
NOTICIAS / NEWS

¿Cómo deben gestionar los científicos sus datos de investigación? Aportaciones desde Maredata



Maredata (2018). *Recomendaciones para la gestión de datos de investigación dirigidas a investigadores*. Autora: Remedios Melero. [Barcelona]: Maredata. (<https://digital.csic.es/handle/10261/173801>)

1. INTRODUCCIÓN

El interés por los datos de investigación se ha incrementado en los últimos años. Existe ya el consenso general sobre la necesidad que los datos científicos sean no tan sólo recogidos, sino también etiquetados con metadatos y almacenados para que pueda ser posible su recuperación y eventual reutilización por parte de toda la comunidad científica.

Esta constatación fue uno de los elementos que impulsaron la creación de Maredata, una red temática financiada por el Ministerio de Economía y Competitividad (CSO2015-71867-REDT) durante el período 2015-2018. Los objetivos propuestos por Maredata fueron aglutinar y consolidar la co-

laboración entre grupos de investigación españoles centrados en el estudio de los datos científicos y también establecer relaciones con aquellos sectores y disciplinas interesados en esta temática. Los miembros de la red procedían de siete instituciones distintas: Instituto de Alimentación y Tecnología de Alimentos (CSIC), INGENIO (CSIC-Universitat Politècnica de València), Universidad de Alicante, Universitat de Barcelona, Universidad Carlos III de Madrid, Universitat Oberta de Catalunya, Universitat de València y Universitat Politècnica de Valencia, que participaban en cinco proyectos de investigación financiados por el Plan Estatal.

Durante los tres años de funcionamiento de la red se han llevado a cabo diversas actividades

de difusión (entre ellas, la organización de cinco seminarios y la creación del portal Maredata), de internacionalización (contactos con Research Data Alliance), de relación con los agentes implicados en la gestión de datos de investigación (investigadores, bibliotecarios), de preparación de nuevas líneas de investigación y, finalmente, se han elaborado unas recomendaciones para la gestión de datos de investigación dirigidas a los investigadores (Maredata, 2018).

En este breve texto vamos a presentar este documento de recomendaciones poniéndolo en el contexto de directrices internacionales similares y comentando su estructura y alcance. Estas recomendaciones se basan en la experiencia de los miembros de Maredata y en la documentación existente relativa a la gestión de datos.

2. ANTECEDENTES

Nadie duda que el principal agente impulsor de la preocupación por los datos de investigación ha sido la Comisión Europea. La primera gran actuación destacable fue la instauración en 2014 de un plan piloto para que los datos científicos de proyectos de investigación financiados por el programa Horizon2020 se difundieran en abierto en siete áreas del conocimiento (ampliándose en 2017 a todos los proyectos). Esto supuso un cambio sustancial en la atención de la comunidad científica hacia esta cuestión y dejó una huella ya imborrable.

En 2015, Carlos Moedas, el comisario europeo de Investigación, Ciencia e Innovación situó el punto de mira un poco más alto, en la denominada "ciencia abierta", la nueva estrategia de la Comisión Europea para impulsar un cambio de modelo en el funcionamiento de la investigación (Moedas, 2015). Se establecieron ocho prioridades o retos para el desarrollo de la ciencia abierta en Europa que incluían, de manera destacada, los datos de investigación junto con otros componentes como: el reconocimiento y los incentivos; los indicadores de investigación y nuevas métricas; el futuro de la comunicación académica; la European Open Science Cloud (EOSC); la integridad de la investigación; habilidades y educación en ciencia abierta, y la ciencia ciudadana.

En 2016 se creó la European Open Science Policy Platform (OSPP), un grupo asesor de alto nivel integrado por expertos con el objetivo de asesorar en el desarrollo y la implementación de políticas de ciencia abierta en Europa. Este grupo elaboró un documento con recomendaciones (OSPP, 2018) que responde a las ocho prioridades anteriores, y también incluyen, por tanto, los datos de investigación.

La European Open Science Cloud (EOSC), un entorno virtual común dado a conocer en 2016, es una infraestructura para que todos los investigadores europeos puedan almacenar, gestionar, analizar y reutilizar datos para la investigación, la innovación y la educación. Se trata de otro elemento fundamental para el desarrollo de la ciencia abierta y los datos de investigación. La idea que subyace en la EOSC es crear un entorno abierto de confianza en el que la comunidad científica pueda compartir y reutilizar datos y resultados de investigación (Comisión Europea, 2016b).

Esta acción impulsora de la Comisión Europea ha tenido buen eco en universidades y centros de investigación, que han generado un buen número de hojas de ruta y directrices para orientar en los procesos de gestión de datos de investigación y, de manera más amplia, a la ciencia abierta. Entre ellas, podemos destacar posicionamientos y orientaciones elaborados por organizaciones universitarias como la League of European Research Universities (LERU, 2013), la Young European Research Universities Network (YERUN, 2017) o la Ligue des Bibliothèques Européennes de Recherche (LIBER, 2018), o por proyectos europeos como OpenAIRE, FOSTER, OpenUp, EDISON, LEARN o RECODE (2014), entre otros, que también han elaborado documentos de este estilo. Todos ellos se han tenido en cuenta para preparar nuestras recomendaciones.

