵ al anquilosamiento en que se encuentra la administración pública, pues a pesar de invertir en herramientas ofimáticas que facilitan la modernización de las entidades públicas, sus estructuras organizacionales continúan siendo las mismas, hecho que impide que el Estado se adecúe a los retos de la cuarta revolución industrial.

Pero este fenómeno de “solucionismo” tecnológico no se limita a la administración pública, sino que se ha extendido a la Rama Judicial, donde en aras de la eficiencia en la administración de justicia se han utilizado herramientas basadas en IA. Así, por ejemplo, el juez primero laboral del Circuito de Cartagena, en Sentencia de 30 de enero de 2023,²⁶ invocando lo dispuesto en la Ley 2213 de 2022, resolvió una acción de tutela en segunda instancia basando su decisión en las respuestas dadas por el conocido ChatGPT,²⁷ de propiedad de OpenAI.

Uno de los aspectos destacables de esta sentencia es que el juez les dio a las respuestas emitidas por el Chat GPT la condición de fuente de derecho por encima de la Constitución, la ley, los reglamentos, la jurisprudencia, la doctrina, etc., y ello a pesar de que sus soluciones al caso²⁸ consistieron en respuestas simples sin

²⁴ Consejo Nacional de Política Económica y Social, Departamento Nacional de Planeación, Documento Conpes 3975 de 8 de noviembre de 2019, 11.

²⁵ Carles Ramió, *Inteligencia artificial y administración pública. Robots y humanos compartiendo el servicio público* (Madrid: Catarata, 2019), 6.

²⁶ Acción de tutela, Radicado 13001410500420220045901.

²⁷ <https://openai.com/product/chatgpt>. Véase Juzgado Primero del Circuito de Cartagena, Sentencia de 30 de enero de 2022, Radicado 13001410500420220045901.

²⁸ El objeto de la solicitud de amparo consistía en determinar si un menor con condición de autista y perteneciente a la población catalogada por el Sisben como pobreza moderada podría ser exonerado del pago de las cuotas moderadoras requeridas por el sistema de salud para usar los servicios requeridos por el accionante.

citación de otras fuentes. Así, más allá de nombrar algunas leyes y sentencias de la Corte Constitucional, la decisión judicial se soportó argumentativamente en las conclusiones de ChatGPT, simplemente por tratarse de un sistema basado en IA, y además, a pesar que dicho sistema no tiene conexión a internet y solo cuenta con entrenamiento e información hasta 2021.

Otro caso similar fue la realización de una audiencia inicial en el marco de un proceso contencioso administrativo, donde el Tribunal Administrativo del Magdalena realizó la diligencia en el *metaverso*. Si bien esta vez no se utilizaron herramientas basadas en IA, se usó la plataforma de Facebook para su celebración, identificando a los intervinientes a partir de su perfil en dicha red social y un mensaje de texto; con esto se buscaba mayor eficiencia en el trámite de los procesos judiciales,²⁹ pero enfocándose solo en que la iniciativa mediática funcionó técnicamente, sin hacer un análisis de ventajas reales para la administración de justicia.

La utilización de la tecnología, y especialmente de la IA, como herramienta de eficiencia ha generado una corriente de pensamiento que surgió a partir de los efectos sociales producidos por la segunda revolución industrial, entre finales del siglo XIX e inicios del XX, donde se hizo posible la producción en masa, gracias a las ventajas derivadas de la electricidad y la cadena de producción,³⁰ momento en el que la tecnología irrumpió en la economía y se empezaron a producir máquinas por máquinas.³¹

En el sector público, esta búsqueda de la eficiencia ha encontrado eco debido a situaciones que se han puesto en evidencia tanto en países con democracias débiles como en los que cuentan con democracias consolidadas, en donde los ciudadanos perciben los sistemas de gobierno como caóticos, contradictorios, lentos e ineficientes, y ello en un mundo cada vez más rápido, tecnológico y competitivo.

La respuesta estatal a esta problemática ha sido acudir a la tecnología para buscar la eficiencia y el mejoramiento de sus resultados. En el caso colombiano, en la última década, el Estado implementó estrategias tendientes a obtener una mayor productividad, eficiencia y bienestar social mediante la utilización de la tecnología en el sector público, sin que se pudiera lograr dicho mejoramiento en todos los sectores del país, razón por la cual se implantaron políticas como las de gobierno abierto, *big data* e IA.

A la luz de lo anterior, es claro que los Estados modernos están buscando solucionar sus problemas de ineficiencia, corrupción e inadecuada prestación de los servicios a su cargo con la utilización de la IA y las tecnologías disruptivas, por lo cual surge la pregunta de si las soluciones tecnológicas son realmente el medio idóneo para resolver problemas estructurales del sistema democrático colombiano.

²⁹ <https://www.youtube.com/live/LXi2TX9OBmQ?feature=share>.

³⁰ Marcos López, *Fundamentos para un derecho de la inteligencia artificial. ¿Queremos seguir siendo humanos?* (Valencia: Tirant lo Blanch, 2020), 28.

³¹ Günther Anders, *La obsolescencia del hombre*. Volumen II. *Sobre la destrucción de la vida en la época de la Tercera Revolución Industrial* (Valencia: Pre-textos, 2011), 21.

