

EVENTOS

2024

XX Workshop de Proyectos Digitales



TÍTULO COMUNICACIÓN.....



crue

Universidades
Españolas

I+D+i

Red de Bibliotecas
REBIUN

PERCEPCIONES DE INVESTIGADORAS E INVESTIGADORES DEL ÁREA DE EDUCACIÓN SOBRE EL DEPÓSITO DE DATOS EN REPOSITARIOS

Celia Martínez-Córdoba ¹

Lidya Groppi-Bosch ¹

Rafael Aleixandre-Benavent ^{1,3}

Andrea Sixto-Costoya ^{1,4}

Inmaculada Chiva-Sanchís ⁵

Adolfo Alonso-Arroyo ^{1,2}

¹ Grupo de investigación UISYS, Unidad de Información e Investigación Social y Sanitaria, Universitat de València. Unidad Asociada a INAEUC. Instituto Interuniversitario de Investigación Avanzada sobre Evaluación de la Ciencia y la Universidad. UC3M-UAM, España

² Departamento de Historia de la Ciencia y Documentación, Universitat de València, España

³ Ingenio (CSIC-Universitat Politècnica de València), España

⁴ Departamento de Trabajo Social y Servicios Sociales, Universitat de València, España

⁵ Departamento de Métodos de Investigación y Diagnóstico en Educación, Universitat de València, España

Introducción

Cada vez se produce una mayor cantidad de datos de investigación y, sin embargo, muchos permanecen ocultos y sin utilizar, desperdiciando su potencial para generar nuevas hipótesis, conocimientos científicos e innovación (Berghmans et al, 2017; Popkin, 2019). El acceso abierto a estos datos es crucial para fomentar el intercambio y el avance científico, y es por ello que resulta de gran importancia la promoción de su almacenamiento, distribución, uso y preservación eficientes (Martínez-Uribe y Macdonald, 2009). En el caso de las Ciencias de la Educación, la falta de prácticas establecidas para compartir datos representa una barrera que afecta al potencial de la investigación y, por consiguiente, al cambio social. En respuesta a estos desafíos, es crucial comprender las condiciones actuales del intercambio de datos en la investigación educativa y desarrollar estrategias para mejorar estas prácticas (Parra-Calderon et al, 2020).

La presente comunicación tiene como objetivo investigar las percepciones acerca del conocimiento y uso de los repositorios para el depósito y uso compartido de datos en el área de Ciencia de la Educación, ya que se plantea necesario modificar los hábitos de investigadores/as y organizaciones, así como abordar la falta de reconocimiento de la importancia del intercambio de datos y aplicar principios sobre las cualidades que deberían cumplir los datos para garantizar su utilidad y fiabilidad.

Metodología

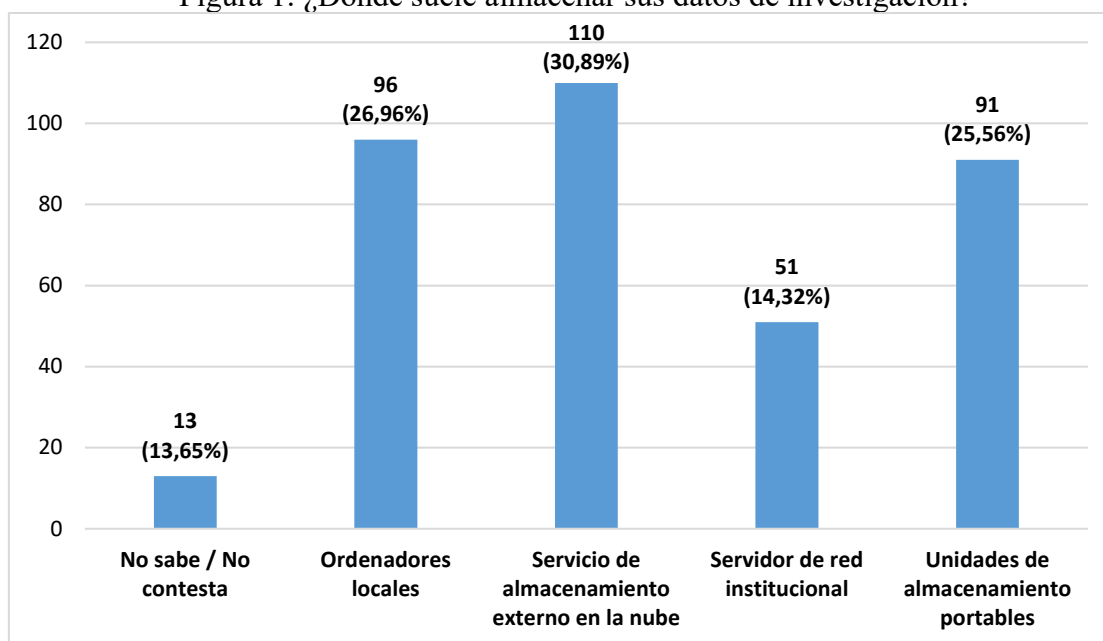
Se llevó a cabo el diseño e implementación de una encuesta anónima y confidencial administrada por correo electrónico al profesorado e investigadores/as del área de Ciencias de la Educación de universidades públicas españolas. El diseño de la encuesta se basó en estudios previos realizados por el grupo investigador sobre el uso compartido

