OFICINA DE SISTEMAS

POLÍTICA DE USO RESPONSABLE DE LA INTELIGENCIA ARTIFICIAL GENERATIVA EN EL DANE

Septiembre de 2026

CONTENIDO

Lista de figuras

Lista de tablas

USO RESPONSABLE DE LA INTELIGENCIA ARTIFICIAL GENERATIVA EN EL DANE

1. Introducción

Como autoridad estadística del país, el DANE tiene la responsabilidad de asegurar la producción de información estadística con criterios de calidad, integridad y transparencia, en cumplimiento de su misión institucional. En este contexto, el uso controlado y supervisado de la inteligencia artificial, se consolida como un habilitador estratégico que mejora la producción de información, optimiza procesos y proporciona herramientas de respaldo para la verificación y el análisis de los resultados antes de su difusión.

No obstante, la incorporación de innovaciones como la inteligencia artificial generativa requiere fortalecer el marco institucional existente mediante disposiciones que orienten su uso responsable, permitan gestionar los riesgos asociados y preserven la calidad, la integridad y la transparencia de la información, así como la confianza en la producción estadística oficial.

En este sentido, la política de uso responsable de la inteligencia artificial generativa del DANE establece un marco institucional de gestión y control para su implementación responsable, definiendo lineamientos, mecanismos de gestión de riesgos y controles para prevenir la materialización de los riesgos identificados y promover un uso consciente fundamentado en los principios definidos en la presente política, en concordancia con el CONPES 4144 y las directrices de la Política de Gobierno Digital.

El presente documento reconoce que los riesgos del uso de IA en el DANE no son estáticos ni completamente previsibles, por lo que su gestión debe ser constante, basada en evidencia y articulada con los procesos de producción estadística. En consecuencia, la presente política es objeto de revisión y actualización periódica, conforme a lo dispuesto en las disposiciones generales de este documento.

Esta política se inscribe en el marco de la Política Nacional de Inteligencia Artificial (CONPES 4144 de 2025), la Política de Gobierno Digital y los referentes internacionales sobre el uso responsable de la inteligencia artificial (UNESCO, OCDE, NIST y el Reglamento de Inteligencia Artificial de la Unión Europea). Se prioriza la inteligencia artificial generativa porque su capacidad de producir contenido nuevo plantea desafíos específicos de veracidad, trazabilidad, transparencia y protección de la

información, que exigen disposiciones particulares frente a los demás sistemas algorítmicos de la entidad.

El documento se estructura en cinco componentes. El primero corresponde a las disposiciones generales, que delimitan el propósito de la política, los principios que orientan el uso responsable de la inteligencia artificial generativa, su alcance y ámbito de aplicación, el marco normativo, las definiciones y el mecanismo de revisión y actualización. El segundo desarrolla la gestión de riesgos y los mecanismos de control, incluyendo la clasificación institucional de los riesgos y sus niveles, la evaluación de impacto algorítmico, los controles aplicables, el esquema de supervisión humana y las restricciones, prohibiciones y consecuencias derivadas del uso de estas tecnologías. El tercero aborda la gestión de datos para el uso de la IA generativa, con las reglas para la construcción de prompts, las condiciones aplicables a proveedores, la validación de las salidas generadas, la transparencia y declaración de uso, y la gestión de sesgos e incidentes. El cuarto establece el modelo de gobierno y seguridad, con sus pilares, la estructura y los roles de gobernanza en los niveles estratégico, táctico y operativo, y la integración de la seguridad de la información. Finalmente, el quinto componente se ocupa del fortalecimiento de capacidades institucionales y de la implementación de la política, orientada por los instrumentos complementarios que la desarrollan.

2. Disposiciones Generales

2.1 Propósito

Establecer la política institucional para orientar el uso responsable, seguro y ético de la inteligencia artificial generativa en el DANE, con el fin de fortalecer la calidad, la integridad, la confiabilidad y la transparencia de la producción estadística oficial, promoviendo la innovación en los procesos institucionales asociados a la generación y gestión de información. La política promueve, asimismo, el aprovechamiento de estas tecnologías para el fortalecimiento de la producción estadística oficial y el cumplimiento de la misión institucional, en los términos de los instrumentos complementarios que la desarrollan.

2.2 Principios para el uso responsable de la inteligencia artificial generativa

El uso responsable de la inteligencia artificial generativa en el DANE se orienta por los siguientes principios, que constituyen el fundamento de la gestión de riesgos, los controles y el modelo de gobierno definidos en la presente política:

Calidad e integridad estadística (IA centrada en datos confiables): el uso de la IA generativa debe preservar la exactitud, coherencia, consistencia, comparabilidad y oportunidad de la información, así como la validez metodológica del proceso estadístico.

Transparencia y trazabilidad (IA explicable, auditable y trazable): todo uso de IA generativa debe ser declarado, documentado y reconstruible, de manera que los productos puedan auditarse y se conozca la contribución de la tecnología a su elaboración.

Supervisión y responsabilidad humana: los resultados generados por IA requieren validación humana; la responsabilidad sobre el dato, el proceso y el resultado no se delega en el sistema, y quien firma o emite un producto responde por su contenido.

Protección de datos, privacidad y reserva estadística (IA segura): el uso de IA generativa debe garantizar el cumplimiento de la Ley 1581 de 2012, la Ley 2335 de 2023 y las obligaciones de confidencialidad y reserva estadística de la entidad.

Proporcionalidad y prevención: los controles se aplican de manera proporcional al nivel de riesgo de cada uso, privilegiando la prevención de la materialización de los riesgos sobre su corrección.

Equidad y no discriminación: el uso de IA generativa debe prevenir sesgos y resultados que generen diferencias injustificadas o discriminatorias entre personas o grupos.

Uso ético: la adopción de la IA generativa se enmarca en el Marco Ético para la IA en Colombia, el marco ético de las operaciones estadísticas y el Sistema de Ética Estadística de la entidad, preservando la autonomía, la explicabilidad, la rendición de cuentas y la legitimidad de las decisiones.

Mejora continua (IA reproducible): la gestión del uso de la IA generativa se revisa y ajusta de manera permanente, con base en evidencia, ante la evolución tecnológica, normativa e institucional, asegurando la reproducibilidad de los procesos en los que interviene.

2.3 Alcance

La presente política se aplica a todos los servidores públicos, contratistas y terceros naturales o jurídicos que desarrollen, implementen o utilicen herramientas o aplicaciones que incorporen capacidades de inteligencia artificial generativa, entendida como aquella basada en modelos que generan contenido nuevo a partir de instrucciones y patrones de aprendizaje.

El alcance cubre las herramientas de inteligencia artificial desarrolladas por la entidad y las proporcionadas por terceros, incluidas las soluciones en la nube o los servicios de acceso público. Comprende igualmente el tratamiento de los datos utilizados como entrada para el entrenamiento, la validación, las pruebas y la operación de los modelos, así como de los contenidos o resultados que estos generen. De la misma forma, comprende el uso de herramientas de inteligencia artificial generativa para el tratamiento y análisis de información geoespacial, cuando estas sean empleadas en el desarrollo de las funciones institucionales del DANE. Finalmente, tiene en cuenta tanto los usos asociados a la producción estadística como los de carácter administrativo y de apoyo, incluidos los procesos de gestión contractual, financiera, administrativa y del talento humano.

El cumplimiento de esta política es obligatorio y busca fortalecer el compromiso del DANE con la seguridad, la protección de datos y la integridad en la gestión de la información, promoviendo un ambiente de confianza y responsabilidad institucional.

Quedan excluidos del alcance de la presente política los sistemas algorítmicos no generativos que el DANE emplea en sus procesos de producción estadística, tales como los modelos de imputación, los sistemas de clasificación automática de actividades económicas y ocupaciones (CIIU/CIUO), los mecanismos de integración de registros (record linkage) y los modelos predictivos de uso institucional.

Esta delimitación obedece a los desafíos específicos de veracidad, trazabilidad y protección de la información que plantea la IA generativa, señalados en la introducción de la presente política.

Estos sistemas se desarrollan en el marco de instrumentos institucionales específicos, tales como lineamientos, guías, procedimientos y metodologías propias de la producción estadística, manteniendo la coherencia con el modelo de gobierno de la inteligencia artificial establecido en la presente política. La extensión del alcance de la presente política a dichos sistemas será evaluada en su primera revisión.

Los aspectos operativos derivados de la presente política —tales como los criterios y el listado de herramientas autorizadas, el catálogo de controles y sus evidencias, la guía para la gestión de prompts, el procedimiento de evaluación de impacto algorítmico y los indicadores de seguimiento— se desarrollan y actualizan mediante lineamientos, guías, procedimientos e instructivos complementarios expedidos por la Oficina de Sistemas, sin que su actualización requiera la modificación de la presente política. Estos instrumentos comprenderán, entre otros, el ciclo de vida y el registro de los sistemas de IA, el uso de la IA en las fases de la producción estadística oficial, la articulación con el gobierno de datos, la interoperabilidad y la arquitectura empresarial institucional, la seguridad específica de los sistemas de IA y las condiciones de adquisición de soluciones tecnológicas con capacidades de inteligencia artificial.

2.4 Ámbito de aplicación

La presente política rige para todas las dependencias del DANE, en sus diferentes sedes, y para todos los procesos estratégicos, misionales, de apoyo, de control y evaluación definidos en el mapa de procesos institucional. En este sentido, sus lineamientos son de obligatorio cumplimiento en todas las actividades en las que se desarrollen, implementen o utilicen herramientas, aplicaciones o sistemas que incorporen inteligencia artificial generativa, asociados a la producción estadística, la gestión institucional, el análisis de información, la atención a usuarios, la generación de contenidos, la gestión administrativa, financiera, contractual, del talento humano y cualquier otra actividad en la que se emplee esta tecnología generativa.

2.5 Normativa

El marco normativo que sustenta esta política comprende las siguientes normas nacionales:

- Constitución Política de Colombia (1991)
  Artículo 13. No discriminación y protección contra sesgos
  Artículo 15. Derecho a la intimidad y hábeas data,
  Artículo 20. Libertad de expresión e información
- Ley 1581 de 2012 - Protección de datos personales y hábeas data.
- Ley 1712 de 2014 - Transparencia y acceso a la información pública.
- Decreto 1078 de 2015 – Decreto Único del sector TIC (Política de Gobierno Digital).
- Marco Ético para la IA en Colombia 2021– Marco de principios para el desarrollo y uso ético de sistemas de inteligencia artificial en Colombia
- Ley 2335 de 2023 - Ley de estadísticas oficiales.
- Directiva Conjunta No. 007 de 2025 – Estándares sobre transparencia algorítmica en el uso de sistemas automatizados por parte de las entidades públicas.
- Sentencia T-067 de 2025 - Transparencia algorítmica y acceso a la información pública en sistemas digitales del Estado
- CONPES 4144 de 2025 - Política Nacional de Inteligencia Artificial

Con relación a la normativa interna del DANE se encuentran:

- Manual de políticas de seguridad de la información del DANE. Diciembre 2024
- Caracterización de producción estadística. Junio 2020
- Caracterización de la gestión de la información y transformación digital. Diciembre 2025
- Marco ético de las operaciones estadísticas y marco ético de los datos del DANE.
- Modelo de Seguridad y Privacidad de la Información (MSPI) – Resolución 02277 de 2025.

Como referentes internacionales se encuentran:

- Recomendación de la UNESCO sobre la Ética de la Inteligencia Artificial (2021) - Principios para una inteligencia artificial ética y centrada en los derechos humanos.
- Recomendación de la OCDE sobre Inteligencia Artificial (2024) - Principios para una inteligencia artificial confiable y transparente.
- Reglamento (UE) 2024/1689 – Ley de Inteligencia Artificial de la Unión Europea (AI Act) - Referente internacional para la clasificación y gestión de los sistemas de IA basadas en el riesgo, enfoque que adopta la presente política.
- Marco de Gestión de Riesgos de Inteligencia Artificial del NIST (AI RMF, 2023) - Referente para la identificación, medición y gestión de riesgos de los sistemas de IA.

