

Motivación, barreras e incentivos para la compartición y reutilización de los datos de investigación. Visión de los investigadores

Motivation, barriers and incentives for research data sharing and research data reuse. Researchers' views

Carolina Navarro-Molina; Remedios Melero

Cómo citar este artículo:

Navarro-Molina, Carolina; Melero, Remedios (2019). "Motivación, barreras e incentivos para la compartición y reutilización de los datos de investigación. Visión de los investigadores". *El profesional de la información*, v. 28, n. 5, e280516.

<https://doi.org/10.3145/epi.2019.sep.16>

Artículo recibido el 12-03-2019
Aceptación definitiva: 09-08-2019



Carolina Navarro-Molina ✉
<https://orcid.org/0000-0003-0713-4100>
Instituto de Agroquímica y Tecnología de Alimentos-CSIC
Catedrático Agustín Escardino, 7.
46980 Paterna (Valencia), España
canamillo@gmail.com



Remedios Melero
<https://orcid.org/0000-0002-1813-8783>
Instituto de Agroquímica y Tecnología de Alimentos-CSIC
Catedrático Agustín Escardino, 7.
46980 Paterna (Valencia), España
rmelero@iata.csic.es

Resumen

Compartir en abierto y facilitar el acceso a los datos de investigación, su uso y reutilización no son prácticas habituales entre los investigadores, a excepción de determinadas disciplinas, como la meteorología, la genética o la astronomía. Las políticas de las instituciones, de algunas editoriales y de las agencias que financian la investigación han provocado que esta cultura de compartir y gestionar los datos de investigación cobre una especial relevancia. Sin embargo, existen barreras como el miedo de los investigadores a que les quiten la primicia y a ser copiados, a que se haga un mal uso de sus datos, a perder el control o infringir la legislación sobre la propiedad intelectual, con las que se enfrenta la comunidad científica para poder hacer que sus datos sean localizables, accesibles, interoperables y reutilizables (datos FAIR). En este trabajo se han analizado los pros y los contras de compartir y reutilizar los datos, basándonos en trabajos previos que han permitido vislumbrar hábitos y opiniones de los investigadores a este respecto.

Palabras clave

Datos de investigación; Compartición de datos; Reutilización de datos; Hábitos de investigadores; Motivación, Barreras.

Abstract

Research data sharing, data reuse and facilitating access to research data are not common researchers' practices, except in certain disciplines, such as Meteorology, Genetics or Astronomy. Institutional, publishers' and funders' policies have made the culture of sharing and managing research data particularly relevant. However, the threat of being scooped, the fear to the misuse of their data, to lose control over their data, or the fear to infringe the legislation on intellectual property, are barriers that scientific community faces when trying to make their data findable, accessible, interoperable and reusable (FAIR data). In this work, we have analyzed the pros and cons of sharing and reusing data, based on previous works in the literature that have allowed us to perceive researchers' habits and perceptions in this regard.

Financiación

Las autoras agradecen al *Ministerio de Ciencia, Innovación y Universidades* la concesión del proyecto de *Plan Nacional* con el código RTI2018-094360-B-I00.

Keywords

Research data; Data sharing; Data re-use; Researchers' habits; Motivation; Barriers.

1. Introducción

El interés y la importancia por compartir los datos de investigación en el ámbito académico se ha incrementado paralelamente al aumento de la producción científica (Tenopir *et al.*, 2011; Schmidt; Gemeinholzer; Treloar, 2015). La evolución tecnológica ha permitido un mejor y más rápido acceso a grandes conjuntos de datos (*big data*) que representan una oportunidad para realizar análisis sobre cuestiones complejas, imposibles de ejecutar hasta ahora (Schmidt; Gemeinholzer; Treloar, 2015) y que se han convertido en esenciales para la toma de decisiones (Van-Panhuis *et al.*, 2014). La disponibilidad, el acceso a los datos y su reutilización, no sólo es importante para asegurar el retorno a la sociedad de la inversión en investigación y poder generar nuevos datos y nuevos productos, sino también para asegurar la perdurabilidad de los mismos (Costello, 2009).

El acceso a los datos, su reutilización y gestión han sido impulsados por las políticas que han llevado a cabo los diferentes agentes implicados en el proceso de la producción científica: entidades y organismos financiadores, gobiernos, universidades, organismos de investigación y editores (Piwowar; Chapman, 2010). Esta situación constata la necesidad de una colaboración integral de la comunidad científica (Stuart *et al.*, 2018), facilitada por un entorno digital en el que los investigadores juegan al mismo tiempo el rol de productores y consumidores de información (Damalas *et al.*, 2018). Entre estas políticas, destacan las de la *Comisión Europea*, que, desde el *Séptimo Programa Marco*, ha abogado por el acceso abierto a las publicaciones y a los datos de investigación. En concreto, respecto a los datos, en el *Programa Marco Horizon 2020 (H2020)* la *Comisión* requiere a los investigadores el depósito en abierto de los datos subyacentes a las publicaciones científicas resultado de los proyectos investigación financiados por Europa (*European Commission*, 2016a). La *CE* también requiere la presentación de un plan de gestión de datos en distintas etapas del proyecto, en donde se describan los datos que se van a generar y cómo van a facilitar que sean datos FAIR. El lema de la *Comisión* respecto a los datos es "tan abiertos como sea posible y tan cerrados como sea necesario". En el caso de que hubiera algún motivo que exima del depósito en abierto debe especificarse en el plan de gestión de datos. Algunos de los motivos para ello serían que los datos fueran sensibles sobre personas, que afecten a la seguridad de una región, o que estén sujetos a derechos de propiedad intelectual o industrial. De forma análoga el *European Research Council (ERC)*, 2017), la *National Science Foundation* (2011), el *Wellcome Trust* (2010), la *Bill & Melinda Gates Foundation* (2015), entre otras entidades financiadoras de proyectos de investigación, requieren el depósito de los datos de investigación y promueven que los datos cumplan con los principios FAIR (*Fair data principles*, 2016; figura 1): *findable* (localizable), *accessible* (accesible), *interoperable* (interoperable) y *reusable* (reutilizable).

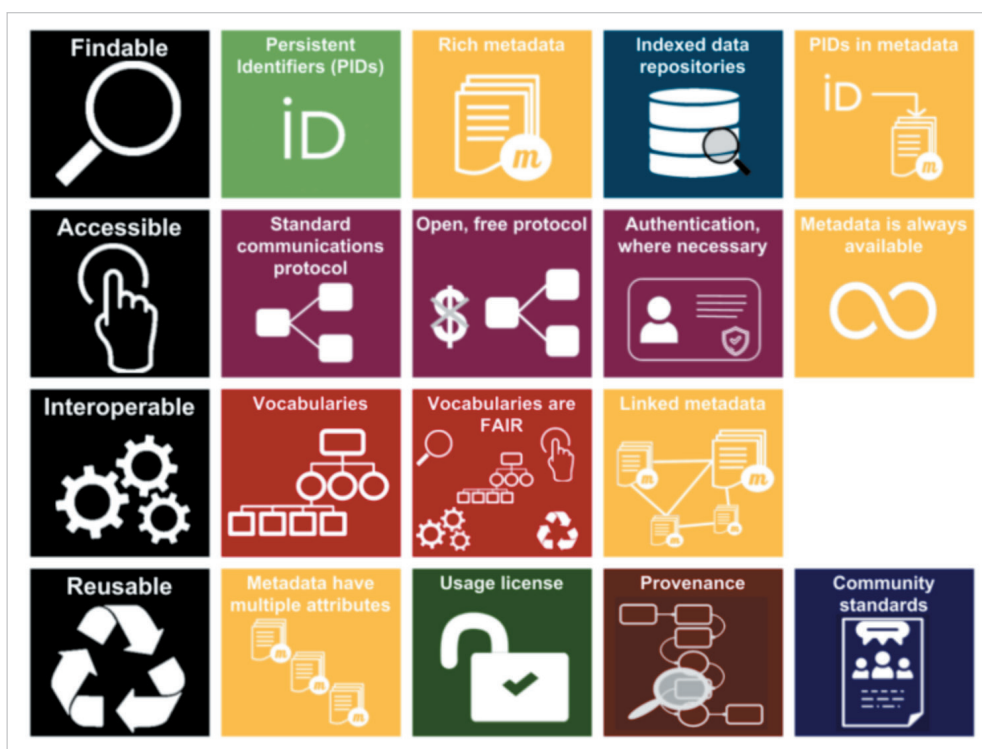


Figura 1. Principios FAIR y sus componentes.

https://www.andis.org.au/_data/assets/image/0011/1416098/FAIR-Data-image-map-graphic-v2-721px.png

La *Comisión Europea* estableció en 2016 (*European Commission, 2016b*) cinco acciones fundamentales en el borrador de la agenda europea para favorecer la ciencia abierta:

- incentivarla y promoverla;
- eliminar barreras que impidan su desarrollo;
- desarrollar infraestructuras que la permitan;
- monitorizar los resultados de la investigación según las políticas de acceso abierto;
- integrar la ciencia abierta en la sociedad.

