



JAVIER DE LA CUEVA

Open Science y FAIR data

Open Science
and FAIR data

Javier de la Cueva
javier.delacueva@javiardelacueva.es
Abogado

Citación: de la Cueva, Javier (2021). "Open Science y FAIR data". *Tábula*, n. 24,
pp. 21-43

Recibido: 15-1-2021. *Aceptado:* 26-2-2021
DOI: <https://doi.org/xxx>

Resumen analítico / Analytic summary

En el presente artículo se hace un repaso del concepto de ciencia abierta y sus componentes, explicando las causas de las que procede la necesidad de su implementación. Se detallan los conceptos desde la perspectiva institucional tanto de la Comisión Europea como de la UNESCO y se describe el desarrollo de la European Open Science Cloud (EOSC). Asimismo, se introducen los principios FAIR (encontrables, accesibles, interoperables y reutilizables) aplicados a los datos científicos, poniéndose de manifiesto las posibilidades de interacción mutua entre la ciencia abierta y la archivística y el enriquecimiento que pudiera surgir de la colaboración entre ambos campos.

CIENCIA ABIERTA | FAIR DATA | METADATOS | EOSC | DATOS CIENTIFICOS

This article reviews the open science concept and details its components, explaining the reasons that makes its implementation necessary. The concept is analysed from the institutional perspectives of both the European Commission and UNESCO, describing the development of the European Open Science Cloud (EOSC). It also introduces the FAIR guiding principles for scientific data management, highlighting the possibilities for mutual interaction between open science and archival science and the enrichment that could arise within the collaboration of both fields.

OPEN SCIENCE | FAIR DATA | METADATA | EOSC | DATA SCIENCE

Una de las discusiones que se está produciendo en la actualidad en el ámbito científico es el de la ciencia abierta o, en la terminología en inglés que se ha impuesto, el de la Open Science. Como se explicará a continuación, se trata de un movimiento propugnado desde organizaciones supranacionales como la UNESCO y la Unión Europea que, en síntesis, busca que la ciencia se comparta, propugnando para ello su mayor apertura a través de la eliminación de los obstáculos que la impiden.

Ahora bien, sea cual fuere el grado de apertura o de publicidad de la actividad científica, de lo que no cabe duda es que la ciencia, para ser tal, ha de tener un soporte documental y ser objeto de unas técnicas archivísticas. La ciencia no se puede entender sin su documentabilidad; ciencia y archivos se hallan íntimamente ligados. Esta relación es recíproca. Tal y como señala Cruz Mundet, “la gestión científica tiene su equivalente en la gestión archivística: el control sistemático de los documentos y de la información” (2011, p. 21). Así pues, la ciencia sirve a la archivística y la archivística sirve a la ciencia,¹ aun cuando tradicionalmente la acepción de archivo se hubiera reservado a aquellos que se formaban en el ejercicio de una actividad administrativa, y sin perjuicio de la nunca acabada discusión sobre si la archivística es una técnica o una ciencia (Heredia Herrera, 1991, pp. 45-46; Cruz Mundet, 1994, pp. 58-64).

La aparición del movimiento de la Open Science es interesante para el mundo archivístico puesto que se están planteando en el seno de la documentabilidad de la ciencia problemas que ya han sido tratados por las personas estudiosas de los archivos. Por ejemplo, actualmente sería difícil, quizás imposible, encontrar una parcela científica donde no se utilicen las Tecnologías de la Información y Comunicación (TIC) como una de sus herramientas. En la aplicación de estas tecnologías, una de las acciones que se realizan en los proyectos de

investigación es la de representar el conocimiento científico mediante taxonomías y ontologías que, a su vez, sirven como las plantillas con las que etiquetar los datos de investigación y vincularlos entre sí. Tal labor de clasificación no es sustancialmente distinta de la que lleva haciendo la archivística desde los archivos de Ugarit (siglos XII a XIII a. C.). Asimismo, el cambio de paradigma de entender la archivística más como una labor dirigida a procesos que a la salvaguardia de objetos (Cook, 2001, p. 4) se alinea con la naturaleza dinámica de la ciencia. De esta manera, la reflexión sobre lo dinámico del archivo es aplicable a las necesidades documentales de los proyectos de investigación. Esta mutabilidad es aplicable especialmente a uno de los componentes de la ciencia abierta, los denominados principios FAIR, acrónimo de *findable*, *accessible*, *interoperable* y *reusable* (encontrable, accesible, interoperable y reutilizable), acrónimo que se aplica a los datos de investigación.