2.6 Definiciones

Alucinación: fenómeno en el que los sistemas de IA generan y presentan con seguridad contenido erróneo o falso en respuesta a las indicaciones. También incluye resultados generados que divergen de las indicaciones u otras entradas, o que contradicen afirmaciones generadas previamente en el mismo contexto. (NIST, 2024)

Análisis espacial: conjunto de técnicas que estudian la localización, distribución y relaciones de los fenómenos en el territorio a partir de información geoespacial.

Ciclo de vida del sistema de IA: conjunto de etapas que comprende un sistema de inteligencia artificial desde su concepción, diseño, desarrollo o adquisición, pasando por su implementación, operación y monitoreo, hasta su actualización o retiro.

Clasificación de coberturas: proceso mediante el cual se identifican y categorizan los tipos de cobertura de la superficie terrestre a partir de imágenes u otras fuentes de datos geoespaciales.

Datos personales: cualquier información vinculada o que pueda asociarse a una o varias personas naturales determinadas o determinables. (Congreso de la República de Colombia, 2012)

Datos sensibles: "(…) a aquellos que afectan la intimidad del titular o cuyo uso indebido puede generar su discriminación, tales como aquellos que revelen el origen racial o étnico, la orientación política, las convicciones religiosas o filosóficas, la pertenencia a sindicatos, organizaciones sociales, de derechos humanos o que promueva intereses de cualquier partido político o que garanticen los derechos y garantías de partidos políticos de oposición así como los datos relativos a la salud, a la vida sexual y los datos biométricos" (Congreso de la República de Colombia, 2012)

Evaluación de impacto algorítmico (EIA): instrumento mediante el cual se identifican, analizan y gestionan los impactos y riesgos derivados de la implementación de un sistema de inteligencia artificial, de manera previa a su uso y a lo largo de su ciclo de vida.

Filtración por prompt (Prompt Injection): técnica mediante la cual se manipulan las instrucciones proporcionadas a un modelo para alterar su comportamiento o inducir la divulgación de información no autorizada. (NIST, 2024)

Gestión de datos: conjunto de prácticas orientadas a asegurar la calidad, la protección, la trazabilidad y el uso adecuado de los datos suministrados a los sistemas de IA generativa y de los resultados que estos producen.

Imágenes satelitales: registros obtenidos por sensores a bordo de satélites que capturan información de la superficie terrestre, utilizados como insumo para la producción y actualización de información estadística y geoespacial.

Información geoespacial: datos que describen objetos, fenómenos o eventos cuya posición está asociada al territorio, mediante coordenadas, direcciones o unidades geográficas de referencia.

Inteligencia Artificial (IA): sistema basado en máquinas que, para un conjunto de objetivos definidos por humanos, puede generar resultados como predicciones, recomendaciones o decisiones que influyen en entornos físicos o virtuales. (ISO & IEC, 2022)

Inteligencia Artificial Generativa: tipo de inteligencia artificial capaz de generar contenido nuevo, incluyendo texto, imágenes, audio, video o código, a partir de patrones aprendidos durante su entrenamiento. (NIST, 2024)

Marca de agua (Watermarking): mecanismo técnico utilizado para identificar o etiquetar contenido generado por inteligencia artificial con el fin de facilitar su trazabilidad y diferenciación frente a contenido generado por humanos. (NIST, 2024)

Modelos de IA: estructura computacional que contiene operaciones lógicas o aritméticas para generar una inferencia o predicción basada en datos de entrada y procesos de entrenamiento previamente definido. Adaptada de (ISO & IEC, 2022)

Modelo de lenguaje de gran tamaño (LLM): modelo de inteligencia artificial entrenado sobre grandes volúmenes de texto, capaz de comprender, procesar y generar lenguaje natural. (NIST, 2024)

Prompt: instrucción en lenguaje natural proporcionada a un sistema de IA generativa para orientar la generación de una respuesta. (NIST, 2024)

Sesgo algorítmico: tendencia sistemática de un sistema de inteligencia artificial a producir resultados que generan diferencias injustificadas o discriminatorias entre personas o grupos debido a sesgos presentes en los datos, el diseño del modelo o el contexto de uso. Adaptado de (NIST, 2022).

Sesgo de datos: distorsión presente en los datos utilizados en el entrenamiento o ajuste de modelos, que afecta su comportamiento y resultados (NIST, 2023).

Supervisión humana (Human in the loop): esquema de supervisión en el que los resultados generados por un sistema de inteligencia artificial requieren validación, intervención o aprobación humana antes de su aplicación. Adaptada de (Presidencia de la República de Colombia, 2021)

Supervisión humana (Human on the loop): esquema de supervisión en el que el sistema de inteligencia artificial opera con monitoreo humano permanente y capacidad de intervención, sin que cada resultado requiera aprobación previa, a diferencia del esquema human in the loop, en el que la validación humana precede al uso del resultado.

Reproducibilidad: capacidad de diferentes expertos para obtener los mismos resultados a partir de los mismos datos. (NIST, 2026)

Reserva estadística: es la obligación legal del DANE y de las entidades del SEN, en el marco de la producción estadística, de garantizar que los datos que impliquen la identificación directa o por deducción de las fuentes primarias o secundarias de personas naturales o jurídicas estén restringidos al público en general, a las entidades públicas y privadas, a los organismos oficiales y a las autoridades públicas. (Congreso de la República de Colombia, 2023)

Riesgo en el uso de IA generativa: posibilidad de que el uso de herramientas de inteligencia artificial generativa afecte la calidad de la información, la validez del proceso estadístico, los derechos de las personas, la seguridad de la información o la reputación institucional.

2.7 Revisión y actualización de la política

La Política de Uso Responsable de la Inteligencia Artificial Generativa en el DANE es un instrumento sujeto a revisión y actualización periódica, por lo que se llevarán a cabo revisiones anuales o extraordinarias cuando se presenten cambios normativos, tecnológicos o institucionales que puedan afectar su vigencia o aplicación. Esta revisión busca incorporar aspectos emergentes derivados de la evolución de la inteligencia artificial y las necesidades institucionales del DANE. La revisión incorpora, como insumo, las lecciones aprendidas derivadas de la implementación de la política. La revisión periódica es liderada por la Oficina de Sistemas, con la participación de las dependencias y de los actores institucionales pertinentes.

Las actualizaciones producto de la revisión a la política no son reconocidas como oficiales hasta tanto no sean presentadas y aprobadas por la instancia estratégica definida en el modelo de gobierno (Comité Institucional de Gestión y Desempeño), conforme al procedimiento establecido para tal fin.

3. Gestión de riesgos y mecanismos de control para el uso de la IA generativa

El DANE adopta un enfoque de gestión de riesgos para el uso de la inteligencia artificial generativa que va más allá de la detección inicial. Se extiende a todo el ciclo de vida del sistema de IA generativa, asegurando que cualquier afectación potencial a la calidad estadística, la comparabilidad de la información, la confidencialidad de los datos y la confianza pública sea identificada, evaluada y tratada oportunamente.

Este capítulo desarrolla, en su orden, la clasificación de los riesgos institucionales asociados al uso de la IA generativa y sus niveles, la evaluación de impacto algorítmico, los controles aplicables, el esquema de supervisión humana y las restricciones, prohibiciones y consecuencias derivadas de su uso.

La clasificación de riesgos definida por el DANE incluye controles que se actualizan en función de la evolución tecnológica, normativa y operativa. Este enfoque se fundamenta en el compromiso de toda la entidad de construir un entendimiento común de las amenazas y oportunidades asociadas al uso de la inteligencia artificial generativa, como requisito para una adopción responsable.

Bajo este enfoque, la política propende porque todos los actores involucrados comprendan y cumplan, sin excepción, los lineamientos establecidos, a fin de prevenir situaciones que puedan afectar negativamente a la entidad y a la calidad de sus resultados.

3.1 Clasificación de riesgos institucionales para el uso de la IA generativa

La gestión de riesgos definida por el DANE se basa en la identificación, análisis, evaluación y tratamiento de riesgos asociados al uso de IA generativa, de manera proporcional al impacto que estos puedan tener en los procesos institucionales y en la producción estadística oficial. En consecuencia, se establece una clasificación de riesgos, con base en los criterios de:

1) **Impacto en la calidad de la información generada**. Grado en que el uso de la IA puede afectar la calidad de la información producida, en términos de exactitud, coherencia, consistencia y oportunidad.
2) **Afectación de la validez del proceso estadístico**. Nivel de alteración de la metodología que respalda cada una de las operaciones estadísticas derivadas del uso de la IA.
3) **Nivel de vulneración ética y derechos de las personas.** Grado de afectación a la privacidad, a la exposición de datos personales y a la discriminación, así como a la autonomía, la explicabilidad,

la rendición de cuentas y la legitimidad de las decisiones, conforme a los principios de la presente política.

4) **Impacto reputacional de la entidad**. Efecto que el uso de la IA puede tener en la confianza, la credibilidad y la imagen institucional del DANE.
5) **Impacto dentro de procesos misionales**. Nivel en el que la IA incide en los procesos de producción, análisis y difusión de la información estadística o en las actividades administrativas o de apoyo.
6) **Nivel de autonomía y toma de decisiones**. Grado en el que la IA opera sin intervención humana y asume decisiones.
7) **Nivel de sensibilidad de los datos utilizados**. Grado en que la IA hace uso indebido de los datos protegidos a cargo del DANE o realiza un tratamiento no autorizado de los mismos.

A partir de estos criterios, la presente política establece la siguiente tipificación de riesgos para el DANE:

1) Riesgos sobre la integridad estadística: posibilidad de que el uso de la IA generativa altere la validez metodológica, la exactitud o la consistencia de los resultados del proceso estadístico.
2) Riesgos de falta de transparencia e interpretación de los modelos: derivados de la opacidad de los modelos, que dificulta explicar, auditar o reproducir la forma en que se generó un resultado.
3) Riesgos por generación de información incorrecta (alucinaciones): producción de contenido erróneo o no verificable presentado con apariencia de validez.
4) Riesgos de discriminación y generación de desigualdades: reproducción o amplificación de sesgos que generan diferencias injustificadas entre personas o grupos.
5) Riesgos de vulneración de derechos: afectaciones a la intimidad, el hábeas data, la autonomía y los demás derechos de las personas.
6) Riesgos sobre la reserva y confidencialidad de la información: exposición de microdatos, datos personales o información amparada por la reserva estadística.
7) Riesgos de acceso y uso no autorizado: uso de herramientas no evaluadas por la entidad o por fuera de las condiciones institucionales definidas.
8) Riesgos de ciberseguridad: amenazas técnicas sobre los sistemas de IA, como la inyección de instrucciones o la filtración de información.
9) Riesgos por errores en delimitaciones, geocodificación, análisis espacial o cartografía generada mediante IA

Adicionalmente, los riesgos se clasifican en los niveles bajo, medio y alto, lo que permite establecer tratamientos diferenciados y determinar el grado de supervisión requerido.

Esta tipificación se fundamenta en los criterios definidos anteriormente y busca cubrir, de manera razonable, las fuentes de preocupación asociadas al uso de la IA generativa en la entidad. La tipificación y los niveles de riesgo cumplen funciones complementarias: los tipos identifican qué puede resultar afectado por el uso de la IA generativa, mientras que los niveles bajo, medio y alto gradúan la magnitud de la afectación potencial. En consecuencia, la clasificación de cada caso de uso comprende la identificación de los tipos de riesgo presentes y la asignación de su nivel, y los mecanismos de la presente política —la evaluación de impacto algorítmico, los controles, la supervisión humana, las restricciones y las prohibiciones— tratan los tipos identificados con una exigencia proporcional al nivel asignado. Para la clasificación de casos dudosos o que combinen varios tipos o niveles de riesgo, la entidad desarrollará una guía o matriz de clasificación como instrumento complementario.

