

Gestión y planificación de datos de investigación para afrontar los requerimientos de Horizonte Europa

Preguntas y respuestas: 28 y 29 de abril 2021

Depósito de datos

- ¿Las publicaciones, debidamente vinculadas, sirven como documentación suficiente para un conjunto de datos? ¿O deben prepararse documentaciones ad hoc?
 - o **En principio, todo proyecto de investigación se basa en datos, ya sean imágenes, gráficos, vídeos, tablas, etc. Estamos hablando de justificar los resultados de un proyecto financiado, por lo que se vinculará a varias publicaciones y estas tendrán que demostrar sus afirmaciones con datos asociados (indicando origen y posibles usos posteriores).**
- ¿Cómo se va abordar la propiedad de los datos en ámbitos como la Biomedicina y la práctica clínica?
 - o **Estos datos tienen tratamiento especial, por un lado debe asegurarse que están anonimizados, que se respetan todos los aspectos éticos para conseguirlos.**
Pueden subirse los metadatos a un repositorio, pero mantener los datos ocultos, si se puede justificar. O durante un periodo concreto.
Pero, además, han añadido una cláusula que, en época de urgencia sanitaria, se aplican unas condiciones distintas.
- Para vincular los datos y las publicaciones ¿hay que depositar los datos previamente? Para así obtener el DOI que se pueda incluir en la publicación ¿no? Porque editar los datos va a ser siempre más fácil que editar las publicaciones.
 - o **Las publicaciones se pueden depositar en un repositorio y añadir un metadato de "relación" para vincularlo con el dataset.**
En ambas tipologías, siempre se pueden subir sucesivas versiones, que demuestren un avance en la investigación.
Otra opción es subir sólo los metadatos, mientras se espera la aceptación del artículo y, una vez recibida, depositar los datos.
- ¿El formato ODS, equivalente a excel en Libre Office, es peor que el CSV?
 - o **El objetivo es que el fichero se pueda volver a consultar en el futuro. El equipo técnico del repositorio donde vaya a depositar sus datos, le indicará el formato más recomendado.**
En cualquier caso, entre los compromisos de conservación de cada plataforma, está el revisar los software utilizados por si hay que realizar alguna actualización y evitar que se conviertan en ficheros inaccesibles. Estas decisiones se toman siempre de manera acorde a nivel internacional.
- ¿Podéis mostrar un ejemplo de datos publicados para que se vea el contenido, formatos etc?
 - o **Claro, puedes consultar el repositorio del Consorcio Madroño:**
<https://edatos.consorcioadrono.es/>; Ejemplo de dataset:
<https://doi.org/10.21950/F9KFQT>

Fíjate que, además de un fichero, todo dataset tiene un documento "Readme.txt" donde se explican todos los detalles sobre los datos depositados.

Algunas colecciones solicitan tener un "libro de visitas" para conocer quiénes están descargando sus ficheros (para ver un ejemplo, hay que intentar descargarse el fichero).

