



# DATOS DE INVESTIGACIÓN, DATOS FAIR Y PLAN DE GESTIÓN DE DATOS

Versión	Fecha	Modificaciones
3	Enero 2024	Consideraciones éticas y protección de datos Formatos de archivo Denominación de archivo Guía para realizar un PGD en Argos

**ÍNDICE**

<b>1. PLAN DE GESTIÓN DE DATOS: VISIÓN GENERAL.....</b>	<b>2</b>
¿Qué es un Plan de Gestión de Datos (Data Management Plan, DMP)? .....	2
¿Para qué sirve un DMP?.....	2
¿Por qué se consideran útiles los DMP? .....	2
¿Dónde encontrar Planes de Gestión de Datos reales?.....	3
<b>2. PRINCIPIOS FAIR PARA LA GESTIÓN Y ADMINISTRACIÓN DE DATOS CIENTÍFICOS.....</b>	<b>4</b>
<b>3. CONSIDERACIONES ÉTICAS Y DE PROTECCIÓN DE DATOS .....</b>	<b>6</b>
<b>4. CONSIDERACIONES AL MOMENTO DE ELEGIR FORMATOS DE ARCHIVO<sup>4</sup> .....</b>	<b>7</b>
<b>5. GENERADORES DE PLANES DE GESTIÓN DE DATOS.....</b>	<b>10</b>
<b>6. GUÍA PARA PGD BASADA EN ARGOS.....</b>	<b>12</b>
<b>7. BIBLIOGRAFÍA.....</b>	<b>21</b>

## 1. PLAN DE GESTIÓN DE DATOS: VISIÓN GENERAL

Este documento está destinado a ayudar a los investigadores e investigadoras a crear sus propios Planes de Gestión de Datos FAIR como requisito en proyectos financiados por distintas entidades como por ejemplo, el programa Horizon Europe de la Unión Europea o el Instituto de Salud Carlos III (ISCIII).

### **¿Qué es un Plan de Gestión de Datos (Data Management Plan, DMP)?**

Un Plan de Gestión de Datos o DMP es un documento formal que describe cómo se van a manejar los datos durante y una vez finalice un proyecto de investigación.

### **¿Para qué sirve un DMP?**

El objetivo principal de un Plan de Gestión de Datos es considerar los muchos aspectos de la gestión de datos: la generación de metadatos, la preservación y el análisis de los mismos antes de que comience el proyecto de forma que los datos estén bien administrados a lo largo de toda su ejecución<sup>1</sup>.

De forma general, un DMP debe plantearse las siguientes preguntas:

1. ¿Qué tipo de datos se van a recopilar y cómo serán descritos?
2. ¿Cómo se almacenarán y mantendrán seguros?
3. ¿Se les permitirá dar acceso una vez finalizado el proyecto? ¿Quién podrá acceder a ellos, en qué condiciones y durante cuánto tiempo?

### **¿Por qué se consideran útiles los DMP?**

Además de ser a menudo un requisito para recibir subvenciones, estos planes que se elaboran antes de iniciar el proyecto resultan útiles para<sup>2</sup>:

- Garantizar que los datos estén en el formato correcto, bien organizado y mejor anotado ahorrando así tiempo y recursos a lo largo de la ejecución del proyecto evitando así reorganizar, formatear o tratar de recordar detalles sobre los mismos.
- Desarrollar una estrategia previa a la ejecución del proyecto acerca de cuestiones como el almacenamiento de datos, la preservación a largo plazo o el manejo de datos sensibles de forma que permita anticipar

- requisitos legales, éticos y comerciales relativos a la divulgación de datos; decidir quién puede tener acceso a los datos a corto y largo plazo.
- Gestionar y documentar los datos a lo largo de su ciclo de vida permite comprender y utilizar los datos en el futuro.
  - Estimar los costes asociados que luego se pueden incluir en el presupuesto final del proyecto.
  - El depósito de datos en un repositorio contribuye a la investigación para que puedan ser reutilizados, evita la duplicación de estudios científicos ya realizados y muestra además su apoyo al acceso abierto (Open Access, OA).

### ¿Dónde encontrar Planes de Gestión de Datos reales?

Cada vez hay más investigadores e investigadoras que publican sus DMP como un resultado más de sus proyectos y lo hacen en repositorios en abierto. A continuación se enumeran varios de estos portales donde se ofrecen varios de planes reales:

1. Zenodo: [data management plan](#)
2. Digital Curation Centre: [DMP online](#)
3. Digital Curation Centre: [example DMPs and guidance](#)
4. [DMP Tool](#)
5. [ARGOS](#)

## 2. PRINCIPIOS FAIR PARA LA GESTIÓN Y ADMINISTRACIÓN DE DATOS CIENTÍFICOS

En 2016, se publicó en *Scientific Data* los “Principios FAIR para la gestión y administración de datos científicos”<sup>3</sup>. En este trabajo, los autores tenían la intención de proporcionar pautas para mejorar la capacidad de búsqueda, accesibilidad, interoperabilidad y reutilización de datos digitales. Los principios, como resultado del aumento en el volumen, complejidad y velocidad de creación de datos, se centran en ensalzar la capacidad de los sistemas computacionales para encontrar, acceder, interoperar y reutilizar datos con ninguna o mínima intervención humana.

