



Manual de administrador

Catálogo de microdatos CKAN

Comisión para el Esclarecimiento de la
Verdad, la Convivencia y la No Repetición

Sistema de Información Misional

LA COMISIÓN PARA EL ESCLARECIMIENTO DE LA VERDAD, LA CONVIVENCIA Y LA NO REPETICIÓN

Comisionados

Francisco José de Roux
Alejandra Miller
Alejandro Valencia Villa
Ángela Salazar
Alejandro Castillejo
Marta Ruiz
Leyner Palacios Asprilla
Carlos Martín Beristain
Lucía González
Carlos Guillermo Ospina
Patricia Tobón Yagarí
Saúl Franco
Alfredo Molano Bravo

Dirección de Conocimiento

Diana Britto
Lina Catherine Mantilla

Sistema de Información Misional

Equipo formulador

Luis Camilo Cárdenas Echeverry
Andrea del Pilar González Peña
Juliana Aristizabal Ruiz
Juan Camilo Murcia Galindo

Diseño, diagramación e ilustración

Juliana Aristizabal Ruiz
Maria Alejandra Torres-Galindo
Leonardo Parra Puentes

CONTENIDO

INTRODUCCIÓN

Nota aclaratoria 5

¿Qué es? 6

GENERAL

Ingreso 8

Carga de información al catálogo de
microdatos 9

DICCIONARIO

Definiciones principales 17

Introducción

01

NOTA ACLARATORIA

Esta serie de guías está dirigida a los equipos de trabajo de la Comisión para el Esclarecimiento de la Verdad, la Convivencia y la No Repetición.

Por razones de conservación y seguridad de la información se recomienda seguir fielmente las indicaciones y criterios para el alistamiento, carga y procesamiento de los recursos digitales (copias finales) producidos o consultados en el desarrollo de los procesos misionales.

Esta guía fue actualizada el 10 de julio de 2022 en su tercera edición y se encuentra sujeta a cambios relacionados con actualizaciones de la plataforma.

¿QUÉ ES?

El catálogo de microdatos -CKAN- es una herramienta para el almacenamiento, disposición y uso de los datos al interior de la Comisión.

CKAN permite visualizar colecciones de bases de datos (registros administrativos), junto con una descripción y diccionarios de datos.

La administración y gestión del catálogo de microdatos es responsabilidad del grupo de analítica, el cual hace parte del Sistema de Información Misional de la Comisión de la Verdad. Los encargados de cargar la información al CKAN son los gestores de información, quienes tienen credenciales específicas para dicho fin.

General

02

INGRESO

1

Ingresa a <http://192.168.1.51/ingresar>

2

Inicie sesión con sus credenciales de acceso.

The screenshot shows the login interface of the VERDAD CKAN system. At the top, there is a navigation bar with the VERDAD logo, a search bar, and links for 'Conjuntos de datos', 'Organizaciones', 'Grupos', and 'Acerca de'. The main content area is titled 'Conectarse' and contains a login form with fields for 'Nombre de usuario:' and 'Contraseña:'. There is a checkbox for 'Recordarme' and a 'Conectarse' button. A sidebar on the left contains links for '¿Olvidó su clave?' and '¿Olvidaste tu contraseña?'. The footer includes 'Acerca de COMISION DE LA VERDAD - CKAN' and 'Gestionado con'.

Ingreso CKAN

En la página de inicio encontrará información general sobre los dataset disponibles (conjunto de datos), organizaciones que han entregado registros a la Comisión así como los temas asociados a cada base de datos.

The screenshot shows the home page of the CKAN system. It features a navigation bar with the VERDAD logo, a search bar, and links for 'Conjuntos de datos', 'Organizaciones', 'Grupos', and 'Acerca de'. The main content area is titled 'Bienvenido a CKAN' and contains a large placeholder image with the text '420 x 220'. Below the placeholder, there is a section for 'Buscar datos' with a search bar and a list of popular tags: 'Etiquetas populares', 'Dominio transversal...', 'Desarrollos institu...', 'Núcleo 2. El Estado...'. The footer includes 'COMISION DE LA VERDAD - CKAN'.