- Si subimos a Zenodo un dataset, ¿tiene sentido depositarlo en otro repositorio? ¿Cómo evitar el riesgo de duplicar los identificadores persistentes?
 - o *Aquí lo ideal es elegir un solo repositorio que, si está bien configurado, será incluido en agregadores regionales, nacionales e internacionales, En tu pregunta, con Zenodo basta, no hay que duplicar el depósito*
- Hay alguna plataforma o página web que ayude a trabajar la interoperabilidad de datos, o que facilite el tema de la interoperabilidad?
 - o *El tema de la interoperabilidad está directamente relacionado con el uso de protocolos que la garanticen y eso es un mundo bastante amplio. No conozco un recurso único en este sentido, pero si me mandas un correo con la pregunta te buscaremos algo.*
En cualquier caso, si su duda se refiere a la capacidad de intercambio de datos entre plataformas o buscadores, si los datos se depositan en un repositorio "de confianza", la interoperabilidad está asegurada.
- ¿El tema de los registros software entonces como queda en todo esto? es decir, ¿esto va a ser negativo? ¿o cómo se gestionará para que se fomente la explotación de los registros software? ¿no es recomendable?
 - o *Según acaba de comentar Teresa Malo, el depósito del software es recomendable. Es mejor consultar la cita mencionada: Fuente: Victoria Tsoukala: Open Science in Horizon Europe: a proposer's primer. April, 21, 2021. Disponible en: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/other/events/20210421/open-science_en.pptx
*"Proporcionar acceso abierto a los resultados de la investigación (por ejemplo, publicaciones, datos, software, modelos, algoritmos y flujos de trabajo) a través del depósito en repositorios confiables". Los repositorios de datos ya están aceptando este tipo de ficheros sin ningún problema.**
- Pero los DOIs del paper y de los datos correspondientes son diferentes ¿verdad?
 - o *Sí, por supuesto, cada artículo tendrá su DOI y cada dataset el suyo, incluso en algunos repositorios se proporciona un DOI diferente a cada fichero de datos.*
- ¿La URJC está en Madroño?
 - o *Sí, estamos UAM, UAH, UC3M, UPM, UNED y URJC*

Embargo

- ¿Los doce meses de embargo es para todos los ámbitos científicos? Esto ha cambiado con respecto a H2020, ¿no? El embargo era de 6 meses máximo (salvo para ciencias sociales y humanidades).
 - o *Efectivamente, hay un máximo de 12 meses de embargo para ciertas áreas, como todo lo relacionado con politécnica.*

- ¿Se ha mantenido el periodo de embargo de 6m para ciencias no humanidades?
 - o **No me consta que se vaya a modificar, pero aún se están actualizando todos los materiales relacionados con Horizonte Europa.**
- No me ha quedado claro si Horizonte Europa permite embargos para el acceso abierto de las publicaciones, ¿podrías comentarlo?
 - o **El embargo máximo admitido es de 12 meses. Además: Si las condiciones de la convocatoria lo imponen en caso de una emergencia pública, los beneficiarios deben (si así lo solicita la autoridad otorgante) depositar inmediatamente cualquier resultado de la investigación en un repositorio y proporcionar acceso abierto a él bajo licencia CC BY, CC 0 o equivalente.**

EOSC

- EOSC ¿serviría también para hacer un DMP como algunas plataformas que ya existen para ello?
 - o **Por lo que yo tengo entendido, de momento no, pero hay iniciativas como ARGOS (proyecto colaborativo entre OpenAire y EUDAT) <https://argos.openaire.eu/home>**

Evaluación

- ¿Cómo se plantea insertar estas nuevas medidas en los sistemas de evaluación? Casi todas las revistas más valoradas y utilizadas para evaluación son híbridas y no cabrían en el Open Science.
 - o **Estamos en una época de transformación global. Se está analizando cómo evaluar todos estos cambios. Una iniciativa interesante es DORA <https://sfdora.org/>**
 - o **Link al webinar de DORA del 13 de mayo 2021 que se ha comentado durante el curso. https://us02web.zoom.us/webinar/register/WN_usEzdbiJQ0WOXTCxcAnucA**
 - o **También puede visitar esta entrada de blog: Evaluación de la investigación – nuevas ideas: <http://biblioteca2.uc3m.es/investigacion/2020/06/10/evaluacion-de-la-investigacion-nuevas-ideas/>**
- ¿Cuál es el posicionamiento del Ministerio y la ANECA al respecto de esta iniciativa OpenScience? Va a ser poco probable que publiquemos en ORE-EOSC si su impacto no es medible y/o admitido por el Ministerio/ANECA. Es una pena, pero aparte de generar impacto difundiendo ideas, nuestra carrera profesional (AKA estabilidad, sueldo, etc) depende de la evaluación de nuestra carrera científica que haga la ANECA y la ANEP
 - o **Hace poco ANECA publicó esta nota de prensa (fecha: 22/02/2021) <http://www.aneca.es/Sala-de-prensa/Noticias/2021/Postura-de-ANECA-en-la-Comision-Open-Science-COS-Gob>. Postura de ANECA en la Comisión Open Science COS-Gob. Destaco el punto 3. ACCIONES ESPECÍFICAS QUE ANECA PUEDE LLEVAR A CABO. En las convocatorias anuales de sexenios de transferencia, dar reconocimiento a las acciones en pro del open knowledge. Valorar en los programas de acreditación del profesorado el esfuerzo por el conocimiento en abierto. Valorar en los programas de acreditación de universidades el esfuerzo por dar a conocer los resultados de la investigación**