Estas acciones sirvieron para que la *Open Science Policy Platform (OSPP)* estableciera en 2018 unas recomendaciones para lograr que los datos de investigación fueran FAIR (*European Commission, 2018a; 2018b*) basadas en:

- incentivos y recompensas;
- nuevas métricas de evaluación;
- futuro de la comunicación científica;
- *European Open Science Cloud (EOSC)*;
- integridad en la investigación;
- competencias profesionales;
- ciencia ciudadana.

El 4 de abril de 2019, el *Parlamento Europeo* aprobó el borrador de la directiva que regulará la reutilización de datos generados con fondos públicos, para facilitar el acceso a los datos del sector público con el objetivo de crear nuevos productos y servicios que contribuyan al desarrollo de una sociedad basada en los datos compartidos (*Parlamento Europeo, 2019*). La novedad más relevante respecto a la anterior directiva es que afecta también a los datos generados de la investigación financiada con fondos públicos y establece que las políticas nacionales deben incluir en sus términos la obligatoriedad de que estos datos estén en abierto por defecto, siempre bajo el citado principio de “tan abiertos como sea posible y tan cerrados como sea necesario”. Establece también los casos en que está justificada la exención del cumplimiento de esta norma cuando se trate de datos personales, que afecten a la privacidad, confidencialidad y seguridad nacional o estén afectados por los derechos de propiedad intelectual de terceros.

En el caso de las revistas, son cada vez más las que implementan políticas y procedimientos por los que solicitan que los trabajos publicados incluyan los datos como material suplementario (publicación enriquecida), bien se depositen previamente en un repositorio de datos relacionado con la disciplina (*ANDS, 2017*) o multidisciplinar (**Herold, 2015; Piwowar; Chapman, 2010; Stodden; Guo; Ma, 2013**), o se publiquen en las revistas denominadas *data journals*, en las que se describen los datos y su forma de obtención, en lugar de exponer los resultados de investigación de forma convencional (**Pampel; Dallmeier-Tiessen, 2014; García-García; López-Borrull; Peset, 2015; Candela et al., 2015**).

En la *Research Data Alliance (RDA)* se ha creado un grupo de interés que ha elaborado seis modelos de tipos de políticas editoriales en función de los requisitos o recomendaciones con respecto a los datos de investigación que subyacen a las publicaciones. Entre las 14 variables que definen esos modelos se encuentran los posibles embargos para abrir los datos, los repositorios donde poder depositarlos y localizarlos, licencias, formatos y estándares (**Hrynaszkiewicz et al., 2019**).

La reutilización de los datos de investigación aporta beneficios no sólo económicos, sino que incentivan nuevos avances y propician oportunidades de co-autoría (**Gonçalves-Curty, 2017**). Sin embargo, a pesar de los beneficios que compartir y reutilizar los datos puedan tener para los investigadores (**Wiley, 2018; Piwowar, 2011; Costello, 2009; Michener, 2015**), en la práctica presenta numerosos desafíos técnicos debidos a su gestión (**Schmidt; Gemeinholzer; Treloar, 2015**) y a las actitudes y hábitos de los investigadores. En genómica, por ejemplo, la tecnología ha salvado las barreras de la distancia para compartir los datos, sin embargo persisten otras derivadas del factor humano (resistencia, competición, hábitos) que ralentizan la puesta en común de los datos (**Fusi et al., 2018**).

Tabla 1. Pros y contras de compartir y reutilizar datos de investigación

Pros	Contras
Se promueve la innovación y generación de nuevos servicios	Miedo al uso fraudulento
Facilita la colaboración entre grupos de trabajo multidisciplinares	Reticencia a compartir por el esfuerzo de haberlos obtenido
Transparencia y fiabilidad de la fuente de datos	Perder la primicia
Facilita la reproducibilidad de los resultados	No saber cómo compartir los datos
Permite verificar la bondad de los datos	Desconocer cómo elaborar planes de gestión de datos
Reduce duplicidad de esfuerzos	Desconocer cómo licenciar
Visibiliza la investigación	Miedo a infringir algún tipo de normativa legal
Al licenciar los datos se pueden establecer las condiciones de uso	Inseguridad en los sistemas de almacenamiento y preservación
Integración de los datos en el ciclo de vida de la comunicación científica y reconocimiento de su impacto	Miedo a la pérdida de control sobre el uso
Participación ciudadana en el uso y generación de datos	Falta de cultura de cómo citar los datos
Retorno a la sociedad de la inversión en investigación	Falta de incentivos y reconocimiento en la carrera investigadora

En salud pública, además de las dificultades técnicas, económicas o de motivación que precisan de soluciones estructurales, existen también consideraciones políticas, legales y éticas que requieren del consenso internacional para propiciar un intercambio de datos fluido (Van-Panhuys *et al.*, 2014).

El debate generado respecto a los pros y contras de compartir y reutilizar los datos de investigación (tabla 1) se ha traducido en la publicación de trabajos en los que se analizan los factores que afectan a los investigadores (Kim; Schuler; Pechenina, 2017), en los que se tienen cuenta los rasgos comunes que afectan a toda la comunidad científica, y también los específicos propios de la disciplina. En este sentido, el objetivo de este trabajo ha sido analizar los trabajos más relevantes publicados sobre los hábitos y actitudes de los autores respecto a compartir, reutilizar y gestionar los datos de investigación, basados en encuestas y entrevistas personales en las que se reflejan estas tres cuestiones.

Los trabajos seleccionados para esta revisión se basan en consultas online dirigidas a investigadores de diferentes áreas de conocimiento y procedencias. Las cuestiones se pueden agrupar en las categorías que se resumen en la tabla 2. División puede servir de modelo de partida en futuros estudios relacionados con la gestión de los datos de investigación.

Tabla 2. Categorías y temas relacionados presentes en las encuestas sobre compartir, reutilizar y gestionar datos de investigación

Categoría	Subcategoría	Ítem
1. Demografía	Personales	Género. Edad. País. Nombre. Correo contacto. Afiliación. Posición. Área de trabajo. Antigüedad
	Trabajo/Especialidad/Organización	
2. Datos de investigación	Formatos	Tipo. Formatos. Software. Tamaño. Metadatos. Gestión de datos. Preservación. Responsable administración de datos. Almacenamiento. Seguridad
	Volumen	
	Metadatos	
	Almacenamiento/Preservación	
3. Entidad financiadora	Seguridad	Entidad y tipo de financiación. Demanda de datos FAIR. Propiedad de los datos. Financiación para el plan de gestión de datos (PGD). Exigencia del PGD. Responsable de la supervisión del PGD
	Fuente de financiación	
	Exigencia de datos FAIR	
	Propiedad de los datos	
	Financiación del <i>Plan de Gestión de Datos</i> (PGD)	
4. Compartir y reutilizar	Exigencia DMP	Importancia y beneficios de compartir datos. Requisitos de las entidades financiadoras y de las editoriales. Reproducibilidad. Quién comparte. Publicaciones vinculadas a los datos en bruto. Reutilización. Permisos. Autoría. Embargos. Propiedad. Licencias. Normas. Fiabilidad. Descubrimiento de datos. Infraestructura
	Responsable de la DMP	
	Razones y beneficios de compartir datos	
	Repositorios	
	Experiencia compartiendo datos	
	Condiciones para compartir los datos propios	
	Condiciones para el uso de datos de otros	
Formas de descubrir los datos		
5. Publicar en abierto	Software	Experiencia. Motivaciones Beneficios de compartir / publicar datos. Actitudes. Adopción de la insignia de datos abiertos. Directrices claras para publicar datos. Infraestructura para acceder /publicar/gestionar datos. Herramientas de gestión de datos. Conocimiento de las herramientas PGD
	Experiencia publicando en abierto	
	Motivaciones para publicar en abierto	
6. Obstáculos	Servicios	Calidad de los datos. Esfuerzos requeridos para compartir datos. Confianza. Inquietudes sobre la autoría. Cómo se recopilan los datos. Procedencia de los datos. Uso de datos en su campo. Disponibilidad de datos
	Compartir y reutilizar datos	
7. Citación / Métricas	Publicar	Creación de una cultura de citación de datos. Impacto de los datos. Estadística. Métricas
	Citación	
8. Licencias	Métricas	Elección de licencias. Propiedad de los datos. Cuestiones legales asociadas al uso y reutilización de los datos
	Elección	
	Conocimiento de la licencia	
9. Formación	Problemas legales	Qué metadatos utilizar. Cómo gestionar los datos. Formación periódica
	Responsable de la formación	
	Temas para la formación	

En la tabla 3 se ha realizado una síntesis de las audiencias, objetivos y principales resultados extraídos de la bibliografía que se refieren a distintos aspectos relacionados con la creación, reutilización, preservación, intercambio de datos, y prácticas habituales de los investigadores respecto a la gestión de los datos de investigación, entre otros. Estos trabajos se han tenido en cuenta para la redacción de los siguientes apartados relativos a las principales motivaciones, barreras y condiciones que afectan a la gestión de datos de investigación y los retos y desafíos que conlleva.