Los principios se refieren a tres tipologías: datos (o cualquier objeto digital), metadatos (información sobre ese objeto digital) e infraestructura.

A continuación se describen los conceptos básicos de los principios FAIR: Findable, Accessible, Interoperable and Reusable.

**Findable** - El primer paso para (re)utilizar datos es encontrarlos. Los metadatos y los datos deben ser fáciles de encontrar tanto para los humanos como para los sistemas informáticos. Para ello es indispensable que tengan asignado un identificador global único y persistente (por ejemplo, DOI).

**Accesible** - Una vez que el usuario encuentra los datos, necesita saber cómo se puede acceder a ellos, posiblemente incluyendo autenticación y autorización. Para ello los metadatos y los datos deben ser accesibles utilizando un protocolo abierto, gratuito y de implementación universal que permita un procedimiento de autenticación y autorización, cuando sea necesario. Es importante mencionar que los metadatos deben ser accesibles, incluso cuando los datos ya no estén disponibles.

**Interoperable** - Por lo general, los datos deben poder integrarse con otros datos. Además, deben poder interoperar con aplicaciones o sistemas de análisis, almacenamiento y procesamiento. Para ello, se considera que los metadatos y datos deben utilizar un lenguaje formal, accesible y de amplia aplicación del conocimiento que representan.

**Reusable** - Los metadatos y los datos deben estar bien descritos para que puedan replicarse y/o combinarse en diferentes entornos. Deben además estar publicados con una licencia de uso de datos clara y accesible y deben además estar asociados con detalle a la autoría de pertenencia.

Herramientas para saber si los datos que estamos trabajando son FAIR<sup>4</sup>:

- Australian Research Data Commons' FAIR data self-assessment: <https://ardc.edu.au/resources/aboutdata/fair-data/fair-self-assessment-tool>
- FAIR AWARE: <https://www.fairsfair.eu/fair-aware>
- F-UJI tool to assess FAIRness of datasets: <https://www.f-ugi.net>

## ¿Dónde depongo mis datos?

Existen repositorios tanto gratuitos como privados a disposición de cualquier usuario, que cumplen con los principios FAIR y donde poder depositar los datos. Recomendación: al subir los datos adjuntar siempre el “readme.txt” (ver [aquí](#)), fichero en el que se describe la información necesaria para que los conjuntos de datos sean comprensibles y reutilizables: autoría, título, descripción, metodología, proyectos financiadores, cobertura temporal y geográfica, derechos de uso y privacidad, etc.

Existe un buscador de repositorios FAIR: [Repository Finder](#). Entre los más utilizados están:

**Zenodo**: Repositorio de acceso abierto desarrollado en 2013 bajo el programa europeo OpenAIRE y gestionado por y alojado en la Organización Europea para la Investigación Nuclear (CERN). Desde 2015 admite también el depósito de conjuntos de datos, permite carga de archivos hasta 50 GB. Proporciona DOIs y asigna metadatos de acceso abierto, licencias, etc.

**DRYAD**: repositorio internacional de acceso abierto de datos de investigación, especialmente datos subyacentes a publicaciones científicas y médicas.

**Figshare**: Repositorio de gestión privada, desde 2012 está financiado por empresa Digital Science. Permite cargar cualquier tipo de resultado de investigación y en cualquier formato y con un tamaño límite de 5 GB, atribuye

DOIs, permite establecer embargos, añadir licencias, enlazar con proyectos financiados, etc.

### 3. CONSIDERACIONES ÉTICAS Y DE PROTECCIÓN DE DATOS

Los **datos confidenciales** o sensibles son datos que deben protegerse contra una divulgación no deseada y su acceso debe ser salvaguardado. La protección de este tipo de datos puede ser necesaria por razones legales o éticas, por cuestiones relacionadas con la privacidad personal o por consideraciones de propiedad.

La UE tiene regulaciones estrictas con respecto a los datos personales (como es el Reglamento general de protección de datos - RGPD) y los datos no personales confidenciales. Ejemplos de casos de datos confidenciales:

- Datos personales: identificadores como nombres o números de identificación, características físicas, fisiológicas, genéticas, mentales, económicas, culturales o sociales, también incluye datos de ubicación de GPS o teléfonos móviles.
- Datos confidenciales: secretos comerciales, investigaciones, datos protegidos por derechos de propiedad intelectual Seguridad: contraseñas, información financiera, seguridad nacional, información militar...
- Combinación de diferentes conjuntos de datos que se pueden combinar en datos confidenciales o personales.
- Datos biológicos: especies (plantas o animales) en peligro de extinción, donde su supervivencia depende de la protección de sus datos de ubicación (comunidad de la biodiversidad)
- Metadatos personales y confidenciales.