de datos y su reutilización, a partir de este diseño previo se realizó una validación lógica de la encuesta basada en un panel de expertos mixto, integrado por especialistas en información y profesionales con experiencia en investigación y docencia en el área educativa. Verificado el análisis lógico y empírico de la encuesta, se implementó la versión definitiva entre la población objeto de estudio a través de LimeSurvey.

Resultados

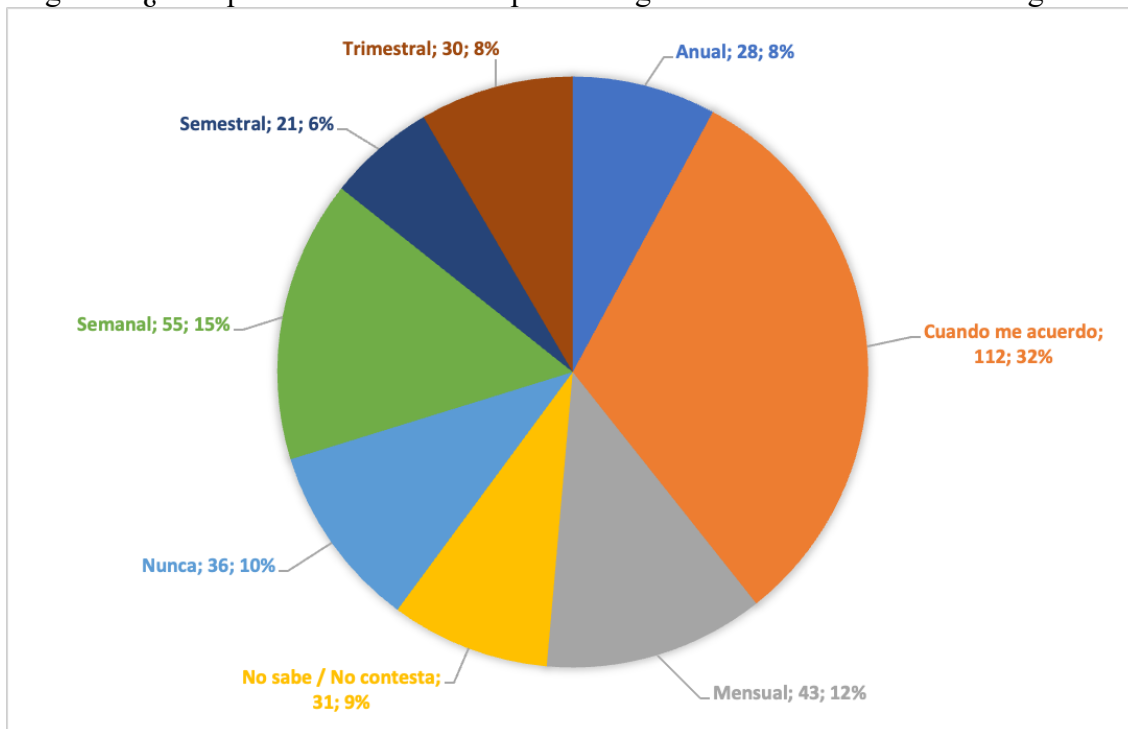
Se obtuvieron un total de 356 encuestas. En primera instancia, se preguntó a los y las profesionales donde suelen almacenar sus datos de investigación. En la Figura 1 vemos como la mayoría, un 30.89%, recurren a servicios de almacenamiento en la nube, mientras que un 26.96% opta por almacenar sus datos en ordenadores locales, y un 25.56% en unidades de almacenamiento portátiles. Por otro lado, una proporción minoritaria utiliza servidores de red institucionales (14.32%). Solo un 3.65% no consideró o no supo responder a esta pregunta.