El contenido de esta ponencia tiene por objeto explicar los orígenes e implantación de la Open Science, de la que se citarán las iniciativas europeas y de la UNESCO, enumerando brevemente sus componentes, y describir en qué consisten los datos FAIR, para de esta manera evidenciar las posibles conexiones entre la archivística y esta novedosa forma de hacer ciencia y cómo pueden ambos campos ayudarse mutuamente.

Open Science

Abordar el concepto de Open Science no es una tarea sencilla y lo que se pueda entender por tales términos depende de la fuente elegida. Tomando como referencia a Antonio Lafuente y a Andoni Alonso, la respuesta es directa, pues la ciencia abierta consistiría simplemente en “mostrar el trabajo que uno hace en ciencia” (Lafuente y Alonso, 2011, p. 40). Estos autores lo vinculan con el *derecho a saber* que el público tiene, derecho cuya legitimidad trae causa del hecho de que las actividades de la comunidad científica le afectan directamente, como se demostró, por ejemplo, en el caso de las vacas locas, donde inicialmente se ocultó a la población el problema con la excusa de no producir alarma. Y afirman que la Open Science no es un sistema novedoso de hacer ciencia, sino que es el mismo sistema de siempre que está reaccionando a los procesos de cierre a los que se ha enfrentado en los últimos años. Para autoras como Eva Méndez, el término Open Science es un pleonismo, puesto que, si la ciencia no es abierta, entonces no es ciencia.²

Si bien la ciencia siempre se ha desenvuelto entre procesos de apertura y de cierre (Eamon, 1986), a principios de este siglo un grupo de personas de los ámbitos académico y bibliotecario señalaron que se estaba produciendo una deriva hacia lo cerrado que no era compatible con la misión de la diseminación del

conocimiento. La declaración de Budapest³ de 14 de febrero de 2002, sobre el acceso abierto a las publicaciones científicas, denunciaba que, si bien existe una herramienta, internet, con la que se puede difundir el conocimiento sin límites, sin embargo, el régimen de propiedad intelectual lo impide ya que los derechos de publicación de los textos científicos se hallan atrapados bajo la potestad de una industria editorial. En el mismo sentido se firmaron las declaraciones de Bethesda⁴ y de Berlín,⁵ respectivamente el 11 de abril y el 22 de octubre de 2003. Estas tres declaraciones abrían el camino en la reivindicación del Open Access o libre acceso a las publicaciones científicas.

La ciencia no se desenvuelve bien en el contexto denunciado por las tres declaraciones fundacionales del Open Access ya que se caracteriza como un conocimiento compartido sobre el que se llega a un consenso. Dificultar el intercambio de ideas y de las demostraciones experimentales no solo atenta contra las necesidades de la actividad científica, sino que limita la posibilidad de su refutación. A la cuestión de la dificultad de producir ciencia en entornos que dificultan la transmisión del conocimiento se une la ilegitimidad de la apropiación de la ciencia cuando los fondos de los que procede son públicos. No tiene justificación alguna que una actividad pagada con los impuestos de la ciudadanía acabe en régimen de oligopolio bajo la propiedad y el control de unos entes privados cuyo interés no es el conocimiento, sino los beneficios de su comercialización (Buranyi, 2017).

Hay muchas otras razones para la apertura de la ciencia, que no serán objeto de esta ponencia, pero sí que ha de señalarse su relación con la dignidad humana. El derecho a la ciencia es un derecho fundamental incluido en los tratados internacionales. En este sentido, el artículo 27 de la Declaración Universal de los Derechos Humanos⁶ estipula el derecho “a participar en el progreso científico y en los beneficios que de él resulten” y el artículo 15 del Pacto Internacional de Derechos Económicos, Sociales y Culturales⁷ dispone que “Los Estados Parte en el presente Pacto reconocen el derecho de toda persona a: [...] b) Gozar de los beneficios del progreso científico y de sus aplicaciones.” La ciencia, por tanto, es patrimonio de la humanidad y no de los grupos editoriales ni, tampoco, únicamente de la comunidad científica.