3. RECOMENDACIONES PARA LOS INVESTIGADORES

En este escenario, es necesario fortalecer y aunar las iniciativas que eviten la redundancia y potencien trabajos coordinados entre grupos de investigación y otros sectores implicados en la gestión y la apertura de los datos de investigación. Por ello, el documento con las recomendaciones quiere ayudar a conseguir que se cumplan los cuatro criterios definidos como «apertura inteligente» (*intelligent openness*) de los datos: que se puedan localizar, utilizar, evaluar y entender (Royal Society, 2012). Es decir, utilizando la sigla de uso universal, que sean FAIR: localizables (*findable*), accesibles (*accessible*), interoperables (*interoperable*) y reutilizables (*reusable*).

El documento presenta un total de 17 acciones que ayudan a optimizar la gestión de datos en el sector de la investigación y se puede consultar en castellano, inglés y catalán-valenciano. Las recomendaciones, además de un enunciado general, incluyen ejemplos e información práctica y complementaria para que los investigadores que necesitan más detalles puedan resolver sus dudas.

Una de las consideraciones que se han tenido en cuenta es que una buena gestión de datos debe incluir criterios, principios y estándares que hagan inteligibles los datos más allá del dominio o disciplina donde se producen.

Para facilitar una visión global del documento, hemos agrupado las 17 recomendaciones en cinco grandes ámbitos:

a) Qué se debe hacer con los datos

1. Adoptar los principios FAIR para que los datos sean Localizables, Accesibles, Interoperables y Reutilizables.
2. Seleccionar los datos de interés para usos futuros.
3. Aplicar los principios FAIR a los metadatos y al plan de gestión de datos.
4. Describir los beneficios y ventajas de haber hecho que los datos de investigación sean FAIR en el informe de evaluación del proyecto.

b) Plan de Gestión de Datos

5. Desarrollar un plan de gestión de datos (PGD) que muestre cómo van a ser tus datos FAIR.
6. Indicar en el PGD si se han consultado o reutilizado datos ya existentes en lugar de obtenerlos de nuevo.

c) Estándares y licencias

7. Utilizar estándares interoperables.
8. Facilitar los datos de investigación de acuerdo con estándares reconocidos por la comunidad científica.
9. Asignar licencias abiertas para promover la reutilización.

d) Difusión, citación y control

10. Facilitar una descripción detallada de las condiciones para que se puedan reproducir los ensayos experimentales.
11. Identificar el repositorio de confianza más adecuado.
12. Verificar los términos o requisitos de las políticas respecto a los datos de investigación.
13. Promover la citación de los datasets como cualquier otro recurso bibliográfico.
14. Cuando se trata con seres vivos, tener en cuenta la legislación vigente sobre los datos.

e) Promoción (Advocacy)

15. Contribuir a la diseminación e intercambio de las buenas prácticas respecto a la gestión de datos entre comunidades científicas.
16. Favorecer la colaboración entre profesionales de la materia y gestores de datos.
17. Propiciar que los investigadores y sus centros avalen el paradigma de la ciencia abierta.

BIBLIOGRAFÍA

- Comisión Europea (2016a). *H2020 programme: guidelines on FAIR data management on Horizon 2020*. Brussels: European Commission, Directorate-General for Research & Innovation. 12 p. http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf
- Comisión Europea (2016b). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: European Cloud Initiative: building a competitive data and knowledge economy in Europe*. Brussels: European Commission. 13 p. <http://ec.europa.eu/transparency/regdoc/rep/1/2016/EN/1-2016-178-EN-F1-1.PDF>
- LERU (2013). *LERU roadmap for research data*. Advice paper n. 14. <https://www.leru.org/files/LERU-Roadmap-for-Research-Data-Full-paper.pdf>
- LIBER (2018). *LIBER Open Science Roadmap*. <http://doi.org/10.5281/zenodo.1303002>
- Maredata (2018). *Recomendaciones para la gestión de datos de investigación dirigidas a investigadores*. Barcelona: Maredata. <https://digital.csic.es/handle/10261/173801>
- Moedas, C. (2015). Open innovation, open science, open to the world. *A new start for Europe: opening up to an ERA of innovation Conference*. Brussels: European Commission. http://europa.eu/rapid/press-release_SPEECH-15-5243_en.htm
- Open Science Policy Platform (2018). *Integrated advice of the Open Science Policy Platform on 8 prioritised open science ambitions*. Brussels: European Commission. <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-policy-platform>
- RECODE Project Consortium (2014). *Policy recommendations for open access to research data*.

London: RECODE. 40 p. http://web.archive.org/web/20180205153933/http://recodeproject.eu/wp-content/uploads/2015/01/recode_guideline_en_web_version_full_FINAL.pdf

Royal Society. Science Policy Centre (2012). *Science as an open enterprise*. London: The Royal Society. 104 p. <https://royalsociety.org/topics-policy/projects/science-public-enterprise/Report/>

Science Europe (2018). *Science Europe guidance document: presenting a framework for discipline-specific research data management: [D/2018/13.324/1]*. Brussels: Science Europe. 46 p. http://www.snf.ch/SiteCollectionDocuments/SE_Guidance_Document_RDMPs.pdf

YERUN (2017). *YERUN statement on open science*. https://www.yerun.eu/wp-content/uploads/2018/05/YERUN_OpenScience_Statement-3.pdf