**IMPORTANTE:** Existen herramientas que ayudan a preparar los datos para su depósito en abierto (herramientas de anonimización, pseudonimización o encriptación como las disponibles en OpenAIRE), pero si ninguna de estas opciones fuera posible, los datos no deben hacerse abiertos y deben archivarse bajo una licencia cerrada. En el caso de que los datos no se pueden exportar ni extraer, si se dan las razones oportunas, rige el "tan cerrado como sea necesario", pero los metadatos del dataset si se puede compartir y subir junto

con un resumen de su contenido. Toda esta información deberá estar recogida en el PGD.

Mencionar también que a la hora de compartir datos de investigación con otras entidades es importante contar con un acuerdo de transferencia de datos o *Data Transfer Agreement* que aseguren que se cumple con todos los requisitos legales.

#### **4. CONSIDERACIONES AL MOMENTO DE ELEGIR FORMATOS DE ARCHIVO<sup>4</sup>**

- Elegir formatos comunes al campo disciplinar al que se está trabajando: Para asegurar la interoperabilidad y la reutilización de los datos.
- Tener en cuenta el tiempo en que se espera conservar los datos: cuanto mayor sea el periodo de tiempo que se desea conservar los datos, mayor será la necesidad de seleccionar formatos abiertos, estandarizados y bien documentados.
- La conversión de archivos puede provocar la pérdida de datos: se han de considerar formatos de multiplataforma común que respondan a estándares específicos para evitar dicha pérdida.
- Verificar los requisitos del repositorio de datos: muchas revistas, archivos y repositorios requieren de formatos específicos al momento de cargar la información

En la siguiente tabla se muestran, según el [UK Data Service](#), los formatos recomendados y aceptados para los distintos tipos de datos existentes:

TYPE OF DATA	RECOMMENDED FORMATS	ACCEPTABLE FORMATS
<b>Tabular data with extensive metadata variable labels, code labels, and defined missing values</b>	SPSS portable format (.por), delimited text and command ('setup') file (SPSS, Stata, SAS, etc.), structured text or markup file of metadata information, e.g. DDI XML file	proprietary formats of statistical packages: SPSS (.sav), Stata (.dta), MS Access
<b>Tabular data with minimal metadata</b>	comma-separated values (.csv); tab-delimited file (.tab), delimited text with SQL data definition statements	delimited text (.txt) with characters not present in data used as delimiters; widely-used formats: MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb), dBase (.dbf), OpenDocument Spreadsheet (.ods)
<b>Geospatial data, vector and raster data</b>	ESRI Shapefile (.shp, .shx, .dbf, .prj, .sbx, .sbnooptional); geo-referenced TIFF (.tif, .tfw); CAD data (.dwg); tabular GIS attribute data; GeographyMarkupLanguage(.gml)	ESRI Geodatabase format (.mdb); MapInfo Interchange Format (.mif) for vector data; Keyhole Mark-up Language (.kml); Adobe Illustrator (.ai), CAD data (.dxf or .svg); binary formats of GIS and CAD packages
<b>Textual data</b>	RichTextFormat(.rtf); plaintext, ASCII(.txt); HypertextMark-upLanguage(.html); widely-eXtensibleMark-upLanguage(.xml) text according to an appropriate Document Type Definition (DTD) or some specific formats: NUD*IST, NVivo and ATLAS.ti schema	used formats: MSWord (.doc/.docx);
<b>Imagedata</b>	TIFF6.0 uncompresssed(.tif)	JPEG(.jpeg,.jpg,.jp2); GIF(.gif); TIFF other versions (.tif,.tiff); RAW image format (.raw); Photoshop files (.psd); BMP (.bmp); PNG (.png); Adobe Portable Document Format (PDF/A, PDF) (.pdf)
<b>Audiodata</b>	FreeLosslessAudioCodec (FLAC) (.flac)	MPEG-1 Audio Layer 3 (.mp3) if original created in this format; Audio Interchange File Format (.aif); Waveform Audio Format (.wav)
<b>Videodata</b>	MPEG-4 (.mp4); OGG video (.ogg, .ogg); motion JPEG 2000 (.mj2)	AVCHD video (.avchd)
<b>Documentation and scripts</b>	RichTextFormat (.rtf); PDF/UA, PDF/A or PDF (.pdf); XHTML or HTML (.xhtml, .htm); OpenDocument Text (.odt)	plaintext (.txt); widely-used formats: MSWord (.doc/.docx), MS Excel (.xls/.xlsx); XML marked-up text (.xml) according to an appropriate DTD or schema, e.g. XHTML 1.0

## **FORMATOS Y ESTÁNDARES PARA METADATOS**

Los metadatos a menudo se guardan como XML u otro formato de lenguaje de marcado (por ejemplo, JSON). Los textos de origen XML son legibles tanto por máquina como por humanos y, además, se pueden transferir a otros formatos.

