

GIGAPP

Estudios / Working Papers

ISSN 2174-9515

NUMERO ESPECIAL
GOBERNANZA DE DATOS E INTELIGENCIA ARTIFICIAL

Vol. 10 Año (2023) Núms. 267-272, págs.329-434

Gutiérrez, Juan David Presentación: Retos de la gobernanza de datos y de inteligencia
Flórez Rojas, María Lorena artificial en el sector público Iberoamericano
(2023-267. págs.329-334)

Flórez Rojas, María Lorena Las evaluaciones de impacto algorítmico: Una visión más allá de
la ética
(2023-268 págs. 335-350)

Casarrubios Blanco, Elena Influencia de las políticas de promoción de la Inteligencia
Artificial de la UE en el proceso modernizador de las
Administraciones Públicas españolas
(2023-269 págs. 351-364)

Gutiérrez, Juan David Adopción de sistemas de decisión automatizada en el sector
Muñoz-Cadena, Sarah público: Cartografía de 113 sistemas en Colombia
(2023-270 págs.365-395)

Jara Iñiguez, Irma Estado de la transparencia y acceso a la información pública en
Ecuador pos-pandemia: ¿Han cambiado las necesidades
ciudadanas?
(2023-271 págs. 396-415)

Gutiérrez, Juan David Lineamientos para el uso de inteligencia artificial en contextos
universitarios (v 5.0)
(2023-272 págs. 416-434)



Grupo de Investigación en
Gobierno, Administración
y Políticas Públicas

GIGAPP Estudios Working Papers es una publicación de la
Asociación Grupo de Investigación en Gobierno,
Administración y Políticas Públicas

www.gigapp.org

Consejo de Dirección

César Nicandro Cruz-Rubio. GIGAPP. España
(Director)

Palmira Chavero-Ramírez. FLACSO. Ecuador
Cecilia Güemes. UAM. España.
José A. Hernández-Bonivento. ICHEM. Chile
Álvaro Ramírez-Alujas. INAP. U. Chile

Comité Editorial

Victoria Alsina Burgues. KSG. USA
Roberto Castellanos Cereceda. UNAM . México
César Nicandro Cruz-Rubio, GIGAPP. España
Cristiana Freitas. UnB. Brasil.
Flavia Galvani Silva. FOG. España
Leandro Grass Preixoto. UnB. Brasil
Efrén Guerrero Salgado. PUCE. Ecuador
José A. Hernández Bonivento. ICHEM. Chile
Juana López Pagán. FOG. España
Diego Pablo Pando. UNSM. Argentina
Erika Rodríguez Pinzón. UNIR. España.



Licencia 4.0 (España) Creative Commons.
Reconocimiento-No Comercial-Sin Obras Derivadas.
<http://creativecommons.org/licenses/by-nc-nd/4.0/es/>

ISSN 2174-9515



Grupo de Investigación en
Gobierno, Administración
y Políticas Públicas

GIGAPP
Estudios / Working Papers

Presentación:

Retos de la gobernanza de datos y de inteligencia artificial en el sector público Iberoamericano

Gutiérrez, Juan David


Universidad de los Andes, Colombia

 juagutie@uniandes.edu.co

ORCID ID: [0000-0002-7783-4850](https://orcid.org/0000-0002-7783-4850)

Flórez Rojas, María Lorena

Groningen University, Países Bajos

 m.l.florez.rojas@rug.nl

ORCID ID: [0000-0002-8264-8227](https://orcid.org/0000-0002-8264-8227)

Resumen

El sector público de Iberoamérica enfrenta nuevos desafíos en relación con la gobernanza de datos y el uso de sistemas de decisión automatizada (SDA). Por una parte, los Estados cada vez recolectan y generan un volumen más grande y variado de datos en el ejercicio de sus funciones. Por otra parte, cada vez más entidades públicas nacionales y subnacionales de Iberoamérica están adoptando SDA para el cumplimiento de sus funciones. Los SDA incluyen algoritmos sencillos (basados en reglas) que automatizan total o parcialmente la toma de decisiones, así como sistemas de inteligencia artificial (IA) basados métodos y técnicas de aprendizajes automatizados (machine learning) y procesamiento de lenguaje natural, entre otros. Estos retos de la gobernanza de datos y de IA en el sector público Iberoamericano son abordados por los artículos que componen este número especial de Estudios Working Papers GIGAPP. Este editorial introduce los cinco textos que abordan diversas temáticas: evaluaciones de impacto algorítmicas; políticas de promoción de la IA; adopción de SDA en el sector público; transparencia y acceso a la información pública; y, políticas de uso de IA en contextos académicos. Los artículos, fueron escritos por académicas y académicos de diferentes nacionalidades e incluyen casos de estudio de Colombia, Ecuador, y España.

Palabras clave

inteligencia artificial, sistemas de decisión automatizada, datos personales, gobernanza, acceso a información pública, transparencia algorítmica.

Abstract

The Ibero-American public sector faces new challenges in relation to data governance and the use of automated decision-making systems (ADS). On the one hand, States are increasingly collecting and generating a larger and more varied volume of data in the exercise of their functions. On the other hand, more national and subnational public entities in Ibero-America are adopting ADSs to carry out their functions. ADS include rule-based algorithms that fully or partially automate decision making, as well as artificial intelligence (AI) systems based on machine learning methods and techniques and natural language processing, among others. These challenges of data governance and AI in the Ibero-American public sector are addressed by the articles included in this special issue of Estudios Working Papers GIGAPP. This editorial introduces the six texts that address various topics: algorithmic impact assessments; AI promotion policies; adoption of ADS in the public sector; transparency and access to public information; and policies for the use of AI in academic contexts. The articles were written by academics of different nationalities and include case studies from Colombia, Ecuador, and Spain.

Keywords

artificial intelligence, automated decision-making systems, personal data, governance, access to public information, algorithmic transparency.

Resumo

O setor público ibero-americano enfrenta novos desafios em relação à governança de dados e ao uso de sistemas automatizados de tomada de decisão (ADS). Por um lado, os Estados estão cada vez mais coletando e gerando um volume maior e mais variado de dados no exercício de suas funções. Por outro lado, cada vez mais entidades públicas nacionais e subnacionais na Ibero-América estão adotando ADSs para o desempenho de suas funções. ADS incluem algoritmos baseados em regras que automatizam total ou parcialmente a tomada de decisões, bem como sistemas de inteligência artificial (IA) baseados em métodos e técnicas de aprendizado de máquina e processamento de linguagem natural, entre outros. Esses desafios de governança de dados e IA no setor público ibero-americano são abordados pelos artigos incluídos nesta edição especial dos Estudios Working Papers GIGAPP. Este editorial apresenta os seis textos que abordam vários tópicos: avaliações de impacto algorítmicas; Políticas de promoção de IA; adoção de ADS no setor público; transparência e acesso à informação pública; e políticas para o uso de IA em contextos acadêmicos. Os artigos foram escritos por acadêmicos de diferentes nacionalidades e incluem estudos de caso da Colômbia, Equador e Espanha.

Palavras-chave

inteligência artificial, sistemas automatizados de tomada de decisão, dados pessoais, governança, acesso à informação pública, transparência algorítmica

Introducción

El sector público de Iberoamérica enfrenta nuevos desafíos en relación con la gobernanza de datos y el uso de sistemas de decisión automatizada. Por una parte, los Estados cada vez recolectan y generan un volumen más grande y variado de datos en el ejercicio de sus funciones. Lo anterior está asociado con cuatro principales retos de política y gestión pública que están interconectados. En primer lugar, que el Estado cuente con una adecuada infraestructura para la explotación de dichos datos para la creación de valor público y la eficaz gestión pública. En segundo lugar, que los datos estén a disposición del público –siempre que no involucren información personal o reservada– para que puedan ser aprovechados para actividades con o sin ánimo de lucro. En tercer lugar, velar por el cumplimiento pleno de las legislaciones de protección de datos y privacidad en el tratamiento de datos personales recolectados por el Estado. Finalmente, construir una gobernanza de los datos públicos que garantice el cumplimiento de los tres objetivos anteriores.

Por otra parte, cada vez más entidades públicas nacionales y subnacionales de Iberoamérica están adoptando sistemas de decisión automatizada (SDA) para el cumplimiento de sus funciones. Los SDA incluyen algoritmos sencillos (basados en reglas) que automatizan total o parcialmente la toma de decisiones, así como sistemas de inteligencia artificial (IA) basados métodos y técnicas de aprendizajes automatizados (machine learning) y procesamiento de lenguaje natural, entre otros.

El desarrollo, adquisición, pilotaje, adopción, implementación y evaluación de los SDA por parte de entidades públicas está asociado con al menos tres desafíos de política y gestión pública. En primer lugar, en el rol de Estado-promotor, generar las condiciones habilitantes para que tanto las entidades públicas como las privadas puedan adoptar efectivamente estas nuevas tecnologías. En segundo lugar, en el rol de Estado-consumidor, propender por la adquisición de herramientas idóneas, que efectivamente contribuyan a resolver problemas públicos o a mejorar la prestación de servicios públicos, respetando tanto estándares de derechos humanos como éticos. Finalmente, en el rol de Estado-regulador, co-crear con múltiples partes un marco normativo adecuado para asegurar que el uso de estos sistemas genere valor público y para prevenir que su uso vulnere derechos fundamentales.

Estos retos de la gobernanza de datos y de IA en el sector público Iberoamericano son abordados por los artículos que componen este número especial de Estudios Working Papers GIGAPP. Este editorial introduce los cinco textos que abordan diversas temáticas: evaluaciones de impacto algorítmicas; políticas de promoción de la IA; adopción de SDA en el sector público; transparencia y acceso a la información pública; y, políticas de uso de IA en contextos académicos. Los artículos, fueron escritos por académicas y académicos de diferentes nacionalidades e incluyen casos de estudio de Colombia, Ecuador, y España.

Aportes de la edición especial a la literatura de gobernanza de datos y la IA en el sector público

El primer artículo titulado “Las evaluaciones de impacto algorítmicas: Una visión más allá de la ética”, de autoría de María Lorena Flórez Rojas, aborda la relevancia del análisis de impacto algorítmico (AIA) como una metodología fundamental para evaluar los riesgos asociados a los juicios de los sistemas de inteligencia

artificial (IA) y su capacidad para respaldar el desarrollo y pilotaje de marcos éticos de IA en la región. El artículo analiza cómo el AIA puede implementarse para asegurar que la IA sea aprovechada de manera responsable y en sintonía con los valores de cada sociedad. El análisis incluye una revisión general del AIA mediante ejemplos internacionales y un contraste con el toolkit ético legal (TELia), desarrollado por la autora, el cual sirve como guía para asegurar el uso ético de la IA en línea con los principios y valores sociales.

La autora reconoce los avances en la integración de principios éticos en algunos sectores industriales, pero a su vez pone de presente los diversos desafíos en la aplicación efectiva de estos principios en los desarrollos de IA. En este sentido, aboga por la adopción más amplia del AIA como una vía para aumentar la confianza en las aplicaciones de IA y fomentar su beneficio social. El artículo subraya la importancia de no limitar la discusión de IA exclusivamente a la ética, sino considerar estos sistemas como parte integral del desarrollo público y privado. Por tanto, la autora destaca la necesidad de implementar el AIA en el desarrollo y uso de la tecnología de IA y, a su vez, insta a los reguladores a establecer reglas y estándares claros para abordar las implicaciones asociadas.

El segundo artículo del número especial, escrito por Elena Casarrubios Blanco, tiene como objetivo realizar un mapeo sobre la política de promoción de inteligencia artificial llevada a cabo por la Unión Europea en el marco del proceso de modernización de las Administraciones públicas en España, en particular, la Administración General del Estado español. La autora hace un recorrido por la evolución tecnológica asociada a la IA y su implementación en la Administración pública española ha seguido un camino paralelo al de la Unión Europea. El texto presenta los puntos generales y principales sobre las políticas regulatorias de la UE y de España en materia de IA como motor de las políticas de promoción de las Tecnologías de la Información y las Comunicaciones. La contribución principal del artículo, titulado "Influencia de las políticas de promoción de la Inteligencia Artificial de la UE en el proceso modernizador de las Administraciones Públicas españolas", radica en presentar una visión general de cómo la implementación de la IA en las administraciones públicas españolas puede apoyar procesos con ciudadanos, empresas y el mismo Estado.

El tercer artículo, escrito por Juan David Gutiérrez y Sarah Muñoz-Cadena, presenta los resultados de una nueva base de datos que identifica y caracteriza 113 sistemas de decisión automatizada (SDA) utilizados por 52 entidades del sector público en Colombia. Los autores caracterizaron cada sistema a partir de 40 variables y más de 300 fuentes de información disponible públicamente. De los 113 sistemas identificados, 111 se encuentran en ejecución (86,7%) o en fase de pilotaje (11,5%). El 93% de los 111 sistemas son usados por el poder ejecutivo y el remanente por entidades de la rama judicial (4%) y órganos de control independientes (3%). Adicionalmente, los autores reportan que tres de cada cuatro SDA fueron implementados por entidades del orden nacional y el resto por entidades subnacionales. A lo largo del documento se mencionan varios ejemplos de los sistemas, cuáles son sus objetivos o cómo buscan aportar a diferentes actividades estatales. Adicionalmente, a partir del análisis de la información de esta nueva base de datos los autores identificaron potenciales riesgos del uso e implementación de estos SDA en particular en lo referente a cuatro tipos de derechos fundamentales: protección de datos personales y la privacidad, derecho de petición y acceso a información pública, derecho a la igualdad y no discriminación, y debido proceso y derecho a la defensa.

El grueso de la literatura sobre adopción de SDA en el sector público se suele concentrar en estudiar casos particulares y, principalmente, casos en países del Norte Global. En ese sentido, esta investigación aporta a la literatura con una cartografía de SDA sobre los que existe información pública de su uso en un país del Sur Global como lo es Colombia. Además, si bien el Estado colombiano tiene tres repositorios con información sobre algunos algoritmos que utiliza para tomar decisiones, los autores demuestran que la información de

estos es incompleta y persiste un subregistro de sistemas. Finalmente, la base de datos que presenta constituye un recurso valioso para futuras investigaciones sobre transparencia algorítmica y sobre el ciclo de vida de los proyectos de SDA en Colombia.

El cuarto artículo del número especial, contribuido por Irma Jara Iñiguez, estudia la evolución de la legislación sobre transparencia y acceso a la información pública en Ecuador e identifica cómo han cambiado las demandas de información pública por parte de la ciudadanía. Además, evalúa la utilidad de la información transparentada por el Estado para efectos de que la ciudadanía participe en actividades de co-gestión pública y en ejercicios de control social.

Adicionalmente, la investigación aborda un caso de estudio puntual: las solicitudes de acceso a información recibidas por la Presidencia de la República de Ecuador en el periodo 2017 - 2021. A partir de dicho caso, la autora encuentra que las solicitudes de información se enfocan en asuntos tales como los informes que respaldan las decisiones o la ejecución de las políticas, información del personal y datos sobre procesos de contratación pública.

Uno de los principales aportes del artículo “Estado de la transparencia y acceso a la información pública en Ecuador pos-pandemia. ¿Han cambiado las necesidades ciudadanas?”, consiste en identificar implicaciones de política pública para garantizar la efectividad de la transparencia activa y pasiva del Estado. Dichas implicaciones no se limitan a cumplir con determinados estándares de transparencia, sino que también incluye medidas estatales que fomenten la capacidad de acceso y de procesamiento por parte de los usuarios de la información.

Finalmente, el quinto texto de este número especial corresponde a los “Lineamientos para el uso de inteligencia artificial en contextos universitarios” contribuido por Juan David Gutiérrez. En el preámbulo que precede a los lineamientos, el autor pone en contexto al lector mediante una breve explicación sobre la historia de su confección y su desarrollo posterior durante el 2023. Los lineamientos están divididos en siete secciones: (1) introducción, (2) objetivos, (3) reglas de uso dentro y fuera del aula, (4) justificación de las reglas, (5) otros recursos para profesores/as, (6) apertura a futuros cambios, y (7) cambios respecto de anteriores versiones de los lineamientos. Aunque los lineamientos son pertinentes para el uso de cualquier tipo de IA, sus reglas se enfocan en el uso de chatbots –tales como ChatGPT, Bing, Claude, y Bard, entre otros– que permiten a los estudiantes acceder a los modelos de lenguaje a gran escala. Los lineamientos son un recurso útil para todos aquellos que precisen de un documento fundamentado, comprensivo, y estandarizado que permita establecer unas reglas claras de trabajo dentro y fuera del aula con el fin de promover un uso informado, transparente, ético, y responsable de la IA.

Futuras agendas de investigación

Este número especial contribuye a abrir nuevas sendas en la literatura sobre la gobernanza de datos y de IA en el sector público Iberoamericano, pero aún quedan muchísimas venas de investigación por explorar y por madurar. Estas áreas se pueden agrupar en cuatro grandes temáticas: el gobierno de datos, el impacto de la IA en los procesos de política pública, la regulación algorítmica, y la economía política de la IA. De un lado, existe un vacío en la literatura Iberoamericana sobre las barreras concretas que enfrenta el sector público en el establecimiento de un gobierno de datos eficiente y equitativo. Además, se requiere un análisis profundo de los obstáculos institucionales y culturales que impactan la gestión adecuada de la información en esta región. Asimismo, es imperativo investigar los marcos de gobernanza de datos en diferentes jurisdicciones para extraer lecciones aprendidas y revisar las mejores prácticas y los errores que pueden evitarse.

Otra área de investigación fundamental está relacionada con el impacto de la IA en los procesos de formulación de políticas públicas sobre IA. Por una parte, hay una brecha en la literatura académica sobre cómo aportan y/o afectan los sistemas de IA los procesos de política pública. Por ejemplo, estamos en mora de evaluar de manera sistemática los tipos de sistemas usados en las diferentes etapas del ciclo de las políticas pública. En esa misma línea, es necesario investigar sobre los requerimientos de calidad de la información para asegurar su uso responsable en las políticas públicas y las dimensiones clave que aseguren la integridad y confiabilidad de los datos empleados en los modelos de IA en cada una de las jurisdicciones teniendo en cuenta sus particularidades.

En tercer lugar, en los últimos años se ha presentado una oleada de proyectos de ley y de propuestas de normas vinculantes en Iberoamérica que directa o indirectamente regulan estas tecnologías. Al respecto, por ejemplo, podría investigarse la emergencia de diferentes modelos regulatorios de IA en las jurisdicciones de la región, así como evaluar los potenciales impactos de los instrumentos regulatorios considerados. En ese sentido, por ejemplo, valdría la pena estudiar el nivel de protección de los derechos fundamentales que ofrecen estas regulaciones en el contexto de la recolección y explotación de datos públicos para salvaguardar la privacidad y seguridad de los ciudadanos. Dado estas tecnologías plantean tanto beneficios como riesgos en diferentes sectores, el aporte de la academia sobre estas materias sería muy importante para contribuir a informar la toma de decisiones en los procesos regulatorios en Iberoamérica.

Por último, existen espacios para contribuir al debate en cuanto a la identificación, análisis y evaluación de las barreras y limitaciones que afectan la planificación, diseño, contratación, implementación y evaluación de proyectos de IA en el sector público. Se insta a investigar soluciones efectivas que mejoren la eficacia de estos proyectos, así como casos de estudio sobre aplicaciones específicas de IA en las ramas ejecutivas, judiciales y legislativas para obtener una visión detallada de su impacto y efectividad. La economía política del uso de IA por parte de los gobiernos también es un área que debe ser explorada para comprender cómo los intereses y poderes económicos y políticos influyen en la adopción y aplicación de estas tecnologías. Todo lo anterior hace indispensable adoptar aproximaciones críticas al uso de IA por parte de los Estados, cuestionando sus implicaciones sociales, éticas y políticas para asegurar un desarrollo equitativo y justo de estas tecnologías en el ámbito público. 81

Las evaluaciones de impacto algorítmico: Una visión más allá de la ética

Flórez Rojas, María Lorena

Groningen University, Países Bajos

 m.l.florez.rojas@rug.nl

ORCID ID: [0000-0002-8264-8227](https://orcid.org/0000-0002-8264-8227)

Documento recibido:

02 septiembre 2022

Aprobado para publicación:

07 agosto 2023

Resumen

Los sistemas que utilizan la inteligencia artificial (IA) serán la norma del futuro. Sin embargo, la intervención humana es crucial a la hora de considerar las cuestiones éticas y legales basadas en los juicios de estos sistemas. La preocupación es sobre cómo usar, desarrollar e investigar la IA con un enfoque social a través de los análisis de impacto algorítmico (AIA). Esta metodología contribuye a una evaluación de los posibles riesgos asociados de la IA. Todas las organizaciones pueden aprovechar esta metodología para aumentar la confianza en las aplicaciones de IA. A pesar del esfuerzo por integrar estas estrategias de autorregulación, en algunas industrias, los principios éticos se quedan cortos en su aplicación. El documento analiza la metodología de AIA y del beneficio social a partir del análisis de un toolkit ético legal que se testeó para hacer un llamado a los reguladores a concretar reglas y estándares sobre las implicaciones de esta tecnología.

Palabras clave

Inteligencia Artificial, Análisis de Impacto Algorítmico, Ética; autorregulación

Resumo

Os sistemas que usam inteligência artificial (IA) serão a norma do futuro. No entanto, a intervenção humana é crucial ao considerar as questões éticas e legais com base nos julgamentos desses sistemas. A preocupação é como usar, desenvolver e pesquisar IA com foco social por meio da análise algorítmica de impacto (AIA). Esta metodologia contribui para uma avaliação dos riscos potenciais associados à IA. Todas as organizações podem aproveitar

essa metodologia para aumentar a confiança nos aplicativos de IA. Apesar do esforço para integrar essas estratégias de autorregulação, em algumas indústrias, os princípios éticos ficam aquém de sua aplicação. O documento analisa a metodologia AIA e o benefício social com base na análise de um kit de ferramentas éticas legais que foi testado para convocar os reguladores a especificar regras e padrões sobre as implicações dessa tecnologia.

Palavras-chave

Inteligência Artificial, Análise Algorítmica de Impacto, Ética; auto-regulação.

Abstract

Systems using artificial intelligence (AI) will be the norm of the future. However, human intervention is crucial when considering the ethical and legal issues based on the judgments of these systems. The concern is about how to use, develop and research AI with a social focus through algorithmic impact analysis (AIA). This methodology contributes to an assessment of the potential risks associated with AI. All organizations can take advantage of this methodology to increase confidence in AI applications. Despite the effort to integrate these self-regulatory strategies, in some industries, ethical principles fall short in their application. The document analyzes the AIA methodology and the social benefit based on the analysis of a legal ethical toolkit that was tested to call on regulators to specify rules and standards on the implications of this technology.

Keywords

Artificial Intelligence, Algorithmic Impact Analysis, Ethics; self-regulation.

1. Introducción

La robótica y la inteligencia artificial (IA) en el ámbito militar y de seguridad son el nuevo paradigma de la defensa nacional, donde el aumento de la intervención contempla los niveles estratégicos en la formación de los ejércitos y la policía del futuro. El uso de artefactos tecnológicos que pueden ser utilizados en un aspecto militar o policial no es ajeno a la discusión actual. Sin embargo, hay cuestiones importantes relacionadas con la ética de la tecnología, la investigación y la formación para preparar a militares, policía, desarrolladores, legisladores y en general a la sociedad para el uso masivo de estos artefactos (Borges, 2017).

A manera de ejemplo, los programas de predicción policial son utilizados por la policía para estimar dónde y cuándo es probable que se cometan los delitos o quién es probable que los cometa. Estos programas funcionan alimentados por datos policiales históricos a través de algoritmos informáticos (Dulce, Ramírez-Amaya and Riascos, 2018). Así, un programa puede evaluar los datos sobre delitos pasados para predecir dónde se cometerán en el futuro, identificando lo que se denominan como 'puntos calientes' en un mapa. Pero, los datos que utilizan estos programas pueden estar incompletos y/o sesgados, lo que conduce a un 'bucle de

retroalimentación', enviando a los agentes policiales a comunidades que ya están injustamente sobre vigiladas.

Uno de los primeros programas para este fin fue adoptado dentro del Departamento de Policía de Los Ángeles (LAPD por sus siglas en inglés), que trabajó con agencias federales en 2008 para explorar enfoques policiales predictivos. LASER hace referencia al '*Programa de extracción y restauración estratégica de Los Ángeles*', que identificaba las zonas en las que es probable que se produzcan actos de violencia con armas de fuego (Uchida and Swatt, 2013). Parte de la operación LASER consistió en crear una lista de 'delincuentes crónicos' que se clasificaban mediante un sistema de puntos. Esta lista se había elaborado a partir de la información recopilada por las diversas patrullas, las tarjetas de entrevistas, las citaciones de tráfico, los informes sobre delitos y los antecedentes penales. A partir de esta lista, el LAPD generaba una serie de boletines los cuales incluían información sobre el historial de detenciones de la persona, sus características físicas, su pertenencia a una pandilla, su situación de libertad condicional, sus órdenes de detención y sus contactos policiales recientes, entre otros datos. Estos boletines se distribuían dentro de los diferentes departamentos para informar a los agentes que patrullaban sobre los individuos que requerían más atención (Davis *et al.*, 2022, p. 187).

Sin embargo, este proyecto se detuvo en 2019 después de que el inspector general del LAPD publicara una auditoría interna la cual encontró problemas significativos con el programa, incluyendo inconsistencias en la forma en que los individuos fueron seleccionados y mantenidos en el sistema (Davis *et al.*, 2022, p. 189). Los documentos evidenciaron cómo los programas policiales basados en datos reforzaron patrones perjudiciales, alimentando la sobrevigilancia de las comunidades negras y latinas. Luego de este suceso, la LAPD ha intentado reestablecer la confianza de la ciudadanía con la policía, de un lado, reconociendo que las estrategias anteriores estaban centradas en la supresión proactiva y por tanto hacían que algunos barrios se sintieran excesivamente vigilados. De otro lado, la LAPD ha afirmado que será "*más transparente y sus procesos más estandarizados, trabajando para colaborar más estrechamente con los residentes locales*" (Bhuiyan, 2021) evidenciando la necesidad de crear confianza digital en estas herramientas a través de la implementación de principios como la transparencia y la auditoría.

Se puede afirmar entonces que los rápidos avances tecnológicos, los cambios en la opinión pública y la mayor atención de los responsables políticos exigen un esfuerzo serio para mitigar los riesgos y aumentar la confianza al desplegar las aplicaciones de IA. La mayor atención a la ética digital y el impacto que la IA podría tener en nuestra sociedad hacen que una evaluación sistemática de los intereses y preocupaciones de las partes interesadas sea esencial para el éxito de las aplicaciones de IA. Para ello, diversos organismos nacionales e internacionales han establecido una metodología sobre análisis de impacto, ya aplicada en otras esferas, como medio ambiente, privacidad y tecnología, pero enfocada en los riesgos de los algoritmos de IA denominada Análisis de Impacto Algorítmica (AIA) (Selbst, 2021; OECD, 2023; OPSI, 2023).

Las evaluaciones de impacto no son nada nuevo para muchas empresas. Pero, para gobernar adecuadamente estos diversos sistemas, las evaluaciones deben tener una estructura dinámica y permitir modificaciones que se adapten al entorno específico de una organización. Cualquier adopción de una tecnología al interior de una organización utiliza algún sistema de análisis de riesgos. Estos análisis difieren en sus métodos, pero no en su alcance. Por ello, y porque los derechos y valores que deben preservarse independientemente de la tecnología utilizada, el modelo que aquí se propone no es una evaluación tecnológica, sino un modelo basado en derechos y valores especialmente en el beneficio social de la IA. Esta metodología proporciona orientación para la

identificación de los aspectos legales y éticos relevantes de las aplicaciones de IA, tales como la confianza, la transparencia, la seguridad, entre otros aspectos. De esta forma, la implementación de esta metodología al interior de las organizaciones para promover los principios éticos de la IA es esencial para innovar de forma responsable (Moss *et al.*, 2020).

A medida que ha aumentado la conciencia de los avances y los riesgos de la IA, la respuesta dominante ha sido un giro hacia estrategias de autorregulación o *soft law* en donde predominan los principios éticos. Así, alrededor del mundo se han publicado cientos de ellos por organizaciones privadas como Google e IBM, hasta diversos gobiernos como el de Canadá o Colombia. A medida que las organizaciones adoptan la IA con mayor rapidez, la implementación de estos principios se considera una de las mejores formas de garantizar que la IA no cause daños involuntarios. Así, podría decirse que las consideraciones éticas nunca han sido tan importantes como ahora (Baum, 2020).

Las cuestiones relativas al control de las posibles consecuencias imprevistas de los futuros avances de la IA han sido consideradas por diversos organismos nacionales e internacionales hasta la fecha. Por ejemplo, en 2018 la Comisión Europea publicó una propuesta de marco regulador. En ella, la Comisión sugería categorizar las aplicaciones de IA en función de su potencial de riesgo. Por ejemplo, un filtro de spam de correo electrónico estaría menos regulado que una herramienta de IA para la atención sanitaria, con ciertos casos de uso completamente prohibidos. En la misma línea, las recomendaciones del Consejo de la OCDE sobre IA, adoptada en mayo de 2019, proporciona un conjunto de principios para el desarrollo responsable y el despliegue de IA digna de confianza (OECD, 2019). Asimismo, a finales de 2021, y adoptado por 193 países, la UNESCO emitió un marco con el objetivo de guiar la construcción de la infraestructura legal necesaria para garantizar el desarrollo ético de esta tecnología, otorgando a los Estados la responsabilidad de aplicarlo a su nivel (UNESCO, 2021).

Sin embargo, se ha dificultado poner en práctica estos principios que se traducen en niveles de riesgo a través del AIA. Fruto de lo anterior, existe una brecha cada vez más marcada entre los principios de alto nivel y la práctica tecnológica. Incluso cuando se reconoce esta laguna y se intentan materializar los principios, traducir conceptos sociales complejos a conjuntos de reglas técnicas resulta difícil. En este sentido este documento busca responder a una pregunta, *¿En qué medida debería priorizarse la aplicación de la metodología del Análisis de Impacto Algorítmico (AIA) como paso inicial hacia el establecimiento de marcos regulatorios sólidos para la inteligencia artificial (IA)?*

Para responder a esta pregunta, este estudio emplea en primer lugar una metodología deductiva a través de la implementación del *Toolkit ético-legal* que fue diseñado y testeado para determinar el impacto ético-legal de los proyectos de IA (*TELia*), teniendo en cuenta los principios éticos establecidos en el Marco ético de IA para Colombia así como su implementación en las organizaciones (Florez Rojas, 2021). Es relevante mencionar, que el *TELia* es fruto de la investigación financiada por el Fondo de Apoyo para Profesores Asistentes (FAPA) de la Universidad de los Andes¹. Además, esta investigación emplea la investigación dogmática legal, en la medida en que esta permite la revisión de doctrina, principios, normas y conceptos alrededor de la implementación de los marcos éticos en el país a través de la metodología de AIA y cómo estas directrices éticas

¹ Se reconoce y agradece la retroalimentación al *Toolkit ético-legal* por parte de actores del ecosistema como Juan David Gutiérrez, María Paula Ángel, Fernando Enrique Lozano, David Niño, Ana Rey y Andrés Rodríguez.

se quedan cortas a la hora de promover un beneficio social. La combinación de estos dos elementos permite un examen exhaustivo de las implicaciones éticas y jurídicas de los sistemas de IA, centrándose en maximizar el beneficio social. Al combinar la doctrina jurídica del beneficio social y la puesta a prueba de un conjunto de herramientas éticas y jurídicas del TELia, la metodología de investigación permite un análisis multidimensional de las consideraciones éticas y jurídicas en torno a la IA.

Así, primero, se describe la metodología de AIA con sus beneficios y se describe el proceso de diseño y pruebas que se realizó con el TELia. Luego, se hace una aproximación sobre las limitaciones que esta metodología tiene frente a su implementación, seguimiento, ambigüedad en las definiciones y prácticas de *Ethics washing*. Seguidamente, se presenta un análisis sobre el deber social de regulación más allá de principios. Finalmente, se presentan algunas recomendaciones y conclusiones de la investigación que destacan algunos enfoques alternativos de la justicia de AI que van más allá de los principios fundamentales haciendo un llamado a la regulación y auditoría.

2. Análisis de Impacto Algorítmico (AIA) y estudio de campo

Los AIA se pueden definir como prácticas emergentes para delimitar la responsabilidad, hacer visibles los daños causados por los sistemas algorítmicos y garantizar que se tomen medidas prácticas para mejorar esos daños (Kaminski and Malgieri, 2020, p. 70). En otras palabras, los AIA son considerados como una herramienta para evaluar las posibles repercusiones sociales de un sistema algorítmico antes de que se empiece a utilizar. Actualmente no existe una sola fórmula de implementación para estas evaluaciones, toda vez que diversos investigadores, responsables políticos y desarrolladores han propuesto varios modelos dependiendo de factores como el área de investigación (v.g. salud, sistema judicial, robótica), la jurisdicción y el tipo de organización proponente (v.g. organizaciones internacionales, industrias, organizaciones de la sociedad civil, academia) (Selbst, 2021, p. 139).

Entre las diferentes propuestas se incluyen aquellas inspiradas directamente en el Marco de Gobernanza y Rendición de cuentas de la Unión Europea (UE), el marco general de protección de datos (RGPD) e incluso el AIA de Canadá. Con base en lo anterior, algunos investigadores señalan que los AIA pueden considerarse como extensiones de las evaluaciones de impacto sobre la privacidad (EIPD), que se utilizan comúnmente para abordar las preocupaciones sobre el adecuado uso de los datos y para cumplir con la reglamentación en esta materia (Decker and Ladikas, 2004, p. 4). Así, en principio estas evaluaciones buscan estimar los *beneficios potenciales*, los riesgos latentes, y los procesos de seguimiento y mitigación.

De esta forma, el término AIA se ha utilizado para referirse a una serie de procesos y documentos que surgen en el contexto de una caja de herramientas en expansión de los métodos de responsabilidad potencial, incluyendo auditorías de algoritmos, hojas de datos, etiquetas, entre otros (Moss *et al.*, 2021, p. 14). En esta línea, los AIA se han propuesto como un método para aumentar la responsabilidad en el diseño y despliegue de los sistemas de IA, que tienen el potencial de aumentar la confianza pública en su uso, reducir el potencial de daño a individuos y grupos, y maximizar su potencial de beneficio (Mantelero, 2018, p. 758). Con ello, la idea general de un AIA es documentar el desarrollo y el impacto de un sistema algorítmico, proporcionando un punto de ventaja para mitigar los posibles daños a las personas y a las comunidades vulnerables.

2.1. Beneficios de los AIA en las organizaciones

Los AIA buscan analizar cómo está diseñado un sistema algorítmico, cómo funciona y las posibles consecuencias de sus resultados o decisiones. Este proceso suele implicar un enfoque multidisciplinar que combina la experiencia técnica con el conocimiento de cuestiones sociales y éticas (Ada Lovelace Institute, 2022, p. 14). Puede incluir una serie de actividades, como la participación de las partes interesadas, evaluaciones y clasificación de los riesgos, así como opciones de seguimiento y auditoría. Así, los AIA son cada vez más importante a medida que los procesos de toma de decisiones se automatizan y se llevan a cabo mediante algoritmos.

A pesar de que actualmente el uso e implementación de los AIA no se encuentran como obligaciones legales, ni para el sector público ni en el privado, y no existe una metodología única y aceptada para su utilización, es importante señalar los principales beneficios de llevar a cabo un AIA. Lo anterior, con el fin de revisar si su adopción en estricto sentido puede ayudar a mitigar tempranamente riesgos. Así, dentro de los principales beneficios de implementar un AIA se encuentran i) identificación temprana de sesgos; ii) mejora la transparencia y rendición de cuentas; y iii) procura por la salvaguarda de los derechos humanos.

En primer lugar, los AIA ayudan a descubrir sesgos y prácticas discriminatorias integradas en los sistemas algorítmicos que las organizaciones buscan implementar (evaluación *ex ante*) o que ya hayan implementado, pero buscan mejorar (evaluación *ex post*) (Ada Lovelace Institute, 2022, p. 15). En este sentido, les permite a organizaciones de todos los sectores evaluar si las tecnologías de IA perpetúan los prejuicios existentes o discriminan injustamente a determinados grupos. A manera de ejemplo, podemos analizar el caso de *Beauty.AI*, que lanzó su concurso de belleza en su primera versión en 2016, financiado por el grupo Youth Laboratories, que utilizaba algoritmos de reconocimiento facial para elegir a los que sus creadores declararon que serían "la primera reina o rey de la belleza juzgados por robots" (Youth Laboratories, 2016). Sin embargo, los resultados fueron criticados por sus sesgos, sobre todo en contra de las personas de tez oscura. Debido a estas críticas, Youth Laboratories volvió a incursionar con un *think tank* denominado *Diversity.AI*, y otro concurso de belleza en su tercera versión. Así, según Georgievskaya, co-fundadora, "esta nueva versión busca incluir un paso que pondrá a prueba los algoritmos para detectar sesgos, especialmente raciales, y todos los demás tipos de sesgos como el género y la edad" (Manthorpe, 2017). En este caso, si la organización hubiera implementado medidas *ex ante* para mitigar los riesgos relacionados con la diversidad y la clasificación de los datos como resultado de un AIA podría haber generado una mejor receptividad por parte de los interesados y evitar pérdidas financieras que se reflejan indirectamente cuando un proyecto no es bien recibido en la esfera pública. Así, al identificar estos problemas, los AIA permiten aplicar los ajustes necesarios y las estrategias de mitigación, lo que conduce a sistemas de IA más equitativos e inclusivos (Benjamins, Barbado and Sierra, 2019, p. 4).

En segundo lugar, los AIA promueven la transparencia y la responsabilidad en el desarrollo y despliegue de la IA generando a su vez confianza en el uso de herramientas de IA y hacia la organización (Moss *et al.*, 2021, p. 27). Sobre el particular, existe un amplio consenso en cuanto a que el acceso público a los AIA y la transparencia en la práctica son objetivos clave para aumentar la apertura y, en consecuencia, la confianza pública en el uso de los sistemas de IA (Reisman *et al.*, 2018, p. 9). Así, los AIA pueden fomentar una cultura de responsabilidad y confianza en la tecnología, toda vez que llevan a desarrolladores, responsables políticos y reguladores a responsabilizarse de las repercusiones sociales de la IA.

En tercer lugar, los AIA sirven como mecanismo para evaluar los daños potenciales que los sistemas de IA pueden suponer para los derechos y el bienestar de las personas. Por ejemplo, pueden ayudar a identificar los riesgos relacionados con la invasión de la privacidad, la discriminación y otros problemas éticos (Kaminski and Malgieri, 2020, p. 72). De esta forma, los AIA ofrecen una oportunidad para abordar de forma proactiva estos riesgos y aplicar salvaguardias necesarias para proteger o mitigar el impacto en los derechos humanos. Así, al llevar a cabo una evaluación exhaustiva de los impactos algorítmicos, los AIA contribuyen al desarrollo de políticas y directrices que priorizan los beneficios sociales al tiempo que minimizan las consecuencias negativas de las tecnologías de IA (Novelli, Taddeo and Floridi, 2023, p. 10).

Con todo lo anterior, se resalta la naturaleza compleja y polifacética de la ética de la IA y la necesidad de un enfoque multidisciplinar, pluralista y orientado a los derechos humanos para abordarla. Con miras a contrarrestar estos beneficios y ponerlos en práctica, es importante implicar a las partes interesadas en el proceso de gobernanza y supervisión de la IA, y garantizar la transparencia y la rendición de cuentas en el desarrollo y despliegue de los sistemas de IA (Flórez Rojas, 2023). En resumen, los AIA aportan diversos beneficios significativos al campo de la ética y la gobernanza de la IA.