3.2 Riesgo nivel bajo

Se considera riesgo de nivel bajo el uso de herramientas o aplicaciones de inteligencia artificial generativa que cumplen funciones como la elaboración de documentos, la organización y sistematización de la información, la automatización de actividades administrativas, contractuales, financieros, de gestión del talento humano y de apoyo del DANE, sin implicación directa en los procesos misionales de producción estadística ni en la generación de resultados oficiales.

Se consideran de riesgo bajo aquellos usos cuya afectación potencial sobre la producción estadística, los derechos de las personas y la información institucional es limitada y puede mitigarse mediante controles básicos.

Entre los usos de nivel bajo de riesgo se encuentran la generación de borradores de documentos, la revisión de redacción y estilo de comunicaciones internas, la traducción de textos, siempre que estos no contengan datos que afecten la intimidad del titular o generen discriminación, la consulta y síntesis de referentes bibliográficos y la generación de contenido gráfico de apoyo, siempre que se reconozca la intervención de la IA en su creación, entre otros.

En este nivel, la IA actúa como herramienta con fines exploratorios o asistenciales, sin capacidad para alterar metodologías estadísticas, transformar datos institucionales ni influir en decisiones técnicas o estratégicas del DANE. El impacto del riesgo sobre los criterios de calidad de la información, validez

metodológica, derechos de las personas y reputación institucional es bajo, siempre que se mantengan las condiciones de uso aquí establecidas.

Para efectos de control institucional, el DANE únicamente autoriza el uso de herramientas de inteligencia artificial generativa que hayan sido evaluadas y aprobadas por la entidad, las cuales pueden utilizarse en modalidades de acceso público o en entornos institucionales habilitados, de conformidad con las condiciones definidas en esta política. Los controles, restricciones y prohibiciones aplicables se desarrollan en los numerales correspondientes de la presente política.

La Oficina de Sistemas es responsable de revisar, actualizar y publicar el listado vigente de herramientas autorizadas, así como las modalidades de acceso y las condiciones específicas para su uso, mediante el procedimiento o lineamiento correspondiente, sin que ello requiera modificar la presente política. La autorización de nuevas herramientas de inteligencia artificial generativa, incluso cuando sean de acceso público o gratuito, se realizará previa evaluación institucional, considerando, como mínimo, los siguientes criterios:

1) Cumplimiento normativo: La herramienta permite cumplir la normativa vigente relacionada con la protección de datos personales, la reserva estadística, la seguridad de la información y la gestión documental.
2) Tratamiento de la información: El proveedor publica las condiciones para el tratamiento, almacenamiento, retención y uso de la información suministrada a la herramienta
3) Configuración de privacidad: La herramienta permite configurar el tratamiento de las interacciones para entrenamiento de modelos u otros usos por parte del proveedor
4) Modalidades de acceso y uso: La herramienta dispone de modalidades de acceso y uso compatibles con los controles definidos por el DANE para cada nivel de riesgo
5) Transparencia del proveedor: El proveedor publica información suficiente sobre las capacidades, limitaciones y condiciones de uso de la herramienta
6) Seguridad: La herramienta dispone de mecanismos que permitan gestionar su uso conforme a las condiciones de seguridad definidas por el DANE
7) Interoperabilidad: La herramienta permite su utilización con la plataforma tecnológica institucional sin afectar las condiciones de seguridad y gobernanza definidas por el DANE.

En todos los casos, es obligatorio asegurar el cumplimiento de la prohibición de ingreso de microdatos según lo establecido en la Ley 2335 de 2023; datos sometidos a reserva estadística; datos personales, de acuerdo con la Ley 1581 de 2012; información clasificada, confidencial o de uso interno; así como información que permita la identificación directa o indirecta de fuentes primarias, secundarias o de

unidades estadísticas. Estas condiciones constituyen un control de protección de la información y deben cumplirse en todos los casos sin excepción.

Adicionalmente, se requiere que los resultados sean validados por el usuario, asegurando la coherencia, precisión y contexto institucional del contenido generado, sin delegar el criterio profesional en la herramienta. Este enfoque es coherente con marcos internacionales basados en el riesgo, como el Reglamento de Inteligencia Artificial de la Unión Europea (AI Act), incluido en el marco normativo de la presente política.

3.3 Riesgo nivel medio

El nivel medio se refiere al uso de sistemas de IA en los procesos misionales y estratégicos del DANE que involucren el procesamiento, depuración, clasificación y análisis exploratorio de la información, así como la automatización de los procesos que conforman la producción estadística. En estos escenarios, la IA no genera los resultados finales, pero sí participa en la elaboración de insumos y en las etapas intermedias de su construcción, lo que conlleva riesgos asociados a la calidad de los datos, la consistencia de los análisis, la introducción de sesgos y la trazabilidad de los procesos.

Se consideran de riesgo medio aquellos usos cuya afectación potencial sobre los insumos y las etapas intermedias de los procesos institucionales es moderada y requiere controles intermedios y supervisión humana continua.

Algunos usos que se encuentran en este nivel son el apoyo al análisis exploratorio de datos, la generación de código o scripts para el procesamiento de información, la clasificación preliminar de datos y metadatos para su uso y difusión, la validación de elementos de calidad asociados a productos intermedios, la automatización de procesos relacionados con la codificación y catalogación de fuentes, la comprobación de la integridad de registros administrativos, entre otros. También se encuentran en este nivel el procesamiento de imágenes satelitales, la clasificación preliminar de objetos geográficos, la generación automatizada de insumos cartográficos y la geocodificación masiva, cuando no formen parte de productos oficiales.

En cuanto al uso de herramientas de inteligencia artificial generativa en este nivel, se prohíbe expresamente el ingreso de microdatos, datos sometidos a reserva estadística, datos identificables o potencialmente reidentificables, datos personales, comunicaciones internas que contengan información sensible, así como información clasificada, confidencial o de uso interno, aun cuando dicha información haya sido sometida a procesos de anonimización o desidentificación.

Con respecto al uso de herramientas de inteligencia artificial generativa contratadas por el DANE, es obligatorio establecer condiciones contractuales que garanticen la no retención de datos (zero data retention), la no utilización de la información para el entrenamiento de modelos, así como el cumplimiento de los lineamientos institucionales de seguridad de la información. Estas condiciones se desarrollan en el numeral de uso de la IA por proveedores y condiciones contractuales de la presente política.

Adicionalmente, el uso de sistemas de riesgo de nivel medio requiere implementar controles intermedios, incluyendo la validación de resultados, la documentación del uso de la IA, los mecanismos de trazabilidad y la supervisión humana continua, en línea con enfoques internacionales basados en el riesgo, como el AI Act de la Unión Europea, asociados a casos de riesgo limitado. Los criterios y condiciones para la implementación de estos controles serán desarrollados mediante los lineamientos, procedimientos o instrumentos que expida el DANE para la implementación de la presente política

3.4 Riesgo nivel alto

Se define como de alto nivel el uso de sistemas de IA generativa en los procesos misionales del DANE, como la generación de información estadística oficial o la producción de resultados orientados a la toma de decisiones institucionales. En el primer caso, la IA se emplea en etapas como el procesamiento, el análisis o el apoyo a la elaboración de propuestas de elementos metodológicos, incluidos el diseño muestral, los factores de expansión, las reglas de imputación, las definiciones conceptuales o los instrumentos de recolección. En el segundo, participa en la generación de insumos analíticos, modelos o recomendaciones que pueden influir en las decisiones estratégicas del DANE. En ambos casos, la IA puede apoyar la exploración metodológica, pero no puede sustituir el juicio estadístico experto. La definición o modificación de dichos elementos corresponde siempre al juicio estadístico experto, conforme a las prohibiciones de la presente política.

Se consideran de riesgo alto aquellos usos cuya afectación potencial sobre los resultados oficiales, los derechos de las personas o la confianza institucional es significativa y exige controles a lo largo de todo el ciclo de vida del sistema y supervisión humana rigurosa.

En este nivel de riesgo, los sistemas de IA generativa intervienen directamente en los resultados y procesos del DANE, con posibles efectos en la calidad de la información, la consistencia de los análisis, la trazabilidad de las decisiones, la protección de los datos y la confianza institucional, lo que representa un alto impacto para la entidad y para los derechos de las personas.

Se consideran usos de alto riesgo:

- Uso de chatbots o asistentes virtuales basados en IA generativa que interactúan directamente con informantes durante operativos de recolección.
- Generación automatizada de metadatos, notas técnicas o documentación metodológica con carácter oficial.
- Generación de resúmenes, análisis o contenidos que alimentan boletines estadísticos, comunicados de prensa u otros productos institucionales.
- Uso de código generado por IA que se incorpora a ambientes productivos sin validación técnica y revisión humana.
- Reescritura o generación asistida por IA de instrumentos de recolección, formularios o fichas metodológicas.
- Generación automatizada de cartografía temática, productos geoespaciales, delimitaciones espaciales o análisis territoriales que formen parte de publicaciones oficiales, instrumentos técnicos, productos de difusión o decisiones institucionales.

Los usos enmarcados en este nivel requieren controles a lo largo de todo el ciclo de vida del sistema de IA, incluyendo evaluación de riesgos, validación metodológica, documentación técnica completa, trazabilidad de las decisiones, supervisión humana rigurosa y monitoreo continuo de desempeño. Estos controles se desarrollan en los numerales de evaluación de impacto algorítmico, controles y supervisión humana de la presente política.

3.5 Evaluación de impacto algorítmico

La política de uso de IA en el DANE establece que la Evaluación de Impacto Algorítmico (EIA) es un requisito obligatorio para los sistemas de IA clasificados en niveles de riesgo medio y alto, con el propósito de analizar y gestionar los impactos derivados de su implementación en los procesos institucionales y en la producción estadística oficial. La evaluación se realiza de manera preventiva en la fase de diseño del sistema de IA y a modo de control, a lo largo de todo el ciclo de vida, asegurando que se mantengan el propósito y el alcance definidos inicialmente y que su implementación no genere riesgos emergentes que afecten la calidad de la información, los compromisos éticos institucionales ni la confianza en la entidad.

Para ello, el DANE adopta el Algorithmic Impact Assessment (AIA) de la UNESCO como instrumento de evaluación, en línea con la recomendación sobre la ética de la inteligencia artificial y lo adapta para

aplicarlo a herramientas o aplicaciones de inteligencia artificial generativa. En este contexto, el alcance de la evaluación incluye su articulación con la fase del proceso estadístico, la identificación de sesgos o errores, la evaluación de su impacto en la consistencia, así como la verificación del cumplimiento normativo relativo a la confidencialidad, la reserva estadística y la protección de datos. Cuando el sistema involucre modelos geoespaciales o procesos de clasificación de imágenes para la identificación de coberturas y objetos, la evaluación incluye la revisión de las variables específicas asociadas a dichos procesos.

Para su aplicación, la evaluación incorpora la definición de niveles aceptables de error o generación de contenido no verificable, los mecanismos de validación humana de las salidas, las condiciones de uso y protección de la información en los prompts, la jurisdicción y residencia de los datos, las condiciones contractuales sobre el uso de la información por parte del proveedor, así como la trazabilidad del proceso desde la generación de la instrucción hasta el uso institucional del resultado. Asimismo, la evaluación asegura que los resultados generados por la IA sean explicables e identificables, especialmente cuando forme parte de productos o procesos institucionales.

El resultado de la evaluación se presenta ante la instancia de gobierno correspondiente —el nivel táctico para los casos de riesgo medio y el Comité Institucional de Gestión y Desempeño para los de riesgo alto—, como parte de los documentos del sistema, para su aprobación. Se debe realizar una nueva evaluación, antes de continuar con la operación del sistema, cuando se presenten cambios en el modelo, en el propósito de uso, en las fuentes de datos, en el proveedor o en el nivel de riesgo identificado. El procedimiento institucional de la evaluación —los responsables de su elaboración, el formato, las variables evaluadas y su periodicidad— se define en el instrumento complementario que expida la entidad. Los contenidos generados que no se incorporen a productos oficiales, tales como prototipos, versiones preliminares o resultados de prueba, deben gestionarse conforme al ciclo de vida del dato y a las Tablas de Retención Documental aplicables, de acuerdo con los lineamientos de gestión documental de la entidad.

