

 <b>Universidad de Santander</b> <b>UDES</b> VIGILADA MINEDUCACIÓN   SNIES 2832	<b>Vicerrectoría de Investigaciones</b>	
	<b>TÉRMINOS DE REFERENCIA CONVOCATORIAS</b> <b>PROYECTOS DE INVESTIGACIÓN</b> <b>VII -FT-027-UDES</b>	Versión:02

## PRESENTACIÓN DE LA CONVOCATORIA

### 1. GENERALIDADES DE LA CONVOCATORIA

<b>Nombre de la convocatoria</b>	Convocatoria Interna para el Diseño y Publicación de Conjunto de Datos con Potencial de Explotación en Inteligencia Artificial 2025 - 2026. Ejecución 2026
<b>Tipo de convocatoria</b>	INTERNA
<b>Página de consulta</b>	<a href="https://udes.edu.co/investigacion/convocatorias">https://udes.edu.co/investigacion/convocatorias</a>
<b>Presentación de la convocatoria</b>	<p>La Universidad de Santander - UDES promueve el fortalecimiento de la investigación como una función sustantiva clave en la formación de profesionales con pensamiento crítico con capacidad para aportar a la solución de problemáticas de contexto local, regionales y globales. Para lo cual la institución impulsa convocatorias internas de investigación, dirigidas a profesores investigadores, para presentar proyectos que ofrezcan alternativas que contribuyan a la solución de problemas en los focos priorizados internamente, a los objetivos de desarrollo sostenible y a su vez permitan el logro de productos resultados de investigación - nuevo conocimiento, formación de talento humano, desarrollo tecnológico, divulgación y apropiación del conocimiento.</p> <p>Desde el año 2022, la Universidad ha venido fortaleciendo sus capacidades en ciencia de datos y ciencia abierta, a través de iniciativas orientadas a la formación de talento humano, la consolidación de repositorios de datos institucionales y el impulso de prácticas responsables en el ciclo de vida de los datos. Esto se hace en consonancia con las políticas nacionales de ciencia, tecnología e innovación en Colombia y los marcos internacionales que promueven la integridad científica, la transparencia, y el cumplimiento de los principios FAIR (por sus siglas en inglés: Findable (Encontrable), Accessible (accesible), Interoperable (interoperable), Reusable (reutilizable)).</p> <p>Los principios FAIR son fundamentales para asegurar que los datos generados en contextos de investigación no sólo sean reutilizables por otros investigadores, sino también accesibles y comprensibles por diversos sectores de la sociedad. Su aplicación permite que los conjuntos de datos sean integrados en plataformas digitales, replicados, utilizados en nuevos modelos de IA, y contribuyan al crecimiento del conocimiento científico global.</p> <p>En el marco del fortalecimiento institucional, la “Convocatoria Interna para el diseño y publicación de conjuntos de datos con potencial de explotación en inteligencia artificial”, tiene como propósito fomentar el desarrollo de proyectos que promuevan la organización, curaduría, anotación (data annotation) y documentación de conjuntos de datos (Datasets) de alta calidad. Es la primera convocatoria ofrecida para promover la publicación de datos de calidad con potencial de uso en proyectos en inteligencia artificial, considerándolos como un insumo destinado a facilitar el desarrollo futuro de modelos de Inteligencia Artificial (IA) o el desarrollo de soluciones</p>



	<p>basadas en IA que respondan a problemáticas específicas en sectores como la salud, la educación, el medio ambiente, la agricultura, entre otros.</p> <p>En este contexto, se reconoce que la estructuración de conjuntos de datos con potencial para aplicaciones en IA constituye una fase crítica dentro del proceso de maduración tecnológica y científica, especialmente cuando estos datos permiten el diseño, entrenamiento, validación o especialización (fine-tuning) de modelos de IA. Por lo anterior la convocatoria contempla, además, proyectos que, adicional a la organización y documentación de datos de calidad, incorporen una proyección hacia o desde la innovación o el desarrollo tecnológico. Esto significa la aplicación de conocimientos previos o resultados de investigación para el diseño de algoritmos, sistemas, prototipos funcionales o modelos iniciales que sirvan como base para futuras soluciones basadas en IA. Para efectos de esta convocatoria, se entiende por desarrollo tecnológico el proceso que conlleva a la materialización de productos, los cuales pueden ser de tipo: esquemas, diseños o prototipos no comercializables, incluyendo proyectos piloto o demostrativos que, sin constituir aún productos industriales, evidencien avances sustanciales en la gestión y disponibilidad de datos aplicables a sistemas inteligentes.</p> <p>Así mismo la convocatoria le da relevancia a la formación de talento humano en pregrado o posgrado, promoviendo competencias en ciencia de datos, ética de la información y pensamiento sistémico. Así mismo la articulación con semilleros de investigación, el trabajo en red con otros campus y con aliados externos.</p> <p>En este sentido, la UDES se compromete a promover una cultura institucional que incentive la curaduría, anotación, normalización, preservación y apertura de datos, como un artefacto tecnológico que conlleva a una práctica ética que potencie la apertura del conocimiento. Esto resulta clave en la formación investigativa y profesional de los estudiantes, quienes deben comprender el valor social de los datos abiertos y su papel en el desarrollo de tecnologías inclusivas, sostenibles y responsables.</p> <p>Esta convocatoria, representa una oportunidad para consolidar prácticas de ciencia abierta dentro de la Universidad, fortalecer el ecosistema de investigación institucional y avanzar hacia una mayor visibilidad y colaboración científica desde y para los territorios.</p>
<b>Objetivo</b>	Fortalecer procesos de organización, enriquecimiento semántico, estructuración, anotación, documentación y publicación de conjuntos de datos, garantizando su alineación con los principios FAIR y su preparación para aplicaciones en Inteligencia Artificial, en un periodo máximo de 12 meses, en proyectos provenientes de cualquier área del conocimiento.
<b>Dirigido a</b>	La convocatoria está dirigida a profesores investigadores tiempo completo o medio tiempo vinculados a grupos de Investigación avalados por la Universidad de Santander.
<b>Requisitos</b>	<p><b>Requisitos del alcance y estructuración de las propuestas</b></p> <ul style="list-style-type: none"> <li>Las propuestas deberán al menos contemplar un proceso de organización, enriquecimiento semántico, estructuración, anotación, documentación y publicación de un conjunto de datos y justificar en su estructuración el potencial uso de estos para el entrenamiento, validación o evaluación de modelos de Inteligencia Artificial, orientados a escenarios de investigación aplicada</li> </ul>



o desarrollo experimental.