Para aumentar la efectividad de los metadatos, los repositorios y la comunidad científica han creado estándares definidos para metadatos, permitiendo vincular y editar metadatos de diferentes fuentes. Un estándar de metadatos simple y generalizado es [Dublin Core](#). Consta de 15 elementos (por ejemplo, <dc: creator> Max Mustermann</ dc: creator>) y varios elementos secundarios.

La [Research Data Alliance](#) (RDA) tiene una lista de estándares de metadatos disciplinarios para datos científicos.

## 5. GENERADORES DE PLANES DE GESTIÓN DE DATOS

Actualmente existen varias herramientas gratuitas online que permiten generar Planes de Gestión de Datos. Desde el IIS Biodonostia se destacan a continuación algunas de estas:

- [DMPOnline](#): herramienta gratuita y *open source* recomendada por la Comisión Europea y contiene las plantillas para generar Planes de Gestión de Datos de diversas agencias financiadoras de investigación. Si, por el contrario, el financiador no se encuentra en la lista o bien no existe un financiador asociado al PGD, se crea una plantilla basada en la plantilla “DCC Template” proporcionada por el Digital Curation Centre y que responde a su propia guía: Checklist for a Data Management Plan. v.4.0. Edinburgh: Digital Curation Centre. Enlace: <http://www.dcc.ac.uk/resources/datamanagement-plans>. Permite agregar a las personas que van a contribuir a la elaboración del PGD.
- [DSW Wizard](#): Herramienta gratuita y *open source* desarrollada por ELIXIR para facilitar la creación de un plan de gestión de datos basado en lo que (actualmente) se consideran «buenas prácticas». Permite utilizar plantillas de distintos financiadores, así como el desarrollo de nuevas plantillas.
- [PGDOnline](#): herramienta creada por el Consorcio de Universidades de la Comunidad de Madrid y de la UNED para la Cooperación Bibliotecaria (Madroño) para la elaboración de Planes de Gestión de Datos, creada a partir de DMPOnline.
- [ARGOS](#): Es una herramienta online, gratuita y *open source* para la creación, gestión, difusión y enlace de un PGD, que hace énfasis en la aplicación de los principios FAIR y en las mejores prácticas para fomentar la accesibilidad a los datos de investigación. Tiene un funcionamiento abierto y colaborativo y ha sido desarrollada de forma conjunta entre OpenAIRE. Ofrece la mayoría de las plantillas necesarias para acceder a los fondos de investigación, así como la necesaria para solicitar los fondos ofrecidos en Horizonte Europa. Con cuenta de administrador existe la

posibilidad de proponer plantillas nuevas a medida de las necesidades de los autores y agencias de financiación.

Los PGD generados en ARGOS se gestionan a su vez como productos de investigación a los que se pueden asignar DOI, licencias, etc. Además, permite que los PGDs se hagan de forma colaborativa entre varias personas.

Todas estas herramientas tratan de integrar explicaciones, recomendaciones y sugerencias para cada sección del Plan y están cada vez más estandarizadas.

## 6. GUÍA PARA PGD BASADA EN ARGOS

Desde el IIS Biodonostia se ha creado una plantilla para la elaboración de un plan de gestión de datos (Horizon 2020 DMP), basada en la herramienta Argos, en ella se explica qué datos completar en cada uno de los apartados.

Importante: muchos de los campos requieren de información ofrecida por el propio repositorio.

### Información principal

Información administrativa como el título, descripción, nombres de los investigadores, organización a la que pertenecen, idioma y contacto.

### Información sobre financiación

Información sobre el alcance, finanziadores y participantes del PGD y el acceso y cuestiones de reutilización para la salida del PGD que está creando.

### Información sobre licencia

Cada PGD puede contener información específica sobre licencias, su visibilidad y disponibilidad, que puede determinar quién puede ver los dataset y cuánto tiempo estará restringido el acceso a los datos.

### Información sobre el dataset

Los datasets se documentan siguiendo plantillas predefinidas con el contenido de la descripción de los datasets. En Argos un PGD puede incluir tantas descripciones de datasets como datasets se documenten.

