

**Planes de Gestión de Datos** (Versión 1, Enero 2020)

Document abreujat i adaptat a la UPC pel Servei de Biblioteques, Publicacions i Arxius de la UPC

Document original: (Doc.CO20/02) (4 RDM\Plans Gestio Dades\PlanesdeGestiondeDatos\_v1\_enero20-ES.docx, 03.02.20) : <http://hdl.handle.net/2072/370510>

# Resumen de los datos



* 1. **Describir los objetivos de los datos generados y/o recolectados y su relación con los objetivos del proyecto. Especificar el tipo y formato de los datos que se generarán/recolectarán**
* Explicar de manera breve el objetivo de los datos generados y/o recolectados y su relación con los objetivos del proyecto.
* Describir el contenido y alcance de los datos. Los datos de investigación se generan por diferentes motivos a través de diferentes procesos y pueden ser de los siguientes tipos:
  + 1. Observacionales: datos capturados en tiempo real (neuroimágenes, datos de muestras, datos de sensores, datos de encuestas, etc.).
    2. Experimentales: datos capturados en equipos de laboratorio (secuencias de genes, cromatogramas, datos de campos magnéticos, etc.).
    3. Simulación: datos generados a partir de modelos de prueba (climatológicas, matemáticas o modelos económicos, etc.).
    4. Derivadas o compiladas: datos reproducibles pero de difícil reproducción (texto y minería de datos, modelos 3D, base de datos compilada, etc.).
    5. De referencia: conglomerado o conjunto de datos (bancos de datos de secuencias de genes, estructuras químicas, portales de datos espaciales, etc.).
    6. Otros. El formato de los datos (texto, numérico, imagen, etc.) también debería indicarse.
  1. **Especificar si se reutilizan los datos existentes y cómo**

Si se reutiliza un conjunto de datos (dataset), especificar la fuente de donde se ha extraído, por ejemplo, de un repositorio relevante. Si se compran o reutilizan fuentes de datos existentes, explicar cómo se han tratado temas como los derechos de autor y la ley de propiedad intelectual. Cuando se crean nuevas fuentes de datos, explicar por qué no se reutilizan las fuentes ya existentes.

* 1. **Especificar el origen de los datos**
* Indicar el origen de los datos, si éstos se generan dentro del proyecto
* Indicar la fuente de donde se han obtenido, si éstos son recolectados.

**1.D Indicar el tamaño esperado de los datos**

Indicar, de manera aproximada, el volumen del conjunto de datos (datasets). Hay que considerar el volumen de los datos en términos de almacenamiento, copias de seguridad, costes y acceso. Estimar el volumen de los datos en MB / GB / TB y el crecimiento gradual para asegurar que el almacenamiento y el soporte técnico adicional puedan ser proporcionados.

**1.E** **Describir la utilidad de los datos: para quienes pueden ser útiles**

Indicar para qué colectivo pueden ser de interés los datos

# Datos FAIR



Los datos tienen que seguir los principios **FAIR**: Findable, Accessible, Interoperable, Reusable

## Datos “findable”: Encontrables (incluyendo los metadatos)

* + 1. **Describir el descubrimiento y la identificación de los datos producidos/usados**
* Descubrimiento de los datos (incluyendo los metadatos): Los metadatos deben documentar cómo se generaron los datos, bajo qué licencia y cómo se reutilizarán. Además, los metadatos ayudan a descubrir nuevos datos y proporcionan el contexto para otras interpretaciones adecuadas.
* Identificar los datos y consultar el mecanismo de identificación estándar: Explicar cómo los datos y los metadatos asignan un identificador permanente (DOI, handle…).

Ejemplo: en la UPC, [CORA. Repositori de Dades de Recerca (RDR)](https://dataverse.csuc.cat/) repositorio de datos de investigación de las universidades catalana i de los centros CERCA, proporciona un identificador permanente DOI a nivel de dataset i a nivel de fichero para la identificación permanente y la citabilidad del dataset.

**2.1.E Especificar qué metadatos se crearán**

Describir los estándares de metadatos y cómo se utilizarán.

## Ejemplo: [CORA.Repositori de Dades de Recerca (RDR)](https://dataverse.csuc.cat/) utiliza metadatos genéricos o específicos por disciplinas favoreciendo la visibilidad de los datos.