- Las propuestas podrán contemplar adicionalmente en el marco de sus objetivos el entrenamiento inicial de modelos profundos o superficiales (pilotaje) a partir del conjunto de datos generado.
- Con el fin de asegurar que los conjuntos de datos generados puedan ser utilizados de manera efectiva en aplicaciones futuras de Inteligencia Artificial, las propuestas deberán contemplar, desde su diseño, un volumen mínimo de registros acorde con las exigencias técnicas de los enfoques de modelado que se proyecta aplicar en el análisis prospectivo. El volumen mínimo de registros se puede asumir de acuerdo con las recomendaciones que se dan en los **requisitos técnicos de los datos**, de los términos de referencia.
- La propuesta deberá incluir al menos un estudiante de nivel de formación (pregrado o postgrado) preferiblemente el estudiante vinculado a un semillero de investigación.
- La propuesta deberá incluir adicional a los productos MINCIENCIAS; como entregables, un informe que contemple el análisis prospectivo sobre el posible uso del dataset para entrenamiento, validación o evaluación de modelos de IA en escenarios futuros.
- Un participante puede vincularse con asignación de máximo 10 horas semanales y mínimo 6 horas semanales, respetando la tipología profesoral. La bolsa de horas para los investigadores que presentan la propuesta puede ser entre 18 y 24 horas semanales dependiendo de la conformación del equipo, con un máximo de tres investigadores. Esta distribución debe estar acorde a las responsabilidades y el cronograma de trabajo.

#### **Requisitos técnicos de los datos**

- Respecto a la fuente de los datos esperada. Se esperan datos ya existentes (investigaciones previas, bases institucionales o gubernamentales, registros administrativos, observatorios, encuestas, entre otros.).

El origen de los datos podrá ser de algunas de los siguientes:

- (i) El conjunto de datos generado de un proyecto terminado o en curso, en donde la institución actuó en calidad de ejecutora o líder de proyecto vinculado a un grupo de investigación, y que se encuentre formalizado en vicerrectoría de investigaciones. Indicar el código y nombre del proyecto formalizado en vicerrectoría de investigaciones.
- (ii) El conjunto de datos generado de un proyecto de extensión terminado o en curso, en donde la institución actuó en calidad de ejecutora o líder de proyecto. Indicar el código y nombre del proyecto formalizado en vicerrectoría de extensión.



- (iii) El conjunto de datos generado de un proyecto de formación en investigación terminado, en donde la institución actuó en calidad de ejecutora o líder de proyecto vinculado a un Semillero de investigación formalizado.
- (iv) Datos institucionales derivados de las funciones sustantivas de la institución que cuenten con el aval de la dependencia proveedora.
- (v) Datos abiertos publicados en bases de datos gubernamentales o generados por la institución y que estén disponibles en abierto.
- Respecto a la naturaleza de los datos. Se esperan datos tabulados, estructurados o semiestructurados (tales como preguntas abiertas en una encuesta) susceptibles de ser convertidos a abiertos y procesables.
  - Datos con trazabilidad y respaldo técnico. Entre estos pueden considerarse ejemplos como: Registros epidemiológicos, encuestas, sensores, bases educativas-institucionales (investigación, enseñanza, extensión), datos abiertos. Si los proyectos de la convocatoria van a trabajar con bases de datos preexistentes como fuente principal de datos, es fundamental establecer criterios técnicos, éticos y metodológicos para considerar que esas bases de datos son válidas, confiables y reutilizables, y que pueden ser sometidas a procesos de curaduría, anotación, documentación y publicación en repositorios.
  - Respecto a la cantidad de datos a anotar. Estos deben estar de acuerdo con el potencial uso del dataset en las fases de entrenamiento, validación o evaluación de modelos de IA en escenarios futuros de investigación aplicada o desarrollo experimental. Este potencial de uso debe estar alineado con el análisis prospectivo
- (I) Para técnicas de aprendizaje superficial (Shallow Learning), el conjunto de datos deberá contar, con un número no inferior a los cientos de muestras/registros (no dimensiones\*) que permitan una convergencia significativa al elegir este tipo de modelos.
- (II) Para técnicas de aprendizaje profundo (Deep Learning), al considerar la naturaleza de estos, se requiere un volumen superior, por lo que el conjunto de datos deberá contar con miles registros.
- (III) En el caso especial, que el conjunto de datos esté compuesto por imágenes y se pretenda aplicar aprendizaje profundo, se aceptarán proyectos que garanticen un volumen mínimo de cientos de imágenes, siempre que se justifique su suficiencia técnica en función del objetivo del modelo y el dominio de aplicación.
- Estas consideraciones permitirán que los conjuntos de datos sean pertinentes, robustos y reutilizables en futuros escenarios de entrenamiento, validación o evaluación de modelos de IA, de acuerdo con estándares ampliamente aceptados en la literatura especializada en ciencia de datos y aprendizaje automático.



- Las bases de datos utilizadas como fuente de los conjuntos de datos a curar, documentar y publicar deberán contar con trazabilidad verificable, permisos de uso o licencias explícitas, características técnicas que permitan su procesamiento, y condiciones éticas que aseguren el cumplimiento de la normativa vigente en protección de datos y confidencialidad. La propuesta deberá incluir descripción detallada de la base, su origen, cobertura, estructura, limitaciones, y el plan de curaduría y documentación técnica que permita su publicación conforme a principios FAIR.

