

	INSTRUCTIVO	PROCESAMIENTO DE INFORMACIÓN OBSERVATORIO DE TIERRAS RURALES	CÓDIGO	INTI-I-005
	ACTIVIDAD	ANÁLISIS DE INFORMACIÓN	VERSIÓN	1
	PROCESO	INTELIGENCIA DE LA INFORMACIÓN	FECHA	11/04/2025

AGENCIA NACIONAL DE TIERRAS - ANT

INSTRUCTIVO PARA EL PROCESAMIENTO DE INFORMACIÓN - OBSERVATORIO DE TIERRAS RURALES

MARZO DEL 2025

	INSTRUCTIVO	PROCESAMIENTO DE INFORMACIÓN OBSERVATORIO DE TIERRAS RURALES	CÓDIGO	INTI-I-005
	ACTIVIDAD	ANÁLISIS DE INFORMACIÓN	VERSIÓN	1
	PROCESO	INTELIGENCIA DE LA INFORMACIÓN	FECHA	11/04/2025

CONTENIDO

CONTENIDO	2
INTRODUCCIÓN	3
1. OBJETIVO	4
2. ALCANCE	4
3.1. Definiciones	4
3.2. Siglas	7
4. GENERALIDADES	8
4.1. MARCO DE REFERENCIA	8
5. NORMATIVIDAD APLICABLE	8
5.1. DIRECTRICES INTERNACIONALES	11
6. INSTRUCCIONES PARA EL PROCESAMIENTO DE INFORMACIÓN – OTR	12
6.1. Paso 01 - de Comprensión del Problema	12
6.2. Paso 02 Entendimiento de los datos asociados	13
6.3. Paso 03 delimitación de diccionario y preparación de datos	13
6.4. Paso 4 Modelado y Análisis Preliminar o de muestras	14
6.5. Paso 5: Evaluación (Verificación del Proceso y Validación de la Calidad)	15
6.6. Paso 6 implementación en pleno o escalabilidad	17
6.7. Paso 7 Elaboración del Informe Final de Procesamiento	17
6.8. Paso 8 Entrega de metadatos	17
6.9. Consideraciones de Excepcionalidad	18

	INSTRUCTIVO	PROCESAMIENTO DE INFORMACIÓN OBSERVATORIO DE TIERRAS RURALES	CÓDIGO	INTI-I-005
	ACTIVIDAD	ANÁLISIS DE INFORMACIÓN	VERSIÓN	1
	PROCESO	INTELIGENCIA DE LA INFORMACIÓN	FECHA	11/04/2025

INTRODUCCIÓN

El marco normativo vigente en Colombia, especialmente contemplando la promulgación de la Ley 2294 de 2023 (Plan Nacional de Desarrollo “Colombia Potencia Mundial de la Vida”) y la consolidación del Sistema Estadístico Nacional (SEN), refuerza la importancia del gobierno de datos como un eje estratégico para las organizaciones públicas colombianas. Desde el Sistema Estadístico Nacional -SEN, y con el gobierno de datos de cada organización se articularía, en cabeza del Departamento Administrativo Nacional de Estadística -DANE, tanto la producción, como la gestión y el uso de la información, garantizando siempre estándares de calidad, accesibilidad y pertinencia que permitan una adecuada y perfectible toma de decisiones basadas en evidencia, accesible también a la ciudadanía. Es en dicho marco que este instructivo para el procesamiento de información y su adecuada implementación resulta crucial para asegurar la interoperabilidad, transparencia y eficiencia en el manejo de datos.

En este contexto, las organizaciones públicas colombianas, como la Agencia Nacional de Tierras -ANT, y desde la Dirección de Gestión del Ordenamiento Social de la Propiedad (DGOSP) y el Observatorio de Tierras Rurales (OTR), enfrentan el desafío de optimizar sus procesos tanto para el uso de la información primaria como secundaria, considerando el imperativo de no realizar múltiples procesos reiterativos o similares y garantizar al mismo tiempo la trazabilidad de los datos. Esto incluye, en escenarios ideales y para procesos consolidados y rutinarios, la incorporación de datos provenientes con origen en registros administrativos internos y externos, sea mediante recepción oficial confidencial o mediante el uso de datos abiertos, los cuales deben ser gestionados teniendo presente siempre un enfoque ético y en línea, progresivamente, con los estándares internacionales. Los estándares que por ahora se recomiendan son los promovidos por el Data Management Body of Knowledge (DAMA-DMBOK) y la Norma Técnica de Calidad NTC PE 1000:2020.

Este Instructivo se acoge a las políticas de la ANT Política de Gestión de Datos e Información y a la metodología de METODOLOGIA CRISP DM que por sus siglas en inglés traduce *Cross-Industry Standard Process for Data Mining* usada en su manual de procesamiento por el SSIT.

Este instructivo se acoge a la política de gestión de datos e información de la ANT, la cual incluye la política "El dato es un activo gobernable" donde menciona que la ANT debe contar con roles y procesos específicos de Gobierno del Dato; adicionalmente, se menciona la identificación de los custodios y consumidores del dato, y ejecución de los requisitos relacionados de cumplimiento y responsabilidad. Por tanto, este instructivo define roles, actividades (pasos), custodios y consumidores para los datos e información que maneja el OTR

El presente documento tiene entonces como objetivo establecer un instructivo claro y estandarizado para el procesamiento de aquella información gestionada por el Observatorio de Tierras Rurales.

Con esta introducción, se espera entonces establecer las bases para un modelo de gestión de datos robusto y sostenible, que permita al OTR desempeñar sus funciones misionales con excelencia y contribuir al desarrollo de políticas públicas del SECTOR RURAL informadas y basadas en evidencia.

	INSTRUCTIVO	PROCESAMIENTO DE INFORMACIÓN OBSERVATORIO DE TIERRAS RURALES	CÓDIGO	INTI-I-005
	ACTIVIDAD	ANÁLISIS DE INFORMACIÓN	VERSIÓN	1
	PROCESO	INTELIGENCIA DE LA INFORMACIÓN	FECHA	11/04/2025

1. OBJETIVO

Describir los pasos a seguir en el procesamiento de la información en el Observatorio de Tierras Rurales (OTR) para el desarrollo de estudios e investigaciones a través de la definición de roles, actividades, pasos, productos esperados y criterios de calidad, garantizando la trazabilidad, interoperabilidad y uso eficiente de los datos conforme a la metodología CRISP-DM y la normativa vigente.

2. ALCANCE

El instructivo aplica para el procesamiento de información gestionada por el Observatorio de Tierras Rurales (OTR) en el marco de las políticas establecidas para el proceso de gestión de información e inteligencia de la Información de la Agencia Nacional de Tierras (ANT). Cubre desde la comprensión del problema hasta la entrega de metadatos, abarcando actividades de análisis, validación, modelado, control de calidad e implementación de modelos analíticos.