2.2. Toolkit ético-legal: Diseño y pruebas

El *TELia* fue diseñado y desarrollado como una guía de autodiagnóstico específicamente adaptada para la evaluación ética y legal de proyectos que incorporan tecnologías disruptivas, con especial atención a la IA. Esta herramienta pretende hacer una valiosa contribución en el ámbito de los AIA, proporcionando una hoja de ruta completa para los diversos sectores que buscan adoptar la IA como una solución para hacer frente a necesidades específicas en Colombia. En esencia, el *TELia* sirve como lista de comprobación para el autodiagnóstico, facilitando la evaluación ética y jurídica a las organizaciones que planean incorporar la IA a sus proyectos (Flórez Rojas *et al.*, 2022, p. 4).

Los AIA, en su mayoría, están diseñados para alcanzar cuatro objetivos principales. En primer lugar, buscan **identificar o capturar el riesgo de un sistema de IA** a través de criterios de control de riesgos, los cuales le permiten a las organizaciones clasificar adecuadamente el nivel de escrutinio necesario para una aplicación específica de IA (Moss *et al.*, 2021, p. 16). Por ejemplo, es probable que un proyecto de automatización del flujo de trabajo genere menos riesgos potenciales que el uso de datos biométricos para el reconocimiento facial dentro de un sistema de predicción policial. Dentro del diseño del *TELia* en su primera iteración se usaron criterios guiados por los principios éticos de Colombia como inclusión, seguridad de los datos, representación, gobernanza, resiliencia transparencia y auditoria y rendición de cuentas (Guío Español, 2021; Gobierno de Colombia, 2022). Esta evaluación busca que los diversos actores involucrados en el uso, diseño e implementación de la IA pudieran analizar estos principios desde su experiencia. Sin embargo, al tratarse de una serie de herramientas que se adoptan de forma voluntaria y autónoma por diversas entidades, no se pretende que estas herramientas solucionen los sesgos o riesgos que los diversos proyectos puedan llegar a tener, sino por el contrario, se busca brindar una serie de puntos clave para tener en cuenta y crear alertas tempranas, así como sugerencias sobre el relacionamiento con los diversos actores que se involucran en el proyecto como usuarios finales, desarrolladores, tomadores de decisiones empresariales, equipo legal entre otros.

En segundo lugar, los AIA buscan cubrir todos los **requisitos del ciclo de vida del desarrollo** (Moss *et al.*, 2021, p. 42). En este sentido, los AIA procuran por identificar riesgos en cualquier etapa del ciclo desde su estrategia y la planificación, pasando por las cuestiones relacionadas con los datos de formación, para llegar finalmente al despliegue (Selbst, 2021, p. 148). Así el *TELiA* plantea tres áreas principales para el estudio detallado: i) Desarrollo y Pruebas; ii) Ético-Legal; y iii) Organizacional. De esta forma, cada área tiene algún tipo de relación con el proyecto, pero, lo más importante es que cada área en realidad debe estar capacitada para señalar algún tipo de riesgo o posibilidad de riesgo ético-legal. Sin embargo, las áreas no constituyen conceptos rígidos preestablecidos, sino que ofrecen una orientación para cada equipo al interior de una organización con el fin de interactuar entre los diferentes actores para crear una verdadera gobernanza de la IA.

En tercer lugar, los AIA buscan **aumentar la responsabilidad a través de un análisis de múltiples partes interesadas**. Una evaluación de impacto exitosa involucra a diversas partes interesadas internas y también puede incluir a representantes externos, como los comités de ética o de revisión de datos (Mintrom and Luetjens, 2016, p. 393). Estos compromisos, que reconocen la compleja influencia que puede tener un sistema de IA, deben generar debates sobre el impacto tanto empresarial como *social* (Castets-Renard, 2021, p. 10). También pretenden abordar la probabilidad de ocurrencia tanto de los beneficios como de los riesgos, así como, la eficacia de los controles diseñados para mitigar riesgos. De esta forma, el *TELiA* recrea un escenario de puntajes para cada una de las preguntas con el fin de evidenciar en qué aspectos el proyecto puede tener algún riesgo ético o legal y retroalimentar el proyecto. El puntaje máximo del diagnóstico preliminar es de 45 puntos. Algunas preguntas no tienen ningún puntaje toda vez que se tratan de preguntas sobre el objeto del proyecto o sobre la proveniencia de los datos lo cual no puede clasificarse entre recomendable o no recomendable en términos de los principios escogidos.

Finalmente, en cuarto lugar, los AIA buscan **facilitar la toma de decisiones** (Moss *et al.*, 2021, p. 42). Este objetivo debe abordar la cuestión de si un modelo debe pasar a producción, determinar si está listo para la transición a las operaciones habituales y decidir si debe continuar como está o ser reorientado, rediseñado o retirado. Aquí una organización puede determinar la auditabilidad del sistema, los requisitos legales adicionales y la eficacia de los controles, así como confirmar que ha logrado un equilibrio adecuado de beneficios y riesgos mitigados (Reisman *et al.*, 2018, p. 19). De esta forma, el *TELiA* entrega un primer diagnóstico sobre los elementos que la organización debe mejorar, dependiendo de una de las 3 tres áreas analizadas. Asimismo, entrega por correo electrónico una serie de recomendaciones generales para proyectos de IA dependiendo del objetivo de este y con base en las respuestas entregadas en el aplicativo. Finalmente, si la organización lo considera necesario se puede hacer una sesión de diagnóstico más especializada.

Teniendo en cuenta los objetivos de los AIA, se testeó el *TELiA* con algunos proyectos universitarios que involucran el uso de IA. El primer proyecto, consistía en analizar datos provenientes del extranjero para identificar una enfermedad en edad temprana que usualmente es más común en niñas que en niños, pero que su detección temprana puede llegar a ser beneficiosa a largo plazo. Sobre este proyecto, lo primero que se identificó fueron los actores involucrados en el desarrollo de este — *estudiantes de la Universidad* —. Con ayuda de algunas preguntas que hacen parte del *TELiA*, ellos mismos identificaron que podrían existir más sujetos involucrados como, por ejemplo, la empresa proveedora de los datos de entrada que usaron, entre otros. Así, con este caso, se puede vislumbrar que el lleno de las preguntas orientadoras brinda tan solo una noción sobre el proyecto, pero no significa esto que las mismas sirvan para modificar el curso del proyecto o incluso cambiar las

prácticas de la industria. Con todo lo anterior, se puede afirmar que los AIA pueden ser un primer paso para identificar los vacíos o falencias de los proyectos de IA a través del escaneo de principios éticos.

3. Limitaciones de la metodología de AIA

Es importante resaltar que los AIA en sí mismos no hacen nada para mitigar los problemas identificados, aunque algunos procesos de evaluación requieren que se documente explícitamente la mitigación de los impactos. Al igual que ocurre con otras formas de evaluación de impacto, cualquier AIA probablemente detectará los daños que deben ser eliminados, pero el proceso en su conjunto debe poder facilitar una toma de decisiones sólida, comprometida y transparente sobre las compensaciones entre los daños potenciales y los beneficios probables (Selbst, 2021). De esta forma, se pueden identificar cuatro principales limitaciones a esta metodología.

En primer lugar, actualmente los AIA no se utilizan de forma generalizada ni en el sector público ni en el privado y no existe una norma única aceptada o una metodología única para su uso. Aunque cada vez hay más consenso sobre la importancia de los principios para el desarrollo y el uso de sistemas de IA, como la responsabilidad, la transparencia y la equidad, las prioridades individuales y la interpretación organizativa de estos términos difieren. Sumado a lo anterior, la falta de coherencia con estos conceptos hace que no todos los AIA estén diseñados para lograr los mismos fines, y el proceso de los AIA dependerá del contexto específico en el que se lleven a cabo (Munn, 2022, p. 3). Para el caso del *TELiA* se entendió que una de las limitaciones más relevantes era el entendimiento conjunto de los principios éticos establecidos en el país para su adaptación en forma de pregunta con miras a determinar un puntaje. Así, las metodologías de AIA carecen de marcos y directrices normalizados, lo que da lugar a variaciones en las prácticas y los resultados de la evaluación, toda vez que, la ausencia de un enfoque unificado puede dificultar la comparabilidad y la coherencia entre las distintas evaluaciones.

En segundo lugar, los AIA son sólo un sustituto de los daños reales que pueden sufrir las personas, es decir, son la previsión hipotética de que pueden ocurrir o no — *riesgos previsibles o imprevisibles* — (Selbst, 2021, p. 130). Describir los daños reales es difícil porque es una verdad que, como muchos aspectos sociales y psicológicos de la humanidad, es difícil de evaluar y representar con palabras, y mucho menos de cuantificar (NIST, 2022, p. 8). Por ejemplo, los diversos esfuerzos por crear AIA buscan medir hasta qué punto su modelo se desvía de su punto de referencia de equidad/representación/seguridad, sin embargo, esas métricas pueden estar basadas en políticas de empresa, sin tener en cuenta el daño emocional o psicológico causado a los individuos, y mucho menos el daño social a comunidades enteras, cuando el algoritmo los reconoce erróneamente en repetidas ocasiones — *no maleficencia de la IA* — (Leikas, Koivisto and Gotcheva, 2019, p. 5). Así, la posibilidad de que las organizaciones confíen demasiado en estos autodiagnósticos puede a su vez conducir a que se limiten a cumplir los requisitos de conformidad en lugar de abordar realmente los desafíos éticos.

En tercer lugar, existe ambigüedad sobre cómo se definen los impactos — *y por quién* —, cómo se evalúan — *y por quién* — y cómo se establecen o no formas sólidas de responsabilidad. A manera de ejemplo, cada organización que lanza una guía de este tipo de análisis tiene intrínseca una visión de su alcance, por ejemplo, en el caso del *TELiA*, la visión es netamente académica y pretende en principio abordar proyectos universitarios. Sin embargo, no está exento de tener ciertas restricciones dependiendo de sus diseñadores y de la entidad financiadora. De esta forma, se puede afirmar que los AIA son construcciones sociales que delimitan ciertos siste-

mas como capaces de causar impactos y estabilizan conceptos abstractos, como los principios de la IA (Metcalf *et al.*, 2021, p. 737). Así, reconocer el carácter socialmente construido de los AIA en una fase temprana del desarrollo ofrece la posibilidad de realizar intervenciones que surjan de una deliberación cuidadosa y políticamente comprometida. Como consecuencia de la complejidad inherente a los sistemas de IA, sus aplicaciones y las consideraciones éticas y sociales implicadas en cada contexto, la colaboración eficaz entre expertos de distintos ámbitos puede verse obstaculizada, ya que pueden tener prioridades, metodologías e interpretaciones de los principios éticos diferentes.

Finalmente, algunas organizaciones se limitan a adoptar AIA sin implementar los mecanismos y herramientas necesarios para cumplir con estos compromisos. A esta práctica se le conoce como *Ethics washing*, en donde las empresas intentan crear una impresión de preocupación por los principios éticos sin un compromiso genuino (Rességuier and Rodrigues, 2020, p. 2). En esta línea, las organizaciones realizan gestos éticos superficiales o iniciativas simbólicas para mejorar su imagen pública sin realizar cambios sustanciales en sus prácticas reales. Por ejemplo, pueden publicar directrices éticas o declaraciones de misión sin aplicar mecanismos eficaces para garantizar una conducta ética en toda la organización (Floridi, 2019, p. 186). Así, es crucial mirar más allá de los gestos superficiales y examinar las acciones sustantivas, la transparencia y los mecanismos de rendición de cuentas existentes para garantizar que los principios éticos estén realmente integrados en los valores y prácticas fundamentales de las organizaciones.

Estas limitaciones ponen de relieve la necesidad de una investigación continua, de esfuerzos de normalización y de una mejora continua de las metodologías de AIA. Para superar estos retos se requiere una colaboración interdisciplinaria, el desarrollo de marcos normalizados, mecanismos de seguimiento sólidos y procesos de auditoría independientes que garanticen que los AIA contribuyen efectivamente al desarrollo y despliegue éticos y responsables de la IA. Sin embargo, ni los AIA ni los principios éticos tienen la fuerza vinculante para sancionar empresas o estados por el indebido uso de estas tecnologías. Esto se debe a que la ética se está utilizando en lugar de la regulación (Mittelstadt, 2019, p. 8). Así, es demasiado pronto para saber hasta qué punto serán comunes los AIA o hasta qué punto serán eficaces para cambiar la forma de la tecnología basada en algoritmos.

4. Más allá de la ética: Implementación regulatoria de los AIA

Los algoritmos de clasificación y predicción ya han demostrado causar daños en el mundo real a los más desfavorecidos, ya sea en el contexto de la justicia penal o de la predicción policial (Bhuiyan, 2021). Los AIA son una gran intervención inmediata, pero sin medidas sólidas de rendición de cuentas, podrían quedarse cortos frente a la dignidad humana y el beneficio social de la tecnología. Tal y cómo se discutió líneas arriba, en ocasiones, se emplean los AIA para asegurar la confianza de usuarios y consumidores, sin embargo, se tiende a confundir la funcionalidad de los marcos éticos con la de los marcos legales (Rességuier and Rodrigues, 2020, p. 3). La ética por sí misma no cuenta con los medios para exigir su cumplimiento de la misma manera en la que el derecho lo hace. De esta forma, la ética no debería implementarse con fines regulatorios, sino de manera analítica y preventiva.

En el mismo sentido, entender cómo se han estructurado las relaciones entre las organizaciones mediante otros tipos de procesos de evaluación de impacto y documentación es fundamental para deliberar sobre los AIA. El éxito en su implementación depende de la gobernanza de IA, entendida como el proceso de gobierno

(corporativo) en el que las reglas y cursos de acción se estructuran, mantienen y regulan, y se asigna la responsabilidad (Young, Bullock and Lecy, 2019). En este sentido, dentro de los procesos de gobernanza de la IA podemos encontrar la implementación de los AIA que tendrán a su vez embebidos los principios éticos de los que se ha discutido.

Al diseñar un AIA, la forma en que se estructuran las relaciones entre las organizaciones es un punto importante de ventaja que debe ser objeto de deliberación antes de formalizar la evaluación del impacto como requisito normativo (Gobierno de Canadá, 2021). El proceso de desarrollo tecnológico existente ya suele incluir documentación, y añadir especificaciones relacionadas con la evaluación de impacto al proceso de documentación existente podría ser mínimamente perturbador, aunque podría diferir entre las nuevas empresas y las organizaciones más maduras. Así, los AIA deben emplearse como herramientas de diagnóstico del contexto en el que se crea, desarrolla o implementa la IA para descubrir puntos ciegos, ampliar el margen de maniobra, y promover la libertad y la autorresponsabilidad (Moss et al., 2021, p. 5). Sin embargo, la adopción de la ética a través de los AIA no puede limitarse a la medición del nivel de riesgo, sino que se requiere de atención constante para detectar riesgos ocultos e incuantificables, y estos indicadores deben apoyarse otras herramientas, como en mecanismos de responsabilidad, documentación, e inventarios de todos los modelos de forma general a todos los actores de la IA.

En suma, los AIA hacen parte de un reto metodológico multifacético por la búsqueda de poner en práctica los principios éticos. Toda vez que la traducción en acciones u omisiones reales implica la especificación de los principios de alto nivel en normas vinculantes a todos los actores de la IA (Munn, 2022, p. 153). Así, Los consensos de alto nivel son alentadores, pero tienen poca relación con la justificación de las normas y los requisitos prácticos propuestos en contextos de uso específicos. Un ejemplo de lo anterior es la adaptación en Canadá sobre la regulación de la toma de decisiones automatizada la cual por ley debe estar precedida de un AIA (CIFRA, 2020). Así, se puede establecer que el diseño e implementación de los AIA que están basados en los principios éticos pueden establecer ideales normativos, pero carecen de mecanismos para hacer cumplir estos valores y principios. Toda vez que los principios no son autoejecutables, y no hay sanciones tangibles cuando son violados o amenazados.

5. Conclusiones

Tras considerar de un lado la creación y testeo inicial del *TELiA*, así como de otro lado los beneficios y las limitaciones de los AIA, puede concluirse que debería darse prioridad a la aplicación de esta metodología como paso inicial hacia el establecimiento de marcos reguladores sólidos para la IA. Sin embargo, si se pretende incentivar la innovación tecnológica protegiendo a su vez los derechos humanos y salvaguardando 'el bien' de la sociedad, se debe identificar en primera medida cuál es el significado cultural y social sobre el "bien" y encontrar la manera de hacer que las múltiples partes interesadas lo persigan, en lugar de perseguir un beneficio individual.


Empero, el grado en que debe darse prioridad a los AIA depende de varios factores. En primer lugar, los beneficios de los AIA son significativos pues ayudan a identificar y abordar los sesgos, la discriminación y los daños potenciales en los sistemas de IA. Al llevar a cabo una evaluación exhaustiva de los impactos algorítmicos, los AIA garantizan que las tecnologías de IA se alineen con los principios éticos y sociales, promoviendo la equidad, la transparencia y la responsabilidad. Así, a través de la aplicación del *TELiA* se busca que las diversas

organizaciones tengan una herramienta gratuita y de acceso abierto para hacer un autodiagnóstico de su proyecto identificando puntos de inflexión para trabajar o mejorar teniendo en cuenta el contexto limitado que abarca esta herramienta. De esta forma, dar prioridad a los AIA como paso inicial permite tomar medidas proactivas para minimizar las consecuencias negativas y salvaguardar los derechos humanos.

En segundo lugar, se deben reconocer las diversas limitaciones de los AIA, toda vez que son procesos complejos que requieren experiencia, recursos y cooperación entre las diversas partes interesadas. La eficacia de los AIA puede verse influida por las dificultades de acceso a los datos, la transparencia algorítmica y la naturaleza dinámica de las tecnologías de IA. A pesar de estas limitaciones, aplicar los AIA como paso inicial brinda la oportunidad de abordar estos retos, testear la metodología y desarrollar directrices para su aplicación.

En tercer lugar, a medida que las tecnologías de IA siguen avanzando y aumenta su impacto social, se hace cada vez más evidente la necesidad de marcos regulatorios sólidos. A pesar de que no todos los países han adoptado algún instrumento regulatorio para la implementación de la IA y en su mayoría han decidido adoptar una especie de *soft law*, los avances en la implementación de AIA son más generalizados a nivel mundial como una herramienta necesaria y previa. Así, los AIA sirven como base crucial para estos marcos éticos, ya que ayuda a comprender y mitigar los riesgos algorítmicos. Al dar prioridad a los AIA, los reguladores pueden recopilar información y pruebas valiosas para informar el desarrollo de reglamentos eficaces que rijan el desarrollo, despliegue y uso de los sistemas de IA.

Por último, es necesario un enfoque híbrido que adopte los AIA como un componente complementario de marcos regulatorios más amplios. Los AIA por sí solos no pueden abordar todas las consideraciones éticas y jurídicas relacionadas con la IA. Por lo tanto, es importante integrar los AIA con otros mecanismos reguladores, como por ejemplo una armonización con las leyes de protección de datos nacionales y los reglamentos específicos del sector. Este aspecto queda representado en el concepto descriptivo del *TELia* que procura por presentar recomendaciones focalizadas y complementarias a la ley general de protección de datos de Colombia, así como algunas recomendaciones en donde ya se han hecho algunos pronunciamientos sobre el uso de IA como en el sector justicia y el sector financiero. Este enfoque holístico garantiza que se maximicen los beneficios de la AIA y que los marcos reguladores cubran los diversos aspectos de la gobernanza de la IA.

En conclusión, la aplicación de la metodología del AIA debe considerarse como paso inicial hacia el establecimiento de marcos regulatorios sólidos para la IA, y que debe tener en cuenta los diversos contextos culturales y sociales en donde se busquen aplicar. Así, los AIA se pueden identificar como una forma emergente de responsabilidad para las organizaciones que construyen y despliegan sistemas de IA. Esta metodología apoya a las organizaciones para entender los impactos como objetos que son relaciones de responsabilidad co-construidas. De esta forma, un enfoque equilibrado, que integre la AIA con otros mecanismos reguladores, garantiza un marco de gobernanza de la IA holístico e integrador. 

Referencias

- Ada Lovelace Institute (2022) Algorithmic impact assessment: a case study in healthcare. Available at: <https://www.adalovelaceinstitute.org/resource/aia-user-guide/> (Accessed: 17 May 2023).
- Baum, S. D. (2020) 'Social choice ethics in artificial intelligence', *AI and Society*. Springer, 35(1), pp. 165–176. doi: 10.1007/s00146-017-0760-1.
- Benjamins, R., Barbado, A. and Sierra, D. (2019) 'Responsible AI by Design in Practice', in *Proceedings of the Human-Centered AI: Trustworthiness of AI Models & Data (HAI) track at AAAI Fall Symposium.*, p. 10. Available at: <https://arxiv.org/abs/1909.12838> (Accessed: 23 August 2021).
- Bhuiyan, J. (2021) 'LAPD ended predictive policing programs amid public outcry. A new effort shares many of their flaws', *The Guardian*. Available at: <https://www.theguardian.com/us-news/2021/nov/07/lapd-predictive-policing-surveillance-reform> (Accessed: 2 September 2022).
- Borges, J. V. (2017) 'Robots and the military: A strategic view', *Intelligent Systems, Control and Automation: Science and Engineering*. Kluwer Academic Publishers, 84, pp. 199–205. doi: 10.1007/978-3-319-46667-5_15/COVER.
- Castets-Renard, C. (2021) 'Human Rights and Algorithmic Impact Assessment for Predictive Policing', Cambridge University Press, pp. 93–110. Available at: <https://papers.ssrn.com/abstract=3890283> (Accessed: 1 September 2022).
- CIFRA (2020) AICan 2020 CIFAR Pan-Canadian AI Strategy Impact Report. Available at: <https://cifar.ca/wp-content/uploads/2020/11/AICan-2020-CIFAR-Pan-Canadian-AI-Strategy-Impact-Report.pdf>.
- Davis, J. et al. (2022) 'Five ethical challenges facing data-driven policing', *AI and Ethics* 2022 2:1. Springer, 2(1), pp. 185–198. doi: 10.1007/S43681-021-00105-9.
- Decker, M. and Ladikas, M. (2004) 'Technology Assessment in Europe; between Method and Impact — The TAMI Project', in *Bridges between Science, Society and Policy*. Springer Berlin Heidelberg, pp. 1–10. doi: 10.1007/978-3-662-06171-8_1.
- Dulce, M., Ramírez-Amaya, S. and Riascos, Á. (2018) 'Efficient allocation of law enforcement resources using predictive police patrolling', in *NIPS 2018 Workshop on Machine Learning for the Developing World*. doi: 10.48550/arxiv.1811.12880.
- Florez Rojas, M. L. (2021) *Toolkit Ético Legal*, GECTI Uniandes. Available at: https://gecti.uniandes.edu.co/qsm_quiz/toolkit-etico-legal-es/ (Accessed: 2 September 2022).
- Flórez Rojas, M. L. et al. (2022) *La adopción de tecnologías disruptivas en las organizaciones a partir de la creación e implementación de un toolkit ético-legal*. Bogotá. doi: 10.1007/s11948-019-00146-8.
- Flórez Rojas, M. L. (2023) 'Pensamiento de diseño y marcos éticos para la Inteligencia Artificial: una mirada a la participación de las múltiples partes interesadas', *Desafíos*. Colegio Mayor de Nuestra Señora del Rosario, 35(1). doi: 10.12804/REVISTAS.UROSARIO.EDU.CO/DESAFIOS/A.12183.

- Floridi, L. (2019) 'Translating principles into practices of digital ethics: five risks of being unethical', *Philos. Technol. Springer Netherlands*, 32(2), pp. 185–193. doi: 10.1007/s13347-019-00354-x.
- Gobierno de Canadá (2021) Responsible use of artificial intelligence (AI). Available at: <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai.html#toc1> (Accessed: 25 August 2021).
- Gobierno de Colombia (2022) Marco ético para la Inteligencia Artificial. Available at: <https://iaeticacolombia.gov.co/> (Accessed: 25 August 2021).
- Guío Español, A. (2021) Consejo Internacional de Inteligencia Artificial para Colombia. Available at: <https://dapre.presidencia.gov.co/TD/CONSEJO-INTERNACIONAL-INTELIGENCIA-ARTIFICIAL-COLOMBIA.pdf>.
- Kaminski, M. E. and Malgieri, G. (2020) 'Multi-layered Explanations from Algorithmic Impact Assessments in the GDPR', *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. New York, NY, USA: ACM. doi: 10.1145/3351095.
- Leikas, J., Koivisto, R. and Gotcheva, N. (2019) 'Ethical Framework for Designing Autonomous Intelligent Systems', *Journal of Open Innovation: Technology, Market, and Complexity Article*, 5(18), p. 12. doi: 10.3390/joitmc5010018.
- Mantelero, A. (2018) 'AI and Big Data: A blueprint for a human rights, social and ethical impact assessment', *Computer Law & Security Review. Elsevier Advanced Technology*, 34(4), pp. 754–772. doi: 10.1016/J.CLSR.2018.05.017.
- Manthorpe, R. (2017) 'The Beauty.AI robot beauty contest is back', *WIRED-Business*. Available at: <https://www.wired.co.uk/article/robot-beauty-contest-beauty-ai> (Accessed: 17 May 2023).
- Metcalf, J. et al. (2021) 'Algorithmic Impact Assessments and Accountability: The Co-construction of Impacts CCS CONCEPTS', in *ACM Conference on Fairness, Accountability, and Transparency (FAcCT '21)*. New York, NY, USA: ACM, pp. 735–746. doi: 10.1145/3442188.
- Mintrom, M. and Luetjens, J. (2016) 'Design Thinking in Policymaking Processes: Opportunities and Challenges', *Australian Journal of Public Administration*. John Wiley & Sons, Ltd, 75(3), pp. 391–402. doi: 10.1111/1467-8500.12211.
- Mittelstadt, B. (2019) 'Principles alone cannot guarantee ethical AI', *Nat. Mach. Intell. Springer Science and Business Media LLC*, 1(11), pp. 501–507. doi: 10.1038/s42256-019-0114-4.
- Moss, E. et al. (2020) 'Governing with Algorithmic Impact Assessments: Six Observations', *SSRN Electronic Journal*. Elsevier BV. doi: 10.2139/SSRN.3584818.
- Moss, E. et al. (2021) 'Assembling Accountability: Algorithmic Impact Assessment for the Public Interest', *SSRN Electronic Journal*. Elsevier BV. doi: 10.2139/SSRN.3877437.
- Munn, L. (2022) 'The uselessness of AI ethics', *AI and Ethics 2022*. Springer, pp. 1–9. doi: 10.1007/S43681-022-00209-W.

- NIST (2022) AI Risk Management Framework: Second Draft Notes for Reviewers. Available at: https://www.nist.gov/system/files/documents/2022/08/18/AI_RMF_2nd_draft.pdf (Accessed: 18 May 2023).
- Novelli, C., Taddeo, M. and Floridi, L. (2023) 'Accountability in artificial intelligence: what it is and how it works', *AI and Society*. Springer Science and Business Media Deutschland GmbH, 1, pp. 1–12. doi: 10.1007/S00146-023-01635-Y/FIGURES/1.
- OECD (2019) *Artificial Intelligence in Society*, Artificial Intelligence in Society. Paris: OECD. doi: 10.1787/eedfee77-en.
- OECD (2023) *Fundamental Rights and Algorithms Impact Assessment (FRAIA)*, Catalogue of Tools & Metrics for Trustworthy AI. Available at: <https://oecd.ai/en/catalogue/tools/fundamental-rights-and-algorithms-impact-assessment-%28fraia%29> (Accessed: 17 May 2023).
- OPSI (2023) *Algorithmic Impact Assessment - Observatory of Public Sector Innovation*. Available at: <https://oecd-opsi.org/toolkits/algorithmic-impact-assessment/> (Accessed: 17 May 2023).
- Reisman, D. et al. (2018) *Algorithmic Impact Assessments: A practical framework for public agency accountability*. Available at: <https://ainowinstitute.org/publication/algorithmic-impact-assessments-report-2> (Accessed: 26 April 2023).
- Rességuier, A. and Rodrigues, R. (2020) 'AI ethics should not remain toothless! A call to bring back the teeth of ethics', *Big Data & Society*. SAGE PublicationsSage UK: London, England, 7(2). doi: 10.1177/2053951720942541.
- Selbst, A. D. (2021) 'An Institutional View of Algorithmic Impact Assessments', *Harvard Journal of Law & Technology* (Harvard JOLT), 35.
- Uchida, C. D. and Swatt, M. L. (2013) 'Operation LASER and the Effectiveness of Hotspot Patrol: A Panel Analysis', *Police Quarterly*. SAGE PublicationsSage CA: Los Angeles, CA, 16(3), pp. 287–304. doi: 10.1177/1098611113497044.
- UNESCO (2021) *Recommendation on the Ethics of Artificial Intelligence*. Paris. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000381137> (Accessed: 18 May 2023).
- Young, M. M., Bullock, J. B. and Lecy, J. D. (2019) 'Artificial Discretion as a Tool of Governance: A Framework for Understanding the Impact of Artificial Intelligence on Public Administration', *Perspectives on Public Management and Governance*. Oxford Academic, 2(4), pp. 301–313. doi: 10.1093/PPMGOV/GVZ014.
- Youth Laboratories (2016) *Beauty.AI 1.0*. Available at: <https://www.beauty.ai/#what> (Accessed: 17 May 2023).

Sobre la autora/ About the autor

María Lorena Flórez Rojas es PhD cum laude Scuola Superiore Sant'Anna en Italia, con Máster en Derecho y Tecnología de la Universidad de Tilburgo en Países Bajos y Abogada de la Universidad de Los Andes en Bogotá, Colombia. Actualmente es Profesora Asistente de la Universidad de Groningen, miembro del grupo de investigación STeP de la misma Universidad e Investigadora externa del Centro CinfonIA y GECTI de la Universidad de los Andes.

URL estable documento/stable URL

<http://www.gigapp.org>

El Grupo de Investigación en Gobierno, Administración y Políticas Públicas (GIGAPP) es una iniciativa impulsada por académicos, investigadores y profesores Iberoamericanos, cuyo principal propósito es contribuir al debate y la generación de nuevos conceptos, enfoques y marcos de análisis en las áreas de gobierno, gestión y políticas públicas, fomentando la creación de espacio de intercambio y colaboración permanente, y facilitando la construcción de redes y proyectos conjuntos sobre la base de actividades de docencia, investigación, asistencia técnica y extensión.

Las áreas de trabajo que constituyen los ejes principales del GIGAPP son:

1. Gobierno, instituciones y comportamiento político
2. Administración Pública
3. Políticas Públicas

Información de Contacto

Asociación GIGAPP.

ewp@gigapp.org

Influencia de las políticas de promoción de la Inteligencia Artificial de la UE en el proceso modernizador de las Administraciones Públicas españolas

Casarrubios Blanco, Elena

Universidad Rey Juan Carlos (URJC), España

✉ e.casarrubiosb@alumnos.urjc.es

ORCID ID: [0000-0003-4301-5387](https://orcid.org/0000-0003-4301-5387)

Documento recibido:

02 septiembre 2022

Aprobado para publicación:

07 agosto 2023

Resumen

El objetivo de este trabajo es exponer la forma en que la Unión Europea ha llevado a cabo la política de promoción de inteligencia artificial, en el marco del proceso modernizador de las Administraciones públicas en España, y más concretamente, de la Administración General del Estado español. La evolución de la innovación tecnológica asociada a esta tecnología data de principios del siglo actual. Su implantación en la Administración pública española lleva un camino paralelo al de la implantación en la Unión Europea, principalmente porque la tendencia de este modelo de innovación es que esté vinculada a los cambios en los ámbitos económico, político y social que sugiere su utilización, por la trasposición de directivas para su implantación (que también es característica de este modelo), y porque se puede considerar una tecnología inspiradora de nuevos avances en la forma de relacionar a las administraciones con ciudadanos y empresas. Son estos argumentos importantes para su consideración como motor de las políticas de promoción actuales de las Tecnologías de la Información y las Comunicaciones, no solo por ser tendencia en el mercado, sino porque de su buena adaptación depende que su gran potencial llegue a ser una realidad en el día a día de las administraciones.

Palabras clave

España, Administración General del Estado, Unión Europea, políticas públicas, Inteligencia artificial.

Resumo

O objetivo deste trabalho é expor a forma como a União Europeia tem levado a cabo a política de promoção da inteligência artificial, no quadro do processo de modernização das Administrações Públicas em Espanha, e mais especificamente, da Administração Geral do Estado espanhol. A evolução da inovação tecnológica associada a esta tecnologia data do início do século atual. A sua implementação na administração pública espanhola segue um caminho paralelo ao da sua implementação na União Europeia, principalmente porque a tendência deste modelo de inovação é que esteja ligado às mudanças nas esferas económica, política e social que a sua utilização sugere, pela transposição de diretivas para a sua implementação (o que também é característico deste modelo), e porque pode ser considerada uma tecnologia inspiradora para novos avanços na forma de relacionar as administrações com os cidadãos e as empresas. Estes são argumentos importantes para a sua consideração como motor das atuais políticas de promoção das Tecnologias de Informação e Comunicação, não só por ser uma tendência no mercado, mas porque a sua boa adaptação depende do seu grande potencial de concretização. administrações atuais.

Palavras-chave

Espanha, Administração Geral do Estado, União Europeia, políticas públicas, inteligência artificial..

Abstract

The objective of this work is to expose the way in which the European Union has carried out the artificial intelligence promotion policy, within the framework of the modernization process of the Public Administrations in Spain, and more specifically, of the General Administration of the Spanish State. The evolution of technological innovations associated with this technology dates from the beginning of the current century. Its implementation in the Spanish public administration follows same path to that of the implementation in the European Union, mainly because the trend of this innovation model is that it is linked to the changes in the economic, political and social spheres that its use suggests, for the transposition of directives for its implementation (which is also characteristic of this model), and because it can be considered an inspiring technology for new advances in the way of relating administrations with citizens and companies. These are important arguments for its consideration due current promotion policies of Information and Communication Technologies, not only because it is a trend in the market, but because its good adaptation depends on its great potential becoming a reality. in the day-to-day administrations.

Keywords

Spain; General State Administration; European Union; public policies; artificial intelligence.

1. Políticas de modernización basadas en las TIC en la Unión Europea

Una de las características comunes de las políticas de promoción de las Tecnologías de la Información y las Comunicaciones (TIC), en la Unión Europea (UE), es la forma en que, a partir de la potenciación de las primeras, la sociedad de la información se ha implantado en las economías europeas. Analizando el planteamiento seguido a la hora de llevar a cabo las políticas públicas¹ relacionadas con estas materias, así como sus efectos colaterales, se puede llegar a conclusiones muy interesantes.

En relación con la promoción de políticas públicas cuyo objetivo era la introducción de la sociedad de la información, el caso europeo –a diferencia de los casos norteamericano, japonés, coreano, y de Singapur– se ha caracterizado por la falta de potenciación del consumo nacional, y por el proteccionismo a este tipo de industrias, dejando que el mercado actuase a su albedrío, aunque definiendo un marco regulatorio en aras de un “servicio universal”². Los efectos de estas políticas han sido la armonización del mercado de telecomunicaciones y servicios asociados desde un punto de vista técnico y de infraestructura, con medidas como la regulación de la actuación de las empresas de telecomunicaciones en un marco de mercado muy competitivo, y la oferta de un marco normativo que ha adoptado, reconocido y definido los estándares técnicos de la industria digital. El fruto de la aplicación de estos programas ha sido la fijación de unos estándares de interoperabilidad y de unas directrices respecto a la comunicación (Casarrubios, 2007).

Respecto a la estandarización, el cumplimiento de las normas armonizadas ha contribuido a garantizar que los productos se ajustasen al derecho de la Unión, a que se alcanzasen objetivos relacionados con la neutralidad climática y con una economía resiliente y circular, y a que se legitimara el trabajo en los foros e instituciones internacionales más importantes. De esta forma, para incrementar la competitividad de las empresas y eliminar los obstáculos al comercio internacional, la UE ha creado una serie de instituciones dedicadas a la normalización técnica, emitiendo estándares que garantizaran la equiparación entre los equipos diseñados bajo sus fronteras, y los del resto del mundo (Casarrubios, 2007)³. También se han establecido métodos y procedimientos necesarios para com-

¹. Las políticas públicas «...representan flujos de prácticas dirigidas intencionalmente a la resolución de algún problema, y en las cuales se transforman distintos tipos de recursos para la producción de bienes, servicios o regulaciones necesarios. Las políticas representan el instrumento principal que tienen los gobiernos para resolver las cuestiones que han problematizado, sin que esté garantizado a priori un resultado favorable. Los gobiernos enfrentan desafíos cada vez más complejos, principalmente por la creciente multicausalidad de los mismos, y por las profundas interdependencias de los factores explicativos» (Bertranou, 2019).

². Se trata de servicios universales de comunicaciones electrónicas ofrecidas a todos los usuarios que lo soliciten. Moore afirma que hay dos tipos de mecanismos a los que los países recurren para elaborar políticas informacionales en relación con la prestación de estos servicios y el desarrollo de la Sociedad de la Información: los que surgen en torno al criterio «capital privado», actuando como agentes facilitadores (obtención de beneficio porque el Estado no interviene ni regula, como por ejemplo Australia, Canadá y Sudáfrica), y los modelos intervencionistas, en los que la política de información está liderada por el Estado, de acuerdo con objetivos fijados con antelación. Este será, por lo tanto, el proveedor de fondos de inversión, el creador de la necesidad (productos y servicios), el propietario de la infraestructura de telecomunicaciones, y el regulador (Moore, 2002).

³ En el nivel técnico existen varios organismos, como el Comité Europeo de Normalización (CEN), el Comité Europeo de Electrotecnia (CENELEC), la Confederación Europea de Administraciones de Correos y Telecomunicaciones (CEPT) y el Instituto Europeo de Normalización de Telecomunicaciones (ETSI). El CEN y el CENELEC se crearon en 1961 y 1962 respectivamente para elaborar normas y armonizar las divergencias entre las normas nacionales. Trabajan junto a la CEPT y

probar y certificar la conformidad de estas normas (sobre todo en el ámbito de la informática y las telecomunicaciones) (Casarrubios, 2007)⁴.

Por otro lado, la promoción de infraestructuras a nivel regional ha pretendido promover la construcción de las redes transeuropeas de transporte, energía y telecomunicaciones, algo a lo que se ha dedicado gran parte del presupuesto de la UE. Para ello se han dispuesto diferentes mecanismos de financiación, unos a fondo perdido (en el caso de las regiones menos desarrolladas), y otros cofinanciados con otros actores como empresas, operadoras tecnológicas, administraciones públicas⁵.

el ETSI. El ETSI fue creado en 1988 emitiendo normas de carácter técnico. En la actualidad centra su actividad en la producción de especificaciones que garanticen la compatibilidad de las redes y servicios de telecomunicación, tanto a nivel territorial (UE), como a nivel funcional (partes comunes de las telecomunicaciones, las tecnologías de la información, y la radiodifusión). En cuanto a los organismos y organizaciones se distinguen el Seminario Europeo de Sistemas Abiertos (EWOS), el Grupo de Promoción y Aplicación de las Normas (SPAG), la Asociación Europea de Fabricantes de Ordenadores (ECMA), la Industria Europea de Telecomunicaciones y de Electrónica Profesional (ECTEL), el Grupo de Usuarios Europeos de MAP, Protocolo de Automatización de la Fabricación (EMUG), el estándar de sistemas operativos Unix, de PTT (X-OPEN), la Confederación de Asociaciones de Usuarios de Ordenadores (CECUA), el Grupo de Altos Funcionarios para la Normalización de las Tecnologías de la Información (SOGITS), el Grupo de Altos Funcionarios para las Telecomunicaciones (SGT) y por último el Comité de Contratación pública en el sector de Tecnologías de la Información (PPG), entre otros. Por último, existe toda una red de organismos formales desarrollando normas. Por ejemplo, el Comité Técnico Consultivo (JTC-1) está formado por miembros de la Organización Internacional de Normalización (ISO) y el Comité Electrotécnico Internacional (IEC). Trabaja únicamente con las Tecnologías de la Información, pero no de las Comunicaciones. De éstas se encarga el Comité Consultivo Internacional de Telégrafos y Teléfonos (CCITT) de la Unión Internacional de Telecomunicaciones (UIT), que es finalmente la que redacta los estándares de la comunicación entre servicios públicos.