- En el caso de utilizar como fuente primaria u origen de datos mayoritario, una o varias fuentes públicas, ya sea proveniente de portales gubernamentales de datos abiertos o instituciones locales o internacionales, la propuesta deberá especificar, el valor agregado y científico añadido por el equipo de investigación al conjunto de datos existente de forma clara y cuantitativa.

#### **Requisitos del equipo de investigadores**

- El investigador principal, en el caso de que la propuesta a presentar derive de un proyecto previo o en curso, el investigador debe haber participado en calidad de investigador principal o coinvestigador y contar con la filiación UDES.

- El investigador principal UDES deben ser profesores de tiempo completo con contrato vigente y estar vinculado al menos a un grupo de investigación institucional.

- Los coinvestigadores UDES pueden ser profesores de tiempo completo o medio tiempo con contrato vigente y estar vinculados a los grupos de investigación institucionales.

- Los profesores que no tengan vinculación a un grupo de investigación institucional al momento de la convocatoria deberán proyectar su vinculación con productos concretos resultados en el marco de la propuesta.

- Todos los investigadores deben estar a Paz y Salvo con la Vicerrectoría de Investigaciones en lo referente a los compromisos establecidos hasta el 2024.

- Los investigadores podrán tener hasta dos proyectos activos (en ejecución) para el año 2026, siempre y cuando la cantidad de horas no supere las condiciones del tipo de contratación.

- El investigador principal de la propuesta debe demostrar experiencia en investigación a través del liderazgo/coliderazgo de proyectos de investigación previos con cumplimiento al 100% y producción científica.

- Al momento de la postulación los investigadores deben tener el CvLAC actualizado y el perfil en el Sistema de Información SIGIIP, con los registros actualizados de proyectos y productos comprometidos con la institución.



· El equipo de investigadores interdisciplinar debe estar conformado al menos por un experto temático y un investigador con experiencia en análisis de datos. Puede hacer parte del equipo un investigador junior (sin experiencia o producción previa o tutor de semilleros). Esta conformación del equipo teniendo en cuenta las necesidades del diseño y publicación del dataset o de las intenciones de fortalecimiento de la formación de investigadores o articulación de esfuerzos.

**Requisitos para la Inscripción**

- Registro y postulación de la propuesta de investigación a través del Sistema Integrado de Gestión de Investigaciones, Indicadores y Proyectos de la Universidad de Santander SIGIIP-UDES. (<http://sigiip.udes.edu.co/>)
- Inscripción de la propuesta de investigación en uno de los ODS al que responda el proyecto. Y si se alinea con alguna de ejes priorizados: Seguridad Alimentaria y Nutricional o Salud Planetaria.
- Aval del comité de investigaciones de la facultad o instituto al cual se vincula.
- Documentos soporte de alianzas si aplica y los soportes que se definen en el numeral 8 – Documentos requeridos en el proceso de Inscripción-.

**Compromisos adquiridos previos al Inicio del Proyecto** (para los proyectos que resulten elegibles de esta convocatoria)

- El equipo de investigadores deberá haber realizado el curso auto gestionable de ética, bioética e integridad científica ofrecido por la institución y cargado el soporte al SIGIIP al finalizar el semestre B del 2025.
- Los miembros del equipo que NO puedan certificar o demostrar sus habilidades ciencia de los datos, deberán realizar, aprobar y certificar un curso de ciencia de datos, provisto por las plataformas virtuales designadas por la vicerrectoría de investigaciones al momento de anuncio de los resultados.

**Consideraciones**

**1. Articulación de la convocatoria con Objetivos de Desarrollo sostenible y focos priorizados**

Aun cuando la convocatoria es abierta al diseño y publicación de datos de cualquier área de conocimiento, desde el año 2022, la Universidad de Santander – UDES ha venido promoviendo el desarrollo de la investigación con propósitos alineados a la contribución de metas de desarrollo sostenible, Misiones declaradas por el Ministerio de Ciencia, Tecnología e Innovación



Minciencias, y focos Institucionales priorizados. En consecuencia, la convocatoria para el diseño y publicaciones de conjunto de datos con potencial de explotación de IA se alinea con estas intenciones institucionales.

El acceso a datos de calidad provenientes de diferentes contextos permite abordar de forma más precisa y contextualizada los retos asociados a los Objetivos de Desarrollo Sostenible, como la salud, el cambio climático, la seguridad alimentaria y la educación.

Así mismo, se articula con las Misiones declaradas por el Ministerio de Ciencia, Tecnología e Innovación Minciencias, contribuyendo específicamente a:

- La Misión de Bioeconomía, Ecosistemas Naturales y Territorios Sostenibles, al posibilitar la generación de datos sobre biodiversidad, bienes y servicios ecosistémicos, y su integración con herramientas tecnológicas para la toma de decisiones territoriales sostenibles.
- La Misión de Derecho a la Alimentación, al aportar datos estratégicos sobre seguridad alimentaria, producción agroecológica, acceso y consumo, que puedan ser base para el diseño de soluciones innovadoras, inclusivas y sostenibles.
- La Misión de Soberanía Sanitaria y Bienestar Social, al fomentar el uso de datos que permitan modelar fenómenos relacionados con la salud pública, el bienestar social y la atención diferencial en territorios.
- La Misión de Energía Sostenible, eficiente y asequible, mediante la generación o documentación de datos sobre consumo energético, acceso a fuentes limpias o iniciativas comunitarias de transición energética, si son incluidas como objeto de análisis prospectivo.
- El Programa Ciencia para la Paz y la Ciudadanía, al promover el acceso abierto a datos que permiten comprender las causas estructurales de conflictos territoriales, la exclusión social o las vulnerabilidades de poblaciones específicas, así como apoyar estrategias de formación ciudadana basadas en evidencia.

Desde el año 2022, la Universidad de Santander – UDES priorizó dos focos que orientan la investigación interdisciplinar hacia la solución de problemas complejos; Seguridad Alimentaria y Nutricional y la Salud Planetaria, los cuales buscan atender desafíos del país en correspondencia con capacidades institucionales para abordarlos de manera integral, articulando el conocimiento desde distintas áreas del conocimiento. Por consiguiente, se considera que existe un potencial de datos provenientes de proyectos interdisciplinarios de convocatorias pasadas con potencial de explotación en inteligencia artificial. Por lo anterior esta convocatoria considera adicionalmente estos focos estratégicos priorizados, sin que esto constituya un factor excluyente para proyectos que no los contemplen.