### CONJUNTO DE DATOS: DESCRIPCIÓN DEL DATASET

**0** **INFORMACIÓN PRINCIPAL: título, descripción, etiquetas y plantilla a utilizar (Horizon Europe)**

**1** **RESUMEN**

- |           |  |  |
|-----------|--|--|
| <b>1a</b> | <b>¿Qué tipo de resultados de investigación está describiendo?</b> | <ul style="list-style-type: none"> <li>• Los datos de investigación son información (particularmente hechos o números) recopilada para ser examinada y considerada, y para servir como base para el razonamiento, la discusión o el cálculo. Acceso abierto a datos de investigación: el derecho a acceder y reutilizar datos de investigación digitales según los términos y condiciones establecidos en el Acuerdo de subvención.</li> <li>• Otros resultados de la investigación se refieren a los resultados que se producen o reutilizan durante</li> </ul> |
|-----------|--|--|

el ciclo de vida de la investigación y la gestión de datos. Pueden ser objetos, instrumentos y materiales en forma digital, como software, flujos de trabajo, protocolos, modelos, etc. Igualmente, en su forma analógica pueden incluir nuevos materiales, anticuerpos, reactivos, muestras, etc.

**1b** ¿Es físico o digital?

**1c** ¿Lo estás generando o reutilizando?

- Además de los datos producidos dentro de su investigación, es posible que desee reutilizar los datos que otros han producido y compartido, en un contexto de investigación diferente.
- Si planea utilizar los datos de otros investigadores en su investigación, hay un par de cosas que debe tener en cuenta, como los derechos de autor de los conjuntos de datos, los permisos proporcionados a través de licencias, los aspectos éticos para la reutilización, etc.
- Más información: <https://www.openaire.eu/can-i-reuse-someone-else-research-data>

**1d** ¿De qué tipo es el conjunto de datos descrito?

- La principal distinción entre tipos de datos es entre datos primarios y datos secundarios. La recopilación de datos puede contener tanto datos primarios como secundarios según la fuente de la que se hayan derivado.
- Los datos primarios son datos que se han recopilado por primera vez y que aún no se han sometido a un procesamiento y/o análisis de datos.
- Los datos secundarios son datos que han sido limpiados, analizados y compartidos por otros (publicados o no publicados) y son los que normalmente se reutilizan.

**1e** ¿Cuál es su formato?

- Los diferentes tipos de datos se adquieren, procesan y almacenan (conservan y/o archivan) de diferentes maneras y pueden ser específicos de la disciplina.

**1f** ¿Cuál es su tamaño esperado?

**1g** ¿Por qué lo está recolectando/generando o reutilizando?

- La recopilación de datos generalmente se encuentra en las etapas iniciales de los ciclos de vida de gestión de datos de investigación para establecer el trasfondo de lo que se necesita (generación de datos), lo que ya está allí (reutilización de datos) y cómo usarlo mejor para cumplir con los objetivos del proyecto (por qué). Aquí puede agregar información sobre el alcance y los objetivos de su proceso de recopilación de datos.

<b>1e</b>	<b>¿Cuál es su origen/procedencia?</b>	
<b>1f</b>	<b>¿Para quién podría ser útil ('utilidad de datos')?</b>	
<b>2</b>	<b>VÍNCULOS ENTRE PRODUCTOS</b>	
<b>2a</b>	<b>¿El producto descrito respalda alguna publicación científica?</b>	<ul style="list-style-type: none"> <li>Crear los vínculos entre las publicaciones científicas y los datos de investigación.</li> </ul>
<b>2b</b>	<b>¿Se proporciona una declaración de disponibilidad de datos junto con la publicación?</b>	<ul style="list-style-type: none"> <li>Explique cómo se controlará y documentará la coherencia y la calidad de la recopilación de datos. Esto puede incluir procesos como calibración, muestras o mediciones repetidas, captura de datos estandarizados, validación de entrada de datos, revisión de datos por pares o representación con vocabularios controlados.</li> </ul>
<b>2c</b>	<b>¿El resultado descrito utiliza o admite algún software?</b>	
<b>3</b>	<b>PRÁCTICAS FAIR</b>	
<b>3.1</b>	<b><u>LOCALIZABLES</u></b>	
<b>3.1a</b>	<b>¿Qué tipo(s) de identificador(es) persistente(s) se utilizan para el conjunto de datos/salida descrito?</b>	<ul style="list-style-type: none"> <li>Los identificadores persistentes y/o permanentes (PID) identifican de forma única objetos, personas, organizaciones y actividades y pueden garantizar que se pueda acceder a la producción científica incluso cuando la URL del sitio web haya cambiado. Los PID se pueden asignar a resultados de investigación, incluidas publicaciones, datos y software/código. Los PID también se pueden asignar a investigadores, muestras, organizaciones y proyectos. Un PID puede estar conectado a un registro de metadatos que describe un elemento en lugar del elemento en sí. Los PID generalmente son proporcionados por repositorios de datos y otras plataformas de depósito.</li> </ul>
<b>3.1b</b>	<b>¿Proporcionará metadatos para el conjunto de datos/salida descrito?</b>	<ul style="list-style-type: none"> <li>Los metadatos son datos sobre datos y son un conjunto esencial de información que describe resultados científicos, en forma de objetos físicos o digitales, en un formato legible por máquina. De acuerdo con el uso esperado, los metadatos pueden recibir diferentes atributos.</li> <li>El tipo más común que permite el descubrimiento y la identificación son los metadatos descriptivos. Los metadatos descriptivos contienen información sobre los aspectos clave necesarios para buscar y encontrar con éxito un producto científico determinado, por ejemplo, por su título, autor/creador, resumen, palabras clave. Además,</li> </ul>