3.6 Controles para el uso de IA generativa

Los riesgos asociados al uso de IA en el DANE se gestionan mediante controles técnicos, organizacionales y procedimentales, de acuerdo con su nivel de riesgo y su vinculación con el proceso estadístico. En consecuencia, deben aplicarse y evaluarse con rigor, garantizando que cualquier desviación sea detectada y reportada oportunamente con el propósito de proteger los principios de calidad estadística, transparencia y responsabilidad institucional.

Los controles definidos en la presente política se establecen como mecanismos verificables de las acciones que el usuario de las herramientas de IA realiza para evitar y mitigar los riesgos identificados. Cada control se establece en función del uso y del nivel de riesgo al que aplica, por lo que su complejidad varía según dicho nivel y su impacto sobre la producción estadística y los demás procesos institucionales, así:

La aplicación de cada control debe quedar debidamente documentada conforme a los lineamientos institucionales y es verificable según el nivel de riesgo del uso.

Controles para el uso de la IA en la elaboración de documentos. El uso de la IA para la generación de documentos institucionales debe describirse en notas metodológicas, mencionando el sistema de IA empleado y el rol que desempeñó en la elaboración del documento, por ejemplo: consulta de referentes, editor de estilo, generador de imágenes, etc. Si el documento contiene un historial de versiones, debe indicarse la participación de la IA en las modificaciones. Adicionalmente, se debe garantizar el respeto de los derechos de autor y de la propiedad intelectual, evitando la reproducción no autorizada de obras protegidas, el uso indebido de contenidos de terceros o la apropiación de la autoría de información generada o derivada sin la debida verificación. En todo caso, quien firma o emite un documento institucional se presume responsable de su contenido, con independencia del apoyo tecnológico empleado en su elaboración, en línea con el criterio del Auto AC739-2026 de la Corte Suprema de Justicia.

Este control se articula con los lineamientos de producción documental, asegurando el cumplimiento de metodologías, regulaciones y estándares en el contexto técnico del DANE. Ante todo, es responsabilidad del usuario del sistema validar el contenido generado.

Controles en el uso de la IA como insumo para los procesos misionales. Cuando la IA se utilice para generar insumos en alguna de las fases del proceso estadístico, se debe validar el resultado mediante contrastes con referentes técnicos y verificar la coherencia, la exactitud y la consistencia de las salidas, entre otros atributos de calidad, especialmente en procesos como la imputación de datos, la clasificación automática o el apoyo en el análisis estadístico. Cuando el insumo corresponda a clasificación de coberturas con imágenes satelitales o detección de objetos realizada mediante scripts generados con IA o soluciones integradas en software licenciado, la validación debe verificar la consistencia de los datos obtenidos conforme a los lineamientos aplicables.

Este control se materializa mediante informes, notas técnicas o registros de validación generados por el profesional que realiza la comprobación. Todo resultado que genere inconsistencia, sesgo, alucinación

o comportamiento imprevisible del sistema debe ser reportado ante la instancia de gobierno correspondiente, indicando detalladamente el evento.

Control de la validación técnica y metodológica de los sistemas de IA. Este control incluye realizar evaluaciones documentadas del desempeño del sistema, evaluar la estabilidad frente a variaciones en los datos y la consistencia de los resultados en ejecuciones repetidas. Adicionalmente, generar los conceptos técnicos relativos a la coherencia del uso de IA con los supuestos, los diseños y los estándares metodológicos del proceso estadístico. Asimismo, producir informes sobre los modelos y los datos utilizados, las configuraciones aplicadas y los resultados obtenidos. Documentar el origen, el tratamiento y el uso de la información, así como los procesos de limpieza, transformación y anonimización que lleva a cabo el sistema de IA. Este control se articula con los lineamientos de gestión de la calidad estadística y de gobierno de datos del DANE. Lo anterior en cumplimiento de la Ley 1581 de 2012 y de la Ley 2335 de 2023.

La evidencia de este control se materializa en actas, informes de validación, fichas de datos y bitácoras de procesamiento, que forman parte de la documentación del sistema, objeto de verificación mediante revisiones focalizadas, con el fin de detectar hallazgos y desviaciones y generar acciones correctivas.

El control de supervisión humana. Todo uso de IA en el DANE debe operar bajo un esquema de supervisión humana, con la asignación explícita de un responsable técnico. El alcance de este control, sus mecanismos y su gradación según el nivel de riesgo se desarrollan en el numeral de Supervisión Humana de la presente política.

Control de herramientas, acceso y uso institucional. El DANE define los lineamientos institucionales sobre las restricciones de acceso, los mecanismos de monitoreo, las herramientas habilitadas y las condiciones bajo las cuales pueden utilizarse los sistemas de IA, promoviendo la articulación con las políticas de seguridad de la información. Este control asegura el uso de los datos en entornos controlados, especialmente durante su procesamiento. El objeto de este control es verificar que el uso de sistemas de IA generativa se realice únicamente mediante herramientas autorizadas por el DANE, conforme a las restricciones de acceso, las políticas de seguridad de la información y las condiciones de uso definidas institucionalmente.

Asimismo, verifica el cumplimiento de las disposiciones mencionadas y brinda acompañamiento para fortalecer las capacidades del talento humano en el uso responsable y ético de la IA. El uso de las herramientas institucionales de IA generativa está condicionado al cumplimiento de los procesos de sensibilización y capacitación definidos por la entidad, cuyo cumplimiento es objeto de seguimiento y verificación.

La aplicación de los controles es obligatoria para todas las dependencias del DANE y su incumplimiento conlleva las consecuencias previstas en la presente política, así como la aplicación de la normativa vigente en los ámbitos administrativo, disciplinario o contractual, según corresponda. En complemento, la verificación de los controles se realiza conforme a los lineamientos del Modelo Estándar de Control Interno (MECI), mediante monitoreo basados en evidencia efectuados por los líderes de proceso y la Oficina de Sistemas en el marco de la gestión del riesgo por líneas de defensa. . Los criterios de verificación se especifican en la Tabla 1 al final de la presente política.

3.7 Supervisión Humana

Para el DANE, la supervisión humana en el contexto de uso de la IA generativa se considera una capacidad institucional que se materializa mediante la aplicación de mecanismos de control orientados a la utilización ética y responsable de la IA, sin comprometer la integridad del proceso estadístico ni la confiabilidad de la información producida, asegurando que el comportamiento del sistema se mantenga estable, coherente y alineado con los criterios metodológicos, técnicos y de calidad definidos por la entidad.

Ante todo, es imprescindible que la responsabilidad sobre el dato, el proceso y el resultado no se delegue en el sistema de IA y que se designe un responsable técnico que asuma la validación del sistema, la interpretación de su comportamiento y la evaluación de su impacto en la calidad de la información. En todos los casos, el DANE debe mantener el control del propósito, el alcance y la continuidad del uso de la IA.

En desarrollo del control de supervisión humana, el responsable técnico designado valida los resultados generados antes de su puesta en funcionamiento, realiza el seguimiento periódico del desempeño del sistema en operación y detecta desviaciones, cambios de comportamiento, inconsistencias y alucinaciones. Este control implica, además, la emisión de conceptos técnicos sobre las decisiones relativas al funcionamiento o a la continuidad del uso de la IA y la gestión de los cambios, justificando

la necesidad que origina la solicitud; la supervisión se materializa mediante las aprobaciones de los sistemas de IA, a través de actas, informes y comunicaciones que evidencien el juicio experto aplicado.

Para efectos de la presente política se distinguen dos esquemas de supervisión: human in the loop, en el que los resultados del sistema requieren validación o aprobación humana antes de su aplicación, y human on the loop, en el que el sistema opera bajo monitoreo humano permanente con capacidad de intervención. El esquema exigible depende del nivel de riesgo del uso. La supervisión humana aplica a todos los usos de la IA generativa, incluidos los administrativos y de apoyo.

La supervisión humana está asociada al nivel de riesgo identificado en la presente política para cada sistema de IA. En los casos de riesgo bajo, la supervisión se realiza mediante la verificación del contenido generado, la revisión de errores, omisiones o interpretaciones inadecuadas, la alineación con el contexto del DANE y la protección de los derechos de autor cuando se empleen referencias documentales elaboradas por terceros, en el marco de un esquema de supervisión tipo *human on the loop*.

Por otro lado, en los casos clasificados como de riesgo medio, la supervisión humana se enfoca en verificar la consistencia, la completitud y la coherencia de los resultados, así como su estabilidad ante ejecuciones repetidas. Esta supervisión incluye la verificación de la coherencia entre variables y unidades de observación, la comparación con series históricas o fuentes de referencia y la aplicación de análisis para detectar valores atípicos y comportamientos irregulares. Asimismo, aborda evaluaciones del impacto de incluir o excluir registros, con el fin de determinar la solidez de los resultados y su coherencia con el diseño metodológico de la operación estadística.

En los casos clasificados como de alto riesgo, la supervisión se centra en el control metodológico y de la calidad estadística, mediante la intervención directa del profesional responsable en la validación de los resultados antes de su uso (*human in the loop*). Más allá de la verificación de los resultados, evalúa el efecto del uso de la IA en el proceso estadístico, considerando aspectos como el sesgo, la variabilidad, la representatividad, el impacto en el diseño estadístico, el marco conceptual y los procesos de tratamiento de datos.

La supervisión también incluye la evaluación de posibles alteraciones en la comparabilidad temporal, territorial o entre dominios de estudio. De igual forma, debe asegurarse de que los procesos en los que interviene la IA sean trazables, reproducibles y debidamente documentados, lo que garantiza la transparencia y la calidad de la información producida. Los resultados que involucren GeoIA para geocodificación, análisis espacial o extracción de elementos sobre imágenes satelitales deben ser validados por personal especializado, y los productos geoespaciales elaborados con apoyo de IA deben

surtir un proceso de supervisión humana que verifique el cumplimiento de los parámetros metodológicos y de calidad establecidos.

La aplicación de IA generativa en este nivel requiere la revisión y aprobación por la más alta instancia de gobierno definida por el DANE para el uso de la IA, así como la formalización de su aplicación mediante documentos técnicos que respalden su uso en la operación estadística y en los resultados obtenidos. Finalmente, el uso de sistemas de IA en los procesos de producción de información debe guardar coherencia con el diseño estadístico definido para cada operación.

3.8 Restricciones en el uso de la IA

El DANE promueve el uso de herramientas o aplicaciones de inteligencia artificial generativa bajo un conjunto de condiciones controladas que delimitan su aplicación respecto al tipo de información que puede ser utilizada, los entornos tecnológicos habilitados, el alcance de los resultados generados y su incorporación en productos o decisiones, de manera que se prevengan riesgos operativos, metodológicos y de seguridad de la información, sin limitar los espacios de exploración y aprovechamiento responsable de estas tecnologías.

Las restricciones al uso de sistemas de IA generativa definidas para el DANE por la presente política son:

Para efectos de la presente política, las restricciones establecen las condiciones bajo las cuales el uso de la IA generativa se encuentra autorizado, mientras que las prohibiciones, desarrolladas en el numeral siguiente, identifican conductas que no pueden realizarse en ningún caso.

1) Emplear exclusivamente las herramientas autorizadas por la entidad conforme a la lista vigente publicada por la Oficina de Sistemas en los canales institucionales definidos y bajo las condiciones definidas para cada nivel de riesgo, evitando el uso de sistemas que no hayan sido evaluados conforme a los criterios institucionales.
2) Limitar el uso de la IA generativa a los procesos y fines autorizados por el DANE sin extender su aplicación a contextos no autorizados o sin la validación correspondiente.
3) Emplear en los sistemas de IA información que no comprometa la confidencialidad, la reserva estadística, la protección de datos personales ni la seguridad de la información del DANE, conforme a lo expresamente prohibido en la presente política. Esta restricción aplica a todos los niveles de riesgo y a todas las herramientas empleadas.
4) Condicionar el uso de los resultados generados por sistemas de IA a procesos de validación técnica y metodológica, conforme a los controles definidos en la presente política, antes de su

incorporación a la producción de información estadística oficial, a la emisión de conceptos institucionales o a la toma de decisiones.