**Seguridad alimentaria y nutricional.** Según el CONPES 113 de 2008, la seguridad alimentaria es



definida como “la disponibilidad suficiente y estable de alimentos, el acceso y el consumo oportuno y permanente de los mismos en cantidad, calidad e inocuidad por parte de todas las personas, bajo condiciones que permitan su adecuada utilización biológica, para llevar una vida saludable y activa”<sup>1</sup>.

En la seguridad alimentaria inciden determinantes sociales que se pueden categorizar como determinantes estructurales (posición socioeconómica, estructura social, y clase social, incluyendo género y etnia) y determinantes intermediarios (circunstancias materiales, por ejemplo, condiciones de vida y de trabajo, disponibilidad de alimentos) que generan un impacto en la salud y el bienestar<sup>2</sup>.

**Salud planetaria.** La salud planetaria está ligada a la salud de los ecosistemas y aborda desafíos globales como, cambio climático, pérdida de la biodiversidad, contaminación ambiental, entre otros, para promover la salud y el bienestar de las personas y del planeta en su conjunto. La Red inVIVO Planetary Health 2018, define la salud planetaria como la vitalidad interdependiente de todos los ecosistemas naturales y antropogénicos (sociales, políticos y de otro tipo)<sup>3</sup>.

En la actualidad existen tres enfoques holísticos e interdisciplinarios orientados a salvaguardar la salud: Una Sola Salud (OneHealth), EcoSalud (Ecohealth), y Salud planetaria (PlanetaryHealth), este último, se ha promovido para indicar “las relaciones dinámicas y sistémicas entre los cambios ambientales globales, sus efectos en los sistemas naturales y cómo los cambios en los sistemas naturales afectan la salud y bienestar de los humanos a múltiples escalas: global, regional y locales ”<sup>4</sup>.

La presente convocatoria considera esta apuesta institucional, dado que busca generar insumos estratégicos para el análisis, la intervención y la toma de decisiones en áreas tales como: la salud planetaria y la seguridad alimentaria. La generación, organización y publicación de datos abiertos, estructurados y alineados con los principios FAIR, permite a la UDES avanzar en la construcción de capacidades científicas orientadas al conocimiento de fenómenos como la inseguridad alimentaria, los impactos del cambio climático o la gestión sostenible del ambiente, la gestión para la obtención de energía renovable y la salud pública.

En el caso de la Seguridad Alimentaria y Nutricional, la recopilación y curaduría de datos puede contribuir significativamente al monitoreo de la disponibilidad, el acceso, el consumo, el aprovechamiento biológico de los alimentos y la calidad e inocuidad alimentaria. Esto a partir de fuentes diversas como encuestas poblacionales, sistemas de salud, bases de datos institucionales y observatorios sociales. Así, la convocatoria permite avanzar hacia sistemas de información que fortalezcan las respuestas de política pública y proyectos de innovación social, con base en evidencia.

De igual forma, en el ámbito de la Salud Planetaria, los conjuntos de datos que integran variables ambientales, sociales, territoriales y de salud pueden ser claves para comprender la interdependencia entre los ecosistemas y la salud humana. La convocatoria incentiva la documentación y organización de datos que permitan caracterizar problemáticas como la contaminación, la pérdida de biodiversidad, los efectos de fenómenos climáticos extremos, las



desigualdades en el acceso a entornos saludables la búsqueda de fuentes energéticas alternativas basada en datos.

---

<sup>1</sup> CONPES 113 de 2008. Política nacional de seguridad alimentaria y Nutricional (2008, 31 de marzo). Departamento Nacional de Planeación <https://www.minagricultura.gov.co/Normatividad/Conpes/Conpes%20113%20de%202008.pdf>

<sup>2</sup> MINSALUD. (2013). Observatorio de Seguridad Alimentaria y Nutricional- OSAN Colombia <https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/PP/SNA/Guia-metodologica-osan.pdf>

<sup>3</sup> Prescott, S.L.; Logan, A.C.; Albrecht, G.; Campbell, D.E.; Crane, J.; Cunsolo, A.; Holloway, J.W.; Kozyrskyj, A.L.; Lowry, C.A.; Penders, J.; et al. The Canmore Declaration: Statement of Principles for Planetary Health. *Challenges* 2018, 9, 31. <https://doi.org/10.3390/challe9020031>

<sup>4</sup> Pongsiri, M. J., Bickersteth, S., Colón, C., DeFries, R., Dhaliwal, M., Georgeson, L., ... Ungvari, J. (2019). Planetary health: from concept to decisive action. *The Lancet Planetary Health*, 3(10), e402–e404. [https://doi.org/10.1016/S2542-5196\(19\)30190-1](https://doi.org/10.1016/S2542-5196(19)30190-1)

## 2. Conceptos orientadores que sustentan la Convocatoria

A continuación, se presentan conceptos fundamentales que sustentan la Convocatoria Interna UDES 2025 para el Diseño y Publicación de Conjuntos de Datos con Potencial de explotación en Inteligencia Artificial.

- **Conjunto de datos (dataset).** Es una colección estructurada de información que puede contener números, texto, imágenes u otros tipos de datos, organizada con el fin de ser consultada, analizada y reutilizada en investigaciones o desarrollos tecnológicos (James, Witten, Hastie & Tibshirani, 2021).
- **La curaduría de datos.** Hace referencia al proceso técnico y metodológico mediante el cual los datos son seleccionados, organizados, validados, documentados y enriquecidos, con el objetivo de asegurar su calidad y facilitar su posterior reutilización (Wilkinson et al., 2016).
- **Principios FAIR.** Corresponden a un conjunto de buenas prácticas que establecen que los datos deben ser Encontrables (Findable), Accesibles (Accessible), Interoperables (Interoperable) y Reutilizables (Reusable). Estos principios buscan promover una ciencia abierta, transparente y colaborativa, asegurando que los datos científicos puedan ser compartidos y utilizados por distintos actores (Wilkinson et al., 2016).
- **Inteligencia artificial (IA).** se define como la rama de la informática dedicada al desarrollo de sistemas capaces de realizar tareas que normalmente requieren inteligencia humana, como el razonamiento, la predicción, el procesamiento de lenguaje o la toma de decisiones (Russell & Norvig, 2020).
- **Aprendizaje automático (machine learning).** Es un subcampo de la IA que permite que los sistemas aprendan a partir de datos sin haber sido programados explícitamente para cada tarea.