<p>los metadatos pueden utilizarse para describir un servicio o un instrumento científico.</p>		
<b>3.1c</b>	<b>¿Qué tipo(s) de metadatos?</b>	<ul style="list-style-type: none"> <li>Describa qué tipo de metadatos se crearán y cómo.</li> </ul>
<b>3.1d</b>	<b>¿Los metadatos usan vocabularios estandarizados?</b>	<ul style="list-style-type: none"> <li>Los vocabularios estandarizados permiten una mayor interoperabilidad entre sistemas. Hay normas de metadatos genéricas y específicas de la disciplina que se han compilado para contener información vital que permita el intercambio teniendo en cuenta las especificidades del trabajo y las demandas en diferentes disciplinas, incluso en diferentes áreas de trabajo dentro de una disciplina.</li> </ul>
<p><b>3.1e</b> Proporcione URL/Descripción de los vocabularios utilizados</p>		
<b>3.1f</b>	<b>¿Se pueden buscar los metadatos?</b>	<ul style="list-style-type: none"> <li>Los metadatos de búsqueda son metadatos indexados por motores de búsqueda y son identificables por rastreadores web.</li> </ul>
<p><b>3.1g</b> ¿Cómo se proporcionan los metadatos de búsqueda?</p>		
<b>3.1h</b>	<b>¿Se proporcionan palabras clave en los metadatos?</b>	<ul style="list-style-type: none"> <li>Las palabras clave y las etiquetas ayudan a optimizar la posibilidad de descubrimiento y luego la posible reutilización de datos.</li> </ul>
<b>3.1i</b>	<b>¿Se pueden recolectar los metadatos?</b>	<ul style="list-style-type: none"> <li>Por lo general, esto lo facilitan los servicios que aplican protocolos estándar para el intercambio de información. Por ejemplo, en los repositorios, el intercambio de metadatos está soportado por el protocolo OAI-PMH.</li> </ul>
<p><b>3.2</b> <u>ACCESIBLES:</u> repositorios, datos y metadatos</p>		
<b>3.2a</b> <b>Repositorios</b>	<b>¿En qué repositorio se depositará el conjunto de datos/salida?</b>	<ul style="list-style-type: none"> <li>Un repositorio de datos, también conocido como archivo de datos, es una colección en línea de conjuntos de datos que se describen y clasifican de una manera estándar que facilita el descubrimiento y la recuperación de datos tanto para humanos como para máquinas.</li> </ul>
<b>3.2b</b> <b>Repositorios</b>	<b>¿El repositorio seleccionado es una fuente confiable?</b>	<p>Los repositorios de confianza asumen un papel central en Horizon Europe para el depósito y el acceso a publicaciones y datos de investigación. Están:</p> <ul style="list-style-type: none"> <li>Repositorios certificados (por ejemplo, CoreTrustSeal, nestor Seal DIN31644, ISO16363) o repositorios disciplinarios y de dominio comúnmente utilizados y respaldados por las</li> </ul>

comunidades de investigación. Dichos repositorios deben ser reconocidos internacionalmente.

- Repositorios de propósito general o repositorios institucionales que presenten las características esenciales de los repositorios de confianza, es decir:

- ❖ presentan características específicas de calidad organizativa, técnica y procedural tales como servicios, mecanismos y/o prestaciones que tienen por objeto asegurar la integridad y autenticidad de sus contenidos, facilitando así su uso y reutilización a corto y largo plazo. Los repositorios de confianza cuentan con disposiciones específicas y ofrecen información explícita en línea sobre sus políticas, que definen sus servicios (p. ej., adquisición, acceso, seguridad del contenido, sostenibilidad a largo plazo del servicio, incluida la financiación, etc.).
- ❖ proporcionar un acceso amplio, equitativo e idealmente abierto al contenido de forma gratuita en el punto de uso, según corresponda, y respetar las limitaciones legales y éticas aplicables. Asignan identificadores únicos persistentes a los contenidos (p. ej., DOI, identificadores, etc.), de modo que los contenidos (publicaciones, datos y otros resultados de la investigación) se referencian inequívocamente y, por lo tanto, se pueden citar. Aseguran que los contenidos vayan acompañados de metadatos suficientemente detallados y de suficiente calidad para permitir el descubrimiento, la reutilización y la cita, y que contengan información sobre la procedencia y la concesión de licencias; los metadatos son accionables por máquina y estandarizados (p. ej., Dublin Core, Data Cite, etc.) preferiblemente utilizando formatos comunes no patentados y siguiendo los estándares de la comunidad respectiva a la que sirve el repositorio, cuando corresponda.