3. DEFINICIONES Y SIGLAS

3.1. Definiciones

Agencia de Desarrollo Rural (ADR): Implementa proyectos para mejorar la productividad y competitividad del sector.

AD - Active Directory almacena información acerca de los objetos de una red y facilita su búsqueda y uso por parte de los usuarios y administradores. Active Directory usa un almacén de datos estructurado como base para una organización jerárquica lógica de la información del directorio.

Analista de datos (Data analyst): es un profesional que analiza datos para ayudar a las empresas a tomar decisiones. El equivalente en español de Data Analyst, serán aquellos expertos en manipulación de Big Data, ya sea por cualificación certificada de estudios o por experiencia usando códigos libres que permitan interoperabilidad como R o Python, y que siendo parte del OTR constituye el agente central de los procesamiento de información. Idealmente debe tener buenos conocimientos de estadística y capacidad para asociar información en visores geográficos. Son actividades relativas al analista de datos las de: a) Recopilar, limpiar y preparar datos para el análisis; b) Identificar ciertos patrones y tendencias en los datos que le fueron encomendados analizar de tal forma que se puedan inferir e influir en estrategias de actuación de la ANT y la institucionalidad del Sector Rural; c) Crear bases de datos, informes y tableros que presenten insights para la mejor comprensión de la información; d) Colaborar con los demás contratistas ejerciendo diferentes roles en el marco de La Mesa Técnica de Datos MTD para definir tanto preguntas de análisis coyunturales como los objetivos de cada procesamiento permanente de información; e) Garantizar tanto la integridad como la calidad de los datos utilizados en los análisis siguiendo las directrices emanadas del MTD como del OD, DA, DQM y el DSO; g) Proporcionar recomendaciones preliminares o definitivas, en acuerdo con el coordinador del OTR, basado en los resultados de los análisis para apoyar la toma de decisiones.

	INSTRUCTIVO	PROCESAMIENTO DE INFORMACIÓN OBSERVATORIO DE TIERRAS RURALES	CÓDIGO	INTI-I-005
	ACTIVIDAD	ANÁLISIS DE INFORMACIÓN	VERSIÓN	1
	PROCESO	INTELIGENCIA DE LA INFORMACIÓN	FECHA	11/04/2025

Analista de Datos Encargado del Tratamiento: persona natural o jurídica, pública o privada, que por sí misma o en asocio con otras, realice el Tratamiento de datos personales por cuenta del responsable del Tratamiento. (Literal d, art. 3, Ley 1581 de 2012). En el marco del presente instructivo son los Analistas de Datos del OTR.

Autorización: consentimiento previo, expreso e informado de quien es Titular de los datos para llevar a cabo el Tratamiento de Datos Personales. (Ley 1581 de 2012, Art. 3°, literal a). (Tomado de POLÍTICA DE PROTECCIÓN DE DATOS PERSONALES, ANT (2024).

Aviso de Privacidad: se refiere al documento (en cualquier formato) que contiene la comunicación generada por el responsable del Tratamiento, y dirigida a quien es Titular de los datos, y que tiene como finalidad informar sobre la aplicación de las políticas de tratamiento de datos personales, dónde pueden encontrarse y qué uso se les dará. (Tomado de POLÍTICA DE PROTECCIÓN DE DATOS PERSONALES, ANT (2024).

Banco Agrario de Colombia: Ofrece servicios financieros a diversos agentes del sector rural en Colombia.

Base de Datos: conjunto organizado de datos personales que sea objeto de Tratamiento. (. (Art. 3°, literal b), Ley 1581 de 2012) (Tomado de POLÍTICA DE PROTECCIÓN DE DATOS PERSONALES, ANT (2024).

Booleano: es un término que se refiere a un dato lógico que solo puede tener dos valores: verdadero o falso.

CDO – Profesional responsable de la gestión de los datos

Clustering (análisis de grupos o agrupamientos) - el clustering sirve para organizar datos en grupos o clústeres basados en similitudes intrínsecas entre ellos. Utilizando algoritmos para identificar patrones y estructuras subyacentes en conjuntos de datos, no tiene la necesidad de usar etiquetas predefinidas.

CRISP-DM (Cross-Industry Standard Process for Data Mining) es una metodología para planificar, organizar e implementar proyectos de minería de datos. Es un modelo de proceso que describe las fases de un proyecto de ciencia de datos.

Corporación Colombiana de Investigación Agropecuaria (AGROSAVIA): Desarrolla investigaciones para innovar en el sector agropecuario.

Data Quality Manager DQM- El Data Quality Manager (DQM) o gerente de calidad de datos es el responsable de garantizar que los datos de una organización cumplan con los estándares de calidad.

Dato: Es una información transformada, gracias a la existencia de un diccionario de datos, en dato y por tanto susceptible de interpretación unívoca, al menos desde los criterios vigentes en el OTR, la ANT y por extensión en la Institucionalidad del Sector Rural.

	INSTRUCTIVO	PROCESAMIENTO DE INFORMACIÓN OBSERVATORIO DE TIERRAS RURALES	CÓDIGO	INTI-I-005
	ACTIVIDAD	ANÁLISIS DE INFORMACIÓN	VERSIÓN	1
	PROCESO	INTELIGENCIA DE LA INFORMACIÓN	FECHA	11/04/2025

Dato personal: cualquier información vinculada o que pueda asociarse a una o varias personas naturales determinadas o determinables. (Ley 1581 de 2012, Art. 3°, literal c). (Tomado de POLÍTICA DE PROTECCIÓN DE DATOS PERSONALES, ANT (2024).

Dato privado: es el dato que por su naturaleza íntima o reservada sólo es relevante para el Titular. (Tomado de POLÍTICA DE PROTECCIÓN DE DATOS PERSONALES, ANT (2024).

Dato público: es el dato calificado como tal según los mandatos de la Ley o de la Constitución Política, y todos aquellos que no sean semiprivados o privados. Son públicos, entre otros, los datos contenidos en “documentos públicos, sentencias judiciales debidamente ejecutoriadas que no estén sometidos a reserva y los relativos al estado civil de las personas”. (Literal f) Artículo 3 Ley 1266 de 2008) (Tomado de POLÍTICA DE PROTECCIÓN DE DATOS PERSONALES, ANT (2024).

Dato semiprivado: es semiprivado el dato que no tiene naturaleza íntima, reservada, ni pública y cuyo conocimiento o divulgación puede interesar no sólo a su Titular sino a cierto sector o grupo de personas o a la sociedad en general. (Literal g) Artículo 3 Ley 1266 de 2008). (Tomado de POLÍTICA DE PROTECCIÓN DE DATOS PERSONALES, ANT (2024).