5) Restringir la modificación de definiciones conceptuales, metodologías, instrumentos de recolección o criterios técnicos del proceso estadístico, como resultado de la IA, a excepción de su uso en el marco de ejercicios de análisis o de exploración controlada, con evaluación de los efectos directos sobre la producción oficial.
6) Limitar el uso de sistemas de IA generativa en procesos automatizados que no cuenten con mecanismos de supervisión humana acordes con su nivel de riesgo, asegurando, en todos los casos, la presencia del juicio experto en la validación, interpretación y uso de los resultados.
7) Restringir el uso de sistemas de IA generativa a aquellos que se encuentren registrados, cuenten con trazabilidad y estén alineados con el modelo de gobierno institucional y con los lineamientos definidos para la gestión de datos, la seguridad de la información y el control interno, de modo que su incorporación se realice de forma controlada, transparente y responsable.
8) Restringir el uso de servicios de inteligencia artificial generativa que no garanticen condiciones adecuadas para el tratamiento de la información utilizada por el DANE, particularmente cuando esta pueda ser procesada fuera del territorio nacional, quedar sujeta a jurisdicciones extranjeras o ser objeto de tratamientos posteriores por parte del proveedor. La evaluación de dichas garantías se realiza conforme a las condiciones establecidas en el numeral de uso de la IA por proveedores y condiciones contractuales.

Estas restricciones son de cumplimiento obligatorio para todas las dependencias y los servidores del DANE y su aplicación se articula con los controles, los mecanismos de supervisión y los lineamientos de gobierno definidos en la presente política. Asimismo, las restricciones indicadas complementan las prohibiciones contenidas en este documento y consolidan un marco integral para el uso ético, responsable y controlado de la inteligencia artificial generativa en el DANE.

3.9 Prohibiciones de uso de IA

El uso de herramientas o aplicaciones de inteligencia artificial generativa en el DANE está sujeto al cumplimiento de las reglas de aplicación obligatoria, sin excepciones, independientemente del nivel de riesgo o del tipo de herramienta utilizada.

Las prohibiciones que establece el DANE son:

- Está prohibido el uso de microdatos, datos protegidos con reserva estadística, datos identificables y datos personales o que afecten la intimidad del titular o generen discriminación en servicios de inteligencia artificial generativa que no hayan sido habilitados por el DANE, independientemente de la aplicación de procesos de anonimización o de desidentificación.
- Está prohibido el uso de contenido generado por herramientas o aplicaciones de inteligencia artificial generativa sin la aplicación de controles de supervisión humana antes de incorporarlo en productos, procesos o decisiones institucionales.
- Está prohibido el uso de herramientas de inteligencia artificial generativa no autorizadas por el DANE, así como su uso para fines distintos de los aprobados.
- Está prohibido el uso de herramientas o aplicaciones de inteligencia artificial generativa para la generación directa de resultados estadísticos oficiales, conceptos institucionales o productos destinados a la toma de decisiones, sin la validación técnica y metodológica correspondiente.
- Está prohibida la incorporación de resultados generados por inteligencia artificial en productos institucionales sin la debida verificación de su calidad, coherencia y consistencia, así como sin la identificación de su origen.
- Está prohibido omitir la declaración del uso de herramientas o aplicaciones de inteligencia artificial en documentos, publicaciones o productos institucionales elaborados, total o parcialmente, con esta tecnología. Está prohibido el uso de herramientas o aplicaciones de inteligencia artificial generativa para modificar o definir elementos metodológicos del proceso estadístico, tales como definiciones conceptuales, diseños muestrales, instrumentos de recolección o criterios técnicos.
- Está prohibida la implementación de asistentes o sistemas basados en inteligencia artificial generativa que interactúen directamente con informantes o fuentes de información sin contar con la Evaluación de Impacto Algorítmico (EIA), aplicable conforme al nivel de riesgo definido en la presente política, y los controles correspondientes.
- Está prohibida la implementación de procesos automatizados basados en inteligencia artificial generativa que operen sin mecanismos de control o que sustituyan el juicio experto en la validación y la interpretación de los resultados.
- Está prohibido omitir el registro, la trazabilidad o la documentación del uso de herramientas o aplicaciones de inteligencia artificial generativa en los procesos y productos del DANE.
- Está prohibido el uso de herramientas de inteligencia artificial generativa para procesar datos transversales que, mediante el cruce de información geográfica, demográfica, económica o social, puedan permitir la identificación de sujetos específicos o vulnerar la reserva estadística.

El incumplimiento de las presentes prohibiciones constituye una infracción a los lineamientos institucionales del DANE y da lugar a las consecuencias definidas en esta política, así como a la aplicación

de la normativa vigente en los ámbitos administrativo, disciplinario o contractual. Adicionalmente, activa los mecanismos de monitoreo y reporte definidos por el DANE a través de lineamientos, instructivos y procedimientos, con el fin de prevenir eventos que comprometan la seguridad de la información, la calidad estadística y la confianza institucional.

3.10 Consecuencias

El uso inadecuado de las herramientas o aplicaciones de inteligencia artificial generativa en el DANE conlleva la aplicación de mecanismos correctivos, disciplinarios, contractuales y de control proporcionales al nivel de riesgo, al impacto generado y a la naturaleza del incumplimiento identificado.

En este contexto, el DANE define las siguientes consecuencias:

- **Suspensión y restricción del uso.** Incluye la suspensión temporal o definitiva del uso del sistema de IA, la restricción de accesos, la desactivación de funcionalidades específicas o la interrupción de procesos automatizados, cuando se identifiquen riesgos que comprometan la calidad de la información, la seguridad de los datos o la confiabilidad de los resultados.
- **Corrección y retractación de resultados**. Cuando el uso de IA haya incidido en la generación o difusión de productos institucionales con errores, inconsistencias o afectaciones a la calidad de la información se deben adelantar las acciones correctivas necesarias para ajustar o retirar dicho contenido, cuando corresponda, y se deberán emitir las aclaraciones o retracciones públicas respecto de los resultados que fueron afectados.
- **Acciones disciplinarias**. La inobservancia de los lineamientos establecidos en la presente política por parte de servidores públicos que genere afectaciones a la integridad o transparencia de información institucional, vulneración de derechos fundamentales (como igualdad, intimidad o hábeas data) o comprometa la ciberseguridad, la reserva de la información, así como cualquier otra conducta que implique un uso irregular de la IA en el ejercicio de la función pública, dará lugar a las investigaciones correspondientes. Esto, bajo la adecuación típica del incumplimiento de deberes, extralimitación de funciones o incursión en prohibiciones consagradas taxativamente en la Ley 1952 de 2019 (Código General Disciplinario) o la norma que la modifique o sustituya.
- **Acciones contractuales**. En el caso de contratistas o terceros, el incumplimiento de las condiciones de uso de la IA dará lugar al estudio para la aplicación de las cláusulas contractuales correspondientes, de acuerdo con el régimen contractual aplicable.
- **Reporte ante instancias de gobierno**. Todo incumplimiento o desviación frente a las disposiciones de la presente política debe ser reportado a las instancias de gobierno definidas por el DANE para

el uso de la IA, a fin de evaluar su impacto, definir medidas adicionales y fortalecer los mecanismos institucionales de control y supervisión.

La aplicación de estas consecuencias debe realizarse garantizando el debido proceso, la trazabilidad de las actuaciones y la articulación con los sistemas de control interno, de manera que se asegure una gestión efectiva de los riesgos asociados al uso de la inteligencia artificial generativa en el DANE.

4. Gestión de datos para el uso de IA

A diferencia de otros sistemas analíticos desarrollados internamente, las herramientas o aplicaciones de inteligencia artificial generativa utilizan instrucciones en lenguaje natural, denominadas prompts, que generan resultados dinámicos cuya lógica interna no es completamente observable ni controlable por el DANE.

En este contexto, la gestión de datos constituye el principal mecanismo de control sobre la interacción entre los usuarios y las herramientas o aplicaciones de inteligencia artificial. En consecuencia, el control institucional no se ejerce sobre el modelo en sí mismo, sino sobre los datos suministrados, las condiciones bajo las cuales se procesan, los resultados obtenidos y la forma en que estos se utilizan en los procesos institucionales. La autorización de herramientas constituye, en este contexto, un control de entrada sobre las condiciones de uso, no un control sobre el modelo.

En este sentido, la gestión de datos en el uso de la IA generativa trasciende la protección de la información y se consolida como un componente esencial del aseguramiento de la calidad estadística y de la gobernanza de las herramientas inteligencia artificial en el DANE. Lo anterior, en coherencia con los principios éticos definidos en la presente política.

Su implementación requiere la aplicación de controles diferenciados definidos en la presente política, en articulación con el marco institucional de gestión de riesgos, sobre los datos de entrada, la relación con proveedores tecnológicos, la validación de resultados, la identificación del contenido generado, la trazabilidad de las interacciones y la gestión de incidentes.

La gestión de datos para el uso de la IA se articula con la Política de Gobierno de Datos de la entidad, en particular con los lineamientos de calidad de los datos, catalogación, metadatos, trazabilidad y linaje, los roles de propietarios y custodios de datos, los activos de información, la interoperabilidad y el intercambio seguro (estándares, API, servicios y catálogos de datos), así como con los dominios y la arquitectura de datos institucional, conforme al instrumento complementario correspondiente.

4.1 Gestión de los prompts

Dentro de las herramientas de inteligencia artificial generativa, el prompt es el principal medio de interacción con el modelo a través del cual se definen su comportamiento, alcance y tipo de respuesta.

Su construcción no se limita al suministro de información, sino que además implica el control de la instrucción que orienta la generación de contenido.

En este contexto, el DANE establece reglas para la construcción de prompts orientadas a mitigar riesgos asociados a la manipulación del modelo, la exposición indebida de información y la generación de resultados no controlados. Las siguientes reglas son de obligatorio cumplimiento:

- Los prompts deben formularse de manera clara, específica y delimitada, evitando ambigüedades o instrucciones abiertas que permitan desviaciones en la respuesta. Incluir el propósito, el contexto y el tipo de salida esperado para obtener resultados controlables y verificables.
- No construir prompts que integren múltiples instrucciones sin jerarquía ni contexto definido, ya que esto incrementa el riesgo de respuestas inconsistentes, incompletas o no alineadas con el objetivo institucional.
- La información incorporada en los prompts debe ser precisa y limitada al propósito para el que fue definida, evitando incluir datos de contexto adicionales que no aporten valor y que incremente el riesgo de exposición innecesaria de información.
- No formular prompts que puedan inducir al sistema a revelar información no solicitada, a reconstruir datos a partir de inferencias o a ampliar el alcance de la respuesta más allá del objetivo definido.
- No utilizar prompts provenientes de textos, documentos o entradas sin validación previa. Se debe mantener el control del propósito del prompt y evitar que instrucciones externas modifiquen su alcance.
- Definir prompts que mantengan coherencia con el objetivo institucional, evitando formulaciones que generen contenido no relevante para el contexto del DANE.
- Cuando las consultas involucren información geográfica o datos geoespaciales, las interacciones con herramientas de IA deben realizarse utilizando únicamente información agregada conforme a los niveles geográficos oficialmente autorizados, evitando el tratamiento de unidades de observación sujetas a reserva estadística.

El usuario de la herramienta es responsable de la formulación del prompt y de los efectos de su uso, por lo que debe velar por que su diseño proporcione resultados controlados, verificables y alineados con los principios de calidad estadística, confidencialidad y uso de la información establecidos por el DANE. Cuando las salidas incorporen datos institucionales, estas deben someterse a las reglas de anonimización y validaciones definidas por el propietario del dato.