2. Inteligencia artificial, derechos humanos y sesgos algorítmicos

A partir de la utilización de los sistemas de *big data* e IA como herramientas para alcanzar la legitimación *ex post* del sistema democrático, es preciso abordar las consecuencias que su uso ha traído en el pasado cercano, haciendo especial énfasis en los sesgos algorítmicos y la afectación que estos causan a los derechos humanos.

Cuando nos encontramos en ambientes digitales es normal toparnos con términos como políticas de privacidad, recolección de datos, *cookies* necesarias, analíticas de rendimiento, sociales, etc., términos que solo se refieren a los datos, nuestros datos, su protección, divulgación y comercialización.

¿Por qué hablar de datos para referirse a IA y derechos humanos? La respuesta es simple: los datos son el combustible que hace funcionar a la IA, pues uno de los objetivos de esta tecnología es el procesamiento del *big data* o los macrodatos, los cuales pueden definirse como las tecnologías que permiten tratar cantidades masivas de datos obtenidos de diferentes fuentes, con la finalidad de darles una utilidad que genere valor y beneficios a partir de su procesamiento.³²

El procesamiento de datos que alimentan los algoritmos de IA ha generado grandes beneficios para la humanidad, como los sistemas de navegación terrestre o el fármaco capaz de acabar con las superbacterias farmacorresistentes descubierto por Regina Barzilay,³³ profesora de IA en el Massachusetts Institute of Technology (MIT). Esta, junto a su equipo, integrado por biólogos e ingenieros de datos, diseñó y entrenó un algoritmo con más de dos mil trescientos compuestos con propiedades antimicrobianas, a fin de identificar cuál de estos podría inhibir el crecimiento de una bacteria nociva para el ser humano. Cuando se anunció el descubrimiento, el *Financial Times* tituló: “La IA descubre antibiótico para tratar las enfermedades farmacorresistentes”.³⁴

Pero, más allá de ser una victoria para la IA, fue un triunfo para la inteligencia humana, pues el éxito del proyecto radicó en la forma de pensar el problema. En efecto, Barzilay no buscó rasgos moleculares que contrarrestaran las infecciones, lo que hubiera significado usar la misma estrategia de los antibióticos tradicionales a los cuales, precisamente, se resistían las bacterias, sino que enfocó su búsqueda en los rasgos que pudieran matar al microorganismo, cosa que un sistema basado en datos, patrones numéricos o aprendizaje automático no puede llegar a hacer. Por este motivo, es el momento de cuestionarnos si el término IA realmente se refiere o no a inteligencia, sumado al hecho de que tales logros médicos han formado en

³² Antonio Mozo Seoane, *Los límites de la tecnología. Marco ético y regulación jurídica* (Madrid: Reus, 2021), 117.

³³ Viktor Mayer-Schönberger, Kenneth Cukier y Francis de Véricourt, *Framers: la virtud humana en la era digital* (Madrid: Turner, 2021), 13.

³⁴ Mayer-Schönberger, Cukier y de Véricourt, *Framers*, 13.

el imaginario colectivo una creencia sobre la infalibilidad y objetividad de la tecnología, creencia que ha justificado su uso en el ejercicio del poder público, como se expondrá más adelante.

Al margen de esos eventos, que se consideran éxitos tecnológicos, es preciso llamar la atención respecto a nuestra relación con los dispositivos conectados a internet o transmisores de datos y el comercio que se ha creado alrededor de estos. Pues no solo hablamos de datos personales, como el número de identificación o la dirección de domicilio, nos referimos al historial médico, la información morfológica, las rutinas de vida, la capacidad económica o la ubicación constante en tiempo real, información que a través de sistemas de IA genera patrones y tendencias que permiten influir en las decisiones de las personas, sus hábitos de consumo o preferencias políticas, para citar algunos ejemplos, con o sin su consentimiento.

Lo anterior produce el primer riesgo al que haremos referencia en estas líneas: se trata de los derechos a la privacidad, al *habeas data* y a la protección de datos personales, reconocidos como fundamentales por el artículo 15 de la Constitución Política de Colombia.

En los términos de la referida norma, la privacidad es entendida como el espacio privado e íntimo de las personas respecto del cual no pueden tener injerencia o acceso las entidades estatales ni los particulares; por su parte, el *habeas data* es el derecho a conocer, actualizar y corregir la información personal que reposa en bases de datos públicas y privadas; mientras que el derecho a la protección de datos personales consiste en la garantía que tienen los ciudadanos de exigir, a los administradores de repositorios de información personal, el cumplimiento de las reglas y los principios referentes a la adecuada recolección, tratamiento, conservación y eliminación de la información a ellos referida.³⁵

Pero, no obstante la protección estatal de rango constitucional, es claro que en materia de *big data* las compañías tecnológicas recolectan datos en tiempo real, sea con el consentimiento de las personas, generalmente obtenido para la prestación real o prometida de bienes y servicios, o sin este, como en el caso del conocido ChatGPT antes aludido, pues pues al momento de la elaboración del presente escrito OpenAI no ha establecido una política de tratamiento de datos, ni se conoce el uso que le da a la información recaudada, pero es claro que recopila y procesa datos de sus usuarios, quienes deben registrarse y verificar su identidad previo a acceder a la referida herramienta.