Datos sensibles: son “aquellos que afectan la intimidad de quien es Titular o cuyo uso indebido puede generar su discriminación, tales como: los 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”. (Art. 5° Ley 1581 de 2012) (Tomado de POLÍTICA DE PROTECCIÓN DE DATOS PERSONALES, ANT (2024).

Diccionario de Datos: Documento en el que los campos de datos son documentados, describiendo el o los archivos del procesamiento de información particular y agrega descripciones precisas para facilitar su entendimiento y trazabilidad en el futuro. Idealmente debe tener los siguientes campos: Nombre del Archivo, Descripción del Archivo, Nombre del Campo, Descripción del Campo, Llave Primaria, Llave Foránea, Campo obligatorio, Dominio o lista de valores, Tipo de datos, Longitud del campo, Regla de validación (en lenguaje natural), Observaciones.

Finalidad: Es la justificación del para qué van a ser utilizados los datos personales, agregados, internos o externos de naturaleza confidencial o con origen en datos abiertos.

Información: Es cualquier variable, cuantitativa o cualitativa que surgió de un procedimiento interno o externo. Procesos internos como procesos administrativos, elevados o no a registros administrativos estandarizados, así como producto de diversos procesos rutinarios o coyunturales de recolección de información. Procesos externos como registros administrativos de otras entidades.

Información Alfanumérica: Es un conjunto organizado de datos procesados, referente a caracteres alfabéticos y/o numéricos.

	INSTRUCTIVO	PROCESAMIENTO DE INFORMACIÓN OBSERVATORIO DE TIERRAS RURALES	CÓDIGO	INTI-I-005
	ACTIVIDAD	ANÁLISIS DE INFORMACIÓN	VERSIÓN	1
	PROCESO	INTELIGENCIA DE LA INFORMACIÓN	FECHA	11/04/2025

Información Documentada: Información que una organización tiene que controlar y mantener, y el medio que la contiene.

Instituto Colombiano Agropecuario (ICA): Vela por la sanidad agropecuaria.

Investigador(a): Es el coequipero del analista de datos, especialmente tratándose de ejercicios de procesamiento de información producida de forma interna, ya sea de forma cotidiana a través de los registros administrativos como de forma ocasional producto del encargo por demanda de un estudio técnico de parte del Director(a) o alguna misional de la ANT. Este rol deberá ser desarrollado por el documento del Sistema Integrado de Gestión Manual de Metodología de Investigación.

Operativo de Recolección de Información: Es el proceso por el cual una serie de contratistas y funcionarios de la ANT adelantan labores coyunturales, es decir ocasionales, de recolección de información en campo, desde las UGT y en comisiones con el fin de validar, aplicar y socializar resultados de aplicación de instrumentos de recolección de información primaria.

Pivote - en bases de datos es una técnica que permite transformar filas de datos en columnas. Esto se hace para organizar los datos de manera que sean más fáciles de leer y analizar.

Steward data - profesional que se encarga de la gestión de datos, su calidad, integridad y seguridad. También se le conoce como administrador de datos.

MINISTERIO DE AGRICULTURA: Ministerio de Agricultura y Desarrollo Rural: Encargado de formular políticas para el desarrollo agropecuario y rural.

Segmentación - La segmentación en estadística es una técnica que consiste en dividir un conjunto de datos en subconjuntos o segmentos. Se puede usar para organizar datos, filtrarlos, o comprender fenómenos.

SEN – Sistema Estadístico Nacional.

3.2. Siglas

C-OTR: Coordinación del OTR

OTR: Observatorio de Tierras Rurales

ANT: Agencia de Tierras Rurales

	INSTRUCTIVO	PROCESAMIENTO DE INFORMACIÓN OBSERVATORIO DE TIERRAS RURALES	CÓDIGO	INTI-I-005
	ACTIVIDAD	ANÁLISIS DE INFORMACIÓN	VERSIÓN	1
	PROCESO	INTELIGENCIA DE LA INFORMACIÓN	FECHA	11/04/2025

4. GENERALIDADES

4.1. MARCO DE REFERENCIA

Data Management Body of Knowledge o Cuerpo de Conocimiento en Administración de datos, constituye un campo y a la vez es una guía exhaustiva desarrollada por DAMA (The Global Data Management International: <https://www.dama.org/cpages/body-of-knowledge>), que proporciona un marco de referencia para la gestión de los datos en todo tipo de organizaciones. Este cuerpo de conocimiento detalla algunas de las mejores prácticas y principios de la gestión de los datos y constituye una de las bases del gobierno de datos en la ANT y el OTR.

El DAMA-DMBOK identifica algunas áreas funcionales clave en la gestión de datos, tales como su gobernanza, arquitectura, calidad, gestión de datos maestros y de referencia, así como la seguridad de los mismos. Cada área funcional tiene su lugar en el gobierno de datos y corresponden al menos con un rol dentro de cada organización que tenga un buen gobierno de datos, ya que con esto se busca coordinar y optimizar todas las dependencias, en el caso de la ANT, las direcciones y subdirecciones para garantizar que los datos sean un activo empresarial estratégico.

El framework, o marco de referencia, PREC es un acrónimo utilizado en el contexto del gobierno y administración de datos y hace referencia a la gestión de múltiples proyectos de procesamiento de información. Aunque el significado específico de cada letra puede variar dependiendo del autor o del contexto en el cual se utilice el framework, lo usual es que cada letra signifique lo siguiente:

P: Proyectos – A través de los cuales se realiza la identificación, planificación y gestión de proyectos de datos dentro del marco de gobierno de datos.

R: Roles – definición de roles y responsabilidades para el gobierno organizacional de datos, asegurando la correcta gestión y uso de la información. En este documento se plantea que los roles en el marco de un buen Gobierno de Datos de la ANT deben ser: el CDO, el AD, el Data Quality Manager DQM, SD, Steward Datos y Data Analyst.

E: Estructuras – diseño y mantenimiento de la estructura organizacional y técnica necesaria para soportar el gobierno de datos de la ANT.

C: Calidad – implementación de procesos y estándares para asegurar la calidad, consistencia y confiabilidad de los datos.

5. NORMATIVIDAD APLICABLE

A continuación, se hace un llamado a la normatividad relacionada con el Observatorio de Tierras Rurales y de otra parte la que respalda las normas para el procesamiento de información:

Ley 160 de 1994: “Por la cual se crea el Sistema Nacional de Reforma Agraria y Desarrollo Rural Campesino, se establece un subsidio para la adquisición de tierras”, entre otros.

