

CARTILLA DE REPOSITORIO  
DE DATOS DE INVESTIGACIÓN

# PAPYRUS

## **Universidad Santo Tomás**

Fray Álvaro José ARANGO RESTREPO, O.P.  
Rector General

Fray Mauricio Antonio CORTÉS GALLEGO, O.P.  
Vicerrector Académico General

Fray Hender Alveiro RODRÍGUEZ PÉREZ, O.P.  
Vicerrector Administrativo y Financiero General

## **CARTILLA DE REPOSITORIO DE DATOS DE INVESTIGACIÓN PAPYRUS**

Centro de Recursos para el Aprendizaje y la Investigación - CRAI  
Abril 2025

Fray Luis Antonio ALFONSO VARGAS, O.P.  
Director General

María del Pilar Florián Escobar  
Directora Técnica

Luz Marina Páez  
Karen Sheraldine Cristancho Luna

## PRESENTACIÓN

El Plan de Gestión de Datos es un documento que describe el ciclo de vida de los datos recopilados, generados y procesados durante el proyecto de investigación. Se especifica la recolección, organización, descripción, almacenamiento, difusión, reutilización y preservación de los datos.

Teniendo en cuenta la importancia de la gestión de Datos de Investigación (GDI), la Universidad Santo Tomás como miembro del Consorcio Colombia, suscribe el Repositorio de Gestión de Datos Papyrus, “repositorio de carácter multidisciplinar, que alberga los conjuntos de Datos científicos de proyectos de investigación de las Instituciones colombianas que pertenecen al Consorcio, con el fin de darles a los datos, visibilidad, garantizar su preservación y facilitar su acceso, recuperación y reutilización a través de la custodia y gestión centralizada, dando cumplimiento a estándares internacionales”(Consortia, 2022).

## Contenido



Búsqueda avanzada

☒ Dataverses (6)

☒ Datasets (0)

☒ Files (0)

### Palabra clave

[Gestión de Datos \(3\)](#)

[Producción \(2\)](#)

[Ciencia Abierta \(1\)](#)

### Año de publicación

[2025 \(1\)](#)

1 de 6 Resultados

↑↓ Orden ▼

#### Directrices de acceso abierto (Crai Santoto)



2025

Cartilla de Repositorio de Datos de Investigación PAPYRUS



#### Principios que rigen la GDI en el CRAI SANTOTO (Crai Santoto)



2025

Cartilla de Repositorio de Datos de Investigación PAPYRUS



#### Compromisos (Crai Santoto)



2025

Cartilla de Repositorio de Datos de Investigación PAPYRUS



#### Contenidos y propiedad intelectual (Crai Santoto)



2025

Cartilla de Repositorio de Datos de Investigación PAPYRUS



#### Aspectos legales de los conjuntos de datos (Crai Santoto)



2025

Cartilla de Repositorio de Datos de Investigación PAPYRUS



#### Información adicional (Crai Santoto)



2025

Cartilla de Repositorio de Datos de Investigación PAPYRUS



#### Glosario (Crai Santoto)



2025

Cartilla de Repositorio de Datos de Investigación PAPYRUS



## 8 DIRECTRICES DE ACCESO ABIERTO

☒  Dataverses (6)

☒  Datasets (0)

☒  Files (0)

### Nombre de autor

[Crai Santoto](#)

### Palabra clave

[Acceso abierto \(2\)](#)

[Ciencia Abierta \(1\)](#)

[Gestión de Datos \(3\)](#)

[Producción académica \(1\)](#)

[Infraestructura \(1\)](#)

[Innovación \(1\)](#)

[Información \(3\)](#)

[Preservación información](#)

[Producción \(2\)](#)

[Producción académica \(1\)](#)

[Publicación científica \(2\)](#)

### Año de publicación

[2025 \(1\)](#)



La Universidad Santo Tomás, a través del CRAI SANTOTO, gestiona actualmente diversas plataformas que facilitan el acceso y la visibilidad de la producción académica institucional

En el marco del desarrollo de la ciencia abierta, y en concordancia con las directrices internacionales de la UNESCO para el acceso abierto, así como con los lineamientos de Colciencias para una política nacional de ciencia abierta en Colombia, la Universidad fortalece sus líneas de acción en infraestructura física, tecnológica y de comunicaciones, así como en vigilancia tecnológica, bibliometría e innovación. Dentro de esta oferta de servicios se incluye la Gestión de Datos de Investigación (GDI)

