MINISTERIO
DE CIENCIA, INNOVACIÓN
Y UNIVERSIDADES

Cofinanciado por
la Unión Europea

AGENCIA
ESTATAL DE
INVESTIGACIÓN

Instrucciones para la elaboración del Plan de Gestión de Datos (PGD)

Para completar la justificación científico-técnica final es indispensable rellenar el Plan de Gestión de Datos

Un Plan de Gestión de Datos es una declaración formal que describe los datos que se utilizan y producen durante el curso de las actividades de investigación. El PGD establece cómo se van a gestionar los datos en un proyecto de investigación, tanto durante el proyecto como una vez que éste ha finalizado. El PGD define la metodología de extracción de la información, la unidad de medida, los procesos de creación, limpieza y análisis, el formato, el software de lectura y procesamiento, las licencias y permisos de acceso y (re)utilización, su sistema de control de versiones, su autoría, su difusión, y su lugar de almacenamiento y preservación a largo plazo. **Este plan se creará al comienzo del proyecto de investigación financiado y, según la naturaleza y evolución del proyecto, puede que no sea un documento definitivo, sino que variará, se completará y/o modificará a lo largo del ciclo de vida de la investigación.**

La Comisión Europea proporciona orientación sobre cómo elaborar un plan de gestión de datos en sus directrices para gestionar datos según el esquema FAIR.

Se puede utilizar cualquier **plantilla o herramientas** para crear el Plan de Gestión de Datos. Se trata de un conjunto de preguntas que debe responderse con un nivel de detalle apropiado al proyecto, para garantizar que **no se pierdan los datos**, sobre todo al finalizar el proyecto de investigación, facilitar **la correcta custodia** desde su producción, **permitir su preservación y validar y reutilizar** los resultados obtenidos. Como recomendaciones el PGD debe incluir la siguiente información:

- **Resumen de los datos**, donde se describa de forma breve y concisa: qué datos se han generado/recogido, qué formatos y estándares se han utilizado, qué valor tienen los datos para otros investigadores, qué datos no se pueden compartir y por qué motivo.
- **Responsabilidades**, indicando las personas responsables y especificando quién determina el acceso a los datos y, en el caso de que exista propiedad intelectual, quién es el titular de los derechos del plan de seguimiento, de la gestión de los datos y de su preservación y conservación.
- **Datos FAIR** (de Findable, Accessible, Interoperable y Reusable)*. Donde se describan las actuaciones realizadas a cabo para que los datos sean localizables, accesibles, interoperables y reutilizables
    - **Datos localizables**, se debe indicar que metadatos se han utilizado, especificando el identificador utilizado para los datos y si se ha utilizado algún control de versión.
    - **Accesibilidad**, se debe indicar cuándo y quién puede acceder a los datos, especificando su disponibilidad y/o los métodos o el software necesarios para acceder a los datos. Indicar si se incluye el software necesario y su documentación, especificando dónde se encuentran depositados los

datos, los metadatos, la documentación y el código, y como se gestiona el acceso en los casos en los que existen restricciones.

- **Interoperabilidad**, describiendo como se facilita la interoperabilidad de los datos mediante el uso de vocabularios controlados.
- **Reutilización**, se debe especificar el grado de reutilización permitido para los datos puestos al alcance de otros investigadores y del público en general. y a partir de qué momento los datos son accesibles para ser reutilizados. Si procede, se debe explicar los motivos y el periodo durante el cual los datos están embargados. Especificar si los datos generados y/o recogidos en el proyecto son reutilizables por terceras partes. Si la reutilización de los datos es restringida, deben explicarse los motivos.

- **Seguridad de los datos****, indicando el tratamiento de la recuperación de los datos, así como el almacenamiento seguro y el traspaso de datos sensibles y las medidas adoptadas para la seguridad de los datos. Esto incluye el almacenamiento y las copias de seguridad, la preservación a largo plazo y los procedimientos para mantener los datos seguros y protegidos, la cobertura del contexto de la revisión ética (describir, si fuera de aplicación, los problemas éticos en la recogida, almacenamiento, procesamiento y archivo de los datos, así como en los procedimientos de aprobación ética relacionados con el proyecto).

* FAIR son las siglas en inglés de Findable, Accessible, Interoperable and Reusable, y responden a la aplicación de estos cuatro principios que harán que los datos de investigación sean localizables, accesibles, interoperables y que se puedan reutilizar. Desde un punto de vista práctico, la implementación de los principios FAIR implica la aplicación de estándares de metadatos y vocabularios asociados al dominio científico de los datos, la utilización de identificadores persistentes, la correcta identificación de licencias y procedencia de los datos, así como el establecimiento de mecanismos de almacenamiento y preservación.

** Si las actividades de investigación incluyen a niños, pacientes, población vulnerable, el uso de células madre embrionarias, cuestiones de privacidad y protección de datos o investigación en animales y primates, deben cumplirse los principios éticos y la legislación nacional pertinente y la de la Unión Europea y debe tratarse en este apartado del plan de gestión de datos