En punto de los riesgos generados por los sistemas de IA, es preciso partir de la estructuración de la teoría de Alan Turing,³⁶ quien evidenció la posibilidad de diseñar

³⁵ Sol Calle, “Tecnologías emergentes y su impacto en la privacidad, ¿doctor Jekyll o señor Hyde?”, en *Disrupción tecnológica, transformación digital y sociedad*. Tomo II. *Políticas públicas y regulación en las tecnologías disruptivas*, ed. por Juan Carlos Henao y Santiago Téllez Cañas (Bogotá: Universidad Externado de Colombia, 2021), 670.

³⁶ Alan Turing, “Computing machinery and intelligence”, *Mind* (1950): 433-460.

máquinas capaces de pensar por sí mismas, tomar decisiones autónomas y hasta improvisar de acuerdo a su programación. Esta probabilidad teórica se convirtió en la realidad que actualmente vivimos, tanto así que los más diversos comportamientos y actividades de la vida diaria –como la relación de las personas con las entidades financieras, las recomendaciones en temas de entretenimiento, las noticias destacadas en redes sociales, la prestación de servicios de salud, el comercio electrónico, la elección de la ruta para trasladarse o la atención a los ciudadanos por parte de las entidades públicas– se realizan a través de algoritmos diseñados para tomar decisiones y para aprender de sus experiencias con el objeto de resolver situaciones futuras sin la intervención humana.³⁷

Ahora bien, pese a lo inteligente que pueda llegar a ser un sistema de IA,³⁸ una de las problemáticas producto de su utilización y que ha generado todo un debate ético se centra en los denominados sesgos algorítmicos, los cuales ponen en jaque la garantía y protección de los derechos humanos más sensibles, como son la dignidad humana, la igualdad, la no discriminación por razones de género, raza o condición socioeconómica, así como el acceso a la administración de justicia y a la libertad.

Para abordar el tema y siguiendo a Belloso,³⁹ un sesgo es *una desviación o inclinación hacia algo*, un preconceito inconsciente fundado en aspectos personales como las preferencias o los prejuicios sociales, económicos o culturales, que han sido aprendidos e interiorizados y nos permiten tomar decisiones automáticas de forma más rápida y minimizando los procesos mentales.

Sin embargo, algo tan natural de la mente humana como son las preferencias y los prejuicios son errores sistemáticos producidos en un muestreo cuando se prefieren unas respuestas sobre otras. Pero, tal como lo plantea Belloso,⁴⁰ ¿por qué se presentan estos sesgos humanos en las decisiones tomadas por la IA?

No es una pregunta fácil de responder, por cuanto los sesgos se pueden producir desde la etapa de diseño, esto es, en la programación del algoritmo, pero también en su entrenamiento; pues, por ejemplo, si se le enseña a identificar rostros humanos con solo un segmento de la población, quienes no cumplan con esas características no serán identificados como seres humanos.

En la etapa de uso y aplicación, los algoritmos hacen sus predicciones a partir de la búsqueda de correlaciones entre los datos disponibles, momento en el cual se pueden producir sesgos por estar implícitos en los datos computados o debido a errores en el entrenamiento que no pueden ser superados por la programación inicial.

³⁷ Alfonso Delgado de Molina Ruis, “Inteligencia artificial, machine learning y deep learning”, en *Regtech y Legaltech: fundamentos y desafíos regulatorios*, dir. por Nydia Remolina (Valencia: Tirant lo Blanch, 2020), 97-114.

³⁸ Piénsese en el ya recordado *Chat GPT* de OpenAI, o en *Bard* de Google o *Copilot* de Windows.

³⁹ Nuria Belloso, “La problemática de los sesgos algorítmicos”, en *Inteligencia artificial y filosofía del derecho*, dir. por Fernando Llano Alonso (Murcia: Laborum, 2022), 47.

⁴⁰ Belloso, “La problemática de los sesgos algorítmicos”, 47.

Finalmente, en las etapas restantes, conocidas como la de validación y la de presentación a los promotores, se pueden generar sesgos financieros, estadísticos o de *marketing*, prefiriendo que el algoritmo tome unas decisiones respecto a otras, a fin de optimizar los resultados perseguidos por quienes financiaron su elaboración.

De acuerdo con Mark Coeckelbergh,⁴¹ la necesidad de discutir estas implicaciones parte de casos puntuales, tal como ocurrió en Estados Unidos, en el sector justicia, con el algoritmo denominado COMPAS, programado para predecir las personas que tenían altas probabilidades de reincidir en la comisión de delitos. Se evidenció que las predicciones identificaban con mayor probabilidad de reincidencia a personas afrodescendientes, situación que también ocurrió en casos de solicitudes de créditos o empleos por parte de esta última población, sin que se pudiera determinar una razón diferente al origen étnico para justificar las decisiones tomadas por la IA.

Adicionalmente, se conoció el caso de la policía predictiva en Estados Unidos y en Bélgica, donde un algoritmo se utilizó para predecir los lugares de mayor probabilidad de ocurrencia de delitos, lo que condujo a que ciertas áreas urbanas fueran estigmatizadas. Idéntica situación ocurre con el reconocimiento facial empleado para vigilancia, entre otras cosas violando la privacidad de las personas y hasta prediciendo sus preferencias sexuales.⁴²

También se ha materializado un riesgo político, pues los algoritmos en redes sociales han difundido discursos de odio, como en el caso de Chatbot Tay de Microsoft en 2016, diseñado para tener charlas lúdicas en Twitter y que, una vez empezó a decidir de forma autónoma, esparció mensajes racistas en esa red social⁴³ y llegó a incidir en las elecciones presidenciales en Estados Unidos en ese mismo año.