Mediante algoritmos, los sistemas pueden identificar patrones, predecir resultados o clasificar información (Hastie, Tibshirani & Friedman, 2009).

- Dentro del aprendizaje automático, se distingue el **aprendizaje superficial (shallow learning)**, que utiliza modelos de baja complejidad como regresiones lineales, árboles de decisión o máquinas soporte vectorial, adecuados para conjuntos de datos de tamaño moderado (Kuhn & Johnson, 2013).
- Por otro lado, el **aprendizaje profundo (deep learning)** hace uso de redes neuronales con múltiples capas ocultas que permiten detectar patrones complejos, tales como Convolucionales, Grafos, Transformers o modelos de corta memoria, entre otros. Esta técnica requiere grandes volúmenes de datos y es ampliamente utilizada en el reconocimiento de voz, visión por computador o procesamiento del lenguaje natural (Goodfellow, Bengio & Courville, 2016).
- **Prototipo.** Es una representación funcional preliminar de un producto, sistema o algoritmo, desarrollado con el fin de validar su utilidad, realizar pruebas iniciales y recoger retroalimentación antes de su implementación a mayor escala (Project Management Institute, 2021).
- **Proyecto piloto.** Es la ejecución en pequeña escala de una solución o intervención tecnológica para observar su funcionamiento en condiciones reales, antes de ser replicada o escalada (OECD, 2005).
- **Ciencia abierta.** Es un movimiento internacional que promueve el acceso libre al conocimiento científico, incluyendo datos, publicaciones, metodologías, herramientas y procesos. Su propósito es democratizar la ciencia, fomentar la transparencia y facilitar la colaboración entre investigadores, instituciones y ciudadanía (UNESCO, 2021).

### 3. Contenido de la propuesta

Los siguientes componentes se requiere que estén cargados en el Sistema Integrado de Gestión de Investigaciones, Indicadores y Proyectos de la Universidad de Santander SIGIIP-UDES. (<http://sigiip.udes.edu.co/>).

- Título del proyecto
- Resumen ejecutivo
- Problema de investigación
- Estado del arte
- Pregunta de investigación
- Objetivo general



- Marco teórico
- Conocimiento que generará el proyecto
- Justificación del proyecto
- Justificación del comité de bioética

**Datos generales**

- Áreas del conocimiento
- Grupos de interés
- Lugar de ejecución
- Objetivos específicos
- Palabras clave
- Referencias bibliográficas
- Registro de Objetivos de Desarrollo Sostenible – ODS
- Riesgos del proyecto
- Solución que aporta el proyecto
- Soportes aportados.

**Integrantes**

- Grupos de investigación
- Semilleros
- Grupos externos
- Integrantes (Investigadores, asesores)
- Unidades académicas/administrativas
- Datos del responsable de la Entidad Ejecutora



**Líneas de investigación**

- Líneas institucionales (Seleccione la línea temática)
- Sublíneas de investigación (Seleccione grupo de investigación y la línea de investigación del grupo a la que se articula el proyecto)
- Foco estratégico

**Metodología**

- Metodología : Metodología de apertura del dato (si aplica)
- Técnicas e instrumentos

**Consideraciones éticas, resultados de sometimientos a comités y otros**

- Consideraciones éticas

**Cronograma del proyecto**

**Presupuesto**

**Productos resultados de la investigación**

- Proyección productos

**Impacto, transferencia e innovación del proyecto**

- Impactos esperados

El Contenido de la propuesta puede ser ampliado en el Anexo 8.

**4. Condiciones inhabilitantes**

**Del proyecto**

- Cuando el proyecto no se ajusta a las condiciones y requisitos establecidos en los presentes



términos de referencia.

- El proyecto no está enmarcado en los focos de la convocatoria.

**De los investigadores.**

- Alguno de los miembros de la alianza no cumple con los términos.
- El proponente no se encuentra a paz y salvo con la vicerrectoría de investigaciones.
- Que el investigador principal se presente con este mismo rol en más de un proyecto
- Que el Investigador Principal tenga más de un proyecto en ejecución para el período en el cual se inicie el proyecto Semestre A del 2026. Lo anterior teniendo en cuenta los requerimientos institucionales en la asignación de horas en las diferentes funciones sustantivas.

**5. Documentos requeridos en el proceso de inscripción**

No.	Documento	Soporte
1	Paz y salvo- investigador emitido por la Vicerrectoría de Investigaciones para cada investigador UDES participante.	Solicitud de paz y salvo mediante el siguiente formulario en línea: <a href="https://forms.office.com/r/JpJvRh5mid">https://forms.office.com/r/JpJvRh5mid</a>
2	Acta de Comité de Investigaciones en la cual se valida y aprueba la propuesta; (Un solo archivo PDF).	Aval (Formato 087) y acta de comité de investigaciones (un solo archivo pdf).
3	Acuerdo de propiedad intelectual firmado por todos los participantes en la propuesta, de acuerdo con el Estatuto de Propiedad Intelectual de la UDES. (Archivo PDF).	Formato 032 Acuerdo de propiedad intelectual.
4	Una carta de confidencialidad por cada uno de los investigadores participantes en un solo archivo	Formato 033 - Acuerdo de confidencialidad.