Open Science en Europa

En la actualidad, la Unión Europea se halla desarrollando la EOSC (European Open Science Cloud) que, a pesar de su nombre, no es una nube que sirve de almacén de los repositorios científicos, sino que se trata de un conjunto de redes de personas, instituciones, infraestructuras y prácticas dedicadas a hacer ciencia. Los inicios de la EOSC pueden remontarse a 2007, cuando la Comisión Europea adoptó la “Comunicación de la Comisión al Parlamento Europeo, al Consejo y al Comité Económico y Social Europeo sobre la información científica en la era

digital: acceso, difusión y preservación {SEC(2007)181}”.⁸ Del contenido de dicha Comunicación puede destacarse el reconocimiento de que “Toda investigación se apoya en los trabajos anteriores y depende de las posibilidades que tengan los científicos de consultar y compartir las publicaciones científicas y los datos de la investigación” (Comisión Europea, 2007). Esta visión suponía el reconocimiento de continuar por las sendas abiertas por las declaraciones sobre Open Access, de las que expresamente se citaba en la Comunicación a la Declaración de Berlín, presumiblemente por ser la única de las tres practicada en el territorio de la UE.

El proceso comenzado en 2007 se potencia en el año 2012 con dos documentos de trabajo de la Comisión, ambos de fecha 17 de julio; se trata de los SWD (2012) 211 final (Comisión Europea 2012a) y SWD (2012) 221 final (Comisión Europea 2012a). El primero de ellos reconoce que existen tres problemas de la documentabilidad de la ciencia que requieren una política para resolverlos: “un acceso subóptimo a las publicaciones de investigación científica”, “un acceso subóptimo a los datos de las investigaciones” y “la ingente marea de datos científicos, que hace necesaria la preservación a largo plazo de esta información” (Comisión Europea 2012a, pp. 4-5). El segundo de los documentos de trabajo plantea qué políticas públicas pueden tomarse para resolver los problemas descritos y, de entre las opciones posibles, se resuelve reforzar el Espacio Europeo de Investigación. De esta manera se traza un plan para seguir los progresos y evaluarlos para recomendar los cambios que fueren pertinentes (Comisión Europea 2012b, pp. 4-6).

Entre julio y septiembre de 2014, la Comisión Europea realizó una consulta pública⁹ para analizar el grado de conocimiento de las personas e instituciones implicadas en actividades científicas sobre el cambio en los sistemas de producir ciencia, sobre los retos que deberían afrontarse ante este cambio, acerca de las acciones necesarias para lograr una mayor competitividad y, finalmente, sobre la denominación de este nuevo modelo. En el resultado de la consulta fue elegida la terminología de Open Science, que es la utilizada a partir de entonces.

Posteriormente y hallándose la UE bajo la presidencia holandesa, se celebraron en Amsterdam el 4 y 5 de abril de 2016 las jornadas “Open Science – From vision to action”, que incluyeron como uno de sus postulados finales la siguiente descripción de la ciencia abierta:

La ciencia abierta trata de cómo los investigadores trabajan, colaboran, interactúan, comparten recursos y difunden resultados. Nuevas tecnologías y datos, una creciente demanda social para afrontar retos societarios de nuestros tiempos y la disposición de la ciudadanía de participar en actividades de investigación conducen a un cambio sistémico hacia la ciencia abierta.¹⁰

La conferencia propugnaba dos retos para el año 2020: por un lado, se planteaba el pleno acceso a todas las publicaciones científicas y, por otro, se postulaba un nuevo acercamiento al tratamiento de los datos de investigación.

