



DNP

CONPES

PARA LA GENTE 4144

INTELIGENCIA ARTIFICIAL Y BIENESTAR SOCIAL

UN COMPROMISO CON EL FUTURO



Departamento Nacional de Planeación



CONPES para la gente 4144

Inteligencia artificial y bienestar social: un compromiso con el futuro¹

Sugerencia de citación: Departamento Nacional de Planeación. (2025). *CONPES para la gente 4144. Inteligencia artificial y bienestar social: un compromiso con el futuro.*

¿Las máquinas pueden pensar? Esta inquietud tan humana, que ha sido considerada desde campos como la ingeniería y la filosofía, entraña algunos de los dilemas que plantea la inteligencia artificial. Así mismo, ciertos temores respecto a los avances tecnológicos, como considerar la posibilidad de que las máquinas tomen conciencia de sí mismas, son temas recurrentes que han sido planteados desde hace décadas por la ciencia ficción.

Sin embargo, el futuro en el que las máquinas podrían llegar a tomar decisiones todavía es una idea distópica². En el punto en el que nos encontramos ahora respecto a los avances de la IA se habla del aprendizaje automático (*machine learning*), que se refiere a cómo los sistemas de IA utilizan estas técnicas para realizar funciones complejas, como en el caso de los automóviles autónomos, que requieren múltiples modelos de aprendizaje para operar de manera efectiva sin intervención humana. Además, estos procesos tienen implicaciones éticas y desafíos asociados con el uso de estos sistemas, especialmente en términos de la comprensión y verificación de los procesos automatizados de toma de decisiones.

Ahora bien, más allá de la ficción, nuestra realidad actual muestra que la inteligencia artificial tiene un potencial enorme que, bien aprovechado, resulta muy beneficioso para nuestras sociedades en diversos sentidos, y es por ello que la publicación del Documento CONPES *Política Nacional de Inteligencia Artificial* resulta tan relevante. La IA puede definirse, a grandes rasgos, como una disciplina de la informática que busca crear sistemas capaces de realizar tareas que normalmente requieren inteligencia humana, como el aprendizaje, el razonamiento y la toma de decisiones, mediante el uso de algoritmos y grandes volúmenes de datos. Abarca diversos enfoques tecnológicos, incluyendo el procesamiento de lenguaje natural, el reconocimiento de imágenes y la automatización de procesos, lo cual puede

¹ En esta edición del CONPES para la Gente se sintetiza el contenido del Documento CONPES 4144 *Política Nacional de Inteligencia Artificial*, aprobado el 14 de febrero de 2025. Disponible en <https://colaboracion.dnp.gov.co/CDT/Conpes/Economicos/4144.pdf>.

² Según la Real Academia Española (2025), una distopía es una “representación ficticia de una sociedad futura de características negativas causantes de la alienación humana”.

transformar múltiples sectores de la economía y la sociedad³.

El objetivo general de esta política es desarrollar las habilidades necesarias para la investigación, el desarrollo, la implementación y el uso ético y sostenible de sistemas de inteligencia artificial, con lo cual se busca promover la transformación social y económica de Colombia para el año 2030. Para ello, se plantean seis objetivos específicos y 106 acciones que deberán ser implementadas por varias entidades, como se detalla en la sección “plan de acción” del documento.

Por ejemplo, una de las líneas de acción busca “monitorear y mitigar el impacto de desplazamiento de mano de obra e inestabilidad laboral por la adopción de sistemas IA” (Departamento Nacional de Planeación [DNP], 2025, p. 98). Para ello, entre 2025 y 2030, el Ministerio de Trabajo, con el apoyo de otras entidades, elaborará y pondrá en marcha una estrategia para salvaguardar los derechos laborales de los empleados frente a la automatización de tareas derivada de la inteligencia artificial, lo cual abarcará la reconversión profesional y la promoción de los derechos laborales esenciales.

Otra de las líneas de acción se orienta a “promover equidad, inclusión y sostenibilidad ambiental por el desarrollo y adopción de la IA” (DNP, 2025, p. 99), para lo cual, entre otras acciones, la Comisión de Regulación de Comunicaciones, en coordinación con el DNP y el Ministerio de Igualdad y Equidad, entre 2025 y 2027, desarrollará e implementará una estrategia de información, concienciación y prevención sobre las violencias de género

que se manifiestan en entornos digitales y que involucran herramientas de inteligencia artificial (como los *deepfakes*⁴, entre otros). Este plan estará dirigido a las autoridades encargadas de proteger los derechos ciudadanos y a la sociedad civil, con el objetivo de promover un uso más seguro y ético de las tecnologías digitales, evitando así la perpetuación de la desigualdad de género y la vulneración de la dignidad de las personas.

