

## **Gestión de datos de investigación y alfabetización en datos como procesos que llegaron para quedarse**

Research data management and data literacy as processes that are here to  
stay

Yusnelkis Milanés Guisado<sup>1\*</sup> <http://orcid.org/0000-0001-5461-9309>

Lourdes Feria Basurto<sup>2</sup> <http://orcid.org/0000-0002-4326-4087>

<sup>1</sup>Universidad Pablo de Olavide, Biblioteca/CRAi. España.

<sup>2</sup>Consultora Independiente en Gestión de Datos, Información y Tecnología, México.

\*Autor para la correspondencia: [ymilanes@gmail.com](mailto:ymilanes@gmail.com)

Recibido: 10/12/2022

Aceptado: 23/12/2022

*¿Qué son los datos más allá de números, variables, algoritmos, representaciones alfabéticas, tablas o gráficas? Los datos son el punto de partida, la base de la investigación científica, la plataforma de estrategias competitivas en las empresas y, en el día a día, se han convertido en componente de decisiones tan familiares como pueden ser el conteo de los pasos que damos, la música que elegimos y los pagos que hacemos. De manera muy simple, los datos son estructuras básicas, en bruto, unidades de*

**Gestión de datos de investigación y alfabetización en datos como procesos que llegaron para quedarse**

*observación que por sí solos no tienen mayor sentido. Cuando ponemos los datos en contexto, aplicamos procesos de agregación, agrupamiento, añadimos atributos, le damos significado, lo convertimos en información, conocimiento y/o sabiduría (Pirámide DIKW). La finalidad es poder tomar decisiones correctas a partir de los datos, en cualquier escenario. Es importante reconocer la omnipresencia de los datos, su fluidez. La realidad es que el incremento de los procesos y capacidades asociados al “big data” han creado nuevas oportunidades para muchas facetas de la sociedad, también para la investigación. Ello ha llevado a la necesidad de crear una mentalidad basada en datos y también una cultura analítica, en tanto se trata de adoptar nuevas tecnologías (Harris, 2012). La International Telecommunication Union identifica como una de las habilidades digitales necesarias en el siglo 21 la información/alfabetización en datos, si bien en su informe más reciente (2022) reporta que de la población mundial que tiene acceso a internet sólo un 25% se puede considerar alfabetizada en datos (Measuring Digital Development, s. f. 2022), lo cual permite observar el largo camino por recorrer para cerrar esa brecha, pensando incluso en las siguientes generaciones de usuarios.*

*La presencia de los datos nos está llevando a imaginar y moldear nuevos servicios y usos de la información y a considerar la toma de consciencia en torno a los usos éticos, posibles sesgos y nuevas realidades, tales como los impactos sociales de la inteligencia artificial. La vida en un entorno de redes sociales, motores de búsqueda, aplicaciones de inteligencia artificial, compras en línea y una miríada de aplicaciones digitales más les hace insumo básico en todos los ámbitos del saber humano y es por ello que desde las ciencias de la Información, la gestión de datos y la alfabetización en datos van tomando cada vez más fuerza como actividades imprescindibles.*

**Gestión de datos de investigación y alfabetización en datos como procesos que llegaron para quedarse**

*En el ámbito científico, los investigadores siempre han gestionado sus datos en el marco de sus proyectos. No hay novedad alguna en ello. Al inicio de cada investigación, debes concebir cómo vas a recoger los datos, instrumentos, procesarlos, realizar el análisis y extraer conclusiones de tu estudio. Sin embargo, los investigadores han hecho esto adoptando prácticas que muchas veces no permiten que la investigación sea un proceso efectivo y reproducible. Se suelen cometer errores como: no comenzar con un fin en mente en materia de qué harás con tus datos en todo el proyecto, no documentar adecuadamente los datos y los procedimientos asociados, no usar diccionarios de datos, no realizar copias máster de los datos brutos, no usar nomenclaturas para los ficheros ni estructuras coherentes y lógicas para ordenar los registros, no compartir datos brutos de calidad para que los resultados puedan ser reproducidos, o compartirlos en formatos propietarios, etc. Esto trae consigo no sólo ineffectividad en la investigación, pérdida de tiempo, sino un impacto en los costes del proyecto.*

*Por otro lado, muchos artículos científicos publicados en revistas indexadas han sido retractados (Dal-Ré, 2020). La estadística es manipulable. Algunos investigadores se obsesionan con demostrar sus hipótesis de investigación. Si los datos no acompañan al artículo, es inviable poder reproducir y verificar la calidad de la investigación publicada. Por ello, cada vez más revistas y editoriales exigen o sugieren el compartir los datos brutos, ya sea como material suplementario del artículo donde las propias revistas y editoriales puedan ejercer de plataformas digitales en las que se almacenan datos; o la posibilidad de subir los datos a un repositorio.*

*Por otro lado, y no menos importante, de la mano de todo el movimiento de acceso abierto a ciencia abierta y ciudadana, los organismos nacionales e internacionales de financiación de la investigación han emitido una serie de requerimientos, políticas y orientaciones sobre la gestión de datos en el contexto de la investigación. Ello requiere*