	INSTRUCTIVO	PROCESAMIENTO DE INFORMACIÓN OBSERVATORIO DE TIERRAS RURALES	CÓDIGO	INTI-I-005
	ACTIVIDAD	ANÁLISIS DE INFORMACIÓN	VERSIÓN	1
	PROCESO	INTELIGENCIA DE LA INFORMACIÓN	FECHA	11/04/2025

Ley 2294 de 2023: Que es la misma ley que establece el Plan Nacional de Desarrollo 2022-2026 "Colombia Potencia Mundial de la Vida", termina realizando unas modificaciones en el relacionamiento del SEN y el gobierno de datos, especialmente para las organizaciones públicas.

Decreto 2363 de 2015: Por medio del cual se crea la Agencia Nacional de Tierras. Dentro de las funciones de la Agencia Nacional de Tierras, el numeral 3 del artículo 4°, se tiene "Implementar el observatorio de tierras rurales para facilitar la comprensión de las dinámicas del mercado inmobiliario, conforme a los estudios, lineamientos y criterios técnicos definidos por la Unidad de Planificación de Tierras Rurales, Adecuación de Tierras y Usos Agropecuarios – UPRA – y adoptados por el Ministerio de Agricultura y Desarrollo Rural", así como "Incorporar a los sistemas información administre la Agencia la información que genere el Observatorio Tierras Rurales" según el numeral 4 del artículo 17.

Así mismo el numeral 5 del art. 4 menciona que la ANT debe "apoyar la identificación física y jurídica de las tierras, en conjunto con la autoridad catastral, para la construcción del catastro multipropósito".

Decreto 902 de 2017: "Por el cual se adoptan medidas para facilitar la implementación de la Reforma Rural Integral contemplada en el Acuerdo Final en materia de tierras, específicamente el procedimiento para el acceso y formalización y el Fondo de Tierras".

Resolución 383 de 2019 del MADR: "por medio de la cual se adoptan los lineamientos para el diseño, implementación, administración y funcionamiento del Observatorio de Tierras Rurales a cargo de la Agencia Nacional de Tierras y se dictan otras disposiciones."

La finalidad de las modificaciones estaba en el sentido de buscan fortalecer la gestión y calidad de la información estadística en el país.

Al consolidar el SEN se busca garantizar una producción de estadísticas oficiales con alta calidad, con pertinencia y en tiempos oportunos. La idea es que se coordine entre las entidades públicas productoras de datos y el Departamento Administrativo Nacional de Estadística (DANE). Este último sería el ente rector del SEN. Se aseguraría así la coherencia y comparabilidad de las estadísticas nacionales.

Otro aspecto importante de las modificaciones tiene que ver con la implementación de **buenas prácticas en el "Gobierno de Datos"** mediante la obligatoriedad de las organizaciones públicas de adoptar estándares internacionales y buenas prácticas en los procesos de gestión de datos. Esto incluye tanto lo relativo a la implementación de las políticas de calidad, como a transparencia y el acceso a la información que debe seguir intensificándose, todo con el objetivo de conseguir mejoras en la confianza que los ciudadanos tienen en las estadísticas oficiales nacionales y facilitar así la toma de decisiones basadas en datos.

Así mismo esta ley versa sobre la **Optimización de Registros Administrativos de tal manera que se consiga** impulsar un uso eficiente de los mismos reconociéndolos como fuente primaria de información estadística. Así se minimiza la duplicidad, incluso los esfuerzos múltiples y reducir costos en la recolección de datos, promoviéndose de esta manera una reutilización de la información existente dentro de las entidades públicas.

	INSTRUCTIVO	PROCESAMIENTO DE INFORMACIÓN OBSERVATORIO DE TIERRAS RURALES	CÓDIGO	INTI-I-005
	ACTIVIDAD	ANÁLISIS DE INFORMACIÓN	VERSIÓN	1
	PROCESO	INTELIGENCIA DE LA INFORMACIÓN	FECHA	11/04/2025

Otro tema que se toca allí es el de **Capacitación y Desarrollo de Capacidades**, puesto que es importante el entrenar al personal de las organizaciones públicas sobre gestión de datos y la producción estadística. Entonces, la ley promueve programas de formación y actualización de tal forma que se asegure que los funcionarios tengan las habilidades para manejar y analizar datos de manera efectiva.

Dicha ley, en relación con el SEN, también plantea la necesidad de un enfoque Diferencial y Participación Ciudadana. Esto se traduce en la incorporación de enfoques diferenciales en la producción de estadísticas, considerando las particularidades de diversos grupos poblacionales. Adicionalmente plantea el fomento de la participación ciudadana en la definición de necesidades de información y en la evaluación de la calidad de los datos producidos.

Ley 1955 de 2019: Modificación de la Ley 1753 de 2015, relativa al Sistema Estadístico Nacional (SEN). En su artículo 155 modifica el artículo 160 de la ley 1753 sobre el SEN...

Ley 1753 de 2015: Creación del SEN. (Plan Nacional de Desarrollo: Todos por un Nuevo País)

Ley 152 de 1994: Establecimiento de procedimientos para la elaboración de planes de desarrollo.

Decreto 2404 de 2019: Reglamentó el SEN en su momento, modificando el Decreto 1170 de 2015.

Decreto 1743 de 2016: Reglamento el artículo 160 de la Ley 1753 de 2015, que crea el SEN.

Decreto 1499 de 2017: Modificación del Decreto 1083 de 2015 sobre el Sistema de Gestión.

Resolución 1418 de 2017: Norma técnica de calidad del proceso estadístico.

Resolución 1419 de 2017: Lineamientos para el proceso estadístico en el SEN.

Norma Técnica de Calidad (NTC PE 1000:2020): Requisitos de calidad estadística para el proceso estadístico en Colombia. La norma NTC PE 1000 expedida en el año 2017 por el Departamento Administrativo Nacional de Estadística (DANE) en colaboración con el Instituto Colombiano de Normas Técnicas y Certificación (ICONTEC), estableció directrices específicas para el gobierno y la gestión de datos dentro del sector público, así como en el privado. Esta norma se centra en la necesidad que las organizaciones gestionen su sistema de información y datos de manera que se garanticen: confiabilidad, accesibilidad y transparencia. Debe destacarse que la NTC PE 1000 da mucha importancia de la alineación de los procesos de gestión de datos con las políticas de protección de datos personales y el cumplimiento de las normativas de privacidad. las organizaciones colombianas no solo adherirse a prácticas de gestión de datos responsables y transparentes, sino también establecer un marco común para la colaboración interinstitucional. Uno de los aspectos clave de la NTC PE 1000 es la necesidad de tener un enfoque centrado en la protección de la privacidad, lo que implica una gestión cuidadosa de los datos personales y sensibles.

	INSTRUCTIVO	PROCESAMIENTO DE INFORMACIÓN OBSERVATORIO DE TIERRAS RURALES	CÓDIGO	INTI-I-005
	ACTIVIDAD	ANÁLISIS DE INFORMACIÓN	VERSIÓN	1
	PROCESO	INTELIGENCIA DE LA INFORMACIÓN	FECHA	11/04/2025

DANE (2017) Código Nacional de Buenas Prácticas del Sistema Estadístico Nacional¹.

DANE (2018) Guía para la anonimización de bases de datos en el Sistema Estadístico Nacional. Implica directrices a manera de manual de privacidad diferencial para intentar anonimizar bases de datos².

5.1. DIRECTRICES INTERNACIONALES

Informes del Foro Mundial de Datos de la ONU (World Data Forum).

Son documentos claves de resultado en la Declaración de Ciudad del Cabo en 2017, la Declaración de Dubái (2018) y la Declaración de Berna (2021). Asimismo, próximamente de la Declaración de la Ciudad de Medellín 2024. Estos documentos resultan relevantes puesto que enfatizan la importancia del fortalecimiento de la infraestructura de datos a nivel mundial para apoyar los avances en materia de ODS. Incluyen llamados a la acción para mejorar la calidad, transparencia y ética en la producción y uso de datos. La Declaración de Berna, en particular, hace un fuerte énfasis en la necesidad de colaboración entre actores gubernamentales, privados y académicos para crear un ecosistema de datos inclusivo y ético.

Principios Fundamentales de las Estadísticas Oficiales (ONU, 1994). Es un documento relevante puesto que dichos principios fundamentales de las Estadísticas Oficiales surgidas de manera multilateral proporcionan la base ética y metodológica para la recolección y diseminación de datos oficiales, con lo cual se podría garantizar que las estadísticas serán: a) imparciales, b) precisas y c) transparentes. Es destacable que para el año 2014, dichos principios fueron reforzados y adaptados al contexto de los ODS, para poner énfasis en la importancia de la independencia estadística y la confianza pública, revalorando el carácter técnico de las mismas y su necesidad de independencia de los gobiernos coyunturales³.

Guía para la transversalización de la perspectiva de género en la producción estadística.

Es otro documento relevante puesto que estableció los lineamientos específicos para garantizar que las estadísticas recojan de forma adecuada las realidades de género y de esta forma se pueda promover la equidad y aportar al cumplimiento de los ODS relacionados.

(CEPAL, 2020) https://cea.cepal.org/12/sites/cea12/files/presentations/cea12_guia-transversalizacion-perspectiva-genero-mexico.pdf

Marco de Monitoreo de los ODS (ONU, 2015)

Este documento es relevante en la medida que como su nombre indica, proporciona un marco fundamental para guiar a los países en sus esfuerzos por avanzar en los 17 Objetivos de Desarrollo Sostenible, estableciendo una serie de indicadores y metodologías para evaluar el progreso en dichos esfuerzos. Así se destaca la importancia

¹

https://www1.funcionpublica.gov.co/documents/34645357/0/Codigo_Nacional_Buenas_Practicas_30072018_final+%281%29.pdf/77c6f040-78ca-b0e7-5f13-71f41387baa1?t=1635180858802

²

<https://www1.funcionpublica.gov.co/documents/34645357/0/Gu%C3%ADa+para+la+anonimizaci%C3%B3n+de+bases+de+datos+DANE+%281%29.pdf/08752339-2839-f188-fcf2-6b1117063785?t=1635180920895>

³ Para consultar: https://unstats.un.org/unsd/dnss/hb/S-fundamental%20principles_A4-WEB.pdf

	INSTRUCTIVO	PROCESAMIENTO DE INFORMACIÓN OBSERVATORIO DE TIERRAS RURALES	CÓDIGO	INTI-I-005
	ACTIVIDAD	ANÁLISIS DE INFORMACIÓN	VERSIÓN	1
	PROCESO	INTELIGENCIA DE LA INFORMACIÓN	FECHA	11/04/2025

de contar con datos de gran calidad, accesibles y desagregados para conseguir una buena toma de decisiones acorde con las realidades de los territorios.

https://www.undp.org/sites/g/files/zskgke326/files/2023-02/undp_co_pub_Marco_Global_Nacional_ODS_Esp_Feb21_2023.pdf

Lo cual llevado al contexto del OTR implica visionar la producción de indicadores relativos a los avances en ODS del Sector Rural en el marco de la Mesa Sectorial de Estadística Campesina, coordinada por el DANE.

Carta de Principios de Datos Abiertos para la Alianza Internacional de Datos Abiertos

Es una carta relevante porque al igual que los Principios Fundamentales de las Estadísticas Oficiales establece principios para que los datos sean abiertos y sean accesibles, fomentando la transparencia, el accountability o rendición de cuentas de la actuación pública basada en evidencia, la responsabilidad pública y el acceso a una información que permita el empoderamiento y participación ciudadana. Si bien esta carta se enfoca en los datos abiertos y no siempre, ni todo dato, en el marco del habeas data podrá ser publicable desde el OTR y por extensión desde la ANT, es de rescatar sus estándares éticos, relevantes para el gobierno de datos del OTR y la ANT.

(Open Data Charter, 2015) https://transparencia.gob.gt/wp-content/uploads/Carta_Internacional_de_Datos_Abiertos2015.pdf

Ética en la Inteligencia Artificial y el Gobierno de Datos

Relevante en tanto sugiere un marco ético que abarca principios y recomendaciones sobre la ética en el uso de datos, especialmente en el contexto de IA. Resalta la importancia de la privacidad, la transparencia y el respeto a los derechos humanos en el gobierno de datos, orientando su implementación con un enfoque ético global.

(UNESCO, 2021) Recuperado de <https://www.unesco.org/es/legal-affairs/recommendation-ethics-artificial-intelligence>

Modelo de Gobernanza de la Infraestructura de Datos para el Estado Colombiano

Acción A1: Plan Nacional de Infraestructura de Datos (PNID)

PolicyLab, Departamento Administrativo de Presidencia de la República, Departamento Nacional de Planeación, Ministerio de Tecnologías de Información y Comunicaciones (2022)

https://gobiernodigital.mintic.gov.co/692/articles-273751_modelo_gobernanza_infraestructura_datos.pdf

6. INSTRUCCIONES PARA EL PROCESAMIENTO DE INFORMACIÓN – OTR

6.1. Paso 01 - de Comprensión del Problema

La implementación del Paso 1, fundamentado en la metodología **CRISP-DM** que se centra en la fase 1 de *Comprensión del problema y negocio*. Este paso garantiza que el Observatorio alinee sus actividades con un marco claro y trazable desde el principio. Donde como mínimo deben estar presente un Analista de Datos “Encargado” y

	INSTRUCTIVO	PROCESAMIENTO DE INFORMACIÓN OBSERVATORIO DE TIERRAS RURALES	CÓDIGO	INTI-I-005
	ACTIVIDAD	ANÁLISIS DE INFORMACIÓN	VERSIÓN	1
	PROCESO	INTELIGENCIA DE LA INFORMACIÓN	FECHA	11/04/2025

un investigador, pero idealmente deben estar al tanto un segundo Analista de Datos que tendrá la responsabilidad de ser el “par” en ciertos pasos como podrá verificar el lector del presente instructivo, también conviene que esté presente la coordinación del OTR.

En esta etapa, el analista de datos y los investigadores deben comenzar por definir los objetivos de los procesos a realizar tras comprender el problema correspondiente a la demanda de información, lo que implica identificar las necesidades de información de la solicitud recibida. Esto conlleva a detallar las variables clave necesarias para cumplir el propósito del análisis, junto con las posibles fuentes disponibles, diferenciando entre registros administrativos, datos abiertos y otros insumos secundarios. Inclusive la solicitud a otras dependencias de la ANT o la implementación de operativos de recolección de información primaria.

Idealmente debería crearse un primer esquema sobre el problema en términos de flujos de datos.

Producto esperado: Acta o Documento de requerimientos, objetivos y alcance del procesamiento de información.

6.2. Paso 02 Entendimiento de los datos asociados

Correspondiente en la **metodología CRISP -DM**, a la fase 2 de *Entendimiento de los datos*. Labor del Analista de Datos Encargado. Esta fase se centra en el análisis del origen de los datos de aquellas variables identificadas como asociadas al problema y que serán utilizadas y inicia el trazo de la ruta o flujo de tratamiento de datos. Para eso se puede partir de varias preguntas, como: “¿La ANT tiene banco de datos al respecto del problema específico a tratar? ¿cómo serán accedidos dichos datos? ¿Cuáles y cuántas fuentes de datos serán empleadas? ¿Cuáles serán los formatos de los datos? ¿Los datos están estructurados?”. A partir de las anteriores preguntas se pasará a la recolección de los datos, teniendo especial cuidado que ninguna información importante quede fuera.

Idealmente debería crearse un primer esquema sobre el problema en términos de flujos de datos.

Productos: a. Forma con cuadro de lista de variables recibidas y/o requeridas para acometer una respuesta efectiva a la demanda de información requerida (Ver anexo 0 de este documento para tomarlo como modelo).

6.3. Paso 03 delimitación de diccionario y preparación de datos.

Corresponde en la metodología CRISP-DM, a la fase 3, de preparación de los datos. Este paso está centrado en la estructuración de la información que será utilizada en el procesamiento analítico. Resulta crucial puesto que establece los fundamentos para la trazabilidad, la calidad y la consistencia de los datos.

La delimitación del diccionario de datos permite estandarizar las variables, definiendo su tipo, formato, dominios válidos y relaciones con otras variables. Así se garantiza parte de la interoperabilidad de los datos en los sistemas del Observatorio de Tierras Rurales (OTR) y facilita la certificación de los registros administrativos, conforme a la **NTC PE 1000:2020** y la **Resolución 1418 de 2017**.

	INSTRUCTIVO	PROCESAMIENTO DE INFORMACIÓN OBSERVATORIO DE TIERRAS RURALES	CÓDIGO	INTI-I-005
	ACTIVIDAD	ANÁLISIS DE INFORMACIÓN	VERSIÓN	1
	PROCESO	INTELIGENCIA DE LA INFORMACIÓN	FECHA	11/04/2025

Actividades clave en este paso son:

- El proceso inicial con la elaboración o actualización del diccionario de datos, donde se definen los nombres de los campos, su tipo de dato (numérico, texto, booleano, etc.), las claves primarias y foráneas, y se documentan las reglas de validación. Con esto se estandarizan los insumos y asegura la consistencia de los registros provenientes de múltiples fuentes (internas, secundarias o abiertas). La validación incluye la detección de valores nulos, duplicados, inconsistencias lógicas y errores de formato.
- Actividades de limpieza y transformación de los datos. Que incluyen, por ejemplo: la imputación de valores faltantes o la normalización de textos, la codificación de aquellas variables categóricas y la creación de nuevas variables derivadas, pivote o instrumentales en el marco del modelo de solución.
- Se genera un flujo de procesamiento que especifica cada etapa de la transformación (siempre y cuando se disponga de tiempo).

Producto esperado: Documento que contenga construcción, actualización o reconocimiento de Diccionario de datos; flujo de procesamiento y protocolo de validación inicial de calidad.

6.4. Paso 4 Modelado y Análisis Preliminar o de muestras

Se enfoca en la construcción, ajuste y validación de modelos analíticos que permitan extraer patrones, predecir comportamientos o en su defecto realizar clasificaciones a partir de los datos procesados en el OTR. Este paso resulta fundamental para transformar la información en conocimiento *estratégico* y aplicable a la toma de decisiones y lo realiza el Analista de datos Encargado. Recordando que corresponde este paso a la fase de **modelado** de la metodología **CRISP-DM**, el cual se centra en la creación de modelos, dependiendo de la necesidad, de carácter: a) predictivo, b) descriptivo o c) de clasificación.

Para todos los modelos se usa desde luego la estadística y las técnicas de minería de datos. Incluyendo, toda vez que sea posible precisión, validez y confiabilidad mediante pruebas y validaciones iterativas. La validez se garantiza con líneas o procesos que confirmen (pruebas) que el código o el procesamiento se está realizando como se espera y la confiabilidad se refiere a que los datos en la base de datos no resultan contradictorios.

Las actividades clave en este paso son cuatro: selección de métodos de modelado, construcción del modelo de analítica, entrenamiento y validación del modelo.

Aunque muy corta, la primera actividad consiste en la **selección del método de modelado**, que depende de la naturaleza del problema y del tipo de dato. Aquí en concreto se trata de escoger entre: aplicar algoritmos de clasificación (árboles de decisión, regresión logística), de predicción (modelos de regresión múltiple) o de agrupación (clustering y segmentación). Este paso, se realiza a veces de forma natural, pero es fundamental que en el informe final quede claro como esta alineada cada variable con los objetivos establecidos en el paso 1 - Comprensión del problema.

	INSTRUCTIVO	PROCESAMIENTO DE INFORMACIÓN OBSERVATORIO DE TIERRAS RURALES	CÓDIGO	INTI-I-005
	ACTIVIDAD	ANÁLISIS DE INFORMACIÓN	VERSIÓN	1
	PROCESO	INTELIGENCIA DE LA INFORMACIÓN	FECHA	11/04/2025

La siguiente actividad en este paso es la **construcción del modelo de analítica de datos**, usando cualquier software especializado como Python, R u otro software propietario de análisis de Big data. En resumen, aquí se definen las rutas y variables independientes o predictoras, así como la variable objetivo, dependiente o variable a explicar. Ahora bien, cuando se trata de simple agrupación no es necesario establecer estas relaciones de dependencia.