y de la transferencia de sus investigadoras/es. Implantar de un Sello internacional de Calidad ANECA en conocimiento en abierto.

Licencias de uso (Creative Commons)

- Si mi publicación es un software. Qué tipo de licencia se recomendaría para que fuese abierto para la comunicad científica pero que se utilizase por una empresa sacar un beneficio. Como se gestionaría para obtener un ingreso de esta empresa para que sirva para invertirlo en nuevas investigaciones?
 - o **Habría que utilizar una licencia Creative Commons de este tipo:**
 - **CC BY: permite reconocimiento + modificar + ganar dinero,**
 - **CC BY SA: permite reconocimiento + modificar + ganar dinero, pero se tiene que compartir con el mismo tipo de licencia,**
 - **CC BY ND SA: permite reconocimiento + ganar dinero, pero se tiene que compartir con el mismo tipo de licencia y no se puede modificar.**

Para marcar que no se puede sacar beneficio económico, hay que incluir el elemento NC (Non comercial)

Recomiendo visitar esta entrada de blog: Aprende a usar e interpretar las licencias Creative Commons:

<http://biblioteca2.uc3m.es/investigacion/2018/05/09/aprende-a-usar-e-interpretar-las-licencias-creative-commons/>

- Todas las licencias que habéis comentado en la primera parte del seminario (CCC y otras) ¿se cumplirían con elaborar el PGD en un repositorio fiable como el del consorcio madroño, zenodo? o ¿se requiere algún otro tipo de licencia/permiso/validación de esos datos?
 - o **Normalmente cada repositorio de datos tiene una política concreta de licencia de datos que otorga una abierta por defecto, del tipo CC By o CC BY NC, y alguna opción de elegir alguna licencia algo más restrictiva para casos especiales. Recuerda el principio de tan abierto como sea posible, tan cerrado como sea necesario.**
 - o **Recomiendo visitar esta entrada de blog: Aprende a usar e interpretar las licencias Creative Commons:**
<http://biblioteca2.uc3m.es/investigacion/2018/05/09/aprende-a-usar-e-interpretar-las-licencias-creative-commons/>

Metadatos

- ¿A qué se refiere con metadatos en publicaciones?,
 - o **Son los campos o epígrafes que describen las publicaciones, ej. Título, autor, editorial. Pero, al tratarse de documentos en formato digital, pasan a formar parte de la descripción interna del fichero, de modo que son intercambiados entre las distintas plataformas, como los agregadores y buscadores. Como ejemplo, al publicar en un repositorio institucional específico, como el de la UC3M, estos metadatos se recogerán de forma automática, sin necesidad de hacer nada más, en:**