Tabla 3. Resumen de trabajos de la bibliografía que utilizan encuestas online con partes implicadas en la generación y gestión de datos de investigación. Principales objetivos, resultados y conclusiones de los mismos

Referencia	Enfoque y audiencia	Objetivos	Conclusiones y resultados
Tenopir et al. (2011)	Multidisciplinar. Profesores ayudantes y titulares (N=1.329)	Estudiar las prácticas de datos de los investigadores sobre acceso, descubrimiento, reutilización, preservación e intercambio de datos.	<ul style="list-style-type: none"> Las barreras para el intercambio efectivo de datos y la preservación están profundamente arraigadas en las prácticas y la cultura del proceso de investigación, así como en los propios investigadores. Los nuevos mandatos para los planes de manejo de datos de la <i>National Science Foundation (NSF)</i> y otras agencias federales, y la atención mundial a la necesidad de compartir y preservar datos podrían llevar a cambios a gran escala. Los programas como <i>DataNET</i>, patrocinado por <i>NSF</i> (que incluye proyectos como <i>DataONE</i>), atraerán la atención y los recursos al problema y facilitarán a los científicos la aplicación de principios sólidos de administración de datos.
Sayogo; Pardo (2013)	Multidisciplinar. Miembros de la <i>Red DataOne</i> (N= 555)	Analizar los desafíos para hacer públicos y preservar los datos de investigación mediante el examen sistemático de los determinantes de los investigadores para compartir sus datos de investigación.	<ul style="list-style-type: none"> El apoyo institucional en las tareas de gestión de datos es determinante para incentivar a los investigadores a publicar sus datos Otro factor que influye es el reconocimiento del trabajo que realiza el creador del conjunto de datos para facilitar el acceso y reutilización, en cumplimiento de algunas políticas institucionales o gubernamentales. El impacto de la publicación de los datos será también un factor determinante para que aumente su distribución.
Kim; Adler (2015)	Ciencias sociales. Investigadores (N= 361)	Localizar los factores individuales, institucionales y de recursos que influyen en los comportamientos respecto al intercambio de datos entre los científicos de ciencias sociales.	<ul style="list-style-type: none"> Los comportamientos de intercambio de datos de los científicos en ciencias sociales se rigen principalmente por motivaciones personales (es decir, beneficio y riesgo percibidos, esfuerzo percibido y actitud hacia el intercambio de datos) y presión normativa percibida. La presión de las agencias de financiación, la presión de las revistas y disponibilidad de repositorio de datos no son factores significativos para influir en el intercambio de datos.
Kratz; Strasser (2015)	Multidisciplinar. Investigadores y gestores de datos (N= 247)	Lograr una mejor comprensión de cómo medir el impacto de los datos, aspecto fundamental para recompensar a sus creadores e incentivar la publicación de datos.	<ul style="list-style-type: none"> A corto plazo, las vistas de páginas y la actividad de las redes sociales no son métricas de datos muy destacadas debido a que no registran mucha actividad en el primer caso, y a la falta de datos que la cuantifiquen en el segundo. Aunque hoy por hoy es un reto, las citas deben ser puestas en valor y recolectadas lo mejor posible. Las descargas deben impulsarse como un medio interesante de obtener datos: tanto razonablemente valioso como razonablemente fácil de medir.
Schmidt; Gemeinholzer; Treloar (2015)	Multidisciplinar. Investigadores, científicos de datos, gestores de datos y tecnólogos (N=1.253)	<p>Identificar actividades clave de datos abiertos en varias comunidades que se ocupan del cambio ambiental global para identificar ejemplos destacados de mejores prácticas desde la perspectiva del usuario.</p> <p>Determinar áreas donde el deseo de compartir de los usuarios podría ser mejorado por nuevos desarrollos.</p> <p>Identificar barreras para "abrir el intercambio de datos" desde la perspectiva del usuario (ya sea como proveedor o como usuario de datos)</p>	<ul style="list-style-type: none"> Ejemplos de buenas prácticas que demuestran una aceptación sustancial del intercambio de datos a través de infraestructuras electrónicas y una necesidad adicional de mejora y consolidación. Las políticas de las entidades financiadoras parecen ser el motivador más importante, lo que respalda la idea de que mandatos más exigentes fortalecerán el intercambio de datos.

<p>Tenopir et al. (2015)</p>	<p>Multidisciplinar. Profesores e investigadores (N=2.344)</p>	<p>Examinar el estado actual del intercambio de datos y comparar con el estudio de referencia de 2009/2010.</p> <p>Examinar las diferencias en las prácticas y percepciones entre grupos de edad, regiones geográficas y disciplinas.</p>	<ul style="list-style-type: none"> • Una mayor aceptación y disposición a participar en el intercambio de datos, así como a un aumento en los comportamientos reales de intercambio. Sin embargo, también hay un mayor riesgo percibido asociado con esto y persisten las barreras específicas para el intercambio. • Hay diferencias entre los grupos de edad: los encuestados más jóvenes son más favorables al intercambio y la reutilización, sin embargo, ponen menos datos disponibles que los encuestados de mayor edad. • Existen diferencias geográficas, que pueden entenderse en parte por las diferencias culturales. • Se observan diferencias entre disciplinas.
<p>Kim; Stanton (2016)</p>	<p>Especializado: Ciencia, tecnología, ingeniería y matemáticas. Miembros de la Facultad e investigadores postdoctorales (N=1.317)</p>	<p>Investigar los factores institucionales e individuales que influyen en las conductas de intercambio de datos de los científicos en diferentes disciplinas científicas.</p>	<ul style="list-style-type: none"> • La presión reguladora por parte de las revistas, la presión normativa a nivel de disciplina y el beneficio percibido en la carrera académica, y el altruismo académico tuvieron relaciones positivas significativas con conductas de intercambio de datos, mientras que el esfuerzo percibido tuvo una relación negativa. • La presión reguladora por parte de las agencias de financiación y la disponibilidad de repositorios de datos a nivel de disciplina no influyen significativamente en las conductas de intercambio de datos.
<p><i>Socrata</i> (2016)</p>	<p>Multidisciplinar. Trabajadores de entidades federales de los Estados Unidos (N=506)</p>	<p>Comprender mejor las actitudes y opiniones actuales sobre el entorno de datos abiertos entre editores y usuarios.</p>	<ul style="list-style-type: none"> • El compromiso con los datos abiertos es amplio, con el gobierno federal a la vanguardia. • La inversión en datos abiertos continúa creciendo en todos los niveles gubernamentales. • Los datos abiertos conducen al desarrollo económico en todo el país. • Los datos abiertos están fuertemente vinculados a la mejora de la eficiencia gubernamental.
<p>Johnson; Chiarelli; Parsons (2016)</p>	<p>Multidisciplinar. Estudiantes, profesores, becarios, personal administrativo y técnico (N=1.185)</p>	<p>Examinar el servicio compartido de datos de investigación piloto del JISC para ofrecer un depósito fácil de los datos con herramientas para su descubrimiento, almacenamiento seguro, archivo a largo plazo y conservación.</p>	<ul style="list-style-type: none"> • El uso en los PGD sigue siendo generalmente bajo. Las razones más comunes para tener un PGD son que constituye una buena práctica de investigación (72%) y que es requerido por el financiador del proyecto (53%). • El 92% de los encuestados realiza una copia de seguridad de al menos algunos de sus datos, generalmente de forma personal con una frecuencia semanal o diaria. • Las soluciones de copia de seguridad más comunes son unidades externas o tarjetas de memoria, almacenamiento de copia de seguridad administrado por la universidad y unidades de nube. • El intercambio de datos de investigación generalmente se maneja mediante el uso de servicios de almacenamiento en la nube (67%), archivos de datos en el correo electrónico (60%) y dispositivos de almacenamiento portátiles (32%).
<p>Gonçalves-Curty et al. (2017)</p>	<p>Multidisciplinar. Académicos, empleados del gobierno y agencia comercial y sin fines de lucro (N=1.015)</p>	<p>Probar la relación entre las creencias y actitudes de los científicos hacia la reutilización de datos y su comportamiento de reutilización de datos autoinformados.</p>	<ul style="list-style-type: none"> • La eficacia percibida y la eficiencia de la reutilización de datos son predictores sólidos del comportamiento de reutilización, y que la importancia percibida de la reutilización de datos corresponde a una mayor reutilización. • Los resultados sugieren que la reutilización de los datos se normalizaría y aumentaría mediante la demostración de su valor.
<p>Berghmans et al. (2017)</p>	<p>Multidisciplinar. Investigadores (N=1.200)</p>	<p>Examinar las prácticas, motivaciones y obstáculos para el intercambio de datos.</p> <p>Descubrir cuáles son las ventajas percibidas de la comparación de datos.</p> <p>Determinar cómo los investigadores comparten datos, sus actitudes hacia la información y por qué pueden ser reticentes a compartir datos.</p>	<ul style="list-style-type: none"> • Las prácticas de intercambio de datos dependen de la disciplina. Las iniciativas de política general hacia los datos abiertos podrían beneficiarse de alentar soluciones de abajo hacia arriba en campos donde los datos abiertos ya forman parte integral del diseño de la investigación. • Aunque el intercambio de datos parece tener un beneficio global, los factores culturales y nacionales plantean un desafío importante para un enfoque único para todos. • El facilitar el acceso y compartición de datos depende de cada disciplina de los investigadores y de las prácticas en su comunidad científica. A eso se suma la normativa propia respecto a la privacidad de los datos y a las licencias de uso. • Tanto en el campo del intercambio de datos intensivo como en el restringido, se deben proporcionar facilidades de capacitación y soporte. • Las revistas de datos siguen siendo un fenómeno relativamente de pequeña escala, pero su popularidad está creciendo rápidamente. • Los datos no siempre se consideran un bien público, sino algo que hay que pagar, lo que podría suponer una amenaza para los datos abiertos. • Cuando se gestionan los datos, a menudo se percibe como una carga, y no como una responsabilidad.