La Política Nacional de Inteligencia Artificial en Colombia busca solucionar el problema de las bajas capacidades del país en términos de investigación, desarrollo, adopción y aprovechamiento ético y sostenible de sistemas de IA. Este enfoque tiene como fin abordar desafíos sociales, económicos y ambientales, tales como la mejora de la seguridad alimentaria, la reducción de la pobreza y el impulso de una economía basada en el conocimiento, garantizando al mismo tiempo que la IA se utilice de forma responsable y beneficiosa para toda la sociedad. Esta política es la hoja de ruta para que Colombia, con un horizonte a 2030, se convierta en un referente regional de IA, para lo cual se destinará una inversión aproximada de 479.273 millones de pesos. Así mismo, esta política busca coordinar esfuerzos entre las entidades públicas y los sectores productivos en torno a seis ejes estratégicos⁵.

Teniendo en cuenta que el Gobierno nacional ha hecho esfuerzos para cerrar brechas entre los territorios y ha planteado diversas rutas para buscar un crecimiento más equitativo y sostenible, en línea con las propuestas del Plan Nacional de Desarrollo 2022-2026, *Colombia, potencia mundial de la vida*, cabe destacar que,

3 Si desea ampliar esta información, en la sección “marco conceptual” de este documento CONPES encontrará una análisis sencillo y detallado al respecto.

4 El término *deepfake* se refiere a los sistemas informáticos que, utilizando técnicas de inteligencia artificial, generan videos manipulados que parecen auténticos. Su grado de realismo puede llegar a ser tan alto que resulta difícil detectar que han sido alterados, lo que puede facilitar la difusión de noticias falsas o que sean utilizados en casos de pornovenganza. Aunque este es su uso inicial, también puede aplicarse a manipulaciones en otros formatos, como el audio (Fundéu RAE, 2023).

5 Los ejes estratégicos en torno a los que se estructura esta política son: (i) ética y gobernanza, (ii) datos e infraestructura, (iii) investigación, desarrollo e innovación (I+D+i), (iv) desarrollo de capacidades y talento digital, (v) mitigación de riesgos y (vi) uso y adopción de la IA.

entre otros asuntos, la implementación de esta política beneficiará a las regiones de varias formas para promover el bienestar social y económico de sus habitantes.

Por ejemplo, el desarrollo de capacidades y la formación en inteligencia artificial generarán nuevas oportunidades de empleo, especialmente en sectores que demandan habilidades digitales y tecnológicas. Así mismo, la adopción de sistemas de IA en servicios públicos permitirá una mayor eficiencia y efectividad, lo que se traducirá en servicios más accesibles y de mejor calidad para la población. Por otro lado, la investigación y el desarrollo en IA facilitarán la creación de soluciones adaptadas a problemas específicos de cada territorio, lo cual fomenta el acceso a tecnología que responda a las necesidades locales.

La política también incluye medidas para mitigar los riesgos de exclusión y asimetrías en el acceso a la tecnología, con lo cual se promueve la igualdad de oportunidades y se favorece a las poblaciones vulnerables. Por otro lado, se fomentará la apropiación social del conocimiento en IA, lo que permitirá a la ciudadanía participar de manera informada en debates públicos y decisiones que afecten su entorno. Así mismo, la implementación de IA en sectores como el medio ambiente ayudará a desarrollar prácticas más sostenibles que beneficien a las comunidades locales a través de la gestión eficiente de recursos.

La IA puede desempeñar un papel fundamental en la reducción de la pobreza en Colombia a través de su aplicación en la agricultura y en el desarrollo socioeconómico. Al facilitar la detección temprana de enfermedades y optimizar los sistemas de riego, la IA puede mejorar la eficiencia y sostenibilidad de la producción agrícola, lo que a su vez puede incrementar la seguridad alimentaria y ofrecer mayores oportunidades para las comunidades vulnerables.

“ El objetivo general de esta política es desarrollar las habilidades necesarias para la investigación, el desarrollo, la implementación y el uso ético y sostenible de sistemas de inteligencia artificial, con lo cual se busca promover la transformación social y económica de Colombia para el año 2030 ”

Además, la IA puede transformar sectores económicos al aumentar la competitividad y diversificar la oferta exportadora del país, lo cual genera nuevas oportunidades laborales y mejora los ingresos. La optimización de procesos productivos mediante la IA no solo conduce a una mayor productividad, sino que también puede resultar en la creación de empleo y la mejora de las condiciones de trabajo. En conjunto, estas capacidades hacen de la IA una herramienta clave para combatir la pobreza y promover un desarrollo más equitativo en Colombia.