En materia administrativa y a pesar de la evolución tecnológica, Miranzo⁴⁴ llama la atención respecto a la necesidad de cuestionarnos sobre si este es el momento para que las administraciones públicas implementen el uso de la IA y la ciencia de datos en el cumplimiento de sus funciones, tales como el diseño de procesos de selección, de criterios habilitantes o de evaluación; y respecto de si tales herramientas podrían aumentar el índice de transparencia de los procesos de contratación estatal, no solo poniendo en evidencia actos de corrupción, sino creando alertas tempranas que condujeran a evitarlos.

Por otra parte, se debe resaltar que un sector de la doctrina ha llamado la atención sobre el riesgo de la utilización de este tipo de tecnologías, pues en materia de IA no es posible determinar las razones o motivaciones que tiene en cuenta una máquina

⁴¹ Mark Coeckelbergh, *Ética de la inteligencia artificial* (Madrid: Cátedra, 2021), 18.

⁴² Coeckelbergh, *Ética*, 19.

⁴³ Coeckelbergh, *Ética*, 20.

⁴⁴ Javier Miranzo, "Inteligencia artificial y contratación pública", en *Administración electrónica, transparencia y contratación pública*, dir. por Isaac Martín y José Antonio Moreno (Madrid: Iustel, 2020), 105-142.

para tomar una determinada decisión, ni la lógica que usa para ello,⁴⁵ lo cual genera una tensión inmediata con algunos principios constitucionales de la función administrativa, como el de transparencia y el de publicidad, y hasta con el debido proceso, si la decisión administrativa es tomada por un sistema de IA.

En la misma línea, Sánchez Graells⁴⁶ resalta el riesgo de utilizar la IA y el *block chain* en la contratación pública, pues dicha tecnología opera a través de la recopilación de datos, los cuales no siempre son de buena calidad o confiabilidad, lo cual podría llevar al Estado a tomar decisiones sobre la base de datos inexactos o erróneos, con el riesgo que ello conlleva, si bien resaltando que tanto la IA basada en *Deep Learning* como el adecuado proceso de recopilación y procesamiento de datos pueden llevar al mejoramiento del proceso de contratación.

Para mitigar esta situación problemática se han propuesto varias metodologías y soluciones, que pasan por regulaciones normativas –como pueden ser leyes, o los reglamentos emitidos por la Unión Europea en la materia–, por propuestas éticas –como las diseñadas por la Unesco y otros organismos multilaterales–, así como por soluciones tecnológicas –como las diseñadas para evitar los sesgos algorítmicos y que pueden incluir hasta la creación de una nueva categoría de derechos fundamentales en el ámbito digital para proteger a las personas potencialmente vulnerables a la discriminación algorítmica–.

Sin embargo, los sesgos algorítmicos se producen principalmente en el momento de su diseño, pues los programadores imprimen sus propias creencias (consciente o inconscientemente); debido a ello, al momento de ser liberados los algoritmos, pueden segmentar o discriminar a individuos pertenecientes a ciertos grupos poblacionales, los cuales pueden notar o no la producción del sesgo. Es lo que ocurre en el caso de los asistentes artificiales *Siri*, *Alexa* y *Cortana*, propiedad de Apple, Amazon y Windows, respectivamente, los cuales tienen nombre y voz femenina, hecho que ha sido catalogado como un sesgo de género imperceptible.⁴⁷ Por las citadas razones es preciso preguntarse: si el avance tecnológico de la IA actual produce esta serie de “errores o mal funcionamiento”, ¿los Estados deberían utilizarla para la toma de decisiones públicas de forma autónoma?

Con ese cuestionamiento en mente, valga proponer un escenario hipotético a fin de invitar al lector a hacer su propia reflexión sobre el particular. En la actualidad existen vehículos autónomos o no tripulados, que tienen como función principal transportar de un punto a otro, de forma segura, a su pasajero sin intervención humana. ¿Qué ocurriría si, durante el trayecto, se atravesara una persona y el vehículo debe decidir si evita el atropellamiento, con una alta probabilidad de que se lesione

⁴⁵ Coeckelbergh, *Ética*, 16.

⁴⁶ Albert Sanchez-Graells, “Data-driven and digital procurement governance: revisiting two well-known elephant tales”, *Communications Law - Journal of Computer, Media and Telecommunications Law* (2019), <http://dx.doi.org/10.2139/ssrn.3440552>.

⁴⁷ Belloso, “La problemática de los sesgos algorítmicos”, 60.

la persona que está en su interior, o arrollar al peatón?, ¿cuáles serían los parámetros éticos que debería valorar el algoritmo? Y si cambiamos el vehículo por un dron que funciona con IA y cuenta con armamento de última tecnología, ¿cuál sería la conclusión ante un dilema ético?

Lo anterior evidencia la necesidad de analizar desde las humanidades el uso seguro de la IA y la conveniencia de someterla a la sociedad, así como la oportunidad de diseñar mecanismos tendientes a la protección de los derechos humanos, cuya lesión puede ocurrir en cualquier momento.