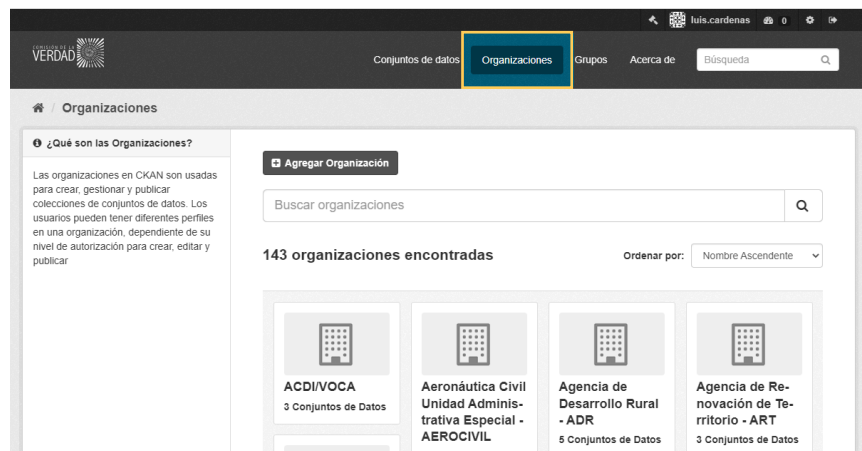
Inicio CKAN

CARGA DE INFORMACIÓN AL CATALOGO DE DATOS

Antes de iniciar la carga de la información al catálogo debe verificarse si la fuente de información de donde proviene el dataset existe dentro de la lista de organizaciones disponible.

1

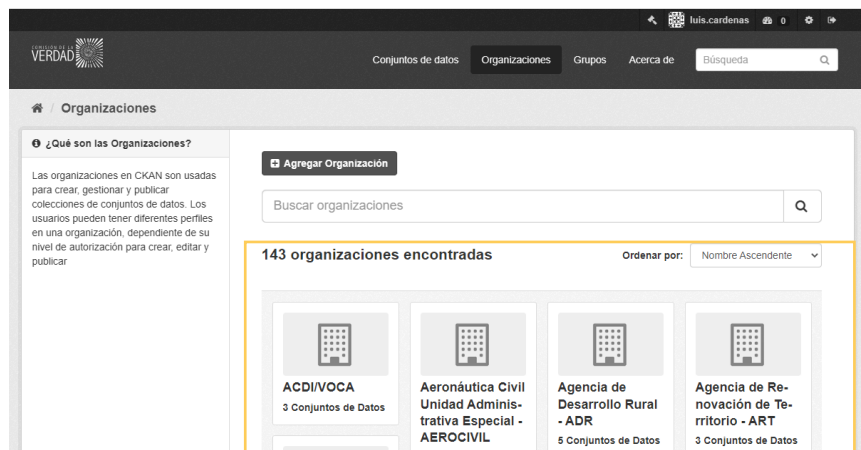
Para realizar esta revisión el usuario debe seleccionar el botón denominado “organizaciones” ubicado en la parte superior derecha.