4.2 Uso de la IA generativa por proveedores y condiciones contractuales que aplican

El DANE establece que la contratación de proveedores que utilicen servicios de inteligencia artificial generativa como parte de sus obligaciones está sujeta al cumplimiento de condiciones técnicas, jurídicas y operativas que garanticen la protección de la información institucional, la integridad del proceso estadístico y la trazabilidad del uso de la tecnología. Estas condiciones aplican igualmente a las soluciones tecnológicas adquiridas que incorporen, o lleguen a incorporar mediante soporte o actualización, capacidades de inteligencia artificial, incluyendo los aspectos de propiedad intelectual, soberanía de los datos y derechos sobre los modelos entrenados con información de la entidad, conforme al instrumento complementario de adquisición.

En este sentido, todo proveedor que procese información del DANE mediante herramientas o aplicaciones de inteligencia artificial debe limitar su tratamiento exclusivamente a la ejecución del objeto contractual, quedando expresamente prohibido su uso para el entrenamiento, ajuste, mejora o evaluación de modelos. En línea con esta restricción, el proveedor debe asegurar que los datos suministrados no sean almacenados de forma persistente ni utilizados para la construcción de historiales, perfiles o repositorios, así como abstenerse de compartir, transferir, replicar o poner a disposición de terceros la información suministrada por el DANE y los resultados derivados de su procesamiento.

Con el fin de salvaguardar la información es indispensable que el proveedor informe sobre la localización geográfica del almacenamiento y procesamiento de los datos, asegurando la aplicación de la normativa colombiana en materia de protección de datos y reserva estadística. En complemento, el DANE debe conocer los controles técnicos implementados por el proveedor, el modelo de inteligencia artificial utilizado y sus condiciones de operación, con el propósito de verificar la confidencialidad, integridad y disponibilidad de los datos, en el marco de la política de seguridad de la información establecida por la entidad.

El DANE verifica el cumplimiento de las condiciones establecidas mediante la revisión de registros y evidencia documental que permitan evaluar el tratamiento de los datos y el cumplimiento contractual. De igual forma, se prohíbe al proveedor recurrir a terceros para la prestación de servicios relacionados con la IA. Para ello, el DANE puede requerir al proveedor certificaciones, informes técnicos o evidencias adicionales que acrediten la no retención de los datos y el cumplimiento de las condiciones pactadas.

Considerando el riesgo inherente a los sistemas de IA, el proveedor debe aplicar mecanismos para la detección, notificación y gestión de incidentes asociados a filtraciones de información, accesos no

autorizados o comportamientos anómalos del sistema, los cuales deben ser reportados al DANE de manera inmediata o dentro de los plazos contractualmente definidos.

El incumplimiento de estas condiciones facultará a la entidad para dar inicio a las medidas contractuales o legales correspondientes, sin perjuicio de las demás acciones a que haya lugar.

4.3 Validación de salidas generadas por IA

Para el DANE, todos los resultados generados por herramientas de inteligencia artificial generativa se consideran datos no confiables hasta que superen un proceso de validación técnica, el cual se fundamenta en los controles y en la revisión humana definidos en la presente política como mecanismos de verificación, orientados a asegurar la calidad, la consistencia y la coherencia exigidas por la entidad. Esta validación aplica tanto a los usos asociados a la producción estadística como a los usos administrativos, de apoyo, de comunicación y de elaboración de documentos institucionales.

En este sentido, la validación de salidas exige comprobar que el resultado generado mantenga coherencia evitando contradicciones, ambigüedades o conclusiones no sustentadas en la información presentada. En complemento, se requiere contrastar los resultados con fuentes oficiales, series estadísticas, marcos conceptuales o documentos técnicos, a fin de verificar su alineación con los criterios definidos por el DANE.

Los datos, cifras, clasificaciones o descripciones técnicas deben verificarse contra fuentes confiables para asegurar su exactitud y precisión, así como su correcta interpretación. A su vez, el contenido generado debe ser coherente con los diseños metodológicos, las definiciones conceptuales y los supuestos estadísticos aplicables de acuerdo con el marco técnico de la operación estadística.

La validación también comprende la identificación de posibles sesgos, la detección de contenido no verificable y de alucinaciones del sistema, así como la evaluación de su impacto en la calidad y la confiabilidad de la información. Este análisis se realiza en función del contexto institucional en el que se utilizará la salida, asegurando el nivel de formalidad requerido.

Todo resultado con componente geoespacial debe ser consistente con el Marco Geoestadístico Nacional y con los estándares adoptados por la Dirección de Geoestadística (DIG). Su validación comprende, además de las verificaciones generales definidas en la presente política, la precisión posicional y temática, la consistencia lógica, la completitud y la exactitud geométrica (topología) de la información geoespacial. El alcance de los productos geoespaciales generados mediante herramientas de

inteligencia artificial debe documentarse y delimitarse, especificando los parámetros técnicos aplicables, tales como la escala de trabajo, la vigencia y calidad de las fuentes de información, los datos de entrenamiento y las métricas de precisión.

Con base en lo anterior, un resultado generado por inteligencia artificial solo podrá considerarse apto para su uso en el DANE cuando cumpla de manera integral con las condiciones de validación definidas y esté debidamente justificado y documentado. En caso contrario, cuando se identifiquen inconsistencias, errores, sesgos o falta de alineación con los referentes técnicos, la salida no puede utilizarse sin aplicar ajustes, realizar validaciones adicionales o reemplazarla por fuentes verificadas.

Finalmente, el proceso de validación debe estar respaldado por registros que evidencien las verificaciones realizadas, los criterios aplicados y las decisiones adoptadas, con el fin de garantizar la trazabilidad, el control y monitoreo del uso de herramientas de inteligencia artificial en el DANE.

4.4 Transparencia y declaración de uso de inteligencia artificial

Dentro de la política de uso de inteligencia artificial generativa del DANE, se establece la exigencia de identificar explícitamente el sistema utilizado y su contribución a la generación de productos institucionales, así como la obligación de documentar su uso dentro del proceso técnico que da origen a los resultados. En este sentido, se requiere elaborar una declaración que especifique el propósito de su utilización, el tipo de tarea realizada y la herramienta empleada, la cual debe incorporarse en el contenido del documento o en sus metadatos, de acuerdo con la naturaleza del producto y los lineamientos derivados de esta política. Los mapas, las visualizaciones territoriales y los productos cartográficos o geoespaciales elaborados total o parcialmente con IA deben declarar explícitamente su uso, los procesos de validación realizados y las aclaraciones o especificaciones pertinentes.

La trazabilidad del uso de la inteligencia artificial permite reconstruir, en la medida de lo posible, las condiciones bajo las cuales se generó el contenido, incluyendo los prompts utilizados, las herramientas empleadas y las transformaciones realizadas, con el fin de respaldar la reproducibilidad, el control institucional y la evaluación de la calidad de los resultados.

Esta declaración no sustituye los procesos de validación técnica ni modifica las responsabilidades del personal experto del DANE respecto de la calidad, coherencia y consistencia de los productos obtenidos. En todos los casos, el contenido generado y utilizado sigue siendo responsabilidad del DANE y su incorporación debe cumplir con los estándares definidos en la presente política.

Omitir la declaración del uso de inteligencia artificial en los casos en que esta sea requerida constituye un incumplimiento de los principios de transparencia y trazabilidad establecidos por la entidad y dará lugar a la activación de los mecanismos de control definidos en este documento y por normativa institucional vigente.

4.5 Gestión de sesgos e incidentes en el uso de IA generativa

Para el DANE, el uso de herramientas o aplicaciones de inteligencia artificial generativa implica la exposición a riesgos asociados a sesgos algorítmicos, errores y eventos operacionales, que pueden afectar la calidad, la imparcialidad técnica, la confiabilidad de los productos y el reconocimiento institucional de la entidad. En este contexto, las desviaciones en los resultados generados por sistemas de IA deben formar parte de los procesos de validación técnica y de monitoreo continuo, especialmente en productos de naturaleza estadística o analítica. Esta gestión comprende tanto las interacciones simples de instrucción y resultado como los usos complejos basados en agentes o en procesos automatizados de múltiples pasos.

Para ello, se establece que la gestión de sesgos e incidentes incluya la aplicación de las siguientes reglas:

1) Realizar la identificación de sesgos e incidentes durante la validación técnica de los resultados y el monitoreo continuo del uso de herramientas de inteligencia artificial generativa.
2) Determinar y analizar las posibles causas del sesgo asociadas a los datos de entrada, las instrucciones (prompts) o el modelo utilizado.
3) Aplicar medidas de mitigación proporcionales al nivel de impacto identificado, incluyendo ajustes en los prompts, refuerzo de las validaciones humanas o restricción del uso del sistema.
4) Registrar y reportar oportunamente los sesgos e incidentes identificados a través de los canales institucionales definidos para la gestión de incidentes.
5) Definir acciones correctivas y preventivas orientadas a evitar la recurrencia de los incidentes identificados.
6) Documentar las acciones realizadas durante la gestión de los sesgos e incidentes reportados.
7) Suspender temporalmente y de manera preventiva el uso del sistema de inteligencia artificial generativa cuando se identifiquen incidentes críticos.

El DANE establece que la gestión de sesgos e incidentes en el uso de la inteligencia artificial generativa debe articularse con el Modelo Integrado de Planeación y Gestión, el Sistema de Seguridad de la

Información y el marco institucional de gestión de riesgos, asegurando la mejora continua y la reducción progresiva de la exposición a riesgos algorítmicos. Para ello, la entidad dispondrá de un mecanismo institucional para que los servidores reporten dudas, incidentes o riesgos identificados durante el uso de herramientas de inteligencia artificial.

5. Gobierno y seguridad en el uso de IA

La incorporación de la inteligencia artificial en la producción estadística fortalece el compromiso institucional del DANE en la generación de resultados oportunos y de calidad, en línea con el rigor técnico, la integridad y la trazabilidad que caracterizan todas las etapas del proceso estadístico. Por ello, las innovaciones aplicadas a los procesos de identificación, diseño, construcción, recolección, procesamiento, análisis, difusión de información y evaluación deben enmarcarse en un modelo de gobernanza que asegure condiciones de control, seguridad y confiabilidad, a la vez que promueva entornos de exploración responsables al interior de la entidad.

En este sentido, el uso de la inteligencia artificial dentro de la entidad requiere la adopción de un modelo de gobierno basado en criterios éticos, de responsabilidad y de control institucional, que permita gestionar su incorporación en los procesos misionales, estratégicos, de apoyo y de evaluación, sin afectar los principios rectores de la producción estadística oficial. Este modelo tiene como propósito avalar y autorizar el uso adecuado de las herramientas de IA generativa dentro del DANE, con el fin de estandarizar los marcos de aplicación de las soluciones desarrolladas, a la vez que se previene el uso indebido y potencialmente peligroso para la entidad. Este modelo se articula, además, con el Modelo Integrado de Planeación y Gestión, la gobernanza del Gobierno Digital, la gobernanza de seguridad de la información, el gobierno de datos, la Arquitectura Empresarial y el proceso de Gestión del Conocimiento y la Innovación. El componente ético del modelo se articula con el Marco Ético para la IA en Colombia, el marco ético de las operaciones estadísticas y el Sistema de Ética Estadística de la entidad.

Para tal efecto, se adopta un esquema de gobierno de carácter federado, que reconoce la autonomía operativa de las dependencias del DANE para la exploración y aplicación de soluciones basadas en inteligencia artificial, bajo lineamientos comunes que aseguren la integridad del proceso estadístico, el cumplimiento normativo y la gestión adecuada de los riesgos asociados. De esta manera, la implementación de la inteligencia artificial generativa no solo se controla y supervisa, sino que también se orienta estratégicamente para fortalecer la cadena de valor del DANE y potenciar sus capacidades institucionales en la producción de información estadística oficial. En todo caso, los proyectos y casos de uso de inteligencia artificial deben ser consistentes con la Arquitectura Empresarial institucional. Este esquema se adopta por cuanto equilibra la autonomía de las dependencias para explorar soluciones con lineamientos comunes de obligatorio cumplimiento; dicha autonomía conlleva los deberes de registro, documentación, gestión de riesgos y reporte definidos en la presente política.

