



Horizon Europe

Data Management Plan Template

Version 1.0
17 Junio 2026

HISTORY OF CHANGES		
Version	Publication date	Changes
1.0	17.06.2026	▪ Initial version
		▪

Action Number: COOP24CIII/00018

Action Acronym: [SENFRA (subproyecto SENFRA-4]

Action title: [Sur8/Shoc2 como nuevo actor en envejecimiento y fragilidad]

Date: [17/06/2026]

DMP version: [1.0]

1. Resumen de datos

El subproyecto SENFRA-4 se centra en el análisis bioinformático e integración de datos ómicos generados por los distintos grupos participantes en el proyecto cooperativo SENFRA. El objetivo principal es identificar mecanismos moleculares asociados a la ausencia de Sur8/Shoc2 y su posible implicación en los procesos de envejecimiento y fragilidad.

El proyecto reutiliza conjuntos de datos generados dentro de la acción colaborativa, incluyendo datos transcriptómicos (RNA-seq de mRNA), datos de expresión de microRNAs (small RNA-seq), datos proteómicos cuantitativos basados en tecnología TMT, datos de fosfoproteómica y datos de oxidación de proteínas obtenidos a partir de organoides intestinales, criptas intestinales y otros modelos experimentales.

Los datos se utilizan para realizar controles de calidad, análisis de expresión diferencial, análisis de enriquecimiento funcional, integración miRNA-mRNA, integración transcriptómica-proteómica y análisis multi-ómicos.

Los formatos manejados incluyen FASTQ, TXT, CSV, XLSX, PDF y scripts desarrollados en lenguaje R. Como resultado de los análisis también se generan matrices procesadas de expresión, tablas de resultados estadísticos, figuras científicas e informes técnicos.

Los datos son generados por los grupos experimentales participantes en SENFRA y transferidos al equipo de SENFRA-4 mediante los mecanismos seguros de intercambio de información disponibles en el Instituto de Salud Carlos III.

El volumen esperado de información se sitúa en el rango de varios cientos de gigabytes, debido principalmente a los datos de secuenciación masiva y proteómica.

Los resultados obtenidos podrán ser de utilidad para investigadores de los ámbitos del envejecimiento, la fragilidad, la transcriptómica, la proteómica y la biología de sistemas.

2. Datos FAIR

2.1. Hacer que los datos sean localizables, incluyendo los metadatos

Los datos se organizan siguiendo estructuras jerárquicas de directorios y nomenclaturas estandarizadas que permiten identificar el modelo experimental, la tecnología utilizada y la fase del análisis correspondiente.

Los metadatos asociados incluyen información sobre el tipo de muestra, modelo experimental, plataforma ómica utilizada, flujo de análisis aplicado, versiones de software empleadas y fecha de procesamiento.

En caso de que los conjuntos de datos sean depositados en repositorios públicos, se utilizarán los identificadores persistentes asignados por dichos repositorios (DOI, accession numbers u otros identificadores equivalentes).

Los metadatos incluirán palabras clave relacionadas con envejecimiento, fragilidad, Sur8, Shoc2, transcriptómica, microRNAs y proteómica para facilitar la localización y reutilización de la información.

Cuando sea posible, los metadatos se proporcionarán en formatos que permitan su indexación y recuperación por parte de repositorios y motores de búsqueda científicos.

2.2. Hacer que los datos sean accesibles

Actualmente los conjuntos de datos permanecen almacenados en la infraestructura informática del Instituto de Salud Carlos III y son accesibles únicamente para los miembros autorizados del proyecto.

Cuando sea necesario para la publicación de resultados científicos, los datos y metadatos asociados podrán depositarse en repositorios científicos reconocidos y apropiados para cada tipo de información.

Los datos que respalden publicaciones científicas se pondrán a disposición de la comunidad científica tan pronto como sea posible, teniendo en cuenta los requisitos editoriales, las necesidades de protección de resultados y los acuerdos establecidos entre los grupos participantes.

Los metadatos descriptivos podrán permanecer accesibles incluso en aquellos casos en los que los datos subyacentes estén sujetos a restricciones de acceso.

Durante la ejecución del proyecto, el acceso a los datos se realiza mediante los mecanismos de autenticación institucionales establecidos por el Instituto de Salud Carlos III.

No se prevé la necesidad de establecer un comité específico de acceso a datos, dado que el subproyecto no maneja datos personales ni información clínica identificable.

2.3. Hacer que los datos sean interoperables

Los datos se gestionan utilizando formatos ampliamente aceptados por la comunidad científica, incluyendo FASTQ, TXT, CSV y XLSX.