Este proceso impulsa a los autores tomasinos a elaborar un Plan de Gestión de Datos, en el cual se describa el ciclo de vida de la información recopilada en sus investigaciones y el tratamiento que debe recibir en cada etapa de desarrollo. De igual manera, el CRAI SANTOTO se encargará de garantizar la custodia y el acceso a dichos datos, de manera que puedan ser utilizados como insumo para nuevas producciones científicas, en el marco de la ciencia abierta, a través del Repositorio de Datos de Investigación (RDI) – Papyrus

## 8 PRINCIPIOS QUE RIGEN LA GDI EN EL CRAI SANTOTO



Para la custodia y el acceso a los conjuntos de datos de investigación, el CRAI SANTOTO establece una serie de principios que orientan la publicación, divulgación y preservación de la información a través del RDI-Papyrus

✓  [Dataverses \(6\)](#)

✓  [Datasets \(0\)](#)

✓  [Files \(0\)](#)

---

**Nombre de autor**

[Crai Santoto](#)

---

**Palabra clave**

[Accesibilidad \(1\)](#)

[Acceso abierto \(3\)](#)

[Almacenamiento de datos](#)

[Indexación \(1\)](#)

[Interoperabilidad \(2\)](#)

[Licenciamiento \(1\)](#)

[Localización \(1\)](#)

[Metadatos \(2\)](#)

[Migración de datos \(2\)](#)

[Organización de datos \(1\)](#)

[Producción científica \(2\)](#)

---

**Año de publicación**

[2025 \(1\)](#)



Los aspectos clave para su desarrollo se organizan en tres grupos:

- »» Administrativo: relacionado con la adquisición, almacenamiento y organización de los datos
- »» Preservación: referido al uso de formatos estándar, los medios de migración y las estrategias de resguardo de la información
- »» Descubrimiento: centrado en los mecanismos de búsqueda y recuperación, el acceso y el licenciamiento que permita poner los datos del repositorio a disposición de la comunidad

Con base en lo anterior, se adoptaron los siguientes principios estándar:

**Principios FAIR (Localización, Accesibilidad, Interoperabilidad y Reutilización de activos digitales):** permiten implementar procesos que aseguren que los datos sean localizables mediante metadatos, identificadores y palabras clave; que sean accesibles a través de licencias y plataformas; que logren interoperabilidad mediante el uso de estándares para indexación y registro; y que fomenten la reutilización de los conjuntos de datos como insumo para nuevas producciones científicas

**Principios DataCite:** tienen como finalidad garantizar la localización y citación de los datos de investigación, en articulación con los principios FAIR. Este proceso se lleva a cabo mediante la asignación de DOIs y metadatos específicos que facilitan el acceso a los conjuntos de datos

**Principios OpenAIRE:** establecen pautas internacionales orientadas a asegurar la accesibilidad abierta de los datos, así como la adopción de políticas de investigación que promuevan la interoperabilidad de los sistemas de información

## COMPROMISOS

✓  Dataverses (6)

✓  Datasets (0)

✓  Files (0)

### Nombre de autor

Crai Santoto

### Palabra clave

Acceso abierto (2)

Almacenamiento (1)

Conjuntos de datos (2)

Copias de seguridad (1)

Datos de investigación (1)

Investigadores (1)

Mantenimiento (1)

Preservación (3)

Semilleros de investigación

Uso de la información (2)

### Año de publicación

2025 (1)



#### Universidad Santo Tomás

- Proveer el acceso y mantenimiento a las plataformas requeridas para la custodia de los datos de investigación
- Garantizar el talento humano para el asesoramiento de almacenamiento y uso de la información contenida en las plataformas



#### Investigadores

- Diligenciar claramente la plantilla de datos y proporcionar los conjuntos de datos necesarios debidamente descritos para su depósito
- Acudir al personal del CRAI para el respectivo resguardo de la información y la permanente actualización de los datos
- Mantener copias de seguridad de los datos que permitan recuperar la información en caso de siniestros en los sistemas de información del CRAI SANTOTO
- Garantizar los documentos soporte de uso libre de la información contenida en los conjuntos de datos