5.1 Modelo de Gobierno del uso de la IA generativa

El modelo de gobierno de la inteligencia artificial generativa se fundamenta en los pilares de confianza, calidad e impacto, que orientan la adopción, el uso y la supervisión de estas tecnologías en la entidad, asegurando su alineación con la misión institucional y con los principios de la producción estadística oficial. Estos pilares concretan, en el modelo de gobierno, los principios definidos en la presente política.

Pilar 1. Confianza

Entendida como la capacidad del DANE para ejercer control sobre el uso de la IA, asegurando su aplicación segura, transparente y responsable, promoviendo la protección de la información, el cumplimiento de la reserva estadística, la gestión oportuna de incidentes y la trazabilidad de los procesos.

Esta capacidad se evidencia cuando la entidad conoce, de manera amplia y suficiente, cómo y dónde se emplea la inteligencia artificial, sin dejar lugar a supuestos ni omisiones, enmarcando su actuar en las restricciones y validaciones que le apliquen y con la seguridad de verificar, con evidencia, los resultados obtenidos. Para que la evidencia sea medible, se definen indicadores orientados a evaluar el grado de control institucional, basados en registros institucionales de los sistemas de IA, reportes de uso, gestión de incidentes, tiempos de respuesta y en la existencia de trazabilidad documentada.

En síntesis, este pilar está enfocado en hacer visible y auditable la IA en los procesos del DANE, como base para la credibilidad institucional de los resultados generados.

Pilar 2. Calidad

Por su parte, el pilar de calidad representa la capacidad institucional del DANE para integrar el uso de la IA en el proceso estadístico sin comprometer su integridad metodológica, salvaguardando la solidez de las validaciones técnicas y la preservación de la calidad de la información en términos de exactitud, consistencia, coherencia, comparabilidad, oportunidad, accesibilidad, relevancia e integridad. En este sentido, la calidad incluye, además, el control del comportamiento de los sistemas de IA y la verificación de sus salidas en línea con los marcos conceptuales de la entidad.

La calidad se materializa cuando los resultados generados por la IA son verificables, reproducibles y técnicamente sustentados y cuando su uso no introduce sesgos, inconsistencias ni alteraciones en las operaciones estadísticas.

Puntualmente, la medición de este pilar se realiza mediante la validación documentada de resultados, la formalización de conceptos técnicos, la evaluación del impacto algorítmico y el seguimiento de métricas de desempeño, como mecanismos que evidencian el aseguramiento de la calidad en el uso de la IA.

Pilar 3. Impacto

El pilar de impacto orienta el uso de las herramientas de inteligencia artificial generativa hacia la obtención de resultados concretos que generen valor público, fortalezcan la eficiencia institucional y contribuyan de manera efectiva al cumplimiento de los objetivos estratégicos del DANE, en coherencia con su misión de producir información estadística oficial de calidad.

El impacto reconoce que el uso de la inteligencia artificial se traduce en beneficios aplicables y sostenibles para el DANE, garantizando que su implementación responda a propósitos claramente definidos desde su fase de formulación y contribuya de manera efectiva a la mejora del desempeño organizacional, evitando su uso en escenarios exploratorios sin un propósito documentado. Esta condición no impide la experimentación controlada ni los pilotos con fines definidos de aprendizaje o innovación, en armonía con el CONPES 4144 de 2025. Para ello, se mide el impacto a través de su aporte a la optimización de tiempos, la reducción de costos, la mejora de la productividad, la promoción de los productos estadísticos, la contribución a la transformación de procesos institucionales y el fortalecimiento de la capacidad de las dependencias para apropiarse de las soluciones implementadas.

Específicamente, es en este pilar donde el uso de la IA en el DANE cobra sentido al traducir la tecnología en beneficios tangibles que contribuyen de manera efectiva a los procesos y a la mejora institucional.

Ilustración 1. Pilares del modelo de gobierno de la IA en el DANE

Creación propia con uso de IA

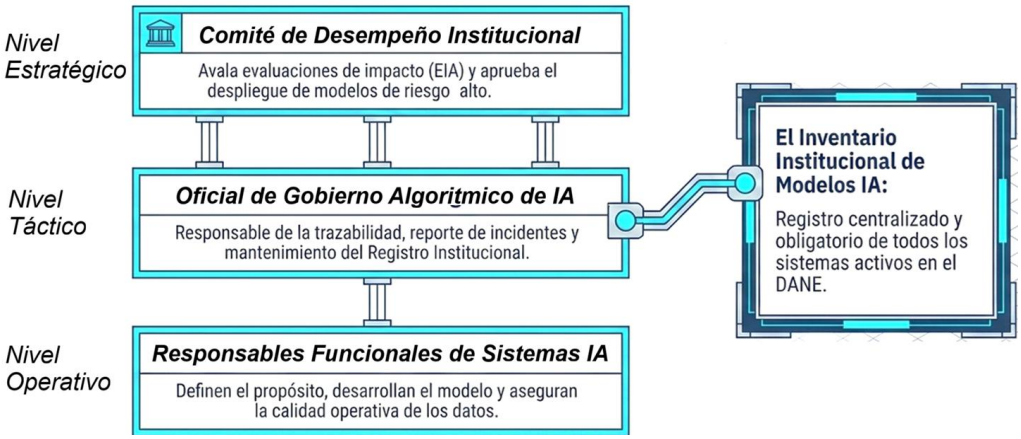
En conjunto, estos pilares garantizan que el uso de la inteligencia artificial no solo responda a criterios de innovación tecnológica, sino que se desarrolle bajo un enfoque de responsabilidad institucional, rigor técnico y generación de valor para la toma de decisiones basada en información estadística confiable.

5.2 Estructura y roles de gobernanza para el uso de la IA

La gobernanza del uso de la inteligencia artificial en el DANE se estructura bajo un enfoque federado que combina la autonomía operativa de las dependencias con mecanismos de coordinación y supervisión estratégica, asegurando la alineación con los principios de confianza, calidad e impacto definidos en la presente política, en el marco del Modelo Integrado de Planeación y Gestión (MIPG), como parte de los mecanismos institucionales de gestión, control interno y mejora continua.

En este sentido, se define una gobernanza jerarquizada en los niveles estratégico, táctico y operativo, que proporciona un control central estratégico y una ejecución distribuida, escalables según el nivel de riesgo asociado al uso de la IA. En términos generales, el nivel operativo identifica y registra los casos de uso, el nivel táctico los clasifica, acompaña y verifica, y el nivel estratégico decide sobre los de alto riesgo, conforme se desarrolla a continuación.

Ilustración 2. Instancias de gobierno de uso de IA en el DANE



Creación propia con uso de IA

Nivel estratégico

Dentro de este nivel se orienta, evalúa y hace seguimiento al uso de la inteligencia artificial en la entidad, tomando decisiones sobre las iniciativas de alto impacto o de alto riesgo para el DANE, con el fin de asegurar su alineación con los objetivos institucionales y con los pilares de confianza, calidad e impacto definidos en la presente política.

Para tal efecto, se establece como órgano responsable al Comité Institucional de Gestión y Desempeño, el cual asume el rol de instancia de gobierno de la inteligencia artificial en el marco de las funciones definidas en la Resolución 1319 de 2018, particularmente en lo relacionado con la recomendación y aprobación de políticas, la evaluación del desempeño institucional y el seguimiento de la implementación de estrategias y planes.

Sin perjuicio de lo indicado en la resolución, son responsabilidades del Comité en temas de IA:

1) Dirigir la implementación de políticas, estrategias y lineamientos relativos al uso de la inteligencia artificial.
2) Evaluar el desempeño del modelo de gobierno de IA en el DANE.
3) Tomar decisiones sobre iniciativas de uso de la IA de alto impacto o de alto riesgo para la entidad.
4) Definir lineamientos para fortalecer el modelo de gobernanza de la IA.
5) Promover el uso ético y responsable de la IA dentro de la entidad.

Conforme a los mecanismos de convocatoria, deliberación y aprobación establecidos en la Resolución 1319 de 2018, las decisiones relacionadas con el uso de la inteligencia artificial se adoptan en el marco de las sesiones del Comité. La Secretaría Técnica del Comité Institucional de Gestión y Desempeño es responsable de incorporar los temas de inteligencia artificial en la agenda institucional, de consolidar la información requerida para la presentación de iniciativas y de custodiar la documentación correspondiente, en concordancia con sus funciones.

Nivel táctico

El nivel táctico es la instancia encargada de articular la implementación del modelo de gobernanza de la inteligencia artificial, garantizando su aplicación consistente en la entidad y a la vez, es responsable de la caracterización y el registro de los casos de uso de inteligencia artificial en el inventario institucional.

A partir de este registro se determina el nivel de riesgo de los casos de uso y se define su tratamiento dentro del modelo jerárquico de la siguiente manera:

- Los casos de riesgo bajo son gestionados directamente por las dependencias, cumpliendo los lineamientos institucionales y con registro obligatorio en el inventario institucional de sistemas de IA. Se abordan en el nivel operativo, con seguimiento en el nivel táctico

- Los casos de riesgo medio incluyen obligatoriamente el acompañamiento a nivel táctico, con seguimiento de su implementación y verificación del cumplimiento de los criterios de calidad y control. El nivel táctico realiza seguimiento a la matriz de riesgos evitando su materialización

- Los casos de alto riesgo se escalan al Comité Institucional de Gestión y Desempeño para su evaluación y decisión. Las iniciativas deben ser puestas en consideración del Comité Institucional de Gestión y Desempeño, para su evaluación y eventual recomendación de aprobación, con base en criterios de riesgo, impacto, calidad metodológica y alineación estratégica.

Este nivel está a cargo de la Oficina Asesora de Planeación, en su calidad de instancia articuladora del Modelo Integrado de Planeación y Gestión y de apoyo técnico al Comité Institucional de Gestión y

Desempeño, actuando como puente entre los niveles estratégico y operativo. En cuanto a sus responsabilidades le corresponde:

1) Administrar y mantener actualizado el Inventario Institucional de Sistemas de IA, como instrumento único de control., el cual registra, como mínimo: el modelo y su versión, el responsable, el proveedor, el nivel de riesgo, los datos utilizados, los controles aplicados y la fecha de aprobación.
2) Gestionar la trazabilidad del uso de la inteligencia artificial.
3) Monitorear los indicadores de los pilares de confianza, calidad e impacto del uso de la IA.
4) Hacer seguimiento de los riesgos e incidentes derivados del uso de la IA.
5) Proporcionar acompañamiento técnico a las dependencias en la implementación de los lineamientos sobre el uso de la IA.
6) Promover el registro institucional de los sistemas de IA como el único instrumento de control.
7) Hacer seguimiento de la implementación de las estrategias y los planes relacionados con la IA.

Este nivel no adopta decisiones sobre la aprobación de iniciativas, pero sí verifica el cumplimiento de los lineamientos definidos en la presente política y emite concepto para su continuidad, ajuste o escalamiento, según corresponda. Cuando los casos de uso involucren información geográfica, cartográfica, geocodificación, imágenes satelitales o análisis espacial, la Dirección de Geoestadística (DIG) participa como instancia técnica consultiva.

Nivel operativo

El nivel operativo está integrado por todas las dependencias del DANE que utilizan herramientas o aplicaciones de inteligencia artificial generativa en el marco de los procesos institucionales y constituye el punto de control primario de su uso, dado que en este nivel se originan los datos, se generan los resultados y se materializan los riesgos asociados a su aplicación.

Por ello, la responsabilidad sobre el uso de sistemas de IA es compartida entre los equipos que aplican estas tecnologías, quienes responden por su adecuada implementación, documentación y validación técnica y los jefes de área, quienes asumen la responsabilidad sobre su aplicación, asegurando el cumplimiento de los lineamientos definidos y la calidad de los resultados generados.