Fecher et al. (2017)	Multidisciplinar. Investigadores, profesores, investigadores y estudiantes (N=1.564)	Averiguar qué impulsa el acceso abierto a los datos de investigación.	<ul style="list-style-type: none"> Independientemente de la disciplina, los investigadores reconocen los beneficios del acceso abierto a los datos de investigación, tanto para su propia investigación como para el progreso científico en general. No obstante, la mayoría comparte sus datos de forma selectiva. El sistema de reconocimiento académico imperante no contribuye a recompensar a los investigadores que comparten sus datos.
Eckes (2018)	Geografía. Investigadores (N=23)	Establecer el conocimiento de la evaluación de los principios FAIR y las prácticas de gestión de datos para los científicos noveles (PhD) en geografía.	<ul style="list-style-type: none"> <i>Dataset</i> de los resultados de la encuesta
Schöpfel et al. (2018)	Multidisciplinar. Directores de laboratorios de investigación (N=432)	Obtener la opinión y el comportamiento de los científicos franceses (nivel directivo superior) sobre la gestión de datos de investigación.	<ul style="list-style-type: none"> Hay diferencias significativas por disciplinas respecto a la producción de datos (tipos), la administración (recursos humanos, financiación y estándares), el intercambio de datos y las necesidades relacionadas. Parece que la gestión y el intercambio de datos no está directamente relacionado con el compromiso de acceso abierto. Con respecto a los principios de FAIR, el documento revela que el 68% de los directores de laboratorio afirman que su producción y gestión de datos cumple con al menos uno de los principios de FAIR. Sólo el 26% cumple con al menos tres principios, y menos del 7% cumple con los cuatro criterios FAIR. Los laboratorios de física nuclear, ciencias sociales y humanidades, y ciencias de la tierra y astronomía están a la vanguardia de otras disciplinas, especialmente en lo que respecta a la capacidad de búsqueda y la disponibilidad de sus datos.
Stuart et al. (2018)	Multidisciplinar. Investigadores (N=7.700)	Comprender la actividad de los investigadores sobre compartir datos durante el ciclo de vida de la investigación. Explorar las actitudes, acciones y desafíos para compartir datos.	<ul style="list-style-type: none"> Los investigadores están compartiendo datos de sus trabajos publicados tanto en repositorios como en información complementaria. Compartir datos como información complementaria ayuda a que estén más disponibles, pero su localización, acceso y la falta de especificación de licencia de uso no contribuyen a que datos sean FAIR. Los hallazgos confirman que los esfuerzos de los investigadores para archivar, publicar y compartir datos continúan siendo obstaculizados por las limitaciones de tiempo y la falta de conocimiento sobre los estándares de datos, los metadatos y la experiencia en gestión, las opciones de repositorio y los requisitos de las agencias financiadoras. Las diferencias temáticas y regionales existen, lo que sugiere que los enfoques específicos pueden ser útiles. Existen desafíos globales comunes que requieren atención concertada: más educación y apoyo para los investigadores, y rutas más rápidas y fáciles para compartir datos de manera óptima.
Zenk-Möltgen et al. (2018)	Sociología y ciencias políticas. Profesores, investigadores y estudiantes (N=446)	Investigar cómo diferentes factores institucionales e individuales afectan al comportamiento de intercambio de datos de los investigadores	
Harper; Kim (2018)	Psicología. Investigadores (N=341)	Investigar los factores de actitud, normativos y de recursos que afectan la adopción de un distintivo de datos abiertos por parte de los psicólogos.	<ul style="list-style-type: none"> El beneficio y el riesgo percibido tuvieron relaciones tanto positivas como negativas con la actitud hacia la obtención de la insignia de datos abiertos. La posibilidad de obtención de un distintivo (insignia) y las normas de intercambio de datos influyen positivamente en las motivaciones de los investigadores en psicología. La disponibilidad de un repositorio de datos y la presión de una revista científica abierta no tenían ninguna relación significativa con la intención conductual de adoptar la insignia de datos abiertos.
Houtkoop et al. (2018)	Psicología. Estudiantes, postdoctorados y profesores (N=600)	Descubrir las barreras percibidas para el intercambio de datos públicos y las posibles medidas para reducirlas.	<ul style="list-style-type: none"> Los datos se comparten con poca frecuencia. Compartir no es una práctica común en la disciplina. Se prefiere compartir bajo petición. Compartir requiere trabajo extra. Compartir requiere nuevas habilidades.

Kim; Oh (2018)	Multidisciplinar. Investigadores (N=221)	Investigar cómo la disciplina, los entornos institucionales y las motivaciones individuales afectan al hecho de que los investigadores depositen sus trabajos en un repositorio institucional.	<ul style="list-style-type: none"> • El miedo a perder la exclusiva en la publicación de los resultados inhibe el autoarchivo de la producción científica de los investigadores en repositorios en abierto.
Zhu (2019)	Multidisciplinar. Investigadores, profesores y becarios (N=1.800)	Explorar el apoyo y herramientas disponibles para facilitar la puesta en común de datos de investigación y qué factores afectan a la práctica de compartir datos.	<ul style="list-style-type: none"> • La mayoría de los académicos reconocen la importancia de compartir datos de investigación, pero en general nunca los habían compartido o reutilizado. • Existen diferencias en el comportamiento respecto al intercambio de datos en función del género, edad, disciplina y posición académica. • El conocimiento de la política de acceso abierto (OA), la experiencia de la publicación en OA, las actitudes hacia la importancia del intercambio de datos y la experiencia en el uso de datos secundarios se asociaron con la práctica del intercambio de datos. • Un pequeño grupo de investigadores utilizó las redes sociales como <i>Twitter</i>, blogs y <i>Facebook</i> para promover los datos de investigación que habían compartido online.

2. Motivación, barreras e incentivos para compartir los datos de investigación

2.1. Motivación para compartir y reutilizar los datos de investigación

La importancia que los datos abiertos tienen para la comunidad científica es ampliamente reconocida. Los investigadores valoran muy positivamente su influencia para hacer avanzar la investigación o evitar duplicar esfuerzos, así como el hecho de que los datos abiertos puedan respaldar algunas aplicaciones a los problemas sociales y permitir a los ciudadanos estar mejor informados (**Schmidt; Gemeinholzer; Treloar**, 2015; **Kim; Stanton**, 2016; **Berghmans et al.**, 2017; **Socrata**, 2016). A pesar de ello, los hábitos de los investigadores todavía precisan de cambios e incentivos para hacer que sus datos estén disponibles en abierto (**Fecher et al.**, 2017). Así, por ejemplo, a pesar de que desde 2005 en el Reino Unido existen políticas de acceso abierto institucionales y de agencias financiadoras de proyectos de investigación, en un trabajo publicado recientemente (**Zhu**, 2019) el 86% de los encuestados coincidían en resaltar la importancia del acceso abierto a las publicaciones y datos de investigación. Sin embargo, tan sólo el 21% de ellos había tenido alguna experiencia depositando sus datos en un repositorio y únicamente el 29% se declaraba usuario de los datos que otros han generado.

Cabe destacar que la difusión y el reconocimiento de los datos de investigación como productos generados de la investigación y equivalentes a otros, como las publicaciones, contribuiría a que los investigadores participaran más en hacer sus datos FAIR (**Schmidt; Gemeinholzer; Treloar**, 2015). Si el beneficio percibido es bajo o nulo, los investigadores no encuentran motivos para compartir sus datos (**Kim; Stanton**, 2016). La citación es un factor importante, y podría ser la principal razón por la que pondrían sus datos a disposición de otros (**Fecher et al.**, 2017). Existe una mayor inclinación a compartir datos cuando se trata de los datos recogidos por los propios autores, y no obtenidos de otra fuente. También se da una mayor predisposición cuando ya se han compartido datos en alguna ocasión (**Zenk-Möltgen et al.**, 2018).