⁴ La publicación del Reglamento (UE) N° 1025/2012 sobre la normalización europea ha derogado la Decisión 87/95/CEE, y ha modificado diversas directivas europeas que incidían en la materia. En el contexto de La «Agenda Digital para Europa» se indicaba que las normas relacionadas con las TIC son fundamentales para la interoperabilidad entre los productos y servicios de TI, de forma que si se mejoraran estas y se promoviera un mejor uso (especialmente por parte de las autoridades públicas cuando adquieran hardware, software y servicios de TIC, con el objeto de promover la eficiencia y reducir la dependencia de proveedores concretos) se obtendría una mejora en el crecimiento económico. En el 2011, la comunicación de la Comisión «Una visión estratégica de las normas europeas: Avanzar para mejorar y acelerar el crecimiento sostenible de la economía europea de aquí a 2020» (Comisión Europea, 2011) apuntaba, en primer lugar, a la creación de un sistema europeo de normalización que apoyara la estrategia Europa 2020 para un crecimiento inteligente, sostenible e integrador. También se incluía un paquete de medidas legislativas y no legislativas resultado de una amplia revisión del sistema europeo de normalización, realizada entre 2008 y 2010, que incluía el «Informe Express» (Informe del Grupo de expertos encargado de la revisión del sistema europeo de normalización, de febrero de 2010 titulado «Standardization for a competitive and innovative Europe: a vision for 2020», (Parlamento Europeo, 2010), dos consultas públicas, el libro blanco «Modernizar la normalización de las TIC en la UE: El camino a seguir» y una serie de estudios en profundidad.

⁵ «...Esta evolución...», (refiriéndose a la influencia de las TIC, que está presente tanto en el incremento de la dotación presupuestaria que se ha hecho para adquirirlas, desde la óptica de las políticas públicas, como en la evolución que la AGE ha realizado para llevar a cabo su proceso modernizador) «...ha seguido en paralelo a la de la gestión presupuestaria de la UE, y su estudio ha venido dando la pauta de la forma en que se han invertido los recursos económicos para llegar a objetivos de desarrollo determinados de antemano (en materia de política agrícola común, de acciones estructurales, de definición del papel de la UE en el ámbito político internacional, de modernización de la administración comunitaria y de la de los Estados miembro, y de desarrollo tecnológico -redes transeuropeas de comunicaciones, I+D, educación...). En España, la forma en que se ha reflejado en la evolución presupuestaria la promoción de las políticas públicas relacionadas con las TIC no es algo claro ya que se distinguen múltiples planes que han provisto fondos para conseguir que España vaya integrándose poco a poco en la economía competitiva de la UE» (Casarrubios Blanco,

El resultado ha sido no sólo la existencia de una infraestructura de comunicaciones, sino también de un mercado que, al tiempo que la ha demandado, la ha sufragado y ha potenciado tanto su consumo interno como externo.

Por último, otra de las contribuciones de la UE ha sido el diseño de las políticas de I+D+i relacionadas con las TIC. Los sucesivos programas marco en Europa han ido dotando cantidades cada vez más importantes a este capítulo para fomentar una cultura de innovación, el establecimiento de un marco jurídico, normativo y financiero favorable, o la articulación de una mejor investigación e innovación a nivel nacional y comunitario⁶. En el caso concreto de las Administraciones públicas españolas, la promoción de las políticas relacionadas con el despliegue de la sociedad de la información se asoció inicialmente a la implantación de normas de comunicaciones que fueran reconocidas por todos los niveles territoriales, estandarizando las normas y creando plataformas de uso común (Casarrubios, 2007). Sin embargo, fueron dando paso a planes concretos relacionados con procesos de modernización en los que los protagonistas comenzaron siendo las administraciones públicas, pero después fueron los ciudadanos y empresas a los que estas políticas públicas iban dirigidos.

Así, la IA se ha posicionado como una de las tendencias más importantes en innovación relacionada con las TIC, con un potencial de integración de anteriores tecnologías ya consolidadas, y de nuevas perspectivas relacionadas con el análisis de datos y de su almacenamiento y tratamiento masivo y securizado.

Algunas notas sobre la innovación tecnológica y gestión de las tecnologías de la información y las comunicaciones en la Administración General del Estado (España), 2023)

⁶ Con anterioridad al primer Programa Marco (1984- 1987) -e incluso durante el desarrollo del mismo- la investigación se centró en los campos del acero y del carbón, la energía nuclear y la agricultura. Hasta 1974 no se habían concebido proyectos de política científica comunitaria, pero a partir de este momento se configuró la estructura de los programas, que cubría seis áreas: energía, recursos, medio ambiente, condiciones de vida y trabajo, servicios e infraestructura, e industria. A comienzos de 1980 se disponía de una gran cantidad de fondos para proyectos de investigación, aunque hasta los 90 las políticas de información continuaron segregadas de los desarrollos del área tecnológica. A partir de los años 90 se revisó la política de TIC, por varias causas: porque la brecha tecnológica con EEUU no había disminuido con las políticas anteriores, porque se buscaba la coordinación e integración de la IDT con otras políticas de la UE, y porque comenzó a surgir una demanda de obtención de beneficios sociales de los recursos empleados en I+D, dejando a un lado el objetivo del incremento de la competitividad. Comenzaron a promoverse actividades como el fomento de una cultura de innovación, el establecimiento de un marco jurídico, normativo y financiero favorable, o la articulación de una mejor investigación e innovación a nivel nacional y comunitario. Surgió entonces el Quinto Programa Marco, que mantuvo como eje principal el fomento de la competitividad industrial (los programas de tecnologías de la información, comunicación y materiales), pero que contó con factores innovadores como la búsqueda de participación de las pequeñas y medianas empresas. Entre las idas y las venidas de las crisis y de la situación económica, el objetivo de conseguir el pleno empleo (asociado principalmente a la investigación y a la innovación) también se acerca y se aleja. En el Sexto Programa Marco (2002- 2006) se introduce la temática de los Ciudadanos y la gobernabilidad en la Sociedad del Conocimiento, incorporando además en la participación de estos programas a las Universidades, a los investigadores, a las Empresas, y las administraciones públicas. A partir del Séptimo Programa también se incrementa considerablemente el presupuesto anual de la UE en materia de investigación. En el Octavo Programa Marco se integran por primera vez todas las fases de la gestión del conocimiento.

2. Algunos apuntes sobre la Inteligencia Artificial

Se trata de un bloque de tendencias en el uso de TIC que, si bien todavía cuentan con pocos casos de implantación (aunque relevantes por su componente innovador), todavía no forman parte del conjunto de herramientas que utiliza y ofrece la Administración General del Estado español (AGE a partir de ahora), sino que más bien se observan como posibilidades con gran potencial. Estas tecnologías innovadoras son la Inteligencia Artificial (desde aquí IA), Big Data o tratamiento de datos masivos, y BlockChain⁷ (Marín Pérez, 2022) o tratamiento de datos encadenados y securizados. Todo ello soportado por tecnología alojada en la nube (on Cloud). De las dos primeras, especialmente de la IA, trata esta investigación.

Los Sistemas de IA son "los programas informáticos (y posiblemente también equipos informáticos) diseñados por seres humanos que, dado un objetivo complejo, actúan en la dimensión física o digital mediante la percepción de su entorno mediante la adquisición de datos, la interpretación de los datos estructurados o no estructurados, el razonamiento sobre el conocimiento o el tratamiento de la información, fruto de estos datos y la decisión de las mejores acciones que se llevarán a cabo para alcanzar el objetivo fijado" (Comisión Europea, 2020). La importancia de esta tecnología radica en su capacidad para interactuar con el entorno, resolviendo -a partir de algoritmos- reglas de comportamiento y autoprogramación, aplicables en los campos de lógica, informática, aprendizaje automático, razonamiento automático, robótica, neurociencia, matemáticas, psicología cognitiva. Es decir, aquellas materias susceptibles de ser traducidas a código informático, para reproducir el comportamiento humano en cuanto a conocimiento, razonamiento y aprendizaje.

Se parte de la premisa de que la máquina de procesamiento de conocimiento más perfecta es la humana, ya que (si bien existen computadoras y robots que pueden procesar grandes volúmenes de información a gran velocidad) todavía no se ha conseguido la generación de un algoritmo que despliegue nuevos algoritmos en base a patrones de necesidad, y de su consiguiente generación de reglas para resolverla. Todo ello en permanente retroalimentación, tal es la complejidad del cerebro humano. Además, esta tecnología tiene un gran potencial -y de aquí radica su importancia transformadora de la sociedad.

Así, se espera que, a partir del tratamiento de grandes cantidades de información en forma de bases de datos, la UE pueda combinar capacidad tecnológica e industrial en una infraestructura digital, y un marco regulatorio. Los diversos avances en esta tecnología, puede suponer cambios para ciudadanos, el sector empresarial y para la misma administración pública. Para los ciudadanos contribuiría a la mejora del sistema de salud, del sistema energético, del de transportes, de los servicios públicos, etc. Para el entorno empresarial podría generar productos y servicios en áreas como maquinaria, transporte, seguridad, ganadería, economía circular y sostenibilidad, salud, y otros sectores con un alto valor añadido como la moda o el turismo. Además, para los servicios públicos esta gran capacidad tendría en cuenta la reducción de costes en la provisión de servicios (transporte, educación, energía, entre otros) sostenibles, con herramientas apropiadas que garantizarían la seguridad de los ciudadanos, sus derechos y su sistema de libertad. Finalmente, para la Administración pública, su-

⁷ De forma sencilla, podríamos explicar el funcionamiento de la Blockchain como un proceso que permite a los nodos en una red proponer, validar y registran cambios de estado en un libro mayor sincronizado/distribuido a través de dichos nodos. Es decir, cada nodo participante cuenta con una réplica de este libro, evitando así que los datos se vean comprometidos en caso de fallo en algún nodo

pone la posibilidad de que los empleados públicos dejen de realizar tareas rutinarias y automáticas y puedan dedicarse a actividades más complejas en las que puedan aportar más valor.

Este es el caso de la automatización robótica de procesos (RPA), la cual parte del aprendizaje automático y del desarrollo de software capaz de tratar parcialmente procesos repetitivos. Como, por ejemplo, abrir correos con archivos adjuntos, conectarse a aplicaciones electrónicas, mover archivos y carpetas, copiar y pegar, llenar formularios, leer y escribir con destino y origen en las bases de datos, recolectar estadísticas de redes sociales, recopilar información de internet, extraer datos estructurados de documentos, realizar cálculos, conectar interfaces de aplicaciones. Todo ello de acuerdo con una serie de reglas predefinidas.

A nivel mundial, las aplicaciones en la vida real de la IA son variadas. Se encuentran en los campos de la biología ejemplos como AlphaFold y DeepMind, de Google, y el Laboratorio Europeo de Biología, con una gran capacidad de secuenciación y de predicción en base al almacenamiento y tratamiento de datos financiados por recursos públicos, que son gestionados por la comunidad científica.

De la medicina hay casos en Indonesia, como el Grupo de Trabajo del Gobierno para la Investigación y la Innovación Tecnológica, que ha estado trabajando en modelos para reforzar los diagnósticos de médicos en la detección del coronavirus. En Croacia, su proyecto de Medico Virtual desarrollado por empresas tecnológicas y por epidemiólogos han podido procesar miles de solicitudes diarias frente a las limitaciones por número y capacidad de los médicos.

En cuanto a la seguridad y la defensa, en varios países, basándose en la recolección de datos relacionados con el reconocimiento facial y con cámaras térmicas situadas en lugares estratégicos, han podido identificar a personas potencialmente infectadas. En el área de defensa, la herramienta GIDE (Experimento de Dominación Global de Información) recopila los datos necesarios para predecir los movimientos de los enemigos en base a gran cantidad de sensores -civiles y militares- repartidos por todo el mundo, y también en el área de seguridad se encuentran ejemplos en la vigilancia de calles, manifestaciones, alertas por terrorismo, etc., por drones que aportan grandes cantidades de información para elaborar patrones de comportamiento (Arenilla, 2021).

Por su parte en España son especialmente relevantes las iniciativas Forecast, herramienta para la toma de decisiones de gestores forestales (con datos del Instituto Geográfico Nacional, Ministerio para la Transición Ecológica y el Reto Demográfico y Copernicus); Ártabro Tender, empleada para encontrar y elegir los concursos públicos que una empresa realmente puede ganar (con datos del Ministerio de Hacienda y Gestión Pública); MarIA, la primera inteligencia artificial de la lengua española (con datos procedentes de la Biblioteca Nacional de España), entre otras (Comisión Europea, 2023).

Junto a la IA existen una serie de conceptos complementarios, y son fundamentales para entender su alcance. En primer lugar, se encuentra el concepto "Big Data", tecnología basada en el procesamiento de grandes volúmenes de información que, partiendo del almacenamiento con acceso veloz, controlado y securizado, trata los grandes volúmenes de datos (tanto estructurados como no estructurados) para dar valor a esta información a través de su capacidad de establecer relaciones entre estas estructuras. En definitiva, ha surgido como solución a la necesidad de utilizar y compartir datos entre empresas privadas, instituciones gubernamentales, organizaciones sociales, medios de comunicación, redes sociales, dispositivos, universidades e instituciones científicas.

"Big Data" comprende tanto el factor técnico como el humano. Del primer factor es importante decir que es necesario que los datos sean abiertos (públicos por legislación y libres para ser reutilizados por terceros), estandarizados y con metadatos asociados (información relacionada con la forma en

que se han producido). Estos datos podrán ser analizados, visualizados, y las predicciones que generen serán susceptibles de generar nuevos datos, conocimiento o servicios. En cuanto al factor humano debe ser capaz de ejecutar procesos y de tomar decisiones a partir de estos datos.

En segundo lugar, el concepto “data mining” resulta relevante para el debate. Sus antecedentes datan de la década de 1970, cuando la introducción del código de barras permitió la captura de un gran volumen de datos relacionados con productos. A partir de los ochenta los CRM (Software de Gestión de Relaciones con los clientes) hicieron posible la captura de los datos de clientes que ofrecían cierto valor a las empresas que invertían en el tratamiento de sus datos (Casarrubios Blanco, 2019). Y estos comenzaron en los noventa a ser almacenados en la nube⁸ de forma más eficiente. Actualmente, la Ciencia de los Datos utiliza técnicas como la estadística, las matemáticas, el aprendizaje automático y las ciencias de la computación. Es esta ciencia la que permite, junto a la obtención de datos fruto de la interrelación de nuevas tecnologías interconectadas (móviles, GPS, tabletas, entre otras) ordenar y establecer relaciones en base a modelos predeterminados.

Por su parte, en las Administraciones públicas, esta tecnología es especialmente importante, ya que estas normalmente disponen de gran cantidad de datos que, de ser tratados, supondrían el fortalecimiento de las relaciones con los ciudadanos y la mejora de la prestación de los servicios públicos, al solucionar uno de los problemas más recurrentes de éstas: el aislamiento de los datos en sistemas estancos, no compartidos, y su tendencia a ser reproducidos en papel y gestionados en la forma tradicional sin demasiado valor añadido. Este es uno de los argumentos empleados para avanzar en una nueva gobernanza de datos basada en la apertura de la información, su posterior integración en el sistema público, su sometimiento a transparencia y control, y su protección jurídica.

En este escenario, los datos serían aportados por las mismas administraciones, las empresas, asociaciones y -por supuesto- por los ciudadanos. A manera de ejemplo, el Banco de Inglaterra, que ha desarrollado bancos de datos para diseñar estrategias de inversión en esta materia y experimentar con las tecnologías y herramientas digitales intentando, sobre todo, evitar un gasto excesivo por decisiones adoptadas en materia de infraestructuras y equipamientos no debidamente contrastadas con las necesidades reales. Este sería el modelo más simple, empleando principalmente técnicas de visualización para comprender mejor la información (Arenilla, 2021).

Del valor que puede apoyar un dato, se distingue el valor descriptivo, predictivo y prescriptivo. El primero se basa en técnicas descriptivas que permiten comprender un conjunto de variables que afectan a un problema, mientras que el segundo, pretende pronosticar una evolución de la realidad a través de modelos estadísticos, técnicas de aprendizaje automático, patrones históricos y grupos de datos actualizados con la probabilidad de una tendencia. El último es el menos desarrollado, e incluye algoritmos de optimización, análisis de decisión multicriterio, reglas de negocio para determinar cuál es la mejor acción actual o futura que debe tomar un agente en una situación, y la posibilidad de disponer de los resultados de los otros dos tipos de análisis. Así, la utilidad de este valor predictivo se

⁸ Esta tecnología se define en la actualidad como un <<... modelo tecnológico que permite el acceso ubicuo, adaptado y bajo demanda en red a un conjunto compartido de recursos de computación configurables compartidos (por ejemplo: redes, servidores, equipos de almacenamiento, aplicaciones y servicios), que pueden ser rápidamente aprovisionados y liberados con un esfuerzo de gestión reducido o interacción mínima con el proveedor del servicio.>> (Mell & Grance, 2011). A partir de este momento las organizaciones no tuvieron por qué preocuparse por su infraestructura, sino por la elección de la tecnología más adecuada a sus necesidades, afrontando su adquisición de forma asequible, y su gestión de forma racional, flexible, y funcional. Los servicios se pudieron dimensionar al alza o a la baja en poco tiempo sin tener que adquirir compromisos a largo plazo.

ha hecho evidente en el caso del análisis de datos de telefonía móvil para prevenir contagios, dar soporte en caso de inundaciones y catástrofes naturales. Por ejemplo, el proyecto “*Billion Prices Project*” analiza millones de datos de agentes minoristas de todo el mundo para realizar proyecciones relacionadas con microeconomía e inflación y permite predecir los precios en un periodo de tiempo (Arenilla, 2021).

Por su parte, en la Administración pública, dado que los datos han de ser abiertos (públicos por legislación y libres para ser reutilizados por terceros), estandarizados y con metadatos asociados (información relacionada con la forma en que se han producido), la tecnología Big Data está especialmente indicada para predecir las necesidades de los ciudadanos, el desarrollo de servicios, la promoción de empleo, los planes de infraestructuras, de eficiencia energética, de consumo de recursos, el diseño de ciudades inteligentes. Es decir, todas aquellas materias que pueden ser analizadas, visualizadas, y cuyas predicciones serán susceptibles de generar nuevos datos, conocimiento o servicios.

Entre los proyectos más conocidos se encuentran el de la Dirección General de Tráfico en España, organismo adscrito al Ministerio del Interior que analiza datos masivamente con el fin de predecir el número de accidentes de tráfico. En este sentido, la iniciativa “IA para la seguridad vial”, junto a otras instituciones⁹, pretende aprovechar el valor de la IA para mejorar el enfoque de sistema seguro para la seguridad vial para todos los usuarios (vehículos, usuarios vulnerables de la vía ya sean motorizados o no motorizados; por ejemplo, peatones, ciclistas, motociclistas, patinetes eléctricos, etc.) (Dirección General de Tráfico, 2023).

Pero sin duda, uno de los casos más interesantes es el de la Agencia tributaria española, que ya utiliza Big Data “...incorporando volumen de ingresos y gastos a los empresarios en Estimación objetiva no agraria (EONA), y en otros usos, como vía para proporcionar más elementos de evaluación útiles para los actuarios que intervienen en la selección y control de contribuyentes, mejorando la gestión de riesgo de fraude y la priorización de las actuaciones” (Gómez López, 2022).

3. Políticas de promoción relacionadas con la IA

El documento “Una Unión que se esfuerza por lograr más resultados” contiene las guías orientadoras de política para la Comisión 2019-2024. En este documento se anuncia la intención de comenzar a legislar la IA en todo lo relacionado con sus implicaciones éticas y humanas. Por su parte, a comienzos del 2020 se publicó el “Libro Blanco sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza” (Comisión Europea, 2020), en el que se definía el objetivo de promover la IA, al tiempo que se abordaban los riesgos de su uso. Para ello se pretendía crear un marco jurídico que garantizara la fiabilidad de esta tecnología en base a los valores y derechos fundamentales de la UE.

Tras este Libro Blanco se realizó una consulta en la que quedó clara la necesidad de una regulación para hacer frente a desafíos y riesgos de un cada vez más solicitado uso de la IA, afrontando la opacidad, la complejidad, el sesgo, cierto grado de imprevisibilidad, y un comportamiento parcialmente autónomo de ciertos sistemas de IA. En octubre de 2020, fueron aprobadas una serie de resoluciones sobre cuestiones como la ética, la responsabilidad civil y los derechos de propiedad intelectual, que

⁹ Esta iniciativa está en consonancia con la Resolución de la Asamblea General de las Naciones Unidas (UN A / RES / 74/299) sobre la mejora de la seguridad vial global, y tiene la colaboración del Grupo de enfoque del evento de lanzamiento sobre conducción autónoma y asistida (FG-AI4AD), y de la Unión Internacional de Telecomunicaciones (UIT),

se complementaron en el 2021 con otras directrices sobre el uso de la IA en el ámbito penal y en los sectores educativo, cultural y audiovisual. Por ejemplo, la Resolución del Parlamento Europeo¹⁰ sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas se alineaba con los principios de proporcionalidad, subsidiariedad y mejora de la legislación¹¹.

El marco reglamentario propuesto tenía como objetivos garantizar la seguridad jurídica para facilitar la inversión e innovación en IA; mejorar la gobernanza y la aplicación efectiva de la legislación vigente en materia de derechos fundamentales y los requisitos de seguridad aplicables a los sistemas de IA; y facilitar el desarrollo de un mercado único para hacer un uso legal, seguro y fiable de las aplicaciones de IA y evitar la fragmentación del mercado. También este esfuerzo se enfocó en “la mejora de infraestructuras y de capacidades tecnológicas, el desarrollo de capacidades en las cadenas de valor digitales estratégicas, acelerar el despliegue de infraestructuras y de muy alta capacidad (en particular la fibra y la 5G) y mejorar la capacidad de la UE para protegerse contra las ciberamenazas, proporcionar entornos de comunicación seguros, especialmente mediante la encriptación cuántica y garantizar el acceso a los datos a efectos judiciales y políticos” (Mineco, 2020).

En cuanto a la legislación aplicada en España, la primera legislación de la IA es la Ley 15/2022, ley integral para la igualdad de trato y la no discriminación, de 12 de Julio. Se trata de una ley que, sin tener que ver específicamente con esta tecnología, define de forma voluntarista el tratamiento de los algoritmos involucrados en la toma de decisiones. El marco al que somete esta norma es el de la Estrategia Nacional de IA, la Carta de Derechos Digitales (Ley 16317/ 2021), y las iniciativas europeas relacionadas con la IA, y lo hace para señalar que las administraciones públicas priorizarán en el marco de sus competencias criterios de minimización de sesgos, transparencia y rendición de cuentas, además de subrayar su potencial para evitar la discriminación. Es importante resaltar que anteriormente, ninguna de las leyes relacionadas con el tratamiento legal de información digital trataba en profundidad de esta tecnología, pero, tampoco lo hace esta norma en detalle, pues únicamente delimita de forma poco precisa los límites de su uso.

¹⁰ Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL)).

¹¹ Comisión Europea, Libro Blanco sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza, COM (2020) 65 final, 2020; Consejo Europeo, Reunión extraordinaria del Consejo Europeo (1 y 2 de octubre de 2020) – Conclusiones, EUCO 13/20, 2020, p. 6; Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas, 2020/2012(INL); Consejo Europeo, Reunión del Consejo Europeo (19 de octubre de 2017) – Conclusiones EUCO 14/17, 2017, p. 8; Consejo de la Unión Europea, Inteligencia artificial: b) Conclusiones relativas al Plan Coordinado sobre la Inteligencia Artificial – Adopción, 6177/19, 2019; Consejo Europeo, Reunión extraordinaria del Consejo Europeo (1 y 2 de octubre de 2020) – Conclusiones, EUCO 13/20, 2020; Consejo de la Unión Europea, Conclusiones de la Presidencia - La Carta de los Derechos Fundamentales en el contexto de la inteligencia artificial y el cambio digital, 11481/20, 2020; Resolución del Parlamento Europeo, de 20 de octubre de 2020, sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas, 2020/2012(INL) ; Resolución del Parlamento Europeo, de 20 de octubre de 2020, sobre un régimen de responsabilidad civil en materia de inteligencia artificial, 2020/2014(INL) ; Resolución del Parlamento Europeo, de 20 de octubre de 2020, sobre los derechos de propiedad intelectual para el desarrollo de las tecnologías relativas a la inteligencia artificial, 2020/2015(INI) ; Proyecto de informe del Parlamento Europeo sobre la inteligencia artificial en el Derecho penal y su utilización por las autoridades policiales y judiciales en asuntos penales, 2020/2016(INI) ; Proyecto de informe del Parlamento Europeo sobre la inteligencia artificial en los sectores educativo, cultural y audiovisual, 2020/2017(INI); Plan de Acción de Educación Digital 2021-2027: Adaptar la educación y la formación a la era digital, (Comunicación de la Comisión COM(2020) 624 final).

Actualmente existe un proyecto de Ley de la Inteligencia Artificial propuesto por la UE en abril de 2021. Se encuentra en proceso de consulta, y se debe aprobar por los estados miembros de la UE y el Parlamento Europeo antes de convertirse en ley. Si es aprobada su fecha de entrada en vigor es posible sea en 2024. Su base sería un sistema de clasificación que determina el nivel de riesgo que una tecnología de IA podría suponer para la salud y la seguridad o los derechos fundamentales de una persona. Por este motivo este marco incluiría cuatro niveles de riesgo: inaceptable, alto, limitado y mínimo¹².

En España, además de la transposición de la ley que actualmente se encuentra en fase de borrador, la elaboración de la presente Estrategia Nacional ha respondido al compromiso de contribuir a que la UE se sitúe como líder en esta materia, compromiso que se ha recogido en la Estrategia Española de I+D+I para el periodo 2020-2025. Esta estrategia tiene siete objetivos estratégicos: promover la excelencia científica e innovación en Inteligencia Artificial, proyectar la lengua española, crear empleo cualificado, transformar el tejido productivo, crear un entorno de confianza para desarrollar la IA, impulsar un debate acerca de los valores humanistas de la IA, y potenciar una IA inclusiva y sostenible.

Asimismo, se han definido seis objetivos o ejes de actuación con acciones prioritarias ejecutables a lo largo del periodo 2020-2025, planteando una serie de líneas de actuación para este periodo. Estos objetivos son impulsar la investigación científica, el desarrollo tecnológico y la innovación en IA, promover el desarrollo de capacidades digitales, potenciar el talento nacional y atraer talento global, desarrollar plataformas de datos e infraestructuras tecnológicas que den soporte a la IA, integrar la IA en las cadenas de valor para transformar el tejido económico, potenciar el uso de la IA en la Administración Pública y en las misiones estratégicas nacionales, y establecer un marco ético y normativo que reforzara la protección de los derechos individuales y colectivos, a efectos de garantizar la inclusión y el bienestar social. Actualmente, en el Mecanismo de Recuperación y Resiliencia (MRR) hay recogidos varios proyectos que presentan las necesidades de administraciones y empresas para obtener beneficios de la IA. Los fondos Next Generation EU se consideran la dotación presupuestaria para la realización de todas estas propuestas.

4. Conclusiones

La implementación de la IA en las Administraciones públicas es uno de los últimos capítulos relacionados con la evolución de las TIC y con su despliegue. Ha contado con una serie de avances técnicos que ya se han materializado en proyectos útiles y con posibilidad de futuro. Pero, también ha contado con la promoción de las políticas de información relacionadas con la implantación de las TIC, de forma que se ha tenido en cuenta el análisis de casos existentes, de necesidades que podría satisfacer, así como de un marco regulatorio que considere tanto a la oferta de servicio como a su gestión, en especial a la parte ética de su gestión. Este interés en crear un marco normativo corresponde a la necesidad de defender a los ciudadanos del uso indiscriminado de datos por parte de las industrias y los estados.

¹² Los sistemas de IA con un riesgo limitado y mínimo (como ejemplo se citan los videojuegos) pueden utilizarse sin más requisitos que las obligaciones de transparencia. Los sistemas que se considera que plantean un riesgo inaceptable (por ejemplo, los sistemas de información gubernamentales, o los de identificación biométrica en tiempo real en espacios públicos) estarían prohibidos con pocas excepciones.

Del mismo modo, la regulación del uso de la IA va a suponer un antes y un después en el proceso de implantación de las TIC en las administraciones. La implicación de los gobiernos en la elaboración de un marco que tome en consideración principios sociales y económicos, y su adaptación a las directivas a las que se traduzca este marco, va a ser también un gran reto. Es interesante que una de las premisas que se deba tener en cuenta para el buen uso de una tecnología (de la información y las comunicaciones en este caso), es la necesidad de un escenario en el que la formación que capacite a los funcionarios públicos en términos de innovación, no discriminación y transparencia sea requisito para que sea utilizada. Tal vez sea esta la piedra angular de un futuro próximo pilar de las TIC, que puede dar lugar a una posible generación de software, hardware, comunicaciones y servicios relacionados. En este sentido, cada vez se ofrecen más iniciativas formativas dedicadas a la orientación técnica en el uso de la IA, aunque menos los que cuentan con contenido relacionado con la ética en su aplicación.

Por otro lado, otro de los factores para tener en cuenta es el grado de participación de los funcionarios en el diseño de métodos de control, así como en el diseño de políticas de promoción coherentes para la construcción de las plataformas dedicadas a la IA.

Sin embargo, todavía no existen organismos dedicados a la supervisión del buen uso de esta tecnología, ni las instituciones actuales se encuentran preparadas para dirimir los conflictos -sobre todo de tipo jurídico- que puedan surgir en este tipo de proyectos. Una vez más, el despliegue de las TIC y la innovación que puede generar en determinada materia lleva una velocidad superior al despliegue normativo que ha de acompañar a su uso posterior.

Referencias

- Alabau, A., & Guijarro, L. (2011). *La Política de Comunicaciones Electrónicas de la Unión Europea*. Valencia: Universidad Politécnica de Valencia.
- Alabau, A. (2004). *La Unión Europea y su Política para el Desarrollo de la Administración Electrónica: Tras los objetivos de la Estrategia de Lisboa*. Valencia: Fundación Vodafone España.
- Arenilla, M. (2021). *La Administración Digital*. Madrid: INAP (Instituto Nacional de Administración Pública).
- Bertranou, J. (2019). El seguimiento y la evaluación de políticas públicas: modelos disponibles para producir conocimiento y mejorar el desempeño de las intervenciones públicas. *MILLCAYAC - Revista Digital de Ciencias Sociales*.
- Casarrubios Blanco, E. (2023). Algunas notas sobre la innovación tecnológica y gestión de las tecnologías de la información y las comunicaciones en la Administración General del Estado (España). *Estudios/ Working Papers Nueva Serie Gigapp*, 248-255.
- Casarrubios Blanco, E. (2019). Las Administraciones Públicas 'en las nubes': tecnología sostenible y universal en un contexto de modernización y de innovación tecnológica. X Congreso de la GIGAPP, *Gobernando el futuro. Agenda 2030. Diez años trabajando para mejorar la gobernanza democrática en Iberoamérica*. Madrid: Gigapp.

- Casarrubios, E. (2007). Estrategias de la Unión Europea para el desarrollo de las políticas de promoción de la Sociedad de la Información. VIII Congreso de la Asociación Española de Licenciados en Ciencias Políticas. Valencia: AECPA.
- Comisión Europea. (2023). Inteligencia artificial y datos abiertos. Obtenido de <https://datos.gob.es>: https://datos.gob.es/sites/default/files/blog/file/infografia-es-va_2_2.docx. Visitado por última vez el 27/07/23.
- Comisión Europea. (2011). Una visión estratégica de las normas europeas: Avanzar para mejorar y acelerar el crecimiento sostenible de la economía europea de aquí a 2020. Bruselas: Comisión Europea.
- Comisión Europea. (2020). Libro blanco sobre la inteligencia artificial- un enfoque europeo orientado a la excelencia y la confianza. Obtenido de <https://ec.europa.eu>: <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52020DC0065>. Visitado por última vez el 27/07/23.
- Dirección General de Tráfico. (2023). <https://www.dgt.es/>. Obtenido de <https://www.dgt.es/>
- Fernández Hernández, C. (2022). La Ley 15/2022 introduce la primera regulación positiva de la inteligencia artificial en España Obtenido de Diario La Ley (<https://diariolaley.laleynext.es>): <https://diariolaley.laleynext.es/dll/2022/07/13/la-ley-15-2022-introduce-la-primera-regulacion-positiva-de-la-inteligencia-artificial-en-espana>. Visitado por última vez el 27/07/23.
- Gómez López, M. (2022). Big Data y gestión del dato en la Administración tributaria: pasado, presente y futuro. Revista de Economía Industrial.
- Marín Pérez, C. (2022). Tecnología Blockchain: origen, funcionamiento y usos. Trabajo de Fin de Grado. Zaragoza: Universidad de Zaragoza.
- Marzal, M. A. (2005). Las Políticas Supranacionales de Información: la Unión Europea. Salvador de Bahía: VI CINFORM, Encuentro Nacional Brasileño de la Ciencia de la Información.
- Mell, P., & Grance, T. (2011). The NIST Definition of Cloud Computing. NIST Special Publication 800-145.
- Mineco. (2020). Estrategia Nacional de IA (ENIA). Madrid: Vicepresidencia 3ª del Gobierno y Ministerio de Asuntos Económicos y Transformación Digital.
- Ministerio de Asuntos Económicos y Transformación; Digital., Secretaría General de Administración Digital. (s.f.). Obtenido de Portal de Administración Electrónica (PAE): https://administracionelectronica.gob.es/pae_Home
- Ministerio de Asuntos Económicos y Transformación; Digital., Secretaría General de Administración Digital. (s.f.). Obtenido de Observatorio de Administración Electrónica (OBSAE): <https://dataobsae.administracionelectronica.gob.es/>. Visitado por última vez el 27/07/23.
- Ministerio de Economía y Empresa. (2019). Plan presupuestario 2020, Reino de España. Madrid: Ministerio de Economía y Empresa, Departamento de Publicaciones.
- Moore, N. (2002). Neo- Liberal or Dirigiste? Policies for an Information Society. In Planificación y evaluación de políticas de Información.

Oficina de Publicaciones de la Unión Europea. (2002). Intercambio de datos entre Administraciones (IDA) en la Unión Europea.

Ortiz Bru, C. M. (s.f.). La política europea en materia de telecomunicaciones: hacia la nueva sociedad de la información. Obtenido de Noticias Jurídicas: <http://noticias.juridicas.com>

Parlamento Europeo. (2010). Normalización para una Europa competitiva e innovadora: una visión para 2020. Bruselas: Parlamento Europeo.

Sobre la autora/ About the author

Elena Casarrubios Blanco es Doctorando en Ciencias Políticas por la Universidad Rey Juan Carlos. Su tema de investigación es el proceso modernizador de las Administraciones públicas españolas, concretamente de la Administración General del Estado, y la influencia que las Tecnologías de la Información y las Comunicaciones han tenido en él. Este tema de investigación también comprende la influencia que la Inteligencia Artificial puede tener en el futuro en el proceso modernizador de las administraciones públicas españolas.

URL estable documento/stable URL

<http://www.gigapp.org>

El Grupo de Investigación en Gobierno, Administración y Políticas Públicas (GIGAPP) es una iniciativa impulsada por académicos, investigadores y profesores Iberoamericanos, cuyo principal propósito es contribuir al debate y la generación de nuevos conceptos, enfoques y marcos de análisis en las áreas de gobierno, gestión y políticas públicas, fomentando la creación de espacio de intercambio y colaboración permanente, y facilitando la construcción de redes y proyectos conjuntos sobre la base de actividades de docencia, investigación, asistencia técnica y extensión.

Las áreas de trabajo que constituyen los ejes principales del GIGAPP son:

1. Gobierno, instituciones y comportamiento político
2. Administración Pública
3. Políticas Públicas

Información de Contacto

Asociación GIGAPP.

ewp@gigapp.org

Adopción de sistemas de decisión automatizada en el sector público: Cartografía de 113 sistemas en Colombia

Gutiérrez, Juan David

Universidad de los Andes, Colombia

✉ juagutie@uniandes.edu.co

ORCID ID: [0000-0002-7783-4850](https://orcid.org/0000-0002-7783-4850)

Muñoz-Cadena, Sarah

Universidad del Rosario, Colombia

✉ sarahm.munoz@urosario.edu.co

ORCID ID: [0009-0003-8545-8036](https://orcid.org/0009-0003-8545-8036)

Documento recibido:

02 septiembre 2022

Aprobado para publicación:

07 agosto 2023

Resumen

La literatura sobre el uso de sistemas de decisión automatizada (SDA) en el sector público ha dado cuenta de su creciente adopción en América Latina. En el caso de Colombia, no existen trabajos que mapeen y caractericen de forma sistemática el uso de SDA por parte del Estado. Este artículo contribuye a cerrar esa brecha en la literatura a partir del análisis de una nueva base de datos sobre 113 SDA de entidades del sector público colombiano, de los cuales 97 se encuentran en funcionamiento (86%), 14 en pilotaje (12%) y dos más que fueron discontinuados. La base de datos caracteriza los sistemas a partir de 40 variables y fue creada con más de 300 fuentes de información pública. Este artículo describe las principales características, usos y potenciales beneficios de estos sistemas en Colombia. Identificamos sistemas implementados por 51 entidades públicas nacionales y subnacionales del poder ejecutivo, así como por entidades de la rama judicial y de organismos de control. El grueso de los 111 sistemas que están en etapa de pilotaje o en ejecución fueron adoptados por entidades públicas del orden nacional (74%) y la gran mayoría fueron implementados por entidades del poder público ejecutivo (93%). Los SDA son usados para realizar una amplia gama de funciones y para apoyar diferentes tipos de actividades estatales, pero más de la mitad de los sistemas son usados para apoyar actividades de los sectores justicia (18%), educación (17%), y medio ambiente (13%). Además, la mayoría de los SDA pueden aportar a la implementación de políticas públicas (el 98%), pero algunos también pueden contribuir con otras fases del ciclo de las políticas públicas (12% al agendamiento, 18% a formulación y 2% a evaluación). Finalmente, el artículo identifica potenciales riesgos asociados al uso de diferentes tipos de SDA por parte entidades públicas.

Palabras clave

sistemas de decisión automatizada, algoritmos públicos, inteligencia artificial, administración pública, Colombia

Resumo

A literatura sobre o uso de sistemas automatizados de decisão (ADS) no setor público tem relatado sua crescente adoção na América Latina. No caso da Colômbia, não existem trabalhos que mapeiem e caracterizem sistematicamente o uso de SDA pelo Estado. Este artigo contribui para preencher esta lacuna na literatura com base na análise de um novo banco de dados de 113 SDAs de entidades do setor público colombiano, dos quais 97 estão em operação (86%), 14 em fase piloto (12%) e mais dois que foram descontinuados. A base de dados caracteriza os sistemas com base em 40 variáveis e foi criada com mais de 300 fontes de informação pública. Este artigo descreve as principais características, usos e benefícios potenciais desses sistemas na Colômbia. Identificamos sistemas implementados por 51 entidades públicas nacionais e subnacionais do poder executivo, bem como por entidades do poder judiciário e órgãos de controle. A maior parte dos 111 sistemas que estão em fase piloto ou em execução foi adotada por entes públicos nacionais (74%) e a grande maioria foi implantada por entidades do poder público executivo (93%). Os SDAs são usados para desempenhar uma ampla gama de funções e para apoiar diferentes tipos de atividades do Estado, mas mais da metade dos sistemas são usados para apoiar atividades nos setores de justiça (18%), educação (17%) e mídia. (13%). Além disso, a maioria dos SDAs pode contribuir para a implementação de políticas públicas (98%), mas alguns também podem contribuir para outras fases do ciclo da política pública (12% para agendamento, 18% para formulação e 2% para avaliação.). Finalmente, o artigo identifica os riscos potenciais associados ao uso de diferentes tipos de SDA por entidades públicas.

