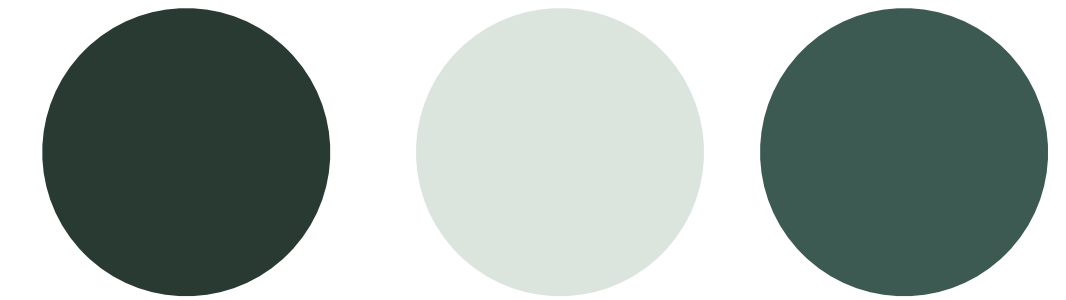




Claves para la Gestión de Datos de Investigación: DMP, Repositorios, Open Data y FAIR

Jose Navarrete

Septiembre - 2026

Introducción

¿Por qué esto ya no es opcional?

UE

Horizonte Europa

Plan de Gestión de Datos obligatorio antes del mes 6, bajo principios FAIR. Regla de fondo: tan abierto como sea posible, tan cerrado como sea necesario.

ES

Ley de la Ciencia

Art. 37 (Ley 17/2022): la investigación financiada con fondos públicos debe depositar sus datos en repositorios de acceso abierto.

AN

Sexenios ANECA

Los datos de investigación ya cuentan como aportación evaluable en la última convocatoria. Esto os toca en la carrera investigadora.

El investigador fantasma



Un proyecto de tres años termina. El investigador principal cambia de universidad. Los datos en bruto quedan en un disco duro externo, en el ordenador de un becario que ya se fue, en una carpeta llamada “definitivo_v2_final”. Nadie recuerda qué significan las columnas de esa hoja de cálculo.

El conocimiento se pierde, aunque el artículo ya esté publicado.

La reproducibilidad imposible



Un grupo quiere replicar un experimento publicado hace cinco años. El artículo dice: “los datos están disponibles bajo petición”. Escriben al autor.. y el correo rebota.

Sin acceso real a los datos, la reproducibilidad es un principio de papel, no un hecho.

El dataset dormido



Una universidad financia un dataset carísimo: encuestas, secuenciación genética, sensores de campo. Se usa una vez, para un artículo, y después duerme en un disco duro.

Podría haber sido la base de tres tesis más si estuviera bien depositado y descrito.

Plan de Gestión de Datos

Un documento vivo, no un trámite

Se presenta al inicio de la investigación y se va modificando a medida que avanza el proyecto. Un DMP mal hecho al principio no se arregla al final — pero un DMP vivo sí se puede corregir sobre la marcha.

En Horizonte Europa:

Se entrega antes del mes 6 del proyecto, y se actualiza a mitad y al final.

Las preguntas detrás de la plantilla

1

Qué datos vas a generar o reutilizar

Tipo, formato y volumen aproximado; si son datos nuevos o ya existentes.

2

Cómo los vas a documentar

Qué metadatos y qué estándar usarás (conecta con Findable).

3

Dónde y cómo se almacenan durante el proyecto

Almacenamiento activo frente a archivo final: no es lo mismo.

4

Quién tiene acceso y bajo qué condiciones

Seguridad y, si hay datos sensibles, referencia al RGPD.

Las preguntas detrás de la plantilla

5

Qué se compartirá y con qué licencia

Tan abierto como sea posible, tan cerrado como sea necesario.

6

Quién es responsable de cada cosa

La sección que menos se rellena bien, y la que más piden los financiadores.

7

Qué pasa con los datos al terminar el proyecto

Preservación a largo plazo: quién la costea.

Cómo hacerlo, en la práctica

Herramientas para elaborar tu DMP

Argos

Recomendado por defecto

OpenAIRE + EUDAT. Gratuito, alineado con FAIR, plantillas por financiador (incluida Horizonte Europa). Asigna DOI al propio plan.

DMPTool

Financiación de EE. UU.

Mantiene actualizados los requisitos de financiadores como la NSF. Alternativa si el proyecto tiene fondos estadounidenses.

Muchas universidades y consorcios tienen su propia plantilla basada en Argos: conviene preguntar en la biblioteca o el servicio de investigación antes de partir de cero.

Referencia genérica, independiente de la herramienta: DCC Checklist for a Data Management Plan (13 preguntas).

Repositorios

Repositorio institucional $\neq$ repositorio de datos

Ya vimos que no toda infraestructura sirve igual para todo. Un repositorio institucional está pensado para producción científica —artículos, tesis—, con metadatos Dublin Core, no para datasets en bruto.

Por ejemplo, en la UJA:

RUJA es el repositorio institucional de producción científica —pensado para publicaciones, no para depositar datasets en bruto.

Tres tipos de repositorio, tres funciones

Institucional

RUJA y equivalentes

Visibilidad institucional y cumplimiento de mandatos de acceso abierto para publicaciones. Recolectado por RECOLECTA/OpenAIRE.

Multidisciplinar de datos

Zenodo, Dryad, Figshare, Dataverse

Acepta cualquier disciplina, asigna DOI automático y usa esquemas orientados a datasets (DataCite).

Temático o especializado

GenBank, GBIF, tDAR...

Cuando existe uno para tu disciplina, suele ser la mejor opción: ya habla el idioma de tu comunidad.