Respecto al apoyo institucional y el papel destacado que juega a la hora de adoptar prácticas de ciencia abierta, **Sayogo y Pardo** (2013) analizaron con más detalle los factores que influían a la hora de compartir en abierto los datos de investigación. Entre sus conclusiones destacan que el apoyo de las instituciones afecta directamente tanto el proceso de gestión de los datos como al reconocimiento por el trabajo añadido al compartir los datos, y por otro lado los investigadores también revelaron cierta suspicacia sobre el uso de licencias. El apoyo institucional requiere de una infraestructura capaz de ajustarse a las características de la disciplina, de los recursos humanos, tecnológicos y de estándares, y de financiación que inciden de forma directa en la producción científica que se comparte en abierto (**Schöpfel et al.**, 2018; **Kim; Schuler; Pechenina**, 2017). Paradójicamente, se dan casos en los que se cuenta con estos servicios de apoyo institucional y, sin embargo, hay un porcentaje bajo de investigadores que los emplean (**Johnson; Chiarelli; Parsons**, 2016).

Las políticas que abogan por el acceso a las publicaciones y a los datos de investigación son motores destacados que invitan o requieren el compartir en abierto. Las políticas pueden ser de organismos financiadores, como algunos de los casos ya mencionados, de sociedades científicas o profesionales, o de las propias instituciones. También juegan un papel relevante las políticas editoriales de las revistas científicas (**Schmidt; Gemeinholzer; Treloar**, 2015).

Respecto a las prácticas de reutilización de datos, tanto la eficiencia como la eficacia percibida, como la importancia que tiene para el avance de la disciplina en particular, son aspectos que incentivan su práctica (**Gonçalves-Curty et al.**, 2017). Sin embargo, debido a los problemas de la multitud de formatos, a la propia generación de los datos de diversas disciplinas o a la falta de comprensión de temas de propiedad intelectual, no contribuyen a su implementación (**Wallis; Rolando; Borgman**, 2013).

El comportamiento de los investigadores respecto al intercambio y el uso de datos también varía según la disciplina. De acuerdo con uno de los primeros estudios llevado a cabo en ese sentido (Tenopir *et al.*, 2011), los investigadores en ciencias atmosféricas y ciencias ambientales y ecología manifestaron en un 87% estar dispuestos a compartir los datos. La mayoría de los autores de otras disciplinas presentes en el estudio, también estaban dispuestos a compartir, aunque con diferencias importantes; en particular, los menos dispuestos fueron los encuestados de informática y medicina. En biología el descubrimiento de los datos es fundamental y es también en esta área donde se registra el porcentaje más alto de datos compartidos (Stuart *et al.*, 2018).

La motivación para hacer que los datos de investigación sean FAIR depende de la disciplina a la que pertenecen los investigadores (Kim; Stanton, 2016). Los investigadores de ingeniería e informática consideraban muy importante para el avance de su área adherirse a estas prácticas; los de ciencias de la tierra y economistas las secundan de forma más discreta (Schmidt; Gemeinholzer; Treloar, 2015). Al igual que en otros casos, los investigadores de ciencias sociales y ciencias políticas se ven influenciados por sus colegas, de forma que es probable que la compartición de datos se convierta en práctica habitual si se formalizan canales en la comunidad científica para hacerlo (Zenk-Möltgen *et al.*, 2018). En las ciencias sociales el impulso en la carrera profesional que supone la publicación en abierto junto con los requisitos institucionales o normativos actúan como principales factores motivadores (Kim; Adler, 2015).

Pese a que las políticas de datos abiertos son consideradas en general como un motor importante, hay diferencias en cuanto a su grado de influencia en función de quién las dicta. Para las ciencias sociales, son determinantes si provienen de organismos financiadores o de las autoridades institucionales; esto mismo ocurriría para las ciencias químicas. Sin embargo, en las áreas de economía, las políticas emitidas por una sociedad científica se imponen a las de las entidades financiadoras, aunque ambas tienen una importancia muy destacable (Schmidt; Gemeinholzer; Treloar, 2015). Estos mismos autores estudiaron el efecto de las políticas editoriales sobre la motivación de los autores. En general, por disciplinas se muestran unánimes en considerarlas motores menos importantes que las mencionadas con anterioridad.

En algunas disciplinas, particularmente en las que se centra la *big science* (física, astronomía, genética), se genera gran cantidad de datos de investigación que pueden estimular su reutilización y aprovechamiento para la generación de nuevo conocimiento partiendo de datos ya existentes o analizar aspectos especialmente complejos. En áreas más especializadas, los datos generados se caracterizan también por una fuerte especificidad, su reutilización es mucho más limitada y necesitan una descripción exhaustiva y precisa para ser interpretados correctamente, por lo que la eficacia de la reutilización puede verse comprometida (Wallis; Rolando; Borgman, 2013).

Otro de los factores que pueden influir es la intención de obtener un distintivo (*badge*) como forma de visibilizar y acreditar una comunicación transparente y abierta de los resultados de investigación. En este sentido existen ejemplos como el de la editorial *Springer Nature* (Pearce, 2018), que en 2018 puso en marcha el *open data badge* para incentivar y recompensar a sus autores por compartir los datos. El *Center for Open Science* facilita una lista de revistas que han adoptado prácticas de la ciencia abierta, entre ellas la de los datos abiertos, y cuyos artículos se distinguen por iconos distintivos de su apertura (figura 2).

El uso de estos iconos parece haber contribuido al aumento del porcentaje de datos disponibles a partir de la fecha de su implantación (figura 3).



Figura 2. Iconos acreditativos de la disponibilidad de los datos de investigación. Fuente: *Center for Open Science* <https://cos.io/our-services/open-science-badges>

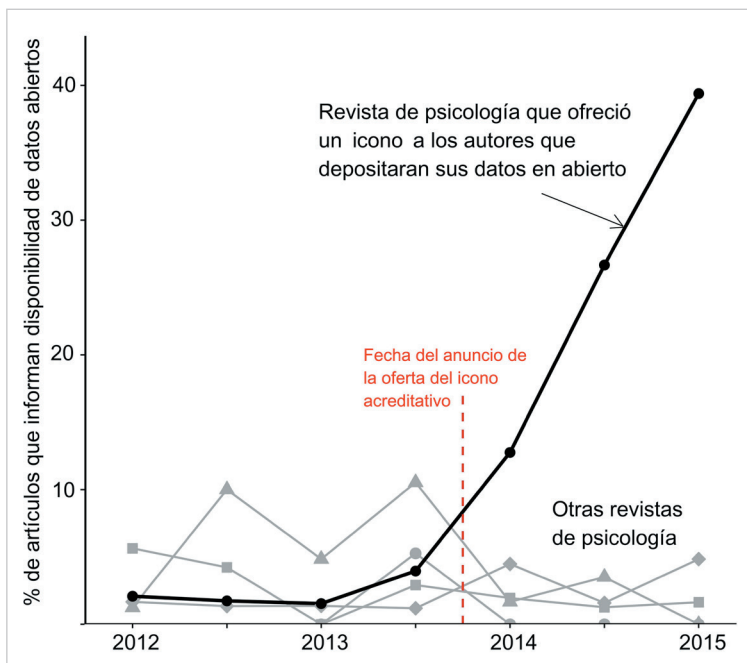


Figura 3. Efecto de la oferta de iconos acreditativos de la disponibilidad de los datos de investigación (Kidwell *et al.*, 2016)

2.2. Barreras para compartir y reutilizar los datos de investigación

Dos de los factores que actúan como barreras para la publicación de datos en abierto y para su reutilización, son:

- el riesgo de un mal uso de los datos,
- el esfuerzo de hacer que los datos sean FAIR.

El temor a que otros investigadores se adelanten en la publicación o a que se comporten de forma poco ética, hace que los investigadores opten por retrasar su compartición hasta que ellos mismos hayan agotado todas las posibilidades de publicar los resultados de la investigación (Kim, 2016).

Existe cierta inquietud ante la posibilidad de que los datos puedan ser malinterpretados (Fecher *et al.*, 2017; Tenopir *et al.*, 2011; Schmidt; Gemeinholzer; Treloar, 2015), y esta preocupación ha aumentado con el tiempo debido fundamentalmente a la mala calidad de los datos (Tenopir *et al.*, 2015). La inversión en tiempo y esfuerzo que es necesario realizar para preparar y compartir los datos de acuerdo con unos estándares de calidad aceptables afecta negativamente al intercambio de datos (Tenopir *et al.*, 2015; Kim; Stanton, 2016). Sólo en algunas áreas los científicos ponen sus datos fácilmente accesibles de forma estandarizada, a pesar de que los argumentos que defienden el gran valor que tienen los datos son compartidos, en general, por todas las disciplinas.

“ Dos de los factores que actúan como barreras para la publicación de datos en abierto y para su reutilización, son el riesgo de un mal uso de los datos y el esfuerzo de hacer que los datos sean FAIR ”

La calidad de los datos puede ser un factor que afecte negativamente a su reutilización y suponer una barrera para estimular su consumo. Los investigadores manifiestan sus reservas acerca de los diversos grados de calidad en los diferentes conjuntos de datos. Así, la diferencia de estándares sobre el modo de recopilar los datos, la diversidad de formatos, las dificultades para comprender, interpretar y acceder a los datos son fuertes obstáculos (Schmidt; Gemeinholzer; Treloar, 2015) y frenan el potencial de los datos para ser reutilizados con otros fines (Wallis; Rolando; Borgman, 2013). También son factores influyentes la creación y preservación de los metadatos, el apoyo facilitado por la institución, y la concienciación de la importancia de los datos (Tenopir *et al.*, 2011).