Figura 1: ¿Dónde suele almacenar sus datos de investigación?



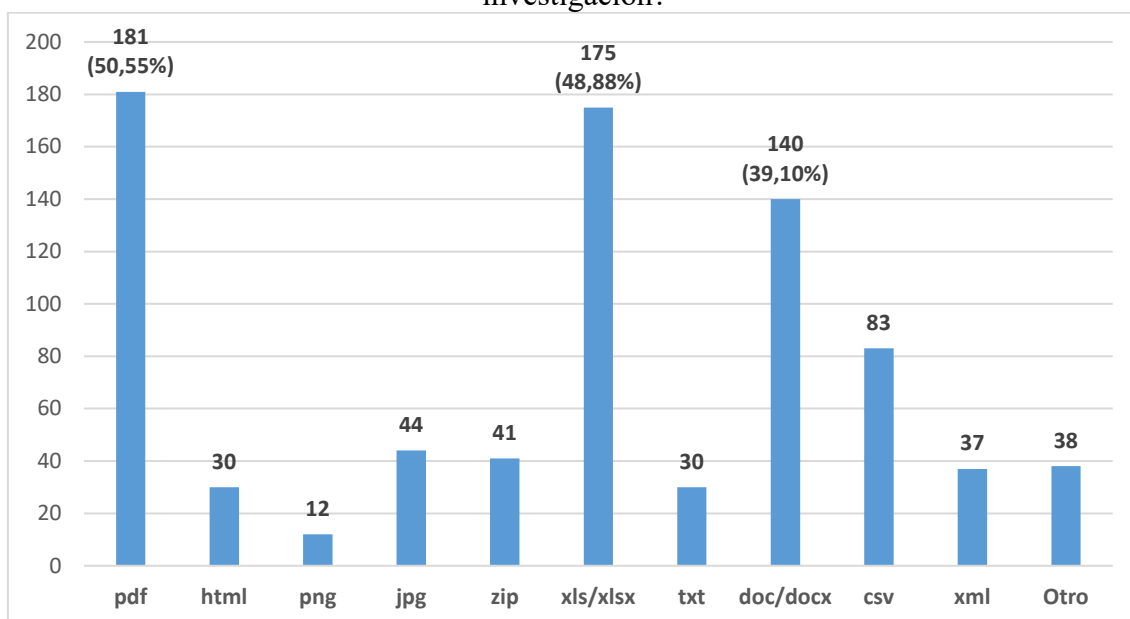
Respecto a las copias de seguridad, los resultados revelan que la mayoría de investigadores/as muestran una variedad considerable en la frecuencia con la que realizan sus copias, con una significativa proporción de un 31.46% que indica hacerlo "cuando se acuerda". El resto de las opciones se encuentra más equilibrado, aunque con ligero despunte de un 15.45% que realiza copias semanales y un 12.07% que las efectúa mensualmente. Las opciones menos escogidas por los/as encuestados/as fueron las copias anuales (7.86%) y semestrales (5.89%). Es reseñable que haya un 10.11% que declara no hacer copias de seguridad nunca (Figura 2).

Figura 2: ¿Con qué frecuencia hace copias de seguridad de sus datos de investigación?



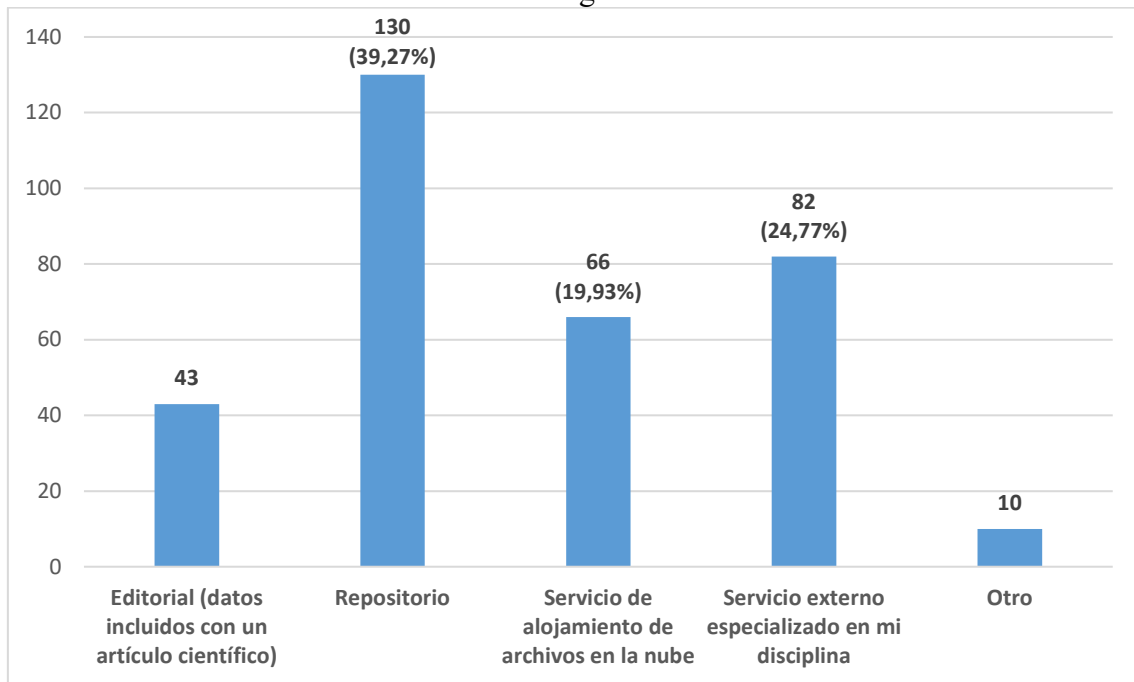
En cuanto al tratamiento de los datos, se obtuvo que un 81,46% utiliza principalmente formatos propietarios, frente a un 12,92% que emplea formatos libres y pequeños porcentajes que o bien no respondieron, o bien utilizan otros formatos. Desglosando estos formatos en los que suelen estar los archivos que contienen los datos brutos de investigación, entre los/as encuestados/as destaca el uso de formatos como .pdf (50.55%), .xls/xlsx (48.88%) y .doc/docx (39.10%). En la Figura 3 vemos cuáles son los archivos empleados. Por otro lado, se observó que los datos recopilados eran diversos en su tipología: datos cuantitativos procedentes de encuestas (68.71%), transcripciones de entrevistas (44.13%), seguidos de datos cualitativos codificados (35.75%) y material pedagógico reutilizable (29.60%), entre otros.

Figura 3: ¿En qué formato suelen estar los archivos que contienen sus datos brutos de investigación?



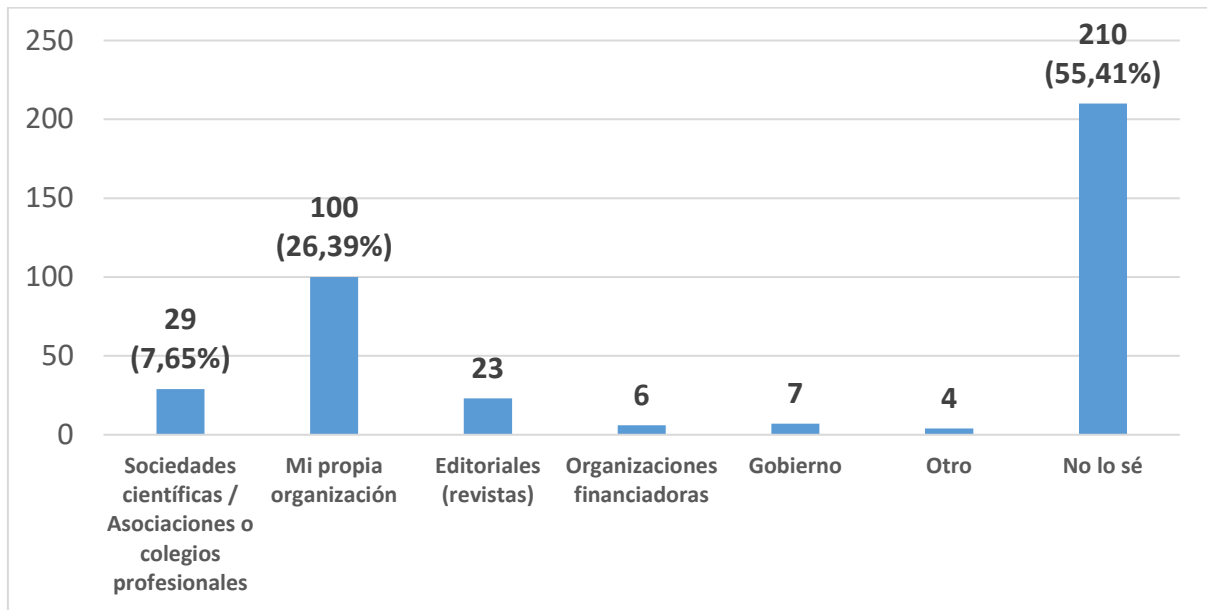
En cuanto a las preferencias a la hora de depositar los datos, en la Figura 4 puede verse que el 39.27% considera los repositorios como el medio óptimo para la preservación y compartición de datos, seguido por servicios especializados en sus respectivas disciplinas (24.77%) y servicios de almacenamiento en la nube (19.93%). Sin embargo, ante la pregunta de si conoce algún repositorio para la preservación de datos en su disciplina, hay un 62.53% que no está familiarizado/a con repositorios especializados para la preservación de datos digitales en su ámbito. Un 24.5% no sabe/no contesta, y un 13% que sí tienen conocimiento de estos repositorios. La mayoría menciona repositorios institucionales o multidisciplinares (Figshare, Zenodo). Por otro lado, únicamente un 4.87% de los/las encuestados/as está al tanto de planes para la creación de un repositorio o archivo de datos en su comunidad.

Figura 4: ¿Dónde considera más adecuado que se preserven y compartan los datos de sus investigaciones?



Como puede verse en la Figura 5, en lo que respecta a guías o recomendaciones para la preservación de datos, el 26.39% considera que provienen de las propias organizaciones, seguido por un 7.65% que menciona a las sociedades científicas. No obstante, un notable 55.41% no está informado sobre quién proporciona estas guías o recomendaciones. En relación con el apoyo recibido por parte de sus organizaciones para la preservación de datos, el 26.28% de los investigadores afirma recibir ayuda, mientras que un 43.20% desconoce si cuentan con este respaldo, y un 30.51% manifiesta no recibirlo. Respecto al término "Principios FAIR", solo un 16% de los investigadores parece estar familiarizado con él.

Figura 5: En su disciplina, ¿quién proporciona guías o recomendaciones para preservar datos?



En una pregunta final destinada a evaluar las razones de mayor peso para la preservación y apertura de datos, los resultados indican que la financiación pública de la investigación como justificación para la divulgación de resultados obtuvo la puntuación más alta (54.10%), seguido por la utilidad de los datos para validar investigaciones (43.47%), la posibilidad de re-análisis de datos existentes (41.64%), y el potencial para fomentar colaboraciones interdisciplinarias e impulsar la innovación educativa (39.21%).

Conclusión

La encuesta realizada en el marco de este estudio arroja luz sobre las prácticas actuales de almacenamiento, gestión y preservación de datos de investigación. Los resultados revelan patrones y tendencias relevantes en cada una de estas áreas, ofreciendo información valiosa para comprender el panorama actual de la gestión de datos de investigación y para identificar áreas que requieren atención y mejora. Del mismo modo, también revelan la existencia de importantes barreras y desafíos en el ámbito de Ciencias de la Educación en lo que respecta al uso y compartición de datos.