- 
- ❖ facilitar la conservación a medio y largo plazo del material depositado. Cuentan con mecanismos o disposiciones para la curación experta y garantía de calidad para la precisión e integridad de conjuntos de datos y metadatos, así como procedimientos para comunicarse con los depositantes cuando se detectan problemas. Cumplen con los criterios de seguridad internacionales y nacionales generalmente aceptados para evitar el acceso no autorizado y la publicación de contenido y tienen diferentes niveles de seguridad según la sensibilidad de los datos que se depositan para mantener la privacidad y la confidencialidad.

---

<b>3.2c Repositorios</b>	<b>Agregar los arreglos apropiados hechos con el(los) repositorio(s) donde se depositará el conjunto de datos descrito</b>
<b>3.2d Repositorios</b>	<b>¿El(los) repositorio(s) asigna(n) conjuntos de datos/resultados con identificadores persistentes?</b>
<b>3.2e Repositorios</b>	<b>¿El(los) repositorio(s) resuelve(n) los identificadores en un objeto digital?</b>
<b>3.2f Repositorios</b>	<b>¿El repositorio admite el control de versiones?</b>
<b>3.2a Datos</b>	<b>¿Cuál es el título del conjunto de datos/salida descrito?</b>
<b>3.2b Datos</b>	<b>¿Cómo se comparte el conjunto de datos/salida?</b>
<b>3.2c Datos</b>	<b>¿Se requieren métodos o herramientas para acceder al conjunto de datos/salida?</b>
<b>3.2d Datos</b>	<b>¿El conjunto de datos/salida descrito está respaldado por un comité de acceso a datos?</b>
<b>3.2e Datos</b>	<b>Especifique cómo se accederá al conjunto de datos/salida durante y después de que finalice el proyecto</b>

---

<b>3.2f Datos</b>	<b>Especifique cuánto tiempo después de que el proyecto haya finalizado, el conjunto de datos/salida estará disponible</b>	Consulte las políticas de los patrocinadores o instituciones que podrían aplicar un embargo sobre el acceso a los datos.
<b>3.2a Metadatos</b>	<b>¿Proporcionará metadatos incluso si el conjunto de datos/salida descrito no se puede compartir abiertamente?</b>	
<b>3.2b Metadatos</b>	<b>¿Bajo qué licencia se proporcionarán los metadatos?</b>	De acuerdo con el Acuerdo de subvención, los metadatos se ponen a disposición abiertamente y se licencian bajo una dedicación de dominio público CC0 .
<b>3.2c Metadatos</b>	<b>¿Proporcionan los metadatos información sobre cómo acceder al conjunto de datos/salida descrito?</b>	
<b>3.2d Metadatos</b>	<b>¿Seguirán estando disponibles los metadatos después de que el conjunto de datos/salida ya no esté disponible?</b>	
<b>3.3 <u>INTEROPERABLES</u></b>		
<b>3.3a</b>	<b>¿Sus (meta)datos utilizan un vocabulario controlado?</b>	Los vocabularios controlados proporcionan terminología estándar en lugar de palabras clave o etiquetas utilizadas para clasificar la información. Ejemplos: taxonomías, ontologías, tesauros.
<b>3.3b</b>	<b>Si usted creó el vocabulario, ¿dónde se puede encontrar?</b>	
<b>3.3c</b>	<b>¿Ha aplicado un esquema estándar para sus (meta)datos?</b>	
<b>3.3d</b>	<b>¿Cuál es la metodología seguida?</b>	
<b>3.3e</b>	<b>¿Qué mejores prácticas de interoperabilidad aprobadas por la comunidad se siguen?</b>	
<b>3.3f</b>	<b>¿El conjunto de datos/salida descrito proporciona referencias calificadas con otras salidas?</b>	Una referencia calificada es una referencia cruzada que explica su intención
<b>3.4 <u>REUTILIZABLES</u></b>		
<b>3.4a</b>	<b>¿Qué licencia reconocida internacionalmente usará para su conjunto de datos/salida?</b>	Hay una serie de licencias que se pueden asignar a los resultados de su investigación. Hay diferentes licencias para software y para publicaciones y datos. Para estos últimos, una práctica común es el uso de licencias Creative Commons que son legibles por máquina.
		Para obtener más información sobre las licencias necesarias para los datos de Horizon Europe, consulte la AGA (artículo 17).