En noviembre de 2018 se celebró el *Top 10 FAIR data things global sprint* organizado por la *Carpentry Library* junto con otras instituciones. Como resultado del evento se publicó en febrero de 2019 el informe *Top 10 FAIR data software things*, cuyo contenido está orientado a facilitar “las 10 cosas” vinculadas con los datos de investigación (licencias, repositorios, metadatos, formatos, revistas, citación, identificadores, etc.) orientadas a diferentes disciplinas, desde la oceanografía a la arqueología. Estas “10 cosas” para cada una de las materias nos dan ejemplo de las diferencias y peculiaridades de cada una de ellas respecto de la tipología de formatos, metadatos, repositorios, vocabularios controlados, etc., que sirven para orientar a los investigadores sobre qué hacer para que sus datos cumplan los principios FAIR, y por lo tanto a vencer las barreras que lo obstaculice (Erdmann *et al.*, 2019).

En el medio académico, la citación juega un papel fundamental en las publicaciones. Sin embargo, en el caso de los datos todavía es poco frecuente citar los ficheros de datos como material de referencia. Por ejemplo, sólo el 17% dice hacerlo en trabajos de ciencias sociales (Kratz; Strasser, 2015). Entre las dificultades señaladas para no hacerlo se encuentran la falta de estandarización para citar los datos publicados, o la poca claridad de cómo hacerlo. Otros motivos también son la falta de apoyo institucional para costear la administración y almacenamiento de los datos. Precisamente tener que pagar por el acceso o por la reutilización de los datos es otro elemento desincentivador (Schmidt; Gemeinholzer; Treloar, 2015). En algunas disciplinas como en el “mapeo de suelos”, las tarifas varían en función del país de origen, lo que en ocasiones conlleva a malentendidos, a los que hay que sumar las diferencias en la legislación nacionales y regionales sobre licencias de uso y privacidad de los datos (Berghmans *et al.*, 2017).

“ En el caso de los datos todavía es poco frecuente citar los ficheros de datos como material de referencia ”

Globalmente, la preparación de los datos para que puedan ser interpretados por otros, los derechos de autor y licencias, el tiempo que se requiere para depositar los datos y los costes, son las dificultades más destacadas a la hora de compartir los ficheros que contienen los datos (Stuart *et al.*, 2018). Según este mismo autor, el tiempo y el esfuerzo que requieren la preparación de los datos son la barrera más importante para los investigadores senior respecto a los investigadores noveles. Para estos últimos influye más el hecho de no saber dónde compartir los datos. El desconocimiento sobre los derechos de autor y las diferencias entre las licencias de uso son una dificultad añadida tanto para los jóvenes investigadores como con aquellos de mayor experiencia (Stuart *et al.*, 2018).

Si estudiamos las principales barreras a nivel de disciplina, hay que señalar que, aunque el deseo por publicar los artículos antes de compartir los datos es una barrera muy importante para todas las disciplinas, se manifiesta con más frecuencia en las ciencias químicas y menos en ingeniería. En las áreas de economía e ingeniería las restricciones legales son los principales obstáculos para disponer los datos en abierto, cuya percepción es compartida por los investigadores de ciencias sociales (Schmidt; Gemeinholzer; Treloar, 2015).

Existe una brecha entre las recomendaciones o políticas respecto al acceso a los datos de investigación y su reutilización, y la práctica de los investigadores. Entre las dificultades encontradas para su implementación, están la falta de claridad en las políticas de datos y factores individuales como el desconocimiento del procedimiento para compartir datos y de los recursos de que disponen para auxiliarles en la puesta en práctica de forma ágil, fácil y práctica (**Zenk-Möltgen et al., 2018**).

En el área de la psicología, compartir datos de investigación de forma abierta no es una práctica común, a pesar de que se haga de facto entre colegas. Como en otros casos referenciados, también se observa la diferencia entre la actitud y la puesta en práctica de los investigadores. Por un lado, compartir datos es algo muy recomendable y de lo que se puede obtener un alto beneficio, y, sin embargo, no es una práctica que los investigadores tengan previsto aplicar en su investigación actual (**Houtkoop et al., 2018**). El esfuerzo de hacer sus datos FAIR es percibido como la mayor barrera para este colectivo, sin embargo, podría atenuarse por la existencia de una normativa clara relativa a la reutilización e intercambio de datos, junto con la práctica de colegas del área que predispone a adoptar comportamientos similares.

Para los investigadores de ciencias físicas el principal obstáculo es organizar los datos, mientras que los derechos de autor y las licencias de uso son la mayor dificultad para las ciencias médicas, seguida de no saber qué repositorio utilizar (**Stuart et al., 2018**). Las áreas en las que las prácticas de compartir datos están más extendidas, tienen resueltos estos problemas de forma más eficaz y la práctica de publicar en abierto se intensifica frente a aquellas en las que no lo están y compartir se restringe a la confianza que se tenga en el receptor de los datos, por lo que su práctica es mucho más limitada.

Los derechos de la propiedad intelectual son también importantes barreras en el campo de las humanidades digitales que han de abordar la cuestión de la concesión de licencias de textos literarios (**Berghmans et al., 2017**).

2.3. Condiciones para estimular e incentivar la reutilización y compartición de los datos de investigación

Los resultados de gran parte de las encuestas consultadas han puesto de manifiesto que la baja citación de los *datasets* es un factor que afecta directamente a las prácticas de compartir y reutilizar los datos (**Kratz; Strasser, 2015; Kim; Stanton, 2016; Van-den-Eynden; Bishop, 2014**) y que recibir una cita formal es esencial para estimular a los investigadores a que compartan sus datos (**Tenopir et al., 2015; Kratz; Strasser, 2015**). En una encuesta interdisciplinaria realizada en 2011, el 95% de los encuestados estuvo de acuerdo en que la cita formal es una condición justa para el intercambio de datos (**Tenopir et al., 2011**). Esta preocupación por la citación viene influenciada por la cultura del impacto a través de las citas, base para el reconocimiento y recompensa de los actuales sistemas de evaluación científica. En ese sentido quizás debería establecerse el *sharing factor*, como sugieren **Pampel y Dallmeier (2014)**, para reconocer el valor de compartir la información y el conocimiento en beneficio de la sociedad en general. Otro modelo novedoso de acreditar el valor de los datos, que no está basado en las citas o descargas, es el propuesto **Missier (2016)**, basado la trayectoria del *dataset* en función de cómo, quien y en que contextos se han reutilizado. Con esta información se construye la *data trajectory* que contiene toda esa información, para ello hay que tener en cuenta que el *dataset* se comporta como un objeto digital encapsulado con formato ORE y dispone de un ID persistente, por ejemplo, un DOI. Esta trayectoria es una gráfica donde se representan entidades (fichero de datos), actividades y actores (humanos, máquinas.).

Otro conjunto de requisitos son los relativos a aspectos relacionados con el control del acceso: quién puede acceder y usar los datos, con qué propósito, cuándo estarán disponibles los datos, así como la confidencialidad y la posibilidad de limitar el intercambio de datos a un grupo determinado cuando se trata de entornos de investigación altamente competitivos (**Van-den-Eynden; Bishop, 2014**). Pese a que el intercambio de datos está muy generalizado entre los investigadores, se practica de forma preferente entre colegas que desarrollan un trabajo conjunto y de forma minoritaria en abierto (**Fecher et al., 2017**). Estas diferencias también se establecen entre disciplinas, según este autor, tan sólo el 9% de los científicos de ciencias naturales dice no haber compartido nunca datos frente al 28% de investigadores de áreas sociales.

Los científicos valoran también la importancia de contar con una infraestructura adecuada para compartir los datos que no suponga un incremento de tiempo y esfuerzo (**Van-den-Eynden; Bishop, 2014**), así como contar con instrumentos que faciliten el acceso, entre ellos los repositorios, que puedan facilitar el intercambio de datos (**Kim; Stanton, 2016**).

Otros aspectos importantes relacionados con las condiciones para sumarse al intercambio de datos son los relativos a cuándo compartir (preferentemente después de la publicación) y a las contraprestaciones que se obtienen por hacer disponibles los datos: formalización de colaboraciones o coautorías en revistas de alto impacto y el intercambio mutuo de los datos (**Damalas et al., 2018**).

Las motivaciones personales, beneficios y riesgos percibidos descritos anteriormente, puede equilibrarse mutuamente, pero la mayor influencia de los riesgos, puede contribuir negativamente de forma directa a la motivación por compartir y reutilizar los datos de investigación, a pesar de su valor intrínseco (**Harper; Kim, 2018**).

3. Consideraciones finales

Como se ha visto en los apartados anteriores todos los temas relacionados con los datos de investigación, su gestión, compartición y reutilización afectan a todas las partes implicadas en la investigación científica, desde quién la financia,

quién la produce, quién la gestiona y quién la difunde. Sin embargo, al contrario que ocurre con las publicaciones académicas cuya existencia supera los tres siglos y son el medio por excelencia para la comunicación científica, en general los investigadores no tienen una cultura de hacer públicos los datos de investigación para su potencial reutilización. Además, el comportamiento del investigador respecto a los datos no es uniforme, se comportan de distinta manera en función de la disciplina, teniendo en cuenta que esta información es totalmente empírica y está sujeta a las respuestas de encuestas o entrevistas personales.