## Datos “accessibles”: Accesibles

* + 1. **Especificar cómo se accederá a los datos (ej. depositando en un repositorio)**

Describir cómo se compartirán los datos, es decir, quién tendrá acceso al conjunto de datos. Se puede crear un procedimiento para que temporalmente los datos se hagan accesibles al resto de miembros del grupo, los socios del proyecto, y al público en general. Hay que indicar si los datos se pondrán en acceso abierto y en qué plazo razonable. Una posibilidad es ofrecer juntamente con las publicaciones. Si hay que establecer un periodo de embargo, es aquí donde hay que especificarlo.

**2.2.C.D.** **Especificar qué métodos o softwares son necesarios para acceder a los datos. Incluir la documentación del software que se necesita para acceder a los datos. Especificar en la documentación qué software es necesario para acceder a los datos**

**2.2.E Especificar dónde se depositan los datos, metadatos, documentación y el código**

Indicar el repositorio en el que los datos, los metadatos, la documentación y el código se almacenarán. Puede ser en el mismo repositorio o en diferentes según el tipo de contenido. Por ejemplo, el código debería depositarse en un repositorio específico. Se puede consultar el documento de [recomendaciones para seleccionar un repositorio de](http://hdl.handle.net/2072/284974) [datos del CSUC.](http://hdl.handle.net/2072/284974)

Es importante usar repositorios que proporcionen a los datos enlaces permanentes (DOI, Handle) para facilitar el descubrimiento y la citación.

## Datos interoperables

## Explicar qué datos y qué vocabularios de metadatos, estándares o metodologías se siguen para facilitar la interoperabilidad

Ejemplo CORA.Repositori de Dades de Recerca (RDR):

El repositorio [CORA. Repositori de Dades de Recerca (RDR)](https://dataverse.csuc.cat/) utiliza el protocolo de interoperabilidad Open Archives Initiative (OAI-PMH) para aumentar la visibilidad de los datos depositados. Utiliza metadatos genéricos o específicos por disciplina.

## Así mismo, para garantizar el acceso directo a los datos, se proporcina un DOI para cada conjunto de datos y fichero alojado en CORA.Repositori de Dades de Rercerca (RDR).

## Datos reutilizables

* + 1. **Especificar cómo se asignan las licencias para permitir la máxima la reutilización de los datos**

Si los datos se ponen a disposición de otros investigadores y del público en general, especificar el grado de reutilización permitido. Se establecerá el grado de reutilización a través de licencias. La CE propone el uso de licencias Creative Commons CC-BY o CC0, pero existen otras.

# Asignación de recursos



* 1. **Estimar el coste para hacer que los datos sean FAIR en el proyecto**

Ejemplo: Los repositorios institucionales y servidores son gratuitos para los investigadores relacionados con el proyecto.

Se tienen que tener en cuenta los costes asociados a la publicación en acceso abierto.

# Seguridad de los datos



* 1. **Tratamiento de recuperación de datos, así como el almacenaje y la transferencia de datos confidenciales**

Describir las medidas adoptadas respecto a la seguridad de los datos. Incluye el almacenaje y las copias de seguridad, la preservación a largo plazo y cómo se mantendrán los datos seguros y protegidos.

# Problemas éticos o legales que pueden afectar la recogida e intercambio de datos



# Describir los posibles problemas éticos durante la recogida, el almacenaje, el procesamiento y el archivo de los datos, junto a los procedimientos de aprobación ética relacionados con el proyecto. Es importante destacar en este punto cualquier aspecto mencionado en el artículo 34 de los acuerdos de financiamiento “[Article 34 — Ethics and research integrity](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/amga/h2020-amga_en.pdf)”.

Si las actividades de investigación involucran a niños, pacientes, población vulnerable, uso de células madre embrionarias, cuestiones de privacidad y protección de datos o investigación en animales y primates, se deben cumplir los principios éticos y la legislación nacional pertinente de la Unión Europea e internacional, y en este punto es donde debe tratarse.

1. **Explicar los recursos utilizados para el desarrollo del plan**

Lista los procedimientos, las políticas o las referencias que se han usado en el desarrollo del plan. Ejemplo: Este DMP ha sido creado con “eiNA DMP”<https://dmp.csuc.cat>