Ahora bien, retomando la reflexión inicial en torno a los riesgos que podría tener el desarrollo y la adopción de la IA, tenemos que estos son reales e incluyen la introducción de sesgos en los sistemas, lo que puede llevar a decisiones discriminatorias que afecten desproporcionadamente a ciertos grupos poblacionales. Estos sesgos pueden surgir cuando los datos utilizados para entrenar algoritmos no son suficientemente diversos o representativos. Además, la complejidad inherente a algunas técnicas de IA, como el aprendizaje profundo⁶, puede dificultar la transparencia y la explicabilidad de las

⁶ El aprendizaje profundo (*deep learning*) es una técnica de IA que permite crear modelos complejos capaces de procesar grandes volúmenes de datos y realizar tareas como el reconocimiento de patrones. Sin embargo, también se destacan desafíos relacionados con la explicabilidad y la transparencia. La explicabilidad de la IA abarca una serie de procesos y técnicas que facilitan a los usuarios humanos la comprensión y confianza en los resultados generados por los modelos de aprendizaje automático. Esta explicabilidad examina el proceso mediante el cual un sistema de IA alcanza un resultado particular y contribuye a definir la transparencia del modelo (IBM, 2024).

decisiones automatizadas, lo que es crítico cuando estas decisiones impactan directamente en la vida de las personas.

Otro efecto no deseado de la IA es la posible amenaza a la seguridad operacional, donde fallas en los sistemas pueden provocar daños a usuarios o terceros, como en el caso de vehículos autónomos o sistemas de diagnóstico médico. Además, el impacto ambiental y social de la IA debe ser monitoreado continuamente, ya que su desarrollo y operación generan una significativa huella de carbono. Así mismo, puede facilitar la creación de contenido desinformativo o malintencionado, lo cual representa un riesgo para la integridad de la información y la libertad de expresión, así como un potencial incremento en la exposición a delitos cibernéticos.

Por lo anterior, en el Documento CONPES 4144 se apunta a la necesidad de desarrollar lineamientos específicos que integren la ética y la responsabilidad en el diseño, desarrollo e implementación de tecnologías de IA. Las prácticas responsables que incluyan la evaluación continua del impacto social, económico y ambiental de los sistemas de IA, así como la educación y formación de los usuarios y desarrolladores, también son esenciales. Igualmente, se menciona la necesidad de regular el tratamiento de datos personales para proteger la privacidad y prevenir abusos, y se enfatiza que el uso responsable de la IA puede ayudar a mitigar sus potenciales efectos adversos, lo cual incluye medidas preventivas para identificar, prevenir y mitigar los efectos negativos relacionados con la seguridad digital, la integridad de la información y otros aspectos cruciales como el medio ambiente y el empleo.

Ahora bien, la probabilidad de que exista un futuro distópico, al estilo de las sagas *Terminator*

o *Matrix*, parece todavía un asunto reservado a la ciencia ficción. Tal vez podamos considerar la posibilidad, más esperanzadora, de entrenar los sistemas de IA para que apliquen *Las tres leyes de la robótica*. Estos principios, que fueron difundidos por el autor de ciencia ficción Isaac Asimov en su colección de relatos sobre robots, tratan temas éticos relacionados con el diseño y empleo de la robótica y señala las normas y pautas que regulan cómo deben comportarse los robots y cómo deben interactuar con las personas:

Ningún robot causará daño a un ser humano o permitirá, con su inacción, que un ser humano resulte dañado.

Todo robot obedecerá las órdenes recibidas de los seres humanos, excepto cuando esas órdenes puedan entrar en contradicción con la primera ley.

Todo robot debe proteger su propia existencia, siempre y cuando esta protección no entre en contradicción con la primera o la segunda ley (Ciudad Seva, 2025).

Por lo pronto, más nos vale ocuparnos de aprovechar los potenciales que nos ofrece esta tecnología y dejar para los cineastas y escritores las consideraciones sobre otros futuros alternativos.

Bibliografía

- Ciudad Seva. (2025). *Las tres leyes de la robótica*. <https://ciudadseva.com/texto/las-tres-leyes-de-la-robotica/>
- Departamento Nacional de Planeación. (2025). *Política Nacional de Inteligencia Artificial* (Documento CONPES 4144). <https://colaboracion.dnp.gov.co/CDT/Conpes/Economicos/4144.pdf>.
- Fundéu RAE. (2023, abril 25). *Ultrafalso, alternativa a deepfake*. <https://www.fundeu.es/recomendacion/ultrafalso-alternativa-a-deep-fake/>
- IBM. (2024, septiembre 6). *¿Qué es la transparencia de la IA?* <https://www.ibm.com/mx-es/think/topics/ai-transparency#:~:text=La%20explicabilidad%20de%20la%20IA,los%20modelos%20de%20machine%20learning>.
- Real Academia Española. (2025). *Distopía*. <https://dle.rae.es/distop%C3%ADa>