#### CRAI

- Proporcionar a los diferentes grupos que conforman la línea de investigación (Investigadores, semilleros de investigación, coautores) un espacio de resguardo de información contenida en los datos de investigación para asegurar su preservación
- Brindar el asesoramiento (Acompañamiento, formación) necesario a los investigadores en la subida, selección, y resguardo de sus conjuntos de datos
- Motivar la promoción y uso de los conjuntos de datos para la producción de nuevo conocimiento en la comunidad de la CRAI SANTOTO

✓  [Dataverses \(6\)](#)

✓  [Datasets \(0\)](#)

✓  [Files \(0\)](#)

**Nombre de autor**

[Crai Santoto](#)

**Palabra clave**

[Acceso \(1\)](#)

[Bases de datos \(2\)](#)

[Citación \(1\)](#)

[Conjuntos de datos \(2\)](#)

[Correlación de datos \(1\)](#)

[Investigación académica \(1\)](#)

[Permanencia \(2\)](#)

[Propiedad de los datos \(1\)](#)

[Propiedad intelectual \(3\)](#)

[Seguridad \(3\)](#)

[Software \(1\)](#)

**Año de publicación**

[2025 \(1\)](#)



- Generar la correlación respectiva entre los conjuntos de datos contenidas en el Repositorio Papyrus y la investigación resguardada en el Repositorio Institucional
- Preservar los conjuntos de datos en el repositorio de acuerdo con el tiempo establecido en la plantilla de datos garantizando su permanencia, acceso y seguridad
- Establecer estrategias de promoción del Repositorio Papyrus en concordancia con la política de ciencia abierta motivando de esta forma la visibilidad nacional e internacional de la producción tomasina
- Velar por la correcta acreditación de propiedad de los conjuntos de datos (Citación)

## CONTENIDOS Y PROPIEDAD INTELECTUAL

El CRAI SANTOTO conservara en el RDI- Papyrus contenidos producidos por la comunidad CRAI SANTOTO que cumplan con los siguientes requerimientos:



Conjuntos de datos que sean parte de una investigación producida por la comunidad de la CRAI y que respondan a un programa académico. Solo se conservarán datos que sean producidos por autores de la comunidad CRAI



| Tipo de dato             | Descripción                              |
|--------------------------|------------------------------------------|
| Contenedores             | tar, gzip, zip, SQL, SPARQL              |
| Bases de datos           | xml, csv, json                           |
| Específicos de programas | Software                                 |
| Código fuente            | Java, C, R, Python                       |
| Video                    | mpeg, avi, mxf, mkv                      |
| Sonido                   | wave, aiff, mp3, mxf flac                |
| Estadísticas:            | dta, por, sas, sav                       |
| Imágenes:                | tiff, jpeg 2000, pdf, dng, gif, bmp, svg |
| Datos tabulares:         | CSV, txt, xls / xlsx, mbd / *.accdb, dbf |
| Texto:                   | xml, pdf / a, html, json, txt            |
| Cualitativos textuales:  | xml, rtf, RDF, OWL, RIS, BIBTex          |
| Archivo web:             | Warc, html                               |

*Nota:* Se recomienda utilizar formatos con estándares abiertos y documentados, no encriptados, sin compresión y usados entre investigadores de su área

✓  Dataverses (6)

✓  Datasets (0)

✓  Files (0)

#### Nombre de autor

Crai Santoto

#### Palabra clave

Aspectos legales (1)

Autenticidad (2)

Autorización (1)

Datos (1)

Derechos patrimoniales (2)

Documentos (3)

Información confidencial

Publicación (2)

Preservación de datos (3)

Promoción (1)

Reproducción (1)

#### Año de publicación

2025 (1)



#### Tipos de datos excluidos del RDI- Papyrus

En el RDI-Papyrus en manejo de la universidad Santo Tomas y el CRAI SANTOTO se abstiene de publicar datos o conjuntos de datos con:

- > Información confidencial
- > Información sin autorización para publicación o divulgación
- > Información que infrinja derechos de terceros/autor
- > Que, contenga elementos de carácter publicitario o pornográfico, ya sea explícito o implícito
- > Que contenga información que constituya acto ilícito alguno
- > Datos que no tengan toda la información clara en la plantilla de datos
- > Documentos diferentes a conjuntos de datos (Artículos- Revistas- libros); Estos últimos se resguardarán en el Repositorio Institucional CRAI