Botón organizaciones

2

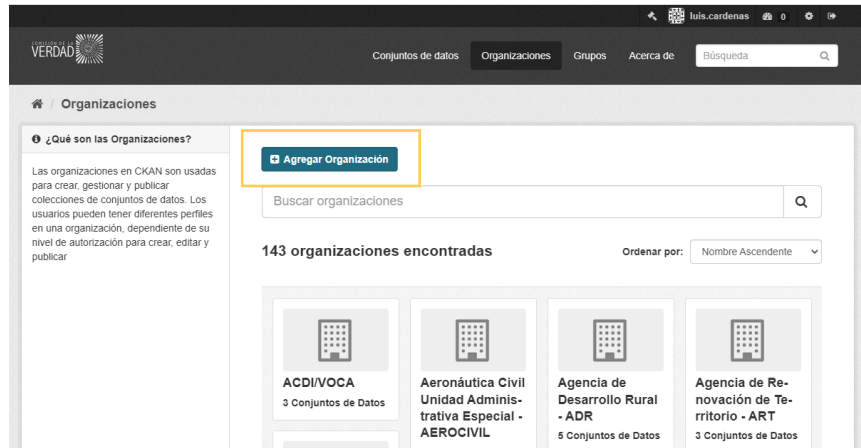
Se desplegará una lista de organizaciones ordenadas alfabéticamente.



Listado de organizaciones

3

Si la organización no está en la lista, el usuario deberá crear una nueva organización haciendo clic en “agregar organización” en el lado izquierdo de la pantalla.



Botón agregar organizacion

4

Al oprimir el icono “agregar organización” se desplegará una ventana con dos (2) recuadros para diligenciar.

Formulario crear una organizacion

- 1 En el recuadro del título de la organización debe escribirse el nombre completo de la entidad fuente de los datos, seguido de un guión y sus respectivas siglas si corresponde.
- 2 Por defecto, en el segundo recuadro aparece la opción “no depende de otra organización”. En caso de que el dataset que esté cargándose pertenezca a un conjunto de entidades que hagan parte de una de una organización más grande debe escogerse la opción. En caso de que esa organización de la cual depende no exista, debe crearse.

5

Una vez se diligencian los dos recuadros, oprima “crear organización”.

6

En la página de inicio, debajo del banner principal se encuentra un ícono “crear dataset”.

7

Al oprimirlo se desplegará una ventana con el campo “crear el dataset”.

- 1 **Título:** Nombre asignado al conjunto de datos. Este corresponde con el nombre de la entrega. De ser un nombre confuso, no acorde o demasiado extenso o genérico el catalogador podrá asignarle uno pero será necesario que en el campo NOTAS escriba la etiqueta “Título asignado por el catalogador”.

Recomendaciones:

Escribir la primera letra del título, los nombres propios y las siglas con mayúscula. Evitar el uso de mayúscula sostenida y el punto final.

Si el nombre posee un subtítulo separarlo usando dos puntos (:). Si se encuentra un guión (-) mantenerlo y dar espacio antes y después de usarlo.

- 2 **Descripción:** Descripción breve de la información contenida en el conjunto de datos. Se recomienda utilizar previamente una aplicación de texto para ajustar ortografía y redacción, destacando los aspectos más importantes del conjunto de datos (Mínimo un párrafo). Dentro del resumen se debe anexar la forma de citación teniendo en cuenta el modelo dado dentro del Manual de Estilo Editorial para las bases de datos.

Estructura de citación

Repositorio del SIM acompañado del ID interno. [Tipo de información estructurada] [Título de la operación estadística o Base de datos], [conjunto de base de datos]. Entidad aportante, fecha de última actualización de la base de datos. Fecha de corte.

- 3 **Etiquetas:** Términos asociados al contenido del conjunto de datos respecto a actores armados, tipos de violencia e hitos.

Recomendación

Los términos deben coincidir con los del Tesauro (lenguaje controlado) por lo que se precisa una verificación de su escritura.

Este campo integra los conceptos o categorías del Tesauro mientras se van escribiendo los términos asociados.

El número de etiquetas tienen un mínimo de uno y un máximo de diez. Es ideal que se use el mayor número de etiquetas posibles para que el etiquetado facilite la recuperación rápida de la información.

- 4 **Licencia:** Condiciones en las que el contenido del conjunto de datos puede ser utilizado, distribuido o reproducido.

- 5 **Organización:** Entidad que elabora y/o entrega la información / conjunto de datos.