Siguiendo esta línea, y no sin dificultades (Burgelman, J. C., 2021), en el año 2016 la Comisión Europea propone la creación de la European Open Science Cloud (EOSC) (Comisión Europea, 2016), cuya misión consistiría en federar las diferentes infraestructuras de investigación repartidas entre disciplinas y Estados miembros de la Unión. En el año 2018 se publica el documento de trabajo para desarrollar la EOSC (Comisión Europea, 2018), realizándose las primeras labores bajo el seno de la EOSC Executive Board desde 2018 hasta 2020 y pasando el testigo a partir del año 2021 a la Asociación EOSC, con sede en Bruselas, que se encarga en la actualidad de su coordinación y desarrollo a través de grupos de trabajo bajo la guía proporcionada por 20 informes previos de la EOSC Executive Board y por un documento, denominado Strategic Research and Innovation Agenda (SRIA) of the European Open Science Cloud (EOSC) (EOSC Executive Board, 2021), que es la guía central.

A pesar de su rol como guía, el documento SRIA no contiene una definición propia de ciencia abierta sino que reenvía a los contenidos de la Wikipedia¹¹ (p. 21) y del proyecto FOSTER¹², para el cual “La ciencia abierta trata acerca de cómo extender los principios de apertura a todo el ciclo de la investigación [...], promoviendo compartir y colaborar tan pronto como sea posible, generando un cambio sistémico en los modos de hacer ciencia e investigación”.^{13,14} La ciencia abierta supone pues un término paraguas que abarca un conjunto de prácticas sobre los diferentes apartados de la actividad científica.

La propuesta que se realiza desde las instituciones europeas se fundamenta en postulados tradicionales propuestos por Robert K. Merton en 1942 en su texto *The Normative Structure of Science* (Merton, 1974, pp. 270-278), según se reconoce explícitamente en el documento SRIA (EOSC Executive Board, 2021, p. 23). Dichos postulados son el universalismo, el comunismo, el desinterés y el escepticismo organizado, cuatro características de la ciencia que siguen vigentes y que suponen una guía para la actual producción y actividad científica. El universalismo mertoniano se refiere a que las aportaciones realizadas para el conocimiento científico no dependen de los atributos personales o sociales de quien las propone y, aun cuando puedan existir reivindicaciones nacionalistas, éstas no influyen en el resultado final. Por su parte, lo que Merton entiende por comunismo hace referencia a que la ciencia genera una titularidad compartida de bienes, dado que los hallazgos que se obtienen mediante la colaboración social se asignan a la comunidad. La reclamación que se hace de los derechos sobre propiedad intelectual se limita al reconocimiento que se le otorga al autor, poniendo Merton como ejemplos el sistema de Copérnico o la ley de Boyle. El postulado del desinterés no trae causa de una previa característica de la personalidad de quienes

hacen ciencia, sino que se fundamenta en el escrutinio público al que la actividad científica se halla sometida; de esta forma se genera una especial integridad. Por último, el postulado del escepticismo organizado consiste tanto en un método como en un mandato institucional: en cuanto al método, se trata de una suspensión temporal del juicio, separando las creencias personales para realizar el análisis y en cuanto a lo institucional, los descubrimientos científicos pueden invalidar dogmas institucionalizados en la sociedad, generando de esta manera conflictos o, en palabras de Kuhn, una necesidad de modificación de los paradigmas imperantes hasta una determinada fecha y su sustitución por otros de carácter novedoso. (Kuhn, 2001).

El Proyecto de recomendación de la UNESCO sobre la Ciencia Abierta

La segunda de las iniciativas a las que se ha de hacer referencia es la de la UNESCO, cuyo Comité ejecutivo aprobó el 7 de agosto de 2019 un plan para realizar una recomendación sobre la ciencia abierta. Tras los trabajos preliminares que incluyeron consultas a las instituciones de los Estados miembros, se presentó el 30 de septiembre de 2020 un borrador de recomendación (UNESCO, 2020). En el texto del documento propuesto se define el término de la ciencia abierta y se enumeran sus componentes:

6. A los efectos de la presente Recomendación, la ciencia abierta se define como un constructo inclusivo que combina diversos movimientos y prácticas con el fin de que los conocimientos científicos estén abiertamente disponibles y sean accesibles para todos, así como reutilizables por todos, se incrementen las colaboraciones científicas y el intercambio de información en beneficio de la ciencia y la sociedad, y se abran los procesos de creación, evaluación y comunicación de los conocimientos científicos a los agentes sociales más allá de la comunidad científica tradicional. Abarca todas las disciplinas científicas y todos los aspectos de las prácticas académicas, incluidas las ciencias básicas y aplicadas, las ciencias naturales y sociales y las humanidades, y se basa en los siguientes pilares clave: acceso abierto al conocimiento científico, infraestructuras de la ciencia abierta, comunicación científica abierta, participación abierta de los agentes sociales y diálogo abierto con otros sistemas de conocimiento.

La definición de la UNESCO se aproxima al concepto de ciencia abierta desde una perspectiva finalista con cuatro objetivos: (1) un conocimiento a disposición global, (2) que dicho conocimiento pueda ser reutilizado también globalmente, (3) el incremento de las colaboraciones científicas e intercambio de información y (4) la apertura de los procesos de creación, evaluación y comunicación de la ciencia a todas las personas que no integren la comunidad científica.

Además, se identifican los elementos constitutivos de la ciencia abierta, que son: (1) el acceso abierto a las publicaciones científicas, (2) datos de investigación abiertos, (3) programas informáticos de código abierto y código fuente abierto (free and open software), (4) equipos informáticos abiertos (open hardware), (5) infraestructuras de investigación compartidas, (6) participación abierta de los agentes sociales, (7) diálogo abierto con otros sistemas de conocimiento y (8) comunicación científica abierta.

El texto de la UNESCO es un borrador en el momento de escribir estas líneas¹⁵ y será objeto de votación en la siguiente asamblea general de la institución prevista para noviembre de 2021. Su redacción ha pasado ya por todos los trámites que pudieran sugerir modificaciones, por lo que, salvo imprevistos de última hora, es difícil que su literalidad se aparte de la aquí transcrita. En todo caso, la aproximación que la UNESCO realiza al fenómeno de la ciencia abierta evidencia la multiplicidad de frentes que se hallan abiertos hoy en día con respecto a la actividad científica. No se trata simplemente de un problema que afecte a quienes trabajan en ciencia, sino a la humanidad en su conjunto, como recientemente estamos comprobando con los efectos de la pandemia. El texto incluye las preocupaciones objeto de las declaraciones de Budapest, Bethesda y Berlín sobre el acceso abierto a las publicaciones científicas, que 20 años después siguen sin haber producido los frutos necesarios, e incluye junto con los antiguos problemas otros contextos de la ciencia donde se necesitan reformas, entre ellos el de los datos de la investigación objeto del próximo apartado.

Los principios FAIR

Tal y como se señaló en la introducción, el acrónimo FAIR (*findable, accessible, interoperable y reusable*) hace referencia a cuatro principios aplicables a la gestión y administración de datos, concretamente a la posibilidad de que los mismos puedan ser encontrados, sean accesibles, interoperables y reutilizables. Se busca que los datos de investigación estén al alcance de cuantas más personas mejor.

El análisis efectuado por la comunidad científica en cuanto a las condiciones en que se hallan los datos de investigación parte de dos premisas fácticas iniciales. En primer lugar, los datos se tratan de forma tanto manual como automatizadamente mediante máquinas. Los dos tipos de tratamiento tienen sus características y se precisa la concurrencia de ambos, puesto que el tratamiento manual permite una mayor sutileza lingüística por parte de quien trata los datos mientras que el tratamiento automatizado permite una mayor complejidad de cálculo, la utilización de datos masivos y una mejor búsqueda, entre otras posibilidades. En segundo lugar, si bien existen bases de datos bien administradas que continuamente enriquecen su contenido mediante nuevas aportaciones y cuya

utilización por parte de la comunidad científica se facilita mediante herramientas a su alcance, también existen otras disciplinas científicas donde la norma es la heterogeneidad de los datos, su dispersión y la dificultad de la reutilización.