Se podrá incluir un pre-print del artículo cuando este aun no esté publicado y se mantendrá oculto del Set de datos

### ASPECTOS LEGALES DE LOS CONJUNTOS DE DATOS



Una vez la información sea ingresada en el Repositorio de Datos:

#### Tipo de datos y formato

- > El autor (es) conservarán los Derechos patrimoniales y morales sobre su obra
- > Entregar a la Universidad y el CRAI SANTOTO los derechos de reproducción y comunicación pública con fines de promoción y preservación de los datos; esto no implica explotación comercial ya que esta es expresa del autor(es)
- > El CRAI SANTOTO no se hace responsable por la autenticidad y veracidad de los datos contenidos en el RDI-Papyrus; esta responsabilidad recae sobre el autor(es)
- > Los autores están en la obligación de expresar cualquier cambio sobre los derechos/obligaciones de uso del conjunto de datos establecidos previamente. Estos deben ser expresados por escrito

- ☒  Dataverses (6)
- ☒  Datasets (0)
- ☒  Files (0)

**Nombre de autor**  
Crai Santoto

**Palabra clave**  
 Alcance de publicación (1)  
 Estándares (1)  
 Conjuntos de datos (2)  
 Correlaciones (1)  
 Derechos de autor (2)  
 Identificador digital (3)  
 Lenguajes COAR (1)  
 Licencias (1)  
 Metadatos (1)  
 Principios FAIR (1)  
 OpenAIRE (1)

**Año de publicación**  
2025 (1)

## Licencias



### Licencia de depósito

Para la publicación del set de conjunto de datos; el autor debe diligenciar y entregar al CRAI SANTOTO:

- Formulario GDI completamente diligenciado. En caso de duda con el diligenciamiento, el personal del CRAI estará disponible para el asesoramiento necesario
- Formato de derechos de autor con la firma de todos los participantes de la investigación
- Los documentos que acrediten el depósito del resultado de investigación de esa investigación en un repositorio institucional o revista electrónica para realizar las correlaciones pertinentes



### Licencia de uso

El licenciamiento de los conjuntos de datos se hará expresamente bajo las licencias Creative Commons 4.0. El autor deberá especificar en la plantilla GDI el tipo de licencia y alcance de acuerdo con la siguiente información



| Tipo de licencia                                        | Descripción                                                                                                                                                                |
|---------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Atribución (CC BY)                                      | La obra puede usarse siempre y cuando se cite al autor                                                                                                                     |
| Atribución - Compartir Igual (CC BY – SA)               | Puede usarse para cualquier proyecto, sea o no comercial, se edite o modifique o no, siempre y cuando se cite al autor y sea publicada bajo la misma licencia CC BY-SA 4.0 |
| Atribución-No Derivadas (CC BY-ND)                      | Sea o no con ánimo de lucro, siempre y cuando se cite al autor y no se haga una obra derivada de ella, es decir no editar la obra                                          |
| Atribución - No Comercial (CC BY-NC)                    | Se puede usar la obra si se cita al autor, pero no en proyectos comerciales                                                                                                |
| Atribución - No Comercial-Compartir Igual (CC BY-NC-SA) | Se puede usar una obra siempre y cuando sea para un proyecto sin ánimo de lucro, se cite al autor y se publique el resultado bajo la licencia CC BY-ND-SA 4.0              |
| Atribución - No Comercial-No Derivadas (CC BY-NC-ND):   | Use una obra siempre y cuando se cite al autor, sea para proyectos no comerciales y no se modifique de ninguna manera                                                      |

De esta forma se garantizará el alcance de publicación y acceso de los conjuntos de datos suministrados por el autor



| Tipo de licencia      | Descripción                                                                                               |
|-----------------------|-----------------------------------------------------------------------------------------------------------|
| Proveedor tecnológico | Consortia Papyrus                                                                                         |
| Estándares            | Directrices de OpenAIRE para Administradores de Repositorios de Literatura 4.0 - Principios FAIR          |
| Esquema               | DataCite - Dublín Core                                                                                    |
| Protocolo             | Citation metadata (DCMI Metadata), Geospatial Metadata (Iso 3166-1), Social Science & Humanities Metadata |
| Identificador digital | DOI, ORCID, SCOPUS                                                                                        |
| Vocabulario semántico | Lenguajes COAR (Confederation of Open Access Repositories)                                                |