En este sentido, es responsabilidad de las dependencias:

1) Definir el propósito, alcance y resultados esperados de los casos de uso de inteligencia artificial, alineados con los procesos institucionales.
2) Registrar los casos de uso en el Inventario Institucional de Sistemas de IA, como requisito para su implementación y seguimiento.
3) Identificar y documentar los riesgos asociados al uso de la inteligencia artificial, especialmente los relacionados con la calidad de la información, la seguridad y la interpretación adecuada de los resultados.
4) Garantizar la calidad, consistencia y trazabilidad de los datos utilizados y de los resultados generados, asegurando su validación técnica y metodológica.
5) Aplicar los lineamientos definidos en materia de seguridad de la información, protección de datos y reserva estadística en todos los usos de la inteligencia artificial.
6) Documentar el uso de la inteligencia artificial en los procesos institucionales, incluyendo las limitaciones y las validaciones realizadas.
7) Reportar incidentes, desviaciones o riesgos materializados generados por el uso de la inteligencia artificial, conforme a los mecanismos institucionales definidos.
8) Atender las recomendaciones y los requerimientos emitidos por los niveles táctico y estratégico.

El nivel operativo gestiona directamente los casos de uso de riesgo bajo, conforme a los lineamientos institucionales y con seguimiento por parte del nivel táctico. En los casos de riesgo medio y alto, debe proporcionar la información, documentación y soporte técnico requeridos para su análisis, validación o escalamiento dentro del modelo de gobernanza.

5.3 Seguridad de la información en el uso de IA

Dentro del modelo de gobernanza de la inteligencia artificial, la seguridad de la información se reconoce como un control obligatorio aplicable a los niveles estratégico, táctico y operativo, así como a todos los casos que utilicen inteligencia artificial, independientemente de su nivel de riesgo.

En este contexto, la presente política no define disposiciones independientes en materia de seguridad, sino que adopta y aplica los lineamientos institucionales vigentes, reconociendo su carácter transversal en los procesos del DANE. Los controles específicos frente a amenazas propias de los sistemas de IA generativa —tales como la inyección de instrucciones (prompt injection), el envenenamiento de datos o del modelo (data/model poisoning), la extracción del modelo (model extraction), la inferencia de pertenencia (membership inference), el uso de herramientas no autorizadas (shadow AI) y la deriva del

modelo (model drift)— se desarrollan en el marco del Modelo de Seguridad y Privacidad de la Información y de los lineamientos complementarios de la presente política.

Su incorporación al modelo de gobernanza se materializa mediante el cumplimiento de:

- **Registro**: Todo caso de uso que implemente sistemas de IA generativa deben identificar los riesgos para la seguridad de la información como requisito para su inclusión en el inventario institucional de sistemas de IA.
- **Evaluación de impacto (EIA):** Los casos de uso sujetos a evaluación deben incluir la valoración de los riesgos para la seguridad de la información, conforme a los lineamientos definidos en el Modelo de Seguridad y Privacidad de la Información. La ausencia de esta evaluación impide su continuidad y su aprobación.
- **Análisis técnico (nivel táctico)**: El nivel táctico verifica el cumplimiento de los lineamientos de seguridad de la información definidos por la entidad frente a los casos de uso de inteligencia artificial, como condición para su continuidad en el proceso y, cuando aplique, su escalamiento a instancias de decisión
- **Decisión (nivel estratégico)**: En los casos de uso de alto riesgo, el Comité Institucional de Gestión y Desempeño incorpora el cumplimiento de los criterios de seguridad como elemento determinante para la evaluación y toma de decisiones sobre la aplicación de la IA.
- **Control de acceso**: El uso de servicios de inteligencia artificial debe realizarse conforme a los mecanismos institucionales de autenticación y control de acceso definidos por la entidad, garantizando la trazabilidad del uso y la gestión adecuada de las credenciales, en concordancia con la política de seguridad de la información. En los casos en que la entidad suministre herramientas de inteligencia artificial licenciadas, administrará las credenciales de ingreso para asegurar su uso o la ejecución de tareas específicas.

La seguridad de la información constituye un elemento habilitador del pilar de confianza, asegurando que el uso de la inteligencia artificial se realice bajo condiciones de control institucional, protección de la información y cumplimiento de la reserva estadística.

5.4 Fortalecimiento de capacidades Institucionales en uso de IA

La definición de la política de uso responsable de la inteligencia artificial no garantiza por sí misma su aplicación efectiva al interior del DANE; por ello, la entidad, en articulación con el Plan Institucional de Capacitación, establece el fortalecimiento de capacidades institucionales para el uso de la inteligencia artificial como un componente fundamental para consolidar la madurez en su adopción, promoviendo un enfoque consciente, ético y responsable que permita a quienes la emplean comprender sus alcances y actuar en consecuencia en el ejercicio de sus funciones.

Para lograrlo, la entidad adopta un enfoque diferenciado de desarrollo de capacidades, que reconoce distintos niveles de apropiación de la inteligencia artificial:

- **Nivel general:** orientado a asegurar que todos los servidores públicos y contratistas comprendan los principios, riesgos, restricciones y responsabilidades asociados al uso de la inteligencia artificial, incluyendo la identificación de usos permitidos y no permitidos, la protección de la información y la necesidad de validación de resultados. Este nivel busca prevenir el uso inadecuado de estas tecnologías, evitando la incorporación de resultados no verificados en productos institucionales o la exposición de información sensible.
- **Nivel especializado:** enfocado en quienes aplican la inteligencia artificial en los procesos misionales, asegurando su capacidad para utilizar estas herramientas conforme a los estándares técnicos y metodológicos del DANE. Este nivel incluye la comprensión de las limitaciones de los sistemas de IA, la validación de resultados generados, la documentación de su uso y la integración de estos en el proceso estadístico sin afectar la calidad, consistencia y trazabilidad de la información.
- **Nivel avanzado:** dirigido a roles con responsabilidad en la validación, supervisión y toma de decisiones sobre el uso de la inteligencia artificial, que deben contar con capacidades para evaluar la conveniencia de los casos de uso, analizar sus riesgos, interpretar sus resultados en contextos complejos y asegurar su alineación con los principios metodológicos, técnicos y éticos de la entidad. Este nivel implica la capacidad de establecer criterios de uso, validar resultados críticos y orientar la toma de decisiones institucionales frente a la incorporación de estas tecnologías.

Este enfoque busca asegurar la apropiación institucional del uso de la IA generativa, reducir los riesgos derivados de su uso inadecuado y fortalecer la capacidad del DANE para integrar la inteligencia artificial en el proceso estadístico sin comprometer la calidad, la seguridad de la información ni la confianza en los resultados. El desarrollo de capacidades incluye competencias en GeoIA para la analítica de datos

geoespaciales y la automatización de procesos, con capacitación continua sobre el uso de herramientas de inteligencia artificial asociadas a información geoespacial. Asimismo, la entidad promoverá acciones de apropiación social del conocimiento, orientadas a que los usuarios de la información y la ciudadanía comprendan el uso de la inteligencia artificial en la producción estadística oficial.

5.5 Implementación de la política

La implementación de la presente política se realizará de manera gradual y planificada, orientada por los instrumentos complementarios que la desarrollan (lineamientos, guías, procedimientos e instructivos), por el plan de fortalecimiento de capacidades y por los indicadores de seguimiento definidos en el modelo de gobierno. La Oficina de Sistemas coordinará su adopción, definirá el cronograma de expedición de los instrumentos complementarios y reportará el avance de la implementación a las instancias de gobierno correspondientes.

Anexo. Nota Metodológica. Declaración de uso de IA

Durante la elaboración de la presente política, se emplearon herramientas de IA generativa como apoyo para las actividades de redacción, revisión de estilo, síntesis de información, consulta exploratoria de fuentes y generación de gráficos. El uso de ChatGPT (OpenAI), Claude (Anthropic), Grammarly y NotebookLM (Google) no sustituyó el análisis técnico y metodológico requerido para la elaboración del documento.

A su vez, todos los contenidos, definiciones, lineamientos, referencias y demás elementos incorporados en esta política fueron revisados, contrastados, validados y adaptados por el equipo responsable de su elaboración, el cual asume la responsabilidad integral sobre el contenido final, su consistencia técnica y su conformidad con el marco normativo e institucional aplicable.

Tabla 1. Anexo - Controles para el uso de la IA generativa en el DANE

Control	Indicador de cumplimiento	Evidencia mínima requerida	Frecuencia de verificación	Responsable de la verificación
Elaboración de documentos con IA	Porcentaje de documentos institucionales que declaran el uso de IA cuando aplicad	Documento con referencia al uso de IA, historial de versiones, notas metodológicas o bibliografía	En cada documento generado o actualizado	Responsable del documento
Uso de IA como insumo en procesos misionales	Porcentaje de resultados generados con IA que cuentan con validación técnica documentada	Informes de validación, notas técnicas, registros de verificación con referentes técnicos	Por fase del proceso estadístico	Responsable del proceso
Validación técnica y metodológica de sistemas de IA	Porcentaje de informes o matrices de evaluación del desempeño del sistema de IA	Informes técnicos, fichas metodológicas, documentación del modelo	Previo a la implementación en la operación estadística	Responsable de la operación estadística
Supervisión humana	Porcentaje de sistemas de IA con responsable técnico asignado y evidencia de validación	Actas de aprobación, conceptos técnicos, registros de seguimiento, correos institucionales	Trimestral o con mayor frecuencia según la criticidad del sistema	Jefe del área o director técnico
Herramientas, acceso y uso institucional	Cumplimiento de las herramientas autorizadas y condiciones de uso definidas por el DANE	Registros de acceso, reportes de monitoreo, configuraciones de seguridad, listados de herramientas autorizadas	Trimestral	Jefe de la Oficina de Sistemas

Bibliografía

Congreso de la República de Colombia. (17 de Octubre de 2012). *Ley 1581 de 2012*. Obtenido de https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=49981

Congreso de la República de Colombia. (03 de 10 de 2023). *Ley 2335 de 2023*. Obtenido de https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=221910

Constitucional, C. (26 de Febrero de 2025). *Sentencia T-067 de 2025* . Obtenido de Relatoria: Sentencia T-067 de 2025

Departamento Nacional de Planeación. (14 de Febrero de 2025). *Documento CONPES 4144*. Obtenido de https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/4144.pdf

ISO & IEC. (2022). *ISO/IEC 22989:2022 Artificial intelligence — Concepts and terminology*. Obtenido de https://www.iso.org/obp/ui/es/#iso:std:iso-iec:22989:ed-1:v1:en

NIST. (Marzo de 2022). *NIST Special Publication 1270*. Obtenido de https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1270.pdf

NIST. (2024). *Generative Artificial Intelligence Profile (NIST AI 600-1)*. Obtenido de National Institute of Standards and Technology: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf

NIST. (2026). *Glossary*. Obtenido de https://csrc.nist.gov/glossary/term/generative_artificial_intelligence

OCDE. (2024). *Recomendación OCDE sobre IA (2019, actualizada 2024)*. Obtenido de Instrumentos Legales: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

Presidencia de la República de Colombia. (2021). *Marco Ético para la IA en Colombia*. Obtenido de Gobierno de Colombia: https://minciencias.gov.co/sites/default/files/marco-etico-ia-colombia-2021.pdf

Procuraduría General de la Nación, Defensoría del Pueblo. (30 de Septiembre de 2025). *DIRECTIVA CONJUNTA No. 007 de 2025*. Obtenido de https://www.defensoria.gov.co/documents/20123/3407303/300925DirectivaConjunta007.pdf

Pública, F. (s.f.). *Ley 1581 de 2012*. Obtenido de Gestor Normativo: https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=49981

UNESCO. (2022). *Recomendación sobre la ética de la inteligencia artificial*. Obtenido de Biblioteca Digital : https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa