

[Título]

Version 0

Financiador

Proyecto

Investigadores

Organizaciones

Plantilla recomendada para proyectos nacionales

Información principal

Título:

Descripción: *[Describa brevemente el contexto y propósito del Dataset.]*

1. Resumen de los datos

1.1 Resumen de los datos

1.1.1 Indique información general sobre el proyecto: Título, descripción breve del mismo y duración.

1.1.2 Indique quiénes son los financiadores del proyecto.

1.1.3 Indique los objetivos de los datos generados y / o recolectados y su relación con los objetivos del proyecto.

1.1.4 Indique el tipo y el formato de los datos que se generarán / recolectarán.

[Si los datos serán observacionales, experimentales, recolectados y el formato de los mismos (hojas de cálculo: xls, ods, csv; imágenes: png, tiff, jpeg; matemáticos: .mat, .m, .nb; ...). Se recomienda el uso de formatos abiertos que permitan su preservación o, si no es posible, formatos ampliamente usados en su disciplina. Puede comprobar la lista de [formatos recomendados por el repositorio e-cienciaDatos](https://www.consorciomadrono.es/docs/formatosArchivosRecomendados.pdf) (<https://www.consorciomadrono.es/docs/formatosArchivosRecomendados.pdf>)]

1.1.5 Indique el origen de los datos que se generarán, recolectarán o se reutilizarán de otras fuentes: encuestas, sensores térmicos, trabajos previos, bases de datos, datos de otros proyectos,...

1.1.6 ¿Cuál es el tamaño esperado de los datos?

[Haga una estimación del almacenamiento necesario para albergar sus datos, por ejemplo, alrededor de 15Mb, unos 200Gb, ...]

2. Datos FAIR

2.1 Los datos y metadatos se deben localizar fácilmente.

2.1.1 Indique el tipo de identificador persistente que se usará para facilitar la cita del dataset depositado en un repositorio.

[Al publicar un dataset, es muy importante que se le asigne un identificador persistente que facilite su cita y localización como el DOI o el handle.

Si es investigador del Consorcio Madroño y publica su dataset en e-cienciaDatos, se le asignará un DOI a sus datasets.]

2.1.2 Describa el sistema que se usará para nombrar los archivos y directorios, así como la estructura de los directorios en el caso de haberlos.

[Las carpetas y archivos con los datos de investigación deben tener nombres suficientemente descriptivos del contenido y deben ser organizados de una manera sistemática, de manera que sean identificables y accesibles para los usuarios.

Para proyectos con muchos datos, puede seguir las [recomendaciones de Standford para la nomenclatura de los ficheros](https://guides.library.stanford.edu/data-best-practices/name-files) de datos (en inglés) (<https://guides.library.stanford.edu/data-best-practices/name-files>).]

2.1.3 Una vez publicado el dataset, ¿podrán publicarse nuevas versiones? Si es así, ¿cómo se nombrarán las distintas versiones?

[Si es investigador del Consorcio Madroño, e-cienciaDatos generará versiones menores (V2.3) cada vez que se modifiquen los metadatos del dataset y mayores si se modifican los ficheros (V3.0).]

2.1.4 Los metadatos ayudan a localizar el dataset a los investigadores que quieran usarlo. ¿Qué esquemas de metadatos se usan en el repositorio en el que se va a depositar el dataset?

[Si es investigador del Consorcio Madroño, los metadatos de e-cienciaDatos se basan en el esquema DDI con la posibilidad, si su proyecto lo necesita, de añadir metadatos geoespaciales, de ciencias sociales, de ciencias de la vida, astronomía o astrofísica. Los metadatos se pueden exportar en los formatos Dublin Core, DDI, DataCite, json, OAI ORE, Shema.org y DDI HTML Codebook.]

2.2 Los datos deben de ser accesibles

2.2.1 Indique si los datos finales de la investigación estarán disponibles en acceso abierto y en qué repositorio se depositarán.

[Se recomienda depositar los datos finales de investigación en repositorios de acceso abierto generalistas como e-cienciaDatos, para los investigadores de las universidades del Consorcio Madroño.

Otros repositorios generalistas que puede utilizar son Zenodo, FigShare, Dataverse, ... o temáticos (PANGEA, Dryad, HEPData, PubChem, ...).

Si va a publicar su dataset en un repositorio temático especializado en un campo o área determinada de conocimiento, puede consultar la [herramienta de re3data de búsqueda de repositorios temáticos](https://www.re3data.org/browse/by-subject/). (<https://www.re3data.org/browse/by-subject/>)

2.2.2 Si los datos no pudieran estar en acceso abierto, indique la razón, quién y cómo podrá acceder a los datos y en qué repositorio se depositarán.

[Indique si los datos estarán codificados y qué tratamiento recibirán. También si tendrán un período de embargo. El embargo es un tiempo desde que se publica el dataset hasta que los datos están públicamente disponibles.]

2.2.3 Indique si el repositorio seleccionado es un repositorio de confianza (trustworthy, trusted). ¿Tiene algún certificado que lo acredite (CoreTrustSeal, ISO 1636, ...)?

[Si es investigador del Consorcio Madroño, e-cienciaDatos tiene el certificado [CoreTrustSeal](https://amt.coretrustseal.org/certificates) (<https://amt.coretrustseal.org/certificates>).]

2.2.4 Indique si tiene un plan de preservación para el mantenimiento de los datos a medio y largo plazo.

[Si es investigador del Consorcio Madroño, e-cienciaDatos tiene un [plan de preservación](#) y garantiza el acceso a los datos durante un mínimo de 20 años.]

2.3 Los datos deben de ser interoperables

[Para la interoperabilidad es necesario usar identificadores persistentes (para datasets DOI, Handle, ...) y esquemas de metadatos estándar, (algunos ejemplos son Dublin Core, DataCite, DDI en ciencias sociales, EML en ecología, dwc para biodiversidad).

Ver apartado 2.1.4.]

2.3.1 Indique qué vocabularios se usarán en la descripción de los datos para materias (Unesco, JEL, MeSH, ...) o identificadores estándar (ORCID o ResearcherID para autores, ROR para instituciones), ...

[Si es investigador del Consorcio Madroño, e-cienciaDatos permite indicar identificadores de autor (entre ellos ORCID) e identificadores de publicaciones relacionadas (entre ellos DOI y handle), así como de instituciones (ROR).]

2.3.2 Indique, si lo conoce, si el dataset estará disponible desde otros portales científicos generalistas o temáticos (Google Scholar, OpenAIRE, Recolecta, ...).

[Si es investigador del Consorcio Madroño, los datasets de e-cienciaDatos son visibles en Recolecta, Harvard Dataverse, Google Dataset Search, OpenAIRE Explore y DataCite Commons, entre otros.]

2.4 Los datos deben de ser reutilizables

2.4.1 Indique las condiciones de acceso a los datos y la licencia para el acceso a los mismos en caso de usar una licencia estándar como las Creative Commons.

[Se recomienda usar una licencia [Creative Commons](#) u [Open Data Commons](#) que permita la máxima reutilización con cita como la CC-BY. En el caso de haber ficheros con distinta tipología, puede haber licencias distintas para algunos de ellos, por ejemplo si se incluye software, se podría usar una licencia Apache o GPL.]

2.4.2 Indique, si lo conoce, el tiempo mínimo que el dataset estará accesible.

Si es investigador del Consorcio Madroño, e-cienciaDatos garantiza el acceso a los datos durante un mínimo de 20 años.

2.4.3 ¿Qué software será necesario para visualizar y reutilizar los datos?

[Puede ser software específico creado ad hoc para el proyecto, hojas de cálculos como Microsoft Excel u Open Office, u otros específicos de la disciplina.]

3. Asignación de Recursos Humanos y Económicos

3.1 Asignación de recursos humanos y económicos

3.1.1 Indique el personal (roles) encargado de generar los datos, tratarlos y preparar el dataset final. Indique quién será el encargado de la gestión de datos del proyecto.

3.1.2 Indique el coste de generar y mantener los datos durante la vida del proyecto si los hubiera (costes directos, indirectos, discos externos, contratación personal auxiliar, preparar los datos para que sean FAIR, ...).

[En e-cienciaDatos el depósito del dataset no le supone ningún coste al investigador del Consorcio Madroño.]

4. Seguridad de los datos

4.1 Seguridad de los datos

4.1.1 Indique los mecanismos para garantizar la seguridad y acceso al dataset (copias de seguridad, evitar acceso no autorizados, ...) tanto durante su gestión como cuando ya se han depositado los datos en un repositorio.

[Explique si se realizarán copias de seguridad y cómo se protegerán los datos de accesos no autorizados, protección de datos sensibles, ... Indique como se realizarán estas tareas antes y después del depósito en el repositorio.]

Si es investigador del Consorcio Madroño, en e-cienciaDatos:

- *Cada fichero tiene sumas de verificación para evitar la corrupción de ficheros.*
- *Incluye un gestor de versiones que permite ver las diferencias entre dos versiones del mismo dataset y el autor del cambio de versión.*
- *Se guarda la identidad del creador del dataset y de cada persona que hace el cambio de versión.*
- *Solo los administradores de un dataset o del repositorio pueden hacer cambios en el mismo.*
- *e-cienciaDatos está situado en el centro de cálculo del campus de la UNED, tras varias capas de firewall.*
- *Físicamente, el centro de cálculo está cerrado con llave y la UNED mantiene un registro de todas las personas que entran y salen en el mismo.*
- *Se realizan copias de seguridad parciales semanalmente y completas mensualmente que se mantienen en localizaciones separadas.]*

5. Aspectos éticos

5.1 Aspectos éticos

5.1.1 Explique los problemas éticos o legales que pueden afectar a la recogida e intercambio de datos.

[Posibles problemas de confidencialidad, de tratamiento de datos sensibles, ... y la forma en que se abordarán estos problemas en caso de existir.]

5.1.2 ¿Es necesario anonimizar datos (indique cuáles en su caso) o crear un comité de acceso a los datos

[Posibles problemas de confidencialidad, de tratamiento de datos sensibles, ... y la forma en que se abordarán estos problemas en caso de existir.]

6. Otros aspectos

6.1 Otros aspectos

6.1.1 Señale cualquier otro aspecto que considere de interés para la gestión del dataset.