## 🔗 INFORMACIÓN ADICIONAL

- ✓ 🔗 Dataverses (6)
- ✓ 📄 Datasets (0)
- ✓ 📁 Files (0)

**Nombre de autor**  
Crai Santoto

**Palabra clave**

- Alcance de publicación (1)
- Archivo (1)
- Datos abiertos (1)
- Estándares (1)
- Conjuntos de datos (2)
- Correlaciones (1)
- Derechos de autor (2)
- Identificador digital (3)
- Lenguajes COAR (1)
- Licencias (1)
- Metadatos (1)
- Principios FAIR (1)
- OpenAIRE (1)
- Repositorio (1)

**Año de publicación**  
2025 (1)



### Plantilla de Gestión de Datos

#### Información administrativa

1. Descripción de los datos
2. Documentación y metadatos
3. Acceso y difusión
4. Re-uso
5. Preservación



### Roles de los participantes de Papyrus

1. Administrator: persona que tiene todos los permisos para todos los dataverse, datasets y archivos
2. Curator: Para datasets, persona que puede editar, editar permisos y publicar datasets
3. Contributor: Para datasets, persona que puede editar después de enviar para revisión
4. Dataverse + Dataset creator: Persona que puede agregar subdataverses y datasets dentro de un dataverse
5. Dataset creator: Persona que puede agregar datasets dentro de un dataverse
6. Dataverse creator: Persona que puede agregar subdataverse dentro de un dataverse
7. File downloader: Persona que puede descargar un archivo publicado
8. Member: Persona que puede ver dataverse y datasets no publicado



### Estructura PGD Universidad Santo Tomás

1. Dataverse: Programa o departamento, para cada programa o departamento se creará un Dataverse
2. Dataset: Grupos de investigación, para cada grupo de investigación se generará un Dataset
3. Datafile: Archivos de cada dataset, cada archivo tendrá asignado un DOI

## GLOSARIO

- ☒  Dataverses (6)
- ☒  Datasets (0)
- ☒  Files (0)

**Nombre de autor**  
Crai Santoto

**Palabra clave**

- Alcance de publicación (1)
- Archivo (1)
- Datos abiertos (1)
- Estándares (1)
- Conjuntos de datos (2)
- Correlaciones (1)
- Derechos de autor (2)
- Identificador digital (3)
- Lenguajes COAR (1)
- Licencias (1)
- Metadatos (1)
- Principios FAIR (1)
- OpenAIRE (1)
- Repositorio (1)

**Año de publicación**  
2025 (1)



**Data/Datos:** Colección de elementos básicos de información producidos, almacenados o procesados por un programa informático

**Datafile/Archivo:** Archivo que contiene texto, gráficos, números u otro tipo de contenido

**Dataset /Conjunto de datos:** Colección de datos relacionados, agrupados o almacenados juntos que fueron producidos dentro de un mismo proyecto, o se enfocan hacia un mismo objeto de estudio

**Dataverse:** Colecciones que pueden contener datasets, datafiles o también otros dataverses

**Datos abiertos:** Son considerados datos abiertos a los que cualquier persona puede acceder. Generalmente su uso está soportado en una licencia que indique el alcance del uso

**Datos compartidos:** Este tipo de datos suelen estar disponibles con contenido de restricción sobre su uso o para un grupo o comunidad específica donde solo los usuarios autorizados tienen acceso

**Datos de investigación:** Definidos según la CEPAL como insumos que se generan o recolectan para sustentar los resultados de una investigación

**Datos restringidos:** Este tipo de datos tienen un tratamiento confidencial por diferentes causas, pero no para uso o libre descarga. Generalmente se incluyen en repositorios para descripción y preservación

**DOI:** Identificador de un objeto digital. Se usa en publicaciones electrónicas

**Metadato:** Dato que describe el contenido, el estado, el origen, entre otras características de datos u otro tipo de información

**Repositorio:** Plataforma donde se almacenan datos

**Rol:** Conjunto predefinido de permisos para realizar ciertas acciones en el dataverse, dataset o datafile. Se puede asignar un rol tanto a un usuario como a un grupo

# crai

---

CENTRO DE RECURSOS PARA EL  
APRENDIZAJE Y LA INVESTIGACIÓN

---