Durante la actividad de **validación del modelo**, lo que procede es la revisión de los indicadores de desempeño, tales como la precisión, la sensibilidad, la especificidad y la exactitud global. Los analistas (encargado y par) revisan las variables y estadísticos o indicadores de control del proceso para evaluar el ajuste del modelo. Si se detecta sobreajuste o subajuste se deben aplicar los cambios pertinentes.

Los resultados y productos del paso 5: principalmente un modelo analítico validado y operativo, materializado en un código o una rutina claramente establecida para el análisis, lo más automatizada posible. Este modelo debe servir para realizar predicciones, clasificaciones o agrupaciones.

Agotado lo anterior, se procede a generar un informe de evaluación del modelo, donde se presenten algunas métricas como estadísticos o indicadores que ayuden a comprender que, si el mismo tiene buen desempeño, es preciso y finalmente válido. Además, puede incluirse un sucinto análisis de posibles riesgos de aplicación e interpretación de los resultados de la implementación del modelo de analítico. Estos resultados del producto 5, contribuyen tanto a la transparencia, reproducibilidad y trazabilidad de los modelos conforme a los lineamientos de la NTC PE 1000:2020 y la Resolución 1418 de 2017.

6.5. Paso 5: Evaluación (Verificación del Proceso y Validación de la Calidad)

Este paso de evaluación corresponde con la fase del mismo nombre en la metodología CRISP -DM. Aquí deben acogerse las indicaciones que progresivamente se emitan sobre gobierno de datos desde las mesas técnicas de datos en la ANT. Fundamental para garantizar que los datos procesados cumplan con los estándares de calidad, trazabilidad y autenticidad exigidos por el Observatorio de Tierras Rurales (OTR), como parte de la ANT, y la normativa vigente, como la **NTC PE 1000:2020** y la **Resolución 1418 de 2017**.

En esta etapa, idealmente el *Analista Par*, aplica pruebas específicas de **duplicación, consistencia e integridad** para evaluar la calidad de los datos y la comenta con la coordinación del OTR. Esta verificación permite identificar errores, corregir inconsistencias y validar la completitud de la información procesada. Al finalizar esta fase, se genera un informe técnico oral o escrito, en función del tiempo con el concepto de autenticidad final de los datos.

Las actividades clave en este paso son tres: revisión de duplicidad, verificación de consistencia y pruebas de integridad.

El paso 5 inicia entonces con la **revisión de duplicados** que consiste en aplicar pruebas para identificar si existen registros repetidos. En la práctica esto implica añadir alguna o pocas líneas al código o códigos. Ahora bien, esto se puede realizar mediante comparaciones de valores únicos (por ejemplo, alguna llave o identificador, típicamente

	INSTRUCTIVO	PROCESAMIENTO DE INFORMACIÓN OBSERVATORIO DE TIERRAS RURALES	CÓDIGO	INTI-I-005
	ACTIVIDAD	ANÁLISIS DE INFORMACIÓN	VERSIÓN	1
	PROCESO	INTELIGENCIA DE LA INFORMACIÓN	FECHA	11/04/2025

nombrado "ID" único, o por combinación de nombre y fecha o nombre y tipo de documento) para garantizar la unicidad de las observaciones. Es fundamental entender que los duplicados también deben ser objeto de análisis, especialmente si se trata de bases de datos producidas internamente, para corregir los procesos de construcción y reporte de estas. Dependiendo de la causa dichos duplicados, en el procesamiento de información se puede: eliminar, consolidar o dejar anotación de la justificación de su permanencia en la base de datos.

La segunda actividad del paso 5 es la **verificación de consistencia**, el cual en esencia es una revisión de las relaciones lógicas entre variables de la base de datos, especialmente si son clave para el procesamiento. Una inconsistencia entre variables en una o varias observaciones podría implicar la necesidad de eliminación de observaciones, siempre y cuando no se afecte la solución del problema. A manera de ejemplo, en registros de predios rurales de adjudicaciones, se comprueba que la "superficie adjudicada" no supere la superficie total disponible y que las fechas de los eventos no sean inconsistentes (otro ejemplo: una fecha de adjudicación posterior a la de cierre). Este análisis se realiza mediante la implementación de **reglas de consistencia** que en la práctica son comparaciones entre variables para establecer si existen igualdades o desigualdades según sea el caso entre dos variables. Ahora bien, no siempre es posible hacer verificación de consistencia, pero siempre que sea posible debe realizarse.

La tercera y última actividad del paso 5 es la **prueba de integridad** que asegura la completitud de la base de datos, dicho de otra manera, la verificación que no se tienen valores faltantes en las variables críticas y que las variables llave como los ID estén completas. Un ejemplo, verificar que todos los registros tengan asignado un **ID único, o con la variable en el formato indicado**, ya que pueden existir algunas observaciones con vacíos en esa columna o con un formato corrompido por error. Algunas veces, por ejemplo, puede quedar con notación científica un número que se requería completo (con todos sus caracteres) probablemente por inadecuada gestión o definición de la longitud (número de caracteres) de la variable previamente. Otro ejemplo, que en las columnas esenciales como "superficie adjudicada" o "nombre del propietario" estén demasiados datos vacíos. Otra forma de realizar pruebas de integridad es verificando las claves foráneas (previas al procesamiento que realiza el OTR), es decir, y que haya coherencia entre las diferentes tablas o bases de datos.

El resultado de este paso será un **Informe de control de calidad y trazabilidad de los datos procesados**, el cual podría incluir las siguientes secciones:

- **Relación de duplicidad de registros:** Incluye enlaces a las bases de datos de registros duplicados detectados, descripción de lo que se realizó para corregir o evitar la alteración de la eficacia del modelo analítico por la existencia de estos duplicados. Es decir, explicar si se eliminaron o consolidaron.
- **Relación de consistencia:** Un escrito sucinto que describa las reglas aplicadas y la proporción de registros que cumplieron cada regla.
- **Relación de integridad:** Un escrito sucinto del nivel (proporción) de completitud de los campos críticos, así como de la validación de las variables llaves y las pruebas de integridad.
- **Concepto de autenticidad final** de los datos, que constituye la validación formal de la calidad alcanzada. Este concepto debería emitirse de acuerdo con los estándares de la **NTC PE 1000:2020**, con base en la trazabilidad de todas las verificaciones realizadas.

