ESTÁNDAR DE DATOS PARA CONJUNTOS DE DATOS DE PROYECTOS DE INVESTIGACIÓN

FECHA DE ACTUALIZACIÓN: MARZO 2026

DESCRIPCIÓN DEL ESTÁNDAR

La adopción de estándares para la publicación de conjuntos de datos permite la interoperabilidad de los datos por sistemas informáticos, además de facilitar el uso de los datos para nuevos análisis y generar nuevo conocimiento.

Para establecer los términos adecuados para describir los conjuntos de datos de proyectos de investigación se revisaron los esquemas de metadatos utilizados por dos de las principales plataformas, dedicadas actualmente a la publicación de este tipo de recursos: la plataforma CKAN que utiliza el esquema Data Catalog Vocabulary (DCAT, https://www.w3.org/TR/vocab-dcat/) y la plataforma de datos de investigación INVENIO que basa su diseño de metadatos en DataCite (https://datacite.org/index.html). Asimismo, se revisaron los vocabularios controlados desarrollados por Schema.org (https://schema.org/), que permiten obtener la mejor recuperación de los conjuntos de datos, así como la descripción de sus contenidos.

Considerando estas fuentes y las necesidades propias de las Plataformas de la DGRU se diseñó el siguiente esquema de metadatos.

CC-BY-4.0 Internacional. https://creativecommons.org/licenses/by/4.0/legalcode.es

Sugerencia de cita: DGRU.2026. Estándar de datos para conjuntos de datos de proyectos de investigación (Versión 2). Dirección General de Repositorios Universitarios. SDI-UNAM. México.

ESTÁNDAR DE DATOS PARA CONJUNTOS DE DATOS DE PROYECTOS DE INVESTIGACIÓN

FECHA DE ACTUALIZACIÓN: MARZO 2026

LISTA DE CAMPOS

Termino	Definición
id	Número consecutivo e irrepetible asignado por la base de datos de la DGRU.
guid	Identificador construido por el código de la institución. Corresponde con el código de la comunidad y el identificador numérico en la base de datos.
dataset_id	Identificador de un grupo de conjuntos de datos o un identificador específico de una comunidad
identifier	Identificador interno de la plataforma.
alternate_identifiers	Lista de identificadores del recurso en otras plataformas. Debe indicarse la cadena del indicador y la plataforma en la que el recurso se identifica.
communities	Nombre de la comunidad a la que pertenece el recurso. Dicho nombre proviene de las comunidades registradas en la base de datos de la DGRU.
title	Nombre o título del recurso. Deber ser un nombre que describa de forma precisa y concisa el tipo de recurso, la temática y la procedencia.
version	Número que identifica la versión del recurso. Se sugiere utilizar cadenas numéricas y seguir algún sistema de versionado como los descritos en https://semver.org
resource_type	Descripción del recurso: base de datos, código, texto, etc.
database	Características técnicas de la base de datos: formato de la base de datos, lenguaje, número de columnas, número de filas, estándar de datos utilizado, última fecha de actualización de los registros.

CC-BY-4.0 Internacional. https://creativecommons.org/licenses/by/4.0/legalcode.es

Sugerencia de cita: DGRU.2026. Estándar de datos para conjuntos de datos de proyectos de investigación (Versión 2). Dirección General de Repositorios Universitarios. SDI-UNAM. México.

ESTÁNDAR DE DATOS PARA CONJUNTOS DE DATOS DE PROYECTOS DE INVESTIGACIÓN

FECHA DE ACTUALIZACIÓN: MARZO 2026

Termino	Definición
creators	Nombre del investigador principal involucrado en la producción del recurso.
contributors	Nombre de las personas o instituciones responsables de la compilación, manejo, distribución o desarrollo del recurso.
date	Lista de las fechas y el tipo de fecha relacionadas con el recurso. Ejemplos: creación: 2019-05-29, actualización: 2019-05-29.
embargo_date	Fecha después de la cual el recurso puede ser público.
files	Descripción técnica de los archivos: nombre del archivo, tipo MIME, tamaño.
subjects	Temas, palabras clave, clasificación o frases que describen al recurso.
description	Información adicional que describen el contenido y la parte técnica de la base de datos, que no se encuentra en ninguna de las categorías indicadas en este estándar.
related_identifiers	Lista de identificadores relacionados con el recurso.
keywords	Palabras clave normalizadas.
language	El idioma principal del contenido del recurso.
license	Licencia de uso del recurso.
references	Lista de referencias utilizadas por el recurso.
access_condition	Condiciones de acceso al recurso.
access_right	Tipo de acceso al recurso: acceso abierto, recurso con fecha de embargo, con acceso restringido, acceso cerrado.
notes	Notas técnicas para el procesamiento del recurso antes de su publicación.
modified	Fecha de la última modificación del registro en la base de datos de origen.

CC-BY-4.0 Internacional. https://creativecommons.org/licenses/by/4.0/legalcode.es

Sugerencia de cita: DGRU.2026. Estándar de datos para conjuntos de datos de proyectos de investigación (Versión 2). Dirección General de Repositorios Universitarios. SDI-UNAM. México.

Termino	Definición
_creation	Fecha de creación del registro en la base de datos de la DGRU.
_creation_user	Nombre del usuario que creó el registro en la base de datos de la DGRU, en caso de no contar con el dato se utilizará el código del dataset_id.
_last_modified	Fecha de última modificación del registro dentro de la Base de datos de la DGRU.
_last_modified_user	Nombre del usuario que realizó la última modificación del registro dentro de la base de datos de la DGRU.
responsable_quality_control	Nombre del usuario responsable del proceso de control de calidad estructural en la CDI-DGRU.

HISTÓRICO DE CAMBIOS

1. Se actualizó DESCRIPCIÓN DEL ESTÁNDAR
2. Se actualizó LISTA DE CAMPOS

CC-BY-4.0 Internacional. https://creativecommons.org/licenses/by/4.0/legalcode.es

Sugerencia de cita: DGRU.2026. Estándar de datos para conjuntos de datos de proyectos de investigación (Versión 2). Dirección General de Repositorios Universitarios. SDI-UNAM. México.