Modelo recomendado

No elegir uno: enlazarlos

Publicación

Repositorio institucional (RUJA)

DOI ↔ DOI

Dataset

Repositorio de datos (Zenodo, Dryad, temático...)

¿Y cómo elijo el repositorio de datos adecuado?

re3data.org — registro global de repositorios de datos, filtrable por disciplina, tipo de contenido y licencia.

Para cuando sí existe un repositorio para tu campo

Un repositorio para cada disciplina

ÁREA	REPOSITORIO	QUÉ CONTIENE
Biología molecular / genética	GenBank	Secuencias de ADN y genómica; gestionado por el NIH (EE. UU.).
Biodiversidad / ecología	GBIF	Registros de presencia de especies a escala global.
Ciencias sociales	ICPSR • CESSDA	Datos de encuestas y estudios sociales; CESSDA es la red europea.
Neurociencia	OpenNeuro	Datos de neuroimagen: fMRI, MEG, EEG.
Ciencias de la Tierra	PANGAEA	Datos geocientíficos y medioambientales.
Arqueología / patrimonio	tDAR	Documentación y registros arqueológicos.
Sin repositorio específico	Zenodo • Dryad	La opción multidisciplinar por defecto.

Antes de subir nada

Qué se sube, y con qué licencia



Metadatos mínimos

Siempre acompañan al dataset: autoría, fecha, formato, cómo se generó.



CC-BY por defecto

La licencia más recomendada: permite reutilización con solo citar la autoría.



Depositar ≠ publicar

El dataset y el artículo son dos objetos distintos, cada uno con su propio DOI.

El repositorio correcto no es el que tenéis más a mano: es el que mejor entiende lo que estáis depositando.

De la teoría a la medida

FAIR: el criterio con el que medimos todo lo demás

F

Findable

Metadatos ricos + identificador persistente (DOI). Localizable en catálogos y buscadores, no solo en un disco duro.

A

Accessible

Recuperable mediante un protocolo estándar, aunque el acceso esté restringido: el “cómo pedirlo” está documentado.

I

Interoperable

Formato abierto y estándar (CSV, no un .xlsx cerrado) con vocabularios controlados en los metadatos.

R

Reusable

Licencia clara (CC-BY) y documentación suficiente para que alguien ajeno al proyecto lo reutilice con garantías.

¿Y cómo se mide esto?

F-UJI

Evaluación automatizada a partir de un DOI: revisa metadatos, formatos y licencias, y genera una puntuación FAIR.

FAIR-Checker

Automatizada, con foco en debilidades semánticas: cómo se expone el dato en la web semántica.

ARDC FAIR Data Self-Assessment Tool

Cuestionario manual, no automatizado: evalúa tus propias prácticas de gestión, no el dataset en sí.

No existe una puntuación FAIR única: la misma dataset puede obtener resultados distintos según la herramienta.

El Consorcio Madroño ya integra F-UJI en e-cienciaDatos, su repositorio de datos de investigación.

Open Data

¿Por qué NO podría estar abierto?

Open Data parte de abierto por defecto. La pregunta no es por qué compartir un dato, sino qué justifica restringirlo. La restricción es la que necesita argumento —no la apertura.

El principio que lo resume:

“Tan abierto como sea posible, tan cerrado como sea necesario.”

Los tres motivos legítimos de restricción

Datos personales

RGPD

Cuando aplican las normas de protección de datos. Conecta con lo ya visto en el DMP.

Propiedad intelectual

Explotación comercial

Expectativa razonable de patentar o explotar comercialmente los resultados.

Confidencialidad

Seguridad

Casos más específicos pero reales, con implicaciones de seguridad.

Un cuarto caso, casi administrativo: si el proyecto no genera datos de investigación, la obligación simplemente no aplica.

Dos formas distintas de restringir

Embargo no es lo mismo que opt-out

Embargo temporal

Se abre más tarde, normalmente vinculado a la publicación del artículo asociado.

Tiene fecha de caducidad prevista.

Opt-out (parcial o total)

Restricción justificada por RGPD, propiedad intelectual o seguridad.

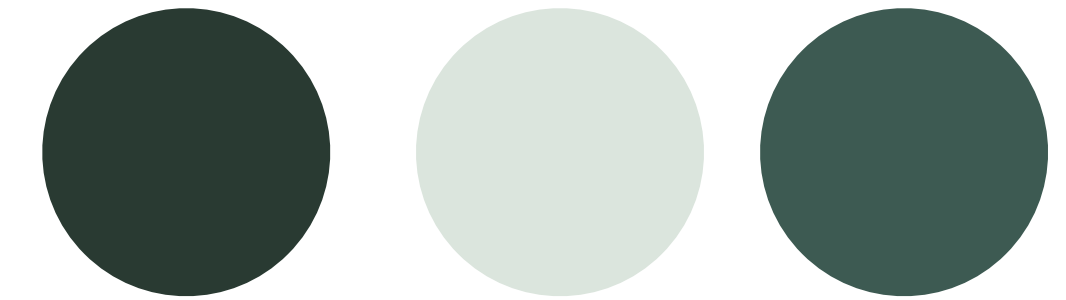
No tiene por qué tener fecha de caducidad.

El matiz importante

Cerrado no es invisible

Aunque el dato no esté abierto, los metadatos —qué es, quién lo tiene, cómo se puede solicitar— sí deberían estar visibles. Esto es lo que evita que “cerrado por necesidad” se convierta en “invisible”. Conecta con lo que ya vimos en FAIR: Accessible no siempre significa descarga libre.

Abierto no es un interruptor de todo o nada: es una decisión justificada, documentada, y siempre revisable.



Claves para la Gestión de Datos de Investigación: DMP, Repositorios, Open Data y FAIR

... Gracias

Jose Navarrete
jcortes@ujaen.es

Septiembre - 2026