		de acuerdo con el Estatuto de Propiedad Intelectual de la UDES (archivo PDF).	
	5	Carta de autorización de uso y almacenamiento de datos personales de cada uno de los investigadores, de acuerdo con modelo MINCIENCIAS (archivo PDF).	Formato 034 - Autorización de uso y almacenamiento de datos personales.
	6	Si aplica, carta de intención de la entidad con quien se realizará la alianza (si aplica), especificando la participación que se realizará en el proyecto y el aporte en especie y en efectivo. (Archivo PDF). (En el caso que el proyecto del cual se derivan los datos hubiese sido generado en alianza con un externo, favor incluir en la carta manifestación explícita de la aprobación para el procesamiento y publicación de los datos (si aplica)). Si la alianza contempla proceso de transferencia tecnológica o de conocimiento, especificar las condiciones de la transferencia y los compromisos acordados.	Modelo de carta de intención. Guía 004 Modelo de carta de aprobación para el procesamiento y publicación de datos.
<b>Evaluación de proyectos</b>	Los proyectos inscritos, dentro de los plazos establecidos para la presente convocatoria y que cumplan con los requisitos, se someterán a los siguientes criterios de evaluación.		
	<b>No.</b>	<b>Criterios de evaluación</b>	<b>Puntaje máximo global</b>
	0	<b>Cohesión de la Propuesta:</b> Se evaluará la coherencia /cohesión de la propuesta a través del todo el documento, título, problemática, justificación, estado del arte, descripción general del conjunto de datos, entregables y perfil del equipo de investigación.	Cumple / No Cumple En caso de no cumplir la propuesta no será evaluada
	1	<b>Calidad de la propuesta.</b> Claridad, coherencia y madurez del conjunto de datos proveniente de investigaciones previas, relevancia del fenómeno	20



		observado o representado.	
	2	Calidad y Cantidad del Conjunto de datos. Considerando los tipos de conjuntos de datos, se evaluará el número de registros y variables/dimensiones /observaciones, nivel de completitud.	30=(2a+2b)
	2a	<p><b>Volumen y Alcance del Conjunto de datos.</b> Evalúa la cantidad de datos disponibles, ajustada al tipo de datos:</p> <p><b>Datos estructurados:</b> número de registros, variables, archivos.</p> <p><b>Datos no estructurados:</b> número de imágenes, horas de grabación, eventos capturados, resolución espacial/temporal.</p> <p>Mínimo (Mencionado en numeral 6 (Requisitos Técnicos de los Datos)): 1–5 pts</p> <p>Medio: 5–10 pts</p> <p>Grande: 10–15 pts</p>	15
	2b	<p><b>Riqueza y Diversidad de Información:</b> Evalúa la complejidad, multidimensionalidad y utilidad del dataset.</p> <p>Monodimensional / Monoclase (solo una variable de interés o clase): 1–5 pts</p> <p>Multivariado o Multiclase (varias variables, tipos de medición, o fuentes): 5-10 pts</p> <p>Multidimensional o multimodal (ej., imagen + texto + sensor), con potencial de reutilización en diferentes disciplinas: 11–15 pts</p>	15
	3	Coherencia de las técnicas e instrumentos. Evalúa la	10



	coherencia entre las técnicas e instrumentos descritos en el componente metodológico para el procesamiento del conjunto de datos inicial en relación con la cantidad de datos.	
4	Experiencia del equipo. Se evalúa la pertinencia del perfil de los investigadores y su experiencia en relación con la naturaleza del conjunto de datos propuesto y el fenómeno o problemática que representa.	5
5	Potencial de Reutilización Interdisciplinaria. Evalúa el grado en que el dataset puede ser útil para investigadores de otras disciplinas, fuera del campo original del proyecto.	15
	Total	80
	<b>Alineación con focos institucionales priorizados.</b> Se evalúa la relevancia y pertinencia con los focos de la convocatoria	4
	<b>Vinculación de estudiantes.</b> Se evalúa la vinculación coherente de estudiantes posgrado/pregrado (semilleros).	4
	<b>Aliados Estratégicos.</b> Se evalúa la vinculación de aliados estratégicos (internos o externos), que aporta al fortalecimiento técnico, la validación del caso de uso o la transferencia de resultados.	2
	<b>Vinculación de Tutor de Semilleros.</b> Se evalúa la vinculación de un tutor de Semilleros de la Universidad de Santander en el equipo de investigadores.	4



	<p><b>Vinculación de un Investigador sin experiencia.</b> Se evalúa la vinculación de un investigador sin previa experiencia o en formación, de la Universidad de Santander al equipo de investigadores.</p>	6
	<p><b>TOTAL</b></p>	20
<p>Los proyectos con un puntaje igual o superior a setenta (80) puntos en la evaluación ingresarán al banco de elegibles y se asignarán los recursos hasta su agotamiento en estricto orden descendente.</p> <p>Las propuestas que estén en el banco preliminar de elegibles deberán solicitar aval al Comité Institucional de Bioética, en el periodo indicado en las actividades de la convocatoria. El concepto bioético favorable es requisito fundamental para que el proyecto sea aprobado y financiado.</p> <p>Los resultados de la convocatoria se publicarán en las fechas establecidas en el cronograma y en la página web de la Universidad de Santander.</p>		
<p><b>Descripción modalidades</b></p>		
<p><b>Descripción cofinanciamiento</b></p>	<p>Cada proyecto podrá solicitar hasta veinte millones de pesos m/cte (\$20.000.000).</p> <p>El monto solicitado deberá ser justificado en el alcance técnico de la propuesta.</p>	
<p><b>Duración</b></p>	<p>Se financiarán proyectos con una duración máxima doce (12) meses.</p> <p><b>Nota.</b> Durante la ejecución del proyecto no se aprueban prórrogas al plazo establecido, sólo se podrá solicitar hasta seis meses de plazo para entrega de publicaciones de artículos indexados.</p>	
<p><b>Resultados esperados</b></p>	<ol style="list-style-type: none"> <li><b>Conjunto de Datos en Formato Abierto:</b> Archivo o conjunto de archivos en formatos abiertos y procesables con la información recolectada de forma ordenada en el caso de ser datos tabulares la primera fila del archivo deberá ser los nombres de las columnas.</li> <li><b>Documentación del Conjunto de Datos:</b> asociada a la descripción de las columnas/observaciones/dimensiones del conjunto de datos, diccionario de datos y metadata y muestras ejemplificadas y documentadas del dataset.</li> <li><b>Documento Técnico:</b> que describa objetivo, justificación, metodología de recolección, proceso de curación y anotación de datos (y metadatos).</li> </ol> <p>o <b>Análisis Descriptivo del Dataset:</b> Variables, distribuciones estadísticas de las dimensiones, anotaciones, codificaciones, entre otros.</p>	