Recomendaciones

Dentro del campo aparecerá una lista desplegable luego es necesario escribir el nombre de la entidad como aparece en el Tesauro. Tener en cuenta las siglas, guiones (-) y demás puntos que

se usen para poder hallarla y verificar su existencia.

Si dentro de la producción del conjunto de datos se ven implicadas dos o más instituciones se deberá separar el nombre de éstas usando un pipeline (|) para separar los nombres.

CREACIÓN: Si dentro de la lista desplegable no se encuentra la organización, la herramienta permite la creación de una. Para ello, se debe dirigir a la parte superior de la ventana en la pestaña Organizaciones e ingresar, allí deberá dar clic en Agregar Organización y completar los campos que aparecerán:

Nombre: Nombre completo de la entidad, tener presente la escritura que aparece dentro del Tesauro. Si no se encuentra en el Tesauro revisar la página oficial de la entidad y escribir el nombre como aparece allí respetando las mayúsculas, siglas, separaciones por guiones (-) o dos puntos (:). Si el nombre de la entidad no se encuentra en el Tesauro sugerir la creación del término en el Formato términos sugeridos 2022 en la pestaña que corresponde a Ckan. Esperar la aprobación de creación del término y entonces continuar con el proceso de creación.

Descripción: Texto corto que describa las principales características de la organización.

Imagen: Logo de la entidad. (Opcional)

Una vez terminado el proceso dé clic en Crear organización y diríjase nuevamente a la pestaña organizaciones verifique que la entidad ha quedado guardada. Proceda a escogerla y continúe con el diligenciamiento del formulario.

- 6 **Visibilidad:** Escoger el tipo de visibilidad que tiene el conjunto de datos. Privada o Pública.
- 7 **Fuente primaria:** Área o dependencia dentro de la Organización que produjo el conjunto de datos.
- 8 **Fecha de producción:** Fecha en la que se cargó la base de datos en CKAN. Deberá usarse el formato ISO 8601 (año, mes, día) separados por guión (-). Es obligatorio que mínimo se cuente con el año de producción.
- 9 **Notas:** Observaciones sobre el estado de la información, situaciones atípicas, incongruencia en fechas, entre otras. Cada una de las notas debe estar separada por un pipeline (|) esto en aras de lograr que la información pueda transferirse al Buscador.

Normalización de las notas:

Título asignado por el catalogador: Cuando se haya asignado un nuevo nombre a la base de datos para el registro dentro de la herramienta.

Rol de otros productores: Si se diligenció el campo de otros productores se debe indicar nuevamente el nombre y el rol de la persona o entidad separado por un guión dentro de este campo.

Rol de Financiación: Si se diligenció el campo de financiación se debe indicar nuevamente el nombre y el rol de la persona o entidad separado por un guión dentro de este campo.

Base Anonimizada: Si los recursos cargados dentro del conjunto de datos pasaron por un proceso de anonimización escribir la etiqueta "Base Anonimizada"

- 10 **Tipo de datos:** tributo que indica la clase de datos que se encuentran dentro de la base de datos.

Lista desplegable:

Registro administrativo: Conjunto de datos que contiene la información de entidades u organizaciones en el cumplimiento de sus funciones o competencias misionales.

Muestreo: Elección de un conjunto representativo de personas o elementos que se extraen de un universo o población para determinar o analizar las características propias del mismo.

Registro de medios: Conjunto de datos registrados por una persona u organización perteneciente a algún medio de comunicación.

Registro jurídico: Conjunto de datos que contienen información de una institución de administración pública relacionados con procesos judiciales.

Censo: Conjunto de datos que dan cuenta del recuento total de individuos que conforman una población estadística.

11 Unidad de Análisis: Entidad principal que está siendo analizada dentro del conjunto de datos.

Lista desplegable:

Individuos: Persona o grupo de personas.

Hechos: Conjunto de violaciones de los Derechos Humanos e infracciones al Derecho Internacional Humanitario de las cuales fueron víctimas las personas en el marco del conflicto armado colombiano. El tesoro contempla 18 hechos victimizantes

Casos: Expediente de un asunto particular

Eventos: Acontecimiento particular ocurrido en el marco del conflicto armado.