Los problemas cotidianos que presenta este sistema heterogéneo y disperso de datos suelen darse en forma de preguntas que se hacen quienes han de utilizarlos: ¿Puedo integrar datos de una tercera parte en mi investigación? ¿Dónde encuentro datos sobre mi disciplina? ¿Existen datos anteriores que yo pueda utilizar? ¿Puedo comercializar mi investigación si he utilizado datos de una tercera parte? En realidad, no se trata de preguntas novedosas, sino que se pueden enmarcar en la tradición que iniciaron Charlotte Hess y Elinor Ostrom cuando en el año 2001 estudiaron los derechos de propiedad intelectual sobre los textos académicos. Frente a un sistema tradicional en el que la propiedad suponía una visión sobre un objeto material, en el mundo digital la facilidad de compartir la información obligaba a replantearse la pregunta sobre en qué consiste la propiedad, en este caso sobre un archivo que es una lista de unos y ceros. La respuesta que Hess y Ostrom obtuvieron de sus colegas de la academia les permitió diferenciar cinco actividades sobre las publicaciones científicas que eran las que les interesaban a quienes investigan: el acceso a las publicaciones, la posibilidad de extracción textos o datos de las mismas, su gestión, la posibilidad de excluir usos y la alienación de las anteriores actividades, esto es, transmitir derechos para efectuar las anteriores actividades (Hess y Ostrom, 2003). De esta manera, ante un sistema de información digital, la granularidad de las diferentes actividades que se podían efectuar sobre el mismo implicaba que la pregunta sobre la *propiedad* de la información ya no era relevante. Lo decisivo para poder construir conocimiento sería, no ya la propiedad sobre un objeto, sino la posibilidad de acceder, extraer, gestionar, excluir y alienar el mismo. Un ejemplo puede ser más clarificador: para gozar de las posibilidades que ofrecen unas webs como las de Wikipedia u Openstreetmap no es necesario ser el propietario de su contenido, puesto que las licencias de las mismas permiten descargar sus bases de datos, extraer de las mismas los datos que se considera, dividirla en cuantas otras bases de datos se tenga por conveniente, transformarlas e incluso venderlas... El derecho de propiedad, tan importante sobre aquellos objetos compuestos por átomos, pasa a un segundo plano cuando de información digitalizada se trata. La capacidad de excluir a otros del uso, que es una de las características principales del derecho dominical sobre un objeto material, pasa a segundo plano cuando se trata de información en soporte electrónico.

Para encontrar soluciones a la heterogeneidad en la que se hallan las bases de datos de uso científico, la comunidad investigadora comenzó a elaborar unos principios que pudieran luego solidificarse en prácticas. Tras un taller en la ciudad de Leiden (Holanda) celebrado en el año 2014 se publicó un borrador de los principios FAIR, borrador que luego se refinó y que dio lugar a un artículo

seminal, *The FAIR Guiding Principles for scientific data management and stewardship* (Wilkinson et al. 2016) en el que se detallan y se desarrollan cada uno de los principios (ver Tabla 1).

To be Findable:	F1. (meta)data are assigned a globally unique and persistent identifier F2. data are described with rich metadata (defined by R1 below) F3. metadata clearly and explicitly include the identifier of the data it describes F4. (meta)data are registered or indexed in a searchable resource
To be Accessible:	A1. (meta)data are retrievable by their identifier using a standardized communications protocol A1.1 the protocol is open, free, and universally implementable A1.2 the protocol allows for an authentication and authorization procedure, where necessary A2. metadata are accessible, even when the data are no longer available
To be Interoperable:	I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. I2. (meta)data use vocabularies that follow FAIR principles I3. (meta)data include qualified references to other (meta)data
To be Reusable:	R1. meta(data) are richly described with a plurality of accurate and relevant attributes R1.1. (meta)data are released with a clear and accessible data usage license R1.2. (meta)data are associated with detailed provenance R1.3. (meta)data meet domain-relevant community standards

Tabla 1. Fuente: Wilkinson et al. (2016)