Palavras-chave

sistemas de decisão automatizados, algoritmos públicos, inteligência artificial, administração pública, Colômbia.

Abstract

Literature review on the use of automated decision systems (ADS, SDA in spanish) in the public sector shows its growing adoption in Latin America. In the case of Colombia, there are no studies that systematically map and characterize the use of ADS by state entities and the national government. This article contributes to closing this gap in the literature based on the analysis of a new database on 113 SDAs of Colombian public sector entities, of which 97 are in operation (86%), 14 in piloting (12%), and two more that were discontinued. Database includes 40 variables and was created with more than 300 sources of public information. This article describes the main characteristics, uses and potential benefits of these systems in Colombia. We identified systems implemented by 51 national and subnational public entities

of the executive branch, as well as by entities of the judicial branch and control agencies. The bulk of the 111 systems that are in the pilot stage or in execution were adopted by national public entities (74%) and the vast majority were implemented by entities of the executive branch (93%). SDAs are used to perform a wide range of functions and to support different types of state activities, but more than half of the systems are used to support activities in the justice (18%), education (17%), and media sectors. environment (13%). In addition, most of the SDAs can contribute to the implementation of public policies (98%), but some can also contribute to other phases of the public policy cycle (12% to agenda setting, 18% to formulation, and 2% to evaluation stage). Finally, the article identifies potential risks associated with the use of different types of SDA by public entities.

Keywords

automated decision systems, public algorithms, artificial intelligence, public administration, Colombia

1. Introducción¹

Gobiernos de todo el mundo actualmente usan algoritmos para tareas sencillas, como procesar bases de datos de usuarios, y para realizar tareas complejas, como predecir la incidencia de enfermedades en una población determinada. La adopción de herramientas tecnológicas que incorporan algoritmos, especialmente aquellos que facilitan decisiones automatizadas, puede contribuir significativamente con los objetivos de los Estados. Por ejemplo, la implementación de sistemas asistidos por inteligencia artificial (IA) por parte de entidades públicas puede aumentar la velocidad y precisión del procesamiento de grandes volúmenes de datos; reducir los tiempos de respuesta de solicitudes y casos; y, eximir al talento humano de tareas rutinarias para reasignarlo a actividades más complejas, entre otros (Wirtz & Müller, 2019).

Sin embargo, el uso de algoritmos en la gestión pública también puede generar riesgos y efectos negativos para los ciudadanos. Wirtz & Müller (2019) argumentan que la implementación de IA por los gobiernos puede implicar la pérdida de predictibilidad de los procesos de la administración pública en la medida en que los sistemas arrojen resultados o decisiones que no sean comprensibles o cuyas justificaciones no sean verificables (el efecto "caja negra"); reducir o eliminar el control de los funcionarios sobre las herramientas que utilizan; y

¹ Los autores agradecen la diligente asistencia de investigación realizada por Michelle Castellanos-Sánchez, estudiante de la Facultad de Estudios Internacionales, Políticos y Urbanos de la Universidad del Rosario. Esta investigación fue financiada con recursos de la Universidad del Rosario y este artículo es una versión más reciente, completa y con datos actualizados de una ponencia que fue presentada por los autores en la XI Conferencia Internacional de Gobierno, Administración y Políticas Públicas 2022 (GIGAPP): ver Gutiérrez y Muñoz-Cadena (2022). Los autores también agradecen a los asistentes a esta conferencia, a la XVII Conferencia Anual de la Red Interamericana de Educación en Administración Pública (INPAE) y al *Section on International and Comparative Administration (SICA), 10th Annual Fred Riggs Symposium* (2023), por sus valiosas preguntas y retroalimentación.

generar riesgos en relación con el tratamiento de datos personales, particularmente aquellos sensibles, y pérdida o afectación de la privacidad, o vulneraciones a los derechos de las personas, entre otros. Por ejemplo, el uso de herramientas con IA en los procesos judiciales podría mantener o exacerbar prácticas discriminatorias de los jueces, reducir la transparencia de los procesos, y erosionar la rendición de cuentas por parte de la administración de justicia (Gutiérrez, 2020). Recientemente, distintas organizaciones y académicos han publicado casos de estudio sobre proyectos de IA en el sector público de Colombia, como es el caso de Cetina et al. (2021), Chenou & Rodríguez Valenzuela (2021), Escobar et al. (2021), Flórez & Vargas (2020), Gutiérrez (2020), y López & Castañeda (2020), Calderón-Valencia et al. (2021), López (2020), Agudelo Londoño (2022), Saavedra (2022), Castaño (2021), entre otros. Sin embargo, la literatura no estudia de manera sistemática la existencia y uso por parte de las entidades públicas colombianas de sistemas de IA u otros sistemas de algoritmos (ej. secuenciales, condicionales etc.) para la toma de decisiones automatizadas.

El objetivo central de esta investigación es identificar y caracterizar los sistemas de decisión automatizada (SDA) usados por el Estado Colombiano. Para tal efecto construimos una nueva base de datos sobre 113 SDA de 51 entidades del sector público colombiano, de los cuales 97 se encuentran en funcionamiento (86%), 14 en pilotaje (12%), uno está suspendido y uno fue discontinuado. (Gutiérrez et al., 2023).² La base de datos caracteriza los sistemas a partir de 40 variables y fue creada con más de 300 fuentes de información pública.

A partir de la base de datos, reportamos las principales características, usos, y potenciales beneficios de estos sistemas en Colombia. Además, caracterizamos el tipo de entidades públicas que adoptaron SDA y los aportes de los sistemas a las etapas del ciclo de las políticas públicas en los que participan dichas entidades. También identificamos potenciales riesgos asociados al uso de diferentes tipos de SDA por parte entidades públicas.

Esta investigación replica parcialmente, para el contexto colombiano, el estudio realizado conjuntamente por el GobLab de la Universidad Adolfo Ibáñez (UAI) y el Consejo para la Transparencia de Chile sobre el estado del arte de los SDA en el sector público chileno (Garrido et al., 2021) y el análisis del catastro de algoritmos públicos creado por el GobLab UAI (2022, 2023). Por otra parte, la investigación emula los esfuerzos realizados por entidades estatales que han construido y publicado repositorios de algoritmos públicos en diferentes ciudades, regiones y países (Gutiérrez, 2023d).³

Este artículo está dividido en cinco secciones, incluida esta introducción. La segunda sección describe la metodología empleada para realizar esta investigación y la siguiente sección explora brevemente las diferentes definiciones de SDA y presenta la literatura de las aplicaciones de SDA en Colombia. La cuarta sección describe y analiza los principales hallazgos identificados a partir de la nueva base de datos de 113 SDA adoptados por el sector público colombiano. La última sección discute las conclusiones de la investigación, esboza implicaciones de política y explica futuras avenidas de investigación asociadas a nuestro estudio de caso.

² La base de datos se puede descargar en el siguiente enlace: <https://doi.org/10.34848/YN1CRT>

³ Un listado de dichos repositorios de algoritmos públicos se puede consultar en el siguiente enlace: <https://forogpp.com/inteligencia-artificial-y-sector-publico/repositorios-y-registros-de-algoritmos/>

2. Metodología

Entre enero de 2022 y junio de 2023 construimos una nueva base de datos sobre SDA piloteados y/o usados por entidades públicas colombianas. Los datos recolectados, provienen de más de 300 fuentes primarias y secundarias que están disponibles públicamente. El principal tipo de fuente que consultamos fue información publicada por entidades públicas a través de repositorios públicos de datos, informes de gestión anual, comunicados de prensa, páginas web y redes sociales. Más concretamente, la base de datos se construyó a partir de 210 fuentes institucionales (68%).

Por otra parte, la base de datos también se nutrió a partir de información secundaria: 45 artículos de prensa (15%), 24 publicaciones académicas (8%), 11 documentos de organizaciones multilaterales (4%), nueve documentos de empresas (3%) y ocho publicaciones de la sociedad civil (2%).

La nueva base de datos documenta 113 SDA piloteados y/o usados por entidades públicas del orden nacional y subnacional y que hacen parte del poder ejecutivo, del poder judicial y de otros organismos autónomos. Esta incluye sistemas adoptados por ministerios, superintendencias, alcaldías, gobernaciones, órganos judiciales, y empresas de servicios públicos de propiedad estatal, entre otras.

La base de datos caracteriza los SDA a partir de 40 variables, las cuales se pueden agrupar en cinco categorías: (i) información básica sobre el SDA, dónde se incluye, nombre del sistema o proyecto, datos de la entidad pública que lo implementa, principales objetivos, estado del sistema, entre otros; (ii) preguntas sobre el tipo de información que requiere, entre las cuales están, si utiliza datos personales; (iii) información sobre el ejecutor y financiador o financiadores del proyecto, además de las cuantías y de donde provienen los recursos; (iv) clasificación del SDA según la función de gobierno de la entidad pública que los adopta, según el sector al que aporta, según el tipo de funciones de la herramienta, según la etapa o etapas del ciclo de políticas públicas al que contribuye, y potenciales aportes a los Objetivos de Desarrollo Sostenible (ODS); y, (v) información sobre las fuentes primarias y secundarias usadas para caracterizar cada sistema.

Una limitación importante de la nueva base de datos consiste en que documenta únicamente sistemas sobre los cuales hay información pública. A partir de nuestra investigación, constatamos que hay entidades del sector público colombiano que usan este tipo herramientas para desempeñar sus funciones pero que no informan públicamente. En particular, esto sucede respecto de sistemas usados para ejercer funciones en materia de seguridad y/o defensa nacional. La base de datos tampoco incluye sistemas respecto de los cuales, con la información disponible, no es posible corroborar o deducir que automatizan algún tipo de proceso o que apoyen o permitan adoptar decisiones de manera automatizada.⁴

Finalmente, la base de datos no incluye sistemas a los que pueden acceder cualquier usuario, público o privado, a través de Internet. Por ejemplo, no incluimos en la base de datos los modelos de lenguaje a gran escala

⁴ Siguiendo el modelo de repositorio construido por el GobLab UAI de Chile, también excluimos de la base de datos los "sistemas que se utilicen sólo para la generación, transmisión, manejo, procesamiento y almacenamiento de datos digitales. Sistemas que se utilicen para manipular o gestionar cantidades físicas o información que se encuentre representada en forma digital. Sistemas que solo sistematizan información a la ciudadanía como observatorios, portales de datos abiertos o que muestran estadísticas agregadas. Sistemas de gestión, transacciones, pago y planificación administrativa." (Garrido et al., 2021, p. 9)

que pueden utilizarse a través de *chatbots*, como *ChatGPT*, y que han sido usados por jueces y magistrados colombianos para adoptar decisiones (Gutiérrez, 2023b, 2023a, 2023c).

3. Marco teórico y revisión de literatura

3.1 Qué se entiende por sistema de decisión automatizada (SDA)

La forma cómo se define el conjunto de herramientas tecnológicas denominado SDA no es trivial para esta investigación por tres razones principales. Primera, las definiciones aunque se usan a diario “ya sean descriptivas o prescriptivas, tienen un efecto prescriptivo” (Price, 2012, p. 1017), trazan límites, por ejemplo, en el caso de las definiciones jurídicas estas “crean restricciones, [...] designan los contextos o circunstancias relevantes [...], contribuyen a proporcionar seguridad jurídica y uniformidad” (Richardson, 2022, pp. 790–791). Por ese motivo no es sorprendente que gran parte de los debates en torno al Proyecto de Ley de IA discutido en el proceso legislativo de la Unión Europea se centró en la definición de “sistemas de inteligencia artificial”.

Segunda, una definición va más allá de la escogencia de una secuencia de palabras. La forma como se define una palabra puede sugerir el modo particular cómo una sociedad o un grupo de personas comprende o interpreta un concepto. Además, la manera como las palabras se construyen socialmente puede tener implicaciones políticas. Por ejemplo, el significado que un formulador de política pública asigne a categorías usadas para describir problemas públicos puede influenciar el tipo de soluciones que posteriormente son aceptadas.

Tercera, los conceptos se construyen, pero también se reconstruyen con el tiempo; en este caso en particular “los contextos sociales e institucionales y la interacción hombre-máquina” (Araujo et al., 2020, p. 612) pueden llevar a reconsiderar aspectos de estos. Incluso el sentido de las definiciones establecidas por la ley no necesariamente es unívoco o estático: “Una definición legal puede –intencionalmente o no– estrechar el lente de un significado o puede expandirlo [...] Pero la definición cambia inevitablemente el significado –a veces sutilmente, a veces radicalmente– aunque intente reflejarlo con precisión” (Price, 2012, p. 1017). Esto ocurre, particularmente, cuando estamos ante definiciones relacionadas con asuntos tecnológicos, donde definiciones muy amplias o demasiado restringidas pueden traer consecuencias en el desarrollo futuro de estas nuevas tecnologías.

La primera advertencia semántica que consideramos pertinente respecto de los SDA es que no hay una única forma de definirlos ni hay consenso en la literatura sobre su significado. Para ilustrar la variedad de acepciones, en el Anexo 1 incluimos un listado de diez diferentes definiciones de SDA que han sido usadas por la literatura académica, la literatura gris o instrumentos gubernamentales.

Por una parte, hay definiciones genéricas, como la establecida por el Gobierno de Canadá que define los SDA como “[c]ualquier tecnología que asiste o sustituye el juicio de los responsables humanos de la toma de decisiones” (Directive on Automated Decision-Making, 2023). Otras definiciones precisan aspectos de los SDA como el tipo de tecnologías que usan (ej. *software*, *hardware*); tipo de técnica del algoritmo (basados en reglas, aprendizaje automatizado); tipo de soporte o apoyo que prestan (ej. predicción, clasificación, identificación etc.); grado de autonomía con el cual opera el sistema (ej. total, parcial); y, tipo de afectación que puede generar a la población (ej. discriminación). Además, es posible distinguir entre sistemas que toman decisiones (sistemas automatizados) y sistemas que ofrecen soporte para la toma de decisiones (sistemas semiautomatizados) (Garrido et al., 2021, p. 42).

La definición de trabajo que usaremos en este artículo para los SDA será la siguiente: *sistemas computacionales basados en datos que toman decisiones o apoyan los procesos de toma de decisiones de entidades públicas*.

Para cerrar esta reflexión semántica, vale la pena recalcar que algunos SDA operan a partir de IA⁵ pero que no todo SDA requiere de IA para funcionar pues la automatización puede operar a partir de algoritmos secuenciales (reglas de negocio que establecen secuencias de órdenes). Por tanto, las tecnologías que componen los SDA son diversas y no se limitan a aquellas que se consideran IA, sino que también incluyen otros sistemas algorítmicos como la automatización robótica de procesos.

3.2 Literatura sobre SDA en el sector público

El uso de algoritmos en el sector público se ha expandido bajo la premisa de que estas herramientas permiten mejorar la prestación de servicios del Estado: a través de procesos o asignación de recursos más eficiente, la equidad en la gobernanza algorítmica o la precisión en los pronósticos (Chenou & Rodríguez Valenzuela, 2021; Zuiderwijk et al., 2021). El *boom* de la aplicación de sistemas basados en IA en el sector público, la creciente disponibilidad de datos y la innovación tecnológica han estado acompañadas de narrativas optimistas en torno a la adopción de estas herramientas (*techno-optimistic discourses*). Dichas narrativas señalan que estos sistemas pueden contribuir a generar entidades gubernamentales más eficientes, veloces y objetivas (Chenou & Rodríguez Valenzuela, 2021). Este mensaje puede ser atractivo para algunos Estados, en particular, en América Latina, donde “las instituciones se asocian tradicionalmente con la parcialidad, la lentitud, [...] la falta de eficacia” o el excesivo papeleo (Chenou & Rodríguez Valenzuela, 2021, p. 60).

En la medida en que los gobiernos han expandido las iniciativas basadas en IA, también se han incrementado las alarmas con respecto a los riesgos, la opacidad en los algoritmos, las implicaciones y los alcances éticos y jurídicos, en particular, los asociados a la protección a la privacidad, la discriminación, el derecho al debido proceso, entre otros (Camacho Gutiérrez & Saavedra Rionda, 2021; Chenou & Rodríguez Valenzuela, 2021; Gutiérrez, 2020; Kuziemski & Misuraca, 2020; Zuiderwijk et al., 2021)

El Estado, en la actualidad, tiene un doble rol como regulador y a la vez promotor y consumidor de IA. En palabras de Kuziemski y Misuraca (2020, p. 1), “un trágico doble vínculo porque sus obligaciones de proteger a los ciudadanos de posibles daños algorítmicos están en conflicto con la tentación de aumentar su propia eficiencia, es decir, de gobernar algoritmos, mientras gobiernan gracias a algoritmos”. Si bien la investigación académica sobre IA y otras tecnologías emergentes no es nueva (Desouza et al., 2020; Zuiderwijk et al., 2021) y se ha incrementado de forma exponencial en las últimas dos décadas,⁶ existe una brecha en la producción de conocimiento acerca de la adopción de estos sistemas en el sector público, los resultados de estas intervenciones, las afectaciones a los ciudadanos y sus limitaciones (Camacho Gutiérrez & Saavedra Rionda, 2021; Reis et al., 2019; Sousa et al., 2019; Zuiderwijk et al., 2021). De hecho, en la revisión de literatura de Sousa et al. (2019), se encontró que, entre 2000 y octubre de 2018, de 1438 artículos académicos publicados sobre es-

⁵ En este artículo se entenderán los sistemas de IA como sistemas computacionales o basados en máquinas, que funcionan con “diversos niveles de autonomía”, y que pueden, “para un conjunto determinado de objetivos definidos por el ser humano, hacer predicciones, recomendaciones o tomar decisiones que influyen en entornos reales o virtuales” (Berryhill et al., 2019, p. 18).

⁶ “El número de publicaciones académicas que hacen foco en la IA se ha multiplicado por nueve desde el año 1996 a la fecha, lo que resulta un incremento anormal comparado con publicaciones en otras ramas científicas” (Estevadeordal et al., 2018, p. 210).

tas tecnologías, solo 59 (4%) se referían a IA aplicada al sector público; 36 de los cuales (61%) fueron publicados entre 2014-2018.

Los anteriores hallazgos evidencian un interés creciente en conocer acerca de las aplicaciones de algoritmos en el sector público. Sin embargo, el grueso de la literatura sobre la implementación de IA por parte de los gobiernos se concentra en analizar casos en el Norte Global (Camacho Gutiérrez & Saavedra Rionda, 2021; Chenou & Rodríguez Valenzuela, 2021). En ese sentido, este artículo pretende contribuir con la literatura mediante el estudio del uso de SDA por parte del sector público en un país del Sur Global.

3.3. Literatura sobre SDA en el sector público colombiano

La literatura sobre el uso de SDA en el sector público colombiano es incipiente y se concentra en casos de estudio sobre herramientas puntuales y la mayoría se enfoca en discutir cuáles son las implicaciones éticas, jurídicas y/o regulatorias de su uso. Por ejemplo, en los artículos de Camacho Gutiérrez et al., (2019), Gutiérrez-Ossa & Flórez Hernández (2020), Gutiérrez (2020, 2023b, 2023a, 2023c), se discuten los desafíos, implicaciones y riesgos de la utilización de herramientas específicas de IA en el sector judicial colombiano.⁷ Entretanto, Flórez & Vargas (2020), analizan “algunas implicaciones legales, regulatorias y éticas de la inteligencia artificial dentro del sector público en Colombia” a partir de cuatro casos: Fiscal Watson (Fiscalía General de la Nación), Prometea (Corte Constitucional), SOFÍA (Dirección de Impuestos y Aduanas Nacionales DIAN) y OCEANO (Contraloría General de la República). Los autores concluyen que ante la “fascinación” por implementar sistemas de IA en el gobierno, “no solo [se] debe contar con mecanismos de auditoría y de control, sino que también debe[n] implementarse a través de etapas de evaluación” (2020, p. 35).

A estos casos de estudio se suman informes producidos por diferentes entidades multilaterales que promueven que los gobiernos creen condiciones habilitantes para el desarrollo y uso responsable de sistemas de IA. Este es el caso de los estudios publicados por CAF (2021; 2022) y la OCDE (2017), en los cuales se resalta cómo Colombia ha avanzado en la creación e implementación de la política de IA y algunas de sus aplicaciones prácticas. En los documentos publicados por el CAF también se describe y analiza la experiencia de dos entidades públicas: el análisis de redes criminales en el contexto de la Procuraduría General de la Nación o ARCPGN (Cetina et al., 2021) y la experiencia de la Superintendencia de Sociedades con SIARELIS (Ortega Rance, 2021).

Finalmente, algunos investigadores han propuesto posibles aplicaciones de IA que podrían utilizarse en el sector público colombiano. Por ejemplo, para predecir posibles casos de corrupción en la contratación estatal (Gallego et al., 2021, 2022), para detectar corrupción en la administración pública municipal (Mojica Muñoz, 2021) o para mejorar el sistema de asignación de esquemas de protección para defensores de derechos humanos en Colombia (Mojica Muñoz et al., 2021). En síntesis, la literatura sobre el uso algoritmos públicos en Colombia incluye casos de estudio puntuales que describen dichos sistemas e investigaciones sobre las implicaciones éticas y legales de la adopción por parte del Estado de algunas de estas tecnologías. Sin embargo, la

⁷ Un artículo de Fedesarrollo también trata sobre la aplicación de IA en el sector judicial posterior a la digitalización de los expedientes judiciales; sin embargo, a diferencia de los antes mencionados, este artículo tiende a una visión más “optimista” que crítica del uso de IA, aunque reconoce que “las sentencias y otras decisiones importantes en la justicia deben ser tomadas por los jueces, no por algoritmos” (Cepeda Espinoza & Otálora Lozano, 2020, p. 5). Entretanto, en el artículo de Castaño (2021) se realiza un análisis de qué tan preparado está el sector judicial colombiano para el uso de IA, cuál es el marco regulatorio y cuáles son algunos ejemplos de herramientas de este tipo que ya se están usando.

literatura no cuenta con trabajos que mapeen y caractericen de manera sistemática la adopción de SDA por el Estado colombiano. En este sentido, esta investigación pretende contribuir al conocimiento sobre el uso de estos sistemas en el sector público en Colombia mediante la creación y análisis de una nueva base de datos de sistemas implementados en el país.

4. SDA piloteados y/o adoptados por el Estado colombiano

En esta sección documentamos brevemente los esfuerzos del gobierno nacional de Colombia por informar a través de repositorios digitales sobre la adopción de sistemas de IA por parte entidades públicas y, posteriormente, presentamos en detalle los hallazgos derivados de nuestra nueva base de datos sobre SDA en el sector público colombiano.

4.1 Repositorios de algoritmos públicos del gobierno nacional

A julio de 2023, el gobierno nacional de Colombia cuenta con tres repositorios de acceso público con información sobre sistemas de IA utilizados por entidades públicas (Tabla 1). Los repositorios son administrados por el Departamento Administrativo de la Presidencia (DAPRE) y por el Ministerio de Tecnologías de la Información y las Comunicaciones (MinTIC). Cada uno contiene diferentes tipos de datos sobre cada sistema reportado. Los repositorios no están interconectados y tienen algunas duplicidades en cuanto a los sistemas registrados en cada uno.

Tabla 1. Repositorios de algoritmos públicos de Colombia

Nombre del repositorio	Entidad a cargo	No. de sistemas registrados ⁸
Dashboard de seguimiento del Marco Ético de Inteligencia Artificial ⁹	DAPRE	6
Datos Abiertos	MinTIC	16
Ejercicios de Innovación Basados en Inteligencia Artificial	MinTIC	4

Fuente: elaboración propia

El número de sistemas que utiliza el Estado colombiano es mucho mayor al registrado en estos repositorios. Este sub-registro quedó constatado con nuestra nueva base de datos, pero antes de construirla teníamos indicios acerca dicha situación. La principal pista sobre dicho sub-registro se hallaba en las respuestas de las entidades públicas que diligenciaron el Formulario Único de Reporte de Avances de Gestión (FURAG).¹⁰ En el 2021, el FURAG fue diligenciado por 2,939 entidades públicas y de estas un total de 233 (8%) contestaron que

⁸ Con corte a 02/09/2022.

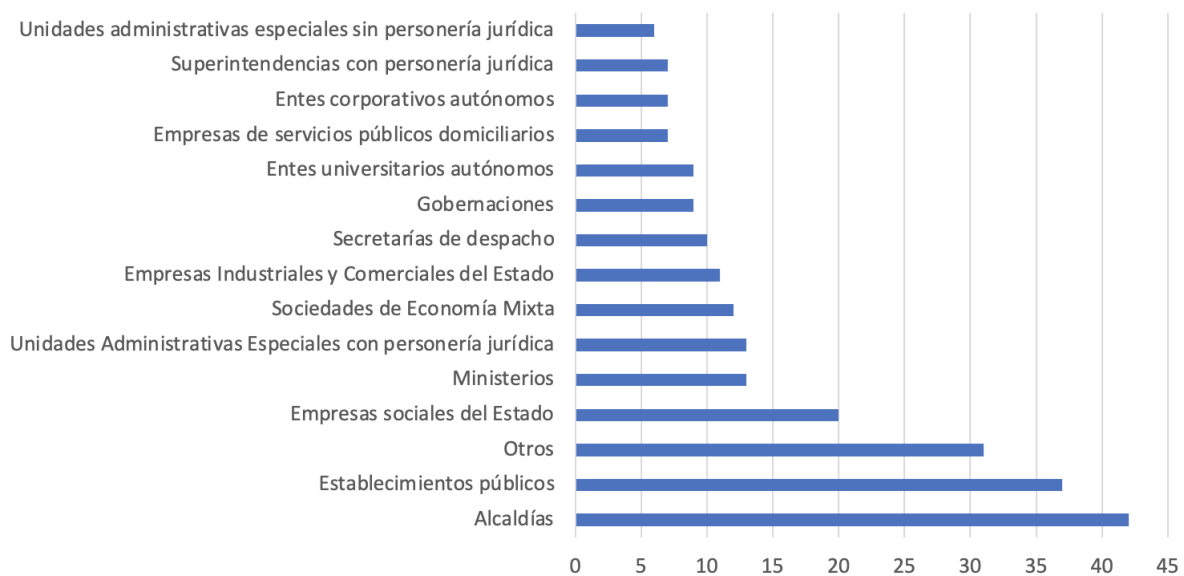
⁹ Desde mayo de 2023 dejó de estar disponible la información sobre estos sistemas que podía accederse, hasta esa fecha, a través de la URL <https://inteligenciaartificial.gov.co/dashboard-IA/>

¹⁰ La información sobre el tipo de preguntas incluidas en el FURAG y la metodología de recolección está disponible en la siguiente plataforma del DAFP: https://www.funcionpublica.gov.co/web/mipg/medicion_desempeno

utilizaban sistemas de IA y/o automatización robótica de procesos (RPA, por su siglas en inglés).¹¹ Más precisamente 172 (74%) reportaron que usaban IA, 116 (50%) que usaban RPA, y 55 (24%) que usaban los dos tipos de tecnologías.¹²

Con respecto al tipo de entidad que reportó el uso de IA y/o RPA, 108 (46%) son del orden nacional y 125 (54%) del orden territorial. En las entidades de orden territorial, 44 (35%) están ubicadas en las tres principales ciudades del país: 25 (20%) están en Bogotá, 12 (9,6%) en Medellín, 7 (5,6%) en Cali. En lo referente a la naturaleza jurídica de las entidades, como se puede observar en la figura 1, el listado está liderado por las alcaldías (42), los establecimientos públicos (37), empresas sociales del estado (20) y ministerios (13).

Figura 1. Naturaleza jurídica de las entidades que respondieron que utilizaban IA o RPA según la encuesta del FURAG¹³



Fuente: elaboración propia

Sin embargo, es posible que algunas de las entidades públicas que respondieron afirmativamente sobre el uso sistemas de IA y/o RPA, realmente no usen este tipo sistemas. Además de que los términos IA y RPA no son usados comúnmente por los funcionarios públicos, el DAFP, que administra el FURAG, no incluyó en su más

¹¹ En este documento se entiende por automatización robótica de procesos (RPA, por sus siglas en inglés) "la aplicación de tecnología y metodologías específicas que se basan en software y algoritmos con el objetivo de automatizar tareas humanas repetitivas [16, 21, 33, 39]. Se basa principalmente en reglas simples y lógica de negocios mientras interactúa con múltiples sistemas de información a través de las interfaces gráficas de usuario existentes [17]. Sus funcionalidades comprenden la automatización de actividades repetibles y basadas en reglas mediante el uso de un robot de software no invasivo, denominado "bot" [27, 29, 38]." (Ivančić et al., 2019, p. 2)

¹² Los datos de las respuestas al FURAG pueden consultarse en la siguiente plataforma: <https://www.datos.gov.co/Funcion-p-blica/FURAG/daed-z4fw/data>

¹³ En la categoría "Otros" se incluyen, entre otros, agencias estatales de naturaleza especial y departamentos administrativos.

reciente glosario estas definiciones para contribuir a que los funcionarios que contestaran contaran con un entendimiento compartido de dichas tecnologías.¹⁴

Para subsanar la brecha de información identificada en los repositorios de algoritmos públicos, construimos la nueva base de datos que caracteriza 113 SDA piloteados y/o adoptados por entidades públicas en Colombia.

4.2 Hallazgos en la nueva base de datos sobre SDA en el sector público colombiano

En la base de datos se registraron 113 SDA del sector público, de las cuales 97 se encuentran en funcionamiento (86%), 14 en pilotaje (12%), uno está suspendido¹⁵ y uno fue discontinuado.¹⁶ Dado que este artículo se concentra en los sistemas que están en fase de pilotaje o en funcionamiento, las estadísticas que se presentan a continuación incluyen 111 sistemas, es decir, excluyen el sistema que está suspendido y el discontinuado.

En lo que resta de esa sección caracterizaremos el tipo de entidades públicas que han adoptado SDA, clasificaremos los sistemas según la línea de sector público al que contribuyen, según el tipo de funciones que realizan y según las etapas del ciclo de políticas públicas al que aportan, y reflexionamos sobre los potenciales riesgos asociados al uso de los SDA mapeados.

4.2.1 Caracterización de las entidades públicas que adoptaron SDA

Las entidades públicas mapeadas en esta investigación que están usando SDA en Colombia son, en su gran mayoría, del poder ejecutivo (93%) y, en menor medida, del poder judicial (4%)¹⁷ o pertenecen a las entidades de control (3%). No encontramos ningún sistema adoptado por el poder legislativo.

En la base de datos se registra que 51 entidades públicas diferentes adoptaron SDA.¹⁸ De los 111 sistemas (ver figura 2), identificamos que 82 (74%) fueron piloteados o implementados por entidades del orden nacional y

¹⁴ El glosario de junio de 2020 (versión 5) del FURAG puede consultarse en el siguiente URL:

https://www.funcionpublica.gov.co/documents/28587410/36200637/Glosario_mipg.pdf/9ff42c08-61a9-e0fa-76b1-1f662cob2202?t=1593207412671

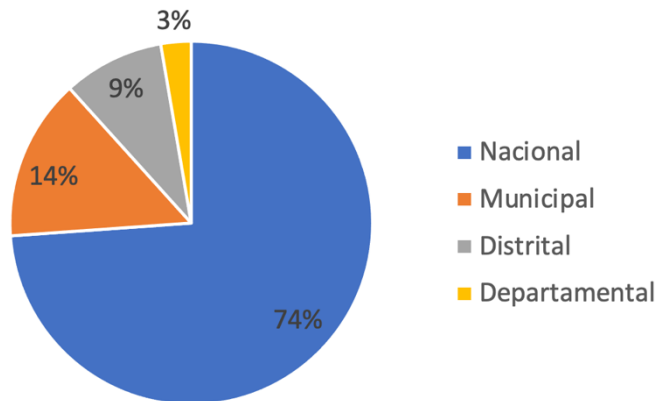
¹⁵ El proyecto “Piloto de alumbrado público Inteligente – Calinteligente” de la Alcaldía de Santiago de Cali, está suspendido porque el Concejo de Cali no aprobó que se creara una empresa mixta para implementar el alumbrado alegando que quien debería asumir dicha responsabilidad era las Empresas Municipales de Cali, Emcali (Noticiero 90 Minutos, 2021).

¹⁶ En agosto de 2020, información publicada sobre el proyecto “Sistema de ciberinteligencia basado en inteligencia artificial para la dirección de inteligencia policial de la Policía Nacional” generó polémica en redes sociales por el proceso que inició la Policía Nacional para la adquisición de un sistema cuyo objetivo era “la obtención de información personal publicada por los usuarios del ciberespacio en páginas web y en redes sociales. La herramienta tecnológica, basada en el uso de algoritmos, consolidaría bases de datos de los ciudadanos para detectar posibles amenazas cibernéticas que atenten contra la Defensa y Seguridad Nacional” (Procuraduría General de la Nación, 2020, p. 1). A raíz de esta situación y del informe de la Procuraduría General de la Nación (2020) se discontinuó el proceso.

¹⁷ Tres de los SDA de la rama judicial fueron implementados por la Fiscalía General de la Nación (FGN): (i) “Watson”, que permite “esculcar denuncias, correlacionarlas y hacer análisis de contexto sobre elementos similares, como el modus operandi, características físicas, tipos de armas, números telefónicos y vehículos, entre otros” (CAF, 2021, p. 244). (ii) El “Perfil de Riesgo de Recurrencia de la Solicitud de Medidas Penitenciarias (Prisma)” que se creó con el propósito de “predecir el riesgo de reincidencia criminal y así apoyar las decisiones de detención preventiva por parte de los fiscales” (CAF, 2021, p. 448). (iii) La “asignación automática de casos a fiscales” que, adicionalmente, también les recomienda actuaciones iniciales a los fiscales (Centro de Innovación Pública Digital, 2022, 2020; Fiscalía General de la Nación, 2019).

29 (26%) del orden territorial (municipal, distrital o departamental).¹⁹ Por ejemplo, uno de los sistemas utilizado por una entidad del orden nacional es el Asistente de Adopciones (ADA) del Instituto Colombiano de Bienestar Familiar (ICBF). Este tiene por objetivo, entre otros, “reducir de 24 a 9 meses los trámites administrativos para la adopción de niños niñas y adolescentes”, y facilitar “a las familias y al equipo psicosocial el proceso” aportándoles, por ejemplo, información en tiempo real sobre el estado de sus solicitudes (ICBF, 2020).²⁰

Figura 2. Porcentaje de sistemas adoptados por entidades del orden nacional o territorial



Fuente: Elaboración propia

En el caso de las entidades del orden territorial, estas se desagregan en iniciativas lideradas por entidades municipales (14%), distritales (9%) y departamentales (3%). Entre los sistemas que implementan estas entidades destacan los de monitoreo ambiental (proyectos en Bogotá²¹, Cali²² y Barrancabermeja²³) y los relacio-

¹⁸ El Servicio Nacional de Aprendizaje (SENA) es la entidad gubernamental del nivel nacional con más SDA en nuestra base de datos (13 sistemas). Los sistemas usados por el SENA pueden dividirse en tres tipos de funciones: (i) mejorar procedimientos internos y rutinarios de la entidad, como, por ejemplo, “PILABOT” que se encarga de automatizar el proceso de cargar las planillas relacionadas con el pago de los aportes de seguridad al Sistema de la Protección Social de los empleados de la entidad; (ii) conectar a los aprendices o egresados de sus programas con las oportunidades laborales que ofrece el mercado, por ejemplo, el “RPA – cerebro de competencias” que tiene por objetivo realizar una búsqueda de perfiles más eficiente e inclusiva, que permita incrementar “la empleabilidad como resultado de búsquedas más asertivas” (SENA, 2021, p. 11); (iii) *chatbots* para acompañar las inquietudes de aspirantes, aprendices y terceros interesados en sus proyectos, tal es el caso del “Chatbot formación” que ofrece “orientación vocacional a los aprendices frente a los programas formativos del SENA, de acuerdo con sus intereses y con su perfil, apoyados en la gestión del conocimiento” (SENA, 2021, p. 12).

¹⁹ En el caso de los SDA usados por diferentes Secretarías de la Alcaldía Mayor de Bogotá, en estas estadísticas se consideran como una sola entidad.

²⁰ La reducción de los tiempos, según explica el ICBF, “solo incluye el proceso de preparación y evaluación hasta que el comité de opciones determine la idoneidad de la familia” (ICBF, 2020).

²¹ “Red de Monitoreo de Calidad del Aire de Bogotá – RMCAB” y el “Sistema Integrado de Modelación de Calidad de Aire de Bogotá (SIMCAB)”.

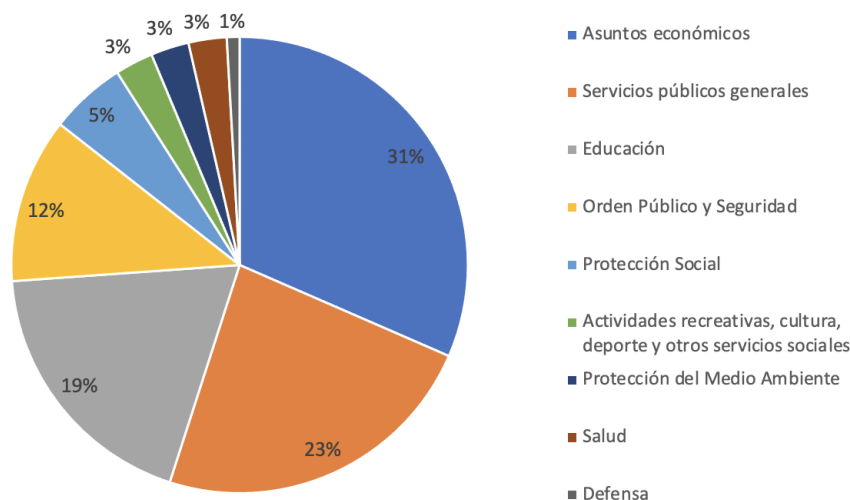
²² Monitoreo Ambiental – Calinteligente.

²³ El “Sistema integral de vigilancia de la calidad de aire para el Distrito de Barrancabermeja” y el “Piloto de estaciones de calidad del aire y meteorológicas de Barrancabermeja” (Santander).

nados con “alumbrado público inteligente” (Bogotá²⁴ y Tunja²⁵). Desde 2021, en el marco de la Política de Gobierno Digital y, más precisamente, de la estrategia de “Ciudades y Territorios Inteligentes”, el MinTIC realiza “acompañamiento especializado” y “asesoría consultiva” con entidades del orden territorial sobre cómo implementar este tipo de iniciativas.²⁶

Aplicando la Clasificación de las Funciones del gobierno (COFOG),²⁷ encontramos que el 74% de los sistemas mapeados fueron adoptados por entidades públicas que realizan tres categorías de funciones: “asuntos económicos” (32%), “servicios públicos generales” (23%) y “educación” (19%) (ver figura 3).