- *el agregador regional de la Comunidad de Madrid: e-ciencia (<https://mcyt.educa.madrid.org/madrid-ciencia-tecnologia/e-ciencia/repositorio-emid>),*
 - *en el agregador nacional: Recolecta (<https://www.recolecta.fecyt.es/>),*
 - *y en el agregador de la UE: OpenAIRE (<https://explore.openaire.eu/>)*
- ¿Lo de los metadatos también tienen que establecerlos los investigadores?
 - *Los metadatos son las etiquetas que describen la publicación o los datos de investigación: Autor, Título, Editorial, Fecha, etc. Tienen el valor añadido de convertirse en parte integrante del fichero (por eso, en lugar de etiquetas, reciben el nombre de "meta-datos"). Es el personal técnico del repositorio el que asigna los metadatos a los distintos elementos. Pero, en ocasiones, los toma de una plantilla donde el investigador indica la información a aportar.*
- ¿Hay algún ejemplo de como adaptar schema.org (Google Datasets Search) a casos de investigación? ¿Cómo se adapta schema.org para la descripción de metadatos en ámbito EU?
 - *Schema.org es un modelo de etiquetado de datos bibliográficos para facilitar su recuperación. Está ampliamente usado por motores de búsquedas comerciales, como Google, y cada vez más por agregadores de repositorios, como DataCite, por lo que su implementación facilita el descubrimiento y accesibilidad a los contenidos de los repositorios. Nos confirman que el recolector a nivel nacional Recolecta (<https://www.recolecta.fecyt.es/>) lo está implementando para asegurar la interoperabilidad con los repositorios de datos. Aquí puede encontrar una "Lista de repositorios que, actualmente, incorporan schema.org": <https://zenodo.org/record/1202178> Hemos localizado la tipología para dataset de schema.org en: <https://schema.org/Dataset> y para proyecto de investigación en: <https://schema.org/ResearchProject>*

Open Research Europe (ORE)

- ¿Qué opináis de la Open Research Europe? ¿Cómo fomentar su uso entre los investigadores hasta que no cambie la manera de evaluar la ciencia (impacto de publicaciones)?
 - *Acaba de arrancar y es un proyecto muy interesante. Seguro que resuelve y agiliza el flujo de trabajo. Pero, es pronto para dar una opinión. Un punto importante es ver cómo mostrarán las métricas de uso y cómo se evaluará el impacto.*
 - *Como ya hemos dicho, nos parece una propuesta muy interesante, porque además garantiza la publicación y el depósito en una misma operación. Su éxito, como todo, dependerá de su uso.*
- ¿La Open Research Europe tiene una forma para subir los metadatos o los datos también?
 - *Sí, toda plataforma requiere usar una descripción "bibliográfica" o metadatos del contenido que aloja.*

- ¿Cómo se van a seleccionar los revisores en ORE? ¿los elegirán los autores o habrá un pool?
 - *En principio, se supone que el peer review está dirigido por el autor, pero es mejor que consulte la web de ORE: <https://open-research-europe.ec.europa.eu/about/policies#publicationcriteria>*
- Si no me equivoco ORE no soporta la publicación de los datos de investigación. ¿Cómo podemos solventar esto si el artículo se puede publicar en ORE y los datos no?
 - *Se pueden depositar por separado publicación y datos, sin ningún problema. La plataforma donde se deposite cada tipo de resultado de investigación proporcionará un identificador único: DOI/handle, que puede ser incluido como metadato de elemento vinculado/relacionado.*
- ¿Sabéis si en algún momento se van a publicar datos en ORE?
 - *No tengo conocimiento de esto. Lo único que encuentro en su web es que sí solicitan que, algunos papers, tengan los datos depositados en alguno de los repositorios que citan: <https://open-research-europe.ec.europa.eu/for-authors/data-guidelines/>*
 - *Parece que ORE tiene muy bien resuelto todo el proceso de publicación, pero tiene que buscar "socios" para el depósito de las publicaciones (ahora sólo con Zenodo) y de los datos (aceptan muchas posibilidades). Entiendo que esto evolucionará, pero es una opinión personal. Gracias.*
- ¿En la plataforma europea se publicarán únicamente resultados provenientes de proyectos financiados por HE o por cualquier otro proyecto también?
 - *Según se indica en ORE: "Open Research Europe is an open access publishing platform for the publication of research stemming from Horizon 2020 funding. All researchers funded by a Horizon 2020 grant can publish original research related to their project at no cost to them." <https://open-research-europe.ec.europa.eu/for-authors/article-processing-charges/>*
 - *Entiendo que, efectivamente, no aceptarán otro tipo de publicaciones financiadas por otros organismos. Pero, habrá que preguntarles directamente a ellos. Esta decisión está relacionada con la financiación que F1000 (quien está detrás de ORE) recibe directamente de Europa.*
- No entiendo muy bien la diferencia y la conexión entre la ORE y la EOSC.
 - *ORE: plataforma que sustituye a una editorial de revistas. Facilita la publicación de papers financiados por Europa, sin coste de APC. Además, incluye el depósito inmediato en el repositorio Zenodo (en el futuro en otros repositorios)*
 - *EOSC: proporciona actualizaciones sobre su gobernanza y actores, los proyectos que contribuyen a su realización, oportunidades de financiación para las partes interesadas de EOSC, políticas europeas y nacionales relevantes, documentos importantes y desarrollos recientes.*
- Por otra parte, ORE al tener un sistema peer review propio, ¿qué impacto puede tener?. Me explico hay convocatorias de organismos públicos (MICINN p.e. con su convocatoria Severo Ochoa) que para evaluar la productividad científica de un investigador se basan en factores de impacto en revistas indexadas, lo que choca con la Open Science. En mi opinión es difícil que iniciativas como ORE prosperen siempre que por las mismas agencias financiadoras que te piden publicar en OA, te pidan a la vez para según que convocatorias, publicar en revistas "tradicionales"/ ¿Los artículos en ORE van a estar indexados? Por ahora para la carrera científica siguen siendo muy importantes las