“ Todos los temas relacionados con los datos de investigación, su gestión, compartición y reutilización afectan a todas las partes implicadas en la investigación científica, desde quién la financia, quién la produce, quién la gestiona y quién la difunde ”

Las entidades privadas y públicas que financian la investigación cada vez más incluyen cláusulas en sus convocatorias que aluden a que los datos generados de la investigación sean datos FAIR. Otro reto será la trasposición de la *Directiva Europea* relativa a los datos abiertos y a la reutilización de la información del sector público, en la que se incluyen los datos de investigación. Estos requisitos harán que en un futuro próximo los investigadores deban cambiar de hábitos respecto a cómo gestionar sus datos. Esto implica vencer esas barreras a perder el control del uso de sus datos, miedo a perder la primicia o el miedo a la copia no reconocida. En síntesis, para evitar esas reticencias, el trabajo de apoyo debe ser de las propias instituciones, facilitando la tecnología, la formación necesaria, y la creación de nuevos perfiles académicos adaptados al análisis y gestión de los datos de investigación: el científico de datos. Existen precedentes en el establecimiento de cuáles serían las competencias para este nuevo profesional, como las señaladas en un informe elaborado por el grupo de trabajo de la Comisión Europea *Working Group on Education and Skills under Open Science* de la *Comisión Europea (European Commission, 2017a)*, entre las que se señalan, aspectos legales, competencias tecnológicas, protección de datos, comunicación y difusión de los resultados.

Otro aspecto no menos importante es el reconocimiento y la recompensa al investigador que contribuye a la ciencia abierta, en la cual se enmarcan los datos abiertos de investigación. En ese sentido existe ya un modelo publicado por el grupo europeo de expertos *Rewards under Open Science* en el que se tienen en cuenta nuevas variables para la evaluación de los investigadores, entre ellas el acceso abierto a las publicaciones científicas y la contribución a que los datos de investigación sean FAIR (*European Commission, 2017b*).

Cómo se generan los datos, su calidad y la de los metadatos que los describen, los planes de gestión de datos, las licencias de uso, entre otros temas, serán aspectos que se deberán tener en cuenta en un futuro próximo en los currículos de formación técnica y científica.

4. Referencias

ANDS (2017). *Data policies and journals*. Australian National Data Service.

<https://www.ands.org.au/working-with-data/publishing-and-reusing-data/data-journals/data-policies-and-journals>

Berghmans, Stéphane; Cousijn, Helena; Deakin, Gemma; Meijer, Ingeborg; Mulligan, Adrian; Plume, Andrew; De-Rijcke, Sarah; Rushforth, Alex; Tatum, Clifford; Van-Leeuwen, Thed; Waltman, Ludo (2017). *Open data: the researcher perspective*.

<https://www.universiteitleiden.nl/en/research/research-output/social-and-behavioural-sciences/open-data-the-researcher-perspective>

Bill & Melinda Gates Foundation (2015). *Bill & Melinda Gates Foundation open access policy*.

<https://www.gatesfoundation.org/how-we-work/general-information/open-access-policy>

Candela, Leonardo; Castelli, Donatella; Manghi, Paolo; Tani, Alice (2015). “Data journals: A survey”. *Journal of the Association for Information Science and Technology*, v. 66, n. 9, pp. 1747-1762.

<https://doi.org/10.1002/asi.23358>

Costello, Mark J. (2009). “Motivating online publication of data”. *BioScience*, v. 59, n. 5, pp. 418-427.

<https://doi.org/10.1525/bio.2009.59.5.9>

Damalas, Dimitrios; Kalyvioti, Grigoria; Sabatella, Evelina; Stergiou, Konstantinos I. (2018). “Open data in the life sciences: The ‘Selfish scientist paradox’”. *Ethics in science and environmental politics*, n. 18, pp. 27-36.

<https://doi.org/10.3354/esepp00182>

Eckes, Annemarie (2018). *Results from survey on: Assessment of awareness of FAIR principles and data management practices for early career scientists (PhDs) in geography*. Dataset.

<https://doi.org/10.17863/CAM.18831>

- ERC (2017). *Open research data and data management plans - Information for ERC grantees by the ERC Scientific Council*. European Research Council.
https://erc.europa.eu/sites/default/files/document/file/ERC_info_document-Open_Research_Data_and_Data_Management_Plans.pdf
- Erdmann, Christopher; Simmons, Natasha; Otsuji, Reid; Labou, Stephani; Johnson, Ryan; Castelao, Guilherme; Vilas-Boas, Bia; Lamprecht, Anna-Lena; Martínez-Ortiz, Carlos; García, Leyla; Kuzak, Mateusz; Martínez, Paula-Andrea; Stokes, Liz; Honeyman, Tom; Wise, Sharyn; Quan, Josh; Peterson, Scott; Neeser, Amy; Karvovskaya, Lena; Lange, Otto; Witkowska, Iza; Flores, Jacques; Bradley, Fiona; Hettne, Kristina; Verhaar, Peter; Companjen, Ben; Sesink, Laurents; Schoots, Fieke; Schultes, Erik; Kaliyaperumal, Rajaram; Tóth-Czifra, Erzsébet; De-Miranda-Azevedo, Ricardo; Muurling, Sanne; Brown, John; Chan, Janice; Quigley, Niamh; Federer, Lisa; Joubert, Douglas; Dillman, Allissa; Wilkins, Kenneth; Chandramouliswaran, Ishwar; Navale, Vivek; Wrihgt, Susan; Di-Giorgio, Silvia; Fasemore, Mandela; Förstner, Konrad; Sauerwein, Till; Seidlmayer, Eva; Zietlin, Ilja; Bacon, Susannah; Hannan, Katie; Ferrers, Richard; Russell, Keith; Whitmore, Deidre; Tim Dennis (2019). *Top 10 FAIR data & software things*. Zenodo.
<https://doi.org/10.5281/zenodo.2555498>
- European Commission (2016a). *Guidelines on FAIR data management in Horizon 2020*. European Commission, Directorate General for Research & Innovation.
http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf
- European Commission (2016b). *Draft European open science agenda*. European Commission, Directorate General for Research & Innovation.
https://ec.europa.eu/research/openscience/pdf/draft_european_open_science_agenda.pdf#view=fit&pagemode=none
- European Commission (2017a). *Providing researchers with the skills and competencies they need to practise open science*. European Commission, Directorate General for Research & Innovation.
<https://doi.org/10.2777/121253>
- European Commission (2017b). *Evaluation of research careers fully acknowledging open science practices. Rewards, incentives and/or recognition for researchers practicing Open Science*. European Commission, Directorate General for Research & Innovation.
<https://doi.org/10.2777/75255>
- European Commission (2018a). *Open Science Policy Platform*.
<https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-policy-platform#>
- European Commission (2018b). *Turning FAIR into reality. Final report and action plan from the European Commission expert group on FAIR data*. European Commission, Directorate General for Research & Innovation.
<https://doi.org/10.2777/1524>
- Fecher, Benedikt; Friesike, Sascha; Hebing Marcel; Linek, Stephanie (2017). "A reputation economy: How individual reward considerations trump systemic arguments for open access to data". *Palgrave communications*, n. 3, pp. 17051.
<https://doi.org/10.1057/palcomms.2017.51>
- Fusi, Federica; Manzella, Daniele; Louafi, Selim; Welch, Eric (2018). "Building global genomics initiatives and enabling data sharing: Insights from multiple case studies". *OmicS: A journal of integrative biology*, v. 22, n. 4.
<https://doi.org/10.1089/omi.2017.0214>
- García-García, Alicia; López-Borrull, Alexandre; Peset, Fernanda (2015). "Data journals: eclosión de nuevas revistas especializadas en datos". *El profesional de la información*, v. 24, n. 6, pp. 845-854.
<https://doi.org/10.3145/epi.2015.nov.17>
- Gonçalves-Curty, Renata; Crowston, Kevin; Specht, Alison; Grant, Bruce W.; Dalton, Elizabeth D. (2017). "Attitudes and norms affecting scientists' data reuse". *PLoS one*, v. 12, n. 12, e0189288.
<https://doi.org/10.1371/journal.pone.0189288>
- Harper, Lindsey M.; Kim, Youngseek (2018). "Attitudinal, normative, and resource factors affecting psychologists' intentions to adopt an open data badge: An empirical analysis". *International journal of information management*, v. 41, pp. 23-32.
<https://doi.org/10.1016/j.ijinfomgt.2018.03.001>
- Herold, Philip (2015). "Data sharing among ecology, evolution, and natural resources scientists: An analysis of selected publications". *Journal of librarianship & scholarly communication*, v. 3, n. 2, pp. 1-23.
<https://doi.org/10.7710/2162-3309.1244>