Instituciones y/o agrupaciones: Conjunto de personas que forman entidades con o sin ánimo de lucro.

Otros: Descripciones de unidades de análisis espacial.

12 País: Indica el país o los países analizados dentro de la base de datos. Si se discrimina una región particular, un territorio específico o una localización específica dentro del país se podrá mencionar dentro del campo resumen.

13 Cobertura Geográfica: Es una categoría que viene por defecto en la plataforma web. Estos temas no se verán en el portal. Sin embargo, como resultan un metadato que podría llegar a utilizarse posteriormente, se eligen de acuerdo a la información que va ser cargada.

14 Frecuencia de actualización: Espacio geográfico en el que se desarrollaron los eventos que describen los datos.

Normalización de la cobertura geográfica:

Teniendo en cuenta la escala en la que se desarrollaron los eventos se deberá poner uno de los siguientes conceptos respetando la escritura dentro del campo.

Nacional: Los datos del conjunto de datos hacen referencia a hechos que abarcan todo el territorio nacional.

Regional: Los datos hacen referencia a dos o más departamentos o porciones de territorio nacional que no contemplan los límites administrativos departamentales.

Departamental: Los conjuntos de datos hacen describen hechos dados dentro de un departamento determinado.

Municipal: Los datos describen eventos dados en un municipio o distrito determinado.

Local: Los datos describen hechos dados dentro de una porción de particular de un municipio o intercepción de estos.

15 Universo: Grupo de personas que son objeto de análisis dentro de la base de datos y a los que se refieren los resultados analíticos. Escoger los términos del Tesauro asociados a “Poblaciones” y “Género”.

16 Investigador principal: La persona, entidad corporativa u organismo responsable del contenido sustantivo e intelectual de la obra..

Recomendación:

El registro de los nombres se debe realizar según la guía de diligenciamiento de Catalogación.

17 Otros productores: Autoridad (persona/organización) que contribuyó a la producción del recurso. Se debe indicar el rol que cumplieron dentro de la elaboración del conjunto de datos. No olvidar describir el rol como se indica en el campo notas.

Si dentro de la producción del conjunto de datos se ven implicados dos o más productores se deberá separar el nombre de estos usando un pipeline (|) para separar los nombres

Ejemplo: GrupoRENACER|UNESCO|

18 Financiación: Describe las fuentes de financiación para la producción de la obra. Al nombrar la entidad o persona que financió el recurso se debe indicar su rol. No olvidar describir el rol como se indica en el campo notas.

Si dentro de la producción del conjunto de datos se ven implicados dos o más financiadores se deberá separar el nombre de estos usando un pipeline (|) para separar los nombres

Ejemplo: Amnistía Internacional|UNESCO|Fiscalía

19 Fecha de recogida de datos: Cobertura temporal de los sucesos / hechos / eventos relatados en la base de datos, relacionados con temas de derechos humanos y conflicto armado. Fecha(s) en que se recogieron los datos que se encuentran dentro del recurso. (AAAA-AAAA)

Recomendación:

Las fechas de este campo corresponden a la temporalidad de los datos según la base, no a los de registro de la base.

20 Nivel de Acceso: Nivel de acceso al documento asignado según el Protocolo de manejo de la información de la CEV.

21 Acceso a la información: Categoría de la calificación del recurso y la justificación legal de la excepción del acceso, según la Ley 1712 de 2014.

Lista desplegable:

1 Inteligencia y contrainteligencia - Ultrasecreto.

2 Inteligencia y contrainteligencia - Secreto.

3 Inteligencia y contrainteligencia - Restringido.

4 Pública reservada - Información exceptuada por daño a los intereses públicos. La defensa y seguridad nacional.

5 Pública reservada - Información exceptuada por daño a los intereses públicos. La seguridad pública.

6 Pública reservada - Información exceptuada por daño a los intereses públicos. Las relaciones internacionales.

7 Pública reservada - Información exceptuada por daño a los intereses públicos. La prevención, investigación y persecución de los delitos y las faltas disciplinarias aseguramiento o se formule pliego de cargos, según el caso.