Tal y como mencionan Wilkinson *et al.*, el tratamiento de datos de forma automatizada persigue como fin que una máquina pueda llevar a cabo las siguientes tareas de la forma más parecida a lo que una persona las realizaría: a) identificar el tipo de objeto digital que está tratando, b) determinar mediante el análisis de los datos o metadatos si el objeto es útil para la tarea que se está llevando a cabo, c) determinar su uso es posible de acuerdo con la licencia, consentimiento o cualquier límite existente sobre el contenido y d) realizar la acción que corresponda. Los cuatro principios FAIR guiarían la adaptación que los conjuntos de datos, las herramientas y los métodos necesitarían realizar para que las

decisiones anteriores pudieran tomarse sin intervención humana. Wilkinson *et al.* cuidan señalar que los principios FAIR no son un estándar, sino un conjunto de intervenciones tecnológicamente agnósticas que puede realizar quien gestione y administre los datos de la forma que entienda más conveniente. Bajo un punto de vista práctico, se trataría de que quien tuviera la posibilidad de alterar un conjunto de datos analizara si tal conjunto cumple con los principios y lo modificase para mejorar sus condiciones de ser hallado, de acceder al mismo, de su interoperabilidad y de su reutilización.

En los diferentes proyectos que integran la EOSC se han abrazado los principios FAIR de forma unánime, existiendo incluso alguno de ellos cuya única razón de ser es ayudar a consolidarlos y estudiar sus derivadas, como pudieran ser la creación de métricas evaluadoras. El propio EOSC Executive Board ha mantenido durante sus dos años de existencia un grupo de trabajo para la implementación de estos principios, grupo que por sí solo ha emitido cuatro informes, a los que hay que sumar otros dos en colaboración con el grupo de trabajo de arquitectura.¹⁶

Por otra parte, el proyecto FAIRsFAIR ha desarrollado 17 indicadores con los que comprobar el grado de utilización de los principios (Devaraju *et al.*, 2020). Los indicadores pueden referirse a datos, metadatos o documentación (licencias, instrucciones, guías que se acompañan al conjunto de datos) y también a aquellos elementos que permitirían la adquisición de características FAIR, como pudieran serlo las infraestructuras necesarias para que dichos datos puedan seguir estando accesibles durante un período de tiempo. No obstante, a pesar de estas métricas, quedarían tres cuestiones pendientes de estudio y análisis. La primera cuestión hace referencia a la calidad de los datos, aspecto del que los principios FAIR nada dicen; la segunda es la ventaja de la que siempre gozarán quienes crean los datos, puesto que su contacto con la realidad estudiada y su decisión sobre los parámetros con los que se efectúa la catalogación les colocan en una posición de privilegio (Paschetto, Borgman y Wofford, 2019). Y por último, el aspecto de cuáles son los datos que merecen ser conservados, lo que supondría tomar una decisión sobre su ciclo de vida que, al tratarse de un formato electrónico, no serían aplicables las viejas normas archivísticas del expurgo sobre valoración, selección y eliminación de los archivos, sino que consistiría en la utilización de los principios FAIR pero en sentido inverso, esto es, dificultar la posibilidad de hallar el conjunto de datos y acceder al mismo, evidenciando en todo caso que dichos datos se hallan obsoletos mediante las pertinentes modificaciones de sus metadatos.

Conclusión

Para concluir la presente ponencia, simplemente señalar que el panorama actual de la ciencia abierta, su documentabilidad y los principios FAIR, presenta interesantes desarrollos que pueden verse beneficiados de los conocimientos y experiencia acumulados por la archivística. Sería muy interesante y productiva una mirada cruzada entre ambas, ciencia y archivística, por el enriquecimiento recíproco que pueden aportarse.