Adicionalmente y como se resaltó líneas atrás, el Estado colombiano está implementando la IA para aumentar sus indicadores de eficiencia en la prestación de los servicios a su cargo, dejando en manos de algoritmos la decisión de asuntos particulares de competencia exclusiva del Estado, sin que se hayan estudiado o siquiera dimensionado los efectos sociales que podría llegar a producir un sesgo algorítmico frente a los ciudadanos; con mayor razón teniendo en cuenta que hablamos de un país diverso, multicultural y pluriétnico, en los términos del artículo 7 de la Constitución, es decir que el universo de personas que se podrían ver afectadas en sus derechos puede ser mayor que en el caso de Europa o Norteamérica.

Por lo anterior, es necesario atender a las experiencias internacionales referentes a posibles situaciones de discriminación, teniendo en cuenta que la tecnología usada en otras latitudes funciona de la misma forma que en Colombia, lo que aumenta el riesgo para los derechos humanos de los ciudadanos destinatarios de las decisiones automáticas.

Además de lo anterior, los sistemas basados en IA, como cualquier otro tipo de propiedad intelectual, se encuentran protegidos por registros y patentes que amparan los derechos de autor de las compañías proveedoras de esta tecnología, razón por la cual la administración pública adquiere licencias de uso mas no la titularidad de los códigos fuente de los algoritmos; esta situación, en la práctica, significa que el Estado paga por entrenarlos y corre con las consecuencias en caso de sesgos algorítmicos o de la producción de daños a terceros, pero la compañía tecnológica propietaria de la IA solo es responsable en los términos del contrato suscrito con la entidad pública contratante, y nada impide que sobre la base del mismo desarrollo cree y comercialice otros sistemas de IA.

3. El enfoque de los derechos humanos frente a los sesgos algorítmicos

Como se ha expuesto líneas atrás, el uso de la IA ha traído consigo una serie de efectos negativos respecto a la garantía y protección de los derechos humanos, pues los llamados sesgos algorítmicos han generado situaciones de discriminación por razones de raza, género, condición social o religiosa, situación que, traída al ejercicio del poder público, implica un riesgo aún mayor, pues no se trata simplemente de

que un grupo poblacional esté siendo discriminado, sino que sería el propio Estado el causante de tal situación.

Bajo esa perspectiva, es necesario analizar cómo, desde el derecho constitucional, se pueden proteger los derechos humanos afectados por las tecnologías disruptivas, poniendo el acento en los sesgos algorítmicos en los cuales hemos centrado esta breve reflexión.

En los términos señalados en los acápites precedentes, el uso de sistemas basados en IA, *big data* e internet de las cosas ha generado una afectación significativa a los derechos a la privacidad, a la intimidad, a la dignidad humana, a la libertad en sus múltiples facetas y al debido proceso. Pues, en efecto y como se afirmó antes, los algoritmos diseñados para determinar si una persona puede acceder a libertad condicional, a un crédito bancario, a beneficios sociales o exenciones de impuestos, en caso de basar sus decisiones en prejuicios de orden racial, étnico, de género o aspectos similares de la esfera individual de las personas, afectan uno o varios de los referidos derechos, situación que se agrava con la dificultad para identificar en tiempo real cuándo se materializa el sesgo; hecho este que, a diferencia del *big data*, aún no puede ser parametrizado o estandarizado a fin de evitarlo o mitigarlo.

En ese sentido, respecto a los efectos producidos por los sesgos algorítmicos, De Asís⁴⁸ propone contrarrestarlos mediante la utilización de la transparencia y publicidad de los datos, la participación en el diseño de aquellos y la formación en ética digital, atendiendo siempre a la diversidad. En esa línea, sostiene que los efectos discriminatorios producidos por el uso de la IA deben abordarse desde la óptica de los derechos fundamentales, más que desde la ética digital, utilizando mecanismos para proteger a los ciudadanos en el entorno digital.

Es preciso resaltar que el estudio de las consecuencias producto del uso de la IA es interdisciplinario, por cuanto estas involucran el derecho, las ciencias computacionales (término que abarca a todas las involucradas) y algunas ciencias sociales como la sociología, la ética y la moral. Así como la solución a la discriminación producto de los sesgos algorítmicos no puede provenir de una sola de estas áreas del conocimiento, y, por tanto, pensar que la discriminación va a desaparecer porque elevemos a una ley positiva algún marco ético o algunas reglas para tratar de eliminarlos es sumamente ingenuo, pues equivale a pensar que la tasa de homicidios se va a reducir a cero por el hecho de tipificar dicha conducta en el Código Penal.

Lo anterior, teniendo en cuenta que la utilización de la IA para tomar decisiones automáticas es un reflejo de la corriente transhumanista, que aboga por el mejoramiento moral de la especie a través de la utilización de la edición genética del cuerpo

⁴⁸ Rafael de Asís, “Ética, tecnología y derecho”, en *Inteligencia artificial y filosofía del derecho*, dir. por Fernando Llano Alonso (Murcia: Laborum, 2022), 29.

humano y de la tecnología para promover cambios culturales, sociales y políticos, a partir de la IA y el *big data*.⁴⁹

Harari⁵⁰ explica la corriente transhumanista a partir de uno de los postulados defendidos por esta: según dicho postulado, el universo se reduce a flujos de datos, y el valor de las cosas y los fenómenos depende de su capacidad para procesarlos o de su contribución a dicho proceso; la ventaja que ofrece esta línea de pensamiento es la identificación *de los datos* como el común denominador de todas las áreas del conocimiento, lo cual permite analizar con los mismos métodos la *Quinta sinfonía* de Beethoven, la burbuja de la bolsa de valores y el virus de la gripe,⁵¹ teniendo en cuenta que siempre son tan solo flujos de datos. Con lo cual se valoran los hechos y la razón dejando de lado las emociones humanas.

