

## 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 Riscos, 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 Andes<sup>1</sup>. 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 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

---

<sup>1</sup> 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.

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 sistemas como capaces de causar impactos y estabilizan conceptos abstractos, como los principios de la IA (Metcalfe *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/](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](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