	INSTRUCTIVO	PROCESAMIENTO DE INFORMACIÓN OBSERVATORIO DE TIERRAS RURALES	CÓDIGO	INTI-I-005
	ACTIVIDAD	ANÁLISIS DE INFORMACIÓN	VERSIÓN	1
	PROCESO	INTELIGENCIA DE LA INFORMACIÓN	FECHA	11/04/2025

6.6. Paso 6 implementación en pleno o escalabilidad

Corresponde a la fase 6 de la metodología CRISP – DM, denominada “implementación o desarrollo”. Aquí se retoma el modelo analítico ya refinado y si fue el caso, se aplica ya no para una muestra sino para el total de unidades de observación que la demanda de información requiere. Dependiendo de las necesidades de la tarea esto puede requerir una work station (computador) con gran capacidad tanto en términos de memoria RAM como de procesador de alta frecuencia. Debe asegurarse desde la coordinación del OTR, en dialogo con dirección DGOSP e Infraestructura tecnológica de la ANT que los analistas de datos encargado y Analista par tengan el computador propicio para evitar prolongaciones excesivas del tiempo de procesamiento de información.

6.7. Paso 7 Elaboración del Informe Final de Procesamiento

El Paso 7 - Creación de un documento integral que sintetiza los resultados, análisis, conclusiones y recomendaciones obtenidos durante el procesamiento de información o aplicación del modelo analítico. El objetivo es comunicar los hallazgos y sugerir implicaciones para la toma de decisiones estratégicas más allá de contar el paso a paso.

El informe puede presentarse de forma **escrita** o, en casos excepcionales (por tiempo), de forma **oral**, apoyado con una **presentación visual sucinta apoyada en gráficos y mapas**. La presentación debe facilitar la comprensión de los resultados clave y permitir la interacción con los usuarios finales.

El Analista de Datos es el encargado de redactar el informe final, que debe incluir mínimo las siguientes secciones: introducción, objetivos, metodología y/o modelo analítico implementado, análisis de resultados, conclusiones y recomendaciones. La redacción se debe facilitar la comprensión de los hallazgos, garantizando la relación directa con los objetivos planteados inicialmente. Como complemento se puede generar una presentación que incluya **gráficos, mapas y tablas**.

El Analista de Datos Par (si por tiempo hubo oportunidad de asignar este rol y este realizó el trabajo) debería incluir un resumen de la revisión y control de calidad basado en los pasos que el presente instructivo de procesamiento indica.

Entre todos los participantes debe realizarse un flujograma sintético del procesamiento realizado para facilitar la comprensión de este.

Productos Finales: Informe final, ii) presentación ejecutiva en PowerPoint o similar, iii) gráficos, mapas y diagramas pertinentes y iv) relación de la trazabilidad de los datos, idealmente con el flujograma final.

6.8. Paso 8 Entrega de metadatos

Acorde con las directrices vigentes de gobierno de datos se realiza entrega de metadatos. Para el dato geográfico, se utilizan las plantillas en formato XML, (Raster, Geoservicio y vector) , basado en las normas ISO: ISO 19115, ISO 19115-2, ISO 19139 Y La Guía de implementación de Metadatos ICDE, Sugiero el Capítulo “Restricciones del

	INSTRUCTIVO	PROCESAMIENTO DE INFORMACIÓN OBSERVATORIO DE TIERRAS RURALES	CÓDIGO	INTI-I-005
	ACTIVIDAD	ANÁLISIS DE INFORMACIÓN	VERSIÓN	1
	PROCESO	INTELIGENCIA DE LA INFORMACIÓN	FECHA	11/04/2025

recurso” en donde se establecen las limitaciones de usos dependiendo del conjunto de datos y la información producida (Página 31).

6.9. Consideraciones de Excepcionalidad

Recolección de información primaria: Si la información requerida no está disponible en bases de datos internas ni externas, se habilita la necesidad de un operativo de recolección. Este proceso implica la activación del paso de diseño de instrumentos de recolección y la implementación de operativos en campo, que se ejecutarán antes del Paso 3. Esto debería realizarse acorde con el manual de metodología de Investigación del OTR.

Acceso a información confidencial: Para los casos en que se deba acceder a registros confidenciales, se activan protocolos adicionales de autorización de acceso, revisión de confidencialidad y medidas de seguridad de acceso. Todos deberían estar en el marco de la Política general de seguridad y privacidad de la información

Procesamiento de grandes volúmenes de datos: Si el volumen de datos supera la capacidad operativa del equipo, se activa el apoyo técnico de la SSIT e infraestructura.

ANEXO 0: Ejemplo de paso 0

UNIDAD DE ANALISIS	VARIABLE (IDENTIFICACIÓN)	Disponible	FUENTE	ULTIMO DATO	PERIODICIDAD	FORMATO	INFORMACIÓN	Proceso en caso de faltante
predio	Valor total terreno	Sí	IGAC	mar-24	Trimestral	.XLSX		
	Valor/ha	No	IGAC					Dividir el valor del terreno entre el área de terreno del predio
	Área terreno	No	IGAC					Cruzar con R2, R1
	NPN	No	IGAC					Cruzar con R2, R1
	Coordenadas	No	IGAC					Cruzar con R2, R1
	Características zona fisiográfica o ZHF	No	IGAC					Cruce de capas
	FMI	Sí	IGAC					
predio	Valor total terreno	Sí	Fincaraiz	mar-24	Web scraping	.XLSX		
	Valor/ha	No	Fincaraiz					Dividir el valor entre el área
	Área terreno	Sí	Fincaraiz					Cruzar con R2 o R1
	NPN	No	Fincaraiz					Cruzar con capa geográfica predial
	Coordenadas	Sí	Fincaraiz					
	Características zona fisiográfica o ZHF	No	Fincaraiz					Cruce de capas geográficas
	FMI	No	Fincaraiz					Cruce SNR

HISTORIAL DE CAMBIOS		
Fecha	Versión	Descripción
09/04/2025	01	Se estandarizan los pasos para el procesamiento de información como insumo de estudios e investigaciones a realizar por el Observatorio de Tierras Rurales de la Agencia Nacional de Tierras.

	INSTRUCTIVO	PROCESAMIENTO DE INFORMACIÓN OBSERVATORIO DE TIERRAS RURALES	CÓDIGO	INTI-I-005
	ACTIVIDAD	ANÁLISIS DE INFORMACIÓN	VERSIÓN	1
	PROCESO	INTELIGENCIA DE LA INFORMACIÓN	FECHA	11/04/2025

APROBACIÓN				
	NOMBRE	CARGO	FIRMA	FECHA
ELABORÓ	Eduardo Reina	Contratista Dirección Acceso a Tierras	ORIGINAL FIRMADO	22/12/2024
REVISÓ	Jhon Alejandro Noriega Talero	Líder Observatorio de Tierras Rurales	ORIGINAL FIRMADO	20/03/2025
APROBÓ	Andrea Silva Porras	Directora Dirección de Gestión de Ordenamiento Social de la Propiedad	ORIGINAL FIRMADO	11/04/2025

La copia, impresión o descarga de este documento se considera COPIA NO CONTROLADA y por lo tanto no se garantiza su vigencia.

La única COPIA CONTROLADA se encuentra disponible y publicada en la página Intranet de la Agencia Nacional de Tierras.