Figura 3. Clasificación de los SDA a partir del primer nivel de COFOG



Fuente: Elaboración propia

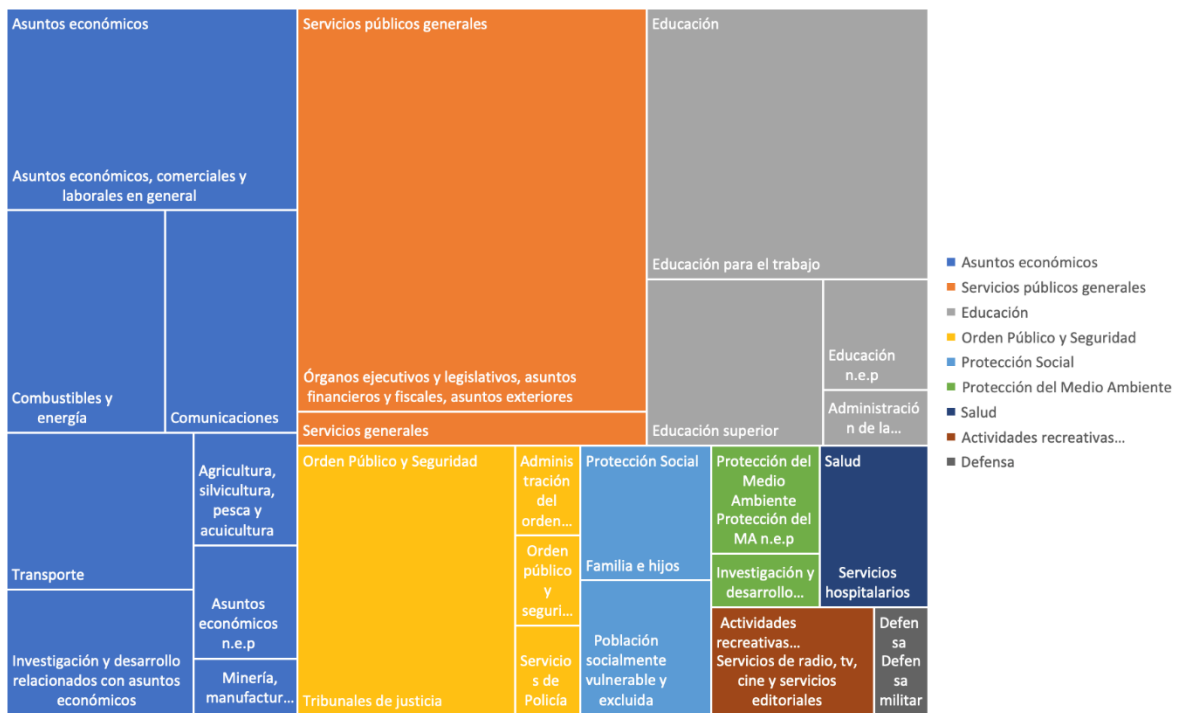
Figura 4. Clasificación de los SDA a partir del segundo nivel de COFOG

²⁴ Piloto de alumbrado público inteligente en Bogotá a cargo de la Unidad Administrativa Especial de Servicios Públicos (Uaesp).

²⁵ “Piloto de iluminación inteligente – Tunja (Boyacá)”.

²⁶ El MinTIC cuenta con 12 plantillas sobre proyectos tipo que las entidades pueden utilizar para formular sus propias iniciativas (MinTIC, 2021), entre los que están los de monitoreo ambiental y el de alumbrado, así que es posible que exista una relación entre la disponibilidad de un modelo/plantilla estructurada y que algunas ciudades los estén llevando a sus territorios.

²⁷ “La Clasificación de las funciones del gobierno (COFOG) es una clasificación detallada de las funciones y los objetivos socioeconómicos que persiguen las unidades del gobierno general por medio de distintos tipos de gasto. Permite identificar el gasto que realiza el Gobierno de acuerdo con las finalidades o las funciones públicas, mostrando la naturaleza de los servicios que prestan las instituciones en nombre del Estado” (DANE, 2020, párr. 1).



Fuente: Elaboración propia

Este hallazgo para el caso colombiano, contrasta con el reportado en el Repositorio de Algoritmos Públicos del GobLab UAI: en el caso chileno de los 75 algoritmos que están registrados (GobLab UAI, 2023), el 61% está relacionados con tres sectores: “salud” (25%), “asuntos económicos” (24%) y “orden público y seguridad” (12%). Por otro lado, en un mapeo de sistemas de IA utilizados por entidades públicas en la Unión Europea, en el cual también se aplica COFOG, se encontró que las tres categorías con más algoritmos mapeados eran: “servicios públicos generales”, “asuntos económicos”, y “orden público y seguridad” (Misuraca et al., 2020, p. 95).

Retornando a nuestra base de datos, la COFOG permite detallar el tipo de función realizada por las entidades públicas que adoptaron SDA a partir de un segundo nivel de categorización, lo cual permite clasificar las entidades a partir de funciones de gobierno más detalladas, como lo ilustra la figura 4

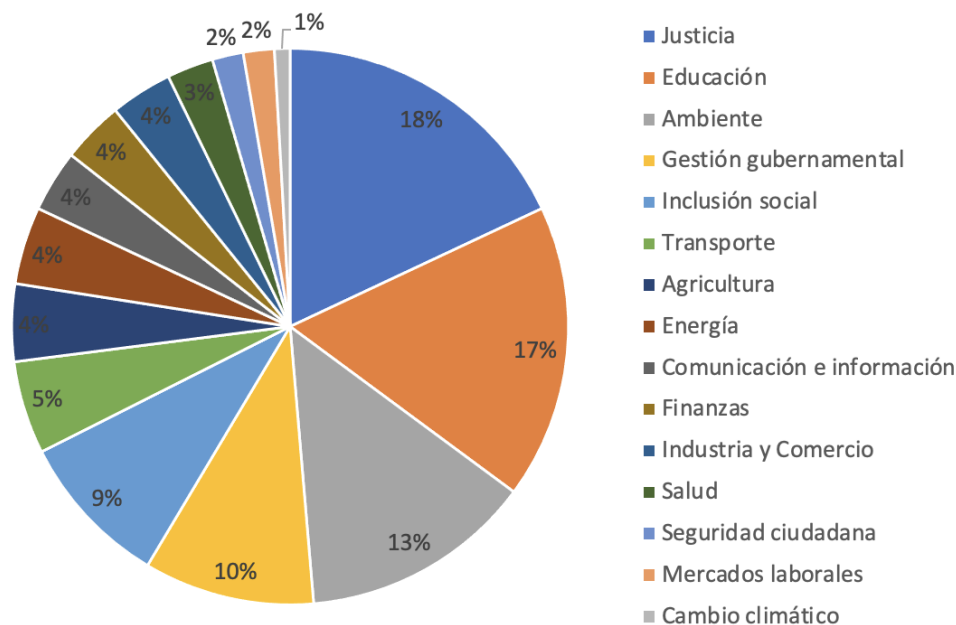
4.2.2 Clasificación de los SDA según la actividad estatal a la que contribuye el proyecto

Aunque COFOG es útil por ser una clasificación estandarizada y usada por diferentes países para reportar sobre las funciones que persiguen las unidades organizacionales de los gobiernos, lo cual facilita la comparación entre jurisdicciones, su principal limitación para esta investigación es que no informa con precisión el tipo de actividad estatal a la cual contribuye cada SDA. Lo anterior, por cuanto las entidades públicas pueden realizar funciones que podrían ser ubicadas en más de una categoría COFOG. Por ejemplo, una entidad pública ubicada en la categoría “asuntos económicos”, como la Superintendencia de Industria de Comercio, puede realizar funciones propias de la rama judicial.

Por lo tanto, para complementar la caracterización de los SDA, utilizamos las categorías utilizadas por el Banco Interamericano de Desarrollo (BID) en el marco de la iniciativa fAIR LAC. Concretamente, el Observatorio

de fAlr LAC desarrolló una clasificación de 18 sectores a los que pueden aportar las iniciativas de IA en América Latina.²⁸

Figura 5. Clasificación de los SDA por tipo de sector según clasificación del Observatorio fAlr LAC del BID



Fuente: Elaboración propia

Con respecto a qué líneas del sector público aportan los SDA bajo la clasificación fAlr LAC, encontramos que 75 de los 111 SDA (68%) están concentrados en cinco categorías: 20 (18%) están relacionados con justicia, 19 (17%) con educación, 15 (13%) con medio ambiente, 11 (10%) con gestión gubernamental, y 10 (9%) con inclusión social (ver figura 5).

Por ejemplo, en la categoría de justicia se encuentra el sistema de la Agencia Nacional de Defensa Jurídica del Estado (ANDJE) denominado “Predicción de sentencias en juicios contra el Estado (Colombia) – Modelo de pronóstico de probabilidad de pérdida”, el cual fue desarrollado por esta entidad y Quantil (una empresa privada) con el objetivo de “estimar la probabilidad de una sentencia desfavorable en un proceso litigioso contra la nación, y recomendar el monto óptimo de un acuerdo basado en las condiciones vigentes del caso” (OECD, 2022b, tbl. 3.3).

En el caso de los sistemas relacionados con educación, identificamos SDA implementados por el Ministerio de Educación Nacional, el Instituto Colombiano para la Evaluación de la Educación (ICFES), el SENA, y el Instituto Colombiano de Crédito Educativo y Estudios Técnicos en el Exterior (ICETEX).

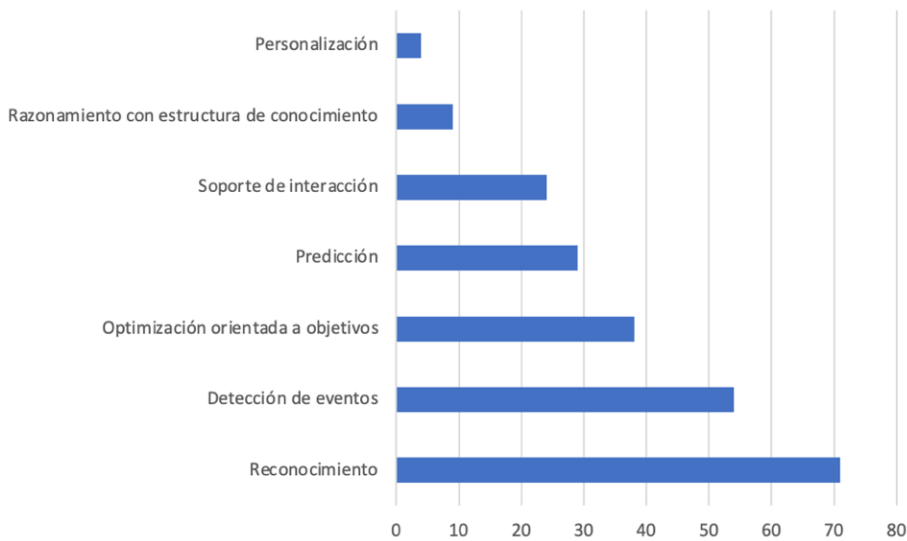
²⁸ El Observatorio fAlr LAC está disponible en el siguiente URL <https://fairlac.iadb.org/observatorio> Si bien utilizamos las categorías del Observatorio fAlr LAC, no tenemos registrado ningún sistema en cuatro categorías: acuicultura, género o diversidad, perspectiva y protección de datos personales. Entretanto, cabe aclarar que incluimos una categoría adicional, que no estaba contemplada en la clasificación del BID: transporte.

4.2.3 Clasificación de los SDA por el tipo de funciones de la herramienta

Para clasificar los SDA según el tipo de función que desarrolla cada herramienta aplicamos la clasificación de la Organización para la Cooperación y el Desarrollo Económicos (OCDE) que identifica siete clases según el tipo de *output* que genera la IA (OECD, 2022a). Es pertinente mencionar que usualmente los sistemas pueden realizar más de una función.

Encontramos que 71 de los 111 SDA (64%) realizan funciones de reconocimiento, 54 (49%) detectan eventos, 38 (34%) se concentran en optimización enfocada en objetivos, 29 (26%) hacen predicción de eventos, 24 (22%) prestan soporte de interacción (en particular, *chatbots*) nueve (8%) realizan razonamiento con estructura de conocimiento y cuatro (4%) buscan personalizar (ver figura 6).

Figura 6. Clasificación de los SDA en Colombia por tipo de funcionalidades



Fuente: Elaboración propia

Por ejemplo, uno de los sistemas que realizan funciones de reconocimiento es el de “Análisis de similitud para identificar procesos con características similares” de la ANDJE, el cual es un “sistema que permite relacionar los casos versus sentencias de unificación de las altas cortes” (El Tiempo, 2022, párr. 5) para ofrecerle a los abogados “un listado de las sentencias textualmente más similares a su caso” (Martínez, 2022, p. 20).

Hay sistemas que pueden realizar más de una función, por ejemplo, “Guardianes de la selva” es un sistema del Instituto de Investigación de Recursos Biológicos Alexander von Humboldt (en asocio con la Fundación Biodiversa Colombia, Rainforest Connection, Corantioquia y Huawei) que busca –a través de sensores que detectan sonidos– reconocer e identificar posibles eventos que dañen los ecosistemas, por ejemplo, deforestación o caza furtiva (Instituto Humboldt, 2023).

Uno de los sistemas que realizan la función de optimización orientada a objetivos es el de “Pilotos semáforos inteligentes” de Armenia (iniciativa desarrollada en el marco de ciudades y territorios inteligentes del MinTIC), el cual “contempla la instalación de modernos semáforos en hasta dos intersecciones viales”, que “funcionarán con un software de data y con cámaras y sensores para medir, entre otros factores, velocidad, número de vehículos y peatones que transitan por el sector en determinado horario” (Alcaldía de Armenia, 2022). En este sentido, se infiere que este sistema además de reconocer patrones de movilidad, también pretende una opti-

mización orientada a objetivos en la medida en que dicha información recolectada “ayudará a mejorar a la hora de tomar decisiones para elevar la calidad de vida de la población y en este caso específico en relación con la movilidad” (Nueva Crónica Quindío, 2022, párr. 4).

Un caso de un sistema a través del cual se busca predecir eventos futuros a partir de datos es el sistema de “Predicción de los pasajeros que usan el Cable Aéreo de Manizales” liderado por la Asociación Cable Aéreo Manizales que “permite predecir o acortar los tiempos de mantenimiento del sistema dependiendo de ese flujo [de pasajeros]” y se espera que “mejor[e] la experiencia del usuario ya que si el modelo predice que hay un alto flujo de pasajeros, el sistema del cable aéreo de Manizales puede aumentar la velocidad de las góndolas para que el usuario esté menos tiempo en la misma” (People Contact, 2021).

En los sistemas que realizan funciones de soporte de interacción se destacan los *chatbots*, por ejemplo, *Voicebot* es un sistema implementado por el ICETEX, “que permite la Automatización del *call center*, donde se interactúa por intenciones, lo que permite que los ciudadanos en primera línea puedan ser atendidos” (ICETEX, 2020). La entidad que lo adoptó estimó que este sistema ha permitido que “se atendier[an] 200.000 personas durante 30 días” (ICETEX, 2020). Por otra parte, uno de los sistemas que realiza funciones de razonamiento con estructura de conocimiento es el “SPCSoil (Métodos de espectroscopia Vis-NIR y aprendizaje automático para diagnosticar propiedades químicas en suelos cañeros colombianos)” de la Corporación Colombiana de Investigación Agropecuaria (Agrosavia).

Por último, uno de los sistemas que realiza personalización es ICFES BOT. Se trata de “una herramienta que interactúa con los estudiantes, y les proporciona información sobre la oferta de programas académicos de educación superior, salarios de los recién egresados, número de profesores de planta y con doctorado en las Instituciones de Educación Superior (IES)” (ICFES, s/f). Además de información sobre financiación de los programas, este algoritmo tiene la particularidad de que “personaliza de acuerdo a las características de los estudiantes, para facilitar el proceso de búsqueda” (ICFES, s/f).

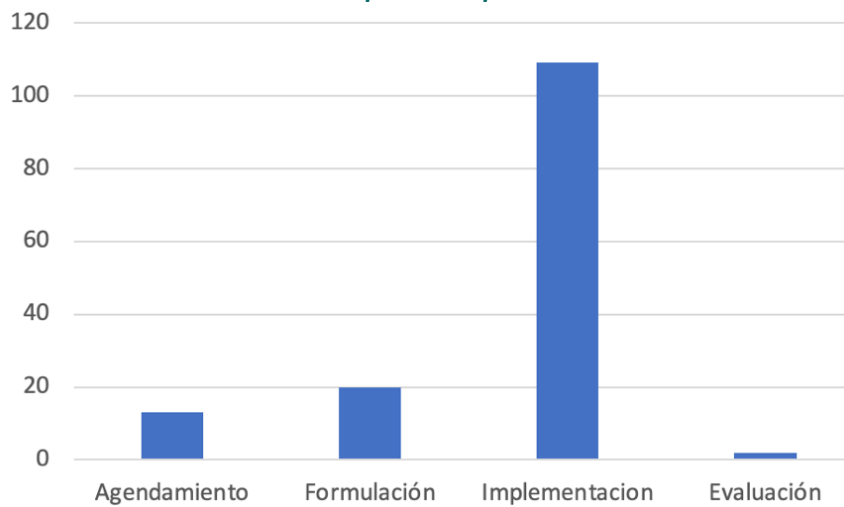
Cerramos esta sub-sección sobre el tipo de actividades que realizan los sistemas bajo la clasificación de la OCDE, comparando los resultados de nuestra base de datos con en el caso chileno. En el Repositorio de Algoritmos Públicos del GobLab UAI la lista está encabezada por las funciones de reconocimiento (29%), predicción (28%), soporte de interacción (16%) y optimización (13%) (GobLab UAI, 2023, p. 52). Por tanto, el mapeo de los sistemas mapeados en Colombia, como se mencionó antes, arroja algunas similitudes con el caso de Chile.

4.2.4 Potenciales aportes de los SDA por etapa del ciclo de la política pública

Los SDA pueden contribuir a realizar actividades de las etapas del ciclo de las políticas públicas (Valle-Cruz et al., 2020).²⁹ En nuestra base de datos identificamos que 13 de 111 SDA (12%) pueden contribuir al agendamiento, 20 (18%) a la formulación, 109 (98%) a la implementación y dos (2%) a la evaluación de las políticas públicas (figura 7).

²⁹ El modelo del ciclo de las políticas públicas es, ante todo, un esquema que nos permite desglosar las políticas públicas a través de una serie de pasos o etapas para facilitar su análisis. Si bien no hay un único modelo del ciclo de las políticas públicas, en general, se suele desglosar en cuatro etapas: agendamiento, formulación, implementación y evaluación de las políticas públicas. Para una definición de cada una de las etapas y reflexiones sobre el ciclo aplicado al caso de Colombia, véase Torres-Melo & Santander (2013), Gutiérrez y Muñoz-Cadena (2023), Gutiérrez y Dajer (2023).

Figura 7 - Clasificación de los SDA por la etapa o etapas del ciclo de la política pública a la cual eventualmente podrían aportar



Fuente: Elaboración propia

A continuación, ofrecemos algunos ejemplos de SDA que contribuyen a dichas etapas del ciclo. En cuanto a la formación de la agenda, por ejemplo, la herramienta *Keepcon* de la Alcaldía de Bogotá busca identificar y dar respuesta a problemáticas planteadas por los ciudadanos en redes sociales. Usando técnicas de análisis de lenguaje natural respecto de las PQRs que se reciben a través de las redes sociales de la Alcaldía, *Keepcon* puede contribuir a mejorar los tiempos de respuesta de los ciudadanos (AgataData, s/f). En la medida en que la herramienta permite identificar ciertos temas que le preocupan a la ciudadanía contribuye a que se agenden temas o problemas públicos sobre los cuales la Alcaldía de Bogotá debería pronunciarse o actuar.

Entre los sistemas que aportan tanto al agendamiento como a la formulación de política pública, se destacan los que utilizan imágenes satelitales, tal es el caso de los sistemas del DNP para la "Identificación de vías terciarias con imágenes satelitales y algoritmos de inteligencia artificial" y el de "Identificación de construcciones en susceptibilidad de inundación debido a su cercanía con las rondas de los ríos".³⁰

En el caso de la implementación, que es la etapa del ciclo de las políticas públicas a las que más aportan los SDA en Colombia, destacaremos dos ejemplos. La herramienta denominada "Automatización Desembolsos del ICETEX" le permite, a esta entidad estatal que otorga créditos educativos, "realizar y gestionar los giros aprobados a los estudiantes que inician y continúan su estudio de educación superior" (ICETEX, 2019, p. 1). La herramienta tiene el propósito de disminuir los tiempos operativos internos para, a su vez, reducir los tiempos de desembolso de los recursos (CPAETD, 2021). Este sistema, estaría, por tanto, relacionado con mejorar los procedimientos al interior de la entidad.

³⁰ El primero, como su nombre lo indica, "busca obtener la línea base de vías terciarias mediante el uso de imágenes satelitales de alta resolución y algoritmos de inteligencia artificial, para reconocer y georreferenciar las vías" (Gobierno Nacional de Colombia, s/f). Entretanto, el segundo sistema pretende, por un lado "generar insumos a las direcciones técnicas para mejorar la toma de decisiones, principalmente en la prevención de riesgos de inundación y la planificación territorial" (GOV.CO, 2021) y, por otro, "anticipar posibles inundaciones para prevenir y evacuar espacios y familias en riesgo" (Misión de Expertos en IA de Colombia, s/f, párr. 5).

Por otra parte, el sistema de “Análítica de redes criminales en el contexto de la Procuraduría General de la Nación (ARCPGN)” puede contribuir a que la autoridad realice sus funciones disciplinarias de manera más efectiva. Este sistema pretende “ayudar a identificar estructuras ilícitas de cooptación institucional, que no operan a partir de sobornos esporádicos sino de procedimientos sistemáticos durante períodos extensos, pero que son susceptibles de pasar desapercibidas frente a los organismos de investigación, judicialización y control” (Cetina et al., 2021; Santiso & Cetina, 2022).

Por último, uno de los sistemas que podría contribuir a realizar seguimiento y evaluación de políticas públicas es “el Sistema de información geográfica Misional y de Apoyo (SIGMA)” de la Unidad Administrativa Especial de Rehabilitación y Mantenimiento Vial de la Alcaldía Mayor de Bogotá D.C. Entre otras funciones, el sistema permitiría “realizar la adecuada gestión de pavimentos de las vías locales de Bogotá” no solo en las etapas de formulación, planeación y ejecución de los proyectos asociados con la conservación de la malla vial, sino también en el seguimiento de estos proyectos, para que se puedan tomar decisiones oportunas (Unidad Administrativa Especial de Rehabilitación, 2022).

4.2.5 Potenciales riesgos asociados al uso de los SDA

Los SDA no solo pueden ofrecer oportunidades para mejorar la prestación de servicios por parte de la administración pública y aportar en los procesos de política pública, sino que también pueden generar riesgos y daños tanto a los usuarios como a terceros. En esta sección destacamos tres tipos de riesgos que pueden generar algunos usos de los SDA mapeados en nuestra base de datos. Identificamos cuatro bloques de derechos fundamentales que pueden ponerse en riesgo: (i) la protección de datos personales y la privacidad, (ii) el derecho de petición y acceso a información pública, (iii) el derecho a la igualdad y no discriminación, y (iv) el del debido proceso y derecho a la defensa.

En primer lugar, identificamos que, de los 111 sistemas mapeados en la base de datos, 64 (58%) usan datos personales. Por ejemplo, *KBoot* es un sistema implementado por la Alcaldía de Medellín para hacer seguimiento a potenciales evasores fiscales a partir de la información que publican en Instagram (CAF, 2021; OECD, 2022b; OECD & OPSI, s/f). Para identificar a las personas físicas detrás de los perfiles, *KBoot* utiliza información como “nombres de usuario, cantidad de seguidores, números de teléfono y cantidad de publicaciones” y contrasta dicha información a través de “la verificación cruzada de nombres contra las bases de datos [de la Tesorería]” y con “información sobre 9.080 usuarios que habían dado un número de contacto” a operadores telefónicos de Medellín (OECD, 2022b, p. 49).

Hay dos principales riesgos que podrían concretarse en relación con los sistemas de requieren datos personales para operar. Por una parte, que los SDA se hayan desarrollado y/o que utilicen datos personales respecto de titulares que no hayan consentido el uso de dicha información. Por otra parte, que se presenten incidentes de seguridad en los cuales terceros accedan sin autorización a datos personales que las entidades públicas deben resguardar en calidad de responsables o encargados del tratamiento de la información.

En segundo lugar, los sistemas pueden afectar negativamente la transparencia y explicabilidad de las decisiones adoptadas por la administración pública. Por ejemplo, esto puede ocurrir cuando los algoritmos usados para tomar decisiones son secretos, como explicaremos en un ejemplo a continuación, o si operan como “cajas negras” respecto de las cuales solo se puede saber qué información entró y qué información produce, pero no es posible determinar el tipo de inferencias que hace el sistema para emitir dichos resultados. La falta de acceso a ese tipo de información podría, por ejemplo, impedir que los ciudadanos afectados por decisiones

discriminatorias –producto de modelos sesgados y/o bases de datos desbalanceadas– puedan ejercer efectivamente las reclamaciones que correspondan.

En relación con las potenciales afectaciones a los ciudadanos producto de la utilización de sistemas de IA en el sector público colombiano, la academia ha identificado casos como el de la implementación del Sistema de Identificación de Potenciales Beneficiarios de Programas Sociales (SISBEN), en su versión más reciente, Sisben IV, que permite perfilar a los individuos y afecta directamente a los de menores recursos (Chenou & Rodríguez Valenzuela, 2021; López, 2020; López & Castañeda, 2020). Precisamente, con este sistema se recogen datos personales³¹, que luego son utilizados para clasificar a la población de acuerdo con sus ingresos y condiciones de vida para, a partir de esta información, determinar si debería o no ser beneficiario de los programas sociales del Estado, como transferencias monetarias condicionadas o subsidios. En los artículos académicos que tratan sobre este sistema (Chenou & Rodríguez Valenzuela, 2021; López, 2020; López & Castañeda, 2020) se resalta la importancia de que los ciudadanos que son “calificados” por un algoritmo cuenten con mecanismos para conocer cuáles fueron los criterios que se utilizaron, dónde pueden “apelar la decisión” en caso de no estar de acuerdo con la calificación asignada, y que sus datos no sean “cedidos” o “utilizados” por terceros.

En tercer lugar, la utilización de SDA que operen a partir de modelos sesgados o que hayan sido entrenados con bases de datos desbalanceadas pueden generar decisiones discriminatorias por razones de género, raza, condición socioeconómica, y pertenencia a grupos étnicos. En Países Bajos, por ejemplo, se han presentado casos de SDA usados por diferentes entidades estatales a través de los cuales excluyeron de manera discriminatoria e injustificada a miles beneficiarios de programas sociales, con el agravante de que muchas de estas personas hacían parte de grupos vulnerables (Peeters & Widlak, 2023).

Por último, el 18% de los sistemas mapeados en nuestra base de datos realizan funciones que están relacionadas con la justicia. Las funciones que cumplen son muy variadas, pero algunos de estos podrían ser usados para tomar decisiones sobre a qué personas u organizaciones investigar e incluso el sentido de algunas decisiones. La literatura ha identificado que el uso de IA en procesos judiciales o administrativos puede poner en riesgo derechos como el debido proceso y el derecho de defensa cuando las partes no tienen conocimiento que dichos sistemas son usados en el proceso o que, aun sabiendo que son usados, no pueden acceder a información sobre cómo operan y qué explica el tipo de recomendaciones o predicciones que generan (Gutiérrez, 2020).

Por ese motivo, Gutiérrez argumenta que es importante realizar “estudios de impacto previos que determinen las potenciales afectaciones relacionadas con aspectos centrales [...] y otros aspectos que no son centrales al proceso [judicial] pero que también incumben a los derechos fundamentales” (2020, p. 514). Además, Camacho y Saavedra resaltan que antes de adoptar herramientas de IA en el ámbito judicial, es necesario,

³¹ Datos que no solo son suministrados por el individuo, sino que también provienen de fuentes externas al sector público. De hecho, en uno de los artículos se detalla como el DNP firmó un Acuerdo de intercambio de información con la empresa multinacional Experian, que maneja Datacrédito, para “utilizar la base de datos de riesgos financieros de Experian para ‘efectuar cruces con la información de la base de datos del SISBÉN’ y mejorar la calidad de la información buscando inconsistencias y ‘con ello obtener mayor precisión respecto de la solvencia de las personas encuestadas’” (López, 2020, p. 19).

entre otras acciones, conocer “las capacidades propias, organizativas, estructurales e individuales de la entidad en cuestión y sus limitaciones” (2021, p. 186).


5. Conclusiones

Este artículo presenta los principales hallazgos derivados de la nueva base de datos sobre SDA de 51 entidades del sector público colombiano. La base de datos identifica 113 SDA que son caracterizados a partir de 40 variables y se basa en más de 300 fuentes de información primaria y secundaria.

A partir del análisis de 111 sistemas que se encuentran en fase de pilotaje o en ejecución en nuestra base de datos destacamos tres principales hallazgos. En primer lugar, encontramos que los SDA han sido adoptados por diferentes tipos de entidades públicas nacionales y subnacionales del poder ejecutivo, así como por entidades de la rama judicial y organismos de control. El grueso de los sistemas en ejecución o pilotaje fueron adoptados por entidades públicas del orden nacional (74%) y la gran mayoría fueron implementados por entidades del poder público ejecutivo (93%). Sin embargo, la base de datos también da cuenta del uso de estos sistemas por parte de entidades del orden territorial (26%) y de entidades que pertenecen a la rama judicial o a entidades de control (7%).

En segundo lugar, el Estado colombiano utiliza los SDA para realizar una amplia gama de funciones y para apoyar diferentes tipos de actividades estatales. Sin embargo, más de la mitad de los sistemas son usados para apoyar actividades de los sectores justicia (18%), educación (17%), y medio ambiente (13%). Además, encontramos que los SDA realizan diferentes funciones (clasificadas por tipo de *output* generado), pero que se concentran en tres tipos de funciones: 64% de los sistemas realizan tareas de reconocimiento, 49% detectan eventos, y 34% realizan optimización enfocada en objetivos. Adicionalmente, encontramos que la mayoría de los SDA pueden aportar a la implementación de políticas públicas (el 98%), pero que algunos también pueden contribuir con otras fases del ciclo de las políticas públicas (12% al agendamiento, 18% a formulación y 2% a evaluación).

En tercer lugar, nuestra base de datos también ilustra cómo cuatro bloques de derechos fundamentales pueden ponerse en riesgo ante determinados usos de SDA por parte del Estado: la protección de datos personales y la privacidad, el derecho de petición y acceso a información pública, el derecho a la igualdad y no discriminación, y el debido proceso y derecho a la defensa.

Además de contribuir con la literatura académica sobre administración pública y tecnologías a partir del mapeo de SDA en el Estado colombiano, esta investigación abre al menos dos futuras líneas de investigación académica. Primero, sobre la observancia del principio de transparencia algorítmica por parte del Estado colombiano y las implicaciones de política sobre las brechas de información acerca de la adquisición, desarrollo, pilotaje, adopción, uso y evaluación de estos sistemas en el sector público colombiano. Segundo, la base de datos permite implementar de manera más rigurosa métodos de selección de caso con el fin de realizar estudios comparados que investiguen todo el ciclo de vida de estos proyectos. Estos estudios podrían abordar preguntas que permitan entender cómo surge la idea o necesidad de contar con estos sistemas, cuáles son los principales retos de gestión pública, cómo se enmarcan estos procesos dentro del contexto organizacional, y cómo pueden desarrollarse sistemas centrados en las personas y con pleno respeto de los derechos fundamentales. 

Anexo 1. Diez definiciones de Sistemas de Decisión Automatizada SDA

Fuente	Definición
(Richardson, 2022, p. 795)	"Definición integral de SDA: 'Sistema de Decisión Automatizada' es cualquier herramienta, software, sistema, proceso, función, programa, método, modelo, y/o fórmula diseñada con o usando computación para automatizar, analizar, ayudar, aumentar, y/o reemplazar decisiones gubernamentales, juicios, y/o implementación de políticas. Los sistemas de decisión automatizados afectan a las oportunidades, el acceso, las libertades, la seguridad, los derechos, las necesidades, el comportamiento, la residencia y/o el estatus mediante la predicción, la puntuación, el análisis, la clasificación, la demarcación, la recomendación, la asignación, la enumeración, la clasificación, el seguimiento, la cartografía, la optimización, la imputación, la inferencia, el etiquetado, la identificación, la agrupación, la exclusión, la simulación, el modelado, la evaluación, la fusión, el procesamiento, la agregación y/o el cálculo."
(Richardson, 2022, p. 795)	"Definición estrecha de SDA: 'Sistemas de Decisión Automatizada' son los sistemas, programas informáticos o procesos que utilizan la computación para ayudar o sustituir las decisiones, juicios o aplicación de políticas del gobierno que afectan a las oportunidades, el acceso, las libertades, los derechos o la seguridad. Los sistemas automatizados de toma de decisiones pueden predecir, clasificar, optimizar, identificar o recomendar."
(WaTech, 2021, p. 5)	"'Sistema de decisión automatizada' o 'sistema' significa cualquier algoritmo, incluyendo uno que incorpore aprendizaje automatizado u otras técnicas de inteligencia artificial, que utilice análisis o cálculos basados en datos para tomar o apoyar decisiones, juicios o conclusiones gubernamentales que causen que un residente de Washington sea tratado de manera diferente que otro residente de Washington en la naturaleza o cantidad de interacción gubernamental con ese individuo, incluyendo, sin limitación, beneficios, protecciones, pagos requeridos, sanciones, regulaciones, requisitos de tiempo, aplicación o proceso."
(Garrido et al., 2021, p. 42)	"Sistema de decisión automatizada (<i>SDA automated decision system</i> o <i>automated decision making</i> ADS o ADM por sus siglas en inglés) son sistemas que asumen o pueden asumir el rol principal o único en la toma de decisiones. Un sistema de decisiones automatizadas es un sistema controlado algorítmicamente, en el que un proceso de decisión se delega parcial o totalmente a otro, el que a su vez toma o propone una decisión automáticamente" [69]. Para efectos de la encuesta elaborada para este estudio, estos fueron definidos como 'aquellos que realizan acciones con cierto grado de autonomía para lograr objetivos específicos. Estos sistemas pueden basarse exclusivamente en software o hardware, proveyendo de un insumo para una decisión realizada por un funcionario público o bien tomar directamente una decisión.'"
(Information Commissioner's Office, s/f)	"¿Qué es la toma de decisiones automatizada? La toma de decisiones automatizada es el proceso de tomar una decisión por medios automatizados sin intervención humana. Estas decisiones pueden basarse en datos objetivos, así como en perfiles creados digitalmente o datos inferidos. Ejemplos de ello son: una decisión en línea para conceder un préstamo; y una prueba de aptitud utilizada para la contratación que utiliza algoritmos y criterios preprogramados."
(Aránguiz Villagrán, 2022, p. 5)	"Un sistema automatizado de toma o soporte de decisiones (ADS por las siglas en inglés de <i>Automated Decision Support</i>) es un sistema computacional que puede, para un determinado conjunto de objetivos definidos por seres humanos, hacer predicciones y recomendaciones o tomar decisiones que inciden en entornos reales o virtuales. Estos sistemas están diseñados para operar con diversos grados de autonomía [2]."

Fuente	Definición
(Araujo et al., 2020, pp. 611–612)	"La toma de decisiones automatizada puede definirse de varias maneras. En sentido estricto, puede describirse como 'decisiones tomadas por medios tecnológicos sin intervención humana' (Comisión Europea 2018, p. 7). Sin embargo, de forma más amplia, puede considerarse como el proceso a través del cual la cantidad –y variedad– cada vez mayor de datos personales 'son posteriormente procesados por algoritmos, que luego se utilizan para tomar decisiones (basadas en datos)' (Newell y Marabelli 2015, p. 4). Por lo tanto, el SDA implica una serie de procesos, desde ayudas para los responsables humanos de la toma de decisiones hasta procesos de toma de decisiones completamente automatizados, en una amplia variedad de contextos."
(Algorithm Watch & Bertelsmann Stiftung, 2019, p. 9)	"La toma de decisiones automatizada y controlada algorítmicamente o los sistemas de apoyo a la toma de decisiones son procedimientos en los que las decisiones se delegan inicialmente –parcial o totalmente– en otra persona o entidad corporativa, que a su vez utiliza modelos de toma de decisiones ejecutados automáticamente para realizar una acción. Esta delegación –no de la decisión en sí, sino de la ejecución– en un sistema basado en datos y controlado algorítmicamente es lo que requiere nuestra atención. En comparación, la Inteligencia Artificial es un término de definición imprecisa que engloba una amplia gama de ideas controvertidas y, por tanto, no resulta muy útil para abordar las cuestiones que nos ocupan. Además, el término 'inteligencia' tiene connotaciones de autonomía e intencionalidad humanas que no deben atribuirse a los procedimientos basados en máquinas. Por otra parte, los sistemas que no se considerarían Inteligencia Artificial según la mayoría de las definiciones actuales, como los simples procedimientos de análisis basados en reglas, pueden tener un gran impacto en la vida de las personas, por ejemplo, en forma de sistemas de puntuación para la evaluación de riesgos."
(Pittsburgh Task Force on Public Algorithms, 2022, p. 8)	"El término 'sistema algorítmico público' se refiere a cualquier sistema, software o proceso que utilice la computación, incluidos los derivados del aprendizaje automatizado u otras técnicas de procesamiento de datos o inteligencia artificial (IA), para ayudar o reemplazar las decisiones, juicios o implementaciones políticas del gobierno que afectan a las oportunidades, el acceso, las libertades, los derechos o la seguridad."
(Directive on Automated Decision-Making, 2023)	"Cualquier tecnología que asiste o sustituye el juicio de los responsables humanos de la toma de decisiones. Estos sistemas se basan en campos como la estadística, la lingüística y la informática, y utilizan técnicas como los sistemas basados en reglas, la regresión, el análisis predictivo, el aprendizaje automatizado, el aprendizaje profundo y las redes neuronales."