**Gestión de datos de investigación y alfabetización en datos como procesos que llegaron para quedarse**

*al investigador poder demostrar tanto en la propuesta del proyecto como durante y una vez concluido, que son capaces de gestionar adecuadamente los datos, con miras a poder compartirllos en abierto en repositorios de confianza y permitir, por tanto, la reproducibilidad de los resultados y el avance del conocimiento científico. Aquí es también relevante las propias políticas en datos que adopten las instituciones. Este entorno de políticas y requisitos, tan conocido como “el palo de la zanahoria”, es un contexto que por un lado presiona, pero también promueve un cambio de rutinas y hábitos en el investigador, que ya está saturado de carga, presión por publicar y también horas lectivas en el caso de los docentes.*

*La labor de las bibliotecas universitarias, centros de documentación en centros de investigación y todo el sistema de investigación es justamente facilitar este proceso, crear estructuras y políticas para “ponérselo fácil al investigador”, concienciando, reduciendo o eliminando posibles barreras, que de manera general no suelen ser pocas. En este sentido, las bibliotecas y los bibliotecarios están viviendo un cambio importante para adoptar nuevos modelos de servicios y desarrollar las competencias necesarias para hacer frente a estos nuevos retos. La IFLA plantea 4 preguntas clave en este sentido: ¿Cómo pueden ayudar las bibliotecas a los investigadores en la Gestión de Datos de Investigación (GDI)?, ¿Qué habilidades necesita el personal de la biblioteca para proporcionar esta ayuda?, ¿Cómo diseñan y ofrecen las bibliotecas servicios de datos?, y ¿Cómo pueden fomentar las bibliotecas la GDI? («Servicios de datos de investigación. especial IFLA journal», 2017).*

*Por un lado, las bibliotecas han de diseñar e implementar servicios de gestión de datos de investigación a la altura de los nuevos requerimientos. Ello no supone copiar recetas posiblemente exitosas para otros centros y universidades. El diseño de un SGDI implica ser innovadores, con soluciones personalizadas, adoptando las propuestas a los recursos*

**Gestión de datos de investigación y alfabetización en datos como procesos que llegaron para quedarse**

*y estrategias de cada centro. Además, necesitamos servicios interconectados (Rebecca Bryant, 2021). Podría resultar ineficiente una propuesta de servicio de GDI donde no se trabajará de la mano del de automatización, y apoyo a la investigación, por ejemplo. Máxime cuando se deben tomar decisiones conjuntas en cuanto a los recursos técnicos, especialmente los repositorios, en el terreno tecnológico, estratégico, financiero e incluso, y no menos relevante, de diseño y aprobación institucional de una política de datos.*

*Un artículo publicado en el 2019 “Research Data Services in Academic libraries: Where we are today?” afirmaba que el 44.1 % de las bibliotecas respondieron no estar involucradas en ningún proyecto de GDI («Research Data Services in Academic Libraries», 2021). Nos animamos a pensar que la realidad pueda ser distinta, sobre todo a futuro. Muchas universidades americanas y europeas son ejemplo de buenas prácticas en este terreno. Si nos detenemos a pensar, estamos presenciando un cambio importante en la transición y madurez de los servicios de gestión de datos de investigación, desde el tradicional acceso y referencia con productos informativos como las libguides, al bibliotecario integrado, ofreciendo servicios enfocados a estudiantes de doctorado, investigadores, grupos de investigación en programas permanentes que integren software, capacitación y asesorías en curación de datos, alfabetización en datos, planes de gestión de datos y principios FAIR, tratamiento de datos sensibles y anonimización, compartir datos en abierto, uso de metadatos, análisis y visualización, etc. Todo ello con el propósito de favorecer el fortalecimiento de habilidades digitales cruciales para aprovechar las tecnologías y los datos como detonadores de innovación y producción científica. Sin dudas, supone un importante espacio para visibilizar la labor e influencia de la biblioteca y el bibliotecario en los resultados e impacto de la universidad y los centros de investigación en la sociedad.*

**Gestión de datos de investigación y alfabetización en datos como procesos que llegaron para quedarse**

*La gestión de datos, como hemos mencionado, ya se encuentra integrada en muchas bibliotecas, al menos en esa fase inicial de una guía como producto de información para orientar a los investigadores y las primeras acciones formativas. Todo ello conlleva una serie de competencias necesarias para su tratamiento tanto como fuente de información propiamente y como considerarse un recurso a ser preservado en los repositorios digitales. En este contexto hay dos retos a desarrollar en el entorno de las bibliotecas: por un lado, la creación de habilidades en datos en los bibliotecarios y posteriormente desarrollar programas de alfabetización en datos para la comunidad universitaria.*

*La alfabetización en datos, podría considerarse como un punto de partida en el desarrollo de habilidades para “acceder, medir, manipular, resumir y presentar datos. El concepto de “alfabetización en datos” es relativamente nuevo y se considera que las diferencias con “alfabetización de información” es nebuloso (Hunt, 2004). Lo cierto es que para los bibliotecarios, que ya hemos desarrollado como base muchas competencias relevantes en este contexto como la orientación al servicio, la gestión de información, la visualización de información, estadística básica, otras tareas en el área de la ciencia de datos conllevan el reciclaje en cuanto a los software a utilizar para el tratamiento, análisis y visualización de datos.*

*Hay una realidad que ciertamente nos anima y que las autoras presenciamos a través de nuestras realidades profesionales, y acciones formativas en materia de datos con bibliotecarios y documentalistas: cada vez los bibliotecarios se forman más en datos, tienen mayor interés en ello y están asumiendo nuevