

El potencial de los datos de investigación abiertos: repositorios y ciencia abierta para la investigación

Paola Carolina Bongiovani



**Serie Cátedra UNESCO en
Gestión de Información en las
Organizaciones**





El potencial de los datos de investigación abiertos: repositorios y ciencia abierta para la investigación es un texto elaborado por **Paola Carolina Bongiovani**, de la Universidad Nacional de Rosario, a partir de los resultados de su tesis doctoral. La autora es una de las grandes especialistas en la gestión de datos abiertos por parte de las instituciones de investigación y de educación superior de toda Latinoamérica, además de ser una referencia en la promoción del acceso abierto en toda la región. Este libro es el segundo de la serie **Cátedra UNESCO en Gestión de Información en las Organizaciones**, publicada en acceso abierto sin coste alguno para los autores.

Cómo citarlo:

Bongiovani, Paola Carolina. (2026). *El potencial de los datos de investigación abiertos: repositorios y ciencia abierta para la investigación*. Editum. Ediciones de la Universidad de Murcia.
<https://doi.org/10.6018/editum.3245>

La imagen de portada fue creada con asistencia de inteligencia artificial mediante ChatGPT (OpenAI), bajo la dirección y supervisión del editor de la serie. Se comparte en [Digitum](#) dentro de la subcolección “Cátedra UNESCO en Gestión de Información en las Organizaciones” bajo licencia <https://creativecommons.org/licenses/by/4.0/deed.es>

DOI: [10.618/editum/3245](https://doi.org/10.6018/editum/3245)

Si el lector detecta algún error en el libro o bien quiere contactar con el autor, puede enviar un correo a publicaciones@um.es



Bongiovani, Paola Carolina. (2026). *El potencial de los datos de investigación abiertos: repositorios y ciencia abierta para la investigación*. Editum. Ediciones de la Universidad de Murcia. <https://doi.org/10.6018/editum.3245>

Se permite la reutilización y redistribución del contenido según los términos de la licencia [Creative Commons Atribución \(CC BY\)](#).

Tabla de contenido

0. Prólogo	5
1. Perspectivas diversas de la ciencia abierta.....	9
1.1 El contexto de la ciencia abierta	9
1.2 Definiciones de la ciencia abierta.....	11
1.3 Responsabilidad en la investigación e innovación	19
1.4 Valores de Merton y Ciencia abierta.....	20
1.5 Preocupación por la equidad e inclusión en la ciencia abierta	26
1.5.1 Perspectiva constructivista.....	28
1.5.2 Perspectiva de la ciencia abierta desde la filosofía de la ciencia	29
1.6 El acceso abierto.....	31
1.7 Propuestas de cambios en el sistema de evaluación científica.....	38
1.8 El acceso abierto en Latinoamérica.....	42
1.8.1 Conocimiento del acceso abierto en Argentina, México y Brasil.....	44
1.8.2 El acceso abierto en Argentina	47
1.8.3 Los investigadores argentinos y el acceso abierto.....	52
1.8.4 El acceso abierto en la Universidad Nacional de Rosario. Repositorio de publicaciones.....	55
1.8.5 Portal Revistas UNR.....	58
1.8.6 Investigaciones sobre acceso abierto en la UNR.....	60
1.9 Los datos de investigación.....	61
1.9.1 Tipos de datos.....	63
2 Compartir datos de investigación.....	67
2.1 El contexto internacional	67
2.2 El contexto argentino	71
2.3 Compartir datos de investigación	71
2.4 Políticas de las revistas	73
2.5 Revistas de datos ('data journals').....	77
2.6 La opinión de los investigadores: incentivos y barreras	77
2.7 Incentivos para compartir y usar datos de investigación	78
2.8 Barreras para compartir datos de investigación (sistemas de evaluación científica).....	87
2.9 Dimensiones y variables en relación con la literatura relevada	92
2.9.1 Aspectos éticos y legales.....	93
3. Los repositorios de datos	104
3.1 Definiciones y tipología.....	104
3.1.1 Tipos de repositorios de datos	105
3.1.2 Repositorios Digitales Confiables (TDR)	106
3.2 Curación de datos de investigación.....	108
3.3 Estándares y calidad de datos. FAIR.....	114
3.4 Replicación, reproducibilidad y reutilización	118
3.5 Necesidad de curadores de datos.....	122
3.5.1 ¿Qué es un curador de datos ('data steward')?.....	123
3.5.2 Curación de datos en bibliotecas y repositorios	125
3.5.3 Capacitación de investigadores.....	131
3.6 Valor de los repositorios datos.....	132
3.7 La autora	134
4. Referencias	135

Tabla de imágenes

Figura 1 Valores y principios para la ciencia abierta.....	10
Figura 2 Valores y principios del Manifiesto de Ciencia abierta y Colaborativa (OCSDNet).	13
Figura 3 Pilares de la ciencia abierta.	16
Figura 4 Hongo de la ciencia abierta.	19
Figura 5 Acceso abierto diamante en las revistas de la UNR.	32
Figura 6 Diagrama de Venn de los niveles de acceso abierto a las publicaciones en función del coste, retención de derechos de autor y revisión del contenido por pares.	33
Figura 7 Medida z-score ponderada y mapa normalizado de las publicaciones en acceso abierto por país.	34
Figura 8 Iniciativas de acceso abierto conocidas en los tres países.	44
Figura 9 Conocimientos y opiniones sobre revistas de acceso abierto asociadas al prestigio en Argentina, Brasil y México.	45
Figura 10 Factores más importantes en la evaluación de las publicaciones de sus pares en Argentina, Brasil y México.	46
Figura 11 Factores más importantes en la evaluación de las publicaciones de sus pares en Argentina, Brasil y México.	46
Figura 12 Factores más importantes en la evaluación de las publicaciones de sus pares en Argentina, Brasil y México.	47
Figura 13 Porcentaje de publicaciones por repositorios.	49
Figura 14 Porcentaje de publicaciones según tipo.	49
Figura 15 Evolución de registros en el cosechador SIUBDU de Argentina..	57
Figura 16 Evolución de cosechas desde el SNRD.	57
Figura 17 Vistas y usuarios al RepHipUNR en 2022.	58
Figura 18 Clasificación de necesidades para el plan de mejoramiento de las publicaciones periódicas).	59
Figura 19 Resoluciones de DOIs de revistas de la UNR 2022-2023..	60
Figura 20 Modelo de ciclo de vida de curación de Digital Curation Centre DCC.	112
Figura 21 Principios FAIR.	115
Figura 22 Reporte de la herramienta de evaluación de datos FAIR automatizada F-UJI.....	118
Figura 23 La autora.....	134

Lista de tablas

Tabla i Cinco Escuelas de pensamiento de la ciencia abierta.	13
Tabla ii Resumen de áreas identificadas de preocupación para la equidad en la ciencia abierta.	27
Tabla iii Resumen de los desafíos percibidos por las agencias de financiamiento y enfoques de solución.	69
Tabla iv Dimensiones y variables del estudio en relación con la literatura relevada.	92
Tabla v Principios TRUST.	106
Tabla vi Actividades de curación de datos y sus acciones correspondientes.	126

0. Prólogo

En un contexto en el que la investigación es cada vez más digital y basada en datos masivos, éstos (los datos) han dejado de ser un subproducto del proceso científico para convertirse en un activo fundamental. No solo constituyen evidencia empírica, sino que son también objeto de gestión, curación, regulación y reutilización. Por ello, es para mí un placer presentar este libro de la **Dra. Paola Bongiovanni**, que aborda una de las cuestiones más relevantes en la evolución contemporánea de la ciencia: el potencial de los datos abiertos en los procesos de producción, validación y comunicación del conocimiento. El libro se inserta claramente en el paradigma de la ciencia abierta, pero lo hace desde una posición razonada que evita tanto la simplificación conceptual como la adopción acrítica de los postulados clásicos de la ciencia abierta, lugares comunes en la teoría, pero difíciles de implementar en la práctica.

Este libro es el resultado de la tesis doctoral de la autora, una investigación cuidada y rigurosa, que además se adorna con una gran precisión conceptual. En lugar de utilizar de manera indistinta “datos abiertos”, “datos FAIR” o “compartición de datos”, la autora establece distinciones claras y operativas, fundamentales para avanzar tanto en la meta-investigación sobre datos, como en el trabajo práctico con ellos en repositorios e infraestructuras de investigación. Esta claridad no es un ejercicio meramente terminológico, sino una condición necesaria para evitar equívocos que, con frecuencia, distorsionan el alcance de la reutilización y la implementación de iniciativas en este ámbito.

Al trabajo riguroso y la claridad terminológica, se suma un enfoque crítico que constituye, a mi juicio, uno de los aportes más relevantes del trabajo. Lejos de presentar los datos abiertos como una solución fácil y universal, el texto examina las condiciones en las que la apertura es posible, así como las tensiones que la atraviesan. La relación entre apertura y protección de datos, la existencia de incentivos contradictorios en la carrera académica o las implicaciones de la reutilización de datos en contextos diversos, son temas que aborda con un nivel de precisión y matización constante, necesario para entender el alcance de los datos abiertos en la investigación contemporánea. Asimismo, la obra pone de relieve que la apertura de datos no es únicamente una cuestión técnica; implica además el

desarrollo de competencias específicas, la aparición de nuevos perfiles profesionales (los ‘data stewards’, a los que ella se refiere en ocasiones como “curadores de datos”), así como la necesidad de estrategias institucionales capaces de integrar la gestión de datos en el ciclo completo de la investigación. Esta dimensión formativa y organizativa a veces se subestima, pero aquí se aborda con la profundidad que merece, reflejando el espíritu del proyecto *Skills4EOSC*¹ en que participó el grupo de investigación OpenScienceLab², al que se incorporó la autora después de su tesis.

Para tratar la complejidad de estos temas, el texto se divide formalmente en tres capítulos, lo que facilita su lectura:

- El capítulo primero está dedicado a las perspectivas diversas de la Ciencia abierta. En el mismo, la autora ofrece un recorrido conceptual riguroso y pluralista por el paradigma de la ciencia abierta, desde sus principales definiciones y marcos normativos internacionales (ej. incluyendo la Recomendación UNESCO de 2021) hasta sus fundamentos filosóficos, sus conexiones con los valores mertonarios y los principios de la Investigación e Innovación Responsables (RRI). Este capítulo no es una mera presentación de la Ciencia abierta como fenómeno homogéneo, es un examen detallado de sus distintas escuelas de pensamiento, sus tensiones internas en torno a la equidad y la inclusión, y dimensión más concreta en el acceso abierto a publicaciones y datos, con especial atención al contexto latinoamericano y argentino.
- El segundo capítulo trata sobre cómo compartir los datos de investigación, adentrándose en las condiciones reales (institucionales, normativas, culturales y éticas) que determinan o condicionan que los investigadores compartan o no sus datos. A través del análisis de políticas de agencias financiadoras, revistas científicas y organismos internacionales, la autora dibuja un panorama en el que los incentivos y las barreras coexisten con una

¹ Skills4EOSC (Skills for the European Open Science Commons) es una iniciativa europea diseñada para abordar las principales carencias de competencias en ciencia abierta y gestión de datos mediante la creación de un ecosistema de formación cohesionado y de alta calidad en toda Europa.

² Grupo de Investigación de la Universidad Carlos III de Madrid dedicado a la metainvestigación para la ciencia abierta; evalúa prácticas científicas, diseña políticas y forma a investigadores para promover la transparencia, la calidad y la colaboración en la investigación científica. Más información en <https://opensciencelab.uc3m.es/>

tensión no resuelta: la distancia entre el mandato formal de apertura y la práctica cotidiana de la investigación, que se alejan profundamente, en muchos casos. Este capítulo refleja además cómo los aspectos éticos y legales, la privacidad, la confidencialidad y los sistemas de evaluación científica emergen como variables críticas que condicionan profundamente la disposición de los investigadores a hacer sus datos accesibles.

- Finalmente, en el tercer capítulo, dedicado a los repositorios de datos, se aborda la dimensión más técnica vinculada a las infraestructuras de la ciencia abierta, particularmente los repositorios de datos, como sistemas técnicos y organizacionales que hacen posible la gestión, curación y reutilización de los datos de investigación. **Paola Bongiovani** examina en este caso, de forma profusa pero clara y sistemática: las tipologías, los estándares de calidad (principios FAIR y TRUST), los procesos de curación de datos y el papel fundamental que bibliotecas y profesionales especializados desempeñan en este ecosistema. Este capítulo pone de relieve que la apertura de los datos no es solo una cuestión de voluntad, política o mandato, sino que requiere de infraestructuras sostenibles, competencias profesionales específicas, así como la formación continua y consistente de investigadores y gestores de datos.

Como codirectora de la tesis doctoral de la que este libro es fruto, tengo el privilegio de haber acompañado de cerca el proceso intelectual que le dio origen. Esta obra no es simplemente la versión editada de una investigación académica, sino el resultado de años de trabajo sostenido, de lecturas atentas y de una capacidad poco común para articular el trabajo práctico desde la reflexión crítica, fundamentada en su trabajo y en la creación del repositorio de datos de la [Universidad Nacional de Rosario](#) (UNR) que dirige la autora. Recuerdo las discusiones en una cafetería de esa ciudad que tiene en su interior una estatua de [Lionel Messi](#) y otra de [Ángel Di María](#). Frente a estos futbolistas rosarinos discutimos, además del contenido metodológico de la tesis, sobre cuestiones conceptuales delicadas que se reflejan en este libro ahora (cómo delimitar el objeto de estudio, cómo evitar la trampa de la descripción aséptica de políticas sin interrogar sus condiciones, etc.). Cada una de esas decisiones está presente, de manera sólida y coherente, en las páginas que siguen.

Que una tesis doctoral se convierta en un libro no es algo automático ni habitual. Requiere del autor una disposición a revisar, ampliar y reescribir el texto con una mirada renovada tras la defensa, orientándolo ya no a un tribunal académico, sino a una comunidad lectora más amplia. En este caso, ese ejercicio permite que distintos perfiles (investigadores, gestores de datos, bibliotecarios, responsables de políticas científicas, etc.) se acerquen al tema de la gestión de datos de investigación. **Paola Bongiovani** ha llevado a cabo esta transformación con la misma seriedad y generosidad intelectual que caracterizaron su trabajo durante la investigación doctoral, convirtiendo su tesis en un texto útil para la comunidad. El resultado es un texto que, estoy convencida, se convertirá en una referencia indispensable para quienes trabajen en el ámbito de los datos abiertos, los repositorios y la Ciencia abierta en el mundo hispanohablante.

En un ámbito donde abundan las declaraciones programáticas y los “brindis al sol” en pos de la Ciencia abierta y los datos abiertos, este libro aporta claridad, evidencia y una lectura honesta de las prácticas que bien conoce la autora. Se trata de una contribución relevante tanto para la comunidad académica como para quienes diseñan e implementan políticas científicas, al ofrecer herramientas conceptuales y analíticas para entender mejor qué significa, en la práctica, trabajar con datos (abiertos) de investigación.

Eva Méndez
Directora de [OpenScienceLab](#)
Universidad Carlos III de Madrid

1. Perspectivas diversas de la ciencia abierta

La investigación e innovación científica cumplen un rol fundamental para encontrar soluciones a los complejos desafíos globales, abarcando desde la atención médica en contextos complejos, como ha sido evidente en la pandemia COVID-19, el cambio climático, el flagelo del hambre y la pobreza, hasta las energías renovables y el uso sustentable de recursos naturales, entre los muchos retos que se nos presentan. Para afrontarlos, contamos hoy con una cantidad sin precedentes de datos, de avances tecnológicos que permiten no solo el almacenamiento y capacidad computacional de procesamiento de datos, sino también herramientas cada vez más potentes para analizar esos datos como la [Inteligencia Artificial](#).

La web y herramientas relacionadas permiten comunicaciones instantáneas entre científicos de todo el mundo, con el potencial de colaboración y distribución inmediata y abierta de los conocimientos. Se espera que la ciencia basada en datos aumente la velocidad de la investigación y su impacto con palpables beneficios para toda la sociedad. La ciencia abierta constituirá un pilar fundamental para afrontar los desafíos actuales a escala global. Este nuevo paradigma está cambiando el modo de hacer ciencia (**Anglada y Abadal, 2018**) implicando mayor transparencia, rigor, responsabilidad y la posibilidad de reproducir las investigaciones, mostrando todo el proceso de investigación, no solo los resultados finales sino también cómo se llegó a esos resultados. El eje principal es el impacto social que se quiere lograr con la ciencia abierta (**Méndez Rodríguez, 2021**).

1.1 El contexto de la ciencia abierta

La ciencia abierta es “conocimiento transparente y accesible que se comparte y desarrolla a través de redes colaborativas” (**Vicente-Sáez y Martínez-Fuentes, 2018**). Esta visión consiste en desarrollar la actividad científica de manera que otras personas puedan colaborar y realizar aportaciones, facilitando el acceso gratuito a los datos, las notas de laboratorio, los métodos y otros elementos del proceso de investigación, con licencias que permitan reutilizarlos, redistribuirlos y reproducirlos (**Bezjak et al., 2017; FOSTER, 2015**). Esta nueva forma de hacer ciencia se basa en principios de inclusión, imparcialidad, equidad y amplia difusión y acceso a todo el proceso de investigación. Su objetivo final es transformar cómo se lleva a cabo la labor investigadora, quiénes participan en ella y cómo se evalúa (esto es el enfoque de Investigación e Innovación Responsables o RRI). En noviembre de

2021 se aprobó la [*Recomendación sobre la ciencia abierta de UNESCO*](#), marco internacional apoyado por 193 países, entre ellos Argentina. En este texto se define la ciencia abierta como:

“... constructo inclusivo que combina diversos movimientos y prácticas con el fin de que los conocimientos científicos estén abiertamente accesibles para todos, se incrementen las colaboraciones científicas y el intercambio de información en beneficio de la ciencia y la sociedad, abriendo los procesos de creación, evaluación y comunicación de los resultados a la sociedad, 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” (UNESCO, 2021).

Los valores de la ciencia abierta la definen como un medio para garantizar la calidad, la integridad y la libertad académica mediante el uso de diversas fuentes de conocimiento y el intercambio de métodos y resultados, facilitando una evaluación rigurosa. Debe constituir un bien público mundial que beneficie a la humanidad y promueva la inclusión, la sostenibilidad y la equidad en el acceso a la ciencia. La equidad y la justicia favorecen la colaboración internacional, el intercambio recíproco y el acceso igualitario al conocimiento, independientemente de la nacionalidad, el género, los ingresos o las circunstancias sociales. Asimismo, la ciencia abierta integra la diversidad de conocimientos, prácticas, lenguas y temas de investigación, adaptándose a las necesidades de la comunidad científica, el público, las comunidades locales, los pueblos indígenas y otros agentes sociales.



FIGURA 1 Valores y principios para la ciencia abierta.

Con el objeto de generar condiciones y acciones para la puesta en práctica de estos valores, se enumeran los principios que la hacen posible: transparencia, control, crítica y reproducibilidad; igualdad de oportunidades; responsabilidad, respeto y rendición de cuentas; colaboración, participación e inclusión; flexibilidad para aceptar la diversidad y sostenibilidad a largo plazo.

Implementar estos valores y principios requiere alinear incentivos, esfuerzos y coordinación de políticas y cambios culturales con varios actores como universidades, agencias financiadoras, editoriales e investigadores: un entorno que favorezca que las prácticas de la ciencia abierta en su quehacer cotidiano. Entre los componentes de la ciencia abierta, se encuentran los datos de investigación abiertos, sobre lo que se centrará este libro.

1.2 Definiciones de la ciencia abierta

Los seres humanos enfrentan desafíos cada vez más complejos, como la pobreza, la inequidad, las pandemias y la degradación ambiental. La ciencia abierta ('open science' en inglés), se vislumbra como una herramienta cada vez más importante para abordar estos problemas, por su potencial para mejorar la investigación científica. Sin embargo, la financiación, realización, comunicación y evaluación de la investigación se mantienen casi sin cambios respecto al pasado. Las agencias de financiamiento, las instituciones y los investigadores continúan perpetuando un sistema de comunicación científica ineficiente, centrado exclusivamente en artículos que, para ser evaluados favorablemente, deben publicarse en revistas de alto [factor de impacto](#). Es sabido que fue desarrollado como una medida de la calidad de las revistas, pero se ha terminado adoptando erróneamente para medir la influencia de los autores individuales en su dominio de investigación. Se contabilizan estas publicaciones y las citas de estas. La forma tradicional de hacer ciencia ya no es efectiva ni suficiente para avanzar el conocimiento con calidad y transparencia (Méndez Rodríguez y Sánchez-Núñez, 2023).

La ciencia abierta propone mucho más que abrir los resultados, es decir, que los artículos, libros y tesis, entre otros resultados, estén disponibles en acceso abierto para todos. Consiste en transformar y abrir cómo se hace investigación, con quiénes, cómo se comparten datos y metodologías (procesos) y cómo ampliar la colaboración entre una amplia variedad de actores sociales, que garantice mayor transparencia, eficiencia, rendición de cuentas a quienes la financian y a la sociedad en su conjunto. Fecher y Friesike (2014) encuentran que la falta de claridad en el

concepto se relaciona con el hecho de que existen diferentes grupos interesados en el tema como investigadores, decisores de políticas públicas, programadores de plataformas, editores y el público en general, quienes tienen diversas perspectivas y entendimientos de lo que significa la ciencia abierta. Estos autores recogen en la **tabla I** lo que denominan “escuelas de pensamiento” de acuerdo con una revisión de la literatura y resumen con sus principales suposiciones y objetivos, así como los grupos involucrados y las herramientas que utilizan. Explican que algunas escuelas comparten ideas similares y se superponen:

- La escuela **democrática** hace hincapié en el acceso al conocimiento científico como un bien público y la necesidad de que el conocimiento financiado con fondos públicos se comparta con la sociedad.
- La escuela **pragmática** se centra en la colaboración en la investigación para mayor eficiencia, cuestión relacionada con el concepto de innovación abierta.
- La escuela **infraestructural**, enfocada en el desafío de implementar plataformas digitales que fomenten esos cambios en las prácticas científicas.
- La escuela **pública** se centra en la idea de que la ciencia sea accesible para todos los ciudadanos, promoviendo tanto la comprensión pública de la ciencia como la participación de los ciudadanos en ella.
- La escuela **de medición** se preocupa por encontrar métodos alternativos para evaluar a los investigadores y los resultados de la investigación.

Escuela	Suposiciones	Agentes de interés	Objetivo Principal	Herramientas y Métodos
Democrática	Acceso al conocimiento está distribuido de manera desigual.	Científicos, ciudadanos, políticos.	Hacer que el conocimiento esté abierto para todos.	Acceso abierto, propiedad intelectual, datos abiertos, código abierto
Pragmática	La creación de conocimiento podría ser más eficiente si los investigadores trabajaran juntos.	Científicos	Abrir el proceso de creación de conocimiento.	Sabiduría de las masas, datos abiertos, código abierto, networking.
Infraestructural	La investigación eficiente depende de las herramientas y aplicaciones disponibles	Científicos, Proveedores de plataformas	Crear plataformas, herramientas y servicios de acceso abierto para científicos	Plataformas y herramientas de colaboración

Escuela	Suposiciones	Agentes de interés	Objetivo Principal	Herramientas y Métodos
Pública	La ciencia debe ser accesible al público.	Científicos, Ciudadanos	Hacer que la ciencia sea accesible para los ciudadanos	Ciencia ciudadana, relaciones públicas científicas, blogs científicos
Medición	Actualmente las contribuciones científicas necesitan medidas de impacto alternativas	Científicos, Políticos	Desarrollar un sistema métrico alternativo para medir el impacto científico	Altmétricas, revisión por pares, citas, factores de impacto

TABLA 1 Cinco Escuelas de pensamiento de la ciencia abierta. **Fuente:** Fecher y Friesike (2014).

La [Red de Ciencia Abierta y Colaborativa para el Desarrollo](#), comunidad en la que participan equipos de investigación, profesionales del desarrollo y actores sociales de 26 países de América Latina, África, Medio Oriente y Asia, publicó el [Manifiesto de Ciencia abierta y Colaborativa: Hacia una ciencia abierta inclusiva por el bienestar social y ambiental](#) (OCSDNet, 2017), poniendo el énfasis en las agendas de desarrollo sostenible, la justicia social y de apertura situada en contexto.



FIGURA 2 Valores y principios del Manifiesto de Ciencia abierta y Colaborativa (OCSDNet).

Abadal y Anglada (2020) analizaron el nacimiento y evolución del concepto de ciencia abierta, situando a **Merton** (1973) y a su *ethos científico*³ entre los primeros sociólogos que plantearon la apertura como “característica institucional fundamental para el avance de la ciencia”. Este concepto evolucionó y los cambios en su denominación han sido acompañados de cambios en su significado. La *e-ciencia* se centraba en la infraestructura de cálculo y comunicación, la *ciencia interconectada* se enfocaba en las redes sociales y la *ciencia 2.0* enfatizaba compartir resultados y recibir comentarios. En contraste, la ciencia abierta implica la apertura de todas las etapas de la investigación científica. A lo largo de los años, ha habido discrepancias sobre qué elementos la componen, con algunos autores considerando solo unos pocos componentes y otros llegando hasta una docena. Se ha considerado que no es un concepto rígido y bien definido, sino más bien un término amplio que abarca diversas prácticas y componentes relacionados con la apertura y accesibilidad de la investigación científica (**Abadal y Anglada**, 2020). Para estos autores, los componentes incluidos en las recomendaciones de la Open Science Policy Platform⁴ (OSPP) son los más reconocidos: incentivos, nuevas métricas de evaluación, acceso abierto, infraestructuras, datos abiertos, integridad en la investigación, formación y ciencia ciudadana.

Tras una revisión sistemática de la literatura, **Vicente-Sáez y Martínez-Fuentes** (2018) la definieron como “conocimiento transparente y accesible que se comparte y desarrolla a través de redes colaborativas”. El término “conocimiento abierto” abarca código, datos, otros resultados, publicaciones, información e ideas. Debe ser transparente, accesible, compartido y desarrollado de forma colaborativa. Dentro los componentes de la ciencia abierta considerados están el acceso abierto, los datos de investigación abiertos, las altmétricas (medidas alternativas al factor de impacto), cuadernos de investigación abiertos y cuadernos abiertos de laboratorio, blogs, bibliografías colaborativas, ciencia ciudadana, revisión abierta, registro previo del diseño del estudio (pre-registro) y software libre.

³ Este autor definió el *ethos científico* como un complejo de valores y normas que guían la conducta de los científicos, enfocados en la búsqueda desinteresada del conocimiento, estableciendo cuatro principios clave: universalismo, comunalismo, desinterés y escepticismo organizado, buscando la integridad y el avance de la ciencia como empresa colectiva, no personal, a pesar de las tensiones con intereses externos.

⁴ Grupo de alto nivel de la Comisión Europea de 2016 para orientar su estrategia de ciencia abierta y hacer la investigación más eficiente y orientada a la sociedad mediante apertura de datos, colaboración y prácticas como acceso abierto, datos FAIR, ciencia ciudadana y nuevas formas de evaluación. Finalizó su actividad en 2020 pero [sus recomendaciones](#) influyen en las políticas europeas, especialmente en *Horizon Europe*: “tan abierto como sea posible, tan cerrado como sea necesario” y el desarrollo de infraestructuras la [European Open Science Cloud](#) (EOSC) para compartir datos de investigación.

El término ciencia abierta comienza a instalarse políticamente en Europa a partir de 2014, concretamente en el discurso del comisario [Carlos Moedas](#) quien promovió una visión de las "[3 O's: Open Innovation, Open Science y Open to the World](#)", respaldada por los estados miembros en el Consejo de 2016 ([Méndez Rodríguez, 2021](#)). La [Declaración de Panamá sobre Ciencia Abierta](#) (2018) resulta de una reunión de representantes de universidades y organizaciones de América Latina y el Caribe, en Panamá durante el [Foro CILAC 2018](#).

Esta declaración destaca la importancia de considerar el conocimiento como un bien común. Se aboga por la promoción de la ciencia abierta, reconociendo su papel en el fortalecimiento de la democracia, la libertad y la justicia social. Además del acceso abierto, en su perspectiva la ciencia abierta se extiende a la colaboración entre academia, ciudadanía, gobierno y empresas.

Este [documento](#) expresa que la apertura científica no solo mejora la eficiencia y la productividad, sino que también construye ciudadanía y fortalece la democracia. Asimismo, se proponen una serie de estrategias de políticas públicas, como creación de políticas nacionales, promoción de publicaciones de acceso abierto, desarrollo de infraestructuras y herramientas abiertas regionales, programas de sensibilización y capacitación, estrategias de implementación de la ciencia ciudadana y políticas centradas en la equidad social y la justicia. Se reconoce que la transición hacia la ciencia abierta requiere un cambio cultural y un compromiso de recursos, pero que los beneficios sociales justifican los esfuerzos necesarios para su implementación en la región de América Latina y el Caribe.

Por su parte, la definición de **UNESCO** (2021) en su recomendación sobre la ciencia abierta, incluye cuatro pilares fundamentales con sus componentes: (1) acceso abierto al conocimiento científico; (2) infraestructuras de la ciencia abierta; (3) participación abierta de los agentes sociales y (4) diálogo abierto con otros sistemas de conocimiento, tal como se observa en la siguiente figura.

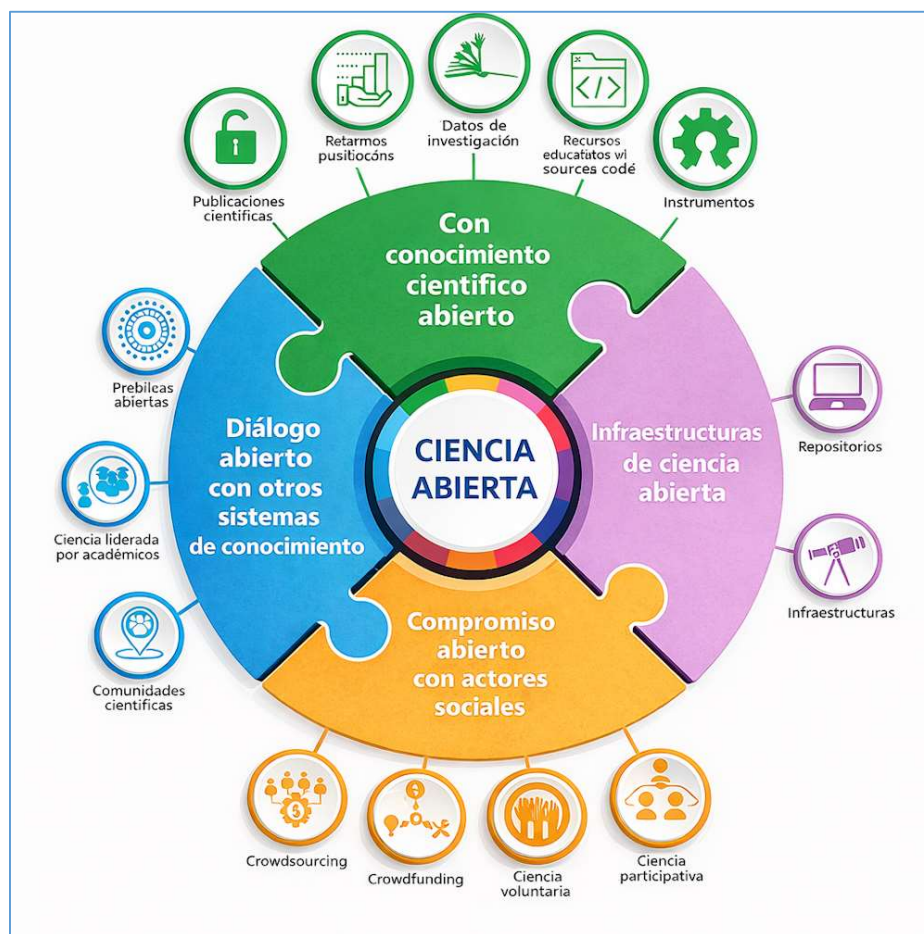


FIGURA 3 Pilares de la ciencia abierta. Fuente: UNESCO (2021).

Dentro del acceso abierto al conocimiento científico (el sector de color verde en la figura) se encuentran:

- El *acceso abierto a las publicaciones científicas*, que incluye artículos revisados por pares, informes de investigación, contribuciones a conferencias, libros, datos de investigación, programas informáticos, códigos fuente, metodologías de trabajo y protocolos y otros materiales que tengan una licencia abierta o estén dedicados al dominio público. Se destaca que los contenidos deben estar depositados en un repositorio en línea abierto, siguiendo estándares aceptados internacionalmente y sostenidos por una institución sólida “que garantice el acceso abierto, la distribución sin restricciones, la interoperabilidad y el archivado a largo plazo”.
- Otro componente de la definición y en el cual se centra este trabajo, son los *datos de investigación abiertos* que incluyen datos en formato digital y analógico, crudos y procesados, junto con los metadatos que los acompañan. Puede comprender índices numéricos, registros textuales,

imágenes, sonidos, protocolos y flujos de trabajo. Los datos tienen que ser curados para ser accesibles, reutilizables, preservados y compartidos de manera abierta por cualquier persona, siempre que se proporcione el debido reconocimiento. Estos datos usan formatos comprensibles y modificables tanto para seres humanos como para sistemas automatizados, siguiendo principios y buenas prácticas FAIR (Fáciles de encontrar, Accesibles, Interoperables y Reutilizables). Destaca la importancia de desarrollar herramientas y protocolos para anonimizar los datos y sistemas de acceso mediado para compartir la mayor cantidad cuando sea posible.

- Los *programas informáticos de código abierto* y *código fuente abierto* es software cuyo código fuente está disponible para el público con licencias abiertas, en un formato fácilmente comprensible y modificable tanto por humanos como por otros sistemas automáticos. Este código debe estar depositado en repositorios abiertos y su licencia ha de permitir posibles modificaciones, creación de obras derivadas y compartir en condiciones abiertas iguales o compatibles. Ejemplos de gran adopción y generación de comunidades internacionales son aplicativos para crear repositorios como [DSpace](#) (publicaciones) y [Dataverse](#) (datos de investigación⁵).
- Los *equipos informáticos abiertos* son objetos físicos cuyas especificaciones de diseño se brindan con una licencia que permite a cualquier persona estudiar, modificar, crear y distribuir el objeto, con el objetivo de facilitar la capacidad de comprender, adaptar y compartir su conocimiento sobre el diseño y funcionamiento del equipo informático.
- En ambos casos, de los programas informáticos de código abierto como de los equipos informáticos abiertos, se crean comunidades para la contribución, atribución y gobernanza, para lograr sostenibilidad, promover la reutilización y evitar la duplicación de esfuerzos.

La recomendación de **UNESCO** (2021) promueve la máxima apertura posible, con restricciones justificadas solo por seguridad nacional, confidencialidad, privacidad, protección de derechos de propiedad intelectual, datos personales, procesos legales, entre otros. Los datos que no están abiertamente disponibles pueden compartirse entre usuarios específicos según criterios de acceso definidos por autoridades competentes.

⁵ Dataverse se instaló en la [Universidad Nacional de Rosario](#). Se comparte con la comunidad la traducción al español tanto de la "[Guía de uso de Dataverse](#)" como del código fuente.

Otros pilares fundamentales son:

- Las *infraestructuras de la ciencia abierta* que abarcan una diversidad de recursos compartidos como laboratorios abiertos, plataformas de ciencia abierta, repositorios abiertos y otros servicios para el acceso, análisis y distribución de datos y literatura científica. Estas infraestructuras permiten atender las necesidades de diversas comunidades y deben cumplir con estándares de interoperabilidad y buenas prácticas.
- La *participación abierta de los actores sociales*: colaboración entre científicos y sociedad más allá de la comunidad científica, involucrando a la ciudadanía en la generación de conocimiento y fomentando la inclusión de perspectivas.
- El *diálogo abierto con otros sistemas de conocimiento* reconoce la amplitud y riqueza de perspectivas en producción de conocimiento y la colaboración entre sistemas de conocimiento diversos y respetando los derechos de aquellos que poseen conocimientos tradicionales.
- Además, *la comunicación de la ciencia* busca difundir el conocimiento científico a través de diversas actividades, como periodismo científico, divulgación, conferencias y redes sociales, promoviendo la confianza de la sociedad en la ciencia y ampliando la participación más allá de la comunidad científica⁶.

La ciencia abierta reúne a una amplia diversidad de actores, sin distinción de origen, género, edad, disciplina, situación socioeconómica, financiación o etapa profesional. Incluye a la comunidad científica y académica, profesionales de la información y la tecnología, instituciones, financiadores, responsables políticos, organizaciones sociales, comunidades, sector privado y ciudadanía en general (UNESCO, 2021).

La equidad y la inclusión son valores de la ciencia abierta para la UNESCO y también indicadores de éxito de su implementación. Para lograr estos objetivos, los países en desarrollo, con la ayuda de financiadores, han de desarrollar infraestructuras que incluya internet de alta velocidad, infraestructura de investigación local y un grupo importante de investigadores capacitados. Si aumenta el acceso al conocimiento, la ciencia abierta generará mayores capacidades científicas en

⁶ Un ejemplo son las [Historias de Datos](#) del repositorio de datos de la [Universidad Nacional de Rosario](#).

comunidades de bajos ingresos, marginadas y en vías de desarrollo (Ali-Khan et al., 2018). UNESCO ha preparado una [Estrategia de Implementación](#) y un [Kit de herramientas](#) para ayudar a todas las partes interesadas en desarrollar políticas, infraestructuras, estrategias de adopción, capacitación, financiamiento y más.

1.3 Responsabilidad en la investigación e innovación

La responsabilidad en la investigación e innovación (RRI) es una perspectiva con un particular énfasis en los proyectos de la UE, promoviendo la responsabilidad y la ética en todas las etapas de la investigación. Este enfoque se alinea con las expectativas y necesidades de la sociedad en relación con la investigación y la innovación, fomentando la creación de procesos inclusivos y sostenibles. En este marco se abordan cuestiones como la promoción de la ciencia abierta, así como temas éticos (integridad de la investigación), la igualdad de género, la participación ciudadana, educación científica para el aprovechamiento del conocimiento y la gobernanza de los procesos científicos e innovadores (Wittrock et al., 2021; Méndez Rodríguez, 2021). Esta profesora representa la ciencia abierta como un hongo: “un símbolo alegre y expresivo de lo que significa hacer ciencia de manera abierta”.

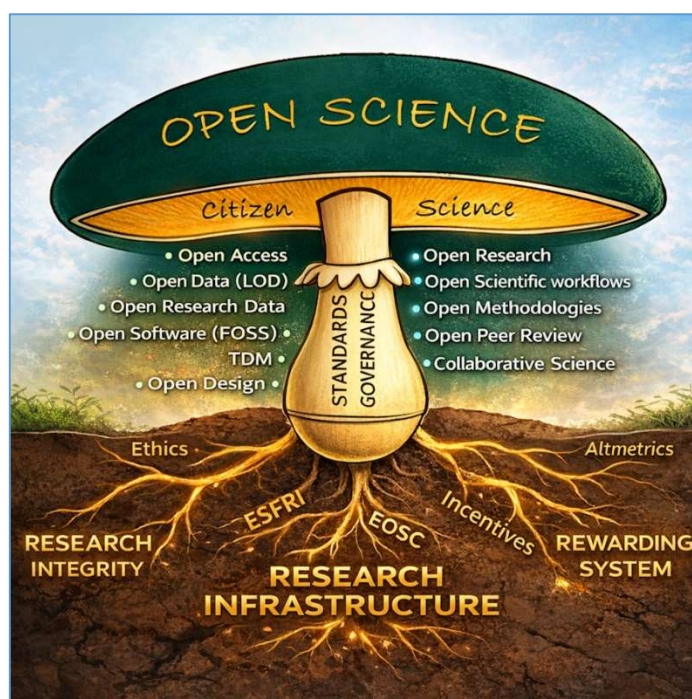


FIGURA 4 Hongo de la ciencia abierta. Fuente: Méndez Rodríguez (2021, 2022). Recreado por Judit Eva's Fazekas-Paragh bajo licencia <https://creativecommons.org/licenses/by/4.0/>

La RRI es el suelo ético y filosófico en el que crece este hongo, proporcionando una base sólida para el desarrollo de la ciencia abierta. La ciencia ciudadana, por otro

lado, se asemeja a las esporas del hongo; son como semillas que permiten que la ciencia se cuestione a sí misma y se reproduzca de nuevas maneras. Finalmente, las raíces del hongo representan las bases necesarias para construir un sistema de sólido, incluyendo las infraestructuras de investigación, la integridad científica y un nuevo sistema de incentivos para los científicos. Esta metáfora ilustra cómo la ciencia abierta se apoya en principios éticos, participación pública y una infraestructura sólida para transformar la forma en que hacemos y compartimos la investigación científica.

1.4 Valores de Merton y Ciencia abierta

Las normas y contranormas de la ciencia son pautas sociales que regulan la investigación científica, la práctica, la publicación y el intercambio de datos entre investigadores en las comunidades científicas. Analizarlas es importante, ya que influyen en cómo los científicos comparten o no sus datos. Aunque no existen normas absolutas que se apliquen a todos los investigadores en todas las disciplinas y épocas, el estudio de las normas tradicionales y contranormas permite entender las prácticas en diversas comunidades científicas.

Merton (1973) identificó cuatro normas de la ciencia: comunalismo, universalismo, desinterés y escepticismo organizado conocidas como su acrónimo en inglés [CUDOS](#) ('Communalism, Universalism, Disinterestedness, and Organised Scepticism'). Estas normas ayudan a entender el funcionamiento de las comunidades científicas. Merton argumentó que estas normas podían "... describir adecuadamente la búsqueda de la ampliación del conocimiento certificado como objetivo institucional de la ciencia" (**Hosseini et al.**, 2022). Las normas mertonianas constituyen caracterizaciones idealizadas y pueden considerarse valores de la estructura cultural de la ciencia. Se entienden como principios a seguir, no como descripciones del comportamiento de los científicos reales.

- *Comunalismo*: Significa que los resultados científicos deben estar disponibles para el público en general y compartidos con todos los miembros de la comunidad científica. Es decir, los resultados de la investigación científica deberían pertenecer a la comunidad que los produce, porque la mayoría de estos resultados se basan en la colaboración entre académicos y en el trabajo de académicos anteriores. Los investigadores deberían brindar a la comunidad científica información sobre sus resultados para que otros científicos puedan replicar o refutarlos. El comunalismo está

en la base de la ciencia abierta ya que promueve el intercambio abierto y gratuito del conocimiento científico, y también fomenta el intercambio de datos de investigación junto con el análisis y los resultados finales. Este valor fundamenta las políticas de ciencia abierta ya que las instituciones científicas deberían respaldar la comunicación completa y abierta de los resultados, que subyace como el valor en la famosa afirmación de [Isaac Newton](#): "si he visto más lejos, es porque estoy parado sobre los hombros de gigantes"⁷.

- *Universalismo*: Implica que la calidad de la investigación se evalúa de acuerdo con estándares científicos objetivos, sin verse afectada por las características personales de los científicos, como su origen social o educativo, no estando sujeto a prejuicios o discriminación. Esto promueve la igualdad de oportunidades en la investigación científica. Como ejemplo del universalismo en la práctica se cita el proceso de revisión por pares ciego.
- *Desinterés*: Los investigadores deben investigar por su valor intrínseco y por el avance del conocimiento y la resolución de problemas. Se espera que los investigadores estén menos interesados en cualquier recompensa personal (como beneficios financieros o reputación personal) que en el desarrollo del conocimiento científico en su comunidad de investigación. El desinterés también evitaría que los investigadores busquen alinear su investigación con oportunidades de financiamiento.
- *Escepticismo organizado*: Bajo esta norma, los investigadores deben revisar los resultados científicos con un grado de escepticismo, incluso sus propios resultados. El trabajo publicado se debe poder replicar, en caso contrario, debe ser rechazado. El escepticismo organizado requiere que se examinen los trabajos de otros investigadores en términos de evidencia empírica y lógica antes de aceptar sus resultados como conocimiento. Esto sustenta el argumento de compartir datos de investigación como evidencia científica.

A estas normas se oponen las "contra normas" de **Mitroff** (1974) para quien no son simplemente lo opuesto a lo propuesto por Merton, sino que en realidad permiten visualizar diferentes motivaciones y formas institucionales que también podrían contribuir a la generación de nuevo conocimiento certificado:

⁷ La famosa frase de **Isaac Newton** no aparece en un libro, sino en una carta dirigida a **Robert Hooke** en 1676. En ella, respondiendo a una crítica, reconocía que sus descubrimientos se basaban en el trabajo de grandes pensadores anteriores como Galileo, Kepler y Copérnico. Se usa comúnmente para ilustrar el progreso científico.

- *Soledad*: Contrarresta el comunalismo y sugiere que los científicos a menudo consideran sus resultados como propiedad protegida y necesitan mantener el secreto para proteger sus derechos. Esto puede llevar a la falta de colaboración y al secretismo en la investigación. De este modo, los resultados científicos no serían de la comunidad sino de los investigadores.
- *Particularismo*: En oposición al universalismo, el particularismo implica que los investigadores pueden ser evaluados según sus antecedentes económicos, sociales, religiosos, políticos, étnicos, etc. lo que puede llevar a favoritismos en la distribución de reconocimiento y oportunidades en la comunidad científica. **Mulkay** (1976) argumentó que la distribución de reconocimiento está sesgada hacia los investigadores de universidades prestigiosas. Un ejemplo es el denominado “*efecto Mated*” según el cual, los investigadores que generan una mejora científica significativa en sus disciplinas tienden a tener más credibilidad de la que deberían tener (**Merton et al.**, 1996). Es decir, la tendencia de los científicos exitosos a recibir más reconocimiento y recursos.
- *Interés propio*: Contraparte del desinterés, el interés propio sugiere que algunos investigadores pueden priorizar las recompensas financieras personales sobre la búsqueda del conocimiento por sí mismo, lo que puede influir en las motivaciones detrás de la investigación. Un ejemplo de interés propio es desarrollar una agenda de investigación basada en oportunidades de financiamiento, en lugar de tener el deseo de perseguir resultados científicos en el área de interés del investigador.
- *Dogmatismo organizado*: Como contra norma frente al escepticismo organizado, esto implica que los científicos pueden aceptar resultados científicos ajenos o propios sin un escrutinio crítico adecuado, lo que podría llevar a la aceptación ciega sin cuestionar su validez.

Tanto las normas tradicionales y contra normas coexisten en las instituciones científicas reales. **Mulkay** (1976) encontró que los científicos utilizan las [normas mertonianas](#) como “vocabularios de justificación, utilizados para evaluar, justificar y describir las acciones de los científicos, pero de forma no institucionalizadas dentro de la comunidad científica, manteniendo una conformidad general”.

Las normas mertonianas presuponen que los científicos las aceptan libremente, aunque en la práctica funcionan como un mecanismo de control social. Por ello, cuando cambian los sistemas de incentivos, también pueden hacerlo los valores. Actualmente, estos incentivos se vinculan con métricas cuantificables, como las

citas y el factor de impacto de las publicaciones, muy valoradas por investigadores e instituciones científicas. Esta situación favorece la adopción de las «contranormas» mencionadas.

La coexistencia de la ciencia abierta y los valores mertonianos en el ámbito académico genera diversas tensiones: publicación en revistas de acceso cerrado, realización de experimentos sin compartir datos, falta de transparencia sobre el código y el software utilizados, diferencias en el comportamiento de los revisores según el tipo de revisión y descripciones metodológicas incompletas. Aunque el conocimiento se difunde, la apertura puede verse limitada por la búsqueda de beneficios personales (Hosseini et al. 2020).

El comunalismo mertoniano y la ciencia abierta promueven la idea de compartir, pero con diferencias importantes. En un principio, el énfasis se ponía en compartir descubrimientos y resultados de la investigación científica, en la ciencia abierta, el “qué” es más amplio porque abarca todo el proceso de investigación incluyendo los datos, protocolos, código fuente, entre otros. Asimismo, se han adoptado licencias abiertas de tipo ‘copyleft’⁸ que van más allá de los derechos de propiedad intelectual tradicionales y patentes, planteándose diferencias también en cuándo y dónde se comparte. Merton (1942) enfatiza la importancia de compartir “importantes descubrimientos”, es decir, hitos específicos de la investigación.

En contraste, la ciencia abierta promueve un enfoque más inmediato de compartir. El conocimiento científico debe compartirse a todos los agentes de manera inmediata o lo más rápidamente posible (UNESCO, 2021). Esta prontitud se relaciona con la idea de aumentar la productividad y eficiencia de la investigación al permitir que otros investigadores utilicen materiales e ideas compartidas sin demoras. Los preprints compartidos durante la pandemia COVID-19 son un ejemplo de esta urgencia por difundir la información y aportar soluciones (Torres-Salinas, 2020).

A pesar de esto, no todos los investigadores respaldan el compartir inmediatamente. En muchos casos se retrasa la difusión de los resultados, métodos o materiales por querer verificar su calidad o proteger las perspectivas futuras de

⁸ Estas licencias son mecanismos legales que permiten el uso, modificación y distribución libre de obras (software, textos, etc.), pero con una condición clave: cualquier versión derivada **debe licenciarse bajo los mismos términos**, asegurando que la libertad de compartir y modificar se mantenga siempre, como un “efecto contagio” (virales). Las más conocidas son la [GNU GPL](#) (fuerte, viral) para software y la familia [Creative Commons con NC/ND](#) (no comercial/sin obras derivadas)

investigación. Cuando los resultados son patentables, la publicación suele retrasarse hasta después de presentar una solicitud de patente. Además, inversiones de tiempo y personal en la creación de resultados, influye en las decisiones sobre cuándo compartir (Hosseini et al., 2020).

La investigación científica es tanto un esfuerzo individual como de una comunidad, lo que lleva a tensiones entre el beneficio individual y el beneficio comunitario de la apertura. La regla de prioridad en la ciencia, donde el primero en hacer un descubrimiento recibe todo el crédito, es un punto conflictivo. El reconocimiento es fundamental para las carreras académicas individuales porque les permite aspirar a mejorar su categoría, obtener financiamiento o un cargo. Cambiar las estructuras de incentivos puede influir en el comportamiento de los investigadores, de hecho, se trata de un desafío clave para el movimiento de ciencia abierta. Diseñar incentivos adecuados en diferentes etapas del proceso de investigación puede fomentar un mayor intercambio de recursos (Heesen, 2017).

Las motivaciones de los investigadores son múltiples. Por ello, circunscribirse a un modelo mertoniano de incentivos deja de lado dimensiones como el apego emocional a las ideas (Mitroff, 1974), los aspectos éticos, la curiosidad, generar confiabilidad entre otros, que deben considerarse al momento de evaluar la apertura de las investigaciones. Si se considera el valor del *universalismo*, las evaluaciones académicas deben basarse en datos objetivos y argumentos sólidos, sin prejuicios de ningún tipo.

Desde la ciencia abierta, la puesta en práctica de la revisión por pares abierta tiende a promover la transparencia porque se revelan las identidades de los autores y revisores. Los informes de revisión se comparten públicamente junto con los artículos, contribuyendo a reducir sesgos en el proceso de evaluación al permitir que la comunidad científica y el público valoren tanto la investigación como la revisión realizada. La revisión por pares abierta es más inclusiva al facilitar la participación de un mayor número de voces y perspectivas en la evaluación académica (Hosseini et al., 2020).

Merton, al plantear el valor del *desinterés*, puso la atención en los marcos institucionales que estimulan y controlan los intereses de los investigadores. **Wunderlich** (1974) describe este concepto como una relación entre un medio (la presión institucional de vigilancia) y un fin (la producción de nuevo conocimiento certificado). Según este autor la vigilancia es un componente crucial del *desinterés*, mientras funcione la vigilancia, la ciencia puede considerarse como desinteresada.

Cuando los medios y el fin no se alinean, el *desinterés* está en riesgo. Las prácticas de ciencia abierta buscan reformar las estructuras normativas de la ciencia promoviendo la transparencia, la apertura y el acceso, con impactos tanto individuales como institucionales. Los investigadores pueden obtener prestigio personal mediante la adopción de prácticas abiertas (Allen y Mehler, 2019) y contribuir, a su vez, al fortalecimiento del reconocimiento institucional. Esto puede favorecer la transformación de las estructuras normativas de la ciencia y de los sistemas de evaluación científica.

Miradas críticas como la de Mirowski (2018), sugieren que las mejoras reales en la investigación bajo las consignas de la ciencia abierta son limitadas y que este movimiento estaría siendo cooptado por intereses comerciales, advirtiendo del peligro de que las grandes empresas terminen como dueñas de las plataformas y datos creados en el marco de iniciativas abiertas, algo contrario a sus fundamentos y principios originales de apertura y colaboración en la investigación.

- El *desinterés*, como valor, implica que la ciencia debe ser objetiva y neutral, libre de intereses e influencias externas. Sin embargo, en la práctica diaria de las investigaciones científicas, múltiples intereses coexisten e influyen en cada situación.
- El *escepticismo organizado* impulsa los avances científicos y evita que se establezcan creencias dogmáticas. La ciencia se beneficia de la actitud crítica de los investigadores, promoviendo la búsqueda constante de evidencias sólidas en lugar de aceptar ideas o teorías sin cuestionarlas.

En el contexto de la ciencia abierta, se han desarrollado herramientas como repositorios de publicaciones y datos, así como canales de comunicación para los investigadores (foros, listas de correo orientadas a disciplinas, canales en redes sociales, congresos en línea, videoconferencias y aplicaciones como *Slack* y *Telegram*) que facilitan la puesta en práctica de este escepticismo.

Sin embargo, es necesaria una alfabetización digital para utilizar estas nuevas herramientas de manera efectiva. Además, se requieren nuevas normas y subnormas para adaptarlas a las necesidades y características de los proyectos científicos actuales, que incluyen innovación abierta y ciencia ciudadana (Hosseini et al., 2020).

1.5 Preocupación por la equidad e inclusión en la ciencia abierta

En el manifiesto hacia una ciencia abierta y colaborativa se plantean los siguientes principios:

1. Promueve el conocimiento como bien común. Permite un patrimonio común de conocimientos en el que cada individuo tiene los medios para decidir cómo se gobiernan y gestionan sus conocimientos para satisfacer sus necesidades.
2. Integra diferentes tradiciones científicas y formas de saber. Reconoce la justicia cognitiva, la necesidad de que coexistan en la producción científica diversas comprensiones del conocimiento.
3. Atiende el rol que tiene el poder y la desigualdad en la producción y distribución de conocimiento. Practica la apertura situada abordando las formas en que el contexto, el poder y la desigualdad condicionan la investigación científica.
4. Crea oportunidades para participar en todas las etapas de la investigación. Defiende el derecho de cada individuo a la investigación y permite diferentes formas de participación en todas las etapas de la investigación.
5. Favorece la colaboración equitativa entre científicos y actores sociales. Fomenta la colaboración equitativa entre científicos y actores sociales y cultiva la co-creación y la innovación social en la sociedad.
6. Incentiva el diseño de mecanismos inclusivos de información y comunicación. Incentiva las infraestructuras inclusivas que empoderan a personas de todas las capacidades para aprovechar y utilizar tecnologías accesibles de código abierto.
7. Mejora el bienestar de nuestra sociedad y de nuestro planeta. Se esfuerza por utilizar el conocimiento como una vía hacia el desarrollo sostenible, equipando a cada individuo para mejorar el bienestar de nuestra sociedad y del planeta.

Dependiendo de las definiciones y prioridades puestas en juego en el contexto de la ciencia abierta, valores como la equidad y la inclusión son susceptibles de entrar en conflicto con otros principios de esta disciplina, influenciados por los diferentes intereses de los actores (investigadores, instituciones de investigación, organismos de financiamiento, gobiernos y editores). Existen discrepancias entre los ideales expresados y las políticas y prácticas que se implementan.

Ross-Hellauer et al. (2022) identificaron varias áreas de preocupación que pueden dificultar la realización de la promesa de equidad en la ciencia abierta.

Aspectos	Área de preocupación (resumida)	Grupos más afectados (resumido)
Factores generales	Altos costos en infraestructura, apoyo y capacitación.	Instituciones y regiones con bajos recursos
	Implementación sujeta a agendas políticas, a veces desiguales.	Regiones sin respaldo o con políticas inequitativas
	Riesgo de mercantilización del conocimiento científico.	Disciplinas y científicos fuera de la lógica neoliberal
Acceso abierto	El modelo APC excluye y segmenta la autoría.	Investigadores e instituciones con menos recursos; también mujeres
	Las publicaciones depredadoras afectan a grupos vulnerables.	Autores de países en desarrollo y noveles
Datos abiertos y FAIR	Las políticas de datos pueden favorecer ciertas disciplinas.	Disciplinas cualitativas e investigadores de contexto sensible
	Las brechas en infraestructura y alfabetización perpetúan desigualdades.	Zonas e instituciones con menos recursos
	Los primeros adoptantes obtienen más citas.	Regiones y autores con menos acceso inicial
Métodos e infraestructura abierta	Transparencia requiere formación y medios adicionales.	Investigadores e instituciones con menos recursos
	Reproducibilidad puede desvalorizar métodos interpretativos.	Disciplinas cualitativas
	Dependencia de plataformas privadas vulnera principios abiertos.	La ciencia en general
Evaluación Abierta	La falta de anonimato puede limitar la crítica sincera.	Investigadores nuevos o en posición vulnerable
	Altmétricas poco confiables, manipulables y anglocéntricas.	Toda la comunidad, especialmente regiones no anglófonas
Ciencia Ciudadana	Participación sesgada y posible extracción de datos.	Público marginado
Interfaces con sociedad, industria y política	Difusión efectiva requiere trabajo traslacional costoso.	Zonas con menos recursos y capacidad institucional
	Relación asimétrica con la industria, enfocada en fines económicos.	Áreas científicas no explotables comercialmente
	Inclusión pública percibida como superficial o desigual.	Público general

TABLA II Resumen de áreas identificadas de preocupación para la equidad en la ciencia abierta.

Fuente: (Ross-Hellauer et al., 2022).

Las políticas de ciencia abierta deberían considerar el contexto de desigualdades de poder e históricas en la generación de conocimiento. Una visión simplista sobre

la apertura del conocimiento corre el riesgo de mantener las ventajas acumulativas de los grupos con más influencia y liderazgo en la producción de conocimiento a lo largo de la historia (Ross-Hellauer et al., 2022). En lugar de un único enfoque, es importante considerar las dinámicas locales y las estructuras de poder que afectan las formas en que los individuos tienden a colaborar (o no) dentro de un contexto particular (Hillyer et al., 2017, OCSDNet, 2017).

Clinio (2019) observa un conflicto entre al menos dos enfoques de ciencia abierta en Latinoamérica. El primero es una visión utilitarista de la ciencia, centrándose en lograr resultados más eficientes, mayor productividad y competitividad. El segundo enfoque se preocupa más por cuestiones de equidad, garantía de derechos, justicia cognitiva y social. Este debate está estrechamente relacionado con el fenómeno del capitalismo académico y cómo las universidades están pasando de ser instituciones que generan conocimiento para el beneficio común a entidades orientadas al mercado, donde la eficiencia es prioritaria.

1.5.1 Perspectiva constructivista

De acuerdo con Knorr-Cetina (1991) las *normas mertonianas* son "cánones" expresados a través de las demandas hechas a los científicos: los hallazgos científicos deben ser publicados (*comunalismo*), las afirmaciones de conocimiento deben ser sometidas a criterios impersonales de evaluación (*universalismo*), los intereses personales deben ser excluidos de los procedimientos científicos adecuados (*desinterés*) y se permite y fomenta la crítica (*escepticismo organizado*). Se parte de la premisa de que la investigación científica es un proceso de construcción, de "fabricación" reflexiva (Knorr-Cetina, 2005).

Desde esta perspectiva, la investigación científica es un proceso socialmente situado y constructivista en el que los investigadores, insertos en comunidades y culturas científicas específicas, participan activamente en la construcción y producción de conocimiento. Esta construcción de conocimiento no se limita a la mera descripción de hechos objetivos, también implica la manipulación de objetos cognitivos, la interacción con la tecnología y la creación de representaciones y modelos para comprender y explicar fenómenos en función de las oportunidades y restricciones del entorno de investigación. La investigación científica se lleva a cabo de acuerdo con normas y valores compartidos en la comunidad científica, pero es inherentemente contextual y dependiente de factores sociales, culturales y materiales. En última instancia, se concibe como un proceso dinámico y

adaptativo en el que los científicos toman decisiones contextualmente informadas para avanzar en la comprensión del mundo natural y social porque no existe una 'perspectiva desde ninguna parte, el conocimiento siempre está situado en un lugar, tiempo, condiciones, prácticas y entendimientos. No existe un conocimiento único, sino múltiples conocimientos (Borgman, 2015, citando a Nancy Van House, 2004). Esta perspectiva situada de la ciencia abierta **es la que se pone de relieve** en este libro.

1.5.2 Perspectiva de la ciencia abierta desde la filosofía de la ciencia

Leonelli (2023) explica que este sistema, basado en la cantidad e impacto de las publicaciones, en el cual los investigadores jóvenes son "entrenados" para priorizar la velocidad sobre el detalle y el rigor, está generando una trayectoria hacia una ciencia cada vez menos interdisciplinaria, ya que la interdisciplinariedad requiere tiempo, colaboración, y una variedad de métodos, entre otros factores.

La ciencia abierta se caracteriza por su enfoque en el intercambio y la transparencia. No obstante, la implementación efectiva de este modelo enfrenta diversas tensiones, oportunidades y retos. Este enfoque es particularmente pertinente en el contexto de las universidades públicas de Argentina, con una infraestructura de investigación muy variable según las áreas y disciplinas, siendo los fondos destinados a la ciencia muy limitados. También se aborda la reproducibilidad en la investigación científica, principio fundamental en la ciencia abierta, planteando que una comprensión estandarizada y limitada de la reproducibilidad pondría en riesgo la diversidad de métodos y enfoques de investigación en diferentes disciplinas. La mayoría de las veces, la experiencia y conocimiento especializado, sumado al contexto social, institucional y cultural de la investigación no se pueden reproducir.

Se considera la heterogeneidad de métodos y contextos de investigación cuando se aplica el principio de reproducibilidad en la ciencia abierta. De acuerdo a esta perspectiva, se destaca el concepto de diversidad epistémica como relevante para la ciencia abierta. Los componentes de esta diversidad influyen en la forma en que se realiza la investigación y se comparten los resultados en diferentes comunidades y contextos. Existen diversas formas de diversidad epistémica relacionadas con:

- Aspectos conceptuales: perspectivas teóricas, supuestos fundamentales.
- Aspectos materiales: objetivos de estudio, Materiales utilizados, procedencia de los materiales.

- Metodológica: métodos, herramientas y estándares (formatos y semántica).
- Infraestructural: niveles de financiamiento y limitaciones, tecnologías de la información y comunicación, y otras tecnologías digitales, lugares para la publicación e intercambio de información, movilidad y transporte, fuentes de financiamiento y compromisos relacionados.
- Sociocultural: sistemas de evaluación de la investigación (local, nacional, internacional), responsabilidad legal y ética, ubicación geopolítica, valores y objetivos, idioma, características socio demográficas de los investigadores (género, clase social, etnia, edad, habilidades físicas).
- Institucional: etapa de la carrera de investigación, dinámicas de poder, apoyo institucional y administrativo, campo de estudio y normas relacionadas, regímenes de propiedad intelectual”.

En lugar de una visión de la ciencia abierta donde el énfasis está puesto en compartir objetos (resultados, datos, etc.), se propone una visión orientada al proceso de la investigación, donde la ciencia contribuya a potenciar las capacidades humanas de actuar y comprender el mundo, y que puedan ser comunicadas y revisadas por otros investigadores. Esto no significa que los objetos creados en la fabricación del conocimiento no son importantes, sino que lo son las formas específicas a través de las cuales el conocimiento activo se abstrae, encapsula y negocia, actuando como anclajes materiales esenciales para la comunicación y la naturaleza interpersonal y social de la investigación científica (**Knorr Cetina, 2005**).

El propósito de compartir estos objetos está subordinado a facilitar las acciones humanas, ya sea en forma de razonamiento teórico o intervenciones, que producen métodos y herramientas para interactuar con el mundo (**Leonelli, 2023**). Su concepto de ciencia abierta se centra en el *proceso*, como el desarrollo de *conexiones juiciosas* entre los investigadores, conexiones generalmente mediadas y constituidas a través de la tecnología y la familiaridad con entornos de investigación específicos, incluyendo participantes humanos y no humanos en esos entornos. Que estas conexiones sean *juiciosas*, requiere evaluar qué conexiones pueden ser relevantes y beneficiosas para el esfuerzo de la investigación en cuestión y evaluar las posibles implicaciones (negativas o positivas) de tales iniciativas. Se propone dejar de lado la idea distributiva de equidad en la ciencia abierta y fomentar un enfoque equitativo orientado a disminuir las desigualdades existentes y fomentar activamente la capacidad de adopción significativa de la ciencia abierta en los contextos vulnerables y los posibles participantes (**Ross-Hellauer et al.; Leonelli, 2023**). Desde esta mirada, la implementación de ciencia

abierta debe empezar por considerar qué es lo necesario para hacer a la investigación más inclusiva, diversa y justa, en lugar de esperar que ese resultado se dé naturalmente a partir de la elegir software, infraestructuras, estándares, plataformas de publicación o cualquier otro recurso tecnológico o institucional “correcto” diseñando para facilitar el acceso a los recursos.

1.6 El acceso abierto

Antes del acceso abierto, las revistas denominadas de “corriente principal” se distribuían bajo el modelo de suscripción. Los investigadores enviaban sus artículos para su publicación (podía tener un costo), otorgaban a las editoriales los derechos de reproducción y los lectores pagaban por acceder a los contenidos, ya sea a través de suscripciones individuales o institucionales a través de las bibliotecas (Bongiovani y Miguel, 2019). Entre los componentes fundamentales que han dado forma a la ciencia abierta, está el acceso abierto a las publicaciones académicas. En la cronología de ese acceso se destacan numerosas iniciativas y proyectos que fueron aportando en su desarrollo. El movimiento en favor del acceso abierto fue definido en una serie de encuentros en los que participaron investigadores, gestores científicos y bibliotecarios.

La [Declaración de Budapest](#) de 2001 (*Iniciativa de Budapest para el Acceso Abierto*), acordó que este modo de acceso implicaba disponibilidad gratuita sin restricciones de la literatura científica, con la condición de tener acceso a internet y de atribuir la autoría, manteniendo los autores el control de forma íntegra de su trabajo. Esta definición recomendaba dos estrategias: **publicar** en revistas de acceso abierto (*vía dorada*) y el **depósito** (archivo) en repositorios (*vía verde*). En la posterior cumbre de Bethesda se estableció que el acceso abierto requería a los autores otorgar derechos de uso sin limitaciones y el depósito de una copia de su obra en un repositorio. En la reunión de Berlín se añadió que el contenido y el software fueran de acceso libre y compatibles, siguiendo el consenso de la comunidad científica en cuanto al patrimonio cultural. Finalmente, en la cumbre de Salvador de Bahía, se abordaron las desigualdades en la producción y distribución del conocimiento, proponiendo el acceso a la información como derecho universal para el avance de la investigación y que los resultados de investigaciones financiadas con fondos públicos se dispusieran en acceso abierto. También se sugirió que los costos de publicación formaran parte de los costos de investigación y se promoviera la creación de revistas locales y repositorios (Salatino y Banzato, 2021).

En la *vía verde* del acceso abierto, se deposita la versión aceptada del manuscrito de un artículo en un repositorio abierto. Ese documento es el borrador final e incluye los cambios realizados como resultado de la revisión por pares, pero no es la versión publicada (el '*postprint*' posterior a la revisión por pares⁹). La versión publicada por una revista científica ('version of record') tiene la marca y el diseño de la editorial en su maquetación. El concepto de *acceso abierto diamante* surge para responder a la apropiación por parte las editoriales comerciales de la *vía dorada*, originalmente referida a las revistas científicas de contenido abierto, sin importar su modelo de negocio. Ahora está más asociado a un modelo de negocios de editoriales que, si bien permiten la lectura en abierto del artículo, cobran al autor unas tasas por publicar (más conocidas por las siglas **APC** de '[Article Processing Charges](#)' en inglés). A diferencia de la *vía dorada*, el *acceso diamante* describe las revistas académicas no comerciales, que no cobran ni al usuario ni al autor y son sostenidas por universidades u otras instituciones de investigación sin fines comerciales (Fuchs y Sandoval, 2013; Becerril-García, 2022), como las revistas de la *Universidad Nacional de Rosario*.



FIGURA 6 Acceso abierto diamante en las revistas de la UNR.

Fuente: Traducido y adaptado para el [portal de revistas UNR](#) de Paula Callan & Sarah Brown.

Lo cierto es que, actualmente, el panorama no deja de ser confuso, porque existen otros modelos de acceso (**figura 6**) como:

- *Acceso abierto híbrido*: Revistas basadas en suscripción que brindan la opción de pagar una tarifa para que la versión publicada de un artículo sea de acceso abierto en la web de la revista: APC. Sus costos son mucho más altos que los de las revistas de acceso abierto (Pinfield, Salter y Bath, 2016 y

⁹ El '*preprint*' es la versión anterior a la revisión por pares, que muchos editores no permiten subir a repositorios, pero que son populares en algunas áreas.

2017, Asai, 2023). Debido a que cobran por el acceso abierto híbrido sin reducir los precios de suscripción, este modelo es ampliamente criticado porque existe doble pago (autor/suscripción).

- *Acceso abierto bronce*: Son los artículos de investigación que una editorial pone a disposición de forma gratuita, ya sea inmediatamente o tras un período de embargo, sin licencia abierta que especifique su uso o reutilización. Por tanto, los artículos pueden colocarse en acceso restringido en cualquier momento. Esto no puede considerarse de acceso abierto.
- *Acceso abierto negro*: Formas no legales de acceso¹⁰.



FIGURA 7 Diagrama de Venn con los niveles de acceso abierto a las publicaciones en función del coste, la retención de derechos de autor y la revisión del contenido por pares.

Hay varios estudios que miden la cantidad de publicaciones que se están liberando en acceso abierto con diversas metodologías (Borrego, 2022). Piwowar et al. (2018, 2019) estudiaron el crecimiento del acceso abierto con datos de Unpaywall¹¹ y encontraron que en 2018 el 28% de los artículos y confirmaron la llamada "ventaja de citas de acceso abierto": los artículos en acceso abierto reciben un 18% más de citas en promedio.

¹⁰ Distribución ilegal de artículos científicos extraídos, generalmente, de revistas de pago a través de plataformas como [Sci-Hub](#) o [LibGen](#), que violan los derechos de autor, ofreciendo un acceso gratuito que sólo es posible vía suscripción.

¹¹ [Herramienta académica de acceso abierto](#) que permite encontrar versiones gratuitas y legales de artículos científicos que normalmente están detrás de una suscripción o compra individual.

También encontraron que en 2019 el 31% de todos los artículos estaban en acceso abierto y que acumulaban el 52% de las visualizaciones. Estimaron que para el año 2025, el 44% de todos los artículos estarán en acceso abierto y el 70% de las visualizaciones serán a esos artículos. Por lo tanto, los artículos cerrados ya no tendrán relevancia (Piwowar et al., 2019).

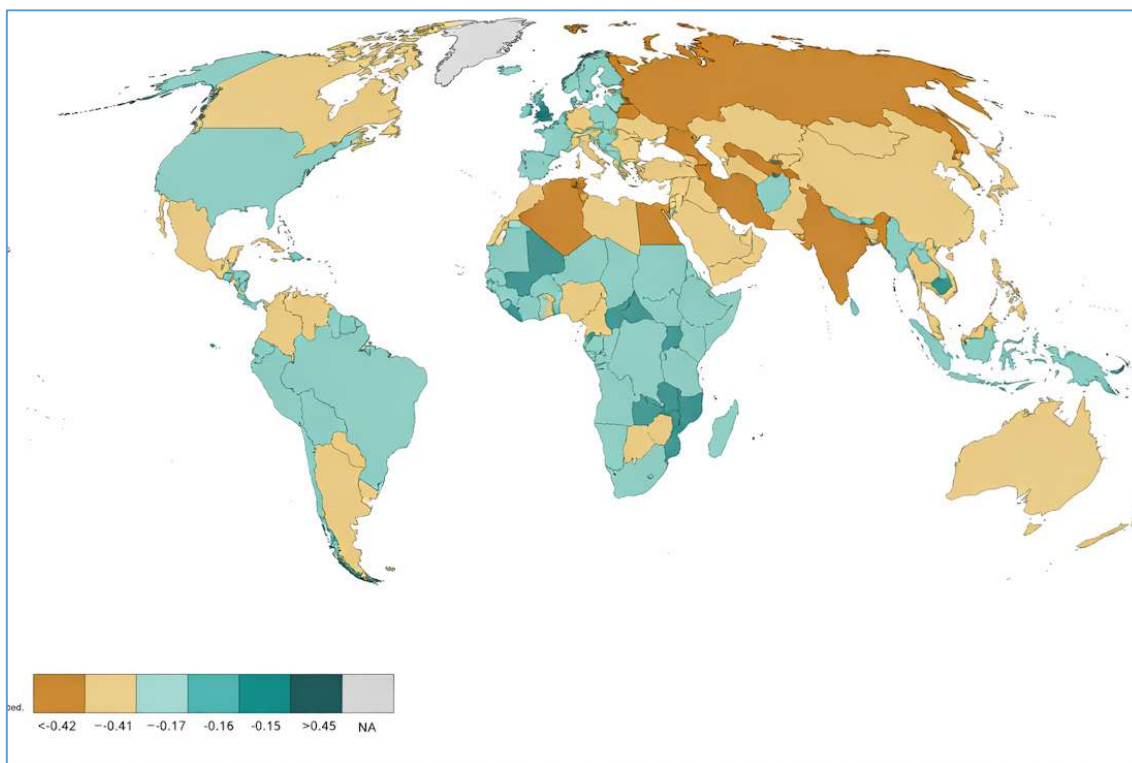


FIGURA 8 Medida *z-score* ponderada y mapa normalizado del número de publicaciones en acceso abierto por país. El rojo indica que un país está por encima del promedio mundial, el azul indica que está por debajo del promedio mundial. El blanco representa el promedio mundial. **Fuente:** (Simard et al., 2022).

Otros autores tomaron como caso de estudio 963 universidades de la edición 2019 del *Leiden Ranking*¹², analizando la disponibilidad en acceso abierto de 4,6 millones de documentos indizados en la [Web of Science](#) y usando [Unpaywall](#) para verificar el acceso (Robinson-Garcia, Costas y Van-Leeuwen, 2020). Encontraron que 77% de los documentos estaban disponibles en acceso abierto verde (repositorios), el 33% en dorado, el 20% en bronce y el 16% en híbrido, con solapamientos entre las vías.

Simard et al. (2022) hicieron un estudio donde analizaron más de 8 millones

¹² [Clasificación universitaria mundial](#) que evalúa el rendimiento de miles de universidades basándose en indicadores bibliométricos (publicaciones y citas científicas) y no en encuestas, proporcionando una visión multidimensional de la investigación, el impacto y la colaboración científica a través de datos objetivos de la [Web of Science](#) o, más recientemente, de datos abiertos (OpenAlex), destacando la calidad y el alcance de la producción investigadora de las instituciones.

artículos publicados entre 2015 y 2019 en [Web of Science](#), verificando en [Unpaywall](#) para ver las diferencias a nivel país en cuanto a publicación en acceso abierto, tomando como referencia el país del primer autor. Calcularon que el 42% de los artículos estaba en acceso abierto (18% en la *vía dorada*, 36% en la *vía verde*, 20% únicamente en repositorios y un 3,7% solo en revistas de acceso abierto), con variaciones según el área temática, siendo las ciencias naturales las que presentaban un mayor porcentaje. También estudiaron la cantidad de referencias que iban hacia publicaciones en acceso abierto y encontraron que el 39% de las referencias iban hacia documentos de acceso abierto, con más de un 30% hacia repositorios.

En su evaluación llevada a cabo después de cumplirse 20 años del acceso abierto, **Anglada y Abadal** (2023) reconocen que, a pesar de los notables avances en la adopción del acceso abierto, todavía estamos lejos de alcanzar la apertura total de la producción científica. No obstante, la dirección es clara hacia el acceso abierto; existen muchos actores involucrados, y será necesario observar cómo se configura este panorama.

Un grupo de grandes editoriales comerciales (*Elsevier, Springer Nature, Wiley y Taylor & Francis*) ha creado un oligopolio que domina más del 50% del mercado editorial (**Larivière et al.**, 2015), generándoles inmensas ganancias. Como resultado, el sistema de comunicación científica **sigue estando sujeto a barreras de pago cada vez más complejas**, lo que restringe el acceso y deja a investigadores de instituciones con bajos recursos sin la posibilidad de publicar, y en muchos casos, sin acceso a la literatura relevante.

A nivel de la Unión Europea, el [Plan S](#), liderado por la **cOAlition S** (2018), tiene como objetivo que todos los artículos científicos financiados con fondos públicos estén disponibles gratuitamente en internet, y busca eliminar o reducir los cargos de procesamiento de artículos. **Brainard** (2021) ofreció datos que permiten dimensionar el impacto potencial de este plan en la infraestructura de comunicación académica: en 2017, el 35% de los artículos publicados en la revista *Nature* y el 31% de los publicados en la revista *Science* fueron financiados por al menos uno de los miembros de esa coalición.

En términos concretos, este plan estableció la obligación para los investigadores europeos financiados por organizaciones e instituciones estatales de publicar sus trabajos en repositorios (*vía verde*) o revistas de acceso abierto (*vía dorada*) a partir

de 2021. Como contrapartida, se estableció el compromiso de otorgar fondos a los financiadores para cubrir los costos de procesamiento de los editores comerciales, es decir, respaldar el modelo comercial de acceso abierto.

Otras voces advierten que la implementación del [Plan S](#) alienta un cambio en las revistas académicas hacia modelos de publicación que imponen cargos por procesamiento de artículos, algo que efectivamente está ocurriendo. Se observa un peligro en la posible reducción de recursos destinados a la infraestructura pública no comercial, esencial para avanzar en los procesos y prácticas de ciencia abierta en América Latina (**Debat y Babini**, 2019). Según estos autores, será necesario esperar una evaluación de los resultados del plan para comprender sus implicancias en los países e instituciones menos privilegiados. Precisamente, en octubre de 2023, se encargó un estudio para evaluar el impacto del [Plan S](#) en el ecosistema global de comunicación académica.

Se está percibiendo que el sistema editorial se adapta al acceso abierto generando el citado oligopolio, concentrando la industria editorial científica en pocas empresas. Usando datos de la base de datos DOAJ¹³ se encontró que la editorial Springer Nature dominaba el 35% de los títulos en OA y [PLOS](#) más de 20% de los artículos (**Rodrigues et al.**, 2020). Se teme que si pagar por publicar es “la norma” para la publicación científica y los editores siguen subiendo sus precios, como se ha visto hasta el momento (**Asai**, 2021, 2023), solo puedan pagar por publicar en revistas consideradas prestigiosas los investigadores con mayor financiamiento, lo cual es un círculo vicioso porque el publicar en esas revistas les genera a su vez más financiamiento. Los investigadores de países con menos recursos no podrán acceder a estas publicaciones (**Brainard**, 2021). Cada vez hay más evidencia de que **las APC generan una nueva barrera** que impide a los investigadores que recién se inician y con menos recursos (países, instituciones, áreas de investigación), publicar sus investigaciones en acceso abierto (**Klebel y Ross-Hellauer**, 2023). **Haustein et al.** (2024) utilizaron un conjunto de datos abiertos con precios anuales de APC de las editoriales Elsevier, Frontiers, MDPI, PLOS, Springer Nature y Wiley, combinándolos con el número de artículos de acceso abierto de estos editores indexados en OpenAlex¹⁴. Con esta información, estimaron que a nivel global se gastó un total

¹³ [Directorio en línea global](#) y sin fines de lucro que indexa y facilita el acceso a revistas académicas y científicas de alta calidad, gratuitas y revisadas por pares, buscando aumentar su visibilidad e impacto, actuando como una “ventana única” para la investigación de acceso abierto.

¹⁴ [Catálogo abierto y gratuito de la investigación mundial](#) que indexa interconectadamente trabajos, autores, instituciones, financiadores y temas de investigación, similar a una biblioteca universal digital

de \$8,349 millones (\$8.968 millones ajustados a dólares de 2023) en APC entre 2019 y 2023. En el año 2023 se calculó que MDPI ingresó \$681.6 millones, Elsevier (\$582.8 millones) y Springer Nature (\$546.6 millones) fueron los principales generadores de ingresos por APC. Además, al ajustar por inflación, **se demostró que el gasto anual casi se triplicó**, pasando de \$910.3 millones en 2019 a \$2,538 millones en 2023. También observaron que las tarifas híbridas superaron a las de acceso abierto total (*'gold'*), y que las tarifas medianas de APC pagadas fueron superiores a las tarifas medianas listadas, tanto para las revistas de acceso abierto como para las híbridas. Estas estimaciones también indican la desconexión entre los precios de los APC y el costo real de la publicación. Es decir, la publicación en acceso abierto tiene costos más bajos (por ejemplo, a través del modelo de *acceso abierto diamante*), liberando así recursos de los presupuestos de investigación (Haustein et al., 2024).

Según la revisión de Borrego (2023), aunque existen revistas de acceso abierto diamante sin tarifas, la mayoría de los artículos abiertos se publican en editoriales comerciales que cobran APC, establecidas según el prestigio atribuido a la revista. Sin embargo, los estudios bibliométricos muestran que las revistas de pago tienen un impacto de citación similar al de otros modelos. Los investigadores suelen mostrarse contrarios al pago de APC, aunque esta actitud varía según la disciplina, la experiencia previa y la disponibilidad de financiación para cubrirlas. Además, las exenciones (*'waivers'*) se utilizan poco y pueden favorecer la incorporación de coautores únicamente para obtenerlas.

El *acceso abierto diamante* generado en países de América Latina con iniciativas como [Redalyc](#), [AmeliCA](#), [CLACSO](#) y [SciELO](#), basadas en la colaboración y sostenimiento por la comunidad, puede ser un modelo alternativo que considerar. Estas iniciativas han criticado fuertemente el modelo APC porque no se basa en los principios originales del acceso abierto sino en obtener ganancias y generar desigualdades (Babini, 2019; Becerril 2022).

Alperfn (2022), investigador y activista del acceso abierto, propone que la única solución sería eliminar del todo las APC, invirtiendo esos fondos en infraestructura, herramientas y servicios compartidos que puedan respaldar varias revistas

inspirada en la histórica [Biblioteca de Alejandría](#). Ofrece datos abiertos para la comunidad académica y el público.

simultáneamente, evitando así el desmantelamiento de un sistema diverso de acceso abierto. Otros autores van más allá y proponen buscar una solución disruptiva que deje de lado al artículo como forma de comunicación científica porque solo genera altos costos y no se adapta a la transformación digital y a la investigación basada en datos (Méndez Rodríguez y Sánchez-Núñez, 2023).

Al cumplirse 20 años de la *Iniciativa de Budapest* (2022), se realizaron recomendaciones que reflejan nuevas prioridades expresadas por la comunidad global en una consulta. Se plantea ahora que *"el acceso abierto no es un fin en sí mismo, sino un medio para alcanzar la equidad, la calidad, la utilidad y la sostenibilidad de la investigación"*. Para alcanzar estos fines se proponen cuatro recomendaciones prioritarias:

1. Alojar los resultados de la investigación en infraestructuras abiertas y controladas por la comunidad que minimice el riesgo de futuras restricciones de acceso o control por parte de entidades comerciales.
2. Reformar el sistema de evaluación de la investigación y de las recompensas para mejorar los incentivos positivos, eliminando obstáculos para el acceso abierto.
3. Favorecer los canales de publicación y distribución inclusivos de manera que nunca se excluya a los autores por motivos económicos.
4. Favorecer modelos controlados por instituciones académicas y organizaciones sin fines de lucro, que no apoyen intereses comerciales concentrados. No considerar los acuerdos de "leer y publicar".

1.7 Propuestas de cambios en el sistema de evaluación científica

En general, la calidad e impacto de la investigación científica y de los investigadores se evalúan por medio de métricas fundamentalmente basadas en contar publicaciones con un efecto conocido como "publicar o perecer" y el conteo de citas. Prevalecen indicadores como el Factor de Impacto (JIF por sus siglas en inglés), que informa de cuántas veces se cita un artículo en esa revista, llevando erróneamente a creer que, si un artículo se publica en una revista de alto JIF, entonces será de buena calidad, no siendo esto necesariamente cierto.

Por el contrario, si un artículo se publica en una revista de bajo impacto, esto no significa automáticamente que sea de baja calidad, es decir, se juzgan los artículos individuales por índice asignado a la revista (Paulus et al., 2018). Estos indicadores

obsoletos son muy criticados por no ser una medida de la calidad e impacto real de una investigación y generando una necesidad de cambio (**Leiden Manifesto** 2015; **Hicks et al.**, 2015). Son indicadores que contribuyen a una concentración desproporcionada de recursos, como prestigio, reconocimiento y financiamiento, en manos de un número limitado de instituciones e investigadores (**Delgado-López-Cózar et al.**, 2021).

En el año 2013 se firma la [Declaración de San Francisco sobre la Evaluación de la Investigación \(DORA\)](#), que marcó un importante movimiento en contra del uso del Factor de Impacto de las revistas y consiguió la firma de 24.096 personas y organizaciones en 164 países fomentando la implementación de buenas prácticas de evaluación. Han surgido nuevas propuestas que promueven esta evaluación responsable de la investigación, como los [Principios de Hong Kong](#) (2019), [SCOPE](#) (2021) que proponen considerar factos más cualitativos como prácticas de ciencia abierta, entre otros. En lugar de basarse en los resultados de investigación (publicaciones) y métricas cuantitativas, se trataría de valorar la labor del investigador y de los equipos de forma más completa.

En el ámbito latinoamericano también se está planteando esta discusión. Ya en 2019 se desarrolló el [Foro Latinoamericano sobre Evaluación Científica](#) (FOLEC-CLACSO) donde se ponen en discusión las políticas y prácticas de evaluación de las actividades científicas. El objetivo es fortalecer enfoques y modelos de evaluación que promuevan el conocimiento como bien común, la colaboración y la sostenibilidad de la ciencia de cara a los desafíos planteados en la región. Entre sus instituciones colaboradoras se encuentra [DORA](#), así como las más importantes iniciativas de ciencia abierta regionales. La [Declaración de San Francisco](#) y el [Manifiesto de Leiden](#) adoptan una perspectiva más global. En cambio, FOLEC, al enfocarse en América Latina y el Caribe, subraya la necesidad de personalizar la evaluación de la investigación para reflejar la rica diversidad cultural y lingüística de la región. FOLEC también resalta la importancia del impacto social en la evaluación de la investigación en América Latina y el Caribe, dando valor a los conocimientos y contribuciones de actores sociales fuera del entorno académico, así como de comunidades que históricamente han sido marginadas como las mujeres y las diversidades (**CLACSO**, 2022).

A finales del año 2022 se creó la [Coalición para el Avance de la Evaluación de Investigación](#) (CoARA por sus siglas en inglés), coalición internacional en la que

participan más de 700 instituciones relacionadas con la investigación (universidades, organismos de financiamiento y agencias de evaluación), así como asociaciones relacionadas a la investigación, sociedades científicas entre otras, cuyo objetivo es mejorar la forma en que se evalúan las investigaciones incorporando mecanismos que se focalizan en el impacto real de la investigación, incluyendo los componentes de la ciencia abierta, y no en el factor de impacto de las publicaciones. Se requieren cambios que tengan en cuenta las capacidades y talentos de los investigadores (Méndez Rodríguez y Sánchez-Núñez, 2023).

Los firmantes de [CoARA](#) se comprometen a tener una perspectiva común que implica que la evaluación de la investigación, los investigadores y las organizaciones de investigación debe abarcar una amplia gama de aportes, prácticas y actividades que maximizan la calidad y el impacto de la investigación. Este enfoque pone en el centro la valoración cualitativa del trabajo de investigación, con la revisión por pares como elemento central de este proceso. Se categorizan en dos grupos: compromisos principales y compromisos de soporte:

- **Compromisos principales:**

- 1) Valorar la variedad de aportes y trayectorias en el ámbito investigador: Este objetivo pretende extender el reconocimiento a una diversidad de actividades y prácticas investigativas en distintas disciplinas, así como valorar las contribuciones más allá de los artículos en revistas y los roles que incluyen gestión de datos, divulgación y comunicación de la ciencia.
- 2) Fundamentar la evaluación de investigación en valoraciones cualitativas: Se enfoca en valorar la calidad a través de la revisión por pares, con un apoyo medido de indicadores cuantitativos. Se subraya la importancia de que la revisión por pares siga principios de rigor y claridad.
- 3) Rechazar el uso indebido de indicadores basados en publicaciones: Se solicita explícitamente dejar de depender de métricas como el Factor de Impacto de las Revistas y el índice H para juzgar la calidad e impacto de los trabajos de investigación.
- 4) Prescindir de rankings en la valoración de la investigación: Se advierte contra el uso de rankings internacionales, que podrían no ser apropiados para la valoración de investigadores, para influir en la evaluación de la investigación.

- **Compromisos de soporte:**

- 1) Asignar recursos necesarios para la reforma de la evaluación investigativa: Esto comprende la asignación de presupuesto o personal necesario para

- realizar cambios, concienciar y apoyar la infraestructura necesaria para prácticas de evaluación transparentes.
- 2) Actualizar y perfeccionar los criterios y herramientas de evaluación: Implica la revisión de los métodos de evaluación de las unidades de investigación como de investigadores, promoviendo la interoperabilidad y asegurando que los criterios sean basados en evidencias y desarrollados en cooperación con la comunidad investigadora.
 - 3) Promover el conocimiento sobre la reforma de evaluación y ofrecer capacitación adecuada. Es crucial que las organizaciones comuniquen con transparencia sus métodos de evaluación y entrenen en ello a los investigadores y evaluadores.
 - 4) Compartir buenas prácticas y experiencias para fomentar el aprendizaje conjunto. Se espera que las instituciones intercambien prácticas y aprendizajes para facilitar un progreso coherente entre las distintas entidades.
 - 5) Informar sobre los avances en la aplicación de los compromisos y los principios: Las organizaciones deben mantenerse al tanto del avance entre ellas y estar dispuestas a recibir evaluaciones de sus comunidades.
 - 6) Basar la evaluación en evidencias sólidas y compartir datos abiertamente. El propósito es garantizar que las decisiones sobre la evaluación estén fundamentadas en evidencias concretas y que los datos estén disponibles para recopilar información y realizar investigaciones, compartiéndolos de manera abierta y respetando la protección de datos personales.

En [CoARA](#) se prioriza el avance colaborativo y a escala mundial, reconociendo que una sola institución o país por sí solo no tiene la capacidad de transformar el sistema. Sin embargo, se considera que cambios significativos son alcanzables si hay una coordinación y un compromiso conjunto de realizar acciones prácticas y realistas. Un ejemplo es el caso de España donde las instituciones de evaluación científica están transitando cambios en su forma de evaluación de acuerdo con los lineamientos de [CoARA](#) (CRUE, 2023). Recientemente, [CLACSO-FOLEC](#) se incorporó a la coalición como observador del consejo de dirección.

1.8 El acceso abierto en Latinoamérica

En América Latina, la investigación científica y la difusión de sus resultados están mayoritariamente respaldadas por recursos públicos. Esto significa que el Estado juega un papel fundamental en las actividades de investigación en la región. Un gran número de investigadores latinoamericanos trabajan en universidades públicas, y estas mismas instituciones son responsables de la publicación de la mayoría de las revistas académicas. Estas publicaciones suelen tener un alcance geográfico local o regional (Babini, 2019; Salatino, 2018).

A diferencia de otras regiones, en América Latina la mayoría de las revistas científicas se conciben y desarrollan en el modelo de acceso abierto *diamante*, derivado tanto del compromiso con la misión pública de las universidades en la región como del reconocimiento de la efectividad del acceso abierto para compartir conocimiento. Se sostienen principalmente con fondos públicos.

En el contexto de las revistas *diamante* a nivel mundial, según Bosman et al. (2011), de once mil revistas de acceso abierto *diamante* relevadas (Directorio DOAJ), 25% se publicaban en Latinoamérica (2810 revistas). Este estudio resalta la importancia de las revistas locales en América Latina, donde más del 75% informaron que al menos la mitad de sus autores son del mismo país que la revista. Al considerar las disciplinas, se destacan las revistas locales de acceso abierto en humanidades y ciencias sociales, medicina y revistas multidisciplinarias. Estas revistas publican en idiomas locales favoreciendo el multilingüismo en la comunicación científica.

Un estudio sobre cómo se sostienen las infraestructuras del acceso abierto no comercial encontró que el 70% de las revistas científicas en América Latina y el Caribe, España y Portugal recibe financiamiento del sector académico siendo un 54,6% de universidades públicas y 15,6% de universidades privadas. De aquellas revistas que reciben financiamiento público, este constituye el 93% del total de su presupuesto (Becerril García, 2021).

Solamente un 4,2% de las revistas de la región adoptaron el APC para sostenerse, cobrando en promedio 196 dólares norteamericanos y comenzando con esta modalidad en los últimos 10 años. Es importante destacar que un 73% de los directores de estas revistas realizan este trabajo como parte de su rol académico, entre otros como dictar clases e investigar. Solamente un 3% de estos directores

está abocado por completo a esa tarea. Otro aporte de este estudio es que se evidencia que la gran mayoría de las revistas utilizan software de código abierto para su gestión editorial.

En América Latina, las comunicaciones científicas son gestionadas por la misma comunidad científica, de forma colaborativa. Incluso antes de estar definido el “acceso abierto” como tal, la mayoría de sus publicaciones ya estaban disponibles de manera abierta en esta región. Desde hace décadas tres iniciativas reconocidas a nivel internacional, [Redalyc](#), [SciELO](#) y [Latindex](#) han desempeñado un papel fundamental en incrementar la visibilidad, la interoperabilidad y la calidad de las revistas científicas en la región (Alperín et al., 2011, 2015). Es decir, el acceso abierto surgió en Latinoamérica con anterioridad a otras partes del mundo, apoyado por las bibliotecas y bibliotecarios de la región (Gómez y Bongiovani, 2012).

Considerando las políticas y legislaciones sobre acceso abierto en este entorno, el depósito de resultados de investigación en repositorios digitales institucionales se establece como el método principal para llevar a cabo estas políticas. Entre 2013 y 2014, se promulgaron o modificaron leyes de acceso abierto en Perú¹⁵, Argentina¹⁶ y México¹⁷, estableciendo la obligación de difundir en acceso abierto los resultados de investigaciones financiadas con fondos públicos a través de repositorios (Babini y Rovelli, 2020). A su vez, se constatan más de 60 políticas institucionales de acceso abierto registradas en el Registro de Políticas de Archivo Obligatorio en Repositorios de Acceso Abierto ([ROARMAP](#)).

Siguiendo con la vía verde del acceso abierto, en la región se destaca el proyecto de red latinoamericana [LA Referencia](#) que ofrece servicios y apoya a organismos nacionales por medio de una plataforma común con estándares de interoperabilidad, cosechando y dando visibilidad a la producción científica generada en las instituciones de educación superior y de investigación científica. Mediante el establecimiento de [acuerdos técnicos](#) para compartir artículos con revisión por pares, informes, tesis de maestría y de doctorado, los nodos de

¹⁵ Ley que regula el repositorio nacional digital de ciencia, tecnología e innovación de acceso abierto. <https://portal.concytec.gob.pe/images/stories/images2013/portal/areas-institucion/dsic/ley-30035.pdf>

¹⁶ Ley 26899: creación de repositorios digitales institucionales de acceso abierto, propios o compartidos. <https://repositoriosdigitales.mincyt.gob.ar/vufind/Content/bibliografia>

¹⁷ Más información en <https://www.lareferencia.info/legacy/noticias/mexico-anuncia-legislacion-nacional-sobre-acceso-abierto-y-repositorios.html>

Argentina, Brasil, Chile, Colombia, Costa Rica, Ecuador, El Salvador, España, México, Panamá, Perú y Uruguay son miembros activos de la red. A través de su buscador se pueden acceder a más de cinco millones de textos completos.

[La Referencia](#) apoya la sostenibilidad de los repositorios de diversas formas como la generación de conocimientos para la actualización de softwares de repositorios, la generación de capacidades locales, la capacitación en diferentes temas relativos a repositorios (por ejemplo, curación de datos), entre otras iniciativas. A su vez, se generan alianzas y acuerdos con diferentes iniciativas de acceso abierto a nivel regional e internacional. La red es miembro de la [Confederación de Repositorios de Acceso Abierto \(COAR\)](#) y colabora con el proyecto europeo [OpenAIRE](#) que actualmente está cosechando los registros de la red (La Referencia, 2023).

Babini y Rovelli (2020) brindan un panorama completo de iniciativas y políticas de acceso abierto en la región que amplía lo aquí expuesto.

1.8.1 Conocimiento del acceso abierto en Argentina, México y Brasil

En el centro de la comunicación científica están los investigadores y evaluadores, quienes pueden generar un cambio en sus prácticas. Bongiovani y Gómez (2015) encuestaron a miembros de comités evaluadores de investigación en ciencias sociales de Argentina, Brasil y México. El resultado mostró que tenían conocimiento de iniciativas latinoamericanas de acceso abierto a revistas, especialmente aquellas de cada país. Por ejemplo, [SciELO](#) es más conocida en Brasil (ver [figura 8](#)).

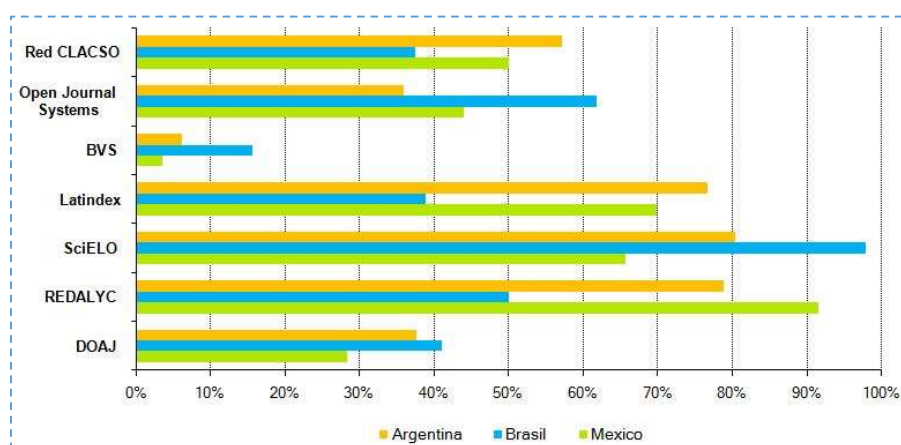


FIGURA 2 Iniciativas de Acceso Abierto conocidas en los tres países. **Fuente:** Bongiovani y Gómez (2015).

Al considerar variables relacionadas con el prestigio de las revistas de acceso abierto se percibió cierto desconocimiento, pero también datos positivos en

relación con ellas. En Argentina, alrededor de un 51%, consideraba falso que tuvieran menos prestigio que las comerciales. En Brasil, un 50%, consideraba falso que tuvieran un valor menor, mientras que un 28% de los investigadores de México lo consideraba verdadero.

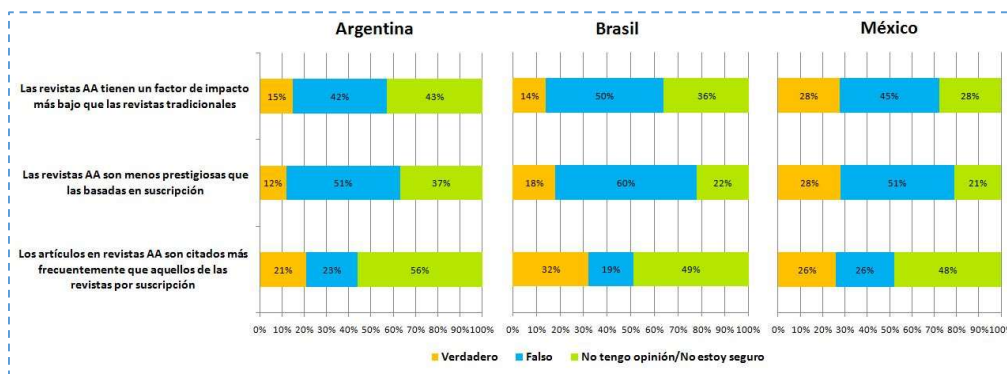


FIGURA 3 Conocimientos y opiniones sobre revistas de acceso abierto asociadas al prestigio en Argentina (n=410), Brasil (n=639) y México (n=208). Fuente: Bongiovani y Gómez (2015).

También se realizaron preguntas vinculadas a la evaluación de las revistas académicas. La primera se enfocó en los factores importantes en la elección que hacen los investigadores al seleccionar la revista donde publicar sus trabajos. La segunda se centró en cómo los investigadores evaluaban las publicaciones de sus colegas. Se encontró que el modelo de acceso abierto, a pesar de ser adoptado hasta llegar a ser ley en varios países, sigue siendo de menor relevancia entre los investigadores cuando publican y evalúan a sus pares.

La revisión por pares y el prestigio de la editorial fueron los aspectos más importantes a considerar. Dependiendo del país, el Factor de Impacto de la revista se consideró más importante en México y en Brasil que en Argentina.

Al existir un gran malestar y rechazo hacia las métricas cuantitativas de evaluación como el Factor de Impacto, es importante tener en cuenta estos estudios que dan un panorama sobre los conocimientos y opiniones de los evaluadores.

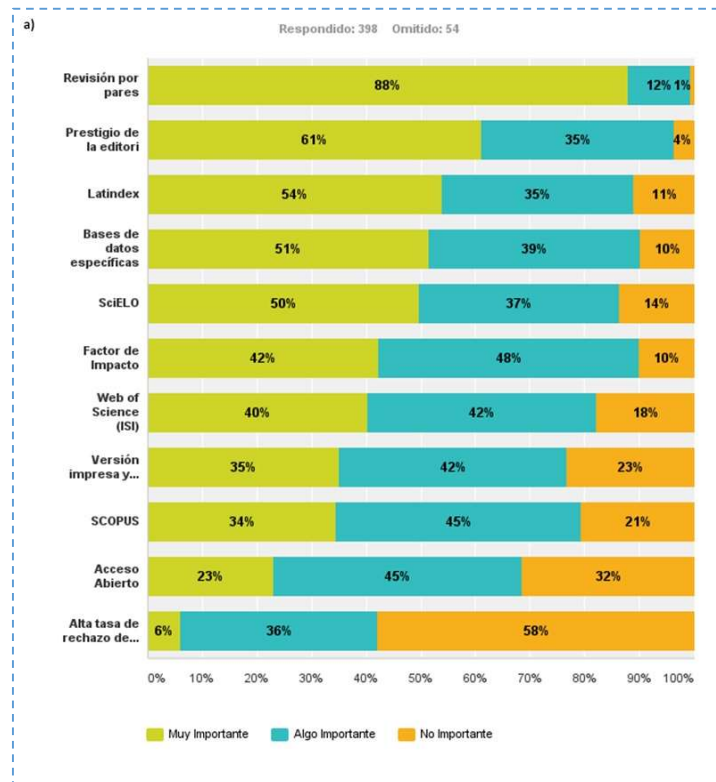


FIGURA 4 Factores más importantes en la evaluación de las publicaciones de sus pares, a) Argentina (n=398); b) Brasil (n=640) y c) México (n=262). **Fuente:** Bongiovani y Gómez (2015).

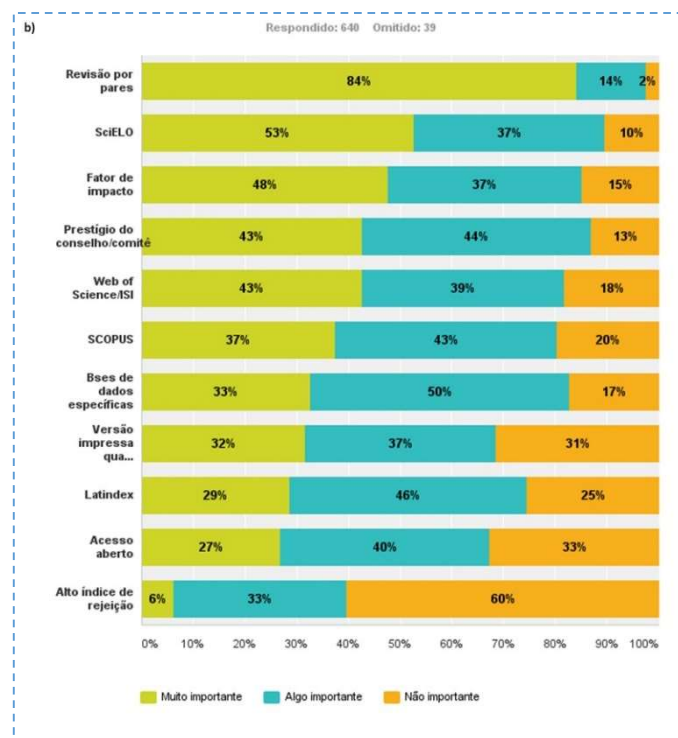


FIGURA 5 Factores más importantes en la evaluación de las publicaciones de sus pares, a) Argentina (n=398); b) Brasil (n=640) y c) México (n=262). **Fuente:** Bongiovani y Gómez (2015).

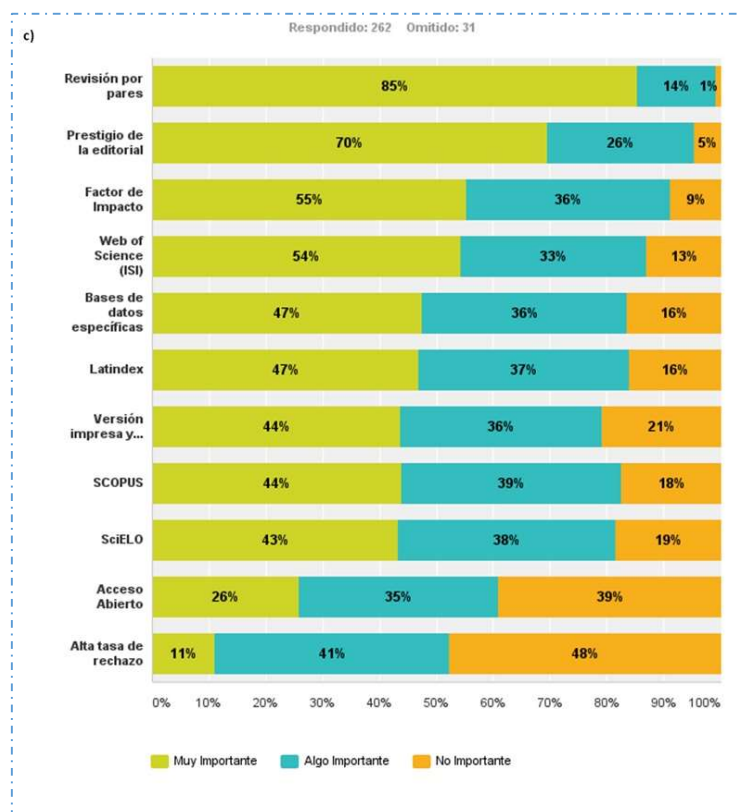


FIGURA 6 Factores más importantes en la evaluación de las publicaciones de sus pares, a) Argentina (n=398); b) Brasil (n=640) y c) México (n=262). Fuente: **Bongiovani y Gómez (2015)**.

1.8.2 El acceso abierto en Argentina

Las políticas de ciencia abierta en Argentina impulsaron fuertemente la *vía verde* del acceso abierto por medio de los repositorios. Se comenzó a gestar en 2009, cuando desde el [Ministerio de Ciencia, Tecnología e Innovación Productiva](#) (MINCyT) se convocó a los responsables de los repositorios digitales entonces en funcionamiento para desarrollar el [Sistema Nacional de Repositorios Digitales](#) (SNRD) en el marco del [Programa de Grandes Instrumentos y Bases de Datos](#).

El SNRD es una red interoperable de repositorios digitales en ciencia y tecnología, que parte del establecimiento de políticas, estándares y protocolos comunes a todos los integrantes del sistema y se apoya en su comité de expertos, para brindar apoyo y asesoramiento a las instituciones adheridas, para la elaboración de políticas institucionales y otras cuestiones relevantes al cumplimiento de los quehaceres del sistema. A través de un portal web brinda acceso a la producción científico-tecnológica disponible a través de los repositorios institucionales adheridos y también gestiona el portal de Datos Primarios en Acceso Abierto de la Ciencia y la Tecnología Argentina ([DACyTAR](#)). Un diagnóstico de la situación, así

como el proceso de articulación y coordinación institucional de los repositorios digitales y la propuesta de *Ley de acceso abierto* en 2010 han sido descritos por **Bongiovani y Nakano (2011)** y por **Lattuada y Giba (2014)**.

La [Ley N° 26.899](#) (2013 y reglamentada en 2016) de *Repositorios Digitales Institucionales de Acceso Abierto* establece la obligación de que la literatura científica y los datos primarios generados en investigaciones financiadas con fondos públicos estén disponibles en acceso abierto, mediante el uso de repositorios digitales gestionados y respaldados por las instituciones que forman parte del sistema científico y tecnológico nacional. Un grupo de investigadoras argentinas (**Miguel et al., 2012**) estudió la disponibilidad real y potencial en acceso abierto de la producción científica de su país entre 2008-2010 a partir de la base de datos de [Scopus](#), [DOAJ](#), [SciELO](#), [Redalyc](#) y otras fuentes de referencia. Encontraron que el 25% había sido publicado en revistas de acceso abierto, con un mayor porcentaje de revistas de Ciencias Sociales y Humanidades; y un 44% podía liberarse potencialmente porque las revistas permitían que se archiven en repositorios, lo cual se consideró positivo para promocionar la vía verde de los repositorios. La investigación de **Fushimi et al. (2021)** de los repositorios y políticas de 65 universidades públicas reveló que estas políticas, con enfoque en la creación y organización del SNRD y la legislación nacional de acceso abierto, habían tenido un impacto significativo, impulsando no solo la consolidación de repositorios existentes en las universidades nacionales, sino en el desarrollo de los repositorios institucionales del país. Desde 2003 que surgió el primer repositorio en el país hasta 2013, cuando se sancionó la Ley 26.899, se crearon 16 repositorios. Luego de este hito, se crearon 30 repositorios nuevos hasta 2019. De los 46 repositorios que encontraron en funcionamiento, 33 se habían adherido al SNRD y el resto estaban en proceso de inclusión. A su vez, se encontró que en general las universidades cuentan con un repositorio centralizado en 32 de 65 universidades relevadas, y además podían tener otros de facultades. En 31 universidades aún no se habían desarrollado repositorios institucionales. En el mismo estudio, se realizaron 40 entrevistas a gestores de repositorios quienes expresaron que, a pesar de los avances, las políticas de acceso abierto se desconocen en muchas áreas de las instituciones y que los docentes e investigadores consideran que depositar sus trabajos en los repositorios es una carga más de trabajo, pidiendo mayor apoyo político institucional. Sin embargo, elaboran diversas estrategias para superar los obstáculos incluyendo campañas de sensibilización (**Fushimi et al., 2021**).

En el año 2023, el SNRD brindaba acceso a 476.913 publicaciones de 48 repositorios cosechados en su [portal](#). Cabe aclarar que no todos los contenidos de los repositorios son recolectados en el portal (**figuras 1.11 y 1.12**), sino solamente aquellos definidos en las directrices del SNRD (2015) que fueron elaboradas por el [comité de expertos del sistema](#).

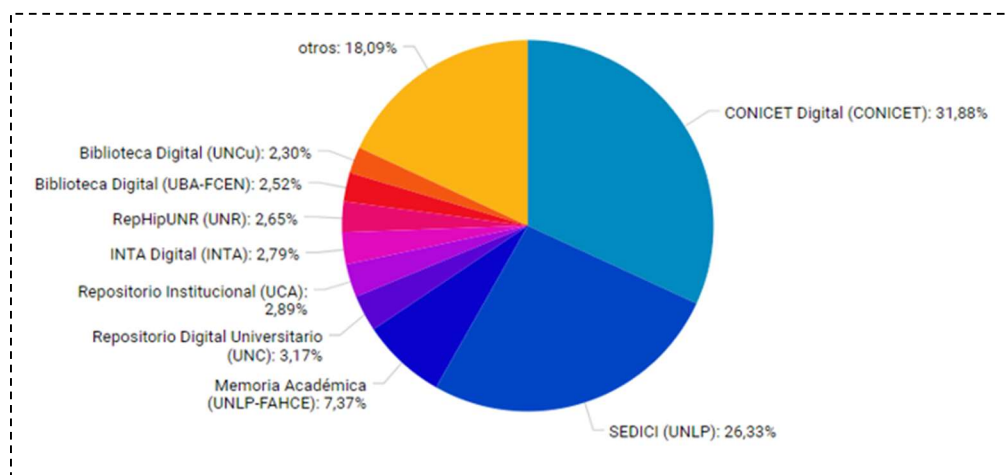


FIGURA 7 Porcentaje de publicaciones por repositorios. Fuente: SNRD.

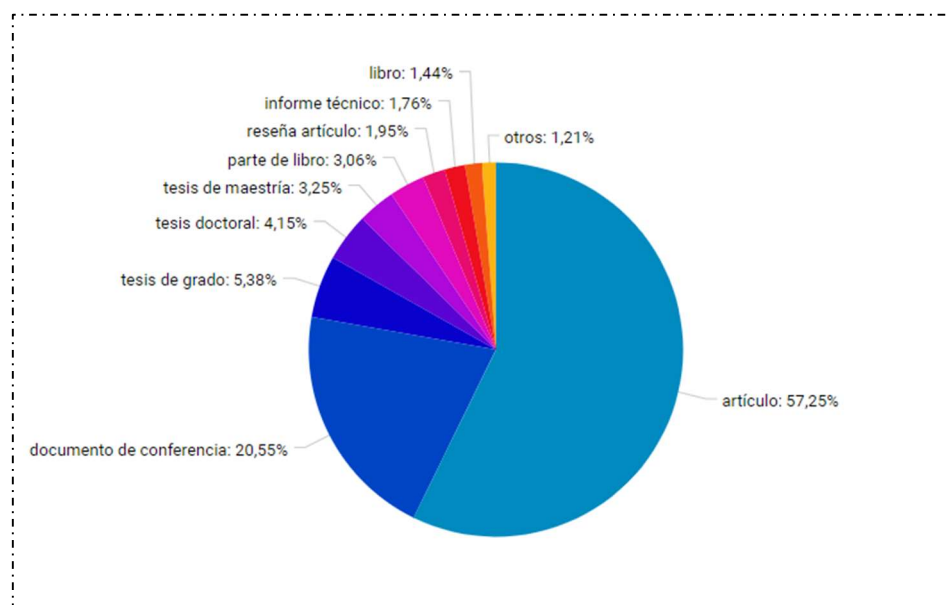


FIGURA 8 Porcentaje de publicaciones según tipo. Fuente: SNRD.

Existe otro cosechador, [SIU-BDU](#), desarrollado por el [Consorcio SIU](#) (Sistema de Información Universitaria) que recopila todos los tipos de documentos presentes en los repositorios nacionales con más de 2 millones de registros en 2023, incluyendo las revistas científicas de universidades que cuentan con protocolo de interoperabilidad OAI-PMH.

Un diagnóstico realizado por el **Comité Asesor en Ciencia abierta y Ciudadana del MINCYT** (2022) sobre las principales dificultades a nivel institucional para el cumplimiento de la Ley 26.899 incluye varios aspectos importantes. Entre ellos se destacan: la falta de reconocimiento y cumplimiento de la normativa por parte de las autoridades, el hecho de que muy pocas instituciones han desarrollado políticas institucionales completas, la prevalencia de acuerdos privados entre investigadores y editoriales, la limitada implementación de repositorios que, en general, no incluyen datos de investigación, la ausencia de Planes de gestión de datos (PGD) y de una gestión efectiva de datos de investigación, la necesidad de presupuesto y apoyo institucional, el requerimiento de un compromiso social y un cambio cultural hacia el acceso abierto, la colaboración intrainstitucional insuficiente, la falta de inclusión de requisitos de acceso abierto en convocatorias de subsidios, la necesidad de que las instituciones puedan reutilizar la información del [Sistema Integral de Gestión y Evaluación \(SIGEVA\)](#), la alineación de políticas editoriales con la normativa, y la promoción de la comunicación entre la comunidad científica y los repositorios.

También se puede ofrecer un panorama sobre las revistas de acceso abierto en Argentina. Al igual que en Latinoamérica, las revistas académicas son gestionadas por instituciones públicas, generalmente universidades. De acuerdo con el directorio del [Portal Latindex](#) (Sistema Regional de Información en línea para revistas científicas de América Latina, el Caribe, España y Portugal), existen en Argentina 1369 revistas electrónicas de acceso abierto, de las cuales 469 cumplen con los criterios de calidad y profesionalización estipulados en el Catálogo 2.0. En [DOAJ](#) encontramos que 378 revistas argentinas indizadas de las cuales 367 no cobran APC, la práctica totalidad.

El [Núcleo Básico de Revistas Científicas \(NBRA\)](#) (CAICYT-CONICET) está constituido por 372 publicaciones científicas y tecnológicas argentinas de acceso abierto consideradas de excelencia donde un gran número son de ciencias sociales y humanidades (274), seguido por ciencias biológicas y de la salud (49), de las ciencias exactas y naturales (29) y 20 revistas de las ciencias agrarias, de ingeniería y de materiales. Estas revistas deben cumplir normas editoriales internacionales, entre los cuales están los parámetros de calidad editorial regionales del [Sistema Latindex Catálogo 2.0](#).

Tras conocerse un convenio bilateral entre Argentina y la Unión Europea, en el cual Argentina entraría al [Plan S](#), [Bongiovani y Miguel](#) (2019) estudiaron, de forma exploratoria, la disponibilidad en acceso abierto de la producción científica argentina en ciencias sociales del año 2017 en la base de datos [Scopus](#), estimado también el pago realizado en APCs. El 61% de trabajos analizados estaban disponibles en acceso abierto, con un 95% de éstos publicados en revistas de acceso abierto sin APC, una proporción significativamente más alta que el promedio mundial. Concluyeron que cumplir con los principios del [Plan S](#) requeriría que los organismos de financiamiento en Argentina asumieran costos alrededor de 900 mil dólares anuales para un pequeño número de artículos y solo para la producción científica en ciencias sociales ([Scopus](#)). Esos costos no estaban contemplados en los financiamientos de investigación.

Más recientemente, [Cuartas et al.](#) (2022) incluyeron todas las áreas del conocimiento y analizaron la producción científica de los investigadores del [Consejo Nacional de Investigaciones Científicas y Técnicas](#) (CONICET), del total de 134.412 producciones publicadas en el período 2013-2020 analizadas (usando las bases de datos Lens, Google Scholar, [Web of Science](#), [Scopus](#), [Unpaywall](#) y [DOAJ](#)), se identificaron 73.271 artículos publicados en revistas con algún tipo de acceso abierto (54,51%). A su vez, se proyectó que el gasto potencial en APC para las publicaciones en las que investigadores argentinos aparecían como coautores, ascendía a más de 31 millones de dólares. Además, se estimó para ese mismo periodo, un gasto en APC de más de 11 millones de dólares para las publicaciones en las que los argentinos figuraban como autores principales. Estos datos generan preocupación dado que, como se ha mencionado antes, los precios de las APC están en crecimiento y se prevé una expansión de estas cifras.

Un portal adicional de CAICYT denominado [Malena](#) registra las políticas de las revistas en relación con el acceso abierto y archivo en repositorios, así como su categorización para la evaluación de publicaciones de investigadores de CONICET. En 2019 se relevaron 31 portales de revistas académicas en 22 universidades, en casi todos los casos usan el software libre [Open Journals System](#) (OJS) ([Fushimi et al.](#), 2019). Asimismo, existe una red nacional de portales de revistas científicas, [RedSara](#), que se organizó para colaborar e intercambiar experiencias entre las personas involucradas y con las problemáticas comunes de las revistas científicas bajo los principios de la ciencia abierta y el software libre. La red organiza capacitaciones, encuentros y genera actividades para compartir capacidades entre los portales miembros. Para avanzar en la ciencia abierta en Argentina se requiere

una significativa mejora en la infraestructura tecnológica y en la disponibilidad de recursos humanos. También se nota falta de sensibilización entre los investigadores sobre los beneficios de cumplir la Ley de acceso abierto y compartir también los datos de investigación. Uno de los factores tiene que ver con los incentivos y recompensas “rara vez las políticas de incentivos premian la difusión en abierto o apoyan las prácticas de ciencia abierta” (**Comité Asesor en Ciencia abierta y Ciudadana MINCYT, 2022**).

Si bien existe una línea de financiamiento para la creación y fortalecimiento de repositorios en el marco del SNRD, han sido pocas las instituciones que han aplicado y recibido financiamiento (**Fushimi et al., 2021**). Esto puede estar relacionado con que, en este financiamiento, se requieren desembolsos efectivos de dinero también desde las instituciones, sumado al personal necesario para gestionar los proyectos financiados. Para apoyar a las revistas editadas a nivel nacional y durante la Semana Internacional del Acceso Abierto de 2023 se lanzó el [“Programa Diamante”](#), una línea de financiamiento que tiene entre sus objetivos apoyar la edición y sostenibilidad de las revistas científicas, especialmente las que están en NBRA, a fin de generar capacidades, mejorar el equipamiento y promover la profesionalización de equipos editoriales. A su vez, esta iniciativa propone desarrollar el *“Portal Diamante: portal de publicaciones científicas argentinas de acceso abierto vía diamante”* en el marco del SNRD y la [Biblioteca Electrónica de Ciencia y Tecnología](#). Se espera que esta línea de financiamiento mejore la localización, visibilidad e impacto de las revistas a nivel internacional, ya que muchas revistas en las instituciones argentinas no cuentan con servicios de asignación del [‘Digital Object Identifier’](#) (DOI) que podría financiarse entre otras necesidades.

1.8.3 Los investigadores argentinos y el acceso abierto

Existen pocos estudios que dan cuenta de las opiniones y experiencias de investigadores argentinos en relación con el acceso abierto. **Bongiovani et al. (2014)** estudiaron las opiniones y prácticas de acceso abierto de los profesores de la [Universidad Nacional de Rosario](#) (UNR). Con respecto a las revistas de acceso abierto, **Bongiovani et al. (2012)** reutilizaron datos correspondientes a respuestas de investigadores argentinos a una encuesta internacional (SOAP) y encontraron que un 73% de los investigadores conocían revistas de acceso abierto en sus campos respectivos, con variaciones de acuerdo con las áreas: agricultura, ciencias

biológicas 84%, medicina (81%), física y astronomía (68%) y ciencias sociales y humanidades (61%). Un alto porcentaje (94%) consideraba que la publicación de artículos en acceso abierto era beneficioso para su área de investigación. Se evidenció que se decía publicar más en acceso abierto en Argentina, un 85% en promedio, que en el estudio internacional (Dallmeier-Tiessen et al., 2011).

San Martín y otros (2015) realizaron una encuesta sobre necesidades y prácticas de difusión en acceso abierto que respondieron más de mil docentes investigadores, principalmente de la UNR y también de la [Universidad Nacional del Litoral](#), la [Universidad Tecnológica Nacional](#) (distritos de Rosario y Santa Fe) y la [Universidad Nacional de Entre Ríos](#). Un 68% estuvo de acuerdo con que les gustaría que sus trabajos estuvieran disponibles de forma gratuita en línea y otro grupo que ya lo estaban compartiendo en acceso abierto (13%) sumaron más de 80% de actitud positiva hacia el OA, solo 3% no estuvo de acuerdo con difundir abiertamente sus trabajos.

El altruismo académico fue la principal motivación del 76% de los encuestados para ofrecer los trabajos en acceso abierto. Un 36% valoró el beneficio para sus estudiantes. Un 73% estaba motivado por el hecho de que sus trabajos fueran más utilizados por profesores, estudiantes y público en general (52%). Sin embargo, solo un 15% habían depositado en repositorios, con un alto porcentaje que desconocía la existencia de repositorios institucionales (44%) (**San Martín et al.**, 2015). Otra investigación en el ámbito de la Facultad de Filosofía y Humanidades de la [Universidad Nacional de Córdoba](#) (UNC) encontró que 61% de los investigadores tenía conocimiento de la existencia del repositorio institucional de la UNC, pero solamente el 38% había depositado trabajos en éste, por lo cual se consideró bajo el nivel bajo de uso del repositorio. Por el lado de las revistas de acceso abierto, se detectó que publicaban más en revistas tradicionales que de acceso abierto y que entre los investigadores nuevos había desconocimiento de revistas en acceso abierto de sus campos temáticos (**Solaro y Lamberti**, 2015).

Bongiovani et al. (2017) encuestaron a 1.086 evaluadores de la carrera científica en universidades de Argentina ([Programa de Incentivos a Docentes Investigadores](#), ahora PRINUAR), sobre sus actitudes y percepciones ante las revistas de acceso abierto. En general, los evaluadores seguían criterios tradicionales. Priorizaban el prestigio y la calidad de la revista y su editorial, seguidas de la pertinencia de la revista para la comunidad o región. En tercer lugar, consideraban el Factor de

Impacto, luego, la inclusión en índices internacionales y siguiendo los índices regionales. Una conclusión de este trabajo fue que los evaluadores aplicaban criterios similares para evaluar la producción de sus pares que los que usaban, en su rol de investigadores, al elegir dónde publicar sus propios trabajos.

El modelo de acceso abierto de las revistas no era un factor determinante en la publicación ni en la evaluación, con variaciones por áreas temáticas (ver **tabla III**). A pesar de existir una ley de acceso abierto, los criterios de evaluación de la producción científica no asignan puntajes a actividades de ciencia abierta el sistema universitario público argentino. Se encontró que los evaluadores que tenían experiencia en publicación en revistas de acceso abierto mostraron una actitud positiva hacia la valoración de esas revistas. Por lo tanto, la valoración positiva de este aspecto depende de los conocimientos y actitudes de los investigadores que forman parte de los comités de evaluación. El 82% de los evaluadores consideró que se debería incluir entre los criterios de evaluación el depósito de las publicaciones en repositorios abiertos, un 27% como obligatorio y un 55% como recomendado (Bongiovani et al., 2017).

La Revista es de Acceso Abierto	Evaluación			Publicación		
	Muy Importante	Algo Importante	No Importante	Muy Importante	Algo Importante	No Importante
Áreas temáticas						
Psicología	41%	47%	12%	47%	42%	11%
Arquitectura	32%	42%	26%	16%	53%	32%
Educación	31%	44%	25%	32%	48%	21%
Agronomía	26%	43%	31%	21%	49%	30%
Economía, administración y contabilidad	25%	28%	47%	21%	48%	30%
Derecho y jurisprudencia	22%	61%	17%	30%	43%	26%
Medicina, odontología y ciencias de la salud	22%	35%	43%	24%	35%	41%
Artes	21%	58%	21%	32%	58%	11%
Historia y geografía	21%	39%	39%	21%	50%	29%
Veterinaria	21%	32%	46%	13%	43%	43%
Literatura y lingüística	18%	42%	40%	26%	49%	36%
Antropología, sociología y ciencias políticas	17%	51%	32%	13%	57%	30%
Ciencias de la tierra, el mar y la atmósfera	14%	36%	50%	11%	38%	51%
Filosofía	14%	45%	41%	10%	38%	52%
Ingeniería	13%	31%	56%	10%	30%	59%
Física, astronomía y geofísica	9%	23%	68%	11%	26%	63%
Química, bioquímica y farmacia	8%	33%	59%	6%	40%	54%
Biología	7%	39%	54%	12%	35%	53%
Matemática	0%	35%	65%	0%	30%	70%

FIGURA 9 Importancia del acceso abierto en la evaluación y publicación por áreas temáticas.

Fuente: Bongiovani et al. 2017.

1.8.4 El acceso abierto en la Universidad Nacional de Rosario. Repositorio de publicaciones

El [Repositorio Hipermedial](#) de la UNR fue uno de los primeros en desarrollarse en el país; se puso en línea en 2008, dependiendo institucionalmente del [Campus Virtual UNR](#) y relacionado con investigadores en TICs en la educación (San Martín et al., 2010), a diferencia de otras universidades nacionales e internacionales, donde los repositorios suelen surgir en las bibliotecas. Además, la UNR fue una de las instituciones que conformaron el SNRD, y estuvo a cargo de la coordinación de este sistema nacional desde 2009 hasta 2011.

Fushimi (2018) profundiza en la implementación de este repositorio con un estudio cualitativo basado en entrevistas. Entre 2011 y 2013, un grupo de investigadores de la UNR, que incluía a la coordinadora del repositorio, participó en un [proyecto PICT-O](#) dedicado a los repositorios digitales de la región Centro-Este, incluyendo a la [Universidad Nacional del Litoral](#) y a las facultades regionales de Santa Fe y Rosario de la [Universidad Tecnológica Nacional](#). Se realizaron personalizaciones al software [DSpace](#) para facilitar el autoarchivo y la gestión de recursos educativos abiertos, se impartieron capacitaciones y se realizaron encuestas para comprender el uso de herramientas digitales, incluido el repositorio, por parte de los investigadores (Bongiovani et al., 2014; San Martín et al., 2015).

A pesar de estos avances, hasta 2016, el repositorio siguió sin estar formalizado institucionalmente, careciendo de un equipo designado, recursos mínimos, y una normativa institucional clara para definir su misión, funciones y políticas, que aparecían en la web, pero no con carácter formal institucional (Fushimi, 2018). Tras un proceso de reuniones, en 2016, el Consejo Superior de la Universidad aprobó la [Resolución N° 1842/2016](#), que formalizó el acceso abierto en la UNR. Allí, conforme a la Ley 26.899 se establece en su artículo 3 que los docentes, investigadores, becarios, no docentes, estudiantes de grado y posgrado deben depositar en el Repositorio Digital Institucional de Acceso Abierto, una copia de la versión final de su producción científica, tecnológica, artística y cultural resultante del trabajo, formación o proyectos financiados total o parcialmente con fondos públicos.

Esta resolución también introducía el requisito para los alumnos de posgrado de autorizar el depósito de su tesis o trabajos finales en el repositorio al presentarla. Este mismo documento estableció la creación de un Comité de Acceso Abierto

compuesto por las principales áreas de gestión de la Universidad, para supervisar la implementación de las normativas y proponer acciones para mejorar su cumplimiento. Además, se dispuso la formación de la Unidad de Gestión de Acceso Abierto, encargada de diseñar y gestionar las tecnologías, dispositivos y procedimientos técnicos relacionados con las iniciativas de acceso abierto. Una vez reglamentada la Ley 26.899 **MINCyT** (2016) que requería a las instituciones implementar una política institucional de acceso abierto, y dado que tenía especificaciones detalladas sobre los elementos que debía contener esta política, este comité decidió que era necesario ampliar la política. Es así como se aprobó la [Resolución de Rector N° 5425/2017](#) que contiene la política de acceso abierto de la universidad, las políticas del repositorio institucional y aprueba la licencia de depósito en el repositorio. Se registró la política en [ROARMAP](#).

En relación con el funcionamiento de la Unidad de Gestión de Acceso Abierto, la [Resolución Rectoral N° 2159/2018](#) estableció su misión, funciones y el organigrama. Dicha normativa indica que se establecerá una coordinación central para el repositorio, así como equipos de trabajo designados por las autoridades de las distintas facultades. Por ello, se generó la necesidad de que cada unidad académica designara a sus respectivos equipos y se organizaron capacitaciones para que pudieran asumir sus funciones. En este sentido, se puede resaltar el destacado trabajo de algunas bibliotecas y bibliotecarios en facultades que han desempeñado su función de manera altamente efectiva, logrando que este repositorio sea una parte integral y cotidiana de la vida de la comunidad académica.

En otras facultades, el trabajo en el repositorio lo asumieron diferentes personas ajenas al personal bibliotecario. Esto no suele sostenerse en el tiempo, con la consiguiente necesidad de capacitar continuamente en las tareas. A lo largo de la vida del [RepHipUNR](#) se han desarrollado numerosas capacitaciones, conferencias y eventos en relación al acceso abierto, que pueden consultarse en la subcomunidad de la [Unidad de Gestión de Acceso Abierto](#) y del [Comité de Acceso Abierto](#). En 2022 se depositaron 1621 documentos al repositorio. En la **figura 16** se puede ver la evolución de los registros disponibles en el buscador SIU desde agosto de 2019 a julio de 2023. También se presenta una evolución de los registros disponibles a través del portal del SNRD entre 2013 y 2023 (**figura 17**). Además del sitio en [DSpace](#), se cuenta con otro de ayuda para el repositorio en <https://recursos-rephip.unr.edu.ar/>. Para aumentar su visibilidad, los contenidos del [RepHipUNR](#) son cosechados por una variedad de buscadores de contenidos científicos entre los

que destaca [LAReferencia](#) donde, a octubre de 2023, se registraban 8856 documentos (artículos, tesis de maestría, tesis de doctorado, informes técnicos, libros y capítulos de libros).

A su vez, [LAReferencia](#) los envía al buscador europeo [OpenAire](#). Más de 16 mil registros de los contenidos del repositorio están visibles en uno de los motores de búsqueda más amplios del mundo, el [BASE](#).

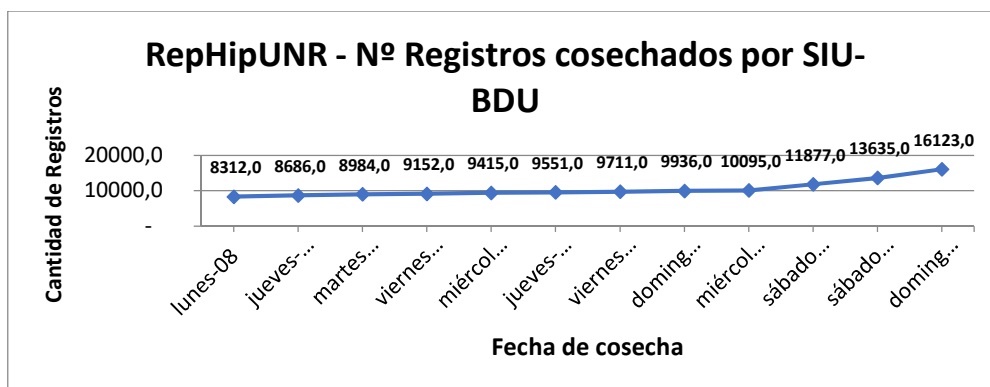


FIGURA 10 Evolución de registros en cosechador SIUBDU de Argentina.
Fuente: elaboración propia con datos del SIU.

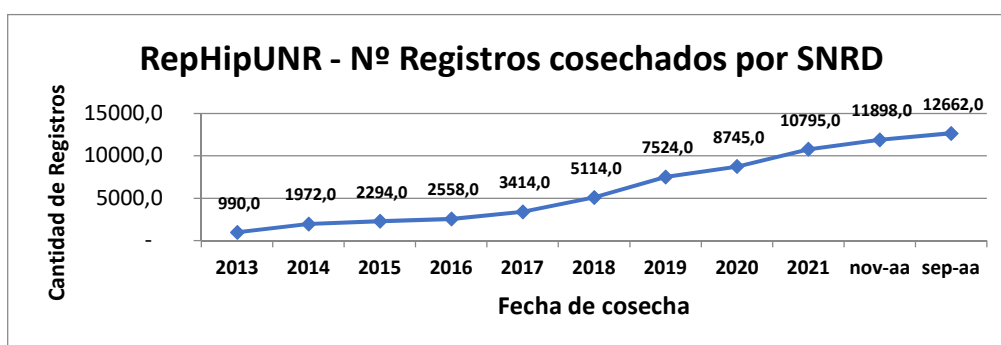


FIGURA 11 Evolución de cosechas desde el SNRD. Fuente: elaboración propia con datos del SRND.

De acuerdo con los datos de Google Analytics, el repositorio contó con 305.221 usuarios y 886.582 páginas vistas en 2022 (**figura 19**). En cuanto al software, se estaba trabajando en la actualización del repositorio en la migración desde la versión 4.7 de [DSpace](#) a la última versión disponible. Se están preparando capacitaciones de usuarios para que puedan usar la nueva interface de carga en el repositorio.

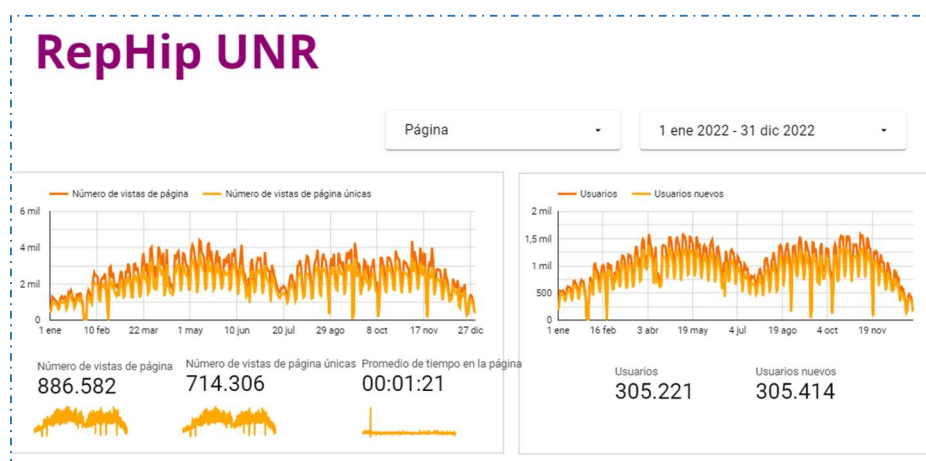


FIGURA 19 Vistas y Usuarios al RepHipUNR en 2022. **Fuente:** Google Analytics.

1.8.5 Portal Revistas UNR

En la UNR, desde el 2018 se vienen impulsando políticas institucionales que buscan mejorar y desarrollar las publicaciones académicas periódicas editadas oficialmente por las unidades académicas. Este esfuerzo tiene como objetivo alcanzar la excelencia académica y científica, así como aumentar la visibilidad y el reconocimiento tanto a nivel nacional como internacional. Además, se busca profesionalizar estas publicaciones para asegurar su calidad y relevancia. Todas las revistas de la UNR son de acceso abierto diamante. Antes de la existencia de estas políticas, cada revista trabajaba de manera independiente, muchas veces usando recursos de los propios investigadores para su sostenimiento. En muchos casos, pedían ayuda a la coordinación del [RepHipUNR](#) para diversas cuestiones relativas a las revistas como la obtención de ISSN, la visibilidad en el repositorio, entre otras. Se evidenciaba la necesidad de coordinar esfuerzos a nivel institucional.

En junio de 2018 se realizó el “Taller de requerimientos” para el [Plan de Fortalecimiento de Publicaciones Periódicas de la UNR](#) donde se relevaron, de forma participativa, aspectos de desarrollo considerados prioritarios para las revistas. Así, las necesidades relevadas se clasificaron en aspectos relativos a la tecnología (por ejemplo, necesidad de actualización del software OJS, de contratar identificadores DOI, software anti plagio), aspectos relativos a los recursos humanos, como contar con personal técnico capacitado, aspectos relativos a capacitación (sobre edición de revistas, indización en base de datos, manejo del OJS, entre otras), aspectos relativos a los derechos de autor (asesoramiento) y diferentes actividades relativas a la indización y visibilidad de las revistas.

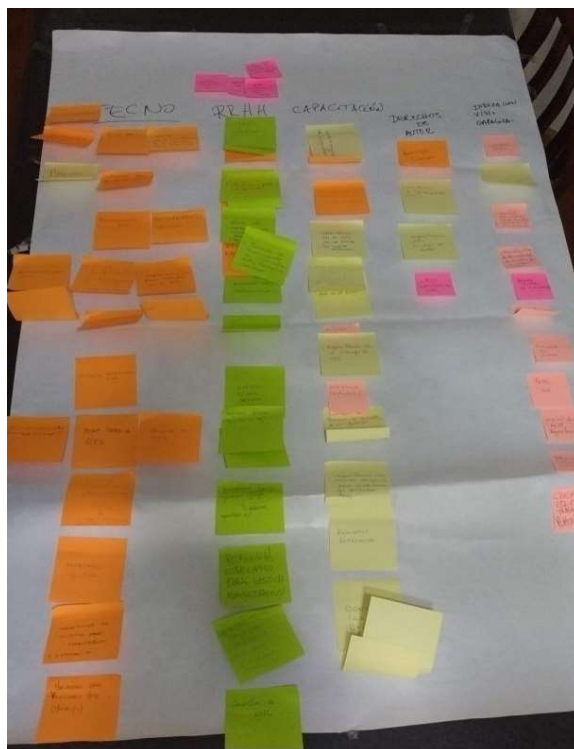


FIGURA 12 Clasificación de necesidades (Taller de requerimientos para el plan de mejoramiento de las publicaciones periódicas). **Fuente:** elaboración propia.

Para empezar a dar respuesta a esas necesidades, también en 2018, se realizaron dos jornadas de capacitación presencial destinada a equipos editoriales de revistas sobre temáticas de gestión, edición, visibilidad y evaluación de revistas científicas a cargo de las docentes [Cecilia Rozemblum](#) y [Carolina Unzuurrúnzaga](#), especialistas en edición científica de la [Universidad Nacional de la Plata](#). Los materiales de esa capacitación están disponibles para todos los editores en el campus virtual de la UNR¹⁸. El [Portal Revistas UNR](#) se inauguró el 18 de junio de 2020 en plena pandemia Covid-19 con una videoconferencia¹⁹ que contó con destacados especialistas a nivel nacional e internacional, pertenecientes a instituciones como [UNLP](#), [MINCyT](#), [CLACSO](#), [RedALyC](#) y [AmeliCA](#), explicaron a la comunidad los beneficios del acceso abierto diamante.

Antes de que comenzara el programa, había 7 revistas en el [NBRC](#)²⁰. En el año 2023,

¹⁸ Espacio en el campus virtual dedicado al Programa de Fortalecimiento de las Publicaciones Académicas Periódicas de la UNR <https://comunidades.campusvirtualunr.edu.ar/course/view.php?id=1983>

¹⁹ La grabación está disponible en <https://www.youtube.com/live/OAOJKuywMPw?si=hQ8-9nMXp86Hw6uV> y las presentaciones están disponibles en el RepHipUNR <http://hdl.handle.net/2133/18362>

²⁰ [Siglas de National Board for Respiratory Care](#), organización sin ánimo de Estados Unidos responsable de desarrollar y administrar programas de certificación profesional en cuidados respiratorios.

el "[Programa de Fortalecimiento de Publicaciones Académicas Periódicas](#)" de UNR contribuyó al mejoramiento significativo de las revistas, alcanzando un total de más de 50, de las cuales 21 son miembros del [NBRC](#) y están visibles en [diferentes bases de datos](#). Para participar las revistas tienen que cumplir ciertos requisitos relacionados con la pertinencia institucional y sostenimiento académico.

El programa, basado en las necesidades de la comunidad, incluye: el portal [Revistas UNR](#), provisión y mantenimiento de infraestructura y una instancia del software OJS para cada revistas que se actualiza permanentemente, asesoramiento integral, plataforma de trabajo colaborativa disponible en el [campus virtual](#) donde se pone a disposición materiales de capacitación, asesoramiento y foros con información relevante a la comunidad de revistas de la UNR, capacitaciones sobre diferentes temas requeridos, asignación de identificadores DOI a los artículos a través de [Crossref](#), provisión de software para prevenir plagio ([Similarity Check](#)), estadísticas, marketing, entre otros servicios. En la **figura 1.17** se pueden ver la evolución de las resoluciones mensuales de DOIs (clics en los DOI) de las revistas que participan en el portal de revistas.

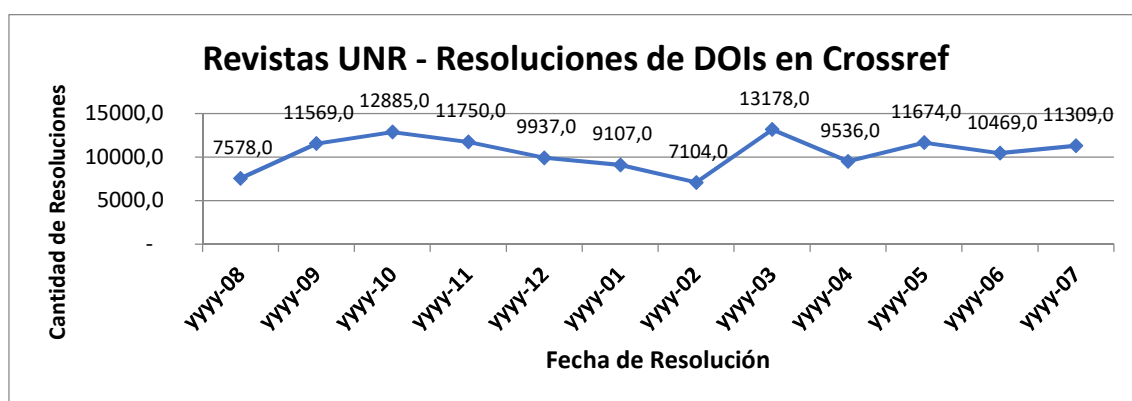


FIGURA 21 Resoluciones de DOIs de revistas de la UNR 2022-2023.

Fuente: elaboración propia en base a datos de Crossref.

1.8.6 Investigaciones sobre acceso abierto en la UNR

Bongiovani et al. (2014) realizaron una encuesta sobre las opiniones y prácticas relacionadas con el acceso abierto entre 783 docentes e investigadores de la UNR e identificaron necesidades en relación con el uso del repositorio institucional. Entre los resultados encontraron que aproximadamente el 80% de los docentes e investigadores de la UNR estaban a favor de la difusión de sus trabajos en acceso abierto. En ese momento, solo el 13% había usado el RI para compartir su

producción en acceso abierto, principalmente debido a que no sabían de su existencia (51,2%). El principal incentivo para que los docentes e investigadores utilizaran el RI sería el aumento en el uso de los trabajos por parte de sus colegas, estudiantes y el público en general. Además, se valoraban las funciones de búsqueda avanzada adaptadas a las necesidades individuales y los informes estadísticos de uso de los documentos proporcionados por repositorio.

En cuanto a la disponibilidad en abierto de la producción, **Bongiovani et al. (2020)** estudiaron la producción científica de la [Facultad de Ciencias Bioquímicas y Farmacéuticas](#) de la UNR, indexada en [Scopus](#) en el período de 2015-2019, usando las bases de datos [Unpaywall](#), [DOAJ](#), [SciELO](#) y los sitios web de las revistas. Además, intentaron calcular los costos asociados a la publicación en revistas de acceso abierto e híbridas que requieren el pago de APC. Se comprobó que el 52% estaba en acceso abierto: 18% en repositorios, 20% en revistas de acceso abierto, 9.7% en revistas bronce, 4.6% en revistas híbridas. Para 140 artículos fue necesario pagar APC, aunque no se contó con información sobre quién realizó el pago, con un costo promedio de \$1800 por artículo y un gasto total de 253 mil dólares estadounidenses, un dato preocupante principalmente porque los subsidios de investigación para proyectos en la Universidad no cubren este tipo de gastos.

1.9 Los datos de investigación

El único consenso entre la comunidad científica es que no existe una definición sobre datos satisfactoria para todos. El [National Institute of Health](#) define a los datos científicos como aquel material registrado y comúnmente aceptado por la comunidad científica como de calidad suficiente para validar y replicar los resultados de la investigación, independientemente de si los datos se utilizan para respaldar publicaciones académicas. También indican en su definición lo que no se incluyen en la misma (**NIH**, 2020).

Es indudable que los datos, como fuente primaria para respaldar la investigación científica y permitir realizar descubrimientos básicos, aplicados o teóricos, son productos fundamentales en los proyectos de investigación. Para que los estudios sean reproducibles y/o replicables, los investigadores y curadores de datos siguen las recomendaciones y buenas prácticas de datos FAIR, con el fin de que sean tan abiertos como sea posible, sin ignorar que en algunos casos es necesario restringir el acceso a los datos debido a limitaciones de seguridad, privacidad,

confidencialidad, entre otras restricciones. Los datos no son objetos naturales, sino representaciones de observaciones, objetos o entidades utilizadas como evidencia en la investigación. Su creación e interpretación dependen del investigador, el contexto y el momento, y suelen constituir un medio para elaborar productos académicos reconocidos, como artículos, libros o comunicaciones científicas (Borgman, 2016). Desde la filosofía de la ciencia, los datos constituyen una categoría relacional aplicada a los resultados de investigación que, en contextos específicos, sirven como evidencia para sustentar afirmaciones de conocimiento. Son una forma de expresar y presentar información producida o incorporada a las prácticas científicas, cuya relevancia depende de las circunstancias de uso. Por tanto, no poseen un valor de verdad intrínseco ni representan directamente los fenómenos, sino que son objetos fungibles definidos por su portabilidad y su utilidad potencial como evidencia (Leonelli, 2015).

En este sentido, la importancia científica de los datos depende del contexto en el que se utilizan: no son representaciones absolutas o directas de los fenómenos que representan y su valor y significado están relacionados con la situación o el contexto de la investigación en la que se aplican. Los datos, como objetos fungibles, pueden ser utilizados o reemplazados por otros datos similares o relevantes en función de la situación o necesidad y no perder su valor como evidencia científica. Los datos no son únicos e inmutables, sino que pueden ser sustituidos por otros que cumplan un propósito similar en un contexto particular de investigación. Por ejemplo, en estudios genéticos, la secuenciación genómica se puede realizar en diferentes equipamientos dependiendo del objetivo de un proyecto y de los recursos disponibles, generando a su vez diversos formatos de datos ([FASTA](#), [FASTQ](#) y otros).

Aunque los datos de investigación tienen un valor [epistémico](#), por su capacidad de contribuir al conocimiento y ayudar en la construcción de teorías científicas sólidas y fundamentadas como si fueran hechos objetivos, en realidad, son creados a través de complejos procesos de interacción entre los investigadores y el mundo. Estos procesos implican el uso de técnicas de observación, dispositivos de medición y manipulación de objetos de investigación. Incluso los datos generados mediante simulaciones computacionales o experimentos están influenciados por las interpretaciones y perspectivas de los investigadores (Knorr-Cetina, 1991; Leonelli, 2015). La pregunta «¿qué son los datos?» solo puede responderse atendiendo a situaciones de investigación concretas, en las que se decide qué objetos sirven

como evidencia para respaldar determinadas afirmaciones. Los datos deben poder circular entre distintos individuos, ya que la ciencia es una actividad social que exige la revisión y validación de los resultados por la comunidad científica. Sin esta posibilidad de comunicación y examen, no pueden funcionar como evidencia científica.

1.9.1 Tipos de datos

Según su forma de recolección, se distinguen:

- *Datos observacionales*: Datos recopilados por observación científica en diversas disciplinas, como estudios de seres vivos, la tierra, el clima y el universo. Para recopilarlos se utilizan instrumentos y suelen derivar de procesos únicos que no pueden reproducirse experimentalmente, por lo tanto, son insustituibles. Para analizarlos, se recurre a técnicas estadísticas, como el modelado longitudinal o la minería de datos. Además, estos datos pueden ser confidenciales en investigaciones sobre poblaciones humanas o de utilidad exclusiva para quien los recopiló. Su valor intrínseco varía según el campo de investigación específico. Algunos ejemplos incluyen observaciones climatológicas, estudios de comportamiento animal, observaciones de flora y fauna, datos médicos (registros clínicos con diagnósticos, tratamientos, etc.). La preservación de estos datos es importante porque son casi imposibles de replicar.

Ejemplo: Vanrell, Sebastián Rodrigo; Chelotti, José; Bugnon, Leandro; Rufiner, Hugo Leonardo; Milone, Diego; Laca, Emilio; Galli, Julio R., 2022, "Dataset: Audio recordings of grazing jaw movements in dairy cattle", <https://doi.org/10.57715/UNR/T7SDAX>, RDA UNR

- *Datos de teledetección*: La teledetección implica la recopilación de información sobre un objeto, área o fenómeno a distancia (desde drones, aviones o satélites). Estos datos pueden incluir diversas mediciones y observaciones de la superficie terrestre o la atmósfera, como temperatura, precipitación, cobertura terrestre, entre otros. Los investigadores utilizan datos de teledetección para estudiar y monitorear cambios ambientales, patrones climáticos y diversos fenómenos relacionados con la Tierra. Dado que implica la observación y medición de eventos y características del mundo real, se clasifica como *datos observacionales*.

Ejemplo: Cotlier, Carlos Gustavo; Brisco, Brian; Mondino, María Cristina; Grasso, Rodolfo; Vicioso, Benito; López, Diego Alejandro Germán; Cornero, Cecilia; Balparda, Laura Rita, 2022, "Dataset: Use of Radarsat-2 ultra-fine images in horticulture-intensive farming: land use detection and crop discrimination", <https://doi.org/10.57715/UNR/YRH5LP>, RDA UNR

- *Datos de encuesta:* recopilados a través de un método sistemático que implica hacer preguntas específicas a una muestra de personas o entidades con el objeto de obtener respuestas e información sobre una variedad de temas. Las encuestas pueden realizarse mediante cuestionarios escritos, entrevistas personales, encuestas telefónicas, en línea o de otras maneras, y son utilizadas en una amplia gama de campos, como la investigación de mercado, la opinión pública, la ciencia social, la salud pública y más.

Ejemplo: Miller, Ezequiel; Grandinetti, Rita María; Rodríguez, Ezequiel, 2022, "Respuestas de iniciativas de Gobierno Abierto en Municipios de Argentina", <https://doi.org/10.57715/UNR/UVRYFZ>, RDA UNR

- *Datos de entrevista:* La información o registros recopilados durante una entrevista realizada a una persona, o varias, como parte de un estudio o investigación. Suelen consistir en las respuestas a preguntas específicas. Las entrevistas pueden ser de varios tipos (estructuradas, semiestructuradas o no estructuradas) y pueden llevarse a cabo en persona, por teléfono, por videoconferencia o incluso a través de cuestionarios en línea.

Ejemplo: Gjerde, Vegard, 2022, "Student interviews on learning strategies and digitalization", <https://doi.org/10.18710/BEYEQR>, DataverseNO

- *Datos clínicos:* Datos relacionados con el diagnóstico de pacientes, datos demográficos, exposiciones, pruebas de laboratorio y datos de salud familiares. Se recopilan en el transcurso de la atención médica de rutina y no se manipulan deliberadamente como en un experimento controlado.

Ejemplo: Parodi, Roberto Leandro; Lagrutta, Mariana; Tortolo, Mauro; Navall, Estefanía; Rodríguez, María S.; Sasía, Gervasio Flavio; De Candia, Lucas F.; Gruvman, Matias A.; Bottasso, Oscar; Greca, Alcides Alejandro, 2022, "Replication Dataset for: A multi-center prospective study of 515 febrile neutropenia episodes in Argentina during a 5-year period", <https://doi.org/10.57715/UNR/KNSPG1>, RDA UNR

- *Datos bibliométricos*: Se utilizan para medir y evaluar la producción científica de un campo mediante el análisis de publicaciones, citas, colaboraciones y otros indicadores. Permiten valorar el impacto y la relevancia de la investigación, identificar tendencias y orientar decisiones sobre financiación y gestión científica.

Ejemplo: Malo de Molina Martín-Montalvo, María Teresa; Bernal Martínez, Isabel; Casaldàliga, Anna; Gómez Castaño, Javier; Guijarro Domínguez, Concepción; Losada Yáñez, Marina; Martínez Gándara, Santiago; Ortiz Uceta, Eva; Pascual del Pobol Valdenebro, Almudena; Puente Bujidos, Isabel de la; Ribes Llopes, Inmaculada; Rovira Fernández, Anna, 2021, "Medición del Acceso Abierto en las universidades españolas y el CSIC (2016-2020)", <https://doi.org/10.21950/2WY6YO>, e-cienciaDatos

- *Datos computacionales*: Datos obtenidos mediante simulaciones, modelos o cálculos informáticos para predecir resultados, analizar escenarios o representar procesos. Su reproducción exige documentar el hardware, el software y los datos de entrada; en algunos casos, también deben conservarse los resultados generados.

Ejemplo: Stenta, Hernán; Riccardim Gerardo; Pedro Basile; Carlos Scuderi, 2022, "Datos de: Modelación matemática hidrológica-hidráulica del escurrimiento superficial en la cuenca del A° Pavón (Santa Fe, Argentina)", <https://doi.org/10.57715/UNR/THC0KS>, RDA UNR

- *Datos experimentales*: Recopilados en experimentos controlados por los investigadores que manipulan una o más variables, observando sus efectos en otras. Implica la prueba sistemática de hipótesis bajo condiciones controladas. La cantidad de datos y documentación suficiente para reproducir el experimento varía según el costo y capacidad de reproducibilidad. Algunos ejemplos incluyen ensayos en los que se administra un nuevo medicamento a un grupo de pacientes experimental y un placebo a otro grupo de control para evaluar su efectividad, experimentos de física, como la prueba del impacto de diferentes fuerzas en el movimiento de objetos en un laboratorio.

Ejemplo: Shing, B., William, K., & Billington, S. (2009). NEES-2007-0422: seismic performance assessment and retrofit of non-ductile RC frames with infill walls. DesignSafe-CI <https://www.designsafe-ci.org/data/browser/public/nees.public/NEES-2007-0422.groups>

- *Registros documentales*. También denominados “datos secundarios”, se refieren a información o datos que han sido recopilados, registrados y documentados previamente por otros investigadores, organizaciones y entidades con fines diferentes a la investigación actual. Estos datos ya existen antes de que un investigador los utilice en su propio estudio y pueden incluir una amplia variedad de fuentes, como informes gubernamentales, registros comerciales, bases de datos académicas, encuestas previas, documentos históricos, entre otros.

Ejemplo: Alexander Adelaar (collector), Alexander Adelaar (recorder), 1987. Pak Nongeng tells about the first four people, food taboos, and various other topics. X-WAV/MPEG. AA1-001 at catalog.paradisec.org.au.
<https://dx.doi.org/10.4225/72/56E97A6DBA4F4>

En el siguiente capítulo se desarrolla un estado de la cuestión relativo a compartir datos de investigación y aspectos que pueden actuar como incentivos y barreras.

2. Compartir datos de investigación

La investigación científica actual está cambiando a un paradigma de ciencia y colaboración más intensiva en el uso de datos, tanto en cantidad como en variedad de estos. La complejidad y multidimensionalidad de los problemas científicos requiere poder acceder a nuevos instrumentos y tecnologías científicas, compartir los altos costos de estas tecnologías, compartir datos y conocimientos, desarrollar enfoques multidisciplinarios para encontrar respuestas satisfactorias a interrogantes globales tales como el cambio climático, la pobreza e inequidad, el agotamiento y la pérdida de biodiversidad, las pandemias, entre otros. De esta forma los datos primarios producidos por la investigación están cobrando una importancia y valor cada vez más importante. Se espera que compartir los datos pueda aportar múltiples perspectivas, ayudar a identificar errores o fraude científico, se usen para capacitar a nuevos investigadores, aumente el uso eficiente de los recursos y se evite generar datos duplicados.

2.1 El contexto internacional

En 2007, las recomendaciones de la OCDE definieron principios y directrices que promovían nuevas prácticas de gestión de datos de investigación, las cuales se esperaba que los países miembros implementaran para potenciar la reutilización de los datos. A partir de estas recomendaciones, la mayoría de las agencias de financiamiento científico internacionales exigen que los datos se gestionen y compartan. Diferentes organismos comprometidos con el proceso de producción científica, tales como organismos financiadores, gobiernos, universidades, editores y otras entidades, han impulsado políticas de gestión, acceso y reutilización de los datos (Piwowar y Chapman, 2010).

Neylon (2017) estudió las políticas de los financiadores de investigación en varios países y regiones, como Estados Unidos, Australia, Reino Unido, Unión Europea, Canadá y el Banco Mundial. Encontró que estas políticas presentan grandes similitudes: se exige la creación de planes de gestión de datos (PGD) en las solicitudes de financiamiento, y se espera que los datos que respaldan los artículos publicados sean accesibles. En muchos casos, se establecen requisitos para el almacenamiento y la disponibilidad de datos durante largos períodos, incluso después de finalizar las subvenciones. No obstante, también se ven diferencias significativas en la implementación de estas políticas. A menudo, se solicitan PGD

sin haber sido evaluados durante el proceso de revisión de las propuestas a financiar, lo que propicia percibir los requisitos como algo desconectado y consideren este requisito como una mera tarea administrativa a cumplir. Desde 2011, todas las propuestas de solicitud de financiamiento presentadas a la [National Science Foundation](#) (NSF) deben contener un plan de gestión de datos de investigación. Otras agencias de financiamiento como NASA, National Oceanographic Data Center (NODC) y National Institutes of Health (NIH), también contemplan la gestión y disponibilidad de datos resultantes.

A nivel europeo, el programa de financiamiento vigente desde 2021 hasta 2027, [Horizon Europe](#), profundiza en lo indicado en [Horizonte 2020](#), que ya exigía depositar los datos primarios y elaborar planes de gestión de datos en los proyectos financiados. Este programa va más allá del acceso abierto a publicaciones y datos al adoptar la ciencia abierta por defecto. Obliga a compartir los datos, como máximo, al publicarse los artículos, conforme a los principios FAIR y respetando los derechos de propiedad intelectual y las obligaciones de explotación, bajo el lema «tan abierto como sea posible y tan cerrado como sea necesario» (**European Commission**, 2021). Para garantizar su cumplimiento, combina obligaciones e incentivos, incluidas sanciones para quienes no proporcionen el acceso abierto requerido, responsabilizando a las instituciones de su producción intelectual. También incorpora métricas de nueva generación para evaluar el impacto de la investigación y el compromiso con la ciencia abierta.

La [Recomendación de la UNESCO sobre la Ciencia Abierta](#), aprobada en noviembre de 2021 y dirigida a los 193 Estados miembros, busca promover una comprensión común de la investigación abierta. Entre sus metas figuran crear un entorno político favorable; invertir en infraestructuras, servicios, recursos humanos y formación; impulsar una cultura de ciencia abierta mediante incentivos aplicables a todo el proceso científico; y fomentar la cooperación internacional y multipartita para reducir las brechas digitales, tecnológicas y de conocimiento. La recomendación establece un compromiso de actuación mundial, nacional, gubernamental e institucional.

Anger et al. (2022), entrevistaron a expertos de las principales agencias de financiamiento internacionales identificando seis retos vinculados con las políticas de datos de estas instituciones: el diseño de políticas claras, el monitoreo del cumplimiento, las sanciones por incumplimiento, los incentivos, el apoyo y las

limitaciones en las propias capacidades de los financiadores, junto con las acciones propuestas por las agencias para superar esos desafíos.

Desafío	Aspectos del desafío	Soluciones propuestas
I - Diseño de políticas y requerimientos para compartir datos.	– Falta de claridad. – Considerar los diferentes estándares de las disciplinas. – Necesidad de esfuerzos compartidos.	– Brindar guías y explicaciones detalladas. – Ofrecer resúmenes de mejores prácticas y aumentar conocimiento sobre las políticas. – Participar en redes de financiadores.
II - Monitoreo del cumplimiento de las políticas de compartir datos	– Falta de seguimiento. – Desaliento de los investigadores. – Falta de información. – Falta de recursos y capacidades.	– Recursos adicionales. – Formas alternativas como cheques automáticos. – Confianza en los investigadores.
III - Sanciones por no cumplir con las políticas de compartir datos	– Falta de mecanismos de aplicación concretos. – Renuencia a aplicar sanciones. – Complejidad de compartir datos. – ¿Cuándo y cómo aplicar sanciones?	– Prescripciones claras y condiciones de concesión del financiamiento. – Retener parte del financiamiento. – Considerar futuras becas. – Esperanza de cambio cultural.
IV - Incentivos para compartir datos	– Falta general de incentivos. – Dificultad para ofrecer incentivos concretos. – Pocas recompensas y reconocimiento por compartir. – Conservadurismo hacia el intercambio de datos.	– Cambiar las métricas de evaluación de subvenciones hacia compartir datos. – Mejor reconocimiento y aceptación. – Esperanza en el cambio cultural.
V - Apoyo y guía para compartir datos	– Financiamiento de infraestructuras para compartir. – Falta de información debido a la falta de retroalimentación y monitoreo. – Desconexión entre el nivel de política y los investigadores financiados. – Desconocimiento de los investigadores sobre el apoyo de financiadores e instituciones.	– Hacerlo más fácil para los investigadores financiados. – Apoyo práctico por parte de organizaciones de investigación. – Los financiadores ofrecen guías generales. – Los financiadores tratan de hacerlo “lo más fácil posible”.
VI - Límites en la capacidad de los financiadores	– Diferentes partes interesadas y estándares comunitarios. – Conflictos con otras partes interesadas y dentro de las agencias de financiación. – Dependencia de otras partes interesadas. – Limitaciones éticas y legales.	– Esfuerzos compartidos por las redes de financiadores. – Aprender de otras agencias de financiamiento. – Colaboración con revistas y organizaciones de investigación. – Estrategias gubernamentales sobre datos de investigación.

TABLA III Resumen de los desafíos percibidos por las agencias de financiamiento y enfoques de solución.

Fuente: traducción de Anger et al. (2022) bajo licencia [CC BY 4.0 Deed](#)

Una preocupación presente en las entrevistas a financiadores, desarrolladores de políticas y proveedores de infraestructuras (Neylon, 2017) era la necesidad de brindar apoyo de diversas maneras, incluyendo sensibilización, educación y capacitación, infraestructura y experiencia. Esto parte de la idea de que es necesario avanzar más allá de simplemente exigir acciones a los investigadores y enfocarse en respaldar un cambio cultural hacia compartir datos en general. Un tema recurrente fue la importancia de ofrecer el tipo adecuado de apoyo en el momento oportuno dentro del proceso de investigación, lo que permitiría integrar perspectivas que de otro modo podrían parecer divergentes. La preocupación central compartida fue encontrar un equilibrio entre imponer requisitos y fomentar un cambio orgánico dentro de las comunidades (Neylon, 2017).

En agosto de 2022, la [Oficina de Políticas Científicas y Tecnológicas](#) (OSTP) de la Casa Blanca emitió un memorando denominado “Garantizar el acceso gratuito, inmediato y equitativo a la investigación financiada con fondos federales” con el propósito de que las investigaciones financiadas por los contribuyentes estuvieran disponibles de inmediato para acceso público libre y sin restricciones (Nelson, 2022). Esta guía insta a las agencias federales a eliminar el embargo de 12 meses para acceder a los resultados —artículos y datos— que deberán depositarse en repositorios designados. Además, el memorando exige que los resultados estén en formatos legibles por máquina, facilitando su uso por investigadores, la minería de textos y datos, el análisis computacional y otras tecnologías.

Por otro lado, incluye una previsión para que los metadatos pertinentes (incluidos los nombres de los autores, las afiliaciones y las fuentes de financiamiento) estén disponibles en el momento de la publicación. Esto exige el uso de identificadores persistentes para los resultados de investigación, los investigadores y las becas. Además, permite a los investigadores incluir los costos asociados con el cumplimiento de las políticas de acceso público (SPARC, 2022). Estaba previsto que las agencias federales adopten estas políticas en 2025. La primera de las agencias federales de Estados Unidos en adoptar este memorando fue la NIH, que a partir de enero de 2023 sigue una política actualizada para gestionar y compartir datos de los proyectos financiados, exigiendo compartir datos con tiempos establecidos y repositorios recomendados, incluyendo los datos de investigaciones que resultaron negativas: drogas o vacunas no efectivas o con problemas de seguridad y efectos secundarios, entre otros.

2.2 El contexto argentino

En Argentina, la [Ley 26.899](#) (2013, 2016) establece que los datos primarios de investigación deberán depositarse en repositorios o archivos institucionales digitales propios o compartidos y estar disponibles públicamente en un plazo no mayor a 5 años del momento de su recolección, de acuerdo con las políticas establecidas por las instituciones. Se permiten excepciones en caso de que los datos primarios estuvieran protegidos por derechos de propiedad industrial, acuerdos previos con terceros, o aquellos que deban mantenerse en confidencialidad, con la justificación de los motivos.

Indart y otros (2021) exploraron la existencia de repositorios y portales de datos en Argentina hallando un total de 25 repositorios y portales, de los cuales 9 se enfocaban específicamente en publicar datos de investigación, mientras que los 16 restantes contenían un número pequeño de conjuntos de datos; 10 de estos últimos eran repositorios principalmente de documentos en universidades. En la realización de este trabajo, aparecieron dificultades en el relevamiento de datos, dado que muchos repositorios no estaban registrados en directorios de repositorios de datos, como [Re3data](#)²¹. Además, algunos de los repositorios disponibles en ese momento dejaron de existir y no figuran en el listado de [Sistemas Nacionales del MINCyT](#), como el [Portal de Datos de Biodiversidad](#) (Sistema Nacional de Datos Biológicos, SNDB), el Portal de Datos del Mar (Sistema Nacional de Datos del Mar, SNDM) y el Portal de Datos Genómicos (Sistema Nacional de Datos Genómicos, SNDG). A finales de 2022, solo doce repositorios institucionales recopilados por el SNRD que contaban con datos primarios entre sus contenidos y eran cosechados por el portal [DACyTAr](#). Se destacó que el [Repositorio de Datos Académicos RDA-UNR](#) era el único repositorio de datos de una universidad en Argentina, miembro del SNRD (**Comité Asesor en Ciencia abierta y Ciudadana MINCYT**, 2022).

2.3 Compartir datos de investigación

La idea de compartir datos generalmente se refiere al acto de ceder datos de forma que puede ser utilizada por otras personas y/o máquinas. Este acto incluye muchos

²¹ [Re3data.org](#) es un registro internacional de repositorios de datos de investigación que permite localizar dónde depositar o reutilizar datasets. Facilita el descubrimiento de infraestructuras de datos según disciplina, acceso y estándares, en el marco de la Ciencia abierta.

medios para intercambiar o publicar datos, pero dice poco sobre la usabilidad de estos datos (Pasquetto, Randles y Borgman, 2017). Como ejemplos de compartir datos en la ciencia se incluyen intercambios privados entre investigadores; publicar conjuntos de datos en sitios web de investigadores o laboratorios; depositar conjuntos de datos en archivos y repositorios, colecciones de dominio específico o colecciones de bibliotecas y adjuntar los datos como materiales complementarios en artículos de revistas, y como novedad, aparecen los artículos de datos ([‘data papers’](#)).

La forma más común de compartir datos es a través de solicitudes directas o como material suplementario en un artículo publicado (Joo & Peters, 2019). Las formas de intercambio varían según el dominio, el tipo de datos, el país, la revista, la agencia de financiamiento y otros factores. Como consecuencia, la capacidad de descubrir, recuperar e interpretar esos datos compartidos es variable.

En la Universidad Nacional de Rosario se observó que el formato impreso era el modo más común usado por los investigadores para compartir sus datos de investigación con otros (64%), compartiéndose también vía correo electrónico (54%), a través de dispositivos electrónicos portátiles (46%) y a través del editor de una revista (36%). Un 13% de los encuestados dijeron que otros investigadores no podían acceder a sus datos (Bongiovani y Martínez Uribe, 2014).

Desde 2008, y como una tendencia que se ha ido profundizando en los últimos años, han surgido revistas especializadas en datos ([‘data journals’](#)), que promueven la práctica en muchos campos de difundir un conjunto de datos a través de un [‘data paper’](#). García-García et al. (2015) realizaron una revisión bibliográfica identificando definiciones de [‘data paper’](#) como un nuevo tipo de comunicación científica: "publicación en una revista cuyo propósito es describir datos en vez de informar sobre una investigación o sus conclusiones" (Chavan y Penev, 2013).

Un artículo de datos describe una colección de datos, un proceso, un software, formatos de archivo, etc., sin el requisito de un análisis novedoso o de conclusiones innovadoras. Describe cuándo, cómo y por qué se recogieron los datos y en qué consiste el producto (Whyte et al., 2013). Los siguientes apartados desarrollan aspectos relativos a compartir datos de investigación en abierto, desde los requerimientos de las revistas, su cumplimiento e impacto, hasta llegar a las revistas de datos y a analizar incentivos y barreras.

2.4 Políticas de las revistas

Una de las formas más habituales en la que los investigadores comparten datos es como material suplementario a un artículo publicado (Joo & Peters, 2019). Las políticas de disponibilidad de datos de las revistas adquieren un rol destacado (Schmidt et 2016). Estas políticas establecen normas y pueden ofrecer incentivos directos para la publicación y citación de datos abiertos. El informe de 2022 de [Digital Science](#) encontró que los requisitos de las revistas o editoriales motivarían a compartir los datos a un 56% de los encuestados (Goode et al., 2022).

Castro et al. (2017) estudiaron las políticas de disponibilidad de datos en revistas de acceso abierto por medio de una muestra aleatoria de publicaciones en diferentes disciplinas registradas en [DOAJ](#). Encontraron una baja prevalencia de políticas y prácticas de datos abiertos en estas revistas de acceso abierto, en contraste con estudios previos sobre revistas comerciales en disciplinas específicas: el 74% de las revistas no tenía políticas de disponibilidad de datos, solo un 6% requería compartirlos, un 4% mencionaba la citación de los datos, y ninguna lo exigía. Un estudio similar se realizó con revistas editadas por instituciones de Brasil y Portugal. A diferencia del anterior, se detectó un mayor porcentaje de revistas (41,67%) con políticas de datos (de Oliveira Carvalho, 2016).

Asociaciones y sociedades científicas han creado guías para ayudar a las revistas en la adopción de políticas de datos. En 2014, se publicó la declaración "[Acceso de Datos y Transparencia de Investigación \(DA-RT\): Una Declaración Conjunta de los Editores de Revistas de Ciencias Políticas](#)," con el objetivo de mejorar la calidad en el compartir datos, la citación y las guías de replicación de investigaciones, impulsada por la [American Political Science Association](#) (APSA) y varios editores. Otro ejemplo son los "Lineamientos para la Transparencia y Promoción de la Apertura" conocidos como las **Guías TOP** (2015), impulsados por el [Centro para la Ciencia abierta](#), que contaba con 5,000 revistas y/o editores firmantes en 2019.

La primera editorial de acceso abierto en establecer una política de datos fue [Public Library of Science](#) (PLOS) en 2014, con el objetivo de facilitar la disponibilidad de los datos y promover la transparencia, liderando así el camino para que otras revistas adoptaran políticas similares. La política de PLOS (Bloom, Ganley y Winker, 2014) incluye su definición de datos que establece que deben compartirse, como mínimo, aquellos utilizados para obtener las conclusiones del manuscrito, junto con sus

metadatos, métodos y la información adicional necesaria para reproducir íntegramente los resultados. Los principales datos descriptivos, métodos y resultados deben figurar también en el texto del artículo, obviando expresiones como «datos no mostrados». Cuando su volumen impida su depósito en repositorios, los autores deben solicitar orientación a la revista. Se tiene una vista actualizada de las revistas con políticas de disponibilidad de datos a través de CHORUS²², organización que en su sitio web compila un [listado actualizado anualmente](#) con enlaces a los sitios web de las políticas de disponibilidad de datos.

En 2020, el grupo de estandarización e implementación de políticas de datos de la [Research Data Alliance](#) (RDA) planteó la necesidad de crear un [marco común para políticas de datos de investigación para todas las revistas y editores](#), con el objeto de evitar confusiones debido a requisitos contradictorios (Hrynaskiewicz et al., 2020). A partir de las políticas existentes, se identificaron 14 características de las políticas de datos de investigación, agrupadas en seis niveles destinados a fomentar buenas prácticas. Incluyen la cita y el depósito de datos, las declaraciones de disponibilidad, los estándares y formatos y la revisión por pares. También se ofrecen pautas de aplicación y textos para las instrucciones a los autores y los procesos editoriales.

Se están empezando a realizar estudios sobre el cumplimiento de las políticas de disponibilidad de datos en revistas, específicamente sobre editoriales y algunas disciplinas (Woods y Pinfield, 2021). Por ejemplo, en el campo de la biología evolutiva y puntualmente en el repositorio [Dryad](#)²³. Thelwall & Kousha (2017) encontraron que las revistas que firmaron la [política conjunta de archivo de datos JDAP](#)²⁴ presentaban un altísimo porcentaje de datos compartidos, concluyendo que en biología evolutiva, compartir los datos se había convertido en la norma.

²² CHORUS es una organización sin fines de lucro que facilita el acceso público y abierto a investigaciones financiadas con fondos públicos, conectando a editores, instituciones y agencias a través de una infraestructura técnica. La plataforma optimiza la búsqueda, identificación y preservación a largo plazo de artículos científicos, garantizando cumplir con las políticas de acceso abierto. Más información en <https://www.chorusaccess.org/about/about-chorus/>

²³ Repositorio digital internacional de acceso abierto y sin fines de lucro, diseñado para alojar datos de investigación científica y permitir que los investigadores compartan y preserven los conjuntos de datos de investigación que respaldan sus publicaciones académicas, facilitando que otros científicos validen, reutilicen y exploren sus análisis. Más información en <https://blog.datadryad.org/about/>

²⁴ La Política Conjunta de Archivo de Datos (Joint Data Archiving Policy o JDAP) es un acuerdo adoptado principalmente en 2011 por un grupo de revistas líderes en el campo de la biología evolutiva y ecológica. Su objetivo es exigir que los datos que respaldan las publicaciones científicas se hagan públicos a través de repositorios, generalmente utilizando [Dryad](#).

Zenk-Möltgen et al. (2018) analizaron las políticas de datos de revistas de sociología y ciencia política entre 2010/2012 y en 2016. En sociología, las revistas con políticas explícitas aumentaron del 71,3 % al 77,2 %, mientras que en ciencia política pasaron del 15,8 % al 44,2 %. En su mayor parte eran meras recomendaciones y pocas exigían compartir los datos o explicaban cómo se verificaba su cumplimiento. Se han realizado estudios con el objetivo específico de verificar la disponibilidad de los datos. **Roche et al.** (2015) analizaron 100 conjuntos publicados en revistas de investigación ecológica y evolutiva con una política fuerte de disponibilidad. Del total, el 56% estaba incompleto, carecía de datos o tenía metadatos insuficientes, y el 64% estaba disponible de una manera que impedía total o parcialmente su reutilización, ya sea por la falta de metadatos esenciales, la presentación de datos procesados en lugar de sin procesar, o el uso de formatos inadecuados (PDF). En definitiva, no cumplían con las políticas establecidas por las revistas. Una de las posibles razones de estos resultados es que los investigadores en estas áreas reciben poca o ninguna capacitación en gestión de datos y pueden no estar familiarizados con las mejores prácticas para el depósito adecuado de datos.

Haciendo uso de una metodología similar en el área de psicología, **Towse, Ellis y Towse** (2021) se encontraron con los mismos problemas de falta de completitud y calidad de los datos y metadatos. Propusieron algunas recomendaciones para las revistas entre las que se encuentran ofrecer pautas claras sobre cómo se espera que los datos estén disponibles y orientaciones prácticas (para mejorar la calidad de los datos, especialmente la interoperabilidad y la reutilización). Otro aspecto importante que destacar es *usar un repositorio de datos para el depósito y recibir asesoramiento*. En sicología los investigadores suelen no tener la capacitación suficiente para compartir sus datos (**Houtkoop et al.**, 2018).

Federer et al. (2018) evaluaron 47.593 artículos publicados en [PLOS One](#) con "declaración de disponibilidad de datos". Solo un 20% de estas declaraciones indicaba que los datos estaban depositados en alguno de los aproximadamente mil repositorios mencionados, siendo el depósito en repositorios el método preferido según la política para compartir datos. El resto de los artículos indicaba que los datos estaban en el propio artículo o en la información suplementaria, aunque no se especifica si esos datos cumplen con la política establecida por la revista.

En la revista *Cognición*, un estudio de series temporales (**Hardwicke et al.**, 2018) halló que el porcentaje de datos compartidos había aumentado del 25 % al 78 %

tras implementar una política de datos abiertos (en cuanto a datos reutilizables se pasó del 22 % al 62%). Los '[Open Data Badge](#)' son una insignia otorgada a los investigadores que comparten datos de investigación de manera abierta. Estas insignias, diseñadas por el Centro de Ciencia abierta (COS), tienen como objetivo fomentar prácticas de ciencia abierta, incluyendo los datos de investigación abiertos y se otorgan como un incentivo visible en las publicaciones. Para conocer un listado actualizado de las revistas que otorgan estos reconocimientos se puede visitar el sitio [TOP Factors](#) ('Transparency and Openness Promotion'). **Harper y Kim** (2018) realizaron un estudio basado en [la teoría del comportamiento planificado](#) analizando los factores que influían en las intenciones de psicólogos para compartir datos y adoptar estas insignias, concluyendo que los factores principales son: (1) los beneficios percibidos para la comunidad (compartir puede mejorar la productividad de la investigación y eso es positivo); (2) los beneficios académicos (la percepción de que compartir puede generar reconocimiento académico a través de citas, autoría y otros reconocimientos); (3) los riesgos académicos percibidos asociados con compartir datos (como la pérdida de oportunidades de publicación, que tuvo un efecto negativo) y (4) la norma de compartir datos (es decir, la creencia de que compartir datos es una práctica predominante y esperada dentro de la comunidad de psicología), la cual influyó positivamente tanto en la intención de compartir datos como en la adopción de insignias de datos abiertos.

Federer (2022) realizó un estudio extrayendo URLs y DOIs de un corpus de 50.000 artículos con declaraciones de disponibilidad de datos publicados en PLOS ONE entre 2014 y 2016. Trató de recuperar los datos asociados de manera automática y manual y halló que el 80 % se recuperaban automáticamente. Posteriormente, se comprobó manualmente una muestra de URLs y DOIs, siendo más alta la recuperación en el caso de los DOIs. Se observó que tener un DOI asociado a los datos y compartirlos en un repositorio, favorecía la disponibilidad y acceso.

La literatura señala que las políticas de disponibilidad de datos son más eficaces cuando se apoyan en infraestructuras y prácticas disciplinares consolidadas, evitan el uso de nuevas herramientas y ofrecen instrucciones claras (**Woods y Pinfield**, 2021). Colavizza et al. (2020), tras analizar 531.889 artículos de PLOS y BMC, comprobaron que aquellos con enlaces identificables a datos depositados en repositorios recibían hasta un 25 % más de citas, lo que puede incentivar su depósito.

Otro caso prometedor es el de la genómica humana, área pionera en el establecimiento de normas para compartir datos, comenzando con el [Proyecto Genoma Humano](#). Reales y Wallace (2023) señalan que la tasa en que se comparten estadísticas resumidas de [estudios GWAS](#)²⁵ sigue siendo baja, limitando el potencial de los avances científicos. Al analizar las citas de los datos del [Catálogo GWAS](#), encontraron que los artículos que comparten datos obtienen un 81.8% más de citas, efecto que se mantiene constante a lo largo del tiempo y que actúa como un incentivo para esa buena práctica, base de la ciencia abierta.

2.5 Revistas de datos ('data journals')

Asimismo, han surgido revistas especializadas en datos ('data journals'), que en muchos campos promueven la difusión de conjuntos de datos a través de 'data papers' (artículos de datos). La primera revista que comenzó a publicarlos fue [Earth System Science Data](#) (ESSD), publicación de acceso abierto que desde 2008 publica descripciones de conjuntos de datos en geociencias. Los conjuntos de datos, en sí mismos, se depositan en un repositorio confiable (Pampel y Dallmeier-Tiessen, 2014). García-García et al. (2015) realizaron una revisión bibliográfica en la que identifican definiciones de 'data paper' como un nuevo tipo de comunicación científica que consiste en una publicación en una revista cuyo propósito es describir datos en lugar de informar sobre una investigación o sus conclusiones.

Un 'data paper' es una publicación académica de un documento de metadatos que permite realizar búsquedas y describe un conjunto de datos accesible en línea en particular, o un grupo de conjuntos de datos, publicado de acuerdo con las prácticas académicas estándar" (Chavan y Penev, 2013). Para Whyte et al. (2013) es un artículo que describe una colección de datos, un proceso, un software, formatos de archivo, etc., sin requerir un análisis novedoso o conclusiones innovadoras. Describe cuándo, cómo y por qué se recogieron los datos y en qué consiste el producto

2.6 La opinión de los investigadores: incentivos y barreras

Zuiderwijk, Shinde y Jeng (2020) identificaron 11 variables que pueden incentivar o inhibir a los investigadores a compartir sus datos en abierto con otros, así como los

²⁵ Estudios en los que se comparan los marcadores de ADN en todo el genoma o material genético completo de una persona.

aspectos que pueden motivar o desmotivar a los investigadores a utilizar datos de investigación abiertos compartidos por terceros:

1. **Antecedentes del investigador:** Aspectos sobre características personales del investigador y su trayectoria de investigación que podrían afectar su comportamiento respecto al uso y al intercambio de datos abiertos.
2. **Requisitos y obligaciones formales:** Políticas de agencias de financiamiento y otros requisitos, como políticas informales.
3. **Motivaciones personales:** Motivaciones intrínsecas tanto para compartir como para utilizar datos de investigación abiertos.
4. **Condiciones facilitadoras:** Factores que pueden facilitar compartir o usar datos de investigación abiertos.
5. **Confianza:** Cómo el nivel de confianza del investigador influye en su disposición a compartir y utilizar datos en abierto.
6. **Desempeño esperado:** Factores que pueden influir en el desempeño de los investigadores al compartir y utilizar datos abiertos de investigación.
7. **Influencia social y afiliación:** Aspectos relacionados con la influencia social y la afiliación que impactan en la decisión de un investigador para compartir y utilizar datos abiertos de investigación.
8. **Esfuerzo:** El esfuerzo necesario para que un investigador comparta o utilice datos de investigación en abierto.
9. **Experiencia y habilidades de los investigadores:** La experiencia previa del intercambio y el uso de datos abiertos, así como las habilidades necesarias para realizar estas actividades, y cómo esto afecta el comportamiento futuro en cuanto al uso y compartición de datos.
10. **Legislación y regulación:** Normas legales y reglamentarias que influyen en el intercambio y uso de datos.
11. **Características de los datos:** Cómo las propiedades de los datos afectan a la disposición de un investigador para compartir y utilizar datos abiertos de investigación.

2.7 Incentivos para compartir y usar datos de investigación

Está reconocido que compartir los datos abiertamente es fundamental para acelerar la investigación científica, aumentando su eficiencia, preservando los datos para que puedan ser reutilizados, además de mejorar la capacidad para resolver problemas globales y específicos. Al tener datos disponibles, las agendas

de investigación pueden orientarse mejor hacia la resolución de problemas detectados con opciones más económicas. Los investigadores pueden ayudar a que los problemas locales sean más visibles y mejor comunicados con datos confiables, permitiendo que otros actores puedan ofrecer sus aportes para desarrollar mejores soluciones (**Arza y Fressoli, 2017**).

La encuesta realizada a investigadores de la UNR encontró que entre los factores que más podrían motivar a compartir datos a través de repositorios digitales serían: permitir colaboraciones y contribuciones de otros, beneficios potenciales para la comunidad de investigación, beneficios potenciales para la sociedad, beneficios demostrables para la institución de pertenencia y beneficios demostrables para la carrera del investigador (**Bongiovani y Martínez Uribe, 2014**).

Navarro-Molina y Melero (2019) revisaron la literatura para identificar motivaciones, barreras e incentivos de los investigadores, así como hábitos y actitudes para compartir y reutilizar datos de investigación. Entre los beneficios están: la promoción de la innovación y creación de nuevos servicios, la facilidad de la colaboración multidisciplinaria, la mejora en la transparencia y confiabilidad de los datos, mejora en la reproducibilidad de resultados y reducción en la duplicación de trabajos. Por otra parte, se establecen claramente las condiciones de uso a través de licencia, los datos son integrados al ciclo de vida de la comunicación científica y permiten analizar su impacto, además de dar visibilidad a las investigaciones, los ciudadanos pueden participar usando y generando datos y se logra mayor rendimiento de los fondos invertidos en investigación.

Chawinga y Zinn (2021), en una revisión sistemática de la literatura, encontraron un alto número de estudios que destacan que compartir datos impulsa el avance de la ciencia, ya que permite que otros investigadores reanalicen e interpreten esos datos utilizando nuevos enfoques o técnicas para generar nuevos conocimientos o avances. Además, compartir datos minimiza los costos de investigación, protege contra el fraude científico al mantener la integridad de los resultados, contribuye a la formulación de políticas que mejoran las economías a nivel nacional, regional y global, y ahorra tiempo a los investigadores.

Sin embargo, compartir los datos en abierto todavía no es la práctica común de la mayoría de los investigadores. Por este motivo se requieren estrategias e incentivos para promover cambios culturales. Se realizaron estudios para detectar

qué factores motivan a los investigadores a compartir sus datos en abierto y se describen los mismos teniendo en cuenta las variables que identificaron **Zuiderwijk, Shinde y Jeng** (2020) con énfasis en la importancia de estos factores. Si tomamos la variable del desempeño esperado, el citado informe de 2022 de *Digital Science* sobre una encuesta respondida por 6798 investigadores en el mundo (**Goode et al.**, 2022) encontró que obtener citación de sus artículos de investigación, como forma de reconocimiento y un mayor impacto y visibilidad de sus artículos fueron de los factores que más motivan a los investigadores a compartir datos. Este factor clave de la citación formal se evidenció en estudios anteriores (**Tenopir et al.**, 2011; **Kratz y Strasser**, 2015; **Tenopir et al.**, 2015; **Fecher et al.** 2017; **Zuiderwijk y Spiers**, 2019). La mayor visibilidad del investigador también es un factor de motivación (**Piwowar, Day y Fridsma**, 2007; **Piwowar y Vision**, 2013).

Otras formas de reconocimiento del creador de los datos son la coautoría en publicaciones, reconocimiento formal de los proveedores de datos (ejemplo una insignia de una revista) y la oportunidad de colaborar con otros (**Sayogo y Pardo**, 2013; **Zuiderwijk y Spiers**, 2019). Compartir los datos en acceso abierto brinda oportunidades de participar en nuevos proyectos de investigación y establecer contactos de diversas disciplinas, además de crear oportunidades de colaborar con académicos con intereses similares y obtener datos para avanzar su propia investigación (**Fecher et al.**, 2015).

Si se dispone de estructuras de reconocimiento adecuadas, ya sea institucional o profesional, los investigadores podrían estar más impulsados a compartir abiertamente sus datos (**Navarro-Molina y Melero**, 2019). En este sentido la Comisión Europea impulsa la adopción de una [Matriz de Evaluación de Carreras de Ciencia abierta](#) (OS-CAM) con criterios de evaluación para las actividades de ciencia abierta que incluye usar los principios de datos FAIR, adoptar estándares de calidad en la gestión de datos abiertos y usar datos abiertos de otros investigadores (**European Commission**, 2017).

Los [datos abiertos](#) permiten que los investigadores muestren sus propios logros y puede considerarse un aspecto de profesionalidad y prestigio, es decir, basarse en los códigos de conducta y ética de la comunidad científica. Este profesionalismo les permite publicar en revistas de alto impacto. Entre los antecedentes del investigador que lo motivan a compartir sus datos se encuentran las prácticas disciplinarias y la cultura de compartir datos (**Fecher et al.**, 2015; **Tenopir et al.**, 2011;

Curty et al., 2017). Por ejemplo, se ha encontrado que los investigadores en el área de biología están más dispuestos a compartir sus datos que los de medicina y farmacia (**da Costa y Lima Leite**, 2019). En biología, el descubrimiento de datos es fundamental, lo que se traduce en una mayor práctica de compartir datos (**Stuart et al.**, 2018). Un estudio sobre las actitudes de los psicólogos encontró que estos investigadores se ven influenciados por las normas de compartir datos dentro de su disciplina (**Harper y Kim**, 2018). Un estudio reciente de **Khan, Thelwall y Kousha** (2023) también encontró diferencias disciplinarias en la forma en que los investigadores comparten sus datos debido a cuestiones culturales: las ciencias físicas, de la tierra y planetarias y las ciencias ambientales son más proclives a compartir, mientras que, en áreas de economía y negocios, medicina e ingeniería son más renuentes.

Pertenecer a una organización que promueve compartir datos de investigación influye positivamente en las actitudes y comportamientos de los investigadores (**Fecher et al.**, 2015; **da Costa y Lima Leite**, 2019; **Zuiderwijk y Spiers**, 2019). Los académicos que se dedican exclusivamente a la investigación, a diferencia de aquellos con obligaciones docentes que ocupan gran parte de su tiempo, tienen más probabilidades de poner sus datos a disposición de otros investigadores. Los investigadores titulares son más propensos a compartir sus datos de investigación que aquellos que aún no lo son (**Fecher et al.**, 2015; **Tenopir et al.**, 2011).

Khan, Thelwall y Kousha (2023) realizaron una encuesta a 3.257 investigadores de todo el mundo concluyendo que, el compartir datos en abierto, así como reutilizarlos sigue aumentando, pero no es práctica generalizada en ninguna área temática, siendo más frecuente entre investigadores con más de 10 años de experiencia. Estos hallazgos refuerzan la asociación entre intercambio de datos y carrera de investigación más extensa, ya señalada por **Gregory et al.** (2020) y **Dorta-González et al.** (2021). También se encontró una relación con la edad de los investigadores, pero de manera inconsistente. En Brasil, **da Costa y Lima Leite** (2019) encontraron que los investigadores más jóvenes son más proclives a compartir abiertamente sus datos, por sus habilidades en el uso de tecnologías y por interés en colaborar con investigadores afines. Por el contrario, los investigadores de 50 años mostraron más interés en compartir datos porque ya han llegado a la titularidad en su cargo (**Tenopir et al.**, 2011). **Kim, Schuler y Pechenina** (2018) identificaron también que los investigadores de mayor edad muestran actitud positiva a compartir datos. En cuanto a las motivaciones

personales, **Zuiderwijk, Shinde y Jeng** (2020) identifican rasgos de carácter de los investigadores, como la apertura a nuevas experiencias, que pueden impulsarlos a compartir, incentivos individuales (como el aprendizaje, la autonomía, y la eficacia) o incentivos sociales (como devolver a la sociedad y fomentar nuevos procesos de aprendizaje). Sentir que están contribuyendo al beneficio público es una gran motivación para compartir datos (**Navarro-Molina y Melero**, 2019; **Goode et al.**, 2022, **Borycz et al.**, 2023).

Uno de los factores que más se han estudiado es cómo influyen los requisitos y obligaciones formales para que los investigadores compartan sus datos. Las agencias de financiamiento actuales alrededor del mundo exigen compartir los datos de investigación a cambio de dar soporte a las investigaciones (**Fecher et al.**, 2015). Algunos estudios encontraron que la presión regulatoria ejercida por las agencias de financiamiento, tenían un efecto significativo en la práctica y la actitud de intercambio de datos de los investigadores (**Kim et al.**, 2018). El servicio [Sherpa Juliet](#)²⁶ registraba 52 mandatos de instituciones que requieren compartir datos como condición para la financiación (**Goode et al.**, 2022).

Las políticas de las agencias financiadoras es un aspecto que no se puede soslayar (**Schmidt**, 2016; **Navarro-Molina y Melero**, 2019, **Zuiderwijk et al.**, 2020). **Woods and Pinfield** (2021) enfatizan que las políticas deben centrarse en cambiar la cultura de la investigación hacia la apertura, no solo a nivel individual. Para ello, brindan recomendaciones para la formulación de estas políticas a largo plazo. Los requisitos establecidos por las revistas para compartir datos, aspecto desarrollado en las políticas de publicación, están entre los principales motivadores a la hora de cambiar de conducta (**Schmidt**, 2016; **Fecher et al.**, 2015, **Goode et al.**, 2022).

Entre las condiciones facilitadoras para las prácticas de compartir datos se encuentran la disponibilidad de infraestructura tecnológica adecuada, sistemas de información apropiados, facilitación de las TICs (internet, computadoras, crecimiento en capacidad de almacenamiento y redes, etc.). Otro factor esencial es la disponibilidad de repositorios de datos para que los investigadores puedan almacenar y compartir sus datos (**Kim y Kim**, 2015; **Kim y Zhang**, 2015; **Borycz et al.**, 2023) y fomentar su reutilización (**Kim y Yoon**, 2017; **Kim**, 2021). Tanto las capacidades de almacenamiento como las de acceso deben ser sostenibles y poder

²⁶ Base de datos sobre las políticas de los financiadores de la investigación y sus requisitos de acceso abierto, publicación y archivo de datos. Actualmente se denomina [Jisc's open policy finder](#).

crecer y operar de manera confiable, en muchos casos la infraestructura tiene que adaptarse a las características de la disciplina, de los recursos humanos, tecnológicos y de estándares, y de financiación (Schöpfel et al., 2018; Navarro-Molina y Melero, 2019, Zuiderwijk et al., 2020).

Otras condiciones facilitadoras encontradas incluyen la asistencia disponible para los investigadores en sus respectivas instituciones, tanto en términos de apoyo técnico como humano (Joo et al., 2017). Especialmente relevante es el rol de los recursos humanos, como asesores y grupos de usuarios y productores de datos, que constituyen un soporte integral para los investigadores (Kim y Yoon, 2017). También se considera esencial contar con un presupuesto continuo y dedicado, apoyo financiero adecuado, y disponer del consentimiento informado o contractual necesario para poder compartir los datos.

En la revisión de Zuiderwijk et al. (2020) se identificó que la confianza comprende diversos aspectos muy importantes al momento de compartir datos. Por un lado, los investigadores quieren lograr la confianza de sus pares, de los usuarios de datos y de la sociedad en general en sus resultados de la investigación, que implica una confianza en sí mismos. Comparten sus datos a fin de mostrar transparencia y credibilidad no solo de sus resultados sino de la metodología y el proceso de investigación, facilitando su interpretación. Otro aspecto importante para establecer confianza es la posibilidad de conectarse con los usuarios y entender qué pueden hacer con sus datos. Algunos repositorios están implementando metodologías para mostrar la reutilización de datos. Es el caso del repositorio [DesignSafe](#)²⁷ publica sus casos de reutilización ([Data Reuse Stories](#)), dando cuenta de las diferentes formas de uso e impacto de esos datos. Garantizar la validez de los datos por parte de múltiples usuarios puede considerarse otro impulsor para compartir abiertamente los datos de investigación.

Al momento de reutilizar datos abiertos, influye la confianza en los investigadores o instituciones que ofrecen los datos (Wallis, Rolando y Borgman, 2013). Se ponen en juego varios tipos y niveles de confianza (Yoon, 2017) en un proceso dinámico, demostrando que la confianza puede formarse, desaparecer, declinar, perderse y recuperarse en el contexto de la reutilización de datos. Cuando se utilizan

²⁷ Es un componente de ciberinfraestructura financiado por la [National Science Foundation](#). Su función principal es proporcionar a la comunidad científica herramientas para gestionar, analizar y publicar datos críticos sobre el impacto de desastres naturales en la infraestructura civil y la sociedad.

repositorios de datos, los investigadores deben poder confiar en que estas infraestructuras de servicios ofrecerán un acceso fácil para los usuarios y garantizarán la preservación a largo plazo de los datos. Por lo tanto, los repositorios de datos necesitan generar esa confianza entre los autores de datos. Esto se logra mediante una infraestructura sostenible y robusta, personal capacitado y el establecimiento de políticas claras tanto para el depósito como para el uso de los datos. Además, la adhesión a buenas prácticas, junto con programas de capacitación y apoyo basados en una buena comunicación con los autores, son también factores facilitadores. Uno de los mecanismos para garantizar la confianza es la [certificación de repositorios CoreTrustSeal](#), que establece requisitos a cumplir para ser considerados repositorios confiables (más detalles en el capítulo "Los repositorios de datos"). En algunos casos, los datos no pueden compartirse en su totalidad o solo parcialmente; en estos casos, la posibilidad de establecer condiciones de acceso a los datos aumenta la confianza del investigador. Si se tiene en cuenta la influencia social y afiliación como impulsores para compartir datos, los investigadores dan cuenta de su capacidad de respuesta social, además de la presión normativa percibida, las normas sociales estándar en su grupo de trabajo, en la disciplina, en la institución o país, las normas subjetivas, la presión de las revistas, la presión de los colegas, la atención mundial a la necesidad de compartir y preservar los datos (por ejemplo en el caso de la Pandemia), además de los códigos de conducta y estándares normativos de colegios profesionales y sus respectivas comunidades. Esto también se verifica en el uso de los datos de otros, ya que, si un colega les recomienda el uso de un conjunto de datos, existe una influencia o presión (si es un supervisor, por ejemplo) para utilizarlos, dependiendo la relación (Harper and Kim, 2018; Zuiderwijk et al., 2020).

Cuando los investigadores perciben que el grado de esfuerzo requerido para compartir los datos es bajo en relación con los beneficios potenciales, como evitar la duplicación de datos, ahorrar tiempo, o tener los datos bien curados para futuras investigaciones, hay más posibilidades de que compartan los datos por el uso optimizado de su tiempo. Si los investigadores cuentan con apoyo para la gestión de los datos durante el ciclo de vida de sus datos, cuentan con equipamiento, software y repositorio de datos, tienen asistencia técnica para compartir los datos es más probable que los compartan. Además, la experiencia previa de haber usado datos de otros investigadores, evitando la necesidad de hacer una recolección de datos propia, lleva también a que se compartan (Zuiderwijk y Spiers, 2019). La facilidad de acceder a los datos, cuando son fáciles de encontrar, cuando la

relevancia de los datos es clara y cuando datos son fáciles de usar, es más probable que exista reutilización de datos.

Tanto la experiencia previa como las habilidades de los investigadores para compartir datos juegan un rol importante (Zenk-Möltgen et al., 2018), ya sea tener acceso a especialistas en datos, acceder a consultas sobre curación y gestión de datos, contar con capacitación y dominio en las habilidades de gestión de datos, los conocimientos sobre metadatos y prácticas relacionadas. Un estudio basado en la TPB sobre compartir datos de redes sociales (Akdeniz et al., 2023) encontró que el comportamiento pasado fue un predictor significativo de compartir datos, demostrando que los investigadores que compartieron sus datos el pasado (experiencia previa) tenían la intención de compartir nuevamente, tanto públicamente, como bajo acceso controlado o bajo solicitud personal. Cuando los investigadores creen que los datos abiertos son herramientas para educar a sus estudiantes o nuevos investigadores y si han tenido experiencias previas exitosas al compartir abiertamente sus datos, tienen mayores posibilidades de incrementar esa práctica (Zuiderwijk et al., 2020).

Lo mismo sucede con la reutilización de datos de otros investigadores. Aquellos investigadores que tienen experiencias pasadas positivas con el uso de datos abiertos podrían estar más motivados para usar datos de investigación abiertos (Zenk-Möltgen et al., 2018). Estar familiarizados con los datos disponibles y encontrar estos datos útiles, tener experiencia en la recopilación de dichos datos y saber cómo gestionarlos lo que podría ahorrarles tiempo en la búsqueda y el uso de datos relevantes para su propia investigación, especialmente si tienen experiencia o conocimientos en datos específicos. Un aspecto fundamental es la capacitación de los investigadores, su capacidad para comprender los datos abiertos y la capacitación formal para encontrar, adquirir y validar datos recopilados por otros. Es muy útil el conocimiento obtenido a través del entrenamiento en la propia disciplina en la reutilización de datos (Zimmerman, 2007).

En cuanto a las legislaciones y regulaciones para compartir datos de investigación, destacan como impulsores las políticas de datos claras y transparentes, las políticas de disponibilidad de datos de las revistas y las políticas institucionales formales. Las legislaciones nacionales, como la [Ley 26.899](#) de Argentina, y los acuerdos o recomendaciones internacionales, como la de la UNESCO, también estimulan esta práctica.

Respecto a las características de los datos, los investigadores muestran mayor disposición a compartirlos cuando existen controles de calidad, buenas prácticas de gestión, identificadores persistentes como los DOI, documentación adecuada, metadatos normalizados y formatos abiertos o interoperables. Asimismo, la ausencia de datos personales, como los de pacientes, reduce las barreras. Es más habitual compartir datos de investigaciones cuantitativas que cualitativas (Cragin et al., 2010), posiblemente porque requieren menos esfuerzo de anonimización y limpieza. También influye la disponibilidad de herramientas que faciliten su utilización.

En España, **González-Teruel et al. (2022)** realizaron entrevistas a investigadores, editores, vicerrectores, responsables de agencias de evaluación y grupos focales de bibliotecarios de repositorios para identificar los factores que favorecen el intercambio de datos de investigación.

FACTORES FAVORECEDORES	I	E	V	B
Personal investigador				
Percepción del beneficio de compartir los datos	●			
Existencia de beneficios curriculares o de reputación académica	●			
Tener confianza personal con quien se comparten los datos	●	●	●	
Existencia de disciplinas en las que compartir y abrir los datos ya se considera un proceso natural		●		
Estímulo a través de la competencia entre investigadores		●		
Apoyo institucional				
Implicación y el compromiso de todos los actores	●		●	●
Existencia de planes de datos previos a nivel institucional	●			
Respaldo y apoyo a los servicios universitarios	●		●	
Marco normativo				
Existencia de un marco normativo al que se ciñen todos los implicados	●	●		
Obligatoriedad del depósito de datos en las convocatorias de financiación de la investigación	●	●		●
Sistema de comunicación científica				
Obligatoriedad impuesta por la revista	●	●		
Fomento de una cultura en la que el depósito de datos sea considerado un reflejo de fiabilidad		●		
I: Investigadores; E: Editores; V: Vicerrectores; B: Bibliotecarios				

FIGURA 22 Factores favorecedores respecto de la apertura de los datos de investigación en España.
Fuente: **González-Teruel et al. (2022)** bajo licencia [CC BY 4.0 Deed](#)

2.8 Barreras para compartir datos de investigación (sistemas de evaluación científica)

Vamos a describir los principales factores que desincentivan a los investigadores a compartir datos, siguiendo las variables identificadas por **Zuiderwijk, Shinde y Jeng (2020)** comenzando por los factores determinantes y continuando con los demás. En cuanto al desempeño esperado, se observa que, en general, salvo algunas excepciones, los sistemas de evaluación y recompensas académicas están basados en las publicaciones más que en la gestión de datos. Las contrataciones, promociones, obtención de fondos y otros aspectos del sistema de recompensas están centrados en la producción de publicaciones con revisión por pares (**Borgman, 2007; Tenopir et al., 2011; Fecher et al., 2017**), lo que favorece prácticas que dificultan el reconocimiento y la recompensa a la ciencia abierta, ampliamente detalladas por **Cole, Reichmann y Ross-Hellauer (2023)**.

Los investigadores no se inclinan a compartir sus datos si no reciben crédito o compensación por esta tarea (**Simmons et al., 2021, Chawinga y Zinn, 2021**), que puede adoptar diferentes formas incluyendo citación adecuada, oportunidades de colaboración, coautoría y reconocimiento formal (**Sayogo y Pardo, 2013; Navarro-Molina y Melero, 2019**). Otros factores relacionados con el desempeño esperado son el temor a perder una ventaja competitiva si terceros publican utilizando sus datos sin reconocer su autoría y la percepción de que los datos no se citan ni valoran adecuadamente. También se identifican riesgos profesionales, como perder financiación, propiedad intelectual lucrativa, oportunidades de comercialización o futuras posibilidades de publicación. Es frecuente el temor a que los datos sean utilizados para realizar análisis adicionales que los propios investigadores han planificado llevar a cabo en el futuro (**Zuiderwijk et al., 2020**).

También preocupa conservar la prioridad del descubrimiento y controlar los resultados hasta su publicación. Los científicos suelen compartir los datos solo después de agotar su uso y publicar sus resultados, pues temen perder oportunidades de intercambio con otros laboratorios. Asimismo, la transparencia puede exponer los datos y análisis a críticas, revelar errores o provocar interpretaciones equivocadas que cuestionen los resultados y dañen su prestigio (**Zuiderwijk et al., 2020**). Esto se vincula directamente con el concepto de confianza para compartir abiertamente los datos de investigación. Las barreras incluyen el temor a perder el control y la integridad de los datos, el miedo a un posible uso no

ético, tanto por la mala interpretación como por el mal uso de los datos, así como su comercialización. Además, existe el temor al daño a la reputación del investigador, la desconfianza en la calidad de los datos de otros investigadores y la incertidumbre sobre quiénes serán los usuarios de los datos. Estos usuarios podrían llegar a sacar conclusiones incorrectas o hacer un uso indebido de los datos. Los investigadores también pueden temer el uso comercial o competitivo indebido de los datos, lo que podría perjudicar su prestigio (Enke et al., 2012; Fecher et al., 2017; Tenopir et al., 2011; Schmidt, Gemeinholzer y Treloar, 2015; Tenopir et al., 2015).

En el informe [‘The State of Open Data 2021’](#)²⁸, elaborado a partir de una encuesta que respondieron 4500 investigadores de todo el mundo, se encontró una alta preocupación sobre el mal uso de los datos (Simmons et al., 2021). Los científicos controlan los datos que ellos generan y se sienten dueños de estos, independientemente de si legalmente les pertenecen. No es siempre sencillo determinar a quién pertenecen los datos porque pueden intervenir agencias de financiamiento, diferentes colaboradores y jurisdicciones. La confianza implícita en el acto de intercambiar datos entre colaboradores cercanos es difícil de replicar cuando esos intercambios se hacen con usuarios desconocidos que accederán a los datos en un futuro, por lo que hacen falta “intermediarios” que ayuden a contextualizar, explicar y curar estos datos. Estos roles pueden ser asumidos por científicos de datos, bibliotecarios, entre otros (Borgman, 2016).

Los investigadores necesitan capacitación en compartir datos, es decir, desarrollar conocimiento, habilidades y competencias de las que carecen. Además, se requiere una infraestructura técnica adecuada para almacenar, gestionar, mantener y recuperar datos. (Chawinga y Zinn, 2021). Una barrera no menor es el esfuerzo que implica para el investigador documentar sus datos para ser usados por otros. Los conocimientos sobre los datos están embebidos en prácticas sociales que tienen que ver con las condiciones de recolección de los datos, los instrumentos utilizados, explicación sobre variables y otra información específica (Borgman, 2016, Navarro-Molina y Melero, 2019). La inversión en esfuerzo individual y tiempo necesario para tener esos datos para compartir, ser preservados y gestionarlos influye en las conductas de compartir datos (Kim, 2013; Kim y Stanton, 2016). Muchas veces, los

²⁸ Sexta edición de una serie de estudios anuales realizados por [Digital Science](#), [Figshare](#) y [Springer Nature](#) para analizar las actitudes y comportamientos de los investigadores globales hacia el intercambio y la compartición de datos.

investigadores deben estructurar el conjunto de datos siguiendo un estándar particular, describirlos de manera completa y documentarlos adecuadamente para que sean reutilizables para otros investigadores. Estos esfuerzos parecen ser mayores para los estudios cualitativos en comparación con los de tipo cuantitativo. Aquí surge otra barrera relacionada con el manejo de la tecnología, ya que interfaces de usuario complejas pueden hacer que el proceso de depositar los datos resulte tedioso (**Zuiderwijk et al., 2020**). Los investigadores necesitan apoyo para que sus datos sean FAIR, aunque este esfuerzo rara vez se contempla en la financiación de los proyectos. Por ello, suelen asumir esta tarea y recurrir a repositorios, editores y bibliotecas, que desempeñan un papel clave para facilitar la disponibilidad abierta de los datos (**Simmons et al., 2021**).

Existen limitaciones relacionadas con legislaciones y regulaciones, como las restricciones de propiedad intelectual, que pueden surgir en casos de datos protegidos por derechos de autor o patentes, o en situaciones donde leyes específicas prohíben la difusión de ciertos datos. En otros casos, los investigadores no pueden reproducir digitalmente los documentos originales sin el permiso de los titulares de los derechos, como sucede en el campo de las humanidades digitales (**Berghmans et al., 2017; Navarro-Molina y Melero, 2019; Chawinga y Zinn, 2021**).

Otro aspecto relacionado son los términos de las licencias. Los investigadores, al enfrentarse a la elección entre una amplia variedad, tienen dificultades de comprensión, esto genera confusión y hace necesario simplificar las opciones disponibles (**Campbell, 2015**). Si los investigadores tienen dudas sobre los aspectos legales de sus datos, es menos probable que los compartan. Estos aspectos junto a los derechos de uso, la confidencialidad y los contratos con industrias o empresas, son factores que también inhiben el intercambio de datos en abierto.

Si se suman cuestiones éticas que exigen el consentimiento de los sujetos estudiados, así como la gestión de datos confidenciales o con información personal, lo que genera preocupaciones sobre la privacidad (**Berghmans, 2017; Chawinga y Zinn, 2021**). En muchos casos, estos factores hacen imposible compartir los datos, como ocurre en situaciones que implican riesgos para la seguridad nacional o para especies en peligro de extinción, entre otros. Puede ser necesario anonimizar los datos para poder difundirlos. Estos aspectos se abordarán en mayor detalle en apartados específicos. Un ejemplo de los aspectos que desalientan a los científicos sociales y de otras disciplinas que trabajan con datos personales son las restricciones relacionadas con las normas de acceso a datos personales y la

confidencialidad en la documentación. Las ciencias sociales son menos acumulativas y el tema de la prioridad es menos importante, ya que los mismos datos pueden ser reinterpretados desde diferentes puntos de vista. En cuanto a los aspectos de propiedad intelectual, factores como la procedencia de los datos (empresas, agencias gubernamentales, etc.) y los términos contractuales para su uso tienen un impacto significativo (**Borgman**, 2016).

En la literatura encontramos motivaciones personales para no compartir, por ejemplo, miedo a que los datos sean poco utilizados, pereza, actitud negativa hacia compartir datos o la limitación de comercializar esos datos si están abiertos (**Zuiderwijk et al.**, 2020). Si se consideran los antecedentes del investigador, se observa que las disciplinas o subdisciplinas tienen su propia cultura. Una encuesta dirigida a más de 1300 investigadores de diferentes países halló diferencias disciplinarias. Por ejemplo, medicina, ciencias de la computación/ingeniería y ciencias sociales eran las que menos compartían datos (**Tenopir et al.**, 2011). Otro caso son los humanistas que suelen tratar de controlar las fuentes de información (archivo, manuscrito, artefacto, etc.) hasta agotar las posibilidades de publicar, como ocurre con los sitios arqueológicos que pueden tener restricciones para ser usados por académicos de los países donde se encuentre el sitio (**Borgman**, 2016).

Si se toma la edad, en relación con los jóvenes investigadores, una encuesta respondida por 3200 investigadores en 14 países (**Gownaris et al.**, 2022) encontró que éstos estaban familiarizados con los datos abiertos y que dos tercios habían usado datos abiertos en su enseñanza o investigación. La encuesta les preguntó acerca de los obstáculos y los jóvenes investigadores mencionaron problemas de ética y confidencialidad, políticas institucionales que desalientan la publicación de datos abiertos y que la práctica es poco común en algunas disciplinas (por ejemplo, en Humanidades y Ciencias Sociales). Las preocupaciones principales fueron la calidad de los datos abiertos y la falta de atribución adecuada y el uso sin un conocimiento detallado del conjunto de datos.

Zuiderwijk et al. (2020) enfatizan que los factores demográficos como la edad, el género, el país, el hecho de ser profesor titular o no, deben analizarse en el contexto de otros factores sociales, organizacionales y culturales más amplios que desempeñan un papel en las decisiones de los investigadores de abrirse. compartir datos de investigación o no.

Los mismos autores señalan que la falta de habilidades, conocimientos y experiencia dificulta compartir datos de investigación en abierto, especialmente cuando existen carencias en gestión de datos y metadatos. La ausencia de experiencia previa o la insatisfacción con experiencias anteriores también pueden actuar como barreras. Por ello, son fundamentales la concienciación, la promoción y la formación sobre los beneficios de compartir datos. En cuanto a sus características, la falta de estandarización, la baja calidad de los datos o metadatos, las deficiencias metodológicas y la inconsistencia descriptiva desalientan su difusión. Asimismo, los problemas de privacidad, los formatos propietarios o difíciles de abrir, el gran volumen de los datos y otras limitaciones técnicas pueden dificultar o impedir que se compartan abiertamente.

En el estudio cualitativo enmarcado en el sistema de ciencia español de **González-Teruel et al. (2022)** se identificaron tanto barreras comunes como otras diferentes: la falta de reconocimiento de los beneficios de compartir datos, la competencia entre investigadores para publicar datos primero y las diferencias entre disciplinas. Además, se mencionó el temor a la pérdida de contexto, lo que podría hacerlos inutilizables para quienes no los generaron.

Otros obstáculos que surgieron fueron el miedo a ser cuestionado, el aumento en la carga de trabajo y la reticencia general de los investigadores a liberar y compartir datos. Asimismo, se señalaron barreras relacionadas con el contexto científico: la ausencia de políticas de datos adecuadas, tanto a nivel local como global, incluyendo la resolución de cuestiones éticas y legales; la necesidad de coordinación entre instituciones de investigación; la incertidumbre sobre el financiamiento de infraestructuras de apoyo y los posibles conflictos derivados de la financiación privada en la investigación.

Si bien existen incentivos para compartir datos, hay aspectos que desalientan estas prácticas considerablemente y los datos no se depositarán en repositorios voluntariamente. Crear la capa de contenidos es responsabilidad de las instituciones y políticos más que de los individuos, la conducta de los investigadores cambiará cuando las políticas ofrezcan más recompensas y cuando las herramientas y servicios ofrecidos mejoren sustancialmente disminuyendo el esfuerzo requerido (**Borgman, 2007**). La clave parecería estar en brindar incentivos potentes para la carrera del investigador.

2.9 Dimensiones y variables de la literatura relevada

La **tabla IV** resume dimensiones y variables de acuerdo con la literatura relevada.

Estudio	Literatura
Dimensión Individual	
Beneficio profesional percibido	Piwowar, Day y Fridsma, 2007; Piwowar y Vision, 2013; Sayogo y Pardo, 2013; Tenopir et al., 2011; Bongiovani y Martínez Uribe, 2014; Kratz y Strasser, 2015; Tenopir et al., 2015; Fecher et al., 2015; Kim y Stanton 2016; Arza y Fressoli, 2017; Fecher et al. 2017; Harper y Kim, 2018; Zuiderwijk y Spiers, 2019; Navarro-Molina y Melero, 2019; Goode et al., 2022; González-Teruel, et al., 2022
Altruismo académico (Ciencia y Sociedad)	Bongiovani y Martínez Uribe, 2014; Kim y Stanton (2016); Arza y Fressoli, 2017; Navarro-Molina y Melero, 2019; Chawinga y Zinn, 2021; Goode et al., 2022
Altruismo académico (Investigadores)	Bongiovani y Martínez Uribe, 2014; Kim y Stanton (2016), Arza y Fressoli, 2017; Navarro-Molina y Melero, 2019; Chawinga y Zinn, 2021; Borycz et al., 2023
Conocimientos y habilidades (Control conductual percibido)	Zenk-Möltgen et al., 2018; Zuiderwijk et al., 2020; Chawinga y Zinn, 202; Akdeniz et al., 2023
Riesgo profesional percibido	Enke et al., 2012; Bongiovani y Martínez Uribe, 2014; Fecher et al., 2017; Tenopir et al., 2011; Schmidt et al., 2015; Tenopir et al., 2015; Borgman, 2016; Berghmans et al., 2017; Harper y Kim, 2018; Navarro-Molina y Melero, 2019; Zuiderwijk et al., 2020; Simmons et al., 2021; Chawinga y Zinn, 2021; Gownaris et al., 2022
Esfuerzo percibido	Bongiovani y Martínez Uribe, 2014; Borgman, 2016; Kim y Stanton, 2016; Harper y Kim, 2018; Navarro-Molina y Melero, 2019; Zuiderwijk et al., 2020; Chawinga y Zinn, 2021; Borycz et al., 2023
Dimensión Institucional	
Presión normativa de la disciplina	Kim, 2013; Fecher et al., 2015; Tenopir et al., 2011; Kim y Stanton, 2016; Curty et al., 2017; Stuart et al., 2018; Harper y Kim, 2018; da Costa y Lima Leite, 2019; Zuiderwijk et al., 2020; González-Teruel, et al. 2022; Khan, Thelwall y Kousha, 2023
Presión regulatoria de la agencia de financiamiento	Fecher et al., 2015; Schmidt et al., 2016; Kim et al., 2018; Navarro-Molina y Melero, 2019; Zuiderwijk et al., 2020; Chawinga y Zinn, 2021; González-Teruel, et al., 2022
Presión regulatoria de las revistas científicas	Enke et al., 2012; Fecher et al., 2015; Kim y Stanton, 2016; Schmidt, 2016; Harper y Kim, 2018; Zenk-Möltgen et al., 2018; Zuiderwijk et al., 2020; Chawinga y Zinn, 2021; Goode et al., 2022; González-Teruel, et al., 2022
Ayuda y Capacitación (Control conductual percibido)	Bongiovani y Martínez Uribe, 2014; Zimmerman, 2007; Navarro-Molina y Melero, 2019; Zuiderwijk y Spiers 2019; Joo et al., 2017; Kim y Yoon, 2017; Zuiderwijk et al., 2020; Simmons et al., 2021; González-Teruel, et al., 2022
Infraestructura y Repositorios (Control conductual percibido)	Kim y Kim, 2015; Kim y Zhang, 2015; Kim y Yoon, 2017; Navarro-Molina y Melero, 2019; da Costa y Lima Leite, 2019; Zuiderwijk, Shinde y Jeng, 2020; Kim, 2021; Borycz et al., 2023
Otros factores	
Edad, antigüedad, etapa del investigador, dedicación.	Tenopir et al., 2011; Fecher et al., 2015; Kim et al., 2018; da Costa y Lima Leite, 2019; Gregory et al., 2020; Dorta-González et al., 2021; Chawinga y Zinn, 2021; Gownaris et al., 2022; Khan et al., 2023
Recursos financieros apropiados	Bongiovani y Martínez Uribe, 2014; Schöpfel et al., 2018; Navarro-Molina y Melero, 2019; González-Teruel, et al., 2022
Características de datos	Bongiovani y Martínez Uribe, 2014; Cragin et al., 2010; Zuiderwijk et al., 2020
Estudios cualitativos vs cuantitativos	Zuiderwijk et al., 2020
Aspectos legales	Campbell, 2015; Borgman, 2016; Berghmans, 2017; Chawinga y Zinn, 2021; González-Teruel, et al., 2022

TABLA IV Dimensiones y variables del estudio en relación con la literatura relevada.

Fuente: elaboración propia.

2.9.1 Aspectos éticos y legales

La experiencia colectiva a nivel mundial durante la pandemia de COVID-19 demostró que el uso y la compartición abierta de datos científicos y técnicos aceleraron el desarrollo de vacunas y tratamientos (Besançon et al., 2019). Para aprovechar al máximo el potencial de los datos abiertos, es necesario superar obstáculos legales. Además, existen diversas restricciones legítimas en el acceso y uso de datos generados públicamente, como la protección de la seguridad nacional, la seguridad pública, la privacidad de las personas, la confidencialidad, la participación del sector privado con derechos de explotación de los datos y los términos de uso exclusivo establecidos por el investigador (períodos de embargo). También hay casos en los que los datos no se comparten abiertamente para evitar daños específicos, como revelar la ubicación de especies en peligro de extinción. Es fundamental considerar estos aspectos para fomentar el intercambio de datos de manera responsable y beneficiosa tanto para la comunidad científica como para la sociedad en general.

Aspectos éticos

Los aspectos éticos en la investigación y en la gestión de los datos son de suma importancia y deben ser considerados en todas las etapas del proceso. Además de las leyes y regulaciones, los investigadores se guían por normas éticas. Las normas éticas se basan en los valores de la disciplina académica y las directrices de las organizaciones profesionales y sociedades científicas. Por ejemplo, en el caso de la publicación en revistas las normas éticas más usadas son las del [Código de conducta y mejores prácticas para la edición](#) de COPE ('Committee on Publication Ethics')²⁹.

Se tienen en cuenta los requerimientos éticos establecidos en el [Código de Nuremberg](#) (BMJ, 1996) que establece diez principios donde se describen los estándares éticos que deben regir la experimentación en seres humanos. Estos principios incluyen el requisito de obtener un consentimiento informado voluntario por parte de los participantes, la necesidad de que los estudios se basen en una justificación científica sólida y la de evitar el sufrimiento o daño físico y mental

²⁹ Este código establece estándares éticos obligatorios para editores, revisores y autores en publicaciones académicas, enfocado en asegurar la integridad editorial, gestionar conflictos de interés y corregir malas prácticas. Se basa en prácticas principales y guías para garantizar la publicación ética.

innecesario a los participantes. El código también enfatiza la importancia de minimizar los riesgos y asegurar que los beneficios de la investigación superen cualquier posible daño. También sirvió como base para posteriores pautas éticas y regulaciones relacionadas con la investigación en seres humanos, incluyendo la [Declaración de Helsinki](#) y el [Informe Belmont](#). Continúa siendo un documento influyente en el campo de la ética médica y la ética de la investigación, destacando la importancia de proteger los derechos y el bienestar de las personas involucradas en la investigación científica.

La [Declaración de Helsinki](#) (World Medical Association, 2013), el [Informe Belmont](#) en Estados Unidos (**National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research**, 1979) y la [Política Federal para la Protección de Sujetos Humanos](#), conocida como “Regla Común” (“**Common Rule**” U.S. Department of Health and Human Services, 1991).

En la Unión Europea, el [Reglamento General de Protección de Datos](#) (GDPR, por sus siglas en inglés) ha generado una mayor conciencia sobre el uso ético de los datos (Voigt y Von dem Bussche, 2017). Este reglamento tiene como objetivo proteger los datos personales de los ciudadanos comunitarios y regular su procesamiento por parte de las organizaciones. Fue implementado en 2018 y se aplica en todos los países miembros de la Unión Europea, así como en aquellos países de fuera de la UE que manejen datos de ciudadanos europeos. Entre sus disposiciones, se incluye la obligación de obtener el consentimiento explícito de los individuos antes de recopilar y procesar sus datos personales, así como la garantía de que los datos sean procesados de manera justa, transparente y segura. También establece el derecho de los individuos a acceder, rectificar y borrar sus datos personales, así como el derecho a la portabilidad de estos.

El GDPR impone responsabilidades a las organizaciones que recopilan y procesan datos personales, exigiendo que adopten medidas adecuadas de seguridad y privacidad, designen un responsable de protección de datos y notifiquen cualquier violación de seguridad que pueda afectar los datos personales. También establece sanciones severas para aquellas organizaciones que no cumplan con sus disposiciones, pudiendo enfrentar multas importantes. Este reglamento ha cambiado la forma en que las organizaciones manejan y protegen los datos personales en la Unión Europea y ha elevado los estándares de privacidad y seguridad de datos en todo el mundo.

En la UNR, al momento de presentar un proyecto de investigación se debe firmar la [Declaración Jurada de Ética para la Acreditación de Proyectos en la Universidad Nacional de Rosario](#) que implica aceptar que se va a llevar a cabo una investigación de acuerdo con el [Código de Nuremberg](#) y la [Declaración de Helsinki](#), con el objeto de respetar los derechos de las personas y salvaguardar su dignidad e integridad. Incluye también el respeto normas éticas relativas al respeto de los animales, así como una aceptación de lo que establece la [Ley N.º 25.326](#) (conocida como *Habeas data*). En esta ley existen excepciones para investigaciones “en la medida que los datos recogidos no puedan atribuirse a una persona determinada o determinable”.

Consentimiento informado

Antes de participar en una investigación, las personas participantes en la misma deben ser plenamente informadas acerca de los objetivos, procedimientos y posibles riesgos de la investigación. Deben otorgar su consentimiento voluntario y estar informados de que tienen el derecho de retirar su participación en cualquier momento. Este procedimiento no está exento de limitaciones. El proceso de consentimiento informado puede presentar dificultades y desafíos, como la comprensión limitada de la información por parte de los participantes, la posibilidad de coerción o presión debido a la relación de poder o influencia de las personas involucradas en la investigación, las dificultades en grupos vulnerables, la necesidad de cambios en el consentimiento y los retos logísticos. Todo lo anterior puede comprometer la voluntariedad y autonomía de la decisión de consentir. Estas dificultades se tratan de abordar mediante la utilización de un lenguaje claro, proporcionando información adicional, garantizando la voluntariedad del consentimiento y utilizando comités de ética y revisión institucional para proteger los derechos de los participantes.

En las investigaciones cualitativas, se ha debatido sobre si es posible lograr un consentimiento verdaderamente informado dada la naturaleza reflexiva y en constante desarrollo de la investigación. Se ha propuesto la implementación del “consentimiento en proceso”, un enfoque en el cual los participantes brindan su consentimiento de manera continua a medida que evolucionan las ideas y preguntas de los investigadores. Sin embargo, otros investigadores abogan por encontrar un equilibrio que proteja a los participantes sin obstaculizar en exceso el proceso de investigación (Mannheimer, 2022).

Un desafío adicional se presenta ante la reutilización de los datos. Al utilizar datos de estudios anteriores, hay quienes argumentan la necesidad de obtener nuevamente el consentimiento de los participantes originales. Sin embargo, existen dificultades prácticas para obtener este nuevo consentimiento. Por ejemplo, puede haber casos en los que no se puede asumir el consentimiento implícito de los participantes originales debido a la sensibilidad de la información. Frente a esto se requiere un juicio profesional para determinar si la reutilización de los datos viola el contrato establecido entre los participantes y los investigadores primarios. Agregado a lo anterior, la necesidad de obtener un consentimiento informado para análisis secundarios puede variar según las características específicas de cada estudio (Mannheimer, 2022).

Las [Directrices Generales de Protección de Datos](#) (GDPR) en la Unión Europea y la Regla Común revisada en los Estados Unidos (**U.S. Department of Health and Human Services**, 1991) establecen normas y directrices para el intercambio de datos y el consentimiento informado. El GDPR requiere que los responsables de datos proporcionen información clara sobre el propósito y el uso de los datos antes de procesarlos. En los Estados Unidos, se ha introducido el concepto de *consentimiento amplio*, que abarca el almacenamiento y uso futuro de los datos. Los Comités de Revisión Institucional³⁰ (IRB por sus siglas en inglés) están generando pautas y lenguaje específico para obtener consentimiento amplio y apoyar la reutilización de datos.

La [Organización de Estados Americanos](#) (OEA) publicó los [Principios Actualizados sobre la Privacidad y la Protección de Datos Personales](#) (2021). En el principio 4 se establece que “los datos personales deberían ser tratados y conservados solamente de manera legítima no incompatible con las finalidades para las cuales se recopilaron. Su conservación no debería exceder del tiempo necesario para cumplir dichas finalidades, de conformidad con la legislación nacional correspondiente. El tratamiento ulterior de datos personales con fines archivísticos, investigación científica e histórica o con fines estadísticos, todos ellos, en favor del interés público, no se consideraría incompatible con las finalidades iniciales. No obstante, dicho tratamiento ulterior seguiría sujeto a este “principio de tratamiento limitado y conservación” y sigue afirmando que los datos personales podrían

³⁰ Estos comités son organismos esenciales dedicados a la revisión ética y científica de las investigaciones que involucran seres humanos. Su objetivo principal es salvaguardar los derechos, seguridad y bienestar de los participantes. Funcionan bajo regulaciones federales estrictas y están supervisados por la [Oficina de Protección de la Investigación en Seres Humanos](#) (OHRP) y la [FDA](#).

conservarse durante períodos siempre con fines archivísticos, de investigación científica e histórica o fines estadísticos, todos ellos en fin del interés público y sin perjuicio de la aplicación de las medidas técnicas y organizativas apropiadas para proteger los derechos y libertades del titular.

Obtener un consentimiento informado es un desafío tanto en la reutilización de datos que han sido anonimizados y también en la investigación social a gran escala. En el caso de los datos cualitativos anonimizados que se han compartido con el fin de su reutilización, a menudo no se puede contactar a los participantes para obtener su consentimiento informado para nuevas investigaciones. En el caso de los datos sociales a gran escala, la magnitud de los datos hace casi imposible obtener el consentimiento informado de cada participante (Mannheimer, 2022).

En cuanto a la curación de datos, el mismo autor observa que los curadores de datos suelen involucrarse en el proceso de investigación una vez que los datos ya han sido recopilados. Recomienda que se pueda establecer contactos entre los investigadores y los curadores de datos al inicio de un proyecto de investigación, dado que los curadores pueden colaborar en la redacción de un lenguaje de consentimiento informado amplio que facilite la futura reutilización de los datos. También pueden promover la formación de grupos de asesores para obtener orientación de la comunidad en relación con el consentimiento, y pueden sugerir el uso de estrategias automatizadas para obtener el consentimiento de los usuarios cuando los investigadores recopilan grandes volúmenes de datos sociales.

Como ejemplo de aspectos éticos que se presentan en la literatura, está el caso del ensayo clínico sobre la aplicación de corticoides antes del nacimiento que se desarrolló en países de bajos y medios ingresos, incluyendo Argentina. Este estudio fue financiado por los [Institutos Nacionales de Salud de Estados Unidos](#) (NIH) con el requisito de compartir los datos. Fue un estudio con resultados negativos que generó dilemas éticos a la hora de decidir compartir los datos por varios motivos: (1) el estudio estaba registrado no solo en NIH sino en también en [Maternal and Newborn Health Registry Study](#) y los investigadores definieron compartir solo los datos de un periodo de tiempo, 2) el consentimiento se solicitó solamente para el análisis primario de los datos y los investigadores no se pusieron de acuerdo sobre si era necesario el consentimiento para análisis subsiguientes; 3) los investigadores principales tenían preocupación sobre la interpretación errónea de los datos por parte de otros investigadores durante los análisis secundarios y decidieron que

deberían participar en el diseño de futuros análisis secundarios para evitar cualquier malinterpretación de los datos (**Barnes et al.**, 2019).

El intercambio de datos relacionados con la salud está regido por numerosos requisitos complejos y en constante evolución, lo que plantea un desafío considerable para crear planes de gestión y compartición de datos a largo plazo para la investigación financiada o realizada por el NIH y otros organismos. Existe preocupación sobre si la anonimización de datos es medida suficiente para proteger la privacidad de los datos personales y se están aplicando controles contractuales que incluyen medidas tendientes a proteger los datos personales entre los cuales se incluyen prohibiciones de reidentificación en datos anonimizados, restricciones en la combinación de conjuntos de datos, requisitos de seguridad para proteger la divulgación inadvertida de datos anonimizados y la prohibición de la divulgación a terceros sin los mismos controles contractuales. Estas medidas buscan prevenir la reidentificación de individuos y garantizar un uso responsable de la información compartida (**Rosati**, 2022).

Besançons et al. (2019) afirman que, en tiempos de crisis, como en el caso de la pandemia, se requiere tener acceso a los datos de investigación y proponen que:

- Los mecanismos de recolección de datos (por ejemplo, cuestionarios, algoritmos de simulación, etc.) debería estar disponible en plataformas dedicadas y, de ser posible, bajo una [licencia Creative Commons CC-BY](#).
- “...si los autores no pueden compartir sus datos brutos, los editores de las revistas deberían ...exigir que un tercero de confianza (que no sea la institución de los autores) examine los datos brutos para establecer la existencia de los mismos y validar los resultados del análisis presentado por los autores”.

Aspectos legales

Conocer los derechos que se aplican a los datos y sus implicaciones es crucial para poder compartir, usar o reutilizar legalmente los datos. Los datos ¿son considerados propiedad intelectual? De acuerdo con **Borgman** (2016), depende de las legislaciones nacionales, de los acuerdos contractuales establecidos entre partes, de las regulaciones de bases de datos, entre otras. La confusión que existe sobre la propiedad de los datos, los derechos de compartir datos, la

responsabilidad de abrir los datos, las condiciones contractuales para poder abrir los datos generan restricciones para compartir datos.

La [Convención de Berna para la Protección de Obras Literarias y Artísticas](#) establece estándares mínimos para el copyright, pero la protección de derechos de autor puede variar en cada país según sus leyes y la interpretación de los tribunales nacionales. Los principios destacados son: la originalidad y creatividad de una obra para ser elegible para protección de derechos de autor, la protección de la expresión y no de las ideas en las que se basa la obra, la fijación de la obra en un medio tangible según la legislación nacional, la protección automática sin necesidad de registro formal y la limitación en el tiempo de protección.

En Argentina, los derechos de autor duran toda la vida de estas personas, a lo que se deben sumar 70 años desde su muerte a favor de sus herederos. Los derechos morales de los autores, que no son transferibles como los patrimoniales, protegen su reputación personal. Aunque existen limitaciones y excepciones a los derechos de autor para los usuarios de obras protegidas, estas suelen ser específicas y dependen del caso. En el caso de los datos de investigación, los derechos de autor no se consideran una solución legal adecuada para implementar los requisitos amplios de reutilización e interoperabilidad de datos de investigación pública (Doldirina et al., 2018).

Estos autores explican que los datos factuales, que representan hechos objetivos y verificables del mundo real, no están protegidos por derechos de autor porque no son una expresión creativa original. La presentación estilística, es decir, la manera en que la información o los datos se presentan, organizan o se les da formato de una manera creativa y original puede ser protegida por derechos de autor, pero si hay solo unas pocas formas de expresar una idea se puede crear lo que se conoce como “fusión de derechos de autor” porque los datos factuales están combinados con su forma y es difícil establecer la originalidad. Las compilaciones de datos o colecciones de hechos pueden ser protegidas por derechos de autor si la selección, coordinación o disposición muestra cierta creatividad humana u originalidad. Sin embargo, el copyright solo abarca la disposición creativa o la selección de elementos de la compilación, no los elementos individuales en sí. En general, las compilaciones de datos, especialmente las de material factual, tienen una protección de derechos de autor “fina” que evita la copia textual y la copia de cualquier expresión creativa en la selección original.

Los metadatos estandarizados, que ofrecen información básica sobre un conjunto de datos, generalmente no están cubiertos por derechos de autor, ya que son seleccionadas para facilitar la búsqueda y la interoperabilidad. Sin embargo, una selección específica y reunión de elementos de metadatos para un propósito particular podría justificar la protección de derechos de autor. La falta de claridad legal sobre qué datos o metadatos están protegidos por derechos de autor plantea desafíos tanto para los proveedores como para los usuarios de datos (**Doldirina et al.**, 2018). Los investigadores pueden estar limitados para compartir datos si estos pertenecen a alguien más. Si utilizan en los datos partes descargadas de bases de datos numéricas, de texto e imágenes con términos de uso que restringen lo que se les permite hacer con el material descargado, es importante poder determinar esos permisos. En el caso de los archivos, muchas veces se exige a los visitantes que acepten restricciones sobre lo que pueden hacer con el material en el archivo (sí pueden fotografiar documentos, entre otros) y después, si pueden compartir estos documentos (**Mannheimer et al.**, 2019).

Los acuerdos contractuales entre organizaciones suelen especificar los permisos y restricciones sobre cómo se pueden manejar los datos. Estas especificaciones pueden incluir cláusulas sobre actualizaciones de datos, controles de acceso, garantías de calidad, cómo se pueden copiar y mostrar los datos, si se pueden difundir, cómo se acreditará la fuente original y quién es responsable de remediar las violaciones de datos (**Grabus y Greenberg**, 2019).

Preguntas que pueden aclarar el panorama de los derechos en juego en una investigación son entender: qué regulaciones y políticas se deben cumplir en función de quiénes financian la investigación y bajo qué términos; qué instituciones colaborarán en el proyecto, si es una o varias universidades y si hay empresas privadas participando con sus términos contractuales y normativas propias. Si en el proyecto participan estudiantes, habrá que ver si es un proyecto con información sensible y si las regulaciones lo habilitan.

En investigaciones cualitativas, los sujetos participantes pueden ser titulares de derechos de autor de sus dichos en una entrevista, por ejemplo. Si en un proyecto participan varias instituciones habrá que ver quién es dueño de los datos que se generan y con quiénes los comparte, entre otros elementos para determinar los derechos (**Andrews**, 2023).

Si se trata de investigaciones basadas en datos compartidos a través de redes sociales, los 'posts' por ejemplo son derechos intelectuales de las personas que los han realizado y licenciados a las empresas de redes sociales de acuerdo con los términos de servicios establecidos. Además, si se trata de datos recolectados en comunidades específicas pueden aplicar regulaciones específicas, como por ejemplo los [Principios OCAP de las Primeras Naciones](#) (Propiedad, Control, Acceso y Posesión), conjunto de principios rectores desarrollados por las [Primeras Naciones en Canadá](#) para afirmar sus derechos inherentes y jurisdicción sobre los datos e información relacionados con sus comunidades (Mannheimer, 2022).

Las excepciones de "uso justo" o "uso legítimo" (en Estados Unidos, Canadá, Europa, Australia y otros países), ofrecen a los investigadores una posible vía para reutilizar datos si no existe una licencia u otra autorización. El "uso justo" garantiza el equilibrio entre los intereses de los propietarios de derechos de autor y los usuarios, evitando que la ley de derechos de autor sofoque la creatividad que pretende fomentar y ofreciendo protecciones importantes a bibliotecas, archivos e instituciones educativas sin fines de lucro, ya que, si creen razonablemente que su uso de una obra con derechos de autor es un "uso justo", no se les pueden aplicar muchas de las sanciones más estrictas en la ley de derechos de autor (Mannheimer, 2022). Se debe tener en cuenta que las limitaciones de este "uso justo" puede variar en diferentes jurisdicciones. La legislación argentina no reconoce el "uso justo" en general ni copia académica o privada en particular.

Interoperabilidad legal

La interoperabilidad legal se refiere a la capacidad de combinar datos de múltiples fuentes sin entrar en conflictos entre las restricciones impuestas por los proveedores de datos (Agosti et al., 2016). Esto significa que no es necesario obtener autorización de los proveedores de datos caso por caso al combinar la información. Para que haya interoperabilidad legal, las condiciones de uso de cada conjunto de datos deben ser claras y fácilmente determinables, preferentemente a través de métodos automatizados. Además, las condiciones de uso de cada conjunto de datos deben permitir la creación y uso de productos combinados o derivados. Los usuarios deben poder acceder y utilizar legalmente cada conjunto de datos sin tener que buscar autorización de los creadores de datos de forma individual, siempre y cuando se cumplan las condiciones acumuladas de uso de todos los conjuntos de datos involucrados.

Privacidad y seguridad de los datos, datos confidenciales

La preocupación por la privacidad de la información puede impedir el intercambio de datos entre investigadores. Los investigadores pueden tener miedo de perder el control o incluso el conocimiento sobre quién tiene acceso a sus datos, así como la forma en que se acceden y se utilizan (Fecher, Friesike y Hebing, 2015). Los principales factores que contribuyen a esta aprehensión son la protección de información de identificación personal (PII, por sus siglas en inglés), como los 18 identificadores de la [Ley de Portabilidad y Responsabilidad del Seguro Médico](#) (HIPAA) en Estados Unidos, la propiedad intelectual y otras categorías de datos sensibles, incluyendo datos indígenas, especies en peligro de extinción o invasoras, datos de enfermedades similares y cuasi-identificadores, como el género, la fecha de nacimiento y el código postal, que, combinados, pueden identificar de forma única la población (Grabus y Greenberg, 2019).

La [Regla de Privacidad de HIPAA](#) no se aplica a todos los actores en la comunidad científica de Estados Unidos, sino a las “entidades cubiertas”, que incluyen a la mayoría de los proveedores de atención médica, planes de salud y organismos de compensación de atención médica. HIPAA no se aplica a universidades (excepto a sistemas de salud y organizaciones proveedoras afiliadas a universidades), muchas organizaciones de investigación y la mayoría de las entidades comerciales en el ámbito de la investigación, como compañías farmacéuticas, de dispositivos médicos y biotecnología. Esto es causa de confusión y genera una demanda de una igualdad de condiciones entre diferentes tipos de organizaciones (Rosati, 2022). Para poder compartir esos datos se deben seguir especificaciones muy estrictas que pueden llevar a tener un set de datos limitado para la investigación. Muchas veces se aplica un método de determinación de expertos, donde un experto estadístico “calificado” determina si el riesgo de que la información pueda ser utilizada, sola o en combinación con otra información disponible, para identificar a pacientes. La otra alternativa que se utiliza cuando se comparten datos a gran escala bajo HIPAA es utilizar un “conjunto de datos limitado”, con información parcialmente anonimizada requiriendo un “acuerdo de uso de los datos” tanto a personal interno como a investigadores externos (Rosati, 2022).

En la [Ley 26.899 de acceso abierto](#) de Argentina y su reglamento, entre sus excepciones dice que, si existen motivos justificados, se podrán no difundir datos

primarios, por ejemplo, en el caso de acuerdos de confidencialidad o de derechos de propiedad industrial. También se definen los datos sensibles que no podrán difundirse. Además, especifica que los datos deben estar en servidores o data centers, preferentemente ubicados dentro del territorio nacional, que garanticen el cumplimiento de lo establecido por la [Ley 25.326 de Protección de Datos Personales](#).

3. Los repositorios de datos

Lynch (2003) considera a los repositorios institucionales (RI) como una serie de servicios ofrecidos por una institución a su comunidad, con el propósito de gestionar y compartir recursos digitales creados tanto por la institución como por sus miembros.

3.1 Definiciones y tipología

Un repositorio de datos es una infraestructura gestionada y dotada de una interfaz de búsqueda que permite registrar, almacenar, organizar, conservar y curar datos y otros objetos digitales. Además de garantizar su acceso y recuperación, reúne materiales seleccionados en distintos formatos para facilitar su localización, consulta y reutilización. Su gestión debe asegurar que estos recursos permanezcan auténticos, fiables, accesibles y utilizables a lo largo del tiempo (**Austin et al.**, 2017).

Según **Kindling et al.** (2017), un repositorio de datos de investigación (RDI) es un sistema técnico y organizacional diseñado para facilitar a los investigadores la gestión, almacenamiento y compartición de sus datos, al mismo tiempo que permite la búsqueda y acceso a conjuntos de datos. Es una parte específica de una infraestructura de información sustentable que ofrece almacenamiento a largo plazo y posibilita el acceso a datos de investigación. Los RDI son una parte esencial de la infraestructura de investigación de diferentes “instalaciones, recursos y servicios relacionados utilizados por la comunidad científica para realizar investigaciones de alto nivel en sus respectivos campos”. Los repositorios de datos de investigación en línea son infraestructuras de bases de datos configuradas para administrar, compartir, acceder y archivar conjuntos de datos de investigadores. Los repositorios pueden ser especializados (disciplinarios) o multidisciplinarios, generales o institucionales. Permiten acceder a productos de investigación más allá del artículo, donde es posible evaluar la evidencia y poder validar la calidad y transparencia por expertos externos (**Uzwyshyn**, 2016).

La misión de los RDI varía desde brindar acceso inmediato a conjuntos de datos hasta procurar la preservación a largo plazo, variando en consecuencia la inversión en la curación de datos. Algunas instituciones dedican un tiempo importante de trabajo profesional a curar cada conjunto de datos antes del depósito, mientras que muchos otros confían en la “autocuración”, con una revisión mínima.

3.1.1 Tipos de repositorios de datos

Pampel et al. (2013) identificaron cuatro tipos de repositorios de datos que sirvieron como base para la organización del [Registro de Repositorios de Datos de Investigación](#) (Re3data), directorio cuyo principal objetivo es orientar a investigadores, organismos de financiación, editores e instituciones académicas sobre los diferentes repositorios de datos de investigación (RDI), tanto en su rol de productores de datos como en el de usuarios. Re3data promueve una cultura de intercambio, mayor acceso y mejor visibilidad de los datos de investigación y otros resultados de investigación (Strecker et al., 2024). Los tipos de repositorios de datos identificados fueron:

- 1) **Institucionales:** Repositorios desarrollados y gestionados por una universidad o un centro de investigación. A nivel universitario son multidisciplinares. Un ejemplo es el [Repositorio de Datos Académicos de la Universidad Nacional de Rosario](#).
- 2) **Disciplinarios:** Repositorios específicos para una disciplina o un conjunto de disciplinas afines. Existe una amplia variedad, por ejemplo [Copernicus Marine Environment Monitoring Service](#) que brinda acceso a datos oceanográficos y [DesignSafe Data Depot Repository](#) (DDR) una plataforma para la curación y publicación de conjuntos de datos generados en el curso de la investigación de ingeniería de desastres naturales.
- 3) **Generalistas/Multidisciplinares:** Repositorios que permiten a cualquier investigador, sin importar país o disciplina, compartir datos. [Figshare](#) y [Zenodo](#) son dos ejemplos. El primero es un repositorio de acceso abierto gratuito, propiedad de [Digital Science](#) (Springer Nature), donde los investigadores pueden preservar y compartir los resultados de sus investigaciones, incluidas figuras, conjuntos de datos, imágenes y videos, ofreciendo opciones de personalización para editores de revistas e instituciones. [Zenodo](#) es un repositorio de acceso abierto general, desarrollado en el programa europeo [OpenAIRE](#) y operado por el [CERN](#). Permite depositar conjuntos de datos, software de investigación, informes y cualquier otro objeto digital relacionado con la investigación. Tras un taller organizado por los [Institutos Nacionales de Salud de Estados Unidos](#) (NIH)

sobre el rol de los repositorios generalistas en la mejora de la capacidad para descubrir y reutilizar datos (recordemos que fue la primera agencia federal que, ya en 2023, exigió el depósito de datos en repositorios), miembros de la [Research Data Alliance](#) realizaron una [comparación de repositorios generalistas](#) con el objetivo de ofrecer una herramienta que los investigadores pudieran utilizar para tomar decisiones informadas sobre la selección de un repositorio (Stall et al., 2020, 2023).

- 4) De **proyecto específico**: Son repositorios creados como resultado de un proyecto específico. Actualmente esta categoría ya no aparece en [Re3data](#) y se encuentran como repositorios disciplinarios.

3.1.2 Repositorios Digitales Confiables (TDR)

Para garantizar la confiabilidad y la preservación a largo plazo de los datos hacen falta repositorios digitales confiables (TDRs, siglas de 'Trusted Digital Repositories'), dotados de una infraestructura sólida, gobernanza sostenible y políticas completas. Los TDRs desempeñan un papel importante al mantener el valor de los datos y deben demostrar capacidades para el acceso y la reutilización a largo plazo.

La certificación es esencial para evaluar la confiabilidad. Existen estándares internacionales como [CoreTrustSeal](#) que certifican repositorios. La transparencia y la evidencia son clave. La confiabilidad requiere auditorías que den cuenta de esa confianza y revisiones periódicas. Los [Principios TRUST](#) (Transparencia, Responsabilidad, enfoque en el Usuario, Sostenibilidad y Tecnología) ofrecen un marco para evaluar y mejorar la confiabilidad de estos repositorios y son relevantes para todos los actores involucrados en la Ciencia abierta (Lin et al., 2020).

Principio	Guía para Repositorios
Transparencia	Ser transparente sobre de los servicios y las colecciones de datos verificables por evidencia pública.
Responsabilidad	Asegurarse de la autenticidad e integridad de los datos y la confiabilidad de los servicios.
Enfoque en el Usuario	Cumplir con las normas y expectativas de gestión de datos de las comunidades de usuarios objetivo.
Sostenibilidad	Mantener los servicios y preservar las colecciones de datos a largo plazo.
Tecnología	Ofrecer infraestructura y capacidades para servicios seguros, persistentes y confiables.

TABLA V Principios TRUST. **Fuente:** Traducción del cuadro de Lin et al. (2020).

Para ayudar a los usuarios a decidir adecuadamente, los repositorios de datos tienen que brindar información clara y accesible sobre su alcance, la comunidad de usuarios a la que se dirigen, sus políticas y capacidades. A fin de cumplir con el principio de *transparencia*, deben indicar claramente su misión y alcance, así como los términos de usos generales del repositorio y de los datos archivados, el período mínimo de preservación digital de los datos y cualquier otra característica o servicio relevante, por ejemplo, el manejo responsable de datos sensibles. Esta comunicación clara permite que se puedan comprender las limitaciones que pueden afectar el servicio (Lin et al., 2020).

La *responsabilidad* en los repositorios requiere cuidar la gestión de los datos y servir a la comunidad de usuarios. Esto implica cumplir con estándares de metadatos, garantizar la calidad y autenticidad de los datos, proporcionar servicios de acceso y gestionar aspectos legales y de seguridad. Los usuarios tienen que poder confiar en que los datos serán completos y duraderos, lo que les permite utilizarlos y citarlos con confianza en sus propias investigaciones. Esta responsabilidad puede tener un respaldo legal o basarse en estándares éticos voluntarios.

La consideración del *usuario* es esencial para un TDR. Por ello, deben adaptarse a las expectativas cambiantes de comunidades de usuarios diversas, incluyendo autores y usuarios de datos, junto a fomentar la descripción completa de los datos y posibilitar comentarios posteriores sobre su calidad. Además, deben aplicar y hacer cumplir los estándares para garantizar la interoperabilidad y reutilización de los datos. Esto incluye implementar métricas, contribuir a directorios o buscadores y la capacidad identificar las expectativas cambiantes de la comunidad y adaptarse a las mismas (Lin et al., 2020).

Centrarse en la *sostenibilidad* es esencial para garantizar el acceso a largo plazo a los datos de TDR para las comunidades presentes y futuras. Esto exige planificar la mitigación de riesgos, asegurar la continuidad de los servicios y prever su recuperación ante desastres u otros cambios. También requiere financiación para conservar los recursos de datos en condiciones óptimas y una gobernanza que garantice su preservación, recuperación, accesibilidad y utilidad futuras. Así, el repositorio podrá adaptarse a las necesidades cambiantes de los usuarios y mantener servicios de calidad a lo largo del tiempo.

En cuanto a la *tecnología*, es esencial usar normas, herramientas y tecnologías relevantes y apropiadas para brindar servicios seguros y confiables. Esto implica usar estándares y tecnologías adecuados para la gestión de datos y contar con medidas de seguridad informática y física para proteger los recursos del repositorio.

Los [Principios TRUST](#) como un medio para facilitar la comunicación con todas las partes involucradas, una guía sobre cómo se puede demostrar que un repositorio es confiable. Un ejemplo de cómo se puede cumplir con estos principios es obtener la certificación internacional que ofrece [CoreTrustSeal](#), basada en los [Requisitos Centrales para Repositorios de Datos Confiables](#) ('Core Trustworthy Data Repositories Requirements'). Este sello se estableció con el objetivo de aumentar la confianza en los repositorios y asegurar que los datos científicos sean gestionados de manera confiable y sostenible a lo largo del tiempo. Los [requisitos de CoreTrustSeal](#) abarcan áreas como la gobernanza y organización del repositorio, políticas y prácticas para la gestión de datos, garantía de acceso y preservación a largo plazo y transparencia en los procedimientos. Obtener este certificado indica que un repositorio cumple con los estándares y prácticas establecidos en el ámbito internacional para el manejo ético y confiable de los datos científicos.

3.2 Curación de datos de investigación

Alcanzar los beneficios asociados a la investigación basada en datos exige garantizar su calidad como aspecto central para la reutilización. En este punto, es necesario describir diferentes aspectos del concepto de curación de los datos, proceso que ayuda a maximizar el potencial de estos, facilitando la investigación, aumentando su calidad y ampliando la base de conocimientos a través de documentación, enlaces y visibilidad. Compartir los datos requiere documentarlos, entre otras razones, porque sin información adecuada sobre cómo interactúan los datos dentro de los contextos en los que se generaron, produjeron, gestionaron e interpretaron, un conjunto de datos es simplemente una cadena de bits donde la interoperabilidad e interpretabilidad de los datos puede estar embebida en numerosos artefactos, instrumentos, esquemas de metadatos, ontologías y otras representaciones (Borgman y Bourne, 2022). Además, el proceso de curación de datos trata de explicar la situación y el contexto de la investigación en la que se generan los datos.

Los beneficios de curar los datos incluyen potenciar su reutilización para nuevas investigaciones, es decir, hacer un uso eficiente de los recursos para generar nueva ciencia; mantener datos de observaciones (astronómicas, climatológicas, satelitales, flora y fauna, etc.) únicos que son imposibles de recrear. En este sentido, habría más datos disponibles para nuevos proyectos de investigación. Contar con datos curados sirve para validar los resultados de investigación y como herramienta de enseñanza y aprendizaje, contribuyendo al bien público. En este proceso también se cumplen con las legislaciones y políticas que exigen la gestión de los datos. Por tanto, cabe entender la curación como la actividad de administrar y promover el uso de datos desde su creación, con el objetivo de ser aptos para el propósito actual y estén disponibles para su descubrimiento y reutilización (en el caso de conjuntos de datos dinámicos, esto significa un enriquecimiento o una actualización continua para mantenerlos adecuados a su propósito). Los niveles más altos de curación también implicarán mantener vínculos con la documentación y con otros materiales publicados (**Lord y Macdonald**, 2003).

Dentro de las actividades necesarias para la curación está el archivo de los datos porque garantiza que estos sean seleccionados y almacenados correctamente, que se pueda acceder a ellos y que su integridad lógica y física permanezca a lo largo del tiempo, incluidas la seguridad y la autenticidad. Asimismo, la preservación es una “actividad dentro del archivo en la cual elementos específicos de los datos se mantienen a lo largo del tiempo para que se pueda acceder a ellos y comprenderlos aun cuando existan cambios tecnológicos. La curación digital es el conjunto de estrategias que facilitan crear y mantener colecciones digitales sustentables, teniendo en cuenta su valor como material para el avance del conocimiento (**Beagrie**, 2006).

De acuerdo con **Esteva** (2017), esta área de estudio dentro de la bibliotecología y ciencias de información adopta elementos de teorías archivísticas (procedencia, autenticidad e integridad, [postcustodialismo](#)), así como también prácticas de bibliotecas digitales (asignación de metadatos, licencias de uso, identificadores únicos digitales) y de gestión de registros digitales (evaluación para retención de registros, flujos de información, registros activos y de archivo). La curación digital incluye las actividades propias de la gestión de datos, pero orientadas a la reutilización de datos para beneficio del conocimiento académico.

Palmer et al. (2013) identificaron hitos y tendencias en la concepción de la curación de datos. Encontraron que se empezó a hablar de ello como disciplina en el contexto de los museos en el año 1994 y, al mismo tiempo, en el Reino Unido se creó el [Joint Information Systems Committee](#) (JISC) que estableció el [Servicio de Datos de Artes y Humanidades](#) (AHDS). En 2002, en el marco del [Sloan Digital Sky Survey](#) (SDSS), ya se planteaba como una actividad que sería realizada mayormente por bibliotecarios.

Al principio, los términos “curación digital” y “curación de datos” se empleaban como sinónimos. El [Programa Educativo sobre Curación de Datos](#) (DCEP) de la [Escuela de Graduados de Bibliotecología y Ciencias de la Información](#) de la Universidad de Illinois en Urbana-Champaign en 2007, la consideraba como un proceso de gestión activa y continua de los datos a lo largo de todo su ciclo de vida de interés y utilidad para el conocimiento. La curación digital es más amplia en el sentido que incluye una amplia gama de material digital y sus actividades curatoriales pueden atender a diferentes comunidades, no necesariamente de investigación. La curación permite el descubrimiento y la recuperación de datos, mantiene su calidad, agrega valor y permite su reutilización a lo largo del tiempo a través de actividades que incluyen la autenticación, el archivo, la gestión, la preservación y la representación (**Johnston**, 2017).

El [ciclo de vida de los datos](#) mapea el proceso de investigación y sus diferentes fases, las etapas, actividades y procesos de curación relacionados a la creación, uso, difusión y preservación de los datos de investigación. Este enfoque es necesario porque los datos digitales son frágiles y susceptibles a los cambios tecnológicos, a lo largo de su ciclo de vida, teniendo en cuenta el impacto potencial de las actividades de curación y preservación en diferentes etapas del proceso de gestión. Es decir, las actividades que se llevan adelante en cada etapa influyen directamente en la capacidad para gestionar y preservar los datos en las etapas posteriores. Si los datos digitales se seleccionan conservando su autenticidad e integridad, se puede lograr la reutilización confiable (**Pennock**, 2007).

En la literatura científica se han propuesto diferentes modelos del [ciclo de vida de los datos de investigación](#). Uno de los más conocidos es el [“Ciclo de vida de la curación digital” de Higgins](#) (**Digital Curation Centre**, 2008) en el Reino Unido. Este modelo permite visualizar las etapas necesarias para la curación y preservación de los datos desde el inicio de la investigación. Se desarrolló como una herramienta

de capacitación para ayudar a los curadores a comprender los procesos involucrados en una curación exitosa y desarrollar metodologías de curación y preservación en las instituciones.

Este enfoque de ciclo de vida para la gestión de datos digitales implica la planificación, creación, evaluación, curación, descubrimiento y uso de estos materiales desde su conceptualización inicial hasta su eventual eliminación o reutilización. El modelo es una representación gráfica (**figura 27**) de este enfoque, que indica las etapas clave requeridas para una curación exitosa y permite la adaptación a diferentes dominios y niveles de granularidad. Higgins considera que este modelo representa las principales etapas necesarias para gestionar y preservar adecuadamente los datos. Su carácter flexible permite adaptarlo a distintos ámbitos y utilizarlo para planificar actividades, asignar responsabilidades, establecer estándares y documentar procesos y políticas. El modelo distingue entre acciones continuas, secuenciales y ocasionales.

El ciclo de vida para la gestión de datos digitales implica la planificación, creación, evaluación, curación, descubrimiento y uso de estos materiales desde su conceptualización inicial hasta su eventual eliminación o reutilización. El modelo es una representación gráfica de este enfoque, que indica las etapas clave requeridas para una curación exitosa y permite la adaptación a diferentes dominios y niveles de granularidad.

El modelo se puede utilizar junto con otros estándares, marcos y modelos de referencia para mejorar el diseño, la gestión y las mejores prácticas del flujo de trabajo para la curación y preservación exitosa de datos digitales, como el [Modelo de Referencia OAIS](#). El enfoque del ciclo de vida de los datos tiene como objetivo garantizar que las etapas esenciales de curación se planifiquen e implementen en secuencia, ya que las acciones que se aplican o no en cada una de las etapas del ciclo, determinan qué tan efectivamente se van a poder gestionar y preservar en las siguientes etapas (Harvey, 2010). Los datos están en el centro del ciclo de vida de la curación e incluyen objetos digitales (simples y complejos) y bases de datos. Los objetos digitales simples son elementos digitales discretos, como archivos de texto, imágenes o archivos de sonido con sus identificadores y metadatos relacionados.

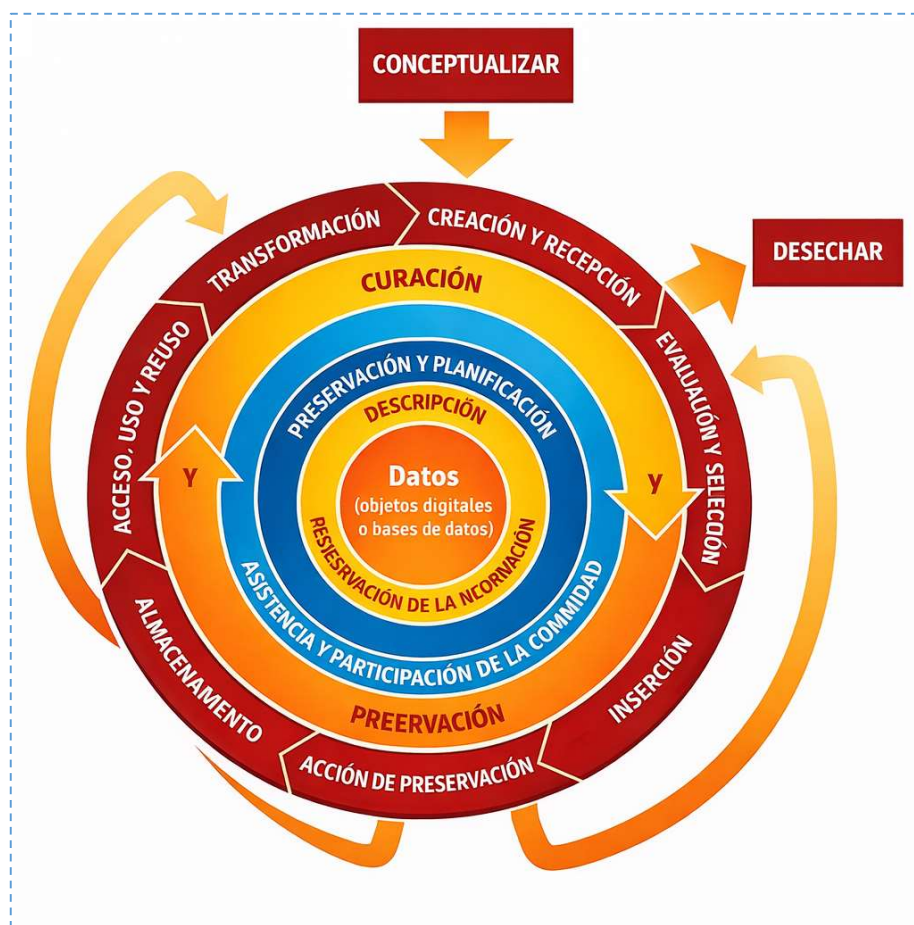


FIGURA 13 Modelo de ciclo de vida de curación de Digital Curation Centre DCC.

Fuente: Traducido de Sarah Higgins Digital Curation Centre.

Por otro lado, los objetos digitales complejos son objetos digitales creados mediante la combinación de varios otros objetos digitales, como sitios web. Las acciones secuenciales en el modelo son: conceptualización, creación/recepción, evaluación/selección, depósito, acción de preservación, almacenamiento y acceso, uso y reúso. Las acciones ocasionales incluyen desechar, reevaluar y migrar.

Las acciones de ciclo de vida completo incluyen asignar metadatos administrativos, descriptivos, técnicos, estructurales y de preservación, utilizando estándares apropiados para garantizar descripción y control adecuados a largo plazo. La planificación de la preservación a lo largo del ciclo de vida de curación de los datos digitales incluye planes para la gestión y administración de todas las acciones del ciclo de vida de curación. La asistencia y participación de la comunidad implica estar al tanto sobre las actividades comunitarias apropiadas y participar en el desarrollo de estándares compartidos, herramientas y software adecuado. Curar y preservar se enfoca en conocer y emprender acciones administrativas y de gestión

planificadas para promover la curación y preservación a lo largo del ciclo de vida de la curaduría.

En general, el modelo ayuda a definir funciones y responsabilidades, crear un marco de estándares y tecnologías, identificar pasos adicionales y garantizar que los procesos y políticas estén adecuadamente documentados. Al ser un modelo “ideal”, este debe modificarse en la práctica y, de acuerdo con las necesidades y contexto, los usuarios ingresan en cualquier etapa del ciclo (**Choudhury, Huang y Palmer, 2020**). Entre los resultados de este estudio, destaca la necesidad de actualizar el [modelo DCC](#) para abordar cambios en la curación de datos y la evolución de la de la tipología de datos, de preguntas de investigación y de procesos computacionales, así como los modos de curación que han surgido debido al auge de la ciencia de datos y los nuevos métodos de investigación intensivos en datos. Las recomendaciones de este estudio consideran:

- Incorporar procesos de curación iterativos, continuos y dinámicos a lo largo de todo el ciclo de vida, en lugar de tratar la curación como una fase única.
- Recomiendan adoptar un enfoque orientado al proceso, organizando el modelo en módulos, desarrollando un enfoque de árbol de decisiones e identificando mecanismos para gestionar posibles sesgos.
- Integrar técnicas de aprendizaje automático y otros métodos automatizados en el proceso de curación, reconociendo la importancia del aprendizaje automático en el procesamiento y análisis de datos.
- Abordar las preocupaciones relacionadas con la equidad, la responsabilidad y la transparencia (FAT por sus siglas en inglés), de los datos y los algoritmos. mediante la incorporación de estos principios en el proceso de curación.
- Reconocer la creciente importancia de las tecnologías de flujo de trabajo y los flujos de trabajo de investigación integrados y acomodarlos en el proceso de curación.
- Tener en cuenta el alcance, la complejidad y la prominencia cada vez mayores de los datos procesables por máquina, y adaptar las prácticas de curación en consecuencia.

Se enfatiza la necesidad de una curación más dinámica, iterativa e integrada con técnicas de aprendizaje automático, tecnologías de flujo de trabajo y principios de equidad, responsabilidad y transparencia, sugiriéndose que un modelo de ciclo de vida actualizado puede guiar las mejores prácticas al estimar costes de curación,

respaldando la integración de la evaluación y la mejora de datos FAT en los servicios de curación (Choudhury, Huang y Palmer, 2020).

3.3 Estándares y calidad de datos. FAIR

Wilkinson et al. (2016) publicaron los [“Principios rectores de FAIR para la gestión y custodia de datos científicos”](#) para ofrecer pautas de mejora de la localización, accesibilidad, interoperabilidad y reutilización (FAIR por sus siglas en inglés) de los datos digitales. Considerando el aumento en el volumen, la complejidad y la velocidad de creación de datos, los principios enfatizan la capacidad de los sistemas computacionales para encontrar, acceder, interoperar y reutilizar datos con mínima o nula intervención humana, dado que los científicos utilizan cada vez más el soporte informático para generar, procesar y gestionar datos. Desde entonces, [varios gobiernos y organizaciones de investigación suscribieron estos principios](#).

En el año 2020 se firmó la [“Declaración de la Sorbona sobre los Derechos a los Datos de Investigación”](#) por parte de la [Asociación de Universidades Americanas](#), la [Alianza de Universidades de Investigación Africanas](#), la [Coordinación de Universidades Francesas de Investigación Intensiva](#), [German U15](#), la [Liga de Universidades Europeas de Investigación](#), [Research Investigation 11](#) de Japón, [Grupo Russell](#) (Reino Unido), el [Grupo de los Ocho](#) de Australia, el [Grupo U15 de Universidades de Investigación Canadienses](#), con el respaldo de la [Alianza Europea de Hospitales Universitarios](#) (EUHA). Representan a más de 160 de las principales universidades de todo el mundo. En la declaración, las instituciones se comprometen a alentar a las universidades y sus investigadores a compartir datos tanto como sea posible; apoyar a las universidades y a sus investigadores para que sus datos sean localizables, accesibles, interoperables y reutilizables (FAIR), promover la curación y el intercambio de datos, el desarrollo de planes de gestión de datos como parte del proceso de investigación; involucrar a las instituciones en el reconocimiento apropiado para los investigadores que hacen que sus datos sean FAIR y los compartan con las licencias de datos abiertos apropiadas; promover que estos principios se integren en las políticas institucionales de datos de investigación; fomentar el establecimiento de programas de capacitación y desarrollo de habilidades que creen un entorno para promover la gestión abierta de datos de investigación en las universidades (Sorbonne Declaration on Research Data Rights, 2020).

Cada uno de los principios FAIR se desarrolla de la siguiente manera (Force11, 2020):



FIGURA 14 Principios FAIR. Fuente: elaboración propia.

También se especifica cómo se podrían aplicar en la práctica cada uno de los elementos en los principios. La localización de los datos usando identificadores únicos y persistentes a nivel mundial (principio F1) es lo más importante ya que es prerequisite para lograr otros aspectos de FAIR.

Los *identificadores persistentes* consisten en un enlace de internet como una URL que se dirige a una página web que contiene el conjunto de datos, los archivos de datos, los metadatos, etc. Los identificadores pueden ayudar a otros investigadores a comprender exactamente lo que quiere decir y permiten que las computadoras interpreten los datos de manera significativa. Además, los identificadores ayudan a otros a citar correctamente los datos al reutilizarlos (GO FAIR, 2020). Estos identificadores deben ser globalmente únicos, usando un servicio de registro que utiliza algoritmos para garantizar la unicidad de los identificadores recién generados. Un ejemplo son los DOIs asignados a través de [DataCite](#) o [Crossref](#). A su vez, los identificadores deben ser persistentes. Los servicios de registro garantizan que se pueda acceder al enlace en el futuro. Existen identificadores únicos para financiadores de las investigaciones ([Crossref Funder ID](#)),

identificadores de investigador ([ORCID](#)), identificador de organización ([ROR](#)), entre otros.

Un segundo elemento clave para crear datos FAIR son los *metadatos completos* (principio **F2**) que incluyen información descriptiva sobre el contexto, la calidad y el estado o las características de los datos. Si estos metadatos son ricos, permiten que un sistema informático realice automáticamente tareas de clasificación y priorización, optimizando el tiempo de los investigadores. **F2** ayuda a las personas a localizar los datos y aumentar la reutilización y las citas. Los metadatos enriquecidos se preparan con diversos usuarios en mente, no se sabe de antemano quién querrá usar los datos o con qué propósito. Para cumplir **F3**, los *metadatos* deben incluir de forma explícita el *identificador* global, único y persistente del conjunto que describen. En los repositorios de datos, este proceso es automático. Hay que tener presente que tener identificadores y descripciones ricas en metadatos no es suficiente para garantizar que se puedan encontrar. Incluso si los recursos de datos son de alta calidad, pueden pasar desapercibidos. Para que se descubran los conjuntos de datos **F4**, hay que dar a conocer su disponibilidad, especialmente a través de la [indexación en buscadores web](#).

Para mejorar la reutilización de datos, se debe adoptar un protocolo que sea de código abierto y gratuito (**A1.1**), permitiendo una fácil implementación en todo el mundo y facilite la recuperación de datos. La 'A' en FAIR no significa "abierto" o "gratis", sino que se refiere a brindar pautas claras sobre la accesibilidad de los datos. Esto significa que incluso los datos privados y protegidos pueden seguir considerándose FAIR siempre que se especifiquen las condiciones de accesibilidad. Idealmente, estas especificaciones deberían ser legibles por máquina, lo que permitiría la ejecución automatizada o alertas para los usuarios (**A1.2**). Con el tiempo, los conjuntos de datos pueden desaparecer debido al costo y a que un repositorio deje de ser sustentable. Como resultado, los enlaces pueden dejar de ser válidos y los usuarios pueden dedicar tiempo a buscar datos a los que ya no se puede acceder. Es mucho más simple y menos costoso almacenar metadatos. Como resultado, el principio **A2** enfatiza que los metadatos deben permanecer disponibles incluso si los datos ya no son compatibles.

El principio **I1** establece que, para permitir la legibilidad de los datos por parte de humanos y máquinas, se debe usar un *lenguaje para la representación del conocimiento* para dar una "comprensión común" de los objetos digitales. El

lenguaje debe tener una especificación formal, ser compartido y accesible, y estar diseñado para la interoperabilidad, por ejemplo, el esquema [Dublin Core](#), entre otros. Para describir los datos o metadatos, se usan *vocabularios* que brindan los términos o conceptos necesarios para representar su contenido. Sin embargo, es importante estar seguros de que estos vocabularios también sean FAIR (I2), para que otros puedan encontrarlos, acceder a ellos, interoperar y reutilizarlos. Esto significa que el vocabulario controlado utilizado para describir conjuntos de datos debe documentarse y resolverse mediante identificadores únicos y persistentes a nivel mundial. Además, esta documentación debe ser fácilmente localizable y accesible para cualquiera que utilice el conjunto de datos. Un ejemplo es el [AGROVOC Multilingual Thesaurus](#) que utiliza tecnologías de la web semántica, vinculando a otros sistemas de organización del conocimiento multilingües y construyendo puentes entre conjuntos de datos.

Una *referencia cualificada* es un tipo específico de referencia cruzada que proporciona información adicional sobre la relación entre los recursos de datos. El propósito de crearlas (I3) es mejorar el conocimiento contextual sobre los datos mediante la creación de conexiones significativas entre ellos. Al describir los datos, es útil especificar relaciones como dependencias, complementariedad y conexiones científicas, por ejemplo, enlace desde el dataset a un artículo que los usa. Además, todos los conjuntos de datos deben citarse correctamente y con sus identificadores únicos y persistentes a nivel mundial.

El principio R1.1 se centra en la *interoperabilidad legal*, lo que significa que los derechos de uso adjuntos a los datos deben describirse claramente para que los usuarios puedan cumplir con las restricciones de licencia. La falta de claridad en el estado de las licencias podría limitar la reutilización de los datos, especialmente a medida que las búsquedas automatizadas se vuelven más comunes. Por lo tanto, las condiciones bajo las cuales se pueden usar los datos deben ser fácilmente comprensibles tanto para humanos como para máquinas.

Se han desarrollado herramientas para evaluar el grado de cumplimiento FAIR. Un ejemplo es [F-UJI](#), desarrollada en el proyecto [FAIRsFAIR](#) ('Fostering FAIR Data Practices In Europe') que especifica el porcentaje de cumplimiento de FAIR con una explicación de cada una de las características cumplidas o no. Debemos tener en cuenta que en el sitio web de la herramienta se explica que el código está en proceso de desarrollo y que los resultados se basan en datos preliminares.

Summary:



Report:

Findable

FsF-F1-01D - Data is assigned a globally unique identifier.

FAIR level: 3 of 3

Score: 1 of 1

Output:

```
{
  "guid": "https://doi.org/10.57715/UNIR/EXIVRO",
  "guid_scheme": "doi"
}
```

Metric tests:	Test:	Test name:	Score:	Maturity:	Result:
	FsF-F1-01D-1	Identifier is resolvable and follows a defined unique identifier syntax (IRI, URL)	1	3	✓
		test target: https://f-uji.net/vocab/identifier			
		condition: any of members of test target			
		tested on: https://f-uji.net/vocab/metadata/property/object_identifier			
		comment: identifier can be given as user input			
	FsF-F1-01D-2	Identifier is not resolvable but follows an UUID or HASH type syntax	0		?
		test target: https://f-uji.net/vocab/identifier/unique			
		specifically: name=uuid, hash			

FIGURA 15 Parte de un reporte de la herramienta de evaluación de datos FAIR automatizada F-UJI.

Fuente: <https://www.f-uji.net/index.php>

3.4 Replicación, reproducibilidad y reutilización

Toca ahora presentar estos conceptos tan importantes en el contexto de los repositorios. Si bien existe cierta confusión (a veces contradicciones) en su uso (Barba, 2018; Plessner, 2018) es importante clarificar a qué nos referimos con “reproducir” y “replicar” la investigación científica y el término de “reutilización” de datos.

- *Reproducibilidad* es la capacidad de repetir el análisis y obtener los mismos resultados utilizando el conjunto de datos original y el código o métodos de análisis empleados por los investigadores originales. Esto implica que toda la información necesaria para reproducir los resultados, como los datos y los scripts de análisis, se comparten abiertamente en un repositorio, permitiendo que otros validen los resultados y construyan sobre el trabajo previo (Peng 2011). La *reproducibilidad* sirve como un estándar mínimo para evaluar afirmaciones científicas, especialmente cuando la replicación

independiente completa no es posible (**Peng**, 2011). Esto es especialmente valioso en el caso de datos computacionales que por su cantidad y complejidad son casi imposibles de replicar.

- *Replicabilidad*, en cambio, es el proceso de realizar un estudio independiente que sigue los mismos procedimientos del estudio original para verificar si se obtienen resultados similares. Para esto puede ser necesario recopilar nuevos datos bajo las mismas condiciones experimentales o metodológicas del estudio original y es considerado el estándar de oro para validar afirmaciones científicas (**King**, 1995) . La replicabilidad puede ser más compleja que la reproducibilidad, ya que en general requiere recursos significativos y, en algunos casos, puede ser prácticamente imposible debido a la naturaleza única de los datos o las condiciones experimentales (**King**, 1995; **Peng**, 2011)

Mientras que la *reproducibilidad* se centra en la transparencia y accesibilidad de los datos y métodos utilizados en un análisis, la *replicabilidad* se preocupa por la generalización y confirmación independiente de los resultados. Ambas ideas son esenciales para el avance de la ciencia y para aumentar la confianza en los resultados informados en la literatura científica. Entre los desafíos para la *reproducibilidad* se incluyen la necesidad de disponer del código de computadora y la adaptación del software utilizado para facilitar este análisis.

No existe una práctica científica que exija reproducibilidad para todas las afirmaciones científicas y plantea la necesidad de una infraestructura integrada para distribuir investigaciones reproducibles. Para superar estos obstáculos, se propone que los investigadores publiquen su código, aunque no sea perfecto, y que los editores de revistas lo exijan como rutina. Se sugiere la creación de repositorios centralizados para todos los datos, metadatos y códigos, lo que facilitaría la reproducibilidad en todas las disciplinas (**Peng**, 2011).

Los conjuntos de datos de replicación, incluyendo la información precisa para replicar resultados empíricos, han de crearse y ponerse a disposición del público, debiendo los autores referenciarlos en sus publicaciones originales. Los conjuntos de datos de replicación se aplican tanto a la investigación cuantitativa como cualitativa. Estos conjuntos de datos deben hacerse públicos y citarse en la publicación original. Si es necesario resguardar la confidencialidad de los datos,

estos pueden no estar públicos. Varios actores pueden fomentar la replicación de las investigaciones, además de los propios autores de los trabajos. Los programas de posgrado pueden exigir que los estudiantes de doctorado presenten los conjuntos de datos de replicación de sus trabajos para fortalecer las habilidades de los estudiantes y fomentar la adhesión a la norma. Los editores de revistas y revisores pueden exigir la adhesión a la norma de replicación. Los resultados publicados en artículos que no se pueden replicar suelen ser menos leídos, citados e investigados, lo que puede ser perjudicial para las carreras de los académicos (King, 1994).

Para que los datos sean verdaderamente útiles, deben ser reutilizados en nuevos contextos o combinados con otros datos para generar conocimiento. A diferencia del uso inicial, la reutilización exige documentación adecuada y, con frecuencia, el código y los modelos científicos asociados. Lo cierto es que se llegue a *reutilizar* un único conjunto de datos en su forma original puede ser complicado, incluso cuando se dispone de la documentación completa y de las herramientas necesarias. Esto se debe a que es preciso comprender en profundidad las razones subyacentes para la recolección de los datos, así como las diversas decisiones tomadas durante los procesos de recolección, limpieza y análisis de los datos. Estos aspectos son esenciales para poder aplicar correctamente los datos en nuevos contextos de investigación. Aún mayor es el desafío de combinar conjuntos de datos en una investigación. Para integrar y comparar conjuntos de datos de manera efectiva, es necesario tener información detallada sobre cada conjunto de datos individual. Se requiere conocer no solo las características técnicas de los datos, sino también los contextos específicos en los que se recolectaron, las metodologías aplicadas, los posibles sesgos y las limitaciones inherentes a cada conjunto de datos. Ese nivel de comprensión es necesario para garantizar que los resultados derivados de la combinación de datos sean válidos y confiables y para que las conclusiones extraídas sean sólidamente fundamentadas. Sin un conocimiento exhaustivo de cada conjunto de datos, las interpretaciones y las inferencias hechas a partir de la integración de datos podrían ser cuestionables (Pasquetto et al., 2017).

Leonelli y Lewandowsky (2022) describen seis formas de *reproducibilidad*, según los diversos contextos en los que se lleva a cabo la investigación científica:

1. *Computacional*: Común en la investigación computacional, implica volver a procesar los mismos datos a través de los mismos algoritmos para

- identificar y corregir errores en el análisis de datos y la programación. Requiere que el código y los datos sean reutilizables.
2. *Experimental directa*: Característica de los experimentos controlados (por ejemplo, ensayos clínicos), esta forma consiste en obtener los mismos resultados mediante la repetición exacta de los mismos métodos y procesos. Depende de protocolos, materiales y entornos estandarizados.
 3. *Indirecta*: Aplicable en experimentos "semiestandarizados" donde no todas las variables pueden ser controladas. Implica comparar datos de diferentes configuraciones para verificar si los resultados son consistentes, confirmando así la solidez de los resultados. Es típica en la investigación exploratoria o al estudiar fenómenos dependientes del contexto.
 4. De *alcance*: Relacionada con experimentos semiestandarizados, esta forma evalúa si los resultados se pueden obtener bajo diferentes condiciones para entender el alcance de los resultados, indicando si están limitados a condiciones específicas y aportando conocimiento sobre los fenómenos estudiados.
 5. *Hipotética*: Cuando los métodos y contextos de investigación no pueden replicarse fácilmente, esta forma se centra en lograr resultados que coincidan con las predicciones basadas en resultados previos, confirmando de manera indirecta estos resultados a través de cadenas de razonamiento sólidas.
 6. *Experiencia reproducible*: En entornos de investigación dinámicos, donde los métodos requieren una adaptación constante, la expectativa es que la experiencia aplicada sea reproducible. Esto significa que investigadores igualmente capacitados deberían obtener resultados comparables, identificando patrones e ideas a pesar de las variaciones de contexto. Esto es especialmente relevante en campos como la antropología, donde la reproducción directa de resultados no es factible.

En el informe de **Leonelli y Lewandowsky** (2022) se reconoce el rol de los gestores de datos como parte integral del ecosistema investigativo, dejando claro que no alcanza con subir datos a un repositorio, sino que su curación es esencial. Que los investigadores y los curadores entiendan estos conceptos es importante, por varias razones ya señaladas en el apartado sobre ciencia abierta y que se enfatizan aquí:

1. **Garantizar la integridad científica**: Si cumplen con estándares y buenas prácticas, los repositorios de datos ofrecen registros verificables y

persistentes de datos y metodologías usadas en las investigaciones.

2. **Facilitar la validación y verificación:** Permiten a otros investigadores validar y verificar los resultados de estudios publicados a través de la reproducción, contribuyendo a la confianza en los hallazgos reportados.
3. **Promover el avance del conocimiento:** Al permitir la reutilización de datos para explorar nuevas hipótesis o aplicaciones, los repositorios de datos amplían el valor de la investigación original y fomentan el descubrimiento científico.
4. **Optimizar recursos:** La reutilización eficiente de datos puede reducir la necesidad de recoger nuevos datos, lo que ahorra tiempo y recursos, y minimiza la redundancia en la recogida de datos.
5. **Establecer transparencia y apertura:** Fomentan una cultura de transparencia y apertura al hacer que los datos y métodos estén accesibles para escrutinio y uso por parte de la comunidad científica y el público.
6. **Favorecer la colaboración interdisciplinaria** al poner a disposición datos que pueden ser integrados y analizados por equipos de diferentes disciplinas.
7. **Cumplir los requisitos** de leyes, de agencias de financiamiento y de revistas.
8. **Mejorar la educación y formación:** Los conjuntos bien curados son recursos valiosos para la educación, permitiendo el aprendizaje mediante la reutilización y análisis de datos reales.
9. **Contribuir a las políticas científicas y de inversión** al demostrar el impacto y el uso de los fondos de investigación, y facilitar la evaluación de los resultados de la investigación basada en datos compartidos y reutilizados.

Para lograr que las investigaciones se puedan reutilizar es necesaria la curación de datos, un rol que asumen los curadores de datos en las instituciones y que se detalla a continuación.

3.5 Necesidad de curadores de datos

En numerosas instituciones académicas, aún queda un largo camino por recorrer en términos de cambios culturales, tanto en los procesos como en la infraestructura relacionada con la gestión y publicación de datos. Esto ha dado lugar a que los datos de baja calidad resulten costosos en términos de inversión en investigación, pérdida de oportunidades estratégicas y menor retorno de la inversión en ciencia y tecnología. A pesar del desarrollo de valiosos métodos y herramientas para el control de la calidad de los datos, estos no han sido

ampliamente adoptados, en parte debido a la falta de roles de gestores de datos claramente definidos y a que los problemas de datos a menudo se consideran cuestiones de infraestructura tecnológica.

Otro problema es que los datos de investigación muchas veces no están vinculados a las publicaciones, lo que dificulta o imposibilita volver a analizar o verificar los resultados. Mientras tanto, existen preocupaciones sobre la seguridad de los datos personales y confidenciales almacenados en servicios en la nube por diversas aplicaciones de software (Wildgaard y Rantasaari, 2022).

Los datos que se comparten con información insuficiente no son confiables y la falta de conocimientos y habilidades para curar los conjuntos de datos a fin de compartirlos (Bongiovani y Martínez Uribe, 2014), se ha identificado como una razón por la cual los investigadores eligen no compartir sus datos (Volk et al., 2014; Kim et al., 2018). Una encuesta a 1,372 investigadores de diferentes países encontró que se requiere una mayor participación institucional en la capacitación y el apoyo a las buenas prácticas de datos, lo que permitiría generar cambios positivos significativos. Es evidente la necesidad de contar con gestores o bibliotecarios de datos, repositorios de fácil acceso para el almacenamiento a corto y largo plazo, y programas educativos que promuevan buenas prácticas de gestión de datos (Tenopir et al., 2020).

3.5.1 ¿Qué es un curador de datos ('data steward')?

Un curador de datos es el profesional encargado de garantizar la calidad, integridad y accesibilidad de los datos y metadatos, de acuerdo con la legislación, las políticas institucionales y los permisos establecidos. Su labor abarca todo el ciclo de vida de la investigación, desde el diseño y la recopilación hasta el procesamiento, análisis, preservación, difusión y reutilización de los datos, contribuyendo también a la reproducibilidad de los resultados (Jetten et al., 2021).

Para garantizar una gestión adecuada de los datos, el curador se debe responsabilizar de la implementación de políticas y estándares basados en las mejores prácticas. Actúa como un enlace entre el departamento de sistemas, los creadores de datos y la institución, y se considera "humanware"³¹, porque la

³¹ Término informático empleado para definir los recursos humanos de un sistema informático, o el hardware y el software diseñado, pensando en la experiencia y la interfaz que le dará el usuario final.

transformación hacia una cultura basada en datos requiere más que solo tecnología. Un curador de datos puede desarrollar, proporcionar y conectar la infraestructura (hardware y software), políticas, mejores prácticas, procesos analíticos y datos. Los administradores de datos pueden especializarse en diferentes áreas, como políticas, infraestructura, análisis de datos o datos disciplinarios.

Los curadores pueden brindar su experiencia en varios roles, ya sea integrados en una organización o como proveedores de servicios genéricos. **Mons** (2018) enfatiza la importancia de desarrollar capacidades y promover la colaboración entre grupos y organizaciones para establecer buenas prácticas de curación de datos como norma. Sin embargo, determinar las responsabilidades, habilidades y requisitos específicos de un curador puede ser complejo. Como iniciativa de apoyo, respaldada por importantes universidades de Estados Unidos, surgió la [Data Curation Network](#)³² (DCN), iniciativa basada en la premisa de que las competencias necesarias para la curación de datos son diversas y abarcan una amplia gama de tipos y formatos específicos para cada disciplina, incluyendo datos geoespaciales, códigos, bases de datos, espectros químicos, imágenes tridimensionales y datos de secuenciación genómica. Dada esta diversidad, ningún repositorio individual puede contar con todos los expertos necesarios para la curación de datos (**Data Curation Network**, 2022).

La red DCN aprovecha los conocimientos de especialistas en curación de datos de diferentes disciplinas y promoviendo oportunidades de capacitación y desarrollo profesional de los curadores de datos. Esa colaboración permite abordar de manera colectiva y eficiente una variedad más amplia de tipos de datos, superando las capacidades que cualquier institución por sí sola podría ofrecer (**Johnston et al.**, 2018).

La red ha elaborado una serie de materiales útiles para profesionales dedicados a la curación de datos tales como: "[CURATE\(D\): Checklist for Data Curation](#)" que contiene listas de comprobación para revisar que los datasets estén correctamente curados. Además, la red creó los "[Data Curation Primers](#)", documentos revisados por pares que se focalizan en un tema específico, un área disciplinaria o tarea como

³² Organización de membresía compuesta por repositorios de datos institucionales y sin fines de lucro que colaboran para avanzar en la investigación abierta. Su objetivo principal es hacer que los datos de investigación sean más éticos, reutilizables y comprensibles.

referencia para curar datos de investigación, por ejemplo, bases de datos, ensayos clínicos, formatos específicos, entre otros, y que están abiertos a la comunidad mundial de curadores de datos.

3.5.2 Curación de datos en bibliotecas y repositorios

Tenopir et al. (2014) analizaron la emergencia de servicios de gestión de datos de investigación en bibliotecas académicas de Estados Unidos y la percepción de los bibliotecarios. Examinaron las 223 respuestas a una encuesta completada por directores de bibliotecas. Los servicios identificados eran diversos e incluían desde servicios de referencia, como ayudar a encontrar un lugar para depositar datos de investigación o brindar ejemplos de planes de gestión de datos, hasta servicios técnicos más prácticos, como aprender a usar un repositorio de datos institucional o ayudar a los investigadores a desarrollar planes de gestión de datos.

El estudio concluía afirmando que las bibliotecas estaban incrementando su apoyo en la planificación, gestión y publicación de datos de investigación como parte de un proceso emergente. **Lee y Stvilia** (2017) realizaron un estudio cualitativo-descriptivo con el objetivo de analizar las prácticas de curación de datos de investigación en los repositorios institucionales y brindar descripciones detalladas de estas. Los datos obtenidos a partir de entrevistas con personal de repositorios de 13 instituciones diferentes mostraron una variedad de tareas relacionadas con los datos de investigación dentro de los RI, incluidas diversas acciones de curación de datos, así como otras actividades relacionadas.

Las actividades de curación de datos se relacionan principalmente con el ciclo de vida de los datos de investigación, que, como se desarrolló previamente, incluye la conceptualización, planificación, creación, carga y publicación. Otras actividades relacionadas facilitan o promueven la gestión de datos y su reutilización a través de los RI, como el análisis de datos, el desarrollo de políticas y la educación, tal como muestra la **tabla VI**.

Actividades	Acciones
Entender las necesidades de curación de datos	• Entrevistar a investigadores • Comunicarse con el personal de IR o de la biblioteca • Consultar con investigadores
Gestionar y compartir datos	• Recibir o transferir archivos de datos • Limpiar datos • Convertir datos a un formato de archivo diferente • Desarrollar y agregar metadatos • Validar datos • Empaquetar datos • Subir y publicar datos en el repositorio
Asegurarse que los datos sean accesibles y reutilizables	• Anotar datos para entidades relevantes • Optimizar datos para motores de búsqueda • Mantener los datos actualizados en el repositorio espejo
Reevaluar datos para la preservación a largo plazo	• Seleccionar conjuntos de datos para preservación a largo plazo
Otras Actividades y Acciones Relacionadas	
Analizar el uso de los datos	• Gestionar estadísticas descriptivas del uso de datos • Brindar a los investigadores datos de seguimiento
Crear políticas e infraestructura administrativa	• Entender las necesidades locales y crear políticas y reglas locales • Construir componentes de infraestructura
Capacitar a las personas sobre la gestión de datos	• Capacitar a bibliotecarios • Educar a investigadores • Ofrecer talleres para herramientas de análisis de datos • Realizar actividades de divulgación sobre curación de datos
Capacitarse de forma permanente	• Aprender las mejores prácticas para la gestión de datos de investigación • Aprender sobre tecnologías futuras

TABLA VI Actividades de curación de datos y sus acciones correspondientes.

Fuente: Lee y Stvilia (2017).

Un primer paso elemental del proceso de curación de datos, son las *entrevistas* con los investigadores para comprender las necesidades de curación específicas. Esto implica dialogar directamente con ellos para entender la naturaleza de sus datos, cómo se generan, cuáles son sus características particulares y qué requisitos

específicos tienen para su gestión y preservación a largo plazo. Una entrevista permite obtener información detallada que no sería accesible a través de métodos menos directos. Otra acción es la *comunicación* con otros miembros del personal encargado de repositorios para coordinar y planificar la gestión y curación de los datos según las necesidades identificadas en esas entrevistas. Esto puede incluir discusiones sobre las capacidades del repositorio, la asignación de recursos para la curación de datos y la creación de estrategias para abordar desafíos específicos que se identifiquen.

La consulta con los investigadores continúa más allá de las entrevistas iniciales, son necesarias para aclarar dudas durante el proceso de curación de datos, ajustar las estrategias según evolucionan los requisitos de los datos, brindar actualizaciones a los investigadores y recibir retroalimentación del proceso. Esto asegura que la curación se ajuste continuamente a las necesidades y expectativas de los investigadores y que los datos estén relevantes y utilizables en el tiempo. La comunicación efectiva y la colaboración entre el personal del repositorio, los bibliotecarios y los investigadores son clave del éxito de estos procesos. Los curadores de datos realizan diversas tareas, dependiendo de las características y contexto de la institución donde se desempeñen (Lee y Stvilia, 2017):

- Consultar con proveedores de datos:
 - Comunicarse con investigadores y otros proveedores de datos para entender sus necesidades de curación de datos.
 - Conectar a los autores de los datos con especialistas de área (bibliotecarios especializados), especialistas en metadatos o el administrador del RI.
- Facilitar la comunicación:
 - Actuar como intermediarios entre diferentes áreas y departamentos dentro de la institución.
 - Facilitar la comunicación técnica entre las partes interesadas.
- Evaluar los datos de investigación:
 - Revisar y evaluar los datos de investigación para determinar si deben seguir siendo mantenidos para la preservación a largo plazo o si deben ser descartados.
- Planificar la estructura de gobernanza de datos:
 - Desarrollar y planificar la estructura de gestión de datos para los repositorios institucionales.

- Mantenerse al día con las tendencias actuales y futuras en la comunidad de curación de datos de investigación.
- Colaborar en la selección de datos para preservación a largo plazo:
 - Trabajar en conjunto con archiveros digitales y bibliotecarios especialistas en la materia para seleccionar conjuntos de datos para la preservación.
- Desempeñar responsabilidades de liderazgo:
 - En algunas instituciones, los curadores de datos realizan tareas típicamente asociadas con el rol de dirección.
- Capacitar y realizar divulgación:
 - Enseñar a otros miembros del personal sobre la gestión de datos de investigación.
 - Realizar actividades de divulgación para los servicios de IR y educar a la comunidad del campus.
- Responder a consultas:
 - Atender preguntas y solicitudes de consultas sobre el uso de IR y la gestión de datos.
- Gestionar metadatos y carga de datos:
 - Ayudar a los autores de datos a crear y agregar metadatos apropiados para sus proyectos de investigación.
 - Cargar datos en los repositorios.

Todas estas tareas dan cuenta del rol polifacético que cumplen los curadores de datos para apoyar las infraestructuras de investigación de los repositorios, asegurando que los datos sean gestionados, preservados y accesibles de acuerdo con las mejores prácticas y necesidades institucionales. **Johnston et al. (2018)** describen procedimientos estandarizados y específicos de DCN que se usan para revisar los datos para cada tipo de archivo, en función de su reutilización futura, empleando habilidades expertas y conocimientos específicos del dominio, proponiendo los pasos de curación **DCN C-U-R-A-T-E**:

1. **Check** (Verificar): Revisar los archivos de datos y leer la documentación.
2. **Understand** (Entender): Intentar entender los datos.
3. **Request** (Solicitar): Pedir al creador de los datos la información faltante o cambios.
4. **Augment** (Ampliar): Ampliar con metadatos para facilitar su localización.
5. **Transform** (Transformar): Cambiar formatos de archivo para reutilización y

preservación a largo plazo.

6. **Evaluate** (Evaluar): Evaluar y calificar el conjunto de datos general por su conformidad con los principios FAIR.

DCN ha desarrollado una amplia documentación y serie de recursos para ayudar en el desarrollo profesional de los curadores de datos, destacando: Lista de Verificación CURATED, Módulos de Aprendizaje en Línea CURATED, Primeros Pasos en la Curación de Datos, Glosario de Curación, Encuesta de Curación de Datos Institucionales, Lista de Herramientas de Software para Curación de Datos, entre otros³³. Específicamente para los curadores de datos que usan el software [Dataverse](#), y basándose en los pasos de DCN, **Cooper et al.** (2021) en Canadá y para la red CARL Portage, desarrollaron una guía de curación para diferentes niveles:

1. Se detallan los pasos mínimos necesarios para publicar con éxito en [Dataverse](#), de manera que el conjunto de datos sea localizable. Por ejemplo, asegurarse que el conjunto de datos haya sido enviado a la colección correcta y que los campos de metadatos requeridos sean precisos.
2. Actividades que mejoran el descubrimiento de los datos y ayudan a asegurar su usabilidad a lo largo del tiempo. Esto puede incluir, por ejemplo, asegurar que los campos de metadatos recomendados estén completos y que el conjunto de datos incluya documentación suficiente para que un usuario con conocimientos similares pueda entender el conjunto de datos, abrir y usar los archivos.
3. Comprende acciones de curación intensivas destinadas a preparar conjuntos de datos para su preservación y mejorar las posibilidades de que los datos y el código puedan ser usados para reproducir o replicar un estudio. Esto significa mejorar la documentación, revisar el contenido de los archivos y el código, y transformar los archivos de datos en formatos adecuados para la preservación a largo plazo.

Cada nivel en la curación de datos demuestra una progresiva complejidad y detalle, empezando con los requisitos elementales para la publicación en [Dataverse](#) (nivel 1), avanzando hacia acciones más complejas destinadas a la preservación y reutilización avanzada (nivel 3). La implementación efectiva de estos niveles por parte de una institución está supeditada por la disponibilidad de recursos y la capacidad de sus servicios de curación de datos (**Cooper et al.**, 2021). Por ejemplo,

³³ Toda esta información y recursos está disponible en <https://datacuration.network/curator-resources/>

en el nivel 3, y referido a los formatos de archivos, se detallan como opciones de verificación por parte de los curadores el identificar los formatos de archivo y el software utilizado para crear los archivos:

- El conjunto de datos contiene formatos de archivo no propietarios o formatos que son amigables para la preservación.
- Los archivos propietarios pueden ser transformados a formatos no propietarios sin perder datos.
- Si los archivos no pueden ser transformados, la documentación describe: ¿El formato de archivo y el contenido de los archivos? ¿El software o instrumento que generó los archivos? ¿El software necesario para ver o usar los archivos? ¿Una opción de software gratuito para acceder a los archivos, si está disponible?
- La documentación describe los datos, las fórmulas utilizadas, los encabezados de columnas, las etiquetas de variables, etc., de modo que los datos sean utilizables.
- La documentación está en un formato preferido (por ejemplo, PDF/A o .txt)."

Cada repositorio define su política de metadatos y documentación requerida y recomendada, así como las directrices para curadores.

Un aspecto importante es la capacitación que necesitan los curadores de datos. **Boté-Vericad et al. (2023)**, a partir de entrevistas a gestores de repositorios de datos, identificaron una serie de competencias y habilidades necesarias para el desempeño efectivo del rol de gestor de repositorios de datos:

- **Técnicas:** habilidades de programación, ingeniería de software e ingeniería informática; competencias en el uso de software especializados, manejar sistemas operativos menos populares como Linux, conocimiento profundo de bases de datos SQL.
- **Análisis y gestión de datos:** relacionadas con la ingesta, procesamiento, interpretación y visualización de datos del sistema; conocimiento de tecnologías de la web semántica.
- **Alfabetización informacional y bibliotecología:** conocimientos de la propiedad intelectual, competencias en recuperación de información; familiaridad con la alfabetización informacional.
- **Transversales:** competencias aplicables a otros campos de naturaleza

multidisciplinaria, habilidades organizativas, experiencia, aptitud para el aprendizaje colaborativo y capacidad para el aprendizaje a lo largo de la vida, dado el cambio constante en el ámbito de la gestión de datos.

Esta serie de competencias se consideran fundamentales para gestionar eficazmente los repositorios de datos, lo que implica no solo el manejo técnico de los datos e infraestructuras, sino también la capacidad de aprendizaje continuo.

3.5.3 Capacitación de investigadores

Uno de los aspectos fundamentales del éxito de los repositorios de los repositorios de datos es la formación de los investigadores en gestión y publicación de datos. La revisión sistemática de la literatura de **Oo et al. (2022)** examinó las características de programas exitosos de formación en gestión de datos de investigación, proporcionando una base útil para instituciones que deseen desarrollar e implementar capacitaciones efectivas en esta materia para su personal y estudiantes, siguiendo las mejores prácticas en el campo. En primer lugar, se destaca la importancia de crear programas de capacitación centrados en las necesidades específicas, personalizando las capacitaciones y contextualizándolas para investigadores que inician su doctorado o para investigadores avanzados en la carrera académica: También encontraron la necesidad de crear actividades interactivas con un enfoque pedagógico educativo.

Con el objeto de poder satisfacer los diferentes niveles de conocimiento y etapas en la carrera investigadora, se han implementado diferentes formas de capacitación con enfoques diversos, tales como sesiones de 15 minutos de café o almuerzo con “Herramientas y Trucos” hasta talleres de día completo. Existen manuales y buenas prácticas de capacitación que han sido implementadas en instituciones, como los recogidos en el libro *Engaging Researchers with Data Management: The Cookbook* (**Clare et al., 2019**) con casos compilados por expertos de la RDA para apoyar la capacitación a investigadores con experiencias que han sido exitosas. También hay manuales específicos sobre cómo incorporar la gestión de datos en el currículum de una institución de educación superior, por ejemplo: **Demchenko et al. (2021)** y **Engelhardt (2022)**. Se ha desarrollado un [catálogo de formación y soporte de gestión de datos de investigación \(RDM\) de EOSC-Pillar](#), colección de recursos en línea para apoyar en la capacitación. También incluye recursos operacionales y fácilmente disponibles para el uso diario que pueden ser utilizados por los

administradores de datos para apoyar a los investigadores. Se puede buscar por etapa del ciclo de vida de la investigación (planificación y diseño, publicación, etc.), por tipo de material (curso, video, juego, guía, herramienta, etc.), público objetivo (investigador, bibliotecario, nivel básico, disciplina, entre otros), formato, licencias, entre otras características. Brinda acceso, por ejemplo, al texto de referencia [Foster Open Science Training Toolkit](#), que cuenta con una [versión en español](#) también.

Es necesario integrar la gestión de datos en la formación académica de grado y sobre todo posgrado, particularmente en los cursos de metodologías de investigación. Esta integración ayudaría a los estudiantes a aprender a “tratar sus datos de investigación como activos valiosos que deben gestionarse, compartirse, descubrirse y reutilizarse” (Borgman y Bourne, 2022). También se plantea tener en cuenta las diferencias disciplinarias, observando que, mientras que en las ciencias sociales es común enseñar gestión de datos en cursos de metodología de investigación, en áreas de las ciencias biomédicas la enseñanza formal de métodos de investigación es menos común. En muchos campos donde el diseño de la investigación se aprende principalmente a través de mentores o directores, los nuevos investigadores dependen de los conocimientos y prácticas de estos mentores, y hay menos mecanismos para desarrollar habilidades sobre estándares y prácticas necesarias para compartir datos. Por tanto, para que no quede a criterio o experiencia de cada mentor, se recomienda incluir la gestión de datos en los trayectos curriculares de doctorados.

Por su parte, las bibliotecas universitarias han asumido un rol fundamental en la capacitación de investigadores en la elaboración de planes de gestión de datos y en capacitaciones específicas, particularmente para cumplir con las políticas de organismos de financiamiento. Sumar a los profesionales de la información es estratégico en la formación, como ya se está haciendo en el [Repositorio de Datos Académicos RDA-UNR](#).

3.6 Valor de los repositorios datos

Las infraestructuras del conocimiento (IC), según Borgman (2020), son sistemas complejos que abarcan una variedad de elementos y procesos, incluyendo aspectos humanos, tecnológicos y organizativos que contribuyen a la generación, intercambio y mantenimiento del conocimiento.

De acuerdo con **Borgman et al. (2019)**, los repositorios de datos digitales representan inversiones a largo plazo que generarán beneficios a lo largo del tiempo. A pesar del costo que implican, al ser intensivos en mano de obra y la dificultad de medida de su impacto, son infraestructuras que deben demostrar constantemente su valor a sus comunidades. Sin embargo, como instituciones de información, su valor principal se encuentra en las futuras generaciones. El valor de los archivos de datos digitales que comparten datos para su reutilización de datos solo puede evaluarse con una perspectiva de largo plazo. Los repositorios de datos, como IC:

- **Ayudan en el almacenamiento y gestión** del conocimiento: Los repositorios de datos almacenan y gestionan cantidad y variedad de datos, que son una forma crítica de conocimiento en muchas disciplinas científicas y académicas.
- **Facilitan la intermediación** de conocimiento: Ofrecen un dispositivo para que investigadores y otros usuarios compartan, accedan y reutilicen datos, lo que fomenta la colaboración y el desarrollo del conocimiento.
- **Requieren componentes tecnológicos y humanos**: Involucran tanto la tecnología (en términos de infraestructura de TI para almacenar, preservar y acceder a datos) como los aspectos humanos (a través de la gestión de datos, curación y soporte a los usuarios).
- **Apoyo a la investigación y educación**: Los repositorios de datos ofrecen recursos esenciales para el aprendizaje, la enseñanza y la investigación científica.
- **Integración con redes de conocimiento más amplias**: Los repositorios de datos se integran (a través de la interoperabilidad, búsqueda y recuperación de información) con redes más amplias de infraestructuras de conocimiento, conectando a investigadores, instituciones académicas, bibliotecas, y otros en el ecosistema del conocimiento.

Los repositorios de datos son ejemplos claros de infraestructuras del conocimiento, como entidades esenciales para el avance y la distribución del conocimiento intensivo en datos y compromiso con las futuras generaciones. La investigación científica actual está cambiando a un paradigma de ciencia y colaboración más intensiva en el uso de datos, tanto en cantidad como en variedad de estos. La complejidad y multidimensionalidad de los problemas científicos requiere poder acceder a nuevos instrumentos y tecnologías científicas, compartir los altos costos

de estas tecnologías, compartir datos y conocimientos, desarrollar enfoques multidisciplinarios para encontrar respuestas satisfactorias a interrogantes globales tales como el cambio climático, la pobreza e inequidad, el agotamiento y la pérdida de biodiversidad, las pandemias, entre otros.

3.7 La autora

Doctoranda en Documentación: Archivos y Bibliotecas en el Entorno Digital en la [Universidad Carlos III de Madrid](#), Master of Library and Information Science (MLIS) por la [University of South Carolina](#) (USA) y Licenciada en Comunicación Social por la [Universidad Nacional de Rosario](#) (UNR).

Es docente e investigadora de la Facultad de Humanidades de la UNR y asesora al Repositorio Institucional. Miembro del Comité de Expertos del Sistema Nacional de Repositorios Digitales (SNRD)

del Ministerio de Ciencia, Tecnología e Innovación Productiva (MINCyT) de Argentina, habiéndose desempeñado como coordinadora de este sistema. Participó como representante técnica para el proyecto Red Federada Latinoamericana de Repositorios Institucionales (LA REFERENCIA) y para la Confederación de Repositorios de Acceso Abierto (COAR). Docente del curso “Diseño e Implementación de Repositorios Institucionales de Acceso Abierto” ofrecido a sus socios el proyecto LA REFERENCIA y Red CLARA. Es coordinadora de la Lista Latinoamericana sobre Acceso Abierto y Repositorios (LLAAR). Dirige proyectos de investigación, publica sobre comunicación científica y acceso abierto y participa activamente en congresos de la especialidad.



FIGURA 16 La autora.

4. Referencias

- Abadal, Ernest. & Anglada, Lluís. (2020). Ciencia abierta: cómo han evolucionado la denominación y el concepto. *Anales de documentación* (Vol. 23, No. 1). <https://doi.org/10.6018/analesdoc.378171>
- Abadal, Ernest; Abad García, María Francisca; Anglada i de Ferrer, Lluís; Boté-Vericad, Juan José; Esteve, Asunción; González-Teruel, Aurora; Labastida i Juan, Ignasi; López-Borrull, Alexandre; Ollé i Castellà, Candela; Melero Melero, Remedios; Rodríguez-Gairín, Josep-Manuel & Santos-Hermosa, Gema. (2023). *Ciencia abierta en España 2023: informe de situación y análisis de la percepción*. <http://hdl.handle.net/2445/200020>
- Agosti, Donald; Alonso Garcia, Enrique; Appleyard, Baden; Bruch, Christoph; Chen, Robert; Clement, Gail; Egloff, Willi; Gruttemeier, Herbert; Hodson, Simon; Llorent, Maria; Minster, J. Bernard & Uhler, Paul F. (2016). *Legal interoperability of research data: principles and implementation guidelines*. RDA-CODATA Legal Interoperability Interest Group. Version 1.0. https://gfzpublic.gfz-potsdam.de/rest/items/item_2042216_2/component/file_2042217/content
- Akdeniz, Esra; Borschewski, Kerrin Emilia; Breuer, Johannes & Voronin, Yevhen. (2023). Sharing social media data: The role of past experiences, attitudes, norms, and perceived behavioral control. *Frontiers in big Data*, 5, 971974. <https://doi.org/10.3389/fdata.2022.971974>
- Ali-Khan, Sarah E.; Jean, Antoine; MacDonald, Emily & Gold, E. Richard (2018). Defining Success in Open Science. *MNI Open Research*, 2(2), 2. <https://doi.org/10.12688/mniopenres.12780.2>
- Anglada, Lluís, & Abadal, Ernest. (2018). ¿Qué es la ciencia abierta? *Anuario ThinkEPI*, 12, 292-298. <https://doi.org/10.3145/thinkepi.2018.43>
- Allen, Christopher & Mehler, David M. (2019). Open science challenges, benefits and tips in early career and beyond. *PLoS biology*, 17(5), e3000246. <https://doi.org/10.1371/journal.pbio.3000246>
- Alexandre-Benavent, Rafael; Vidal-Infer, Antonio; Alonso-Arroyo, Adolfo; Peset, Fernanda & Ferrer Sapena, Antonia. (2020). Research data sharing in Spain: Exploring determinants, practices, and perceptions. *Data*, 5(2), 29. <https://doi.org/10.3390/data5020029>
- Alperín, Juan Pablo; Fischman, Gustavo E. y Willinsky, John. (2008). Open access and scholarly publishing in Latin America: Ten flavours and a few reflections. *Liinc em Revista*, 4(2). <https://purl.stanford.edu/kj464qf4719>
- Alperín, Juan Pablo; Fischman, Gustavo E. y Willinsky, John. Scholarly communication strategies in Latin America's research-intensive universities. *Educación Superior y Sociedad*, 16(2). (2011). <https://doi.org/10.54674/ess.v16i2.659>

- Alperín, Juan Pablo y Fischman, Gustavo E. (2015). *Hecho en Latinoamérica: acceso abierto, revistas académicas e innovaciones regionales*. Clacso. <https://biblioteca.clacso.edu.ar/clacso/se/20150722110704/HechoEnLatinoamerica.pdf>
- Alperín, Juan Pablo. (2022). Why I think ending article-processing charges will save open access. *Nature*, 610(7931), 233-233. <https://doi.org/10.1038/d41586-022-03201-w>
- Andrews Mancilla, Heather; Teperek, Marta; van Dijck, Jasper; Den Heijer, Kees; Eggermont, Robbert; Plomp, Esther; Turkyilmaz-van der Velden, Yasemin & Kurapati, Shalini. (2019). On a quest for cultural change-surveying research data management practices at Delft University of Technology. *LIBER Quarterly*: 29(1), 1-27. <https://doi.org/10.18352/lq.10287>
- Andrews Mancilla, Heather. (2023). *Programa de Capacitación sobre Gestión de Datos de Investigación en la Práctica Científica*. LA Referencia, Research Data Alliance, CONARE. <https://www.lareferencia.info/es/blog-masonry/noticias/item/310-programas-de-capacitacion-sobre-datos-de-investigacion-2023>
- Anger, Michael; Wendelborn, Christian; Winkler, Eva C. & Schickhardt, Christoph. (2022). Neither carrots nor sticks? Challenges surrounding data sharing from the perspective of research funding agencies. *Plos one*, 17(9), e0273259. <https://doi.org/10.1371/journal.pone.0273259>
- Anglada, Lluís & Abadal, Ernest. (2023). Open access: a journey from impossible to probable, but still uncertain. *El profesional de la información*, 32(1). <https://doi.org/10.3145/epi.2023.ene.13>
- Arza, Valeria & Fressoli, Mariano. (2017). Systematizing benefits of open science practices. *Information Services & Use*, 37(4), 463-474. <https://doi.org/10.3233/ISU-170861>
- Asai, Sumiko. (2021). An analysis of revising article processing charges for open access journals between 2018 and 2020. *Learned Publishing*, 34(2), 137-143. <https://doi.org/10.1002/leap.1334>
- Asai, Sumiko. (2023). Determinants of article processing charges for hybrid and gold open access journals. *Information Discovery and Delivery*, 51(2), 121-129. <https://doi.org/10.1108/IDD-09-2021-0098>
- Austin, Claire C.; Bloom, Theodora; Dallmeier-Tiessen, Sünje; Khodiyar, Varsha K.; Murphy, Fiona; Nurnberger, Amy; Raymond, Lisa; Stackhouse, Martina; Tedds, Jonathan; Vardigan, Mary & Whyte, Angus. (2017). Key components of data publishing: using current best practices to develop a reference model for data publishing. *International Journal on Digital Libraries*, 18, 77-92. <https://doi.org/10.1007/s00799-016-0178-2>

- Babini, Dominique. (2019). La comunicación científica en América Latina es abierta, colaborativa y no comercial. Desafíos para las revistas. *Palabra Clave* (La Plata), 8(2), e065. <https://doi.org/10.24215/18539912e065>
- Babini, Dominique & Rovelli, L. (2020). *Tendencias recientes en las políticas científicas de ciencia abierta y acceso abierto en Iberoamérica*. CLACSO-Fundación Carolina. https://www.memoria.fahce.unlp.edu.ar/art_revistas/pr.16667/pr.16667.pdf
- Balparda, Laura Rita; del Valle, Héctor; López, Diego Alejandro Germán; Torralba, María Cecilia; Tazzioli, Florencia; Ciattaglia, Brunela; Vicioso, Benito; Peña, Hector; Delorenzi, Dardo Javier; Solís, Tania. (2023). *Datos de: Huella Urbana de la Ciudad de Rosario, Santa Fe, Argentina* [Dataset]. RDA UNR. <https://doi.org/10.57715/UNR/EXIVRO>
- Bangani, Siviwe & Moyo, Mathew. (2019). Data sharing practices among researchers at South African Universities. *Data Science Journal*, 18, 28-28. <https://doi.org/10.5334/dsj-2019-028>
- Barba, Lorena A. (2018). Terminologies for reproducible research. arXiv preprint. <https://doi.org/10.48550/arXiv.1802.03311>
- Barnes, Karen I.; Canario, Julio Arturo; Vernekar, Sunil S.; Goudar, Shivaprasad S.; Espinal, Roberto; Merson, Laura & Cheah, Phaik Yong. (2019). Equitable data sharing: challenges and suggestions for ways forward. *Wellcome Open Research*, 4, 172. <https://doi.org/10.12688/wellcomeopenres.15425.1>
- Beagrie, Nell. (2006). Digital curation for science, digital libraries, and individuals. *IJDC*. <https://doi.org/10.2218/ijdc.v1i1.2>
- Becerril-García, Arianne y Córdoba-González, S. (2021). La infraestructura que sostiene el acceso abierto no comercial en América Latina, el Caribe, España y Portugal. Resultados de la encuesta regional a revistas científicas. 2021. En *Conocimiento Abierto en América Latina. Trayectoria y Desafíos*. CLACSO. <http://hdl.handle.net/20.500.11799/112626>
- Becerril García, Arianne. (2022). Favorecer los canales de publicación y distribución inclusivos de manera que nunca se excluya a los autores por motivos económicos: el Acceso Abierto "verde" y "diamante" en América Latina en el marco de BOAI20. *Tramas Y Redes*, (3), 327-337. <https://doi.org/10.54871/cl4c315a>
- Beigel, Fernanda. (2014). Publishing from the periphery: Structural heterogeneity and segmented circuits. The evaluation of scientific publications for tenure in Argentina's CONICET. *Current sociology*, 62(5), 743-765. <https://doi.org/10.1177/0011392114533977>

- Beigel, Fernanda; Gallardo, Oswaldo & Bekerman, Fabiana. (2018). Institutional expansion and scientific development in the periphery: The structural heterogeneity of Argentina's academic field. *Minerva*, 56, 305-331. <https://doi.org/10.1007/s11024-017-9340-2>
- Besançon, Lonni; Peiffer-Smadja, Nathan; Segalas, Corentin; Jiang, Haiting; Masuzzo, Paola; Smout, Cooper; Billy, Eric; Deforet, Maxime & Leyrat, Clémence. (2021). Open science saves lives: lessons from the COVID-19 pandemic. *BMC Medical Research Methodology*, 21(1), 1-18. <https://doi.org/10.1186/s12874-021-01304-y>
- Berghmans, Stephane; 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*. Doctoral dissertation. <https://hdl.handle.net/1887/72152>
- Bezjak, Sonja; Conzett, Philip; Fernandes, Pedro L.; Görögh, Edit; Helbig, Kerstin; Kramer, Bianca; Labastida, Ignasi; Niemeyer, Kyle; Psomopoulos, Fotis; Ross-Hellauer, Tony; Schneider, René; Tennant, Jon; Verbakel, Ellen; Clyburne-Sherin, April; Brinken, Helene & Heller, Lambert. (2017). *Manual de capacitación sobre Ciencia abierta*. Foster. <https://doi.org/10.5281/zenodo.2583101>
- Berry, William, D., & Feldman, Stanley. (1985). *Multiple regression in practice* (No. 50). Sage. <https://archive.org/details/multipleregressi0000berr>
- Bezuidenhout, Louise. M.; Leonelli, Sabina; Kelly, Ann H., & Rappert, Brian. (2017). Beyond the digital divide: Towards a situated approach to open data. *Science and Public Policy*, 44(4), 464-475. <https://doi.org/10.1093/scipol/scw036>
- Bloom, Theodora; Ganley, Emma & Winker, Margaret. (2014). Data access for the open access literature: PLoS's data policy. *PLoS medicine*, 11(2), e1001607. <https://doi.org/10.1371/journal.pmed.1001607>
- BMJ. (1996). *The Nuremberg Code* (1947). <https://doi.org/10.1136/bmj.313.7070.1448>
- Bongiovani, Paola Carolina y Nakano, Silvia (2011). Acceso abierto en Argentina: la experiencia de articulación y coordinación institucional de los repositorios digitales en ciencia y tecnología. *E-colabora*, 1(2), 163-179. https://www.academia.edu/download/70461228/Acceso_Abierto_en_Argentina_a_La_experiencia20210928-10717-1j6rhhr.pdf
- Bongiovani, Paola Carolina; Diana Gómez, Nancy Diana & Miguel, Sandra. (2012). Opiniones y hábitos de publicación en acceso abierto de los investigadores argentinos. Un estudio basado en los datos de la encuesta SOAP. *Revista Española de Documentación Científica*, 35(3), 453-467. <https://doi.org/10.3989/redc.2012.3.903>

- Bongiovani, Paola Carolina; Guarnieri, Griselda; Babini, Dominique & López, Fernando A. (2014). Acceso abierto en la Universidad Nacional de Rosario: Necesidades y prácticas de los docentes/investigadores. *Información, cultura y sociedad*, (30), 13-33. <https://www.redalyc.org/articulo.oa?id=263031224002>
- Bongiovani, Paola Carolina & Martínez-Urbe, L. (2014). Necesidades de gestión de datos científicos en Argentina. El caso de la Universidad Nacional de Rosario. En *Anais das sessões temáticas e pôsters. Conferência Internacional Acesso Aberto, Preservação Digital, interoperabilidade, Visibilidade e Dados Científicos*. Biredial Istec 2014. <http://hdl.handle.net/10183/108316>
- Bongiovani, Paola Carolina & Gómez, Nancy Diana. (2015). Conocimientos y Opiniones sobre acceso abierto en Argentina, México y Brasil. En: Alperín, et al. *Hecho en Latinoamérica: acceso abierto, revistas académicas e innovaciones regionales*. Clacso. <http://hdl.handle.net/2133/6386>
- Bongiovani, Paola Carolina; Miguel, Sandra & Hernández-Pérez, Tony. (2017). Actitudes y percepciones de los evaluadores de la carrera científica en Argentina sobre la publicación en acceso abierto. *Revista Española de Documentación Científica*, 40(2). <https://doi.org/10.3989/redc.2017.2.1404>
- Bongiovani, Paola Carolina & Miguel, Sandra. (2019a). ¿Cuán abierta es la producción científica de los investigadores argentinos de Ciencias Sociales? *Palabra clave*, 9(1). <https://doi.org/10.24215/18539912e080>
- Bongiovani, Paola Carolina; Pérez Bigot, Alejandro y Voras, Claudia. (2019b). La institucionalización del acceso abierto en la Universidad Nacional de Rosario. En *La cultura de los datos: Actas del II Congreso Internacional de la Asociación Argentina de Humanidades Digitales Universidad Nacional de La Plata*. Facultad de Humanidades y Ciencias de la Educación. <http://hdl.handle.net/2133/13572>
- Bongiovani, Paola Carolina; Díaz Pacífico, F., Ferreyra, S. D. C., & Freán, P. (2020). *Disponibilidad en abierto de la producción científica de la UNR incluida en bases de datos. El caso de la Facultad de Ciencias Bioquímicas y Farmacéuticas*. [Comunicación en congreso]. XIV Jornadas de Ciencia, Tecnología e Innovación. Universidad Nacional de Rosario. <http://hdl.handle.net/2133/19380>
- Bongiovani, Paola Carolina; Esteva, María; Díaz Pacífico, Fernando; Gutierrez, Ana Paula; Balparda, Laura Rita; Mestre, Diego; Salazar, Analía y Freán, Paulina. (2022). Planificación e Implementación del Repositorio de Datos Académicos de la Universidad Nacional de Rosario (RDA-UNR): una metodología colaborativa e interdisciplinaria. *Conferencia Internacional BIREDIAL-ISTEC*. RepHipUNR <http://hdl.handle.net/2133/24576>
- Bongiovani, Paola Carolina; Salazar, Analía y Freán, Paulina. (2023a). *Implementación de PIDs en América Latina. Repositorio De Datos Académicos RDA-UNR dataverse.unr.edu.ar*. <https://doi.org/10.5281/zenodo.7860470>

- Bongiovani, Paola Carolina. (2023b). *Capacitación ¿Cómo publicar datos de investigación en el Repositorio de Datos Académicos RDA-UNR?* RepHipUNR <http://hdl.handle.net/2133/26454>
- Bongiovani, Paola Carolina. (2023c). *Capacitación Módulo 2 Traé tus datos. Cómo publicar Datos de investigación en el Repositorio De Datos Académicos RDA-UNR*. RepHipUNR <https://rehip.unr.edu.ar/handle/2133/26503>
- Bongiovani, Paola Carolina. (2023d). *Datos de encuesta sobre factores individuales e institucionales que influyen en la intención de compartir datos de investigación en la Universidad Nacional de Rosario. RDA UNR*. <https://doi.org/10.57715/UNR/HLH331>
- Bongiovani, Paola Carolina; Salazar, Analía; Freán, Paulina; Díaz Pacífico, Fernando; Gutiérrez, Ana Paula; Quintana, Dolores; Balparda, Laura Rita; López, Diego; Stenta, Hernán y Santucci, Natalia. (2023). Trabajo en equipo y compromiso, claves para la implementación exitosa del Repositorio de Datos Académicos RDA-UNR. *Semana de Ciencia abierta LA Referencia 2023*. <https://hdl.handle.net/2133/29041>
- Borgman, Christine. L. (2012). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, 63(6), 1059-1078. <https://doi.org/10.1002/asi.22634>
- Borgman, Christine. L. (2016). *Big data, little data, no data: Scholarship in the networked world*. MIT press. <https://doi.org/10.7551/mitpress/9963.001.0001>
- Borgman, Christine. L., Scharnhorst, A., & Golshan, M. S. (2019). Digital data archives as knowledge infrastructures: Mediating data sharing and reuse. *Journal of the Association for Information Science and Technology*, 70(8), 888-904. <https://doi.org/10.1002/asi.24172>
- Borgman, Christine. L. (2020). *Knowledge infrastructures in past, present, and future tense*. eScholarship. <https://escholarship.org/uc/item/5v73333z>
- Borgman, Christine. L. & Bourne, P. (2022). Why It Takes a Village to Manage and Share Data. *Harvard Data Science Review*, 4(3). <https://doi.org/10.1162/99608f92.42eec111>
- Borrego, Ángel. (2022). Indicadores de medición del acceso abierto: fuentes y herramientas. *Anuario ThinkEPI*, 16. e16a18. <https://doi.org/10.3145/thinkepi.2022.e16a18>
- Borrego, Ángel. (2023). Article processing charges for open access journal publishing: A review. *Learned Publishing*, 36(3), 359-378. <https://doi.org/10.1002/leap.1558>

- Borycz, Joshua; Olendorf, Robert; Specht, Alison; Grant, Bruce; Crowston, Kevin; Tenopir, Carol & Sandusky, Robert J. (2023). Perceived benefits of open data are improving but scientists still lack resources, skills, and rewards. *Humanities and Social Sciences Communications*, 10(1), 1-12. <https://doi.org/10.1057/s41599-023-01831-7>
- Bosman, Jeroen; Frantsovåg, Jan Erik; Kramer, Bianca ; Langlais, Pierre-Carl & Proudman, Vanessa. (2021). *OA Diamond Journals Study. Part 1: Findings*. Zenodo. <https://doi.org/10.5281/zenodo.4558704>
- Boté-Vericad, Juan José; Carballo-García, Ana; Bautista-Villaescusa, Mónica; Healy, Sharon. (2023). Research data repositories in the RDM cycle: challenges and strengths for curators/data stewards. *AIBstudi* 63(3). <https://doi.org/10.2426/aibstudi-13892>
- Brainard, Jeffrey. (2021). Open access takes flight. *Science*. <https://doi.org/10.1126/science.371.6524.16>
- Braun, Virginia & Clarke, Victoria. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101. https://www.researchgate.net/publication/235356393_Using_thematic_analysis_in_psychology
- Braun, Virginia & Clarke, Victoria. (2012). Thematic analysis. In *APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological*. (pp. 57-71). American Psychological Association. <https://doi.org/10.1037/13620-004>
- Braun, V. & Clarke, Victoria. (2013). Successful qualitative research: A practical guide for beginners. *Successful qualitative research*. Sage. https://www.researchgate.net/publication/256089360_Successful_Qualitative_Research_A_Practical_Guide_for_Beginners
- Braun, Virginia & Clarke, Victoria. (2022). Conceptual and design thinking for thematic analysis. *Qualitative Psychology*, 9(1), 3. <https://doi.org/10.1037/qup0000196>
- Brinken, Helene; Kuchma, Iryna; Kalaitzi, Vasso; Davidson, Joy; Pontika, Nancy; Cancellieri, Matteo; Melero, Remedios; Kastelic, Danjama Borba, Filomena; Lenaki, Katerina; Toelch, Ulf; Zourou, Katerina; Konth, Petr, Schmidt, Birgit & Rodrigues, Eloy. (2019). A Case Report: Building communities with training and resources for Open Science trainers. *LIBER Quarterly*, 29(1). <https://doi.org/10.18352/lq.10303>
- Budapest Open Access Initiative. Recomendaciones en su 20º Aniversario (2022). <https://www.budapestopenaccessinitiative.org/boai20/boai20-spanish-translation/>

- Camkin, Jeff; Neto, Susana; Bhattarai, Basundgara; Ojha, Hermant; Khan, Shazbaz; Sugiura, Ai; Lin, Jiaying; Nurritasari, Fitrie Atviana & Karanja, Joseph Muirui. (2022). Open Science for Accelerating the Sustainable Development Goals: Status and Prospects in Asia and the Pacific. *Frontiers in Political Science*, 4, 878761. <https://doi.org/10.3389/fpos.2022.878761>
- Candela, Leonardo; Castelli, Donatella; Manghi, Pablo & Tani, Alice. (2015). Data journals: A survey. *Journal of the Association for Information Science and Technology*, 66(9), 1747-1762. <https://doi.org/10.1002/asi.23358>
- Castro, Eleni; Crosas, Mercé; Garnett, Alex; Sheridan, Kasey & Altman, Micah. (2017). Evaluating and promoting open data practices in open access journals. *Journal of Scholarly Publishing*, 49(1), 66-88. <https://doi.org/10.3138/jsp.49.1.66>
- Carroll, John B. (1953). An analytical solution for approximating simple structure in factor analysis. *Psychometrika*, 18(1), 23-38. <https://doi.org/10.1007/BF02289025>
- Cerda-Cosme, Roxana, & Méndez, Eva. (2023). Analysis of shared research data in Spanish scientific papers about COVID-19: A first approach. *Journal of the Association for Information Science and Technology*, 74(4), 402-414. <https://doi.org/10.1002/asi.24716>
- Champieux, Robin; Solomonides, Anthony; Conte, Marisa; Rojevsky, Svetlana; Phuong, Jimmy; Dorr, David A.; Zampino, Elizabeth; Wilson, Adam; Carson, Matthew B. & Holmes, Kristi. (2023). Ten simple rules for organizations to support research data sharing. *PLOS Computational Biology*, 19(6), e1011136. <https://doi.org/10.1371/journal.pcbi.1011136>
- Choudhury Sayeed; Huang Calhong & Palmer Carole L. (2020). Updating the DCC curation lifecycle model. *International Journal of Digital Curation* 15(1): 1-12. <https://doi.org/10.2218/ijdc.v15i1.721>
- Chawinga, Winner Dominic & Zinn, Sandy. (2019). Global perspectives of research data sharing: A systematic literature review. *Library & Information Science Research*, 41(2), 109-122. <https://doi.org/10.1016/j.lisr.2019.04.004>
- Clare, Connie; Cruz, María; Papadopoulou, Elli; Savage, James; Teperek, Marta & Wang, Yang. (2019). Engaging researchers with data management: The cookbook. *Open Book Publishers*. <https://www.openbookpublishers.com/books/10.11647/obp.0185>
- Clinio, Anne. (2019). Ciência aberta na América Latina: duas perspectivas em disputa. *Transinformação*, 31, e190028. <https://doi.org/10.1590/238180889201931e190028>
- cOAlition S. (2018). Plan S. <https://www.coalition-s.org/why-plan-s/>

- Cohen, Jacob; Cohen, Patricia; West, Stephen G. & Aiken, Leona S. (2002). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates. <https://doi.org/10.4324/9780203774441>
- Colavizza, Giovanni; Hrynaszkiewicz, Iain; Staden, Isla; Whitaker, Kirstie, & McGillivray, Barbara. (2020). The citation advantage of linking publications to research data. *PloS one*, 15(4), e0230416. <https://doi.org/10.1371/journal.pone.0230416>
- Cole, Nicki Lisa; Reichmann, Stefan & Ross-Hellauer, Tony. (2023). Toward equitable open research: stakeholder co-created recommendations for research institutions, funders and researchers. *Royal Society Open Science*, 10(2), 221460. <https://doi.org/10.1098/rsos.221460>
- Cooper, Alexandra; Steeleworthy, Michael; Paquette-Bigras, Ève; Clary, Erin; MacPherson, Erin; Gillis, Louise & Brodeur, Jason. (2021). Creating Guidance for Canadian Dataverse Curators: Portage Network's Dataverse Curation Guide. *Journal of eScience Librarianship*, 10(3). <https://doi.org/10.7191/jeslib.2021.1201>
- Conzett, Philip. (2020). DataverseNO: A national, generic repository and its contribution to the increased FAIRness of data from the long tail of research. *Ravnetrykk* <https://doi.org/10.7557/15.5514>
- Conzett, Philip. (2023). DataverseNO README File Template - General (2.3). Zenodo. <https://doi.org/10.5281/zenodo.7828073>
- Comité Asesor en Ciencia abierta y Ciudadana. MINCYT, Argentina, Ministerio de Ciencia, Tecnología e Innovación de Argentina (2022). *Diagnóstico y lineamientos para una política de ciencia abierta en Argentina*. <https://bicyt.conicet.gov.ar/fichas/produccion/12291633>
- Cragin, M. H., Palmer, C. L., Carlson, J. R., & Witt, M. (2010). Data sharing, small science and institutional repositories. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 368(1926), 4023-4038. <https://doi.org/10.1098/rsta.2010.0165>
- Crosas, Mercè. (2011). The Dataverse Network®: An Open-Source Application for Sharing, Discovering and Preserving Data. *D-Lib Magazine*, 17(1/2). <https://doi.org/10.1045/january2011-crosas>
- Crosas, Mercè. (2020). Fair Principles and Beyond: Implementation in Dataverse. *Septentrio Conference Series*, 2, Article 2. <https://doi.org/10.7557/5.5334>
- CRUE Universidades Españolas (2023). *La Coalición para el Avance de la Evaluación de la Investigación (CoARA) aprueba la propuesta de reforma de Crue, ANECA y CSIC para España*. <https://www.crue.org/2023/07/coara-aprueba-la-propuesta-de-reforma-de-crue-aneca-y-csic-para-espana/>

- Cuartas, Gabriel Vélez; Beigel, Fernanda; Quintero, Diego Restrepo; Uribe Tirado, Alejandro; Gutiérrez Gutiérrez, Gerardo; Pallares, César; Soto Herrera, Diego Alejandro & Gallardo, Osvaldo. (2022). *La producción argentina en acceso abierto y pagos de APC*. CONICET. <https://www.conicet.gov.ar/wp-content/uploads/INFORME-CONICET-Argentina-Publicaciones-y-Pagos-de-APC-REVISADO-2023.pdf>
- Curty, Renato Gonçalves; Crowston, Kevin; Specht, Alison; Grant, B. W. & Dalton, Elizabeth D. (2017). Attitudes and norms affecting scientists' data reuse. *PloS one*, 12(12), e0189288. <https://doi.org/10.1371/journal.pone.0189288>
- Cushing, Amber L. & Shankar, Kalpana. (2019). Digital curation on a small Island: a study of professional education and training needs in Ireland. *Archives and Records*, 40(2), 146-163. <https://doi.org/10.1080/23257962.2018.1425135>
- da Costa, Michelli Pereira & Lima Leite, Fernando César (2019). Factors influencing research data communication on Zika virus: a grounded theory. *Journal of Documentation*, 75(5), 910-926. <https://doi.org/10.1108/JD-05-2018-0071>
- Dallmeier-Tiessen, Suenje; Darby, Robert; Goerner, Bettina; Hyppoelae, Jenni; Igo-Kemenes, Peter; Kahn, Deborah; Lambert, Simon; Legenfelder, Anja; Leonard, Chris; Mele, Salvarore; Nowicka, Malgorzata; Polydarotou, Panayiota; Ross, David; Ruiz-Pérez, Sergio; Schimmer, Ralf; Swaisland, Mark & Van Der Stelt, Wim. (2011). *Highlights from the SOAP project survey. What scientists think about open access publishing*. arXiv preprint. <https://doi.org/10.48550/arXiv.1101.5260>
- Data Curation Network. (2022). *Data Curation Network Strategic Framework 2022-2025 v1.0*. University of Minnesota Twin Cities. <https://hdl.handle.net/11299/226182>
- David, Romain; Mabile, Laurence; Specht, Alison; Stryeck, Sarah; Thomsen, Mogens; Yahia, Mohamed; Jonquet, Clement; Dollé, Laurent; Jacob, Daniel; Bailo, Daniele; Bravo, Elena; Gachet, Sophie; Gunderman, Hannah; Hollebecq, Jean Eudes & Ioannidis, Vassilios. (2020). FAIRness literacy: the Achilles' heel of applying FAIR principles. *CODATA Data Science Journal*, 19(32), 1-11. <https://doi.org/10.5334/dsj-2020-032>
- Malo de Molina, Teresa (2019). E-cienciaDatos, el repositorio de datos de investigación del consorcio madroño. *RUIDERAE: Revista de Unidades de Información*. <https://revista.uclm.es/index.php/ruiderae/article/view/2170>
- de Oliveira Carvalho, Teila. (2016). A influência das revistas científicas de acesso aberto para o depósito e publicação dos dados de pesquisa. *Cadernos de Biblioteconomia, Arquivística e Documentação*, 2. <http://hdl.handle.net/20.500.11959/brapci/66626>

- Debat, Humberto & Babini, Dominique. (2019). Plan S in Latin America: A precautionary note. *PeerJ Preprints*, 7, e27834v2. <https://doi.org/10.7287/peerj.preprints.27834v2>
- Declaración de Budapest (2001). <http://www.soros.org/openaccess/read.shtml>
- Declaración de Bethesda (2003). <http://www.earlham.edu/~peters/fos/bethesda.htm>
- Declaración de Berlín (2003). http://openaccess.mpg.de/67627/Berlin_sp.pdf
- Declaración de Panamá sobre Ciencia abierta (2018). <https://forocilac.org/declaracion-de-panama-sobre-ciencia-abierta/>
- Delgado-López-Cózar, Emilio; Ràfols, Ismael y Abadal, Ernest. (2021). Letter: A call for a radical change in research evaluation in Spain. *El Profesional de La información*, 30(3). <https://doi.org/10.3145/epi.2021.may.09>
- Demchenko, Yuri; Stoy, Lennart; Engelhardt, Claudia & Gaillard, Vinciane. (2021). *D7. 3 FAIR Competence Framework for Higher Education (Data Stewardship Professional Competence Framework)*. <https://doi.org/10.5281/zenodo.5361917>
- Devaraju, Anusuriya; Huber, Robert; Mokrane, Mustapha; Herterich, Patricia; Cepinskas, Linas; de Vries, Jerry; L'Hours, Herve; Davidson, Joy & Angus White. (2020). *FAIRsFAIR Data Object Assessment Metrics* (0.4). <https://doi.org/10.5281/zenodo.4081213>
- DiMaggio, Paul J., & Powell, Walter W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American sociological review*, 147-160. <https://doi.org/10.2307/2095101>
- Directrices para proveedores de contenido del Sistema Nacional de Repositorios Digitales. (2015). Buenos Aires: SNRD. Ministerio de Ciencia, Tecnología e Innovación Productiva. https://repositoriosdigitales.mincyt.gob.ar/files/Directrices_SNRD_2015.pdf
- Doldirina, Catherine; Eisenstadt, Anita; Onsrud, Harland & Uhler, Paul. (2018). Legal approaches for open access to research data. *LawArXiv Papers, Santa Barbara*, 1-52. <https://doi.org/10.31228/osf.io/n7gfa>
- Dorta-González, Pablo; González-Betancor, Sara M. & Dorta-González, María Isabel (2021). To what extent is researchers' data-sharing motivated by formal mechanisms of recognition and credit? *Scientometrics*, 126(3), 2209-2225. <https://doi.org/10.1007/s11192-021-03869-3>
- Esteva, María. (2017). *Material didáctico del módulo 1. Seminario "Curación de colecciones de datos de investigación". Especialización en Gestión de Información Científica y Tecnológica*. Universidad Nacional de La Plata. Facultad de Humanidades y Ciencias de la Educación. [Video]. <https://youtu.be/oNnD89LEUfw?si=gjKwZQo-8lGCuala>

- Esteva, María; Jansen, Craig; Arduino, Pedro; Sharifi-Mood, Mahyar; Dawson, Clint N. & Balandrano-Coronel, Josue. (2019). Curation and publication of simulation data in DesignSafe, a natural hazard engineering open platform and repository. *Publications*, 7(3), 51. <https://doi.org/10.3390/publications7030051>
- Esteva, María. (2020). Por qué y para qué. Desafíos y beneficios de la gestión de datos abiertos de investigación. Capacitación *Gestión de datos de investigación en la UNR* <https://rephip.unr.edu.ar/server/api/core/bitstreams/4c14ee72-9453-4634-8d67-2229b444706c/content>
- Edwards, Paul N. (2010). *A vast machine: Computer models, climate data, and the politics of global warming*. MIT Press.
- Edwards, Paul N.; Jackson, Steven J.; Chalmers, Melissa K.; Bowker, Geoffrey C.; Borgman, Christine L.; Ribes, David; Burton, Matt & Calvert, Scout. (2013). *Knowledge Infrastructures: Intellectual Frameworks and Research Challenges*. Ann Arbor: Deep Blue. <http://hdl.handle.net/2027.42/97552>
- Engelhardt, Claudia; Biernacka, Katarzyna; Coffey, Aoife; Cornet, Ronald; Danciu, Alina; Demchenko, Yuri; Downes, Stephen; Erdmann, Christopher; Garbuglia, Federica; Germer, Kerstin; Helbig, Kerstin; Hellström, Margareta; Hettne, Kristina; Hibbert, Dawn; Jetten, Mijke; Karimova, Yulia; Kryger Hansen, Karsten; Kuusniemi, Mari Elisa; Letizia, Viviana; McCutcheon, Valerie; McGillivray, Barbara; Ostrop, Jenny; Petersen, Britta; Petrus, Ana; Reichmann, Stefan; Rettberg, Najla; Reverté, Carmen; Rochlin, Nick; Saenen, Bregt; Schmidt, Birgit; Scholten, Jolien; Shanahan, Hugh; Straube, Armin; Van den Eynden, Veerle; Vandendorpe, Justine; Venkataram, Shanmugasundaram; Wiljes, Cord; Wuttke, Ulrike; Yeomans, Joanne & Zhou, Biru. (2022). *How to be FAIR with your data: A teaching and training handbook for higher education institutions*. Universitätsverlag Göttingen. <https://doi.org/10.5281/zenodo.6674301>
- Enke, Neelo; Thessen, Anne; Bach, Kerstin; Bendix, Jörg; Seeger, Bernhard & Gemeinholzer, Birgit. (2012). The user's view on biodiversity data sharing—Investigating facts of acceptance and requirements to realize a sustainable use of research data—. *Ecological informatics*, 11, 25-33. <https://doi.org/10.1016/j.ecoinf.2012.03.004>
- Enwald, Heidi; Grigas, Vincas; Rudžionienė, Jurgita & Kortelainen, Terttu. (2022). Data sharing practices in open access mode: a study of the willingness to share data in different disciplines. *Information Research*, 27(2). <https://doi.org/10.47989/irpaper932>
- European Commission (2017). *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 (2021). Horizonte Europa https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe_en
- Everitt, Brian S. y Wykes, Till. (2001). *Diccionario de Estadística para Psicólogos*. Barcelona: Ariel.
- Faniel, Ixchel M.; Barrera-Gomez, Julianna; Kriesberg, Adam & Yakel, Elizabeth. (2013). A Comparative Study of Data Reuse Among Quantitative Social Scientists and Archeologists. *IConference*, 797–800. <https://doi.org/10.9776/13391>
- Federer, Lisa. (2018). Defining data librarianship: a survey of competencies, skills, and training. *Journal of the Medical Library Association: JMLA*, 106(3), 294. <https://doi.org/10.5195/jmla.2018.306>
- Federer, Lisa M., Belter, Christopher W.; Joubert, Douglas J.; Livinski, Alicia; Yan-Ling, Lu; Snyders, Lyssa N & Thompson, Holly. (2018). Data sharing in PLOS ONE: an analysis of data availability statements. *PloS one*, 13(5), e0194768. <https://doi.org/10.1371/journal.pone.0194768>
- Federer, Lisa M. (2022). Long-term availability of data associated with articles in PLOS ONE. *Plos one*, 17(8), e0272845. <https://doi.org/10.1371/journal.pone.0272845>
- Fecher, Benedict; Friesike, Sascha, & Hebing, Marcel. (2015). What drives academic data sharing? *PloS one*, 10(2), e0118053. <https://doi.org/10.1371/journal.pone.0118053>
- Fecher, Benedict; Friesike, Sascha; Hebing, Marcel, & Linek, Stephanie. (2017). A reputation economy: how individual reward considerations trump systemic arguments for open access to data. *Palgrave Communications*, 3(1), 1-10. <https://doi.org/10.1057/palcomms.2017.51>
- Force11. (2020). *Guiding Principles for Findable, Accessible, Interoperable and Re-Usable Data Publishing Version B1.0*. <https://force11.org/info/guiding-principles-for-findable-accessible-interoperable-and-re-usable-data-publishing-version-b1-0/>
- Fushimi, Marcela. (2018). Desarrollo de repositorios digitales institucionales en las universidades nacionales en Argentina, período 2004-2015. [Tesis de Maestría]. Universidad Nacional de Quilmes. <http://ridaa.unq.edu.ar/handle/20.500.11807/887>
- Fushimi, Marcela; Pené, Mónica G.; Unzurrunzaga, Carolina y Sanllorenti, Ana. (2021). *Situación de los repositorios digitales en las universidades nacionales argentinas de gestión pública (2019)*. Universidad Nacional de La Plata. IIHCS. <http://www.memoria.fahce.unlp.edu.ar/libros/pm.4722/pm.4722.pdf>

- García-García, Alicia; López-Borrull, Alexandre & Peset, Fernanda. (2015). Data journals: eclosión de nuevas revistas especializadas en datos. *El profesional de la información*, 24(6), 1699-2407. <https://doi.org/10.3145/epi.2015.nov.17>
- Gefen, David; Straub, Detmar & Boudreau, Marie Claude. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the association for information systems*, 4(1), 7. <https://doi.org/10.17705/1CAIS.00407>
- Gerwig, Katherine J. (2017). Current Outreach and Marketing Practices for Research Data Repositories. In L. Johnston, *Curating Research Data, Volume 1: Practical Strategies for Your Digital Repository*. Chicago: Association of College and Research Libraries. https://www.ala.org/acrl/sites/ala.org/acrl/files/content/publications/booksanddigitalresources/digital/9780838988596_crd_v1_OA.pdf#page=222
- Gómez, Nancy y Bongiovani, Paola Carolina. (2012). Open access and A2K: Collaborative experiences in Latin America. En Lau, Jesús; Tammara, Anna Maria & Bothma, Theo J.D. *Libraries driving access to knowledge* (Vol. 151). <https://e-archivo.uc3m.es/rest/api/core/bitstreams/10c4b477-b6b5-4268-a8d9-705f44a8f547/content>
- González-Teruel, Aurora; López-Borrull, Alexandre; Santos-Hermosa, Gema; Abad-García, Francisca; Ollé, Candela & Serrano-Vicente, Rocío. (2022). Drivers and barriers in the transition to open science: the perspective of stakeholders in the Spanish scientific community. *El Profesional de la información*, 31(3). <https://hdl.handle.net/10609/145540>
- Goode, Gregory; Hahnel, Mark; Zhou, Yuanchun; Jiang, Lulu; Chandramouliswaran, Iswar; Hafez, Amy; Paine, Taunton; Gregurickm Susan; Simango, Samuel; Palma Peña, Juan Miguel; Murray, Holly; Cannon, Matt; Grant, Rebecca; McKellar, Kate & Day, Laura. (2022). *The State of Open Data 2022*. <https://doi.org/10.6084/m9.figshare.21276984.v5>
- Gownaris, Natasha J.; Vermeir, Koen; Bittner, Martin Inmanuel; Gunawardena, L., Kaur-Ghumaan, S., Lepenies, R., ... & Zakari, I. S. (2022). Barriers to Full Participation in the Open Science Life Cycle among Early Career Researchers. *Data Science Journal*, 21(1). <https://doi.org/10.5334/dsj-2022-002>
- Grabus, Sam & Greenberg, Jane. (2019). The landscape of rights and licensing initiatives for data sharing. *Data Science Journal*, 18, 29-29. <https://doi.org/10.5334/dsj-2019-029>
- Gregory, Kathleen; Groth, Paul; Scharnhorst, Andrea & Wyatt, Sally. (2020). Lost or Found? Discovering Data Needed for Research. *Harvard Data Science Review*, 2(2). <https://doi.org/10.1162/99608f92.e38165eb>
- Hair, Joseph F.; Black, William C.; Babin, Barry J. & Anderson, Rolph E. (2019). *Multivariate data analysis, Eighth Edition*. Cengage Learning EMEA

- Hardwicke, Tom E.; Mathur, Maya B.; MacDonald, Kyle; Nilsonne, Gustav; Banks, George C.; Kidwell, Mallory C.; Mohr, Alice Hofelich; Clayton, Elizabeth; Yoon, Erika J.; Tessler, Michael Henry; Lenne, Richie L.; Altman, Sara; Long, Bria & Frank, Michael C. (2018). Data availability, reusability, and analytic reproducibility: evaluating the impact of a mandatory open data policy at the journal *Cognition. R Soc Open Sci.*, 15;5(8):180448. <https://doi.org/10.1098/rsos.180448>
- 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*, 41, 23-32. <https://doi.org/10.1016/j.ijinfomgt.2018.03.001>
- Harvard University. The Institute of Quantitative Social Science. (2022). *Guía de Uso Proyecto Dataverse*. <http://hdl.handle.net/2133/23695>
- Harvey, Ross. (2010). *Digital curation: a how-to-do-it manual*. New York: Neal. <https://archive.org/details/digitalcurationh0000harv>
- Haustein, Stefanie; Schares, Eric; Alperín, Juan Pablo; Hare, Madelaine; Butler, Leigh-Ann & Schönfelder, Nina. (2024). Estimating global article processing charges paid to six publishers for open access between 2019 and 2023. *arXiv preprint arXiv:2407.16551*. <https://doi.org/10.48550/arXiv.2407.16551>
- Hernández-Pérez, Tony & García-Moreno, María Antonia. (2013). Datos abiertos y repositorios de datos: nuevo reto para los bibliotecarios. *Profesional de la información*, 22(3), 259-263. <https://doi.org/10.3145/epi.2013.may.10>
- Hernández Sampieri, Roberto; Fernández Collado, Carlos & Baptista Lucio, Pilar. (2010). *Metodología de la investigación*. Mc Graw Hill educación. https://www.uv.mx/personal/cbustamante/files/2011/06/metodologia-de-la-investigaci%C3%83%C2%B3n_sampieri.pdf
- Heesen, Remco. (2017). Communism and the Incentive to Share in Science. *Philosophy of Science*, 84(4), 698-716. <https://doi.org/10.1086/693875>
- Higgins, Sarah. (2008). Draft DCC curation lifecycle model. *International Journal of Digital Curation*, 2(2). <https://doi.org/10.2218/ijdc.v2i2.30>
- Hicks, Diana; Wouters, Paul; Waltman, Ludo; De Rijcke, Sarah & Rafols, Ismael. (2015). Bibliometrics: the Leiden Manifesto for research metrics. *Nature*, 520(7548), 429-431. <https://doi.org/10.1038/520429a>
- Hillyer, Rebecca; Posada, Alejandro; Albornoz, Denisse; Chan, Leslie & Okune, Angela. (2017). Framing a situated and inclusive open science: emerging lessons from the open and collaborative science in development network. In *Expanding perspectives on open science: Communities, cultures and diversity in concepts and practices* (pp. 18-33). IOS Press. <https://escholarship.org/uc/item/2dj1h8nh>

- Hodonu-Wusu, James Oluwaseyi; Noorhidawati, A. & Abrizah, A. (2020). Malaysian researchers on open data: the first national survey on awareness, practices and attitudes. *Malaysian Journal of Library and Information Science*, 25(2), 1-20. <https://doi.org/10.22452/mjlis.vol25no2.1>
- Hofmann, Bjern. (2022). Open science knowledge production: Addressing epistemological challenges and ethical implications. *Publications*, 10(3), 24. <https://doi.org/10.3390/publications10030024>
- Hong Kong principles for assessing researchers. World Conference on Research Integrity. (2019). <https://www.wcrif.org/guidance/hong-kong-principles>
- Hosseini, Mohammad; Senabre Hidalgo, Eric; Horbach, Serge P. J. M.; Güttinger, Stepahn & Penders, Bart. (2022). Messing with Merton: The intersection between open science practices and Mertonian values. *Accountability in Research*, 1-28. <https://doi.org/10.1080/08989621.2022.2141625>
- Houtkoop, Bobbie Lee; Chambers, Chris; Macleod, Malcon; Bishop, Dorothy V. M.; Nichols, Thomas E. & Wagenmakers, Erik Jan. (2018). Data sharing in psychology: A survey on barriers and preconditions. *Advances in methods and practices in psychological science*, 1(1), 70-85. <https://doi.org/10.1177/2515245917751886>
- Howell, George F. & Buck, Jeffrey M. (2012). The adult student and course satisfaction: what matters most? *Innovative higher education*, 37, 215-226 <https://doi.org/10.1007/s10755-011-9201-0>
- Hrynaskiewicz, Iain; Simons, Natasha; Hussain, Azhar; Grant, Rebecca, & Goudie, Simon. (2020). Developing a research data policy framework for all journals and publishers. *Data Science Journal*, 19(1). <https://doi.org/10.5334/dsj-2020-005>
- Huijpen, Kim & Aspaas, Per Pippin. (2021). Recognition & Rewards in the Netherlands. *Open Science Talk*, (38). <https://doi.org/10.7557/19.6326>
- Ibarzabal, Juan Ignacio; Lagrutta, Mariana; Pendino, Juan Carlos; Parodi, Roberto Leandro. (2022). *Dataset Relación entre el Aumento de la Amplitud de Distribución del Diámetro Eritrocitario (Red Cell Distribution Width o RDW) y la morbilidad postquirúrgica*. [Dataset]. <https://doi.org/10.57715/UNR/OJ4BUD>
- Indart, Camila; Ríos Hilario, Ana Belén & Rodríguez Bravo, Blanca. (2021). Relevamiento de repositorios y portales de datos abiertos de investigación en la Argentina. *Información, cultura y sociedad*, (45), 59-74. <https://doi.org/10.34096/ics.i45.10421>
- Jetten, M., Grootveld, M., Mordant, A., Jansen, M., Bloemers, M., Miedema, M., & van Gelder, C. W. (2021). Professionalising data stewardship in the Netherlands. *Competences, training and education. Dutch roadmap towards national implementation of FAIR data stewardship. National Programme Open Science (NPOS)*.

- Johnson, Rob, Parsons, Tom & Chiarelli, Andrea. (2016). *Jisc Data Asset Framework Toolkit 2016*. <http://doi.org/10.5281/zenodo.177876>.
- Johnston, Lisa R. (2017). *Introduction to Data Curation. Curating Research Data. Volume One: Practical Strategies for Your Digital Repository*. ACRL. https://www.ala.org/sites/default/files/acrl/content/publications/booksanddigitalresources/digital/9780838988596_crd_v1_OA.pdf
- Johnston, Lisa R.; Carlson, Jake; Hudson-Vitale, Cynthia; Imker, Heidi; Kozlowski, Wendy; Olendorf, Robert; Stewart, Claire; Blake, Mara; Herndon, Joef; McGeary, Timothy M. & Hull, Elizabeth. (2018). Data curation network: A cross-institutional staffing model for curating research data. *IJDC* <http://dx.doi.org/10.2218/ijdc.v13i1.616>
- Joo, Soohyung & Peters, Christie. (2020). User needs assessment for research data services in a research university. *Journal of Librarianship and Information Science*, 52(3), 633-646. <https://doi.org/10.1177/0961000619856073>
- Joo, Yeon Kyoung & Kim, Youngseek. (2017). Engineering researchers' data reuse behaviours: a structural equation modelling approach. *The Electronic Library*. <https://doi.org/10.1108/EL-08-2016-0163>
- Joo, Soohyung, Kim, Sujin & Kim, Youngseek. (2017). An exploratory study of health scientists' data reuse behaviors: Examining attitudinal, social, and resource factors. *Aslib Journal of Information Management*, 69(4), 389-407. <https://doi.org/10.1108/AJIM-12-2016-0201>
- Khan, Nushrat; Thelwall, Mike & Kousha, Kayvan. (2023). Data sharing and reuse practices: disciplinary differences and improvements needed. *Online Information Review*. <https://doi.org/10.1108/OIR-08-2021-0423>
- Kim, Youngseek. (2013). *Institutional and Individual Influences on Scientists' Data Sharing Behaviors*. [Dissertation]. Syracuse University. https://surface.syr.edu/it_etd/85
- Kim, Youngseek, & Zhang, Ping. (2015). Understanding data sharing behaviors of STEM researchers: The roles of attitudes, norms, and data repositories. *Library & Information Science Research*, 37(3), 189-200. <https://doi.org/10.1016/j.lisr.2015.04.006>
- Kim, Youngseek, & Kim, Sujin. (2015). Institutional, Motivational, and Resource Factors Influencing Health Scientists' Data-Sharing Behaviours. *Journal of Scholarly Publishing*, 46(4), 366-389. <https://doi.org/10.3138/jsp.46.4.05>
- Kim, Youngseek & Burns, C. Sean. (2016). Norms of data sharing in biological sciences: The roles of metadata, data repository, and journal and funding requirements. *Journal of Information Science*, 42(2), 230-245. <https://doi.org/10.1177/0165551515592098>

- 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*, 67(4), 776-799. <https://doi.org/10.1002/asi.23424>
- Kim, Youngseek. (2017). Scientists' data reuse behaviors: A multilevel analysis. *Journal of the Association for Information Science and Technology*, 68(12), 2709-2719. <http://doi.org/10.3886/E100404V1>
- Kim, Youngseek. (2017). Fostering scientists' data sharing behaviors via data repositories, journal supplements, and personal communication methods. *Information Processing & Management*, 53(4), 871-885. <https://doi.org/10.1016/j.ipm.2017.03.003>
- Kim, Jeonghyun; Schuler, Eric R. & Pechenina, Anna. (2018). Predictors of data sharing and reuse behavior in academic communities. In *Knowledge Discovery and Data Design Innovation: Proceedings of the International Conference on Knowledge Management (ICKM 2017)* (pp. 1-25). https://doi.org/10.1142/9789813234482_0001
- Kim, Youngseek & Nah, Seunghan. (2018). Internet researchers' data sharing behaviors: An integration of data reuse experience, attitudinal beliefs, social norms, and resource factors. *Online information review*, 42(1), 124-142. <https://doi.org/10.1108/OIR-10-2016-0313>
- 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*, 44(6), 824-832. <https://doi.org/10.1016/j.acalib.2018.09.013>
- Kim, Youngseek. (2021). A study of the roles of metadata standard and data repository in science, technology, engineering and mathematics researchers' data reuse. *Online Information Review*, 45(7), 1306-1321. <https://doi.org/10.1108/OIR-09-2020-0431>
- Kim, Youngseek & Oh, Jong Sir. (2021). Researchers' article sharing through institutional repositories and ResearchGate: A comparison study. *Journal of Librarianship and Information Science*, 53(3), 475-487. <https://doi.org/10.1177/0961000620962840>
- Kim, Youngseek. (2022). Data sharing by biologists: A comparative study of genome sequence data and lab experiment data. *Library & Information Science Research*, 44(1), 101139. <https://doi.org/10.1016/j.lisr.2022.101139>
- Kindling, Maxi; Pampel, Heiz; van de Sandt, Stephannie; Rücknagel, Jessika; Vierkant, Paul; Kloska, Gabriele; Witt, Michael; Schirmbacher, Peter; Bertelmann, Roland & Scholze, Frank. (2017). The Landscape of Research Data Repositories in 2015: A re3data Analysis. *D-Lib Magazine*, 23(3/4). <https://doi.org/10.1045/march2017-kindling>

- Klebel, Thomas & Ross-Hellauer, Tony. (2023). The APC-barrier and its effect on stratification in open access publishing. *Quantitative Science Studies*, 4(1), 22-43. https://doi.org/10.1162/qss_a_00245
- Knorr-Cetina, Karin. (1991). Merton's Sociology of Science: The First and the Last Sociology of Science? *Contemporary Sociology*, 20(4), 522-526. <https://doi.org/10.2307/2071782>
- Knorr-Cetina, Karin. (2005). La fabricación del conocimiento. *Un ensayo sobre el carácter constructivista y contextual de la ciencia*. Buenos Aires: Universidad Nacional de Quilmes Editorial. <https://ridaa.unq.edu.ar/handle/20.500.11807/595>
- Kratz, John E. & Strasser, Carly. (2015). Making data count. *Scientific data*, 2(1), 1-5. <https://doi.org/10.1038/sdata.2015.39>
- Larivière, Vincent; Haustein, Stefanie & Mongeon, Philippe. (2015). The oligopoly of academic publishers in the digital era. *PloS one*, 10(6), e0127502. <https://doi.org/10.1371/journal.pone.0127502>
- Latour, Bruno. (1987). *Science in action: How to follow scientists and engineers through society*. Harvard university press.
- Leiden Manifesto for Research Metrics. (2015). <http://www.leidenmanifesto.org>
- Leonelli, Sabina. (2015). What counts as scientific data? A relational framework. *Philosophy of Science*, 82(5), 810-821. <https://doi.org/10.1086/684083>
- Leonelli, Sabina & Lewandowsky, S. (2022). *The Reproducibility of Research In Flanders: Fact Finding And Recommendations. KVAB Thinkers' Report 2022*. https://kvab.be/sites/default/rest/blobs/3679/tw_reproducability.pdf
- Leonelli, Sabina. (2023). *Philosophy of Open Science* (Elements in the Philosophy of Science). Cambridge University Press. <https://doi.org/10.1017/9781009416368>
- Lattuada, Mario José y Giba, Gabriela Silvia. (2014). Una introducción a la Ley de Repositorios Abiertos para Ciencia y Tecnología. *Debate Universitario CAEE-UAI*, 2(4), 79-99. <http://www2.biblio.unlp.edu.ar/jubiuna/capacitacion/taller-sobre-politicas-de-acceso-al-conocimiento-cientifico-en-las-uunn/4588-24144-1-PB-2.pdf>
- Lee, Dong Joon & Stvilia, Besiki. (2017). Practices of research data curation in institutional repositories: A qualitative view from repository staff. *PloS one*, 12(3), e0173987. <https://doi.org/10.1371/journal.pone.0173987>
- Ley 26.899 (2013). Creación de Repositorios Digitales Institucionales de Acceso Abierto, Propios o Compartidos. https://repositoriosdigitales.mincyt.gob.ar/files/Boletin_Oficial_Ley_26899.pdf. Reglamento de la Ley 26.899 (2016). https://repositoriosdigitales.mincyt.gob.ar/files/Boletin_Oficial_Resolucion_753.pdf

- Lin, Dawei; Crabtree, Jonathan; Dillo, Ingrid; Downs, Robert R.; Edmunds, Rorie; Giaretta, David; De Giusti, Marisa; L'Hours, Hervé; Hugo, Wim; Jenkyns, Reyna; Khodiyar, Varsha; Martone, Maryann E.; Mokrane, Mustapha; Navale, Vivek; Petters, Jonathan; Sierman, Barbara; Sokolova, Dina V.; Stockhouse, Martina & Westbrook, John. (2020). The TRUST Principles for digital repositories. *Scientific Data*, 7(1), 144. <https://doi.org/10.1038/s41597-020-0486-7>
- Lord, Philip & Macdonald, Alison. (2003). e-Science curation report. *Data curation for e-Science in the UK: an audit to establish requirements for future curation and provision*. <https://digitalpreservation.gov/news/2004/e-ScienceReportFinal.pdf>
- Lynch, Clifford A. (2003). Institutional repositories: essential infrastructure for scholarship in the digital age. *portal: Libraries and the Academy*, 3(2), 327-336. <https://doi.org/10.1353/pla.2003.0039>
- Mason, Claire M.; Box, Paul J. & Burns, Shanae. M. (2020). Research data sharing in the Australian national science agency: Understanding the relative importance of organisational, disciplinary and domain-specific influences. *Plos one*, 15(8), e0238071. <https://doi.org/10.1371/journal.pone.0238071>
- Majid, Shaheen; Foo, Schubert & Zhang, Xue. (2018). Research data management by academics and researchers: Perceptions, knowledge and practices. In *Maturity and Innovation in Digital Libraries: 20th International Conference on Asia-Pacific Digital Libraries, ICADL 2018, Hamilton, New Zealand, November 19-22, 2018, Proceedings 20* (pp. 166-178). Springer International Publishing. https://doi.org/10.1007/978-3-030-04257-8_16
- Marušić, Ana; Bošnjak, Lana & Jerončić, Ana. (2011). A systematic review of research on the meaning, ethics and practices of authorship across scholarly disciplines. *Plos one*, 6(9), e23477. <https://doi.org/10.1371/journal.pone.0023477>
- Mannheimer, Sara; Pienta, Amy; Kirilova, Dessilava; Elman, Colin & Wutich, Amber. (2019). Qualitative data sharing: Data repositories and academic libraries as key partners in addressing challenges. *American Behavioral Scientist*, 63(5), 643-664. <https://doi.org/10.1177/0002764218784991>
- Mannheimer, Sara. (2022). Data curation for qualitative data reuse and big social research: *Connecting communities of practice*. Humboldt Universitaet zu Berlin (Germany). <https://doi.org/10.7191/jeslib.2021.1218>
- Melero, Remedios & Navarro-Molina, Carolina. (2020). Researchers' attitudes and perceptions towards data sharing and data reuse in the field of food science and technology. *Learned Publishing*, 33(2), 163-179. <https://doi.org/10.1002/leap.1287>
- Méndez Rodríguez, Eva. (2021). Open Science por defecto: La nueva normalidad para la investigación. *Arbor: Ciencia, pensamiento y cultura*, 197(799), 2. <https://doi.org/10.3989/arbor.2021.799002>

- Méndez Rodríguez, Eva & Sánchez-Núñez, Pablo. (2023). Navigating the Future and Overcoming Challenges to Unlock Open Science. In *Ethics and Responsible Research and Innovation in Practice: The ETHNA System Project* (pp. 203-223). Cham: Springer Nature. https://doi.org/10.1007/978-3-031-33177-0_13
- Merton, Robert K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago press.
- Miguel, Sandara, & de Moya Anegón, Félix. (2009). La ciencia argentina bajo la lupa de los indicadores cuantitativos: una mirada crítica de la realidad científica argentina. Universidad Nacional de la Plata. <https://memoria.fahce.unlp.edu.ar/library?a=d&c=libros&d=jpm1553>
- Miguel, Sandra; Gómez, Nancy-Diana & Bongiovani, Paola Carolina. (2012). Acceso abierto real y potencial a la producción científica de un país. El caso argentino. *El Profesional de la Información*, 21(2), 146-154. <https://doi.org/10.3145/epi.2012.mar.04>
- Ministerio de Ciencia, Tecnología e Innovación. (2022). *Diagnóstico y herramientas para la Política de Ciencia abierta en Argentina*. <https://back.argentina.gob.ar/ciencia/comite-ciencia-abierta/diagnostico-y-lineamientos?>
- Mirowski, Philip. (2018). The future (s) of open science. *Social studies of science*, 48(2), 171-203. <https://doi.org/10.1177/0306312718772086>
- Mitroff, Ian I. (1974). Norms and counter-norms in a select group of the Apollo moon scientists: A case study of the ambivalence of scientists. *American sociological review*, 579-595. <https://doi.org/10.2307/2094423>
- Mons, Barend. (2018). *Data stewardship for open science: Implementing FAIR principles*. CRC Press. <https://doi.org/10.1201/9781315380711>
- Nash, Jack. (2021) 'I ain't afraid of no myth' – busting the myths on data sharing. *Wellcome Open Research*. <https://blog.wellcomeopenresearch.org/2020/08/20/i-aint-afraid-of-no-myth-busting-the-myths-on-data-sharing/>
- National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). *The Belmont report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. <https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html>
- National Institutes of Health. (2020). *Final NIH Policy for Data Management and Sharing (NOT-OD-21-013)*. <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html>

- Navarro-Molina, Caroñina & 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*, 28(5). <https://doi.org/10.3145/epi.2019.sep.16>
- Nelson, Alondra. (2022). *Ensuring Free, Immediate, and Equitable Access to Federally Funded Research*. Memorandum for the Heads of Executive Departments and Agencies. Washington, DC: Executive Office of the President Office of Science and Technology Policy. <https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-access-Memo.pdf>
- Neylon, Cameron. (2017). Compliance Culture or Culture Change? The role of funders in improving data management and sharing practice amongst researchers. *Research Ideas and Outcomes*, 3, e21705. <https://doi.org/10.3897/rio.3.e21705>
- Nunnally, Jum C. & Bernstein, Ira H. (1994). *Psychometric Theory*. New York. NY: McGraw-Hill.
- OCSDNet Red de Ciencia abierta y Colaborativa para el Desarrollo. (2017). *Manifiesto de Ciencia abierta y Colaborativa: Hacia una ciencia abierta inclusiva por el bienestar social y ambiental*. <https://ocsdnet.org/wp-content/uploads/2015/04/Manifiesto-Infographic-Spanish-1.pdf>
- OECD (2007). *Principles and guidelines for access to research data from public funding*. https://www.oecd.org/en/publications/oecd-principles-and-guidelines-for-access-to-research-data-from-public-funding_9789264034020-en-fr.html
- ODI (2017). *Open data essentials*. <https://open-data-essentials.learndata.info/>
- Oo, Chery Zin.; Chew, Adrian W.; Wong, Adeline L. H.; Gladding, Joanne & Stenstrom, Cecilia. (2022). Delineating the successful features of research data management training: A systematic review. *International Journal for Academic Development*, 27(3), 249-264. <https://doi.org/10.1080/1360144X.2021.1898399>
- Palmer, Carole; Weber, Nicholas M.; Renear, Allen H. & Muñoz, Trevor. (2013). Foundations of data curation: The pedagogy and practice of "purposeful work" with research data. *Archives Journal*, Vol 3. <http://hdl.handle.net/2142/78099>
- Pampel, Heinz; Vierkant, Paul; Scholze, Frank; Bertelmann, Roland; Kindling, Maxi; Klump, Jens; Goebelbecker, Hans Jurgen; Gundlach, Jens; Schirmbacher, Peter & Dierolf, Uwe. (2013). Making research data repositories visible: the re3data.org registry. *PloS one*, 8(11), e78080. <https://doi.org/10.1371/journal.pone.0078080>
- Pampel, Heinz & Dallmeier-Tiessen, Sünje. (2014). Open research data: From vision to practice. In: Bartling, S., Friesike, S. (eds). *Opening science: The evolving guide on how the Internet is changing research, collaboration and scholarly publishing*, 213-224. https://doi.org/10.1007/978-3-319-00026-8_14

- Pasquetto, Irene V.; Randles, Bernadette M. & Borgman, Christine L. (2017). On the Reuse of Scientific Data. *Data Science Journal*, 16. <https://doi.org/10.5334/dsj-2017-008>
- Pasquetto, Irene V. (2018). *From open data to knowledge production: Biomedical data sharing and unpredictable data reuses* [Doctoral dissertation]. UCLA. <https://escholarship.org/uc/item/4nd5j5jp>
- Pasquetto, Irene. (2019). Do scientists reuse open data? [Blog Post]. *Caos by Sage Bionetworks*. <https://medium.com/voices-from-the-open-science-movement/by-irene-pasquetto-bac8bc02afdc>
- Paulus, Frieder M.; Cruz, Nicole & Krach, Sören. (2018). The impact factor fallacy. *Frontiers in psychology*, 9, 1487. <https://doi.org/10.3389/fpsyg.2018.01487>
- Peng, Robert D. (2011). Reproducible research in computational science. *Science*, 334(6060), 1226-1227. <https://doi.org/10.1126/science.1213847>
- Pennock, Maureen. (2007). Digital curation: A life-cycle approach to managing and preserving usable digital information. *Library & Archives Journal*, 1. https://www.ukoln.ac.uk/ukoln/staff/m.pennock/publications/docs/lib-arch_curation.pdf
- Peset, Fernada; Aleixandre-Benavent, Rafael; Blasco-Gil, Yolanda & Ferrer-Sapena, Antonia. (2017). Datos abiertos de investigación. Camino recorrido y cuestiones pendientes. *Anales de Documentación*, 20(1). <https://doi.org/10.6018/analesdoc.20.1.272101>
- Pilat, Dirk & Fukasaku, Yukiko. (2007). OECD principles and guidelines for access to research data from public funding. *Data Science Journal*, 6, OD4-OD11. <https://doi.org/10.2481/dsj.6.OD4>
- Pinfield, Stephen; Cox, Andrew M. & Smith, Jen. (2014). Research data management and libraries: Relationships, activities, drivers and influences. *PLoS One*, 9(12), e114734. <https://doi.org/10.1371/journal.pone.0114734>
- Pinfield, Stephen; Salter, Jennifer & Bath, Peter A. (2016). The “total cost of publication” in a hybrid open-access environment: Institutional approaches to funding journal article-processing charges in combination with subscriptions. *Journal of the Association for Information Science and Technology*, 67(7), 1751-1766. <https://doi.org/10.1002/asi.23446>
- Pinfield, Stephen; Salter, Jennifer & Bath, Peter A. (2017). A “Gold-centric” implementation of open access: Hybrid journals, the “Total cost of publication,” and policy development in the UK and beyond. *Journal of the Association for Information Science and Technology*, 68(9), 2248-2263. <https://doi.org/10.1002/asi.23742>

- Piwowar, Heather; Day, Roger. S. & Fridsma, Douglas. B. (2007). Sharing detailed research data is associated with increased citation rate. *PloS one*, 2(3), e308. <https://doi.org/10.1371/journal.pone.0000308>
- Piwowar, Heather & Chapman, Wendy W. (2010). Public sharing of research datasets: a pilot study of associations. *Journal of informetrics*, 4(2), 148-156. <https://doi.org/10.1016/j.joi.2009.11.010>
- Piwowar, Heather & Vision, Todd J. (2013). Data reuse and the open data citation advantage. *PeerJ*, 1, e175. <https://doi.org/10.7717/peerj.175>
- Piwowar, Heather; Priem, Jason., Larivière, Vincent; Alperín, Juan Pablo; Matthias, Lisa; Norlander, Bree; Farley, Asheley; West, Jevin & Haustein, Stefanie. (2018). The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ*, 6, e4375. <https://doi.org/10.7717/peerj.4375>
- Piwowar, Heather; Priem, Jason, & Orr, Richard. (2019). The Future of OA: A large-scale analysis projecting Open Access publication and readership. *BioRxiv*, 795310. <https://doi.org/10.1101/795310>
- Plesser, Hans E. (2018). Reproducibility vs. replicability: a brief history of a confused terminology. *Frontiers in neuroinformatics*, 11, 76. <https://doi.org/10.3389/fninf.2017.00076>
- Plomp, Esther; Dintzner, Nicolas; Teperek, Marta & Dunning, Alastair. (2019). Cultural obstacles to research data management and sharing at TU Delft. *Insights*, 32(1). <https://doi.org/10.1629/uksg.484>
- Priya, Arya. (2021). Case study methodology of qualitative research: Key attributes and navigating the conundrums in its application. *Sociological Bulletin*, 70(1), 94-110. <https://doi.org/10.1177/0038022920970318>
- R Core Team (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna. <https://www.R-project.org/>
- Reales, Guillermo & Wallace, Chris. (2023). Sharing GWAS summary statistics results in more citations. *Communications Biology*, 6(1), 116. <https://doi.org/10.1038/s42003-023-04497-8>
- Robinson-Garcia, Nicolás; Costas, Rodrigo & van Leeuwen, Thed N. (2020). Open Access uptake by universities worldwide. *PeerJ*, 8, e9410. <https://doi.org/10.7717/peerj.9410>
- Rod, Alisa B.; Hervieux, Sandy & Lee, NuRee (2023). We will meet you where you are. *The Development and Evaluation of Tailored Training for the Management of Research Data*. <https://www.ala.org/acrl/sites/ala.org.acrl/files/content/conferences/confsandpreconfs/2023/WeWillMeetYouWhereYouAre.pdf>

- Rodrigues, Rosângela Schwarz; Abadal, Ernest & De Araújo, Breno Kricheldorf Hermes. (2020). Open access publishers: The new players. *PloS one*, 15(6), e0233432. <https://doi.org/10.1371/journal.pone.0233432>
- Roche, Dominique G.; Kruuk, Loeske E. B.; Lanfear, Robert & Binning, Sandra A. (2015). Public data archiving in ecology and evolution: how well are we doing? *PLoS biology*, 13(11), e1002295. <https://doi.org/10.1371/journal.pbio.1002295>
- Rosati, Kristen B. (2022). Legal Compliance and Good Data Stewardship in Data Sharing Plans. *Harvard Data Science Review*, 4(3). <https://doi.org/10.1162/99608f92.5ff070bf>
- Ross-Hellauer, Tony; Reichmann, Stefan; Cole, Niki Lisa; Fessler, Angela; Klebel, Thomas & Pontika, Nancy. (2022). Dynamics of cumulative advantage and threats to equity in open science: a scoping review. *Royal Society Open Science*, 9(1), 211032. <https://doi.org/10.1098/rsos.211032>
- Salatino, Maximiliano. (2018). Más allá de la indexación: Circuitos de publicación de Ciencias Sociales en Argentina y Brasil. *Dados*, 61(1), 255-287. <https://dx.doi.org/10.1590/001152582018152>
- Salatino, Maximiliano & Banzato, Guillermo. (2021). Confines históricos del acceso abierto latinoamericano. En *Conocimiento abierto en América Latina: Trayectoria y desafíos*, 79-115. <https://doi.org/10.2307/j.ctv2v88f34>
- Salazar, A.; Freán, P.; Chiappero, G. Quintana, D.; Bongiovani, P.C.; Alfieri, A. (2023). *Implementación de Identificadores Persistentes (PID) para aumentar el impacto del Repositorio De Datos Académicos RDA-UNR*(dataverse.unr.edu.ar). XVII Jornadas de Ciencia, Tecnología e Innovación. Universidad Nacional de Rosario. <https://jornadasctei.unr.edu.ar/implementacion-de-identificadores-persistentes-pid-para-aumentar-el-impacto-del-repositorio-de-datos-academicos-rda-unr-dataverse-unr-edu-ar/>
- San Martín, Patricia Silvana; Guarnieri, Griselda; Rodríguez, Guillermo Luján; Bongiovani, Paola Carolina & Sartorio, Alejandro Roberto. (2010). El dispositivo hipermedial dinámico Campus virtual UNR. RepHipUNR <http://hdl.handle.net/2133/1390>
- San Martín, Patricia Silvana; Bongiovani, Paola Carolina; Casali, Ana & Deco, Claudia. (2015). Study on Perspectives Regarding Deposit on Open Access Repositories in the Context of Public Universities in the Central-Eastern Region of Argentina. *Scholarly and Research Communication*, 6(1). <https://doi.org/10.22230/src.2015v6n1a145>
- Sayogo, Djoko Sigit & Pardo, Theresa A. (2013). Exploring the determinants of scientific data sharing: Understanding the motivation to publish research data. *Government information quarterly*, 30, 19-31. <https://doi.org/10.1016/j.giq.2012.06.011>

- Schmidt, Birgit; Gemeinholzer, Birgit & Treloar, Andrew. (2016). Open data in global environmental research: The Belmont Forum's open data survey. *PloS one*, 11(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*, Vol. 52 No. 2, pp. 248-265. <https://doi.org/10.1108/DTA-01-2017-0005>
- Schöpfel, Joachim & Azeroual, Otmane. (2021, April). Rewarding Research Data Management. In *Companion Proceedings of the Web Conference 2021* (pp. 446-450). <https://doi.org/10.1145/3442442.3451367>
- Sierra Bravo, Restituto. (1981). *Ciencias Sociales. Análisis Estadístico y Modelos Matemáticos*. Paraninfo.
- SCOPE Framework for Research Evaluation. (2021). <https://inorms.net/scope-framework-for-research-evaluation/>
- Shankar, Kalpana & Cushing, Amber. (2016). Creating an evidence-based digital curation curriculum for Ireland: Case study at University College Dublin. *An Leabharlann*, 25(2). <http://hdl.handle.net/10197/8124>
- Scott, W. Richard. (2004). Institutional theory. En G. Ritzer *Encyclopedia of Social Theory*(408-414). Thousand Oaks, CA: Sage
- Scott, W. Richard. (2013). *Institutions and organizations: Ideas, interests, and identities*. Sage publications.
- Simard, Marc André ; Ghiasi, Gila ; Mongeon, Philippe & Larivière, Vicent. (2022). National differences in dissemination and use of open access literature. *Plos one*, 17(8), e0272730. <https://doi.org/10.1371/journal.pone.0272730>
- Simons, Natasha; Goodey, Greg; Hardeman, Megan; Clare, Connie; Gonzales, Sara; Strange, Damon; Smith, Graham; Kipnis, Daniel; Iida, Keisuke; Miyairi, Nobuko; Tshetsha, Veliswa; Ramokgola, Mankaleme; Makhera, Pfano & Barbour, Ginny. (2021). *The state of open data 2021*. Digital Science. <https://doi.org/10.6084/m9.figshare.17061347.v1>
- Phillips, Dan & Smit, Michael. (2021). *How To Write A (Good) Data Description: Developing Best Practice*. <https://doi.org/10.5281/zenodo.4709835>
- Sorbonne Declaration on Research Data Rights (2020). <https://sorbonnedatadeclaration.eu/>
- SPARC (2022). *Fact Sheet: White House OSTP Memo on Ensuring Free, Immediate, and Equitable Access to Federally Funded Research*. <https://sparcopen.org/our-work/2022-updated-ostp-policy-guidance/fact-sheet-white-house-ostp-memo-on-ensuring-free-immediate-and-equitable-access-to-federally-funded-research/>

- Stall, Shelley; Martone, Maryann E.; Chandramouliswaran, Iswhar; Federer, Lisa; Gautier, Julian; Gibson, Jennifer; Hahnel, Mark; Larkin, Jennie; Pfeiffer, Nicole; Sedora, Brian; Sim, Ida; Smith, Tim; Van Gulick, Anna E.; Walker, Erin; Wood, Julie; Zaringhalam, Maryam & Zigoni, Alberto. (2023). *Generalist Repository Comparison Chart* (3.0). Zenodo. <https://doi.org/10.5281/zenodo.7946938>
- Stenta, Hernán; Riccardi, Gerardo; Basile, Pedro & Scuderi, Carlos. (2022). *Datos de: Modelación matemática hidrológica-hidráulica del escurrimiento superficial en la cuenca del A° Pavón (Santa Fe, Argentina)* [Data set]. RDA UNR. <https://doi.org/10.57715/UNR/THC0KS>
- Strecker, Dorothea; Axtman, Alexandra; Bertelmann, Ronald; Pal, Jiban K; Cousijn, H; Lindenmann, Iris; Neidiger, Charlotte; Wiit, Michael; Ulrich, Robert R.; Upmeier, Urne; Weisweiler, Nina Leonie & Wright, Sarah J. (2024). *Metadata Schema for the Description of Research Data Repositories: version 4.0*. <https://doi.org/10.48440/re3.014>
- Stuart, David; Baynes, Grace; Hrynaszkiewicz, Iain; Allin, Katie; Penny, Dan; Lucraft, Mithu & Astell, Mathias. (2018). *Practical challenges for researchers in data sharing*. <https://doi.org/10.1002/leap.1184>
- Suhr, Bettina; Dungal, Johanna & Stocker, Alexander. (2020). Search, reuse and sharing of research data in materials science and engineering—A qualitative interview study. *PloS one*, 15(9), e0239216. <https://doi.org/10.1371/journal.pone.0239216>
- Tabachnick, Barbara G.; Fidell, Linda S. & Ullman, J. B. (2013). *Using multivariate statistics*. Boston, MA: Pearson
- Tammaro, Anna Maria; Matusiak, Krystyna K.; Sposito, Frank Andreas; Pervanm Ana & Casarosa, Vittore. (2017). Understanding roles and responsibilities of data curators: an international perspective. *Libellarium: journal for the research of writing, books, & cultural heritage institutions*, 9(2). <https://doi.org/10.15291/libellarium.v9i2.286>
- Tang, Rong & Hu, Zhan. (2019). Providing research data management (RDM) services in libraries: Preparedness, roles, challenges, and training for RDM practice. *Data and Information Management*, 3(2), 84-101. <https://doi.org/10.2478/dim-2019-0009>
- Tenopir, Carol; Allard, Suzie; Douglass, Kimberly; Aydinoglu, Arsev Umur; Wu, Lei; Read, Eleanor; Manoff, Maribeth & Frame, Mike. (2011). Data sharing by scientists: practices and perceptions. *PloS one*, 6(6), e21101. <https://doi.org/10.1371/journal.pone.0021101>
- Tenopir, Carol; Sandusky, Robert J.; Allard, Suzie & Birch, Ben. (2014). Research data management services in academic research libraries and perceptions of librarians. *Library & Information Science Research*, 36(2), 84-90. <https://doi.org/10.1016/j.lisr.2013.11.003>

- 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*, 10(8), e0134826. <https://doi.org/10.1371/journal.pone.0134826>
- Tenopir, Carol; Christian, Lisa., Allard, Suzie., & Borycz, Josh. (2018). Research data sharing: Practices and attitudes of geophysicists. *Earth and Space Science*, 5(12), 891-902. <https://doi.org/10.1029/2018EA000461>
- Tenopir, Carol; Rice, Natalie M.; Allard, Suzie; Baird, Lynn; Borycz, Josh; Christian, Lisa; Grant, Bruce; Olendorf, Robert y Sandusky, Robert. J. (2020). Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. *PloS one*, 15(3), e0229003. <https://doi.org/10.1371/journal.pone.0229003>
- The Turing Way Community. (2022). *The Turing Way: A handbook for reproducible, ethical and collaborative research*. <https://doi.org/10.5281/zenodo.7625728>
- Thelwall, Mike & Kousha, Kavian. (2017). Do journal data sharing mandates work? Life sciences evidence from Dryad. *Aslib Journal of Information Management*. <https://doi.org/10.1108/AJIM-09-2016-0159>
- Torres-Salinas, Daniel. (2020). Ritmo de crecimiento diario de la producción científica sobre Covid-19. Análisis en bases de datos y repositorios en acceso abierto. *El profesional de la información*, 29(2). <http://hdl.handle.net/10481/66008>
- Towse, John N.; Ellis, David A. & Towse, Andrea S. (2021). Opening Pandora's Box: Peeking inside Psychology's data sharing practices, and seven recommendations for change. *Behavior Research Methods*, 53, 1455-1468. <https://doi.org/10.3758/s13428-020-01486-1>
- Unal, Yurdagül; Chowdhury, Gobinda; Kurbanoglu, Serap; Boustany, Joumana & Walton, Geoff. (2019). Research data management and data sharing behaviour of university researchers. *Information Research: an international electronic journal*, 24(1). <https://informationr.net/ir/24-1/isic2018/isic1818.html>
- University of Cambridge (2023). *Data champions*. <https://www.data.cam.ac.uk/intro-data-champions>
- Uzwysyn, Ray. (2016). Online Research Data Repositories: The What, When Why and How. *Computers in Libraries*, 36(3), pp. 18-21. <https://www.infotoday.com/cilmag/apr16/Uzwysyn--Research-Data-Repositories.shtml>
- UNESCO. (2021). *Recomendación de la UNESCO sobre la Ciencia abierta*. https://unesdoc.unesco.org/ark:/48223/pf0000379949_spa
- Universidad Nacional de Rosario (2016). *Resolución de Consejo Superior N° 1842/2016*. RepHipUNR <http://hdl.handle.net/2133/12981>

- Universidad Nacional de Rosario (2017). *Resolución N° 5425/2017*. RepHipUNR <http://hdl.handle.net/2133/12982>
- Universidad Nacional de Rosario (2018). *Resolución N° 2159/2018*. RepHipUNR <http://hdl.handle.net/2133/12983>
- Universidad Nacional de Rosario. (2020). *Capacitación "Gestión de datos de investigación en la UNR"* [Video]. YouTube. <https://youtu.be/RCgnARvtx8?si=DJIdGvyl0cZ6dX9g>
- U.S. Department of Health and Human Services. (1991). Federal policy for the protection of human subjects ("Common rule"). <https://www.hhs.gov/ohrp/regulations-and-policy/regulations/common-rule/index.html>
- Van House, Nancy A. (2003). Digital libraries and collaborative knowledge. *Digital library use: Social practice in design and evaluation* (271-95). <https://doi.org/10.7551/mitpress/2424.003.0015>
- Van den Eynden, Veerle; Knight, Gareth; Vlad, Anca; Radler, Barry; Tenopir, Carol; Leon, David A.; Manista, Frank; Whitworth, Jimmy & Corti, Louise. (2016). Survey of Wellcome researchers and their attitudes to open research. <https://doi.org/10.6084/m9.figshare.4055448.v1>
- Vicente-Sáez, Rubén & Martínez-Fuentes, Clara. (2018). Open Science now: A systematic literature review for an integrated definition. *Journal of business research*, 88, 428-436. <https://doi.org/10.1016/j.jbusres.2017.12.043>
- Voigt, Paul & Von dem Bussche, Axel. (2017). The EU general data protection regulation (GDPR). *A Practical Guide, 1st Ed., Cham: Springer International Publishing*, 10(3152676), 10-5555. <https://doi.org/10.1007/978-3-319-57959-7>
- Volk, Carol J.; Lucero, Yasmin & Barnas, Katie. (2014). Why is data sharing in collaborative natural resource efforts so hard and what can we do to improve it?. *Environmental management*, 53, 883-893. <https://doi.org/10.1007/s00267-014-0258-2>
- 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*, 8(7), e67332. <https://doi.org/10.1371/journal.pone.0067332>
- Wildgaard, Loma & Rantasaari, Jukka. (2022). *RDA Professionalising Data Stewardship - Data Stewardship Landscape Initial Report*. <https://doi.org/10.15497/RDA00076>

- Wilkinson, Mark D. ; Dumontier, Michel ; Aalbersberg, IJsbrand Jan ; Appleton, Gabrielle; Axton, Myles; Baak, Arie; Blomberg, Niklas; Boiten, Jan-Willem; da Silva Santos, Luiz Bonino; Bourne, Philip E.; Bouwman, Jildau; Brookes, Anthony J.; Clark, Tim; Crosas, Mercè; Dillo, Ingrid; Dumon, Olivier; Edmunds, Scott; Evelo, Chris T.; Finkers, Richard; Gonzalez-Beltran, Alejandra; Gray, Alasdair J.G.; Groth, Paul; Goble, Carole; Grethe, Jeffrey S. & Mons, Barend. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3(1), 1-9. <https://doi.org/10.1038/sdata.2016.18>
- Wittrock, Christian; Forsberg, Ellen Marie; Pols, Auke ; Macnaghten, Philip & Ludwig, David. (2021). *Implementing responsible research and innovation: Organisational and national conditions* (p. 120). Springer Nature. <http://dx.doi.org/10.1007/978-3-030-54286-3>
- Woltman, Heather; Feldstain, Andrea; MacKay, J. Christine., & Rocchi, Meredith. (2012). An introduction to hierarchical linear modeling. *Tutorials in quantitative methods for psychology*, 8(1), 52-69. <https://www.tqmp.org/RegularArticles/vol08-1/p052/p052.pdf>
- Woods, Helen Buckley & Pinfield, Stephen. (2021). Incentivising research data sharing: a scoping review. *Wellcome open research*, 6. <https://doi.org/10.12688/wellcomeopenres.17286.1>
- World Medical Association. (2013). *Declaration of Helsinki*. <https://www.wma.net/what-we-do/medical-ethics/declaration-of-helsinki>
- Wunderlich, Richard. (1974). The scientific ethos: A clarification. *British Journal of Sociology*, 373-377. <https://doi.org/10.2307/589402>
- Yoon, Ayoung & Kim, Youngseek. (2017). Social scientists' data reuse behaviors: Exploring the roles of attitudinal beliefs, attitudes, norms, and data repositories. *Library & Information Science Research*, 39(3), 224-233. <https://doi.org/10.1016/j.lisr.2017.07.008>
- Yoon, Ayoung. (2017). Data reusers' trust development. *Journal of the Association for Information Science and Technology*, 68(4), 946-956. <https://doi.org/10.1002/asi.23730>
- Zenk-Möltgen, Wolfgang Zenk; 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*, 74(5), 1053-1073. <https://doi.org/10.1108/JD-09-2017-0126>
- Zhu, Yimei. (2020). Open-access policy and data-sharing practice in UK academia. *Journal of Information Science*, 46(1), 41-52. <https://doi.org/10.1177/0165551518823174>
- Zimmerman, Ann. (2007). Not by metadata alone: The use of diverse forms of knowledge to locate data for reuse. *International Journal on Digital Libraries*, 7, 5-16. <https://doi.org/10.1007/s00799-007-0015-8>

- Zuiderwijk, Anneke & Spiers, Helen. (2019). Sharing and re-using open data: A case study of motivations in astrophysics. *International Journal of Information Management*, 49, 228-241. <https://doi.org/10.1016/j.ijinfomgt.2019.05.024>
- Zuiderwijk, Anneke; Shinde, Rhytima & Jeng, Wei. (2020). What drives and inhibits researchers to share and use open research data? A systematic literature review to analyze factors influencing open research data adoption. *PLoS one*, 15(9), e0239283. <https://doi.org/10.1371/journal.pone.0239283>

El potencial de los datos de investigación abiertos: repositorios y ciencia abierta para la investigación

Paola Carolina Bongiovani

La autora.

Paola Carolina Bongiovani, es una investigadora y profesional de la información argentina especializada en ciencia abierta, comunicación científica y acceso abierto. Desarrolla su actividad principalmente en la Universidad Nacional de Rosario (Argentina), vinculada a la gestión de información académica, bibliotecas universitarias y proyectos de investigación. Su trabajo se centra en el análisis de los sistemas de evaluación científica, los modelos de publicación académica y las infraestructuras abiertas de conocimiento. Ha participado en iniciativas nacionales e internacionales para la promoción de políticas de ciencia abierta en América Latina, con un compromiso sostenido con la equidad y la apertura del conocimiento científico.

El libro.

El manuscrito analiza el potencial de los datos de investigación abiertos en el marco de la ciencia abierta, con especial atención al papel de los repositorios, las políticas de acceso abierto y los principios FAIR. Examina los factores técnicos, éticos, legales y culturales que condicionan la apertura y reutilización de los datos, así como las tensiones asociadas a los sistemas de evaluación científica, desde una perspectiva crítica centrada en el contexto latinoamericano.

La serie editorial

Esta serie aspira a difundir trabajos de valor científico llamados a convertirse en textos de referencia para el estudio de la gestión de información. Está abierta a textos teóricos originales y a resultados de investigación con impacto y transferencia social. La serie promueve una investigación rigurosa y una divulgación científica de calidad en el ecosistema informacional actual. Las obras se someten a una evaluación académica que garantiza su solvencia. Aplicando los principios de la ciencia abierta, los textos se publican íntegramente en acceso abierto en el portal de libros abiertos de [EDITUM](#) y se depositan en el repositorio institucional [DIGITUM](#), con el respaldo de la [Cátedra UNESCO en Gestión de Información en las Organizaciones](#).