De acuerdo con esta corriente de pensamiento, podemos confiar completamente en los datos y los algoritmos para resolver nuestros problemas y mejorar nuestros resultados por medio de la automatización, prefiriendo la IA sobre las capacidades humanas.

También confían en que la tecnología suministre un proceso de toma de decisiones racional, objetivo, imparcial, que los humanos somos incapaces de realizar, razón por la cual confían en mantener el poder mientras la IA se encarga de las tareas operativas cotidianas.⁵² Bajo esa perspectiva, limitar la solución a esta problemática al diseño de marcos éticos o a la expedición de normas del nivel que sea no va a mitigar la discriminación algorítmica, así dichas reglas busquen el reconocimiento de nuevos derechos y garantías en el ámbito digital; pues, así como la IA es un área interdisciplinaria del conocimiento, la solución debe tener ese mismo enfoque, y no abordar la cuestión desde varias ópticas pone en riesgo la efectividad de los remedios que se intente aplicar.

A título de ejemplo, si se crearan nuevos derechos humanos como única acción correctiva de los sesgos algorítmicos, la sobreproducción de derechos disminuiría su efectividad, así como la de los mecanismos para su protección, pues se produciría un efecto similar a la inflación en términos económicos, tal como lo afirma de Asís.⁵³

Belloso,⁵⁴ a su turno, explica los otros mecanismos que se han propuesto para la supresión de sesgos algorítmicos en sistemas de IA, consistentes en la participación de personas potencialmente susceptibles de ser víctimas de discriminación a fin de evitar que en la etapa de desarrollo se puedan presentar dichos prejuicios.

⁴⁹ Antonio Diéguez, *Cuerpos inadecuados. El desafío transhumanista a la filosofía* (Barcelona: Herder, 2021), 81.

⁵⁰ Yuval Noah Harari, *Homo Deus. Breve historia del mañana* (Barcelona: Penguin Random House, 2021), 400.

⁵¹ Harari, *Homo Deus*, 400.

⁵² Mayer-Schönberger, Cukier y de Véricourt, *Framers*, 26-27.

⁵³ de Asís, "Ética, tecnología y derecho", 38.

⁵⁴ Belloso, "La problemática de los sesgos algorítmicos", 53.

Por otra parte, están las contramedidas para combatir los prejuicios algorítmicos, que consisten en incluir las políticas antidiscriminación dentro de las normas del aprendizaje automático, lo que es aplicable a cualquier sistema de IA.

Dichas medidas, también se centran en el proceso de recolección y tratamiento de los datos y en el entrenamiento algorítmico a partir de estos, es decir que, en las etapas de entrenamiento y presentación de los algoritmos, se debe hacer una revisión tendiente a evidenciar la presencia de sesgos, con la participación de los grupos potencialmente susceptibles de ser discriminados.

Ahora bien, teniendo en cuenta que los sesgos se pueden producir en cualquier momento, Belloso propone someter los sistemas de IA a una auditoría algorítmica periódica, tendiente a diagnosticar sesgos y proponer acciones correctivas; que es lo que hace Eticas Consulting, la compañía que diseñó la primera *Guía de Auditoría Algorítmica* con enfoque en los sesgos.⁵⁵

En ese orden de ideas, si bien el ámbito de los derechos humanos es la óptica desde la que se deben abordar los efectos del uso de la IA, al ser estos los afectados con la producción de sesgos, no es posible afirmar que la creación de nuevos derechos, la educación en ética, las metodologías para auditorías algorítmicas, las herramientas antis sesgos, el diseño de políticas u otros remedios tecnológicos sean la solución para mitigar la discriminación algorítmica, pues la IA y los sesgos que se pueden producir cuando ella está en operación no son iguales en todos los casos: en efecto, no es lo mismo una máquina que decide si se otorga un crédito bancario que una que debe resolver un conflicto u otra que fija precios para competir en el mercado.

Es decir que, en la actualidad, la única forma para determinar si un enfoque, una regulación o una herramienta tecnológica puede evitar, o al menos alertar de la producción de un sesgo algorítmico es probarlo directamente en la población destinataria de las decisiones que se pretenden automatizar, lo cual no es otra cosa que experimentar con humanos, generando automáticamente una afectación a la dignidad humana.

Conclusiones y reflexiones finales

A manera de conclusión, valga llamar la atención respecto a la utilización de los sistemas basados en IA para realizar tareas propias de las competencias del Estado, hecho que evidencia la falta de legitimación del sistema democrático en Colombia y el intento de los gobiernos por legitimarlo *ex post*, mediante estrategias basadas en redes sociales, *big data* e IA.

Con el auspicio de organismos multilaterales como la OCDE y el Banco Interamericano de Desarrollo (BID), Colombia diseñó la política pública de transformación digital del Estado, con la finalidad de volcar a los sectores público y privado

⁵⁵ Belloso, “La problemática de los sesgos algorítmicos”, 53.

a la realización de sus competencias y actividades mediante el uso de tecnologías disruptivas, para lograr mayores índices de eficiencia en su gestión.