La falta de prácticas establecidas representa una limitación significativa que afecta el potencial de la investigación en esta área. A pesar de que la mayoría de los/as investigadores/as considera que es importante depositar los datos y conocen la existencia de repositorios institucionales, una minoría conocen la existencia de repositorios especializados en el depósito de datos. Tampoco conocen los planes de sus instituciones respecto al depósito y gestión de datos, y la mayoría desconocen los principios FAIR de encontrabilidad, accesibilidad, interoperabilidad y reutilización, por lo que resulta evidente la necesidad de modificar hábitos, tanto de los investigadores como de las organizaciones, así como abordar la falta de reconocimiento de la importancia del intercambio de datos, siendo crucial aplicar principios que garanticen la utilidad y fiabilidad de estos. No obstante, es alentador observar que existe un reconocimiento de la importancia de la preservación y apertura de datos, especialmente cuando la investigación está financiada públicamente.

Para fomentar el depósito y reutilización de los datos brutos de investigación, deberían implementarse diversas medidas complementarias que permitan superar las barreras técnicas, económicas, éticas, legales y motivacionales que lo dificultan. Una de las medidas más importantes es la realización de campañas de formación a los investigadores en las que, además de informarles del amplio abanico de ventajas y oportunidades que supone compartir datos, se dé respuesta a sus preocupaciones, la mayoría relacionadas con los aspectos legales y de propiedad intelectual, como el uso indebido de los datos por terceros o su apropiación sin el debido reconocimiento a los propietarios de los datos. Estas campañas deben incluir también la formación técnica básica sobre qué tipos de datos pueden o deben compartirse y cuáles son los formatos más apropiados para el depósito, de manera que puedan reutilizarse posteriormente sin barreras. Otra medida necesaria es proporcionar el apoyo técnico preciso en todos los procesos relacionados con la gestión de los datos. Este apoyo deberían proporcionarlo las instituciones y convendría que corriera a cargo de profesionales especializados en planes de gestión de datos. Sería conveniente que las instituciones convocaran ayudas específicas para subvencionar los planes de gestión de datos, o que las convocatorias de proyectos contemplaran poder destinar una dotación presupuestaria a este propósito. Además, los especialistas en gestión de datos deberían contribuir, en contacto directo con los investigadores de cada área de la ciencia, a que los datos depositados sean realmente útiles y cumplan los principios FAIR establecidos por la comunidad internacional (encontrables, accesibles, interoperables y reutilizables). Dado el esfuerzo que supone la gestión adecuada de los datos, una medida que puede mejorar considerablemente su práctica es el reconocimiento a quienes la cultiven, ya sea en forma de créditos curriculares o con otros tipos de recompensas académicas, o incluso con incentivos económicos.

Referencias bibliográficas

- Berghmans, S; Cousijn, H; Deakin, G; Meijer, I; Mulligan, A; Plume, A; Rijcke, S; Rushforth, A; Tatum, CT; Van Leeuwen, T; Waltman, L. (2017). Open Data: the researcher perspective – survey and case studies. <https://doi.org/10.17632/bwrnfb4bv.1>
- Martinez-Urbe, L; Macdonald, S. (2009). User Engagement in Research Data Curation. <http://www.era.lib.ed.ac.uk/handle/1842/3206>
- Parra-Calderon, CL; Sanz, F; McIntosh, LD. (2020). The Challenge of the Effective Implementation of FAIR Principles in Biomedical Research. *Methods of Information in Medicine*, 59(4/5), 117-118. doi: 10.1055/s-0040-1721726
- Popkin, G. (2019). Data sharing and how it can benefit your scientific career. *Nature*, 569(7756), 445–447. <https://doi.org/10.1038/d41586-019-01506-x>



Este trabajo se ha beneficiado de una ayuda del Ministerio de Ciencia e Innovación del Gobierno de España. Agencia Estatal de Investigación. Proyectos de Transición Ecológica y Digital 2021. TED2021-131057B-I00

 **crue** Universidades
Españolas

I+D+i

Red de Bibliotecas
REBIUN