Fuente: elaboración propia

Referencias

- AgataData. (s/f). ¿Qué hacemos? - AgataData. Recuperado el 25 de octubre de 2022, de <https://agatadata.com/queHacemos.html>
- Agudelo Londoño, S. (2022). Inteligencia Artificial en el proceso penal: Análisis a la luz del fiscal Watson. *Pensamiento Jurídico*, 1(54). <https://revistas.unal.edu.co/index.php/peju/article/view/96091>
- Alcaldía de Armenia. (2022, junio 6). \$2.000 millones pondrían a Armenia a la altura de grandes capitales inteligentes—Alcaldía de Armenia. <https://www.armenia.gov.co/atencion-al-ciudadano/noticias/2-000-millones-pondrian-a-armenia-a-la-altura-de-grandes-capitales-inteligentes>
- AlgorithmWatch & Bertelsmann Stiftung. (2019). Automating Society: Taking Stock of Automated Decision Making in the EU (1a ed.). AW AlgorithmWatch gGmbH. www.algorithmwatch.org/automating-society
- Aránguiz Villagrán, M. (2022). Auditoría algorítmica para sistemas de toma o soporte de decisiones. BID. <https://publications.iadb.org/publications/spanish/viewer/Auditoria-algoritmica-para-sistemas-de-toma-o-soporte-de-decisiones.pdf>
- Araujo, T., Helberger, N., Kruikemeier, S., & de Vreese, C. H. (2020). In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI & SOCIETY*, 35(3), 611–623. <https://doi.org/10.1007/s00146-019-00931-w>
- Berryhill, J., Heang, K. K., Clogher, R., & Keegan McBride. (2019). Hola, Mundo: La inteligencia artificial y su uso en el sector público (Vol. 36). <https://doi.org/10.1787/726fd39d-en>
- CAF. (2021). Experiencia: Datos e Inteligencia Artificial en el sector público. CAF. <http://scioteca.caf.com/handle/123456789/1793>
- Calderon-Valencia, F., Perez-Montoya, J.-J., & De Morais, F. S. (2021). Sistemas de IA en la experiencia del Supremo Tribunal Federal Brasileño y la Corte Constitucional Colombiana: Análisis Prospectivo. *Law, State and Telecommunications Review*, 13(1), 143–169. <https://doi.org/10.26512/lstr.v13i1.35614>
- Camacho Gutiérrez, O. L., Castañeda Gómez, J. D., & Saavedra Rionda, V. P. (2019). Colombia. Enthusiasm and complexity: Learning from the "PROMETEA" pilot in Colombia's Judicial System. En APC, Sida, & Article 19, *Global information society watch 2019: Artificial intelligence: Human rights, social justice and development* (pp. 113–116).
- Camacho Gutiérrez, O. L., & Saavedra Rionda, V. P. (2021). El acercamiento del sector público a la IA: la divergencia entre la expectativa y la realidad. En G. A. Dobratinich (Ed.), *Derecho y nuevas tecnologías* (1a ed.). La Ley.
- Castaño, D. (2021). Marco analítico exploratorio—República de Colombia. En C. Aguerre & G. Bustos Frati (Eds.), *Preparación del sector judicial para la inteligencia artificial en América Latina* (compilación) (pp. 3–100). Centro de Estudios en Tecnología y Sociedad (CETyS) de la Universidad de San Andrés. https://www.researchgate.net/publication/360849506_Preparacion_del_sector_judicial_para_la_inteligencia_artificial_en_America_Latina_compilacion

- Centro de Innovación Pública Digital. (2022). Estrategia de Seguimiento: Soluciones basadas en Inteligencia Artificial 2020. https://gobiernodigital.mintic.gov.co/692/articles-210416_recurso_8.pdf
- Centro de Innovación Pública Digital. (2020). Ejercicios de innovación basados en Inteligencia Artificial. <https://gobiernodigital.mintic.gov.co/portal/Centro-de-Innovacion-Publica-Digital/Proyectos/#data=%7B%22filter%22:%22412572%22,%22page%22:1%7D>
- Cepeda Espinoza, M. J., & Otálora Lozano, G. (2020). Modernización de la administración de justicia a través de la inteligencia artificial. FEDESARROLLO. https://www.repository.fedesarrollo.org.co/bitstream/handle/11445/3980/Repor_Julio_2020_Cepeda_y_Otálora.pdf?sequence=3&isAllowed=y
- Cetina, C., Garay Salamanca, L. J., Salcedo-Albarán, E., & Vanegas, S. (2021). La analítica de redes como herramienta de integridad: El caso de la Procuraduría General de la Nación en Colombia (Policy Brief # 22) [WorkingPaper]. CAF. <https://cafscioteqa.azurewebsites.net/handle/123456789/1675>
- Chenou, J.-M., & Rodríguez Valenzuela, L. E. (2021). Habeas Data, Habemus Algorithms: Algorithmic Intervention in Public Interest Decision-Making in Colombia. *Law, State and Telecommunications Review*, 13(2), 56–77. <https://doi.org/10.26512/lstr.v13i2.34113>
- CPAETD. (2021). ¿Cómo vamos avanzando en la transformación digital? DAPRE.
- DANE. (2020). Clasificación de las funciones del gobierno (COFOG). <https://www.dane.gov.co/files/sen/nomenclatura/cofog/COFOG-AC.pdf>
- Desouza, K. C., Dawson, G. S., & Chenok, D. (2020). Designing, developing, and deploying artificial intelligence systems: Lessons from and for the public sector. *Business Horizons*, 63(2), 205–213. <https://doi.org/10.1016/j.bushor.2019.11.004>
- El Tiempo. (2022, junio 20). Las cuatro innovaciones para la defensa en demandas contra la nación. *El Tiempo*. <https://www.eltiempo.com/justicia/servicios/las-cuatro-innovaciones-para-la-defensa-en-demandas-contra-la-nacion-681660>
- Estevadeordal, A., Beliz, G., Estevez, E., Ovanessoff, A., Plastino, E., Rao, A., Diamond, P., Barral, W., Petrus, G., Donaldson, D., Vashistha, A., Vashistha, A., Herrera, M., Heymann, D., Mira, P., Chesñear, C., Lakhani, K., Tinn, P., Lin, M., ... Gillam, M. (2018). *Algoritmolandia: Inteligencia artificial para una integración predictiva e inclusiva de América Latina* (1a ed.). Planeta. <https://publications.iadb.org/publications/spanish/document/Revista-Integración--Comercio-Año-22-No-44-Julio-2018-Algoritmolandia-inteligencia-artificial-para-una-integración-predictiva-e-inclusiva-de-América-Latina.pdf>
- Fiscalía General de la Nación. (2019). Informe Rendición de cuentas (2018-2019) Fiscalía General de la Nación. <https://www.fiscalia.gov.co/colombia/wp-content/uploads/FGN.-Informe-de-gestio%CC%81n.-Versio%CC%81n-final.pdf>
- Flórez Rojas, M. L., & Vargas Leal, J. (2020). El impacto de herramientas de inteligencia artificial: Un análisis en el sector público en Colombia. En C. Aguerre (Ed.), *Inteligencia Artificial en América Latina y el Caribe*.

- Ética, Gobernanza y Políticas. CETyS Universidad de San Andrés. <https://guia.ai/wp-content/uploads/2020/05/GECTI-El-impacto-de-herramientas-de-inteligencia-artificial.pdf>
- Gallego, J., Prem, M., & Vargas, J. F. (2022). Predicting politicians' misconduct: Evidence from Colombia. *Data & Policy*, 4. Cambridge Core. <https://doi.org/10.1017/dap.2022.35>
- Gallego, J., Rivero, G., & Martínez, J. (2021). Preventing rather than punishing: An early warning model of malfeasance in public procurement. *International Journal of Forecasting*, 37(1), 360–377. <https://doi.org/10.1016/j.ijforecast.2020.06.006>
- Garrido, R., Lapostol, J. P., & Hermosilla, M. P. (2021). Transparencia algorítmica en el sector público. GOB LAB UAI. Consejo para la Transparencia. <https://goblab.uai.cl/wp-content/uploads/2021/10/ESTUDIO-TRANSPARENCIA-ALGORITMICA-EN-EL-SECTOR-PUBLICO-GOBLAB-CPLT-final....pdf>
- Gobierno Nacional de Colombia. (s/f). Dashboard de seguimiento al Marco Ético de Inteligencia Artificial. Recuperado el 27 de diciembre de 2022, de <https://inteligenciaartificial.gov.co/dashboard-IA/>
- GobLab UAI. (2022). Repositorio de algoritmos públicos de Chile. Primer informe de estado de uso de algoritmos en el sector público. Universidad Adolfo Ibáñez (UAI). <https://goblab.uai.cl/wp-content/uploads/2022/02/Primer-Informe-Repositorio-Algoritmos-Publicos-en-Chile.pdf>
- GobLab UAI. (2023). Repositorio Algoritmos Públicos. Informe Anual 2023. Universidad Adolfo Ibáñez (UAI).
- GOV.CO. (2021). Datos Abiertos. Datos Abiertos. <https://herramientas.datos.gov.co/taxonomy/term/501?page=0>
- Gutiérrez, J. D. (2020). Retos éticos de la inteligencia artificial en el proceso judicial. En ICDP (Ed.), *Derecho Procesal. #NuevasTendencias. XLI Congreso Colombiano de Derecho Procesal*. Instituto Colombiano de Derecho Procesal (ICDP) y Universidad Libre. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4011179
- Gutiérrez, J. D. (2023a). ¿Están los jueces en capacidad de usar modelos de lenguaje a gran escala (LLMs)? *Revista EXCEJENCIA*, 7, 10–15.
- Gutiérrez, J. D. (2023b). Hablemos sobre el uso de ChatGPT para redactar decisiones judiciales. *La Silla Vacía*. <https://juangutierrez.co/publicaciones/otras-publicaciones/#:~:text=Hablemos%20sobre%20el%20uso%20de%20ChatGPT%20para%20redactar%20decisiones%20judiciales>
- Gutiérrez, J. D. (2023c, febrero 23). ChatGPT in Colombian Courts: Why we need to have a conversation about the digital literacy of the judiciary. *VerfBlog*. <https://verfassungsblog.de/colombian-chatgpt/>
- Gutiérrez, J. D. (2023d, julio 6). Repositorios y registros públicos de algoritmos. *Foro Administración, Gestión y Política Pública*. <https://forogpp.com/inteligencia-artificial/repositorios-y-registros-de-algoritmos/>
- Gutiérrez, J. D., & Dajer, D. M. (2023). Pensamiento de diseño y procesos de política pública. *Desafíos*, 35(1), 1–27.

- Gutiérrez, J. D., & Muñoz-Cadena, S. (2022). Sistemas de decisión automatizada implementados por el Estado Colombiano. 32. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4294032
- Gutiérrez, J. D., & Muñoz-Cadena, S. (2023). Assessing Government Design Practices from a Human-Centered Perspective: Case Study of an Improved Cookstoves Program in Colombia. *Desafíos*, 35(1), 1–38.
- Gutiérrez, J. D., Muñoz-Cadena, S., & Castellanos-Sánchez, M. (2023). Sistemas de decisión automatizada en el sector público colombiano (Versión V1) [Data set]. Universidad del Rosario. <https://doi.org/10.34848/YN1CRT>
- Gutiérrez-Ossa, J. A., & Flórez-Hernández, I. C. (2020). Inteligencia Artificial (IA) Aplicada en el Sistema Judicial en Colombia. *Derecho y Realidad*, 18(35), 53–80. <https://doi.org/10.19053/16923936.v18.n35.2020.9638>
- ICBF (Director). (2020, agosto 27). Conoce cómo funciona el Asistente de Adopciones (ADA). <https://www.youtube.com/watch?v=iL7mfoSwxx8>
- ICETEX. (2019). Gestión de accesos y retiro de servicios (Gestión de servicios tecnológicos Versión 11). <https://web.icetex.gov.co/documents/20122/144117/Gestion+de+acceso+y+retiro+de+servicios+v11.pdf>
- ICETEX. (2020, 12). Café con Oficina Comercial y de Mercadeo- 15 de diciembre de 2020 (Gestión del Conocimiento e innovación). <https://web.icetex.gov.co/documents/20122/301689/Caf%25C3%25A9%2Bcon%2BOficina%2BComercial%2By%2Bde%2BMercadeo.docx>
- ICFES. (s/f). Icfes-Bot—Icfes. Recuperado el 27 de diciembre de 2022, de <https://www2.icfes.gov.co/icfes-bot>
- Information Commissioner's Office. (s/f). What is automated individual decision-making and profiling? <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/automated-decision-making-and-profiling/what-is-automated-individual-decision-making-and-profiling/#id2>
- Instituto Humboldt. (2023, febrero 23). Escuchar para resguardar los bosques. <http://www.humboldt.org.co/es/boletines-y-comunicados/item/1813-escuchar-para-resguardar-los-bosques>
- Ivančić, L., Suša Vugec, D., & Bosilj Vukšić, V. (2019). Robotic Process Automation: Systematic Literature Review. En C. Di Ciccio, R. Gabryelczyk, L. García-Bañuelos, T. Hernaus, R. Hull, M. Indihar Štemberger, A. Kő, & M. Staples (Eds.), *Business Process Management: Blockchain and Central and Eastern Europe Forum* (Vol. 361, pp. 280–295). Springer International Publishing. https://doi.org/10.1007/978-3-030-30429-4_19
- Kuziemski, M., & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *ScienceDirect*. <https://doi.org/DOI:10.1016/j.telpol.2020.101976>

- López, J. (2020). Experimentando con la pobreza: El SISBÉN y los proyectos de analítica de datos en Colombia. Fundación KARISMA. <https://web.karisma.org.co/wp-content/uploads/download-manager-files/Experimentando%20con%20la%20pobreza.pdf>
- López, J., & Castañeda, J. D. (2020). Automatización, tecnologías digitales y justicia social: La experimentación con la pobreza en Colombia. En *Inteligencia Artificial en América Latina y el Caribe. Ética, gobernanza y políticas*. CETyS Universidad de San Andres. <https://proyctoguia.lat/wp-content/uploads/2020/05/Lopez-Casta%C3%B1eda-Automatizacion-tecnologias-digitales-y-justicia-social-la-experimentacion-con-la-pobreza-en-Colombia.pdf>
- Martínez, C. A. (2022, junio). Analítica de datos en el sector público: El poder de los datos. Quinto Encuentro de equipo transversal de gestión de conocimiento y la innovación. https://www.funcionpublica.gov.co/documents/37183536/41306203/2022-0616_Presentacion_completa_quinto_gescoi.pdf/b827772a-b3b1-ece8-a29c-3979ae971ecd?t=1655501285963
- MinTIC. (2021). Ciudades y Territorios Inteligentes [Gubernamental]. Gobierno Digital. <https://gobiernodigital.mintic.gov.co/portal/Iniciativas/Ciudades-y-Territorios-Inteligentes/>
- Misión de Expertos en IA de Colombia. (s/f). Casos de Uso | AprendeIA. Recuperado el 15 de agosto de 2022, de https://aprendeia.org/casos_de_uso
- Misuraca, G., Van Noordt, C., & Boukli, A. (2020, septiembre 23). The use of AI in public services: Results from a preliminary mapping across the EU. *Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance (ICEGOV 2020)*, Athens, Greece. <https://doi.org/10.1145/3428502.3428513>
- Mojica Muñoz, K. S. (2021). Inteligencia Artificial para Detectar Corrupción en la Administración Pública Municipal de Colombia. *Universidad de los Andes, Facultad de Economía, CEDE*, 31, 70.
- Mojica Muñoz, K. S., Algarra Saavedra, P. A., & Peña Duarte, J. (2021). Artificial Intelligence to Prevent the Assassination of Human Rights Defenders.
- Noticiero 90 Minutos. (2021, diciembre 9). Se hundió proyecto de "Cali Inteligente" en el Concejo. *Noticiero 90 Minutos*. <https://90minutos.co/cali/se-hundio-proyecto-cali-inteligente-concejo-09-12-2021/>
- Nueva Crónica Quindío. (2022, junio 13). Armenia ahora será una ciudad inteligente con semáforos tecnológicos 800. *Cronica del Quindío*. <https://cronicadelquindio.com/noticias/ciudad/armenia-ahora-sera-una-ciudad-inteligente-con-semaforos-tecnologicos-800>
- OCDE. (2017). Revisión de Gobierno Digital de Colombia. Hacia un sector público impulsado por el ciudadano. Hallazgos clave. OECD Publishing. [https://www.oecd.org/gov/digital-government/Digital%20Gov%20Review%20Colombia%20\[Esp\]%20def.pdf](https://www.oecd.org/gov/digital-government/Digital%20Gov%20Review%20Colombia%20[Esp]%20def.pdf)
- OECD. (2022a). OECD Framework for the Classification of AI systems. OECD. <https://doi.org/10.1787/cb6d9eca-en>

- OECD. (2022b). Uso estratégico y responsable de la inteligencia artificial en el sector público de América Latina y el Caribe. Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/governance/uso-estrategico-y-responsable-de-la-inteligencia-artificial-en-el-sector-publico-de-america-latina-y-el-caribe_5b189cb4-es
- OECD & CAF. (2022). The Strategic and Responsible Use of Artificial Intelligence in the Public Sector of Latin America and the Caribbean. OECD Public Governance Reviews, OECD Publishing. 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
- OECD, & OPSI. (s/f). Tracking potential tax evaders on Instagram. Observatory of Public Sector Innovation. Recuperado el 27 de diciembre de 2022, de <https://oecd-opsi.org/innovations/tracking-potential-tax-evaders-on-instagram/>
- Ortega Rance, G. (2021). La digitalización como estrategia para la calidad regulatoria: Aprendizajes de la iniciativa de transformación digital de la Superintendencia de Sociedades de Colombia (Policy Brief # 29) [WorkingPaper]. CAF. <https://cafsciotea.azurewebsites.net/handle/123456789/1784>
- Peeters, R., & Widlak, A. C. (2023). Administrative exclusion in the infrastructure-level bureaucracy: The case of the Dutch daycare benefit scandal. *Public Administration Review*, n/a(n/a). <https://doi.org/10.1111/puar.13615>
- People Contact (Director). (2021, septiembre 24). Predicción de pasajeros | Cable aéreo | Manizales. <https://www.youtube.com/watch?v=xQOY30ouxzM>
- Pittsburgh Task Force on Public Algorithms. (2022). Report of the Pittsburgh Task Force on Public Algorithms. University of Pittsburgh. https://www.cyber.pitt.edu/sites/default/files/pittsburgh_task_force_on_public_algorithms_report.pdf
- Price, J. F. (2012). Wagging, not barking: Statutory definitions. *Cleveland State Law Review*, 60(4).
- Procuraduría General de la Nación. (2020). Informe relatoria actuación preventiva -Sistema de Ciberinteligencia basado en inteligencia artificial para la DIPOL. https://apps.procuraduria.gov.co/gp/gp/anexos/ciberseguridad_informe_relatoria.pdf
- Reis, J., Santo, P. E., & Melão, N. (2019). Artificial Intelligence in Government Services: A Systematic Literature Review. En Á. Rocha, H. Adeli, L. P. Reis, & S. Costanzo (Eds.), *New Knowledge in Information Systems and Technologies* (Vol. 930, pp. 241–252). Springer International Publishing. https://doi.org/10.1007/978-3-030-16181-1_23
- Richardson, R. (2022). Defining and Demystifying Automated Decision Systems. *Maryland law review* (1936), 81(785).
- Saavedra, S. (2022). Technology and State Capacity: Experimental Evidence from Illegal Mining in Colombia. Red Investigadores de Economía. Documento de Trabajo. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3933128

Santiso, C., & Cetina, C. (2022). DIGIntegridad: La transformación digital de la lucha contra la corrupción. CAF. <https://cafscioteqa.azurewebsites.net/handle/123456789/1901>

SENA. (2021, noviembre 24). Proyectos Transformación Digital. <https://www.sena.edu.co/es-co/transparencia/FURAG2/Forms/AllItems.aspx>

Sousa, W. G. de, Melo, E. R. P. de, Bermejo, P. H. D. S., Farias, R. A. S., & Gomes, A. O. (2019). How and where is artificial intelligence in the public sector going? A literature review and research agenda. *Government Information Quarterly*, 36(4), 101392. <https://doi.org/10.1016/j.giq.2019.07.004>

Torres-Melo, J., & Santander, J. (2013). *Introducción a las Políticas Públicas: Conceptos y Herramientas Desde la Relación Entre Estado y Ciudadanía*. Instituto de Estudios del Ministerio Público, Procuraduría General de la Nación.

Directive on Automated Decision-Making, (2023). <https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592>

Unidad Administrativa Especial de Rehabilitación. (2022, mayo 17). Transformación digital. SIGMA. <https://tic.bogota.gov.co/sites/default/files/imagenSimple/UNIDADDEMANTENIMIENTOVIAl.pdf>

Valle-Cruz, D., Criado, J. I., Sandoval-Almazán, R., & Ruvalcaba-Gomez, E. A. (2020). Assessing the public policy-cycle framework in the age of artificial intelligence: From agenda-setting to policy evaluation. *Government Information Quarterly*, 37(4), 101509. <https://doi.org/10.1016/j.giq.2020.101509>

WaTech. (2021). Automated Decision Making Systems Workgroup Charter. <https://watech.wa.gov/sites/default/files/public/privacy/ADS%20Workgroup%20Charter%20Final%207-2021.pdf>

Wirtz, B. W., & Müller, W. M. (2019). An integrated artificial intelligence framework for public management. *Public Management Review*, 21(7), 1076–1100. <https://doi.org/10.1080/14719037.2018.1549268>

Zuiderwijk, A., Chen, Y.-C., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly*, 38(3), 101577. <https://doi.org/10.1016/j.giq.2021.101577>

Sobre el autores/ About the authors

Juan David Gutiérrez es Profesor Asociado de la Escuela de Gobierno Alberto Lleras Camargo de la Universidad de los Andes. PhD en Política Pública de la Escuela de Gobierno de la Universidad de Oxford. Investiga sobre política pública, inteligencia artificial, y gobernanza de los recursos naturales. Sarah Muñoz-Cadena es estudiante de la maestría en Economía de las Políticas Públicas de la Universidad del Rosario e investigadora en Policéntrico. Politóloga y profesional en Gobierno y Asuntos Públicos con estudios complementarios en periodismo de la Universidad de los Andes. Investiga sobre gobernanza de la inteligencia artificial y design thinking.

URL estable documento/stable URL

<http://www.gigapp.org>

El Grupo de Investigación en Gobierno, Administración y Políticas Públicas (GIGAPP) es una iniciativa impulsada por académicos, investigadores y profesores Iberoamericanos, cuyo principal propósito es contribuir al debate y la generación de nuevos conceptos, enfoques y marcos de análisis en las áreas de gobierno, gestión y políticas públicas, fomentando la creación de espacio de intercambio y colaboración permanente, y facilitando la construcción de redes y proyectos conjuntos sobre la base de actividades de docencia, investigación, asistencia técnica y extensión.

Las áreas de trabajo que constituyen los ejes principales del GIGAPP son:

1. Gobierno, instituciones y comportamiento político
2. Administración Pública
3. Políticas Públicas

Información de Contacto

Asociación GIGAPP.
ewp@gigapp.org

Estado de la transparencia y acceso a la información pública en Ecuador pos-pandemia: ¿Han cambiado las necesidades ciudadanas?

Jara Iñiguez, Irma

Instituto de Altos Estudios Nacionales (IAEN), Ecuador

✉ Irma.jara@iaen.edu.ec

ORCID ID: [0000-0002-6342-4068](https://orcid.org/0000-0002-6342-4068)

Documento recibido:

02 septiembre 2022

Aprobado para publicación:

07 agosto 2023

Resumen

En Ecuador la primera ley orgánica de transparencia y acceso a la información pública estuvo vigente en el periodo 2004-2023. Para este periodo, no existe evidencia identificada por el ente garante o la academia sobre la utilidad de la información publicada para la incidencia ciudadana en la co-gestión o el control social. En este escenario, se aprobó una nueva ley que considera varios aspectos del modelo de la OEA, pero que aún presenta vacíos para responder a las actuales necesidades. El estudio tiene como propósito identificar los cambios en las demandas de información en el último quinquenio que -bajo la nueva ley- podrían ser atendidos desde la transparencia focalizada. Para el efecto, se aplica una investigación descriptiva con enfoque cualitativo. Se realiza un análisis documental y la revisión de las solicitudes de acceso para la Presidencia de la República. Se identifica que actualmente se requiere información específica y útil para la toma de decisiones.

Palabras clave

Acceso a la información pública, Legislación, Transparencia activa, Transparencia focalizada, Transparencia pasiva

Resumo

No Equador, a primeira lei orgânica sobre transparência e acesso à informação pública vigorou no período 2004-2023. Para este período, não há evidências identificadas pela entidade garantidora ou pela academia sobre a utilidade das informações publicadas para a advocacia cidadã em cogestão ou controle social. Nesse cenário, foi aprovada uma nova lei que considera vários aspectos do modelo da OEA, mas ainda apresenta lacunas para atender às necessidades atuais. O objetivo do estudo é identificar as mudanças nas demandas de informação nos últimos cinco anos que, sob a nova lei, poderiam ser atendidas por meio de transparência focada. Para tanto, aplica-se uma investigação descritiva com abordagem qualitativa. É realizada uma análise documental e análise dos pedidos de acesso à Presidência

da República. Identifica-se que atualmente são necessárias informações específicas e úteis para a tomada de decisão.

Palavras-chave

Acesso à informação pública, Legislação, Transparência ativa, Transparência focada, Transparência passiva.

Abstract

The first organic law on transparency and access to public information was in force in Ecuador in the period 2004-2023. During this period, there is no evidence identified by national transparency's responsible office nor the academy on the usefulness and potential contribution of the published information to foster citizen advocacy in co-management or social oversight. In this scenario, a new law was approved that considers various aspects of the OAS model, but still has gaps to respond to current needs. The purpose of the study is to identify the changes in the demands for information in the last five years that -under the new law- could be addressed through a focused transparency strategy. For this purpose, a descriptive investigation with a qualitative approach is applied. A documentary analysis and review of access requests for the Presidency of the Republic is carried out. Based on the analysis, we argue that specific information is currently required for decision making.

Keywords

Access to public information, Legislation, Active transparency, Targeted transparency, Passive transparency

Introducción

En Ecuador, en mayo de 2004, se publicó la Ley Orgánica de Transparencia y Acceso a la Información Pública, LOTAIP¹ (Congreso Nacional, 2004); y en 2008, con la nueva Constitución de la República, se creó la Función de Transparencia y Control Social, entre otras, para promover e impulsar el control de las entidades y organismos del sector público, y de las personas naturales o jurídicas del sector privado que presten servicios o desarrollen actividades de interés público para que los realicen con transparencia.

En este marco, Ecuador sí evidencia avances a la fecha respecto a la transparencia y su promoción a través del uso de las plataformas electrónicas y otras herramientas, entre las que se destacan: La homologación de las páginas web de las entidades del gobierno, con su acceso a información mínima de sujetos obligado ante la LOTAIP; el portal de compras públicas², que muestra en detalle el

¹ Registro Oficial, Suplemento N° 337 del 18 de mayo de 2004

² <https://www.compraspublicas.gob.ec/ProcesoContratacion/compras/>

costo de una obra o servicio, y el de contrataciones abiertas³; el portal Ecuador en cifras⁴ en el que se encuentra información de carácter estadístico y cuenta con un banco de datos; y el sistema de gestión documental, QUIPUX, que es un registro fiable de las comunicaciones hacia y desde lo público.

Además, en el 2020, se emitió la política de datos abiertos y la guía de datos abiertos, entre otros propósitos, para fortalecer la transparencia gubernamental (MINTEL, 2021). Así también, se dispone del portal de datos abiertos en cuyo catálogo se encuentran más de 3212 recursos y 895 conjuntos de datos abiertos de 95 instituciones⁵ y la plataforma de diálogo 2.0⁶, como una herramienta de participación ciudadana.

Por otro lado, sin embargo, como resultado de la evaluación y el análisis normativo y práctico de cada uno de los pilares de integridad por parte del Consejo de Participación Ciudadana y Control Social (CPCCS, 2019) se obtuvo en resumen, en cuanto a la temática de transparencia, que los sujetos obligados a cumplir con la LOTAIP: inobservan la normativa respecto a la difusión de la información pública que producen; no han institucionalizado procedimientos para atender de manera oportuna las solicitudes de acceso a la información pública ingresadas; difunden y entregan información pública incompleta, desactualizada y fuera de los plazos contemplados en la ley; no implementan programas de difusión y capacitación para los servidores públicos y la ciudadanía.

Adicionalmente, se anota que las organizaciones políticas no transparentan la información relacionada al uso detallado de los fondos públicos asignados, así como de los fondos privados y que el Consejo Nacional Electoral no ha publicado la información relacionada al gasto electoral de los procesos de 2017 y 2018 (CPCCS, 2019).

En cuanto a la ciudadanía, se anota que esta no hace uso de la información pública como insumo para incidir en la toma de decisiones de interés público y ejercer el control social (CPCCS, 2019), y esto se explicaría debido a que la transparencia activa actualmente no se corresponde con las necesidades y demandas o que estos no son fáciles de procesar.

En este contexto, con avances y retrocesos, la Defensoría del Pueblo realizó una propuesta para una nueva Ley de Transparencia y Acceso a la Información Pública, luego de un ejercicio de co-creación con diversas entidades públicas, organizaciones de la sociedad civil y la academia, la cual se planteó en el marco de la ley 2.0 de transparencia y acceso a la información pública, propuesto por la Organización de Estados Americanos (OEA), enfatizando en las nuevas tecnologías de información y comunicación para facilitar el ejercicio del derecho (Defensoría del Pueblo, 2021).

En esta línea, es importante anotar que las plataformas digitales pueden mejorar la eficiencia del sector público y democratizar el acceso para su uso; no obstante, para efectivizarlas, se debe avanzar en el reconocimiento del valor de los datos, la regulación de su uso y reúso, tanto a través de tecnologías como de la inteligencia artificial (Cetina, 2021).

Por otro lado, en un diagnóstico sobre la inteligencia artificial (IA) se determinó que los proyectos implementados son escasos y muy básicos, aun cuando se acepta que la IA tiene mucho potencial en el marco de políticas públicas y en concordancia con las necesidades de los actores del ecosistema

³ <https://datosabiertos.compraspublicas.gob.ec/PLATAFORMA>

⁴ <https://www.ecuadorencifras.gob.ec/>

⁵ <https://www.datosabiertos.gob.ec/>

⁶ <https://aportecivico.gobiernoelectronico.gob.ec/>

digital. Pero, se espera poder identificar más iniciativas en el Ecuador como parte del proceso de elaboración de la Estrategia Nacional de IA (MINTEL, 2021).

Ahora, en la práctica, a raíz de la pandemia, muchos de los países comenzaron a publicar también los datos epidemiológicos sobre el coronavirus en sus páginas web con el propósito de responder mejor a la situación (Álamo et al., 2020) y, además, como una forma más eficaz de informar a la ciudadanía (Díez y Melero, 2022). No obstante, al mismo tiempo, en la región existió un bajo nivel de apertura de datos, ya que se registró un promedio de 40,38 de 100 puntos de calificación en el Barómetro de Datos Abiertos y Ecuador se encuentra dentro de dicho nivel (Fumega et al., 2021).

Por tanto, para el caso de Ecuador resulta de interés indagar en varios periodos, incluyendo el de la pandemia, sobre los cambios en las demandas de información y discutir sobre la utilidad de la información publicada para la incidencia ciudadana en la co-gestión o el control social y para que -bajo la nueva ley- estos requerimientos puedan ser atendidos desde la transparencia focalizada (TF). El concepto de la TF favorece la comprensión de las dinámicas de funcionamiento de la transparencia en áreas clave para la gestión pública, considerando tanto la efectividad y calidad de los canales o mecanismos de acceso, como las capacidades e intereses de los usuarios (Olabe y Cruz, 2011).

En este marco, el presente estudio tiene como objetivo analizar el avance y grado de coherencia entre lo dispuesto en la normativa y las necesidades actuales de la ciudadanía, a partir de las solicitudes de acceso (transparencia pasiva) en un caso seleccionado como relevante. Esto permitirá justificar la necesidad y guiar la implementación de mecanismos de transparencia focalizada.

En primer lugar, se revisa la definición, tipos y propósitos de la transparencia, así como las condiciones controlables a través del diseño de las políticas como perspectiva de análisis del estudio. Después, se explica el método seguido. Más tarde, se presentan los resultados que describen la evolución normativa como marco para la nueva propuesta ecuatoriana, aspectos de tecnología, orientaciones de los planes de desarrollo, se identifica la evolución de las demandas ciudadanas y se la discute desde la perspectiva de análisis, la oportunidad instrumental y contextual. Se finalizará con las conclusiones.

Perspectiva de análisis

Se entiende por transparencia efectiva aquella que permite un control eficaz de la acción pública, promueve la efectividad de las políticas públicas y con ello, genera confianza en el Gobierno (Villoria, 2021).

Por otra parte, la transparencia activa se basa en aportar bases de datos compatibles y reutilizables, e información que permita conocer y controlar el desempeño de lo público; actualmente, se pueden encontrar a través de plataformas integradas que proporcionan acceso a datos abiertos con la expectativa de que al procesarlos generen valor (Villoria, 2014; 2021). No obstante, en la práctica, cuando el cumplimiento de las obligaciones de publicidad activa y la inclusión de datos no incluyen la verificación de su calidad, sus efectos no son destacables (Villoria, 2018).

Al referirnos a la transparencia es importante resaltar su valor instrumental y contextual. De mejor forma, estos valores se evidenciaron en la pandemia, ya que se destacó el rol de las tecnologías basadas en datos para la toma de decisiones en salud pública y para la recuperación económica. Esto porque, para reducir el impacto sanitario de la crisis, se requerían los datos para: el rastreo y pronóstico de la propagación del COVID-19, la investigación científica y el diseño de estrategias, lo que lle-

vó a impulsar la gestión y apertura de esta tipología de datos (Cetina, 2021; Ball, 2020). Además de que, en contextos de gran demanda informativa, la transparencia es fundamental pues permite tomar decisiones informadas y monitorear el poder (Sierra, 2020).

Ahora bien, para fortalecer las capacidades digitales de los países se debe fortalecer la apertura de datos en el sector público, sin olvidar la adopción de estándares de ética sobre todo en el ámbito de la aplicación de la IA, así como la transparencia de los sistemas de IA y el derecho a la privacidad que tienen los ciudadanos (Cetina, 2021).

Aparecen así nuevos aspectos que deben ser considerados en el marco normativo de la transparencia y en los esquemas de la gobernanza de datos. Por ejemplo, cuando las autoridades públicas utilizan la toma de decisiones algorítmicas estos procesos de implementación deben incorporar mecanismos de transparencia en relación con los procesos algorítmicos, características y limitaciones, para evitar resquebrajar la ética y ralentizar el proceso de transformación digital.

Esta temática, entonces, es una discusión imprescindible en el actual contexto y en el proceso de construcción de leyes y regulaciones, ya que la disponibilidad de datos que demandan el big data y la IA tiene un rol esencial, pero también, se requiere evitar interferencia con otros derechos e intereses legítimos (Zlotnik, 2019, p. 27; Valero y Cerdá, 2020). En este sentido, no debemos olvidar que la gobernanza de datos es principalmente la gestión del ciclo de vida de los datos, lo que involucra aspectos de administración de los datos como activos, así como la responsabilidad y rendición de cuentas frente a su uso y gestión (GPAI, 2020).

En este marco, para la transparencia existen dos enfoques en la región, uno está relacionado con la sanción y puesta en práctica de leyes de acceso a la información que se conoce como el enfoque general; mientras que, el segundo está ligado al diseño y a la implementación de sistemas o mecanismos en sectores o áreas de intervención específicas, que se construyen a partir de la legislación existente, siendo este último el enfoque sectorial o focalizado y en el que se evidencia mayor facilidad de los usuarios para interpretar los beneficios concretos (Olabe y Cruz, 2011).

A su vez, las iniciativas a partir del enfoque general requieren que se atiendan durante la construcción de leyes, el aspecto relacionado con la entidad reguladora encargada de vigilancia y garantía de la ley, y el establecimiento y delimitación de excepciones, sin dejar de atender principalmente el interés público (Olabe y Cruz, 2011). Además, en su implementación debe existir un amplio conocimiento por parte de la sociedad, para lo cual es importante que esta participe en su proceso de construcción, caso contrario el desconocimiento ligado a la desconfianza en el órgano legislativo produce que por un lado la ciudadanía no exija su cumplimiento y, por otro lado, que se desmotive a su utilización.

Entre otros, lo acontecido en la pandemia reveló que cada sector tiene vulnerabilidades y dinámicas de funcionamiento propias, las cuales determinan los contenidos y formas de intervención para promover el acceso a la información pública. Es ahí cuando la transparencia focalizada (TF) favorece la comprensión y es útil al enfocarse en una audiencia en específico, a partir de las capacidades e intereses de los usuarios, con mecanismos simples y eficientes que faciliten el acceso, y el uso de dicha información (Olabe y Cruz, 2011; Fung, Graham y Weil, 2007).

Pero las capacidades e intereses de los usuarios no se mantienen estáticos, sino que cambian y evolucionan, al mismo tiempo que cambian o se ajustan las políticas públicas. Por ello, el marco normativo debe ajustarse en la medida que evoluciona la sociedad. Se necesita redefinir las obligaciones de divulgación de transparencia activa. Lo que actualmente se publica no es suficiente, no es útil. No

solo las formas de acceso, sino los contenidos, los enfoques y necesidades de información sectorial también han cambiado y la normativa debe reflejar la lógica contextual en cuanto a TF.

Entonces, el punto de arranque para la TF es la comprensión de las necesidades, prioridades y capacidades de los usuarios de la información. Ahora bien, las necesidades de información han cambiado por un lado debido a las actuales oportunidades tecnológicas, y, por otro lado debido al aumento del involucramiento ciudadano en los asuntos y las políticas públicas (Meijer y Thaenes, 2009).

Respecto a lo indicado, en Ecuador al existir el compromiso sobre la implementación de un modelo de gobierno abierto, entonces, se parte de la premisa de que se basa en los principios de participación ciudadana y uso de datos a partir de las posibilidades que ofrece la innovación tecnológica. Dicho de otro modo, en este contexto resulta imprescindible tener en cuenta las exigencias e incidencias del Gobierno Abierto (Valero y Cerdá, 2020).

La comprensión de estas necesidades es entonces un primer paso para mejorar la transparencia activa y, para ello, además del diálogo permanente que permite identificar estas demandas, resulta útil observar y analizar cómo han evolucionado los requerimientos de acceso a la información pública, ya que estos serían la voz y la evidencia – al menos formal- de estas necesidades.

Se precisa diseñar un modelo institucional e integrado basado en la transparencia y el acceso a la información conforme a las exigencias del Gobierno Abierto y, en particular, el principio de datos abiertos por defecto (Valero y Pardo, 2012, p. 459; Valero y Cerdá, 2020), de manera que se impulse definitivamente la co-creación de servicios digitales.

La información debe disponerse fácilmente y sin sesgo, debe tener interés para los ciudadanos, en lugar de publicar o disponer datos inservibles o no usables (Villoria, 2014; Arenas, 2018). Aunque, es justo aquello lo que suele darse, de ahí que se debe clarificar quién está detrás de los datos y su estructura (Etzioni, 2010; Arenas, 2018).

Metodología

El análisis documental se ha realizado para identificar y analizar los tópicos preseleccionados en los artículos científicos esenciales extraídos de la búsqueda y el marco legal nacional e internacional sobre transparencia y acceso a la información pública.

En relación con las técnicas de recolección de información, además de la búsqueda y extracción documental aplicada durante todo el proceso, se trabajó con base en un esquema construido a partir de los antecedentes teóricos y empíricos de las políticas de transparencia y acceso a la información pública en el país. El esquema está basado en un estudio diacrónico del marco legal e institucional.

Entre las fuentes relevantes de información respecto a las acciones de transparencia desarrolladas se encuentran la Defensoría del Pueblo, en particular los informes anuales publicados y los datos del sistema integrado de gestión defensorial; y los informes del Consejo de Participación Ciudadana y Control Social.

Se parte de que la comparación como método es un aporte al control empírico de los fenómenos políticos que incorpora elementos históricos, políticos, económicos y sociales para comprender la evolución de las demandas ciudadanas, así como los rasgos que en distintos contextos inciden en el rol de las organizaciones y la institucionalidad. Por ello, se selecciona como caso a la Presidencia de

la República y la unidad de análisis son las solicitudes de acceso a la información recibidas durante el periodo 2017-2021; ya que, dicho periodo abarca el periodo pre y pos pandemia ya relevado como importante, pero así mismo, por los rasgos que inciden en esta institución central y su rol preponderante durante la toma de decisiones.

Resultados y Discusión

Evolución normativa y propuesta ecuatoriana

Entre los Principios de Lima, adoptados por la Comisión Interamericana de Derechos Humanos en el año 2000, se incluyen los de acceso a la información como derecho humano, el acceso a la información en una sociedad democrática, transparencia y desarrollo, obligaciones de las autoridades, legislación sobre acceso a la información, excepciones al acceso a la información y protección legal al acceso a la información.

Respecto a transparencia y desarrollo, se indica que el acceso a la información es indispensable para el escrutinio y el debate sustentado sobre las acciones de gobierno, como condiciones para la transparencia en la gestión de las entidades públicas y para evitar la corrupción, así como otros abusos del poder. Se destaca que este derecho permite también alcanzar mayores niveles de participación en la vida política, económica, social y cultural en un país (OEA, 2000).

En esta línea, en el año 2006 la Corte Interamericana de Derechos Humanos, en el caso Claude Reyes y otros marcó un hito al reconocer en la Corte que el acceso a la información es un derecho humano que forma parte del derecho a la libertad de expresión. La información pertenece a las personas, no es propiedad del Estado. Entonces, en este sentido, el Estado debe adoptar las disposiciones legislativas o de otro nivel que fueren necesarias para promover y asegurar el reconocimiento y ejercicio de este derecho, además debe actuar con la debida diligencia en la promoción del acceso a la información, identificar a quienes deben proveer la información, prevenir los actos que lo nieguen y sancionar a sus infractores (OEA, 2007).