La IA ha sido usada en Colombia para la toma de decisiones administrativas y judiciales, sin embargo, no se han realizado estudios o análisis sobre el impacto de dicha tecnología en los ciudadanos, destinatarios de esas decisiones automáticas; pero, eso sí, se ha generado en el imaginario colectivo la creencia de que la IA es segura, confiable, objetiva y más eficiente que los seres humanos, como lo demostrarían sus éxitos en el campo de la medicina, lo que la haría prácticamente incuestionable.

Algunos efectos que estas tecnologías han causado en la población implican la vulneración de sus derechos humanos, pues las garantías mínimas, como la privacidad, la intimidad, el *habeas data*, la dignidad, la igualdad, la libertad, el debido proceso y sus mecanismos de protección, se han visto gravemente afectados por las grandes compañías tecnológicas, las cuales aprovechan la relación entre las personas y los dispositivos digitales para recaudar y procesar sus datos, a fin de comercializarlos e influir en todos los ámbitos de sus vidas, con o sin el consentimiento de los titulares de la información.

Asimismo, los derechos humanos se han venido poniendo en riesgo con el uso de sistemas de IA, pues el uso de dicha tecnología en los sectores públicos y privados ha generado que la toma de decisiones automáticas incluya sesgos raciales, de género y socioeconómicos, entre otros, los cuales han afectado a grupos poblacionales en su dignidad y su patrimonio, por el solo hecho de su pertenencia a ellos.

En el sector público se han puesto en funcionamiento múltiples iniciativas para modernizar y aumentar los índices de eficiencia estatal con el uso de algoritmos, los cuales en algunos casos (como los citados en este escrito) han discriminado a personas consideradas como minorías, situación que se ha intentado explicar por los expertos en la materia señalando que al momento de programar los algoritmos las creencias y los prejuicios de los desarrolladores quedan implícitas en aquellos, dado que, en muchos casos, los desarrolladores pertenecen a grupos poblacionales propensos a excluir a grupos vulnerables.

Otro momento de aparición de los sesgos es el procesamiento de los datos, pues en muchos casos la calidad de la información está permeada por errores que, al momento de generar las parametrizaciones y estandarizaciones necesarias para producir las predicciones, los patrones o tendencias, arrojan conclusiones erróneas y, en consecuencia, la toma de decisiones equivocadas o injustas.

Las vulneraciones de los derechos humanos producto de los sesgos algorítmicos han sido un fenómeno ampliamente estudiado, y sobre esa base se han propuesto múltiples formas de mitigar sus efectos, tales como el diseño de marcos éticos que pretenden evitar los sesgos a partir de un catálogo de reglas y buenas prácticas que deberían seguir los programadores.

Otro mecanismo para evitar los sesgos consiste en vincular a personas pertenecientes a grupos poblacionales potencialmente susceptibles de ser discriminados, a fin de evitar que en el diseño inicial se incluyan prejuicios; así como de la inclusión de

políticas antidiscriminatorias en las reglas de programación y el procesamiento de datos mediante *big data*, y sometiendo su funcionamiento a auditorías periódicas, con la creación de nuevos derechos humanos para la era digital, la educación en ética y la transparencia de los datos, entre otras iniciativas que apuntan al mismo objetivo.

Sin embargo, teniendo en cuenta este panorama, se evidencian situaciones sumamente problemáticas para la sociedad actual, pues se están liberando al público en general, sistemas de IA sin ningún tipo de control o entrenamiento, para ser usados por la población sin siquiera saber cuál es el riesgo que se corre por esta interacción.

El Estado ha venido implementando el empleo de esta tecnología para el cumplimiento de sus funciones, y con la pretensión de hacerlo al margen de los sesgos. Pero el hecho es que, cuando las administraciones públicas usan los sistemas de IA, no son propietarias del código fuente de los algoritmos, pues estos están salvaguardados por registros y patentes que protegen los derechos de autor de las compañías proveedoras de la tecnología; lo cual significa que el Estado paga por el uso y además entrena los algoritmos y corre con las consecuencias en caso de errores técnicos, sesgos o daños a terceros, mientras la compañía tecnológica sigue siendo la única propietaria de la IA, lo cual abre el espacio para otro debate.

Finalmente, es preciso reiterar una de las reflexiones hechas a lo largo del escrito, en el sentido de que, si esta tecnología se encuentra en plena etapa de desarrollo, en la actualidad no es posible determinar todos los efectos que puede producir su uso; además de que las consecuencias conocidas ya han sido en algunos casos nefastas para la garantía y el respeto de los derechos humanos. Así las cosas, ¿no deberíamos cuestionarnos si, a pesar de que se puedan tomar decisiones públicas automáticas desde el punto de vista técnico, es necesario dar primero este debate, antes de usar la IA en el sector público, en atención al riesgo que se genera para los derechos humanos, evitando con ello que el Estado los vulnere de forma automática (siendo tales derechos, dicho sea de paso, el resultado de múltiples luchas sociales y políticas de la humanidad)?

A lo mejor, hasta tanto no se tenga definido el uso seguro de la IA por parte del Estado, se deba considerar que aún no es el momento para emplearla como herramienta en el cumplimiento de las funciones estatales, aun cuando técnicamente sea posible realizarlas en menor tiempo y con menor esfuerzo y costo, puesto que la búsqueda de la eficiencia administrativa y judicial no puede justificar que se ponga siquiera en riesgo la garantía de los derechos humanos de los ciudadanos.