4. **Documento con estudio prospectivo:** Adicionalmente se espera que los investigadores realicen un estudio prospectivo de la explotación del conjunto de datos en el cual se deberá incluir. Este análisis facilitará el aprovechamiento transversal del dataset en distintas áreas del conocimiento.

o **Análisis de modelos de IA.** Describir qué modelos de inteligencia artificial existentes ofrecen un buen desempeño con el dataset propuesto; sin depender de un caso de uso específico, garantizando su valor en tareas diversas como clasificación, predicción o extracción de patrones, permitiendo establecer comparaciones objetivas a través de métricas estandarizadas ampliamente reconocidas.

o **Análisis Exploratorio** de casos de uso en los cuales el conjunto de datos provee/complementa información.

**5. Participación en (1) evento científico.**

NOTAS:

- Toda la documentación generada deberá estar en inglés como primer idioma y en español (opcional).
- Para los entregables 2, 3 y 4, la vicerrectoría de investigación proporcionará plantillas especializadas, las cuales los investigadores podrán utilizar para el cumplimiento del entregable.

**Resultados Diferenciadores:** Los siguientes pueden ser resultados opcionales a la propuesta presentada. Si bien no son obligatorios, serán considerados como elementos prioritarios en el proceso de evaluación.

6. Producto de Formación de Recurso humano de Pregrado o Postgrado.
7. Publicación en revista indexada.
8. Proceso de Apropiación Social, teniendo en cuenta la política nacional de apropiación social de conocimiento<sup>5</sup>

**Características del dataset**

- Se espera la aplicación de procesos de curaduría técnica: limpieza, validación, estandarización, enriquecimiento, documentación y anotación.
- Los datos deben cumplir los principios FAIR (Encontrables, Accesibles, Interoperables, Reutilizables).



· Los conjuntos de datos generados deberán generar valor agregado, comprensión y entendimiento del fenómeno en estudio, organizados de manera coherente cuando incluyan múltiples dimensiones teniendo en cuenta los requerimientos técnicos del dato descritos en el numeral 7 de los términos de referencia. Se requiere que sean anónimos y éticamente gestionados, especialmente si contienen datos sensibles, cumpliendo con la normativa vigente. Así mismo, se esperan que estén balanceados dimensionalmente para evitar sesgos y, cuando sea posible, incluir variables diversas (numéricas, categóricas, temporales, geoespaciales). Se espera una documentación rigurosa, que incluya descripción general, diccionario de datos con ejemplos, metadatos y una declaración explícita de alineación con los principios FAIR, que asegure su reutilización y accesibilidad futura.

· **De las publicaciones:** Se valorará la generación de productos asociados al dataset, como 'data papers' o artículos de reflexión sobre la metodología de recolección, anotación y curaduría de los datos y la publicación en revistas indexadas, se recomendará seleccionar medios afines al área temática del proyecto y al tipo de producto generado.

· **De la formación de estudiantes:** Se valorará la proyección de Trabajo de pregrado finalizado con acta de sustentación o estudiante de posgrado vinculado con trabajo de investigación relacionado con el proyecto, y estudiantes de semillero de investigación con evidencia concreta de participación en el proyecto.

Si bien el objetivo principal de esta convocatoria es promover la publicación de datos abiertos como parte del compromiso institucional con la ciencia abierta, la Universidad y la Vicerrectoría de Investigación reconocen que existen casos especiales en los que la publicación de los datos no es posible por razones éticas, legales, de secreto empresarial o de confidencialidad. En tales situaciones, las solicitudes de excepción deberán ser debidamente justificadas y serán evaluadas por el Comité Central de Investigaciones, el cual determinará la viabilidad de mantener los datos en acceso restringido.

<sup>5</sup> Documento de Política Pública de Apropiación Social del Conocimiento en el marco de la Ciencia, Tecnología e Innovación.2021 Minciencias

**Propiedad intelectual**

La propiedad intelectual rige lo contemplado en el Acuerdo No 030 del Consejo Académico de la Universidad de Santander, disponible en el siguiente enlace:

[https://udes.edu.co/images/la\\_universidad/documentos/estatuto\\_de\\_propiedad\\_intelectual.pdf](https://udes.edu.co/images/la_universidad/documentos/estatuto_de_propiedad_intelectual.pdf)

**Criterios de selección**

**Criterios de priorización**

1. Se priorizarán propuestas que presenten conjuntos de datos con mayor cantidad de registros, variables, dimensiones u observaciones, y que evidencien un alto nivel de completitud y organización inicial. Esta priorización se basa en la mayor viabilidad de obtener resultados significativos en los procesos de curaduría, enriquecimiento, anotación y documentación técnica.