Bibliografía

- Buranyi, S. (2017, junio 27). "Is the staggeringly profitable business of scientific publishing bad for science?" *The Guardian*. <http://www.theguardian.com/science/2017/jun/27/profitable-business-scientific-publishing-bad-for-science>
- Burgelman, J. C. (2021). "Politics and Open Science: How the European Open Science Cloud Became Reality (the Untold Story)". *Data Intelligence. Volume 3, Issue 1 Winter 2021*. https://doi.org/10.1162/dint_a_00069
- Comisión Europea (2007). Comunicación de la Comisión al Parlamento Europeo, al Consejo y al Comité Económico y Social Europeo sobre la información científica en la era digital: acceso, difusión y preservación {SEC (2007)181} <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A52007DC0056>
- Comisión Europea (2012a). <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=SWD:2012:0221:FIN:ES:PDF>
- Comisión Europea (2012b). <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=SWD:2012:0211:FIN:ES:PDF>
- Comisión Europea (2016). 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. COM (2016) 178 final. http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=15266
- Comisión Europea (2018). Commission Staff Working Document. Implementation Roadmap for the European Open Science Cloud SWD (2018) 83 final. https://ec.europa.eu/research/openscience/pdf/swd_2018_83_f1_staff_working_paper_en.pdf
- Cruz Mundet, Pedro (2011). Manual de archivística. Fundación Germán Sánchez Ruipérez.
- Devaraju, A.; Huber, R.; Mokrane, M.; Herterich, P.; Cepinskas, L.; de Vries, J.; L'Hours, H.; Davidson, J.; White, A. (2020). FAIRsFAIR Data Object Assessment Metrics (Version 0.4). <https://zenodo.org/record/4081213>
- EOSC Executive Board (2021) Strategic Research and Innovation Agenda (SRIA) of the European Open Science Cloud (EOSC) https://www.eosc.eu/sites/default/files/EOSC-SRIA-V1.0_15Feb2021.pdf

- Gallego Domínguez, Olga y López Gómez, Pedro (1989). Introducción a la archivística. Gobierno Vasco, Servicio Central de Publicaciones.
- Hess, C., & Ostrom, E. (2003). "Artifacts, Facilities, and Content: Information as a Common-pool Resource". *Law & Contemporary Problems*, 66 (Winter/Spring 2003), 111-145.
- Kuhn, Thomas S. (2001). *La estructura de las revoluciones científicas*. Fondo de Cultura Económica.
- Lafuente, Antonio y Alonso Puellas, Andoni (2011). *Ciencia expandida, naturaleza común y saber profano*. Universidad Nacional de Quilmes.
- Merton R. K. (1974). *The sociology of science: Theoretical and empirical investigations*. University of Chicago Press.
- Pasquetto, Irene V., Christine L. Borgmann y Wofford Morgan F. (2019). "Uses and Reuses of Scientific Data: The Data Creators' Advantage". *Harvard Data Science Review* 1.2. DOI: 10.1162/99608f92.fc14bf2d
- UNESCO. (2020). First draft of the UNESCO Recommendation on Open Science. <https://unesdoc.unesco.org/ark:/48223/pf0000374837.locale=en>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Mons, B. (2016). "The FAIR Guiding Principles for scientific data management and stewardship". *Scientific Data*, 3(1), 160018. <https://doi.org/10.1038/sdata.2016.18>

¹⁵ Julio de 2021.

¹⁶ Pueden accederse en el siguiente enlace: <https://www.eoscsecretariat.eu/eosc-governance/eosc-executive-board-outputs>

Notas

¹ Gallego Domínguez y López Gómez distinguen entre dos tipos de archiveros, los *científicos* y los *técnicos* (1989, p. 38).

² En entrevista personal con el autor.

³ <https://www.budapestopenaccessinitiative.org/read>

⁴ <https://legacy.earlham.edu/~peters/fos/bethesda.htm>

⁵ <https://openaccess.mpg.de/Berlin-Declaration>

⁶ <https://www.un.org/es/about-us/universal-declaration-of-human-rights>

⁷ <https://www.ohchr.org/SP/ProfessionalInterest/Pages/CESCR.aspx>

⁸ <https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52007DC0056&from=EN>

⁹ <https://ec.europa.eu/digital-single-market/news/final-report-science-20-public-consultation>

¹⁰ Traducción propia.

¹¹ https://en.wikipedia.org/wiki/Open_science

¹² <https://www.fosteropenscience.eu/>

¹³ <https://www.fosteropenscience.eu/content/what-open-science-introduction/>

¹⁴ Traducción propia.