Bibliografía

- ANDERS, Günter. *La obsolescencia del hombre*. Volumen II. *Sobre la destrucción de la vida en la época de la Tercera Revolución Industrial*. Valencia: Pre-textos, 2011.
- BARONA VILLAR, Silvia. *Algoritmización del derecho y de la justicia. De la inteligencia artificial a la Smart Justice*. Valencia: Tirant lo Blanch, 2021.

- BELLOSO, Nuria. “La problemática de los sesgos algorítmicos”. En *Inteligencia artificial y filosofía del derecho*, dir. por Fernando LLANO ALONSO. Murcia: Laborum, 2022.
- CALLE, Sol. “Tecnologías emergentes y su impacto en la privacidad, ¿doctor Jekyll o señor Hyde?”. En *Disrupción tecnológica, transformación digital y sociedad*. Tomo II. *Políticas públicas y regulación en las tecnologías disruptivas*. Editado por Juan Carlos HENAO y Santiago TÉLLEZ CAÑAS. Bogotá: Universidad Externado de Colombia, 2021.
- COECKELBERGH, Mark. *Ética de la inteligencia artificial*. Madrid: Cátedra, 2021.
- CONSEJO NACIONAL DE POLÍTICA ECONÓMICA Y SOCIAL, República de Colombia, Departamento Nacional de Planeación. Documento Conpes 3975 de 8 de noviembre de 2019, “Política Nacional para la Transformación Digital e Inteligencia Artificial”. <https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/3975.pdf>.
- DE ASÍS, Rafael. “Ética, tecnología y derecho”. En *Inteligencia artificial y filosofía del derecho*, dir. por Fernando LLANO ALONSO. Murcia: Laborum, 2022.
- DEL ÁGUILA, Rafael. “La política: el poder y la legitimidad”. En *Manual de ciencia política*, coord. por Rafael DEL ÁGUILA. Madrid: Trotta, 2000.
- DELGADO DE MOLINA RUIS, Alfonso. “Inteligencia artificial, machine learning y deep learning”. En *Regtech y Legaltech: fundamentos y desafíos regulatorios*, dir. por Nydia REMOLINA. Valencia: Tirant lo Blanch, 2020.
- DIÉGUEZ, Antonio. *Cuerpos inadecuados. El desafío transhumanista a la filosofía*. Barcelona: Herder, 2021.
- FERRARI, Vincenzo. *Funciones del derecho*. Bogotá: Universidad Externado de Colombia, 2014.
- HABERMAS, Jürgen. *Ciencia y técnica como “ideología”*. Madrid: Tecnos, 2007.
- HARARI, Yuval. *Homo Deus. Breve historia del mañana*. Barcelona: Penguin Random House, 2021.
- LAFONT, Cristina. *Democracia sin atajos*. Madrid: Trotta, 2021.
- LÓPEZ, Marcos. *Fundamentos para un derecho de la inteligencia artificial. ¿Queremos seguir siendo humanos?* Valencia: Tirant lo Blanch, 2020.
- MAYER-SCHÖNBERGER, Viktor, Kenneth CUKIER y Francis DE VÉRICOURT. *Framers: la virtud humana en la era digital*. Madrid: Turner, 2021.
- MIRANZO, Javier. “Inteligencia artificial y contratación pública”. En *Administración electrónica, transparencia y contratación pública*, dir. por Isaac MARTÍN y José Antonio MORENO. Madrid: Iustel, 2020.
- MOZO SEOANE, Antonio. *Los límites de la tecnología. Marco ético y regulación jurídica*. Madrid: Reus, 2021.
- ORGANIZACIÓN PARA LA COOPERACIÓN Y EL DESARROLLO ECONÓMICOS (OCDE). *The Strategic and Responsible Use of Artificial Intelligence in the Public Sector of Latin America and the Caribbean*, 22 de marzo de 2022. <https://www.oecd-ilibrary.org/governance/>

the-strategic-and-responsible-use-of-artificial-intelligence-in-the-public-sector-of-latin-america-and-the-caribbean_1f334543-en.

PADRÓN, Floralba. “Presentación”. *¿El Estado constitucional en jaque?* Tomo I. *Los retos del componente democrático*. Bogotá: Universidad Externado de Colombia, 2018.

PINZÓN, Mario. *Aproximaciones al Análisis Económico del Derecho*. Bogotá: Universidad Externado de Colombia, 2010.

RAMÍO, Carles. *Inteligencia artificial y administración pública. Robots y humanos compartiendo el servicio público*. Madrid: Catarata, 2019.

RÚA, Carlos. “La legitimidad en el ejercicio del poder político en el Estado social de derecho. Una revisión desde el caso colombiano”. *Ius et Praxis* 19, n.º 2 (2013). <http://dx.doi.org/10.4067/S0718-00122013000200004>.

SANCHEZ-GRAELLS, Albert, “Data-driven and digital procurement governance: Revisiting two well-known elephant tales”. *Communications Law - Journal of Computer, Media and Telecommunications Law* (2019). <http://dx.doi.org/10.2139/ssrn.3440552>.

SCHWAB, Klaus. *La Cuarta Revolución Industrial*. Barcelona: Debate, 2018.

TURING, Alan. “Computing machinery and intelligence”. *Mind* (1950): 433-460.

WEBER, Max. *Economía y sociedad*. México: Fondo de Cultura Económica, 2014.