Las anotaciones de genes, transcritos, microRNAs y proteínas utilizan nomenclaturas y bases de datos de referencia reconocidas internacionalmente.

Los análisis bioinformáticos se desarrollan principalmente mediante lenguaje R utilizando herramientas ampliamente empleadas en transcriptómica, proteómica y análisis estadístico.

Cuando procede, se mantienen referencias a bases de datos biológicas y recursos de anotación externos que facilitan la interoperabilidad, comparación e integración de resultados con otros estudios.

2.4. Incrementar la reutilización de los datos

Todos los análisis realizados se documentan mediante scripts, parámetros de ejecución, informes de análisis y documentación metodológica.

La documentación generada incluye información sobre los procedimientos de control de calidad, métodos estadísticos empleados, versiones de software utilizadas y descripción de los flujos de trabajo bioinformáticos.

En aquellos casos en los que los datos se compartan públicamente, se utilizarán las condiciones de reutilización y licencias establecidas por los repositorios correspondientes y por las políticas institucionales aplicables.

La procedencia de los datos se documenta a lo largo de todo el flujo analítico, desde la recepción de los datos originales hasta la generación de los resultados finales.

Los procedimientos de aseguramiento de calidad incluyen verificaciones de integridad de ficheros, controles de calidad específicos para cada tecnología ómica, revisión de metadatos y validación estadística de los resultados obtenidos.

3. Otros resultados de investigación

Además de los datos de investigación, el proyecto genera otros resultados científicos digitales que forman parte esencial de los análisis realizados y de su reproducibilidad.

Entre ellos se incluyen:

- Scripts desarrollados en lenguaje R para análisis transcriptómicos, de microRNAs y proteómicos.
- Flujos de trabajo bioinformáticos reproducibles.
- Procedimientos de integración multi-ómica.
- Informes técnicos.
- Figuras científicas y visualizaciones.
- Documentación metodológica y de análisis.

Estos materiales se almacenan junto con los datos asociados en la infraestructura informática del Instituto de Salud Carlos III y constituyen un elemento fundamental para garantizar la reproducibilidad y reutilización de los resultados obtenidos.

4. Asignación de recursos

Las actividades de gestión de datos se realizan utilizando la infraestructura informática ya disponible en el Instituto de Salud Carlos III.

No ha sido necesaria la contratación de infraestructuras externas específicas para almacenamiento, procesamiento o conservación de los datos.

La responsabilidad de la gestión de los datos recae en el equipo investigador del subproyecto SENFRA-4 y en su investigador principal.

La preservación a largo plazo se garantizará mediante los sistemas institucionales de almacenamiento y respaldo del Instituto de Salud Carlos III y, cuando corresponda, mediante el depósito de conjuntos de datos y resultados asociados en repositorios científicos reconocidos.

5. Seguridad de los datos

Todos los análisis se realizan en equipos informáticos y servidores institucionales gestionados por el Instituto de Salud Carlos III.

La transferencia de datos se lleva a cabo mediante mecanismos seguros de intercambio de información proporcionados por la institución.

El acceso a los datos está restringido exclusivamente a los investigadores autorizados que participan en el proyecto.

Los datos, scripts y resultados analíticos se almacenan en sistemas sujetos a las políticas institucionales de seguridad, copia de respaldo y recuperación establecidas por el Instituto de Salud Carlos III.

El subproyecto no maneja datos personales ni información identificativa de participantes humanos.

6. Aspectos éticos

El subproyecto SENFRA-4 no genera ni procesa datos personales.

Las actividades desarrolladas se limitan al análisis bioinformático de datos ómicos generados por los grupos experimentales participantes en el proyecto cooperativo.

No se recogen ni gestionan muestras humanas dentro del subproyecto.

Todas las actividades se realizan de acuerdo con la normativa institucional del Instituto de Salud Carlos III aplicable a la investigación y a la gestión de información científica.

7. Otros aspectos

El proyecto sigue los procedimientos institucionales establecidos por el Instituto de Salud Carlos III para la gestión, almacenamiento, transferencia y conservación de datos científicos.

Asimismo, se aplican las políticas de seguridad de la información vigentes en la institución y las directrices establecidas por el proyecto cooperativo SENFRA para el intercambio de información entre los grupos participantes.

Actualmente no se utilizan procedimientos adicionales de gestión de datos distintos de los establecidos el Instituto de Salud Carlos III y el propio proyecto SENFRA.