métricas como el factor de impacto y el índice h que proporcionan las publicaciones en las revistas "tradicionales"

- **Efectivamente, esto es lo que inquieta de esta plataforma. Estamos pendientes de cómo lo van a resolver, pero aún es pronto para conocerlo, pues están en la fase inicial de recabar papers para publicar. La evaluación del CV del investigador también está en vía de revisión, en paralelo a la evolución de todo lo relacionado con Open Science. Le recomiendo que visite esta entrada de blog: Evaluación de la investigación – nuevas ideas: <http://biblioteca2.uc3m.es/investigacion/2020/06/10/evaluacion-de-la-investigacion-nuevas-ideas/>
Quizá le interese este webinar: DORA: <https://sfdora.org/> recognizes the need to improve the ways in which the outputs of scholarly research are evaluated. webinar on recognizing and rewarding data 13 may. 2021 05:00 p. m.**

Plan de Gestión de datos (PGD)

- O sea que recomendáis usar DMP online para preparar el plan?
 - **Es útil, porque en realidad funciona como un un tutorial y así lo completas correctamente**
- ¿Está también en inglés? Para poder usarlo con socios de proyectos de otros países.
 - **La herramienta está solo en español, en cuanto a ayudas, campos, etc., pero el plan debe redactarse en inglés. Si quieres usar una herramienta inglesa, en realidad PGDonline es la traducción del DMPonline del Digital Curation Center de UK**
- ¿Es fácil enlazar el PGD que hagamos en el PGD del consorcio madroño al paper/los paper relacionados y publicados en nuestro repositorio institucional?
 - **Lo que obtienes es un pdf que puedes usar como quieras.**
- ¿Podéis dar un esqueleto de elementos a incluir en el borrador de plan de gestión de datos para una RIA / EIC Pathfinder de H Europe? Teniendo en cuenta que se tiene 1/2 ó 1 página ...
 - **El Digital Curation Center tiene una check list que puede resultar útil: https://www.dcc.ac.uk/sites/default/files/documents/resource/DMP/DMP_Checklist_2013.pdf**
- Dada la limitada extensión del borrador del DMP en las propuestas de H Europe. ¿Cuáles diríais que son los elementos más importantes que no pueden faltar? ¿Sería uno de ellos definir claramente los datos que se van a compartir?
 - **SI, es muy importante describir los datos, pero también qué tratamiento se va a dar a los datos en el ciclo de vida del proyecto.**
 - **Horizon 2020 ofrece una plantilla: Horizon 2020 FAIR Data Management Plan (DMP) template (https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm) estaremos pendiente de la plantilla de Horizonte Europa, por si hay cambios.**