<b>3.4b</b>	<b>¿Qué métodos de reutilización y/o reproducibilidad se siguen?</b>	Proporcionar material complementario que valide el análisis de datos y facilite la reutilización de datos
<b>3.4c</b>	<b>¿Proporcionará el conjunto de datos/salida descrito en el dominio público?</b>	
<b>3.4d</b>	<b>¿Tiene la intención de garantizar la (re)utilización por parte de terceros después de que finalice su proyecto?</b>	A pesar de dónde se almacenan los datos para las actividades diarias durante la vida útil del proyecto, hay ciertas opciones para elegir entre el almacenamiento de datos que garantiza la reutilización en la ejecución del registro. Los más populares son los archivos de datos, que de otro modo se encuentran como repositorios. Dependiendo del volumen de datos, el archivo puede estar sujeto a pequeñas tarifas.
<b>3.4e</b>	<b>¿Está bien documentada la procedencia?</b>	La procedencia de los datos, o el linaje de los datos, es un tipo de metadatos que rastrea la historia de un objeto digital en el tiempo y sus relaciones con otras entidades. También puede consultar: <a href="https://casrai.org/term/provenance-metadata/">https://casrai.org/term/provenance-metadata/</a>
<b>3.4f</b>	<b>¿Qué procedimientos documentados para el aseguramiento de la calidad tiene implementado?</b>	Los procedimientos de garantía de calidad están estrechamente relacionados con RDM y están destinados a validar que los datos estén limpios (sin duplicados ni inconsistencias), libres de errores, bien estructurados y representados en el análisis. Los controles de garantía de calidad pueden facilitarse mediante el uso de herramientas o scripts.
<b>4 ASIGNACIÓN DE RECURSOS</b>		
<b>4a</b>	<b>¿Cuál será el costo de hacer JUSTO el resultado descrito?</b>	<ul style="list-style-type: none"> <li>La gestión de datos puede resultar costosa, especialmente cuando su planificación no ha sido suficiente desde el inicio del proceso de investigación. El costeo de la gestión de datos incluye, por ejemplo, el uso potencial de servicios y herramientas patentados o el esfuerzo adicional necesario para realizar tareas específicas o incluso para desarrollar herramientas desde cero.</li> </ul>
<b>4b</b>	<b>¿Cómo se cubrirá este costo?</b>	<ul style="list-style-type: none"> <li>Tenga en cuenta que los costos relacionados con la gestión de resultados/datos de investigación son elegibles como parte de la subvención de Horizon Europe (si cumplen con las condiciones del Acuerdo de subvención)</li> </ul>
<b>4c</b>	<b>Identificar las personas que serán responsables y su(s) papel(es) en la</b>	<ul style="list-style-type: none"> <li>Proporcione los nombres y las responsabilidades de las actividades de gestión y administración de datos de los investigadores o administradores de</li> </ul>

	<b>gestión del producto descrito</b>	datos que se realizan a lo largo del proyecto para el resultado descrito
<b>5</b>	<b>SEGURIDAD</b>	
<b>5a</b>	<b>¿Qué medidas de seguridad se siguen?</b>	<ul style="list-style-type: none"> <li>• Describa qué disposiciones existen para la seguridad de los datos</li> </ul>
<b>5b</b>	<b>¿Qué condiciones cumplen las medidas de seguridad?</b>	<ul style="list-style-type: none"> <li>• Describa las provisiones de datos, incluida la recuperación de datos, así como el almacenamiento/archivo seguro y la transferencia de datos confidenciales.</li> </ul>
<b>5c</b>	<b>¿Cómo preservará el conjunto de datos/resultados descritos a largo plazo?</b>	<ul style="list-style-type: none"> <li>• Describa las políticas de curación y preservación seguidas para el contenido del repositorio.</li> </ul>
<b>6</b>	<b>ASPECTOS ÉTICOS</b>	
<b>6a</b>	<b>¿Hay algún problema ético o legal que pueda tener un impacto en compartir el conjunto de datos/resultado descrito?</b>	
<b>6b</b>	<b>¿El conjunto de datos/salida descrito contiene información confidencial?</b>	
<b>6c</b>	<b>¿El conjunto de datos/salida descrito contiene datos personales?</b>	
<b>7</b>	<b>OTROS</b>	
		<b>¿Utiliza otros procedimientos para la gestión de datos?</b>

## 7. BIBLIOGRAFÍA

1. Burnette, M; Williams, S; Imker, H. "From Plan to Action: Successful Data Management Plan Implementation in a Multidisciplinary Project". *Journal of EScience Librarianship.* 5 (1): e1101. 2016. doi:10.7191/jeslib.2016.1101.
2. MIT Libraries.
3. Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data.* 2016 Mar 15;3:160018. doi: 10.1038/sdata.2016.18. Erratum in: *Sci Data.* 2019 Mar 19;6(1):6. PMID: 26978244; PMCID: PMC4792175.
4. "Gestión de Datos de Investigación en el marco de la Ciencia Abierta" Alicia Fátima Gómez, 2022.