Este contexto legal ha llevado a su vez a desarrollar los respectivos marcos institucionales en la región para asegurar la transparencia y el cumplimiento del derecho de acceso a la información pública. Es así que, en Ecuador, en mayo de 2004 se publica la Ley Orgánica de Transparencia y Acceso a la Información Pública que establece que se difundirán a través de un portal o página web la información mínima actualizada a disposición del público, y que se desplegarán los medios necesarios implementados en la misma institución para tales efectos. Además, se le encarga a la Defensoría del Pueblo el seguimiento al cumplimiento de la Ley en cuanto al acceso, la promoción, vigilancia y garantías.

Durante el año 2010, se aprobó la Ley Modelo Interamericana sobre Acceso a la Información con el propósito de proporcionar a los Estados el marco legal necesario para garantizar el derecho de acceso a la información, así mismo, se presenta una guía de implementación (OEA, 2010). En las definiciones de este modelo ya se identifican diferencias con la Ley expedida años atrás en Ecuador. Es así que, la definición de información pública en la LOTAIP de 2004 se centraba en documentos, mientras que el modelo diferencia documentos e información.

Por otro lado, se evidencia diferencia en cuanto a la forma en la que deben presentarse las solicitudes de acceso a la información y los plazos de respuesta y prórrogas, siendo en ambos casos más flexible el modelo OEA.

Una década después, en marzo del año 2020, conforme Resolución de la Asamblea General de la OEA⁷ se presenta la propuesta de Ley Modelo Interamericana 2.0 sobre acceso a la información pública 2.0, misma que aborda temas identificados como prioritarios, que fueron sobre: régimen de excepciones, transparencia activa, gestión documental, órganos garantes, partidos políticos e información pública en el marco del poder judicial (OEA, 2020).

La nueva ley modelo destaca los principios *pro homine* e *in dubio pro actione*, es decir que persigue en su aplicación la interpretación más favorable al ejercicio del derecho protegido y enfatiza en un claro, detallado y preciso régimen de excepciones que deberán estar definidas por ley, ser legítimas y estrictamente necesarias en democracia. Se amplía el detalle de aspectos a publicar respecto a los servidores públicos, incluidas las declaraciones patrimoniales, sanciones recibidas, y detalle de permisos, licencias, comisiones, etc. Se destaca además que quienes realicen las solicitudes de acceso, a diferencia de lo dispuesto en la ley ecuatoriana vigente a esa fecha, pueden realizar solicitudes en forma anónima y sin tener que justificar las razones la petición.

Sin duda, en la nueva ley modelo se evidencia que los avances tecnológicos de los últimos años marcan una diferencia en cuanto a los diferentes tipos de registro que las instituciones públicas pueden disponer para una facilidad de acceso al ciudadano, y recoge la evolución en las necesidades ciudadanas y resultados que se han obtenido en cuanto al acceso a la información en los últimos años.

Por su parte en Ecuador, en el año 2020, la Defensoría del Pueblo, que posee iniciativa legislativa, presentó a la Asamblea Nacional el proyecto de ley sustitutiva a la LOTAIP, propuesta que para su elaboración contó con un proceso de co-creación en el que participaron diversos actores (Defensoría del Pueblo, 2020a). Adicionalmente, otros actores también presentaron propuestas ante la Asamblea, por lo que, dicha instancia dispuso la unificación de los proyectos de ley en materia de transparencia y acceso a la información pública, que se estaban tratando en las diferentes comisiones especializadas permanentes de la Asamblea Nacional. Actualmente, una vez completado el proceso legislativo correspondiente, en la nueva LOTAIP se definen los siguiente tipos:

“Transparencia Activa: (...) la obligación de las instituciones del sector público y de los demás sujetos establecidos en esta Ley, de mantener de forma permanente en el portal de información o sitio web, la información actualizada, suficiente y relevante, sin que sea necesario requerimiento alguno, por parte de autoridad competente o de las personas.

Transparencia Colaborativa: (...) la obligación que tienen las Instituciones del sector público y demás sujetos establecidos en esta Ley, de publicar información que surja de espacios de colaboración en los que la ciudadanía presente sus necesidades de información en base en sus demandas e intereses, bajo los principios del gobierno abierto y el Estado abierto.

Transparencia Focalizada: (...) la obligación que tienen las instituciones del sector público y demás sujetos establecidos en esta Ley, de no limitar la publicación de información a la mínima obligatoria establecida en la Ley, sino de publicar de manera proactiva, información y datos adicionales que puedan ser requeridos desde la ciudadanía, con estrategias de liberación en formato abierto relacionada con cuestiones específicas, cuyo propósito es mejorar el conocimiento sobre algún problema público, con el objeto de forta-

⁷ Resolución de la Asamblea General de la OEA AG/RES. 2905 (XLVII-O/17) sobre el Fortalecimiento de la Democracia, párrafo ix en la que se solicita al Departamento de Derecho Internacional la actualización o ampliación de la Ley Modelo Interamericana sobre Acceso a la Información Pública.

lecer el proceso de toma de decisiones ante situaciones complejas y una adecuada rendición pública de cuentas.

Transparencia Pasiva: (...) la obligación que tienen las instituciones del sector público y los demás sujetos establecidos en esta Ley, de responder a las solicitudes de información pública, previa solicitud de la interesada o interesado" (Asamblea Nacional, 2023: Art.4).

Se destaca también que la nueva LOTAIP ya incorpora entre los principios el *In Dubio Pro Actione*⁸ y el *In Dubio Pro Petitor*⁹. Además, se incluye el derecho a solicitar información sin tener que justificar las razones, y se amplía en algunos elementos las obligaciones de transparencia activa, el registro de activos de información que contenga información solicitada con frecuencia, así como otras de carácter específico sectoriales. En cuanto a la publicación de la información ya se dispone el formato de datos abiertos. Todo esto en concordancia con el modelo OEA.

No obstante, aunque se aclara que la transparencia activa no debe atenerse a disposiciones legales por debajo de la Ley, ni limita a los sujetos obligados a aplicar los preceptos de la transparencia focalizada, no queda clara la obligatoriedad y espacios para este tipo de transparencia que favorece la participación y toma de decisiones.

Transparencia y era digital

Los adelantos tecnológicos y, en general, la era digital han contribuido a generar transformaciones en los comportamientos y han incrementado la demanda de información por parte de las personas. De hecho, el funcionamiento tradicional de la administración pública se ha transformado, enfatizando en las prácticas de gobierno electrónico para responder a estas demandas mediante el incremento de las capacidades de almacenamiento, interoperabilidad, disposición, accesibilidad y difusión de la información, eliminando -o minimizando- con ello las restricciones creadas por los funcionarios públicos en cuanto al acceso.

En este contexto, consecuentemente, la Ley Modelo 2.0 se enmarca en que los sujetos obligados deberán permitir de la forma más amplia el acceso la información, con el propósito de que sea posible su interoperabilidad en un formato de datos abiertos y, también determinar las estrategias para la identificación, generación, organización, publicación y difusión adecuada de dicha Información, a fin de que sea reutilizable (OEA, 2020).

En la nueva LOTAIP, respecto a la calidad de la información, se especifica que será publicada o entregada en formato de datos abiertos permitiendo su recirculación y reutilización. Además de la información mínima que se debe colocar como transparencia activa, también se debe colocar en datos abiertos lo requerido de forma complementaria al Ministerio de Finanzas, las actas de votaciones en la Asamblea Nacional, los planes de trabajo de candidatos y los resultados de los procesos electorales por parte de la función electoral.

⁸ En caso de duda respecto de la interpretación de las normas, ésta debe resolverse atendiendo a la interpretación más favorable al derecho de acción, al derecho del interesado.

⁹ La interpretación de las disposiciones en la norma debe ser efectuada, en caso de duda, siempre en favor de la mayor vigencia y alcance del interesado.

Planes de Desarrollo y Transparencia

Otra forma de entender el estado de la transparencia en Ecuador y cómo han evolucionado las necesidades ciudadanas es destacar lo que prescriben los planes de desarrollo. Es así que, se destaca que a pocos años de la emisión de la LOTAIP, el Plan Nacional de Desarrollo 2007-2010, con un nuevo mandato en la temática de análisis, estableció entre sus objetivos el de garantizar el acceso a la participación pública y política entre cuyas metas de acción se encontraban las de promover el acceso a la información pública; propiciar la rendición de cuentas de los funcionarios públicos y la lucha contra la corrupción; promover mecanismos de codecisión y cogestión pública; y, generar información oportuna y precisa sobre las condiciones de vida de los ecuatorianos (Senplades, 2007).

En el nuevo período de Gobierno, el Plan Nacional de Desarrollo 2009-2013, en su objetivo diez referido a garantizar el acceso a la participación pública y política, identifica entre sus políticas la de fortalecer, generar e innovar formas de control social y rendición de cuentas a los mandantes; garantizar el libre acceso a información pública oportuna; y promover el desarrollo estadístico y cartográfico para la generación de información de calidad (Senplades, 2009).

Por su parte, el Plan Nacional para el Buen Vivir 2013-2017, en su objetivo uno referido a consolidar el Estado democrático y la construcción del poder popular, identifica entre sus políticas la de consolidar la participación ciudadana en los procesos de elaboración de políticas públicas y en el relacionamiento Estado-sociedad; y fortalecer los mecanismos de control social, la transparencia de la administración pública y la prevención y la lucha contra la corrupción (Senplades, 2013).

El Plan Nacional de Desarrollo 2017-2021, en su eje tres, más sociedad mejor Estado, y su objetivo ocho prescribe la promoción de la transparencia y la corresponsabilidad para una nueva ética social, y dentro de este se destaca la política de fortalecer la transparencia en la gestión de instituciones públicas y privadas y la lucha contra la corrupción, con mejor difusión y acceso a la información pública de calidad (Senplades, 2017).

Finalmente, el Plan de creación de oportunidades 2021-2025, dentro de su eje institucional incluye el objetivo quince que se enfoca en fomentar la ética pública, la transparencia y la lucha contra la corrupción¹⁰, y sus políticas que buscan fomentar la integridad pública y la lucha contra la corrupción en coordinación interinstitucional efectiva entre todas las funciones del Estado y la participación ciudadana, además de impulsar el Gobierno Abierto que propicie la transparencia y el acceso de información oportuna y cercana a la ciudadanía (Secretaría Nacional de Planificación, 2021).

Calidad de la Transparencia y pandemia

Al 2020, según informe de la Defensoría del Pueblo conforme el catastro de 1560 entidades públicas y privadas que reciben fondos públicos; 1459 entidades presentaron el informe anual, de acuerdo con la LOTAIP, es decir, se cumplió en un 94 % (Defensoría el Pueblo, 2021); sin embargo, este informe no aborda la calidad de la información. Es así que, se evidencia que durante la pandemia se exhortó al Estado ecuatoriano a que cumpla la obligación de garantizar el derecho de acceso a la información sobre los sucesos y procesos de interés general relacionados con esta emergencia sanitaria, sin que existan restricciones de información, a excepción de aquella confidencial y/o reservada;

¹⁰ <http://oportunidades.planificacion.gob.ec/Plan2125/eje-instituciones/>

y, que para este efecto se garantice el acceso universal a las tecnologías de la información y comunicación (Defensoría el Pueblo, 2020b).

En ese contexto, también se emitieron dictámenes correctivos a diversas entidades, especialmente en el marco de la emergencia sanitaria por COVID-19. Entre los principales está el remitido al Centro de Operaciones de Emergencias Nacional (COE) con recomendaciones para ser aplicadas como ajustes de la sección correspondiente en el portal institucional del Sistema Nacional de Gestión de Riesgos y Emergencias¹¹ y de la página www.coronavirusecuador.com. Se solicitó publicar información relacionada con: inscripciones de defunciones registradas, a escala nacional; defunciones clasificada por COVID-19 y otras causas; información sobre los procesos de contratación pública que gestiona directamente el COE nacional y sus integrantes en el marco de la emergencia y el estado de excepción (Defensoría el Pueblo, 2021).

El otro dictamen relevante fue el remitido al Ministro de Salud Pública por los reiterados incumplimientos frente a la LOTAIP, solicitándole que aplique las medidas necesarias, y que se publique en el portal web institucional la información mínima obligatoria que se establece en la Ley, y de ser el caso, se emitan sanciones aplicables. También, debido a la implementación de las jornadas especiales en el sector público por la crisis sanitaria en el país, se emitió una directriz para el cumplimiento de la transparencia activa en cuanto a la publicación de las remuneraciones en el sector público en una matriz homologada (Defensoría el Pueblo, 2021). Cabe resaltar, la importancia de la información, ya que mientras desde el 2016 hasta antes de la pandemia en Ecuador se registraban 15.669 teletrabajadores, durante la emergencia sanitaria el total de teletrabajadores se incrementó a 442.305 (MINTEL, 2021).

En lo específico, en cuanto a los datos de mortalidad y morbilidad de la población, se instó al Estado ecuatoriano a que se implemente un componente unificado de información sobre las cifras del Servicio Nacional de Gestión de Riesgos y Emergencias, el Ministerio de Salud Pública y el Registro Civil, a fin de establecer la realidad del impacto de las defunciones causadas por la COVID-19. Así mismo, se solicitó al Estado la depuración de la información para evitar los subregistros existentes de casos de la COVID-19, y con ello generar políticas públicas alineadas con los objetivos nacionales (Defensoría el Pueblo y UTE, 2020).

Se requería información clara para evitar desórdenes y desmanes que complicaban el control sanitario y de seguridad que era provocado por la desinformación. En este sentido, también es posible evidenciar cómo la transparencia y el acceso a la información tienen un carácter instrumental al habilitar otros derechos, como en este caso el derecho a la salud.

¿Por qué no tenemos datos públicos y si los tenemos no los usamos? Problemas para la Gobernanza de datos

En abril de 2020 la palabra coronavirus arrojaba más de 9.300 millones de resultados de búsqueda en Google (Cetina, 2020). Esta demanda de información requería ser atendida, pero en su momento existía un escaso conocimiento del COVID-19, generándose entonces desinformación y más letalidad. Cada Estado debía desarrollar la capacidad de generar y disponer de datos confiables para la toma de decisiones del Gobierno basada en evidencia, pero así mismo, para alertar a la ciudadanía y posibilitar el ejercicio del escrutinio de lo público, la colaboración de actores y la innovación pública.

¹¹ www.gestionderiesgos.gob.ec

No obstante, conforme lo indicado en los mismos informes de la Defensoría del Pueblo de Ecuador, los datos cuando existían tenían problemas de calidad y utilidad. Entonces, tenemos un doble problema, por un lado, la escasez de datos y por otro, la calidad de datos (confiabilidad y usabilidad). Desde ambos lados, esta situación incide en la insatisfacción de las necesidades ciudadanas y más aún cuando estas evolucionan, mientras que la oferta activa de datos del sector pública se detiene en un mínimo establecido hace casi dos décadas, en el mejor de los casos.

Por ello, el Barómetro de Datos Abiertos afirma que existen tres áreas de política pública que deben atenderse para que los datos sean útiles en el ciclo de la política pública y para la construcción de una gobernanza: apertura por defecto, infraestructura de datos y datos con un propósito (Cetina, 2021). Es en esta última área, la que interesa en este estudio. Porque, para que los datos tengan un propósito se requiere trabajar en colaboración con el usuario final, pero así mismo se requiere que este usuario, en plena era digital, tenga formas de acceso a los datos.

En Ecuador, para el período 2021 al 2025, la perspectiva es que la penetración de internet móvil y fijo a nivel nacional tenga un incremento de 9,92 y, en este sentido, teniendo una línea base al 2021 de 75,36% de parroquias rurales que cuentan con servicio de acceso a internet, se pretende al 2025 alcanzar un 77,21% ; en cuanto al servicio móvil avanzado, mientras al 2021 el 74,15% de parroquias rurales se encuentran conectadas, se pretende al 2025 alcanzar un 79%; y, el porcentaje de hogares con servicio de internet fijo en las zonas rurales busca pasar de 22,12% al 2021, a un 30,68% al 2025 (MINTEL, 2022). Se observa así que, al 2025 las brechas siguen estando vigentes de forma ostensible, si observamos sobre todo las zonas rurales.

Esto es aún más grave si tenemos en cuenta que, los planes de datos y los dispositivos con internet no son asequibles para las personas pobres de la región. En promedio, el costo de un plan de datos de solo 1GB representa el 2,7 % del ingreso familiar mensual (o entre el 8 % y el 10 % para el quintil inferior en algunos países). Además, el costo del teléfono inteligente básico más barato disponible representa entre el 4 % y el 12 % del ingreso familiar promedio en gran parte de la región (Drees-Gross y Zhang, 2021).

Esto nos muestra y enfatiza la conformación de élites que puede acceder y procesar los datos para realizar sus aportes a las políticas públicas; pero a su vez que, mientras por un lado las tecnologías nos ayudan a cada vez más exigir nuevos datos públicos de forma activa, por otro lado, esta mirada de datos necesarios tiene el sesgo de lo urbano, de la élite y de las demandas desde sus situaciones, por lo tanto aún es muy difícil hablar de un arreglo entre la demanda y la oferta de datos públicos.

Evolución de las necesidades ciudadanas para el caso

De la Tabla 1, se registra que el 12.67% del total de solicitudes corresponde al requerimiento de acceso a informes que respaldan las decisiones o la ejecución de las políticas públicas o que sustenten la toma de decisiones. Entre estos: los informes de deuda, de préstamos, de concesiones, de situación ante el COVID-19, etc. Al respecto, se evidencia que esta constante necesidad ciudadana está incluida parcialmente en la nueva LOTAIP en su artículo 19 sobre transparencia activa respecto a información sobre las concesiones, contratos de créditos, pasivos, planes y programas; aunque no se específicamente informes completos.

Se destaca que durante el periodo 2017-2021, el 11,98% de las solicitudes corresponden a información del personal, como: acciones de personal, documentos de respaldo de autorizaciones de viaje, comisiones de servicios, contratos de servicios profesionales, desvinculaciones, informes de concur-

sos, respaldos de viáticos, información de sumarios administrativos, evaluaciones de desempeño, etc. Cabe indicar que esta información solicitada no se incluye en los portales, ya que se encuentran fuera del mínimo dispuesto por la LOTAIP como transparencia activa vigente a dicha fecha.

Tabla 1. Solicitudes de Acceso a la Información Pública. Presidencia de la República 2017-2021

Tipo de Información/ Categoría	Porcentaje %
Informes	12,67
Personal	11,98
Contratación Pública	7,81
Procesos de pago	6,08
Investigaciones	5,20
Gestión Administrativa	4,34
Documentos ingresados	4,17
Documentos emitidos	2,95
Contraloría	2,60
Actas	0,87
Base Legal Interna	0,52
Agenda	0,52
Audiencias	0,17
Otras	40,10

Fuente: Defensoría del Pueblo (2022). Elaboración: Autora

Mientras que, de estos, en la nueva LOTAIP, en relación con esta identificación se agrega: “*Un detalle de los funcionarios que gocen de licencia de servicio y de comisión de servicio*” (Asamblea Nacional, 2023: Art. 19.4). Queda claro, el interés de la ciudadanía en indagar sobre la legalidad y legitimidad de diversas acciones que se relacionan con la gestión de personal. En ese sentido, resulta importante su inclusión en la nueva ley, ya que esto contribuirá a que los ciudadanos mejoren su confianza en la conformación del servicio público, y en la ejecución del gasto al que corresponde este rubro.

Respecto al rubro de 7,81% en la categoría de contratación pública, lo que se solicitan son los expedientes de la fase de adjudicación y los respaldos de la fase contractual. Cabe indicar en este punto, que, aunque desde la derogada LOTAIP ya se establecía su publicación, no obstante, esta información no se publica en detalle. Los formatos homologados para obtener dicha información remiten al ciudadano a un link de la página del sistema oficial de contratación pública de Ecuador, SOCE, en el cual si bien existe información valiosa, esta no corresponde a los expedientes completos del proceso.

Otra consideración es que lo indicado se publica dentro del portal de cada entidad a mes vencido, de otro modo es necesario acudir directamente a la página del SOCE, pero conociendo el funcionamiento del sistema y algunos códigos para agilizar la búsqueda. Se verifica también que, pese a que las fechas se encuentran vencidas para la fase de ejecución de contrato, no se encuentra actualizada la información.

No obstante, cabe indicar que, a partir de 2020, el Servicio Nacional de Contratación Pública, SERCOP, implementó, en el marco del primer plan de acción de Gobierno Abierto de Ecuador, la plataforma de contratación abierta, a través de la implementación del estándar de datos.

En esta misma línea, el 6.08% corresponde a solicitudes que requieren el respaldo de procesos de pago. Por lo que, se evidencia la necesidad de la ciudadanía de conocer la legalidad y comprobar la

eficiencia en el uso de recursos públicos. El 5,21% corresponde a información de investigaciones en curso, principalmente en lo relacionado con casos de gran impacto público y mediático.

Un dato que alarma es el 40,10% en categoría de 'Otras' solicitudes, mismas que al estar descritas de una forma muy general en el sistema de seguimiento no permiten su identificación exacta, respecto a si se refieren o no a solicitudes de acceso a la información. De las que sí se obtuvo alguna referencia, se identifica que se refieren a registros internos, resoluciones, grabaciones de enlaces ciudadanos, etc.

Complementariamente, de la demanda anual, que se muestra en la Tabla 2, se observa la consistencia del requerimiento ciudadano a la máxima autoridad del Ejecutivo sobre el acceso a los informes que evidencian el cumplimiento de sus facultades. Así mismo, se verifica la disminución del total de solicitudes, en el marco de la pandemia, aspecto atribuido a las diferentes preocupaciones de la ciudadanía en tan compleja situación, pero, a su vez, su especificidad respecto a los requerimientos ante el Ejecutivo en su afán de indagar en las causas de los resultados económicos y sociales.

Tabla 2. Demanda anual de solicitudes de Acceso a la Información Pública.
Presidencia de la República

Año	Total Solicitudes	Categoría 1 ¹²	%	Categoría 2 ¹³	%
2017	35	Personal	14,29	Documentos emitidos	8,57
2018	218	Personal	16,59	Investigaciones	10,14
2019	216	Informes	13,89	Contratación Pública	11,11
2020	57	Informes	21,05	Personal	7,02
2021	51	Informes	29,41	Contratación Pública	7,84

Fuente: Defensoría del Pueblo (2022). Elaboración: Autora

Se evidencia de la revisión de la ley vigente, que se han incluido algunas de las necesidades ciudadanas como respuesta a la evolución, mismas que también se encuentran planteadas en la propuesta de ley modelo OEA 2.0. Sin embargo, asociado con las necesidades identificadas durante la pandemia se anota, por ejemplo, en la ley vigente una obligación que es necesaria, pero que no se define adecuadamente, y que, por tanto, sería vulnerable a la inclusión de generalidades, esta indica "*Políticas públicas o cualquier información que afecte a un grupo específico, en todas sus fases*" (Asamblea Nacional, 2022: Art.19.21).

Otros aspectos que están considerados parcialmente son los relacionados con la necesidad de identificar las demandas ciudadanas frecuentes y debido a aquello exigir su publicación como transparencia activa o focalizada. Así, en la ley vigente se identifica lo siguiente:

"20. Registro de Activos de Información, que contenga información solicitada con frecuencia, y otra información complementaria que de carácter obligatorio deban cumplir los sujetos obligados, de acuerdo con los parámetros establecidos por la Defensoría del Pueblo;

24. Otra información que la entidad considere relevante para el ejercicio del derecho de acceso a la información pública, la participación ciudadana y el control social, en especial la que permita

¹² La primera más alta excluyendo la categoría de "otras"

¹³ La segunda más alta excluyendo la categoría de "otras"

el seguimiento a la consecución de los Objetivos de Desarrollo Sostenible.” (Asamblea Nacional, 2023: Art.19).

Esto permitiría, por un lado, atender las demandas ciudadanas reconociendo su evolución, pero, además identificar necesidades específicas hacia una transparencia focalizada. Se sostiene, sin embargo, que la ley no anota el detalle de obligaciones para grupos específicos que permitan exigir esta información desde los sujetos obligados.

Conclusiones

El ejercicio de los derechos de ciudadanía requiere el acceso a la información, por lo que el Estado debe disponer de los instrumentos que asegure tal ejercicio. Además, en el contexto actual, la demanda de transparencia se ha visto acelerada debido a varios factores, entre los que se encuentran la pérdida de legitimidad debido a los resultados evidenciados durante la pandemia. Esta afirmación es consistente con el hallazgo del estudio sobre las solicitudes de acceso a la información que se centran en informes que muestren los sustentos que respalden las decisiones del gobierno y su desempeño.

Entonces, surge la necesidad de gestionar mejor la información y datos en el sector público, para dar respuesta a las demandas sociales de transparencia e incidir en la calidad de los servicios públicos. Así, la transparencia se debe estructurar como una respuesta a los ciudadanos para participar en las actividades públicas, en la toma de decisiones y en la gestión de asuntos públicos; y, además, como un modo de reducir la corrupción.

Si se requiere que la transparencia cumpla con su propósito, entonces, no tiene sentido tener una gran cantidad de documentos almacenados, si no existe una ley que los haga accesibles, de modo que fomente la consulta, colaboración y la responsabilización. Además, la demanda de transparencia es un reclamo básico para avanzar hacia formas más participativas en un modelo de Gobierno Abierto como el adoptado por Ecuador desde el 2018.

Pero, mientras no se procesen los requerimientos ciudadanos, labor que hoy en día se facilita con el uso de las tecnologías de información y comunicación y la inteligencia artificial, se seguirán evidenciando las brechas entre las necesidades de información y lo que el sector público entrega como transparencia activa. Esto da como resultado, la falta de información sobre la ejecución y gestión de las políticas públicas; y en general el déficit en la calidad de datos publicados en los sitios web.

En este sentido, a partir de la transparencia focalizada se deben crear incentivos para que se suministre información a un grupo de usuarios, a fin de que estos puedan tomar decisiones informadas. Pero además, para que se genere valor agregado, se debe partir de una adecuada identificación de los riesgos que se desea reducir o mitigar, así como mejorar las capacidades de acceso y procesamiento de los usuarios de información.

Por otro lado, considerando la existencia de múltiples actores en el ciclo de las políticas de transparencia se deben solucionar las asimetrías de información, también, desarrollando mecanismos de coordinación entre los actores involucrados, es decir una adecuada gobernanza. 🌐

Referencias

- Álvarez-Flores, E. P., Núñez-Gómez, P., & Mañas-Viniegra, L. (2021). Efecto híbrido en la demanda del profesional publicitario: Un reto ante la transformación digital en la publicidad. *Palabra Clave*, 24(2), e2424-e2424. <https://doi.org/10.5294/pacla.2021.24.2.4>
- Antunes, C., & Sebastiao, S. P. (2020). Desafíos éticos en las redes sociales en línea: La producción de contenidos y la opinión de los profesionales. *Cuadernos.info*, 46, 222-248. <https://doi.org/10.7764/cdi.46.1473>
- Arenas, M. 2018. "Participación y Transparencia en la Sociedad digital". En: Sánchez, M y D. Fernandez (Coord). *Apuntes sobre la Transparencia*. UCM, ACREDITRA.
- Asamblea Nacional. 2023. Ley Orgánica de Transparencia y Acceso a la Información Pública. Registro Oficial No.245, 7 de Febrero 2023
- Ávila, C., Semería, A., Malache, L. (2011). Ejercicio de las relaciones públicas: el "lado oscuro de la fuerza" [Review of Ejercicio de las relaciones públicas: el "lado oscuro de la fuerza"]. *UNIVERSIDAD VERDAD / No 55, No 55, p. 89-100.*
- Aydin, B. O., Gürbüz, S., & Dugan, O. (2020). Public relations in the networked publics. *Public relations in the networked publics* (pp. 1-284) doi:10.3726/b16513 Retrieved from www.scopus.com
- Ball, W. P. 2020. "Open Data, Open Science and Transparency in the time of COVID 19". *Radical Statistics*, 127, 4-9
- Bizberge, A., & Segura, M. S. (2020). Los derechos digitales durante la pandemia COVID- 19 en Argentina, Brasil y México. *Revista de Comunicación*, 19(2), 61-85. <https://doi.org/10.26441/RC19.2-2020-A4>
- Castillo Esparcia, A., & Smolak Lozano, E. (2017). Relaciones públicas digitales: Análisis de las estrategias de comunicación de los think tanks. *Obra digital*, 13, 59-80. <https://doi.org/10.25029/od.2017.158.13>
- Cetina, C. 2020. Tecnología para la integridad en tiempos del COVID-19. Policy Brief #8. CAF.
- Cetina, C. 2021. Gobernanza de datos y capacidades estatales para la pospandemia. Policy Brief # 28. CAF.
- Congreso Nacional. 2004. Ley Orgánica de Transparencia y Acceso a la Información Pública. Registro Oficial Suplemento 337 de 18-may.-2004
- CPCCS. 2019. Propuesta del Sistema Nacional de Transparencia y Lucha contra la Corrupción. Subordinación Nacional de Transparencia,
- Cuenca-Fontbona, J., & Puentes-Rivera, I. (2019). Relaciones públicas avanzadas: Confianza. 7.
- Cuenca-Fontbona, J., Matilla, K., & Compte-Pujol, M. (2020a). Transformación digital de los departamentos de relaciones públicas y comunicación de una muestra de empresas españolas. *Revista de Comunicación*, 19(1), 75-92. <https://doi.org/10.26441/RC19.1-2020-A5>

- Defensoría del Pueblo y Universidad UTE. 2020. Informe temático sobre la situación de los derechos humanos durante la declaratoria del estado de excepción por la emergencia sanitaria por COVID-19 en Ecuador en 2020.
- Defensoría del Pueblo. 2020a. Noticias. Defensor del Pueblo presenta a la Asamblea Nacional el Proyecto de Ley Sustitutiva a la Ley Orgánica de Transparencia y Acceso a la Información Pública.
- Defensoría del Pueblo. 2020b. Defensoría del Pueblo exhorta a que se garantice el acceso a información veraz, verificada, oportuna y contextualizada sobre la pandemia de Covid-19. Boletín
- Defensoría del Pueblo. 2021. Informe anual de rendición de cuentas 2020.
- Defensoría del Pueblo. 2022. Reporte General de literal B. Detalle de las solicitudes de acceso a la información y el trámite dado a cada una de ellas. Sistema Integrado de Gestión Defensorial.
- Díez-Garrido, M., y Melero L., M. 2022. "La transparencia como medicina. La comunicación de la pandemia a través de los portales de datos abiertos de las autonomías". *Estudios sobre el Mensaje Periodístico* 28 (1), 621-637. <https://dx.doi.org/10.5209/esmp.77172>
- Drees-Gross, F y Zhang, P. 2021. El escaso acceso digital frena a América Latina y el Caribe ¿Cómo solucionar este problema?. Banco Mundial Blogs. <https://blogs.worldbank.org/es/latinamerica/el-escaso-acceso-digital-frena-america-latina-y-el-caribe-como-solucionar-este>
- Etzioni, A. 2010. "Is Transparency the Best Disinfectant?". *Journal of Political Philosophy*, Volume 18, Issue 4: 389-404.
- Ferrari, M. A., & Durán, A. M. (2019). Relaciones Públicas y Sustentabilidad: estudio en organizaciones brasileñas y ecuatorianas. *Obra digital: revista de comunicación*, (16), 29-41.
- Fumega, S., Scrollini, F., Zapata, E. 2021. ¿Cuán abiertos están los datos públicos? El barómetro de datos abiertos de América Latina y el Caribe 2020. Caracas: CAF. Retrieved from: <https://scioteca.caf.com/handle/123456789/1710>
- Fung, A., M. Graham y D. Weil. 2007. *Full Disclosure: The Perils and Promise of Transparency* Cambridge, Reino Unido: Cambridge University Press.
- Gökalp, E., Şener, U., & Eren, P. E. (2017, October). Development of an assessment model for industry 4.0: industry 4.0-MM. In *International Conference on Software Process Improvement and Capability Determination* (pp. 128-142). Springer, Cham.
- GPAI. 2020. "Data Governance Working Group A Framework". Paper for GPAI's work on Data Governance. Available: <https://gpai.ai/projects/data-governance/gpai-data-governance-wg-report-november-2020.pdf> 11. Guidelines for AI Procurement.
- Grunig, J. E., & Hunt, T. T. (1984). *Managing public relations*. Holt, Rinehart and Winston.
- LCM, (2022). Álamo, T., Reina, D.G., Mammarella, M., y Abella, A. 2020. Open data resources for fighting covid-19. arXiv preprint arXiv:2004.06111.

- Meijer, A., y Thaens, M. 2009. "Public information strategies: Making government information available to citizens". *Information Polity*, 14(1,2), 31-45. doi:10.3233/ip-2009-0167
- MINTEL. 2021. Agenda Digital Ecuador 2021-2025. Quito-Ecuador
- MINTEL. 2022. Plan de Servicio Universal 2022-2025. Quito -Ecuador
- Navarro, C., Moreno, A., Molleda, J. C., Khalil, N., & Verhoeven, P. (2020). The challenge of new gatekeepers for public relations. A comparative analysis of the role of social media influencers for european and latin american professionals. *Public Relations Review*, 46(2) doi:10.1016/j.pubrev.2020.101881
- OEA . 2007. Estudio Especial sobre el derecho de acceso a la información. Relatoria espacial para la libertad de expresión. En:
<http://www.cidh.oas.org/relatoria/section/estudio%20oespecial%20sobre%20el%20derecho%20de%20acceso%20a%20la%20informacion.pdf>
- OEA. 2000. Principios de Lima. En: <https://www.oas.org/es/cidh/expresion/showarticle.asp?artID=158&IID=2>
- OEA. 2010. Ley Modelo Interamericana sobre Acceso a la Información. Documento presentado por el Grupo de Expertos sobre Acceso a la Información coordinado por el Departamento de Derecho Internacional, de la Secretaría de Asuntos Jurídicos, de conformidad con la resolución AG/RES. 2514 (XXXIX-O/09) de la Asamblea General)
- OEA. 2020. Propuesta de Ley Modelo Interamericana 2.0 sobre Acceso a la Información Pública 2.0. 96º. Período ordinario de sesiones, OEA/Ser.Q. CJI/doc. 607/20
- Olabe, P. B., y Vieyra, J. C. 2011. Acceso a la información y políticas de transparencia focalizada. Banco Interamericano de Desarrollo.
- Pereira Villazón, T., Portilla Manjón, I., & Rodríguez Salcedo, N. (2019). Big data y Relaciones Públicas. Una revisión bibliográfica del estado de la cuestión. *Revista de Comunicación*, 18(1).
<https://doi.org/10.26441/RC18.1-2019-A8>
- Permatasari, A. N., Soelistiyowati, E., Suastami, I. G. A. P. P., & Johan, R. A. (2021). Digital Public Relations: Trend and Required Skills. *Jurnal ASPIKOM*, 6(2), 373. <https://doi.org/10.24329/aspikom.v6i2.836>
- Secretaria Nacional de Planificación. 2021. Plan de Creación de Oportunidades 2021-2025.
- Senplades. 2007. Plan Nacional de Desarrollo 2007-2010.
- Senplades. 2009. Pan Nacional del Buen Vivir 2009-2013.
- Senplades. 2013. Plan Nacional del Buen Vivir 2013-2017..
- Senplades. 2017. Plan Nacional de Desarrollo 2017-2021 Toda una Vida.
- Sierra Rodríguez, J. 2020. "Periodismo, comunicación institucional y transparencia: aprendizajes de la crisis sanitaria del Covid-19". *Revista de Comunicación y Salud*, 10(2), 569-591.

- Thelen, P. D. (2021). The status of public relations research addressing latin america: A content analysis of published articles from 1980 to 2020. *Public Relations Review*, 47(4) doi:10.1016/j.pubrev.2021.102079
- Vega, C. R., & Rodríguez, J. C. (2019). Barreras de las pymes en colombia para la implementación de estrategias y herramientas de marketing digital. 38.
- Valero Torrijos, J. y Pardo Lopez, M.M. 2012. "Institutional Backing and PSI Reuse: Is the EU Going Too Far or Just Going in the Wrong Way?". *Masaryk University Journal of Law and Technology (MUJLT)*, 6(3), pp. 455-470.
- Valero, J. y Cerdá, I. 2020. "Transparencia, acceso y reutilización de la información ante la transformación digital del sector público: enseñanzas y desafíos en tiempos del COVID-19." *Eunomía. Revista en Cultura de la Legalidad*, 19, pp. 103-126.
- Villena, E. (2018). El Influencer de moda como sujeto de relaciones públicas: recursos y herramientas, Redmarka. *Revista de Marketing Aplicado*, vol 01, núm. 022, 115- 128.
<https://doi.org/10.17979/redma.2018.01.022.4938>
- Villoria, M. 2021. "¿Qué condiciones favorecen una transparencia pública efectiva? Artículo de revisión". *Revista de Estudios Políticos*, 194, 213-247. doi: <https://doi.org/10.18042/cepc/rep.194.08>
- Villoria, M. 2018. "El reto de la transparencia". *Anuario de Transparencia Local*, 1, 15-41.
- Villoria, M. 2014. "La transparencia como política pública en España: algunas reflexiones". *EUNOMÍA. Revista en Cultura de la Legalidad*, 85-103.
- Westerman, G.; Bonnet, D.; McAfee, A. (2014). *Leading digital: Turning technology into business transformation*. Boston, MA: Harvard Business Review Press. ISBN: 978 1 62527 247 8
- Zhou, A., & Xu, S. (2022). Digital public relations through the lens of affordances: A conceptual expansion of the dialogic principles. *Journal of Public Relations Research*, 1-19.
<https://doi.org/10.1080/1062726X.2022.2046585>.
- Zlotnik, A. 2019. "Inteligencia Artificial en las Administraciones Públicas: definiciones, evaluación de viabilidad de proyectos y áreas de aplicación". *Boletín*, 84, pp. 24-32. *Latin American Communication Monitor*. <https://latincommunicationmonitor.com/>

Sobre las autora/ About the author

Irma Jara Profesora-Investigadora de la Escuela de Gobierno y Administración Pública del Instituto de Altos Estudios Nacionales. Doctora en Administración por la Universidad Andina Simón Bolívar, y MBA de la Universidad de Québec en Montréal. Realizó un posdoctorado en América Latina en el orden global. Sus líneas de investigación son: Gestión Pública y Transparencia; Rotación de Directivos Públicos.

URL estable documento/stable URL

<http://www.gigapp.org>

El Grupo de Investigación en Gobierno, Administración y Políticas Públicas (GIGAPP) es una iniciativa impulsada por académicos, investigadores y profesores Iberoamericanos, cuyo principal propósito es contribuir al debate y la generación de nuevos conceptos, enfoques y marcos de análisis en las áreas de gobierno, gestión y políticas públicas, fomentando la creación de espacio de intercambio y colaboración permanente, y facilitando la construcción de redes y proyectos conjuntos sobre la base de actividades de docencia, investigación, asistencia técnica y extensión.

Las áreas de trabajo que constituyen los ejes principales del GIGAPP son:

1. Gobierno, instituciones y comportamiento político
2. Administración Pública
3. Políticas Públicas


Información de Contacto

Asociación GIGAPP.
ewp@gigapp.org

Lineamientos para el uso de inteligencia artificial en contextos universitarios (v 5.0)

Gutiérrez, Juan David

Universidad de los Andes, Colombia

 juagutie@uniandes.edu.co

ORCID ID: [0000-0002-7783-4850](https://orcid.org/0000-0002-7783-4850)

Resumen

Este instrumento ofrece una orientación a profesores y estudiantes para que el uso de inteligencia artificial (IA) en contextos universitarios sea informado, transparente, ético, y responsable. Los Lineamientos proponen reglas, principios y ejemplos que son aplicables para las actividades pedagógicas realizadas dentro y fuera del aula. El texto está precedido de un preámbulo que busca poner en contexto al lector mediante una breve explicación sobre la historia de la formulación y desarrollo de los Lineamientos. El instrumento está dividido en siete secciones: (1) introducción, (2) objetivos, (3) reglas de uso dentro y fuera del aula, (4) justificación de las reglas, (5) otros recursos para profesores/as, (6) apertura a futuros cambios, y (7) cambios respecto de anteriores versiones de los lineamientos. Aunque los lineamientos son pertinentes para el uso de cualquier tipo de IA, sus reglas se enfocan en el uso de chatbots –tales como ChatGPT, Bing, Claude, y Bard, entre otros– que permiten a los estudiantes acceder a los modelos de lenguaje a gran escala. El objetivo de los lineamientos es convertirse en un recurso útil para quienes precisen de un documento fundamentado, comprensivo, y estandarizado que permita establecer guías de trabajo claras respecto del uso de IA en la universidad.