- Según mi experiencia el DMP en proyectos H2020 debe confeccionarse al inicio del proyecto como un esbozo, y es un entregable obligatorio a los 6 meses y con cada uno de los informes científicos del proyecto, que suelen ser a la mitad y al final del mismo. Además cada vez que haya un cambio significativo se tiene que volver a enviar. ¿Sabéis si esto se mantiene en HE?
 - ***Si, el tema de la gestión de datos es muy similar entre Horizonte 2020 y Horizonte Europa. De hecho, en Horizonte 2020 se inició como piloto en algunas áreas temáticas y se generalizó a partir del 2017. Horizonte Europa mantiene el mismo mandato. Es posible, sin embargo, que en algunas convocatorias concretas introduzcan alguna variante significativa***

Publicación

- En relación al open access, ¿no habría que gestionar un poco más que el acceso a las publicaciones en abierto sea más económico? es decir, publicar en abierto normalmente tiene un coste y suele ser un coste que no muchos grupos pueden asumir. Está claro que si tienes un europeo tienes financiación pero previamente a ello, debes tener publicaciones y es complicado.
 - ***Tiene toda la razón. Afortunadamente, este tema está cambiando. Estamos en un momento clave y está en manos del propio investigador proponer al editor un addendum al firmar la publicación del paper. Hay modelos como OPEN ACCESS PUBLISHING AGREEMENT: https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/oa-pilot/h2020-aa-guide-model-for-publishing-a_en.pdf Y Copyright toolbox: <https://sites.google.com/site/copyrighttoolbox/> Scholar's copyright addendum engine: <http://scholars.sciencecommons.org/>***
 - ***Además, cada vez se están firmando más "acuerdos transformativos", pregunte en su institución por si existe algún acuerdo que le evite pagar las APC.***
 - ***Copio esta noticia reciente: Crue y el CSIC firman con Elsevier un acuerdo que permitirá la publicación de artículos de los investigadores españoles en acceso abierto: <https://www.crue.org/2021/03/convenio-crue-csic-elsevier/>***

Repositorios de confianza

- El repositorio de la Consejería de Sanidad también estaría incluido en los repositorios validos?
 - ***Pues no estoy segura, pero supongo que aún no será un "repositorio confiable" porque esto es algo muy nuevo, y pocos repositorios lo son por ahora, pero si permite cumplir con el mandato de depósito y publicación en abierto***
- Para buscar datos, ¿hay que buscarlos en cada uno de los repositorios individualmente? ¿no se puede hacer una búsqueda general en un agregador de repositorios y que en este nos indiquen un enlace al repositorio donde se han depositado dichos datos y del que podemos descargarlos?
 - ***Hay algunos buscadores de repositorios de datos y otros buscadores de datasets (conjuntos de datos). Si lo que buscas son datos, deberás usar los***

segundos. . Por ejemplo. [Mendeley Data](#) de Elsevier, que busca datos en varios repositorios.

Zenodo

- ¿Es Zenodo válido como repositorio único (datos, publicaciones, software de código abierto y otros outputs) en proyectos H Europe?
 - o ***Sí, es un repositorio multidisciplinar, que acepta varias tipologías de resultados de investigación. Luego hablaremos de él.***
- ¿Utilizando un repositorio válido (como Zenodo) se cumple automáticamente con el FAIR data?
 - o ***Sí, Zenodo es reconocido como repositorio "de confianza".***
- Para una institución pequeña que no tiene repositorio institucional, ¿recomendáis usar Zenodo como "repositorio institucional"?
 - o ***Sí, Zenodo es considerado "repositorio de confianza".***
- En Zenodo creo que se puede vincular con el propio DOI del paper, el DOI de los datos correspondientes ¿tenéis más información al respecto?
 - o ***SI, en general en todos los repositorios de datos se puede incluir el identificador persistente (DOI o Handle) del artículo y viceversa, desde el artículo, se puede incluir como metadato el Identificador persistente de los datos.***

¡Muchas gracias a todos por las preguntas y observaciones!