- Houtkoop, Bobby L.; Chambers, Chris; Macleod, Malcolm; Bishop, Dorothy V. M.; Nichols, Thomas E.; Wagenmakers, Eric-Jan** (2018). "Data sharing in psychology: A survey on barriers and preconditions". *Advances in methods and practices in psychological science*, v. 1, n. 1, pp. 70-85.
<https://doi.org/10.1177/2515245917751886>
- Hrynaszkiewicz, Iain; Simons, Natasha; Hussain, Azhar; Goudie, Simon** (2019). *Developing a research data policy framework for all journals and publishers*. Figshare. Preprint.
<https://doi.org/10.6084/m9.figshare.8223365.v1>
- Johnson, Rob; Chiarelli, Andrea; Parsons, Tom** (2016). *Data asset framework (DAF) survey results 2016*. Figshare.
https://figshare.com/articles/Data_asset_framework_DAF_survey_results_2016/3796305
- Kidwell, Mallory C.; Lazarević, Ljiljana B.; Baranski, Erica; Hardwicke, Tom E.; Piechowski, Sarah; Falkenberg, Lina-Sophia; Kennett, Curtis; Slowik, Agnieszka; Sonnleitner, Carina; Hess-Holden, Chelsey; Errington, Timothy M.; Fiedler, Susann; Nosek, Brian A.** (2016). "Badges to acknowledge open practices: A simple, low-cost, effective method for increasing transparency", *PLoS biology*, v. 14, n. 5, e1002456.
<https://doi.org/10.1371/journal.pbio.1002456>
- Kim, Youngseek; Adler, Melissa** (2015). "Social scientists' data sharing behaviors: Investigating the roles of individual motivations, institutional pressures, and data repositories". *International journal of information management*, v. 35, n. 4, pp. 408-418.
<https://doi.org/10.1016/j.ijinfomgt.2015.04.007>
- Kim, Youngseek; Oh, Jong-Sir** (2018). "Disciplinary, institutional, and individual factors affecting researchers' depositing articles in institutional repository: An empirical analysis". *The journal of academic librarianship*, v. 44, n. 6, pp. 824-832.
<https://doi.org/10.1016/j.acalib.2018.09.013>
- Kim, Jeonghyun; Schuler, Eric R.; Pechenina, Anna** (2017). "Predictors of data sharing and reuse behavior in academic communities". In: Alemneh, Daniel G.; Allen, Jeff; Hawamdeh, Suliman. *Knowledge discovery and data design innovation. Proceedings of the Intl conf on knowledge management (ICKM 2017)*, pp. 1-25. ISBN: 978 981 3234 47 5
https://doi.org/10.1142/9789813234482_0001
- Kim, Youngseek; Stanton, Jeffrey, M.** (2016). "Institutional and individual factors affecting scientists' data-sharing behaviors: A multilevel analysis". *Journal of the Association for Information Science and Technology*, v. 67, n. 4, pp. 776-799.
<https://doi.org/10.1002/asi.23424>
- Kratz, John E.; Strasser, Carly** (2015). "Making data count". *Scientific data*, n. 2, article n. 150039.
<https://doi.org/10.1038/sdata.2015.39>
- Michener, William K.** (2015). "Ecological data sharing". *Ecological informatics*, v. 29, pp. 33-44.
<https://doi.org/10.1016/j.ecoinf.2015.06.010>
- Missier, Paolo** (2016). "Data trajectories: Tracking reuse of published data for transitive credit attribution". *International journal of digital curation*, v. 11, n. 1.
<https://doi.org/10.2218/ijdc.v11i1.425>
- National Science Foundation (2011). *Increasing access to the results of research funded by the National Science Foundation*.
<http://www.nsf.gov/pubs/2015/nsf15052/nsf15052.pdf>
- Pampel, Heinz; Dallmeier-Tiessen, Suenje** (2014). "Open research data: From vision to practice". *Semantic scholar*.
https://doi.org/10.1007/978-3-319-00026-8_14
- Parlamento Europeo (2019). *Resolución legislativa, 4 de abril de 2019, sobre la Directiva del Parlamento Europeo y del Consejo relativa a la reutilización de la información del sector público (versión refundida)*. COM(2018)0234 – C8-0169/2018 – 2018/0111(COD).
http://www.europarl.europa.eu/doceo/document/TA-8-2019-0352_ES.html
- Pearce, Rebecca** (2018). "Springer Nature launches open data badges pilot". *Research in progress blog*, 8 October.
<https://blogs.biomedcentral.com/bmcblog/2018/10/08/springer-nature-launches-open-data-badges-pilot>
- Piwowar, Heather A.** (2011). "Who shares? Who doesn't? Factors associated with openly archiving raw research data". *PLoS one*, v. 6, n. 7, e18657.
<https://doi.org/10.1371/journal.pone.0018657>
- Piwowar, Heather A.; Chapman, Wendy W.** (2010). "Public sharing of research datasets: a pilot study of associations". *Journal of informetrics*, v. 4, n. 2, pp. 148-156.
<https://doi.org/10.1016/j.joi.2009.11.010>

- Sayogo, Djoko-Sigit; Pardo, Theresa A.** (2013). "Exploring the determinants of scientific data sharing: Understanding the motivation to publish research data". *Government information quarterly*, v. 30, supp. 1, S19-S31.
<https://doi.org/10.1016/j.giq.2012.06.011>
- Schmidt, Birgit; Gemeinholzer, Birgit; Treloar, Andrew** (2015). "Open data in global environmental research: The Belmont forum's open data survey". *PLoS one*, v. 11, n. 1, e0146695.
<https://doi.org/10.1371/journal.pone.0146695>
- Schöpfel, Joachim; Ferrant, Coline; André, Francis; Fabre, Renaud** (2018). "Research data management in the French National Research Center (CNRS)". *Data technologies and applications*, v. 52, n. 2, pp. 248-265.
<https://doi.org/10.1108/DTA-01-2017-0005>
- Socrata (2016). *Open data benchmark study report*.
<https://www.tylertech.com/products/socrata>
- Stodden, Victoria; Guo, Peixuan; Ma, Zhaokun** (2013). "Toward reproducible computational research: An empirical analysis of data and code policy adoption by journals". *PLoS one*, v. 8, n. 6, e67111.
<https://doi.org/10.1371/journal.pone.0067111>
- Stuart, David; Baynes, Grace; Hrynaszkiewicz, Iain; Allin, Katie; Penny, Dan; Lucraft, Mithu; Astell, Mathias** (2018). *Whitepaper: Practical challenges for researchers in data sharing*. Figshare.
<https://doi.org/10.6084/m9.figshare.5975011.v1>
- Tenopir, Carol; Allard, Suzie; Douglass, Kimberly; Aydinoglu, Arsev U.; Wu, Lei; Read, Eleanor; Manoff, Maribeth; Frame, Mike** (2011). "Data sharing by scientists: Practices and perceptions". *PLoS one*, v. 6, n. 6, e21101.
<https://doi.org/10.1371/journal.pone.0021101>
- Tenopir, Carol; Dalton, Elizabeth D.; Allard, Suzie; Frame, Mike; Pjesivac, Ivanka; Birch, Ben; Pollock, Danielle; Dorsett, Kristina** (2015). "Changes in data sharing and data reuse practices and perceptions among scientists worldwide". *PLoS one*, v. 10, n. 8, e0134826.
<https://doi.org/10.1371/journal.pone.0134826>
- Van-den-Eynden, Veerle; Bishop, Libby** (2014). *Sowing the seed: Incentives and motivations for sharing research data, a researcher's perspective. A Knowledge Exchange report*.
<http://www.knowledge-exchange.info/projects/project/research-data/sowing-the-seed>
- Van-Panhuis, Willem G.; Paul, Proma; Emerson, Claudia; Grefenstette, John; Wilder, Richard; Herbst, Abraham J.; Heymann, David; Burke, Donald S.** (2014). "A systematic review of barriers to data sharing in public health". *BMC public health*, n. 14, pp. 1144.
<https://doi.org/10.1186/1471-2458-14-1144>
- Wallis, Jillian C.; Rolando, Elizabeth; Borgman, Christine L.** (2013). "If we share data, will anyone use them? Data sharing and reuse in the long tail of science and technology". *PLoS one*, v. 8, n. 7, e67332.
<https://doi.org/10.1371/journal.pone.0067332>
- Wellcome Trust (2010). *Data, software and materials management and sharing policy*.
<https://wellcome.ac.uk/funding/guidance/policy-data-software-materials-management-and-sharing>
- Wiley, Chistie** (2018). "Data sharing and engineering faculty: An analysis of selected publications". *Science & technology libraries*, v. 37, n. 4, pp. 409-419.
<https://doi.org/10.1080/0194262X.2018.1516596>
- Zenk-Möltgen, Wolfgang; Akdeniz, Esra; Katsanidou, Alexia; Naßhoven, Verena; Balaban, Ebru** (2018). "Factors influencing the data sharing behavior of researchers in sociology and political science". *Journal of documentation*, v. 74, n. 5, pp. 1053-1073.
<https://doi.org/10.1108/JD-09-2017-0126>
- Zhu, Yimei** (2019). "Open-access policy and data-sharing practice in UK academia". *Journal of information science*. First online.
<https://doi.org/10.1177/0165551518823174>