8 Pública reservada - Información exceptuada por daño a los intereses públicos. El debido proceso y la igualdad de las partes en los procesos judiciales.

9 Pública reservada - Información exceptuada por daño a los intereses públicos. La administración efectiva de la justicia.

10 Pública reservada - Información exceptuada por daño a los intereses públicos. Los derechos de la infancia y la adolescencia.

22 **Autorización de acceso:** Persona o equipo de contacto para obtener acceso a la ubicación del conjunto de datos.

23 **Idioma:** Escoger de la lista desplegable el idioma en el que se encuentra escrita la base de datos dentro del conjunto de datos.

8

El dataset no puede ser creado si no se vinculan recursos. Los recursos son las bases de datos o la documentación de cada dataset.

1 **Título del recurso y descripción del recurso:** Nombre de la base de datos o de los archivos de documentación y su descripción.

2 **Tipo de recurso:** De la lista desplegable se elige el tipo de archivo al que corresponde cada recurso. (archivo de datos, documentación código)

3 **Fichero:** Este ícono sirve para cargar la información. De acuerdo al peso del recurso los tiempos de carga varían.

4 **Formato:** En este recuadro se escribe el formato de cada recurso (xlsx, PDF, Word, etc).

5 **Licencia:** Al igual que con la creación del dataset, el recuadro de licencia despliega los 4 niveles de acceso de la información. Se debe elegir una opción de acuerdo a la tabla de niveles de acceso de la Comisión de la Verdad.

9

El paso final para cargar la información se resume en las opciones que aparecen una vez se finaliza el diligenciamiento de todos los elementos citados anteriormente.

1 **Seguir y agregar recursos:** Se oprime esta opción cuando el dataset que vamos a cargar relaciona múltiples recursos.

2 **Subir y dejar en borrador:** Se oprime esta opción cuando la información contenida en el dataset es de reserva o se trata de las bases originales. Aunque queda guardada en la plataforma no es accesible para los usuarios básicos.

3 **Publicar:** Una vez se termina el diligenciamiento de los formularios, se oprime esta opción. Se publican solo los dataset públicos y los que contengan bases de datos anonimizadas.

Diccionario

03

DEFINICIONES PRINCIPALES

Término	Definición
Base de datos	Conjunto de datos pertenecientes a un mismo contexto y almacenados sistemáticamente para su posterior uso al interior de la Comisión.
Datasets	<p>Conjuntos de datos que contienen información del diccionario de datos y registros administrativos entregados por entidades a la Comisión de la Verdad.</p> <p>En este espacio, el usuario encontrará la base original o anonimizada y el diccionario de datos de la información a consultar.</p> <p>Por cada datasets, el usuario encontrará una pequeña descripción de la información, datos generales y etiquetas de la base, temas asociados a la Comisión e información de actualización de la base.</p>
Diccionario de datos	Grupo de metadatos que describen el contenido informativo de las variables de un recurso o base de datos.
Fuente de Información	Es la persona, entidad u organización que recibe o conoce datos personales de los titulares de la información, en virtud de una relación comercial o de servicio o de cualquier otra índole y que, en razón de autorización legal o del titular, suministra esos datos a un operador de información.
Tesauro	Es una lista de palabras o términos controlados, empleados para representar los conceptos empleados por los miembros de la Comisión de la Verdad.

Canales de comunicación

Catálogo de datos CKAN: <http://datos.comisiondelaverdad.co>

Correo institucional SIM: sim@comisiondelaverdad.co



Comisión para el Esclarecimiento de la Verdad la Convivencia y la No Repetición

Sistema de Información Misional