Palabras clave

inteligencia artificial, inteligencia artificial generativa, modelos de lenguaje a gran escala, ChatGPT, competencias digitales, educación universitaria.

Resumo

Este instrumento oferece orientações a professores e alunos para que o uso da inteligência artificial (IA) em contextos universitários seja informado, transparente, ético e responsável. As Diretrizes propõem regras, princípios e exemplos aplicáveis às atividades pedagógicas desenvolvidas dentro e fora da sala de aula. O texto é precedido por um preâmbulo que busca contextualizar o leitor por meio de uma breve explanação sobre a história da formulação e desenvolvimento das Diretrizes. O instrumento está dividido em sete seções: (1) introdução, (2) objetivos, (3) regras de uso dentro e fora da sala de aula, (4) justificativa das regras, (5) outros recursos para professores/ assim, (6) abertura para mudanças futuras e (7) mudanças de versões anteriores das diretrizes. Embora as diretrizes sejam relevantes para o uso de qualquer tipo de IA, suas regras se concentram no uso de chatbots – como ChatGPT, Bing, Claude e Bard, entre outros – que permitem aos alunos acessar modelos de linguagem em grande escala. O objetivo das diretrizes é tornar-se um recurso útil para aqueles que precisam de um documento bem fundamentado,

abrangente e padronizado que permita estabelecer guias de trabalho claros sobre o uso da IA na universidade.

Palavras-chave

inteligência artificial, inteligência artificial generativa, modelos de linguagem em grande escala, ChatGPT, habilidades digitais, educação universitária.

Abstract

This instrument offers guidance to professors and students so that the use of artificial intelligence (AI) in university contexts is informed, transparent, ethical, and responsible. The Guidelines propose rules, principles and examples that are applicable to the pedagogical activities carried out inside and outside the classroom. The text is preceded by a preamble that seeks to put the reader in context through a brief explanation of the history of the formulation and development of the Guidelines. The instrument is divided into seven sections: (1) introduction, (2) objectives, (3) rules of use inside and outside the classroom, (4) justification of the rules, (5) other resources for teachers/ thus, (6) openness to future changes, and (7) changes from previous versions of the guidelines. Although the guidelines are relevant to the use of any type of AI, its rules focus on the use of chatbots—such as ChatGPT, Bing, Claude, and Bard, among others—that allow students to access language models through big scale. The objective of the guidelines is to become a useful resource for those who need a well-founded, comprehensive, and standardized document that allows clear work guides to be established regarding the use of AI at the university.

Keywords

artificial intelligence, generative artificial intelligence, large-scale language models, ChatGPT, digital skills, university education.

Preámbulo: El proceso de formulación de estos lineamientos¹

La primera vez que accedí al modelo *GPT-3*, desarrollado por la empresa *OpenAI*, fue el 15 de abril de 2022. Como era previsible, comencé a probarlo solicitando que escribiera ensayos sobre mis temas de interés, tales como el significado de los términos “política pública” e “instrumento de política pública”, y sobre la evolución del análisis de políticas públicas en los países en desarrollo. Al principio hice las consultas en inglés y luego probé en castellano, encontrando que en ambos casos el sistema

¹ (Nota del Editor). El equipo editorial informa que este documento no es un artículo académico, sino una propuesta de lineamientos preparados por el Dr. Gutiérrez, profesor. de la Uniandes y miembro de la red GIGAPP, para contribuir con su difusión a un uso más adecuado de la IA en espacios de docencia universitaria. Decidimos publicarlo aquí por su gran relevancia académica. Celebramos y agradecemos al Dr. Gutiérrez, coordinador de este número, por aceptar la invitación para su publicación en este medio y animamos a los lectores a contribuir a su mejora. Sugerimos su visionado electrónico usando Acrobat Reader o similares, debido a que contiene insertados varias URLs de recursos en texto.

producía un puñado de párrafos con respuestas articuladas a partir de cada uno de los *prompts* que introducía.²

Mi primera reacción fue de asombro por la capacidad del modelo *GPT-3* de producir respuestas que, si bien no eran brillantes, no estaban nada mal para empezar. Pero pronto advertí otra potencial consecuencia: si mis estudiantes pudieran acceder a estos sistemas, necesitaría una nueva herramienta para poder detectar el texto sintético. Como expliqué en un ensayo recientemente publicado en la [Revista de Educación Superior en América Latina](#), en mis clases procuro realizar actividades que le permitan a mis estudiantes fortalecer su capacidad de pensar y escribir por sí mismos.³ Razón por la cual [me preocupó](#) que algunos de mis estudiantes se limitaran a copiar y pegar respuestas de estos modelos como *GPT-3* para producir los memorandos y reportes de política pública de mis clases.

El desafío que había avizorado en mis primeras interacciones con *GPT-3* se materializó el 30 noviembre de 2022, fecha en la cual la empresa *OpenAI* lanzó [ChatGPT](#) y abrió de par en par las puertas del “corralito” en el que mantenía sus modelos de lenguaje a gran escala (LLMs). En enero de 2023, *ChatGPT* ya había alcanzado la cifra de [100 millones de usuarios activos al mes](#), lo cual la convirtió en la aplicación digital de más rápido crecimiento en la historia.

El 19 de enero de 2023 compartí con mis colegas de la Universidad del Rosario el primer borrador del instrumento que denominé “Política de uso de sistemas de inteligencia artificial en la asignatura” y posteriormente hice pública la [versión 3.0 de la política](#). A finales de febrero, a partir de la retroalimentación que había recibido, publiqué una [nueva versión en Castellano](#) (versión 4.2) y la [versión traducida al Inglés](#). Estas últimas fueron las versiones que piloteé en mis clases y en los talleres realizados con estudiantes y profesores durante el primer semestre de 2023. A la fecha, las diferentes versiones de mis políticas han sido descargadas más de 15,000 veces.

Después de haber probado el instrumento en mis clases y talleres, produje la versión aquí publicada que aprovecha enseñanzas de su puesta en práctica y las valiosas recomendaciones de colegas, estudiantes, e incluso personas que me escribieron al correo o a través de mis redes sociales. Esta versión 5.0 del instrumento se denomina “Lineamientos para el uso de inteligencia artificial en contextos universitarios”. Los Lineamientos proponen reglas, principios y ejemplos que son aplicables para las actividades pedagógicas realizadas dentro y fuera del aula.

El instrumento está dividido en siete secciones: (1) introducción, (2) objetivos, (3) reglas de uso dentro y fuera del aula, (4) justificación de las reglas, (5) otros recursos recomendados para profesores/as, (6) apertura a futuros cambios, (7) cambios respecto de anteriores versiones de los lineamientos, y (8) Licencia de uso y agradecimientos. Aunque los lineamientos son pertinentes para el uso de cualquier tipo de IA, sus reglas se enfocan en el uso de *chatbots* –tales como *ChatGPT*, *Bing*, *Claude*, y *Bard*, entre otros– que permiten a los estudiantes acceder a los LLMs.

Actualmente no existen herramientas confiables para distinguir textos producidos por humanos respecto de textos sintéticos. Esto dificulta detectar casos en los cuales los estudiantes transcriban las

² Los modelos de lenguaje utilizan como insumo el lenguaje natural y lo procesa para generar lenguaje natural. El “prompt” es el insumo introducido por el usuario en forma de texto –en clave de consulta o instrucciones– para que el *chatbot* responda a través de texto.

³ El ensayo “[Aprender a pensar y escribir por uno mismo en la era de los modelos de lenguaje a gran escala](#)”, puede consultarse en el Número 14 de la Revista de Educación Superior en América Latina de 2023.

respuestas producidas por un LLM y las hagan pasar como propias. Este es solo un ejemplo de los enormes retos pedagógicos que plantea el acceso masivo a sistemas de IA. En este contexto, el principal objetivo de los “Lineamientos para el uso de inteligencia artificial en contextos universitarios” es constituir un recurso útil para todos quienes precisen de un documento fundamentado, comprensivo, y estandarizado que permita establecer unas guías de trabajo claras respecto del uso de IA en la universidad.

Lineamientos para el uso de inteligencia artificial en contextos universitarios

Documento elaborado por [Juan David Gutiérrez \(juagutie@uniandes.edu.co\)](mailto:juagutie@uniandes.edu.co).

Profesor asociado de la [Escuela de Gobierno](#) de la Universidad de los Andes.

Versión 5.0 (15/08/2023). [Licencia C.C. BY 4.0](#).

1. Introducción

<p>1.1 Usamos herramientas de inteligencia artificial diariamente</p>	<p>De manera consciente o inconsciente, prácticamente todos usamos algún tipo de inteligencia artificial (IA) en nuestra vida diaria, en particular sistemas basados en modelos de lenguaje.</p> <p>Por ejemplo, usamos modelos de lenguaje cuando estamos escribiendo un mensaje de texto o un correo electrónico y la herramienta de “autocompletar” nos sugiere cómo terminar la oración. También usamos modelos de lenguaje para apoyarnos en la realización de actividades académicas, tales como sistemas de traducción automatizada o programas que detectan errores gramaticales, entre otros.</p> <p>Además, actualmente, todo tipo de plataformas web y aplicaciones móviles permiten acceder a herramientas de IA para generar textos y contenidos audiovisuales.</p>
<p>1.2 Algunas IA pueden apoyar procesos de aprendizaje, investigación, enseñanza, y otras actividades profesionales.</p>	<p>Existen usos adecuados de dichas tecnologías, que no solo pueden resultar útiles para apoyar procesos de aprendizaje, enseñanza, e investigación en la universidad sino también en el plano profesional.</p> <p>Por ejemplo, diferentes herramientas de IA hoy en día contribuyen en las actividades de todas las etapas de los procesos de política pública y, en países como Colombia, más de 50 entidades del sector público han adoptado sistemas de decisión automatizada para apoyar la realización de sus funciones.</p>
<p>1.3 Pero ciertos usos de la IA son riesgosos, por tanto, su uso debe ser informado, transparente, ético y responsable.</p>	<p>Las herramientas de IA no son idóneas para realizar cualquier tipo de actividad y determinados tipos de uso pueden ser contraproducentes para el proceso pedagógico. Algunos usos de la IA pueden generar riesgos tanto para los usuarios como para terceros. Además, como se explica en el punto 4 de estos lineamientos, es indispensable ser conscientes de las implicaciones éticas, ambientales y de derechos humanos asociadas a la utilización de estas herramientas.</p> <p>Precisamente, debido a estos riesgos algunas universidades y profesores han publicado lineamientos, principios, políticas, directrices y guías sobre el uso de IA en el plano académico (ver listado en punto 5).</p>

<p>1.4 No todos nuestros/as estudiantes son “nativos digitales” ni quieren adoptar nuevas tecnologías.</p>	<p>La actitud de las personas respecto de las tecnologías es diversa y compleja, no se debe asumir que todos nuestros estudiantes tienen avidez por adoptar nuevas tecnologías.</p> <p>Por ejemplo, en los sondeos realizados en talleres con estudiantes sobre el uso de modelos de lenguaje de gran escala (LLMs) en la universidad, el autor de este documento ha encontrado que estos sistemas generan en ellos/as sentimientos muy diversos: miedo, intriga, emoción, incertidumbre, curiosidad, desconfianza, euforia, preocupación, asombro, expectativa, angustia, tranquilidad, confusión, fascinación, precaución, cuestionamiento, rechazo, desasosiego y confort.</p> <p>Además, la capacidad de las personas para adoptar nuevas tecnologías es heterogénea y depende de diferentes factores económicos, sociales y culturales, entre otros. Todo lo anterior implica que los/as profesores/as deberíamos implementar estrategias de enseñanza diferenciadas según las necesidades de nuestros estudiantes y que no debemos forzar el uso de nuevos sistemas de IA.</p>
<p>1.5 Una orientación es necesaria para el uso de IA en contextos universitarios</p>	<p>Estos lineamientos ofrecen una orientación básica a profesores/as y estudiantes universitarios sobre cómo usar IA, particularmente modelos de lenguaje de gran escala (LLMs), de manera informada, transparente, ética y responsable. Además, en el punto 5 de esos lineamientos podrán encontrar otros recursos para profesores/as y un glosario.</p>

2. Objetivos de estos lineamientos

<p>2.1 Contribuir con la alfabetización digital</p>	<p>Contribuir con el aprendizaje de conocimientos básicos sobre el uso de IA en contextos universitarios –con énfasis en los modelos de lenguaje de gran escala (LLMs). Este objetivo incluye la enseñanza sobre sus beneficios y riesgos, sus implicaciones para la sociedad, y el desarrollo de competencias digitales críticas.</p>
<p>2.2 Promover el uso responsable de IA</p>	<p>Promover el uso responsable y ético de estas tecnologías en los procesos de aprendizaje y para su futuro uso en la vida profesional.</p>
<p>2.3 Prevenir la vulneración de la integridad académica</p>	<p>Prevenir situaciones en las cuales los estudiantes consciente o inconscientemente incurran en deshonestidad académica.</p>

3. Reglas de uso de IA dentro y fuera del aula universitaria

<p>3.1 Regla general</p>	<p>Está permitido el uso de IA como herramienta de soporte para realizar las diferentes actividades de aprendizaje. Los parámetros para el uso de estas herramientas que se describen a continuación distinguen entre herramientas de “bajo riesgo” y de “alto riesgo”, según los riesgos que la respectiva IA genere para el proceso pedagógico, para los usuarios y para terceros.</p>
<p>3.2 Reglas para uso de la IA de bajo riesgo</p>	<p>Podrán usar libremente herramientas de IA de bajo riesgo, es decir aquellas que les permitan corregir o revisar contenidos generados por los propios estudiantes o aquellas que les permitan recolectar y procesar datos. Por ejemplo, se consideran de bajo riesgo las herramientas de corrección gramatical, de traducción, de transcripción audio-texto, y de búsqueda de información en Internet, entre otras.</p> <p>Se recomienda que cuando dichas herramientas permitan realizar labores significativas (ej. traducción), se indique en la sección que corresponda (ej. metodología) sobre el uso de dichas herramientas. Si tienen dudas sobre cómo clasificar el riesgo de uso de una IA específica, se sugiere consultar al profesor/a.</p>
<p>3.3 Reglas para uso de las IA de alto riesgo (LLM)</p>	<p>Son de alto riesgo, las herramientas de IA que se usen para generar texto sintético. Es decir, los LLM que pueden accederse a través de chatbots, tales como ChatGPT, Bard, Bing, y Claude, etc. Los/as estudiantes solo podrán incluir en sus entregas individuales y colectivas contenidos generados por IA cuando cumplan con 4 requisitos:</p> <ol style="list-style-type: none"> <li data-bbox="480 1227 1425 1442"> <p>i. Uso informado Que previo a su uso hayan investigado quién o qué empresa desarrolló la herramienta, cómo fue desarrollada, cómo funciona, qué funciones puede desempeñar, cuáles son sus términos de uso, qué hace la empresa con la información que recolecta del usuario, y qué limitaciones y/o riesgos presenta para el usuario y/o para terceros.</p> <li data-bbox="480 1464 1425 1680"> <p>ii. Uso transparente Que indiquen detallada y expresamente qué herramienta usaron y cómo la usaron en la sección de metodología o en la que consideren pertinente. Este es un requisito que actualmente solicitan revistas científicas y casas editoriales como <i>Nature</i> y <i>Cambridge University Press</i>, a quienes entregan manuscritos para revisión de pares.</p> <li data-bbox="480 1702 1425 1980"> <p>iii. Uso ético En relación con el uso del LLMs, que sea posible distinguir qué fue escrito o producido directamente por ustedes y qué fue generado por una herramienta de IA. Respecto de este punto, aplican las reglas de citación generales, por ejemplo, usar comillas si incluyen apartes literales además de la referencia correspondiente. Para citar texto sintético producido por LLMs, pueden consultar las reglas sugeridas por el equipo de estilo de APA.</p>

	<p>Las infracciones a esta regla, particularmente respecto del segundo requisito, serán abordadas bajo un enfoque similar al que aplica en casos de plagio. Además, <i>ChatGPT</i> no es una persona, luego no puede ser considerado su coautor. Algunas revistas científicas ya han actualizado sus políticas de publicación para aclarar que los modelos de lenguaje como <i>ChatGPT</i> no satisfacen sus criterios de atribución de "autoría". Por otra parte, el uso ético de sistemas de IA –en general, no solo LLMs– debería estar informado por cuatro principios desarrollados por la bioética: <i>beneficencia</i> (consecución de beneficios), <i>respeto a la autonomía de las personas</i> (incluye la obtención de consentimiento si hay terceros involucrados), <i>justicia</i> (especialmente la no discriminación) y <i>no maleficencia</i> (no producir daño a otros de manera intencionada o por imprudencia).</p> <p>iv. Uso responsable Se recomienda que el uso de estas IA se circunscriba a etapas tempranas de la investigación, para inspirar o sugerir rumbos, no para producir contenido que luego sea incluido en sus entregas. En todo caso, si optan por usar información obtenida a través de un LLM, deben acreditar que contrastaron rigurosamente dicha información con otras fuentes confiables dado que estos sistemas tienden a ofrecer información imprecisa, errónea, y falsa. El uso responsable también implica evitar que la utilización del sistema cause daños al usuario o a terceros. Por ejemplo, no se recomienda incluir en las consultas realizadas a los <i>chatbots</i> datos personales o información confidencial propia o ajena porque una vez introducida dicha información se pierde control de quien puede acceder a la misma y se genera el riesgo de que sea conocida por terceros no autorizados.</p>
<p>3.4 Usos plausibles de los LLMs</p>	<p>i. Explorar Explorar nuevos asuntos y buscar inspiración, es decir, usar el <i>chatbot</i> en las etapas más tempranas de la investigación.</p> <p>ii. Automatizar tareas básicas Apoyar ejercicios que impliquen el reconocimiento y/o reproducción de patrones en textos, por ejemplo: traducir, resumir, modificar el tono y estilo del texto –por ejemplo, para que la redacción texto sea más clara o asequible a determinadas audiencias–, corregir ortografía y gramática, conversión audio–texto, programación, etc. Tampoco pueden fiarse y se debe revisar con cuidado el <i>output</i> del LLM.</p> <p>iii. Escritura trivial Apoyar tareas de escritura poco importantes. Para decidir si algo es trivial o no, piensen en la siguiente heurística: ¿qué ocurriría si mi interlocutor se entera que utilicé un LLM para escribir el mensaje?</p> <p>iv. Sparring Probar argumentos o ideas propias a través de la interacción con el <i>chatbot</i>, por ejemplo, solicitando al sistema que responda con críticas o contraargumentos a los planteamientos del estudiante.</p>

<p>3.5 Usos no recomendados de los LLMs</p>	<ul style="list-style-type: none"> i. Usar el <i>chatbot</i> como principal o única fuente para buscar información factual o técnica No recomendado pues <u>los LLMs no son herramientas de búsqueda fiables</u>. Además, los buscadores de Internet son más eficaces y confiables para esa tarea. En todo caso, dado que hacer búsquedas en los <i>chatbots</i> puede ser percibido como una manera de obtener información más fácil y rápida, se recomienda que la información obtenida por ese medio siempre sea cuidadosamente contrastada con fuentes confiables. ii. Uso de texto sintético para asuntos sustanciales No recomendado por la tendencia de los LLMs a producir respuestas con información <u>imprecisa, errónea y falsa</u>. En todo caso, si lo hiciera, debe contrastarse cada punto con otras fuentes confiables. iii. Realizar cálculos matemáticos No recomendado pues los LLMs <u>no están entrenados para realizar cálculos matemáticos</u> y, por ende, con frecuencia fallan en responder operaciones aritméticas sencillas.
<p>3.6 Acompañamiento del profesor y desarrollo de competencias digitales</p>	<p>El profesor dispondrá tiempo de la clase para explicar qué es la IA, qué son los modelos de lenguaje a partir de los cuales funcionan herramientas como ChatGPT, qué oportunidades y riesgos generan para los trabajos académicos y profesionales, y cuáles son las implicaciones éticas, ambientales y de derechos humanos asociadas al uso de estas herramientas.</p> <p>El profesor siempre estará disponible para aclarar el alcance de estos lineamientos, para discutir y cocrear estos lineamientos, y para resolver preguntas puntuales sobre el uso de IA.</p> <p>Además, el profesor procurará realizar actividades que contribuyan con la alfabetización digital de los estudiantes, en particular para el desarrollo y fortalecimiento de las competencias digitales tales como la habilidades para evaluar críticamente los beneficios y riesgos de las nuevas tecnologías.</p>
<p>3.7 Supervisión</p>	<p>El profesor usará diferentes fuentes para identificar potenciales situaciones de deshonestidad académica a través del uso de IA, pero no depender de las herramientas disponibles para detectar texto sintético, dado que actualmente ninguna de esas herramientas es fiable.</p> <p>En todo caso para evitar riesgos asociados a la integridad académica, la principal estrategia que usará el profesor será pedagógica no punitiva.</p>

4. ¿Por qué es necesario el uso informado, transparente, ético y responsable de IA en la universidad?

El uso de herramientas de IA debe ser informado, transparente, ético y responsable por –al menos– cuatro motivos, que se explican aquí en detalle:

<p>4.1 Los LLMs no siempre son confiables</p>	<p>Porque sus <u>respuestas NO siempre son confiables</u> a pesar de que, por ejemplo, una IA generadora produce textos que tienen <u>apariencia de ser convincentes</u>. Los sistemas basados en LLMs como <i>ChatGPT</i> <u>no funcionan con la precisión</u> de otras herramientas usadas en ambientes de aprendizaje como las calculadoras.</p> <p>De hecho, <i>ChatGPT</i> tiende a incluir en sus respuestas <u>información falsa o fantasiosa</u>. El <i>chatbot</i> incorporado por Microsoft a su <u>explorador de Internet, Bing</u>, y el <u>chatbot de Google, Bard</u>, también han presentado los <u>mismos tipos de problemas</u>. Estos sistemas <u>no distinguen lo verdadero de lo falso</u>. ¿Por qué ocurre esto?</p> <p>Los LLMs hilan palabras a partir de inferencias probabilísticas de los datos con los cuales fueron entrenados, pero no tienen la capacidad de entender lo que producen ni asocian significados a las palabras que emiten (son “<u>loros estocásticos</u>”).</p>
<p>4.2 Hay riesgos de que ciertos usos afecten negativamente los procesos de aprendizaje</p>	<p>Porque el uso de herramientas generadoras de texto puede desestimular la motivación de los estudiantes para <u>escribir y pensar por su propia cuenta</u>.</p> <p>Vale la pena reiterar que muchas de las actividades de aprendizaje buscan desarrollar habilidades cognitivas y que estos lineamientos buscan evitar que algunas <u>herramientas de IA se conviertan en mecanismos de plagio automatizadas</u>.</p>
<p>4.3 El riesgo de que los usuarios traten la IA como si fuera un humano</p>	<p>Por el riesgo de que los usuarios, consciente o inconscientemente, traten la conducta de la IA como si fuera humana (<u>Efecto Eliza</u>). Por ejemplo, las herramientas basadas en <u>LLMs no “entienden” los textos que producen</u>, simplemente imitan patrones de lenguaje a partir de la síntesis de grandes volúmenes de datos con base en los cuales generan las secuencias de palabras.</p> <p>Este problema de <u>antropomorfizar las máquinas</u> se exagera por cuanto algunos <i>chatbots</i> han producido <u>respuestas violentas o peligrosas que podrían llevar a las personas a realizar conductas dañinas</u>.</p>
<p>4.4 Implicaciones</p>	<p>Por las implicaciones éticas y de derechos humanos asociadas al uso de ciertos</p>

éticas y de
derechos humanos

sistemas de IA dado que:

- i. Algunas herramientas tienden a reproducir o amplificar estereotipos derogatorios y discriminatorios asociados al género, raza, etnia o discapacidad;
- ii. Las tecnologías podrían haber sido desarrolladas a partir de la violación masiva de derechos de autor;
- iii. Nuevas formas de colonialismo a través de la extracción no consentida de información de comunidades históricamente marginalizadas;
- iv. Algunas herramientas habrían sido desarrolladas en contextos de explotación laboral;
- v. El desarrollo y operación de dichos sistemas genera una huella de carbono considerable; y,
- vi. La potencial violación a los derechos de privacidad y de protección de datos personales de quienes las usan.

Para profundizar en los daños asociados al uso de IA Generativa, se recomienda consultar el reporte "Ghost in the machine: Addressing the consumer harms of generative AI" (2023) del *Norwegian Consumer Council* y el reporte "Generating Harms: Generative AI's Impact & Paths Forward" del *Electronic Privacy Information Center*.

5. Otros recursos recomendados para profesores/as

5.1 Glosario de definiciones clave

A continuación, encontrarán definiciones de trabajo de tres términos tecnológicos utilizados en estos lineamientos:

- a) [Sistema de Inteligencia Artificial \(IA\)](#): "sistema basado en máquinas que puede, para un conjunto determinado de objetivos definidos por el ser humano, hacer predicciones, recomendaciones o tomar decisiones que influyen en entornos reales o virtuales. Los sistemas de IA están diseñados para funcionar con diversos niveles de autonomía".
- b) [Modelo de lenguaje](#): algoritmos que procesan, analizan y producen el lenguaje natural en forma escrita o verbal.
- c) [Modelo de lenguaje a gran escala \(LLM\)](#): algoritmos entrenados a partir de una vasta cantidad de información cuya principal función es predecir el hilo de palabras más probable dado el contexto que lo antecede o le sigue.

5.2 Otros lineamientos sobre uso de IA en la universidad.

Recomiendo consultar lineamientos, políticas, y principios de otras universidades, profesores/as y otras organizaciones, tales como:

- [University of Washington](#)
- [Hertie School](#),
- [University of Tartu](#)
- [Yale University](#)
- [University of Helsinki](#)
- [Technical University of Munich](#)
- [University College London](#)
- [UNESCO](#)
- [Departamento de Educación del Reino Unido](#)
- [Boston University](#)
- [Warwick University](#)
- [Universidad de Buenos Aires \(IALAB\)](#)
- [Universidad del Desarrollo](#)
- [Russell Group](#)
- [listado de Lance Eaton](#)

5.3 Literatura introductoria sobre LLMs.

- a) Para una introducción sobre cómo fueron desarrollados y cómo funcionan los LLMs (en particular, GPT-3), les recomiendo este [texto de Timothy B. Lee y Sean Trott](#).
- b) Para una explicación en Castellano sobre cómo funcionan los LLMs y cuáles son sus principales riesgos, recomiendo el texto "[El riesgo de los grandes modelos lingüísticos: convertirse en los estocásticos](#)", por Esther Sánchez García y Michael Gasser, que resume el artículo "[On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? □](#)" de Emily M. Bender, Timnit Gebru, Angelina McMillan-Major y Shmargaret Shmitchell.

5.4 Recursos para la enseñanza.

- a) Quienes estén interesados en los retos que los LLMs generan para enseñar a escribir en las universidades les sugiero consultar [el](#) documento titulado "[Adaptando la escritura universitaria a la era de los Modelos de Lenguaje a Gran Escala como ChatGPT: Próximos pasos para educadores](#)" de Anna Mills y Lauren M. E. Goodlad.

- b) Con el mismo fin, también recomiendo los materiales contenidos en "[AI Text Generators and Teaching Writing: Starting Points for Inquiry](#)" de Anna Mills. El artículo de Anna Mills, Maha Bali y Lance Eaton titulado "[How do we respond to generative AI in education? Open educational practices give us a framework for an ongoing process](#)" documenta y reflexiona cómo profesores de diferentes universidades del mundo han adoptado nuevas prácticas pedagógicas para responder a los retos generados por la IA Generativa.
- c) Para una reflexión sobre cómo integrar IA en procesos de formación en ciencias humanas, recomiendo este [blog post](#) de [Bastien Bosa](#).
- d) Para una reflexión sobre qué estrategias pedagógicas pueden adoptarse para [evaluar estudiantes "en la era de ChatGPT"](#), recomiendo el artículo de M. Paz Prendes Espinosa, María del Mar Sánchez Vera y Víctor González Calatayud.
- e) El MIT ofrece de manera gratuita su programa [DAILY Curriculum](#) que contiene materiales para enseñar conocimientos básicos sobre IA a adolescentes.

5.5 Aproximación crítica al uso de IA en entornos académicos.

- a) Recomiendo consulta el [listado de recursos de la profesora Maha Bali](#) sobre desarrollo de alfabetización crítica en IA.
- b) Para una aproximación crítica sobre el uso de LLMs para al proceso de investigación y revisión bibliográfica, recomiendo el texto de "[Nadie escribiendo y nadie leyendo: los generadores de texto con inteligencia artificial y la ciencia que queremos](#)" de Ariel Guersenzvaig y Javier Sánchez Monedero.
- c) También recomiendo leer una [reflexión](#) de los profesores Mohammad Hosseini, Lex Bouter, y Kristi Holmes en la cual argumentan que no es deseable adoptar herramientas de IA en la educación sin reflexionar antes sobre sus efectos y sesgos y sin haber adoptado medidas que mitiguen sus riesgos.
- d) El artículo "[How ChatGPT robs students of motivation to write and think for themselves](#)" de Naomi S. Baron discute cómo el uso de LLMs puede afectar negativamente los procesos creativos de estudiantes.
- e) El [blog post](#) de Dimitrinka Atanasova reflexiona sobre cómo los LLMs pueden reducir las barreras para que estudiantes e investigadores escriban en Inglés (cuando este no es su primer idioma), pero que sus efectos respecto de la inclusividad cultural en la educación superior son más ambiguos.
- f) También es valiosa la propuesta de Kathryn Conrad que aboga por una "[Carta de Derechos de la IA para la Educación](#)" que propone derechos para profesores y estudiantes que sirvan de base para la formulación de políticas y medidas protectoras.

6. Lineamientos abiertos al cambio

La IA es un conjunto de herramientas que se transforma rápidamente, razón por la cual estos lineamientos permanecerán abiertos a futuras evaluaciones, modificaciones y revisiones. A lo largo del semestre académico el profesor abrirá espacios para discutir estos lineamientos con los/las estudiantes y, de ser necesario, podrán introducirse modificaciones a partir de ejercicios de cocreación.

7. Cambios introducidos a la presente versión de los lineamientos

Se introdujeron cambios de forma y de fondo respecto de la versión anterior de este instrumento ([v. 4.2](#)) que se resumen a continuación:

7.1 Cambios de forma

- La anterior versión se denominaba "Política de uso de inteligencia artificial en la asignatura", el nuevo título es más amplio puesto que el instrumento es útil para ámbitos académicos que van más allá de las actividades en clase (ej. actividades de investigación).
- Introduce nuevas fuentes que pueden consultarse en forma de hipervínculos a lo largo del texto.

7.2 Cambios de fondo

- En esta nueva versión de los lineamientos introduje una nueva subsección con usos plausibles y usos no recomendados de los LLMs por parte de estudiantes.
- Además, en esta versión incluí una nueva sección de recursos para profesores/as que incluye: ejemplos de otros lineamientos, literatura básica sobre LLMs, recursos para la enseñanza, y literatura crítica sobre uso de IA en contextos educacionales.
- Por otra parte, la anterior versión de este instrumento indicaba que el profesor usaría sistemas de detección de texto sintético, pero he eliminado esas alusiones en la presente versión porque los sistemas actualmente disponibles producen un [alto nivel de falsos positivos y negativos](#). Por este motivo, por ejemplo, el *Teaching Center* de la Universidad de Pittsburgh publicó un [comunicado de prensa](#) en el cual informaba que no avalaba ni promovía ninguna de las herramientas actualmente disponibles.


8. Infográfico

Se elabora para esta versión un infográfico para promocionar su conocimiento, uso y difusión. Este infográfico ha sido realizado por el equipo editorial GIGAPP y ha sido revisado y curado por el autor para esta publicación. Al final de este documento se reproduce el infográfico aunque pueden acceder a él desde [este enlace](#).

9. Licencia de uso y agradecimientos

Estos lineamientos están publicados bajo una [licencia Creative Commons Atribución 4.0](#), luego si usted es un/a educador/a interesado/a en formular sus propias políticas o directrices, puede compartir y adaptar estos lineamientos siempre que realice la correspondiente atribución e indique si realizó cambios.

Agradezco a mis colegas y demás personas que aportaron sus comentarios y críticas a versiones iniciales de este documento y también a quienes han participado en las charlas y talleres que he realizado sobre el uso de modelos de lenguaje de gran escala en contextos universitarios

Agradezco a quienes quieran enviarme sus comentarios, críticas y sugerencias que por favor me escriban a juagutie@uniandes.edu.co. 

Sobre el autor/ About the author

Juan David Gutiérrez es Profesor Asociado de la Escuela de Gobierno Alberto Lleras Camargo de la Universidad de los Andes. PhD en Política Pública de la Escuela de Gobierno de la Universidad de Oxford. Investiga sobre política pública, inteligencia artificial, y gobernanza de los recursos naturales.

Inteligencia artificial *en el aula universitaria*

Basado en documento de Gutiérrez, J.D. (2023)
"Lineamientos para el uso de inteligencia artificial en contextos universitarios (v 5.0)"



¿Sabías que?

● Chat GPT no es la única IA

Usamos herramientas de inteligencia artificial (IA) diariamente. Chat GPT cuenta con 100 millones usuarios activos a enero 2023.

● IA generativas

Algunas IA generativas LLM pueden apoyar procesos de aprendizaje, investigación, enseñanza, y otras actividades profesionales.

● Hay riesgos en su uso

Ciertos usos de la IA generativa son riesgosos, por tanto, su uso debe ser informado, transparente, ético y responsable.

● ¿Nativos/as digitales?

No todos nuestros/as estudiantes son "nativos digitales" ni quieren adoptar nuevas tecnologías.

● Lineamientos

Es necesaria una orientación específica para el uso de IA generativa en contextos universitarios



Uso de la IA en el aula universitaria



Lineamientos

Uso de IA en el aula universitaria

1. Uso informado

Que previo a su uso, los/as usuarios investiguen:

- ✓ Qué empresa desarrolló la herramienta
- ✓ Cómo fue desarrollada
- ✓ Cómo funciona y qué funciones puede desempeñar
- ✓ Términos de uso
- ✓ Qué hace la empresa con la información que recolecta
- ✓ Qué limitaciones y/o riesgos presenta

Uso de IA en el aula universitaria

2. Uso transparente

Se debe indicar en la sección metodológica o similar, de forma detallada y expresa, qué herramienta(s) IA se usaron y cómo fueron usadas.

- ✓ La transparencia genera confianza en el o la lectora.
- ✓ Es pertinente para cualquier lector/a saber cómo se usó el sistema de IA en la preparación del texto.



Lineamientos



Uso de IA en el aula universitaria

Uso ético

Que sea posible distinguir qué fue escrito o producido directamente por los estudiantes y profesores y qué fue generado por una herramienta de IA. La correcta citación de IA es necesaria como fuente.

- ✓ Chat GPT no puede considerarse como un autor
- ✓ Cuando se utiliza texto sintético producido por IA éste tiene que ser citado. Ya existen formatos de citación disponible para estos casos.

Uso de IA en el aula universitaria

Uso responsable

Se recomienda que su uso se circunscriba a etapas tempranas de la investigación, para inspirar o sugerir rumbos, y no para producir contenido que luego sea incluido en sus entregas.

Se debe acreditar que se contrastó rigurosamente dicha información con otras fuentes confiables.

También debemos evitar que con su uso se vulneren los derechos de otros (p.ej. cuando se introduce información personal de terceros en las consultas al sistema).



Saber qué no hacer



Uso de IA en el aula universitaria

Cómo no usarla

- ⊗ Como fuente de información principal.
- ⊗ Como herramienta de búsqueda fiable.
- ⊗ Sin contrastar la información obtenida.
- ⊗ Para responder a cuestiones sustanciales.
- ⊗ Para realizar cálculos matemáticos.



Grupo de Investigación en
Gobierno, Administración
y Políticas Públicas



Para saber más ...

Te invitamos a conocer, compartir y usar este documento de Lineamientos. Está bajo licencia [Creative Commons 4.0](#)

La versión 5.0 ha sido publicada como documento especial en la Revista Estudios/Working Papers de GIGAPP (2023) en donde te ofrecemos muchos enlaces y documentos web para saber más.

Cómo citar este recurso:
Gutiérrez, J. D. (2023). Lineamientos para el uso de inteligencia artificial en contextos universitarios. *GIGAPP Estudios Working Papers*, 10(267-272), 416-428.
Recuperado a partir de:
<https://www.gigapp.org/ewp/index.php/GIGAPP-EWP/article/view/331>



Dr. Juan David Gutiérrez

Profesor Asociado.
Escuela de Gobierno de la Universidad de los Andes.

 juagutie@uniandes.edu.co

Contacto

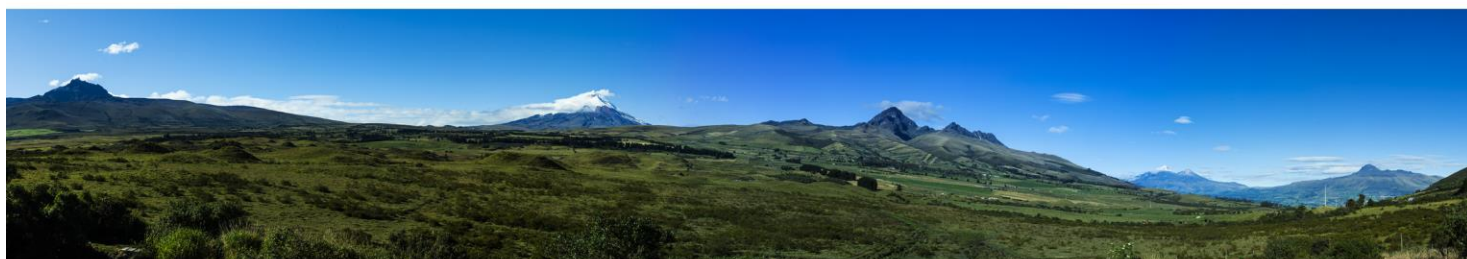
La Serie **GIGAPP Estudios Working Papers** es un espacio de divulgación científica sobre avances de investigación y estudio en materia de gobierno, administración y políticas públicas.

Su propósito principal es contribuir al intercambio y debate de ideas, servir de plataforma para el fortalecimiento de las labores de investigación aplicada en estas materias y poner al alcance de la comunidad académica, investigadores, estudiantes y público en general interesado, reflexiones y contenidos del más alto nivel con el objeto de promover nuevas miradas respecto del perfeccionamiento y mejora de nuestros sistemas políticos, gobiernos y administraciones públicas.

GIGAPP Estudios Working Papers (nueva serie) es una publicación de la Asociación Grupo de Investigación en Gobierno, Administración y Políticas Públicas GIGAPP. (ISSN 2174-9515).

Se aceptan para evaluación trabajos inéditos en castellano, portugués e inglés.
Envío de originales ewp@gigapp.org

Consulte las normas para la presentación de originales en la web
<http://www.gigapp.org/index.php/publicaciones/working-papers>



El Verano en los Andes: El Sincholagua, el Cotopaxi, El Rumiñahui, los Illinizas, el Corazón. Foto hecha en las Laderas del Pasochoa. Imagen cortesía del Efrén Guerrero @auraneurotica

GIGAPP
Estudios / Working Papers

ISSN 2174-9515