	<p>2. Proyectos que contemplen desde su formulación la vinculación de aliados estratégicos (internos o externos), que puedan aportar al fortalecimiento técnico, la validación del caso de uso o la transferencia de resultados.</p> <p>3. Proyectos que incluyan en el equipo de investigadores al menos un tutor de semilleros de investigación.</p> <p>4. Proyectos que vinculen al menos un semillero adscrito a un grupo de investigación de la Universidad de Santander.</p> <p>5. Se priorizarán proyectos que involucren la formación de estudiantes de pregrado o posgrado mediante su participación en la gestión del conjunto de datos.</p> <p>6. Se valorará la proyección de publicaciones científicas derivadas del dataset en revistas indexadas en Scopus o WoS, preferiblemente en cuartiles Q1 o Q2. Asimismo, se priorizarán iniciativas que contemplen procesos de apropiación social del conocimiento, en línea con la Política Nacional de Apropiación Social y la Guía Institucional UDES.</p>
<p><b>Rubros financieros</b></p>	<p><b>Rubros financiables por parte de la UDES especie</b></p> <ul style="list-style-type: none"> <li>- Personal científico. - Hasta 24 horas distribuidas</li> <li>- Equipos de cómputo y software. Corresponde a la capacidad instalada en la institución y al software licenciado con los que cuentan los programas académicos.</li> </ul> <p><b>En efectivo:</b></p> <ul style="list-style-type: none"> <li>- Costos de publicación y asignación de DOI</li> <li>- Participación en eventos científicos nacionales o internacionales</li> <li>- Curso certificado en ciencia de datos</li> <li>- Recursos para apropiación social y visualización de datos</li> <li>- Costo de Procesamiento de Artículos - APC</li> </ul> <p>Para un total Financiable en recursos frescos de máximo: \$20.000.000</p> <p>Rubros no financiables por parte de la UDES en esta convocatoria</p>



- Salidas de campo.
- Bibliografía.
- Servicios técnicos externos para limpieza, anotación o curación de datos
- Salarios de personal no vinculado a la UDES.
- Desarrollo de páginas WEB.
- Otros, infraestructura tecnológica y gastos de registros de propiedad intelectual y pagos de licencias de Software.
- Gastos suntuarios o personales.
- Estudios de factibilidad de proyectos (asesoría jurídica, financiera, etc.).
- Adquisición de vehículos automotores.
- Compra de equipos de laboratorio.
- Servicios públicos y personal administrativo.
- Adquisición de equipos de cómputo, muebles y enseres, terrenos, dotación.
- Adecuaciones físicas.
- Gastos de administración del proyecto o comisión alguna.
- En general, no se deben realizar gastos que no contribuyan a cumplir los fines del proyecto.
- Servicios técnicos de digitación, tabulación, recolección de información o similares.
- Servicios técnicos de personal en los cuales la institución tenga la capacidad humana, tecnológica y científica para ejecutar, elaborar y desarrollar la actividad.
- Gastos como el pago de impuestos propios de la operación de las entidades, el cubrimiento de deudas, costos fijos y gastos de funcionamiento, y demás gastos recurrentes de operación.
- Hechos cumplidos.
- Servicios Técnicos.



	<p><b>Nota 1.</b> El investigador principal podrá realizar modificación en los rubros del presupuesto, como máximo una vez después de aprobado el presupuesto y sujeto a las siguientes condiciones: a) siempre que sean necesarios y justificados, para cumplir con los fines de la investigación; b) realizarlo dentro del tiempo de ejecución de la propuesta; c) solicitar oportunamente la aprobación del cambio a la Vicerrectoría de Investigaciones y de la Vicerrectoría Administrativa y Financiera; d) siempre que no correspondan a falta de planeación; y e) no violar las limitaciones aquí establecidas, relacionadas con el pago de personal.</p> <p><b>Nota 2.</b> Sólo se podrán ejecutar recursos dentro del término de duración de la propuesta</p> <p><b>Nota 3.</b> Toda contrapartida deberá estar formalizada mediante una carta de intención para efectos del proceso de evaluación interna. En caso de ser aprobado el proyecto, se deberá suscribir, previo al desembolso UDES, una carta de entendimiento entre las partes comprometidas en la cofinanciación de este, firmada por los representantes legales correspondientes</p>
<b>Fecha de cierre</b>	2025-12-19
<b>Sede que origina la convocatoria</b>	Campus Bucaramanga

## 2. TIPOS DE PROYECTO

Tipo proyecto
Proyectos de Investigación científica aplicada
Proyectos de desarrollo tecnológico

## 3. RETOS DE LA CONVOCATORIA

Retos de convocatoria
No aplica

## 4. LÍNEAS TEMÁTICAS

Líneas temáticas

## 5. MODALIDADES

Modalidad de la convocatoria
No detalladas en esta convocatoria

## 6. SEDES A LAS QUE ESTÁ DIRIGIDA LA CONVOCATORIA

 <b>Universidad de Santander</b> <b>UDES</b> VIGILADA MINEDUCACIÓN   SNIES 2832	<b>Vicerrectoría de Investigaciones</b>	
	<b>TÉRMINOS DE REFERENCIA CONVOCATORIAS</b> <b>PROYECTOS DE INVESTIGACIÓN</b> <b>VII -FT-027-UDES</b>	Versión:02

Nombre sede	Ciudad	Departamento
Extensión Bogotá D.C.	Bogotá D.C.	Cundinamarca
Campus Bucaramanga	Bucaramanga	Santander
Campus Cúcuta	Cúcuta	N. de Santander
Campus Valledupar	Valledupar	Cesar

## 7. ACTIVIDADES DE LA CONVOCATORIA

Nombre actividad	Fecha inicial	Fecha Fin
Apertura de la convocatoria	08/12/25	12/19/25
Periodo de recepción postulaciones en SIGIIP-UDES	08/12/25	10/03/25
Periodo de revisión de requisitos (Compleitud)	10/06/25	10/10/25
Periodo de subsanación de requisitos (Compleitud - en caso de requerirse)	10/14/25	10/16/25
Evaluación de viabilidad técnica y presentación ante Comité Central de Investigaciones	10/20/25	11/14/25
Publicación de resultados definitivos.	11/21/25	11/21/25
Aval presupuestal y Comité de Bioética	11/24/25	12/19/25
Cierre de la convocatoria	12/19/25	12/19/25

## 8. ENTIDADES EXTERNAS QUE PARTICIPAN

--

## 9. UNIDAD ACADEMICA VINCULADA

Nombre	Descripción
Vicerrectoría de Investigaciones	Para mayor información dirigirse a la Vicerrectoría de Investigaciones de la Universidad, bloque Arhuaco, piso 2, teléfono 6516500 Ext: 1021 - 1022, campus Bucaramanga. Email de contacto: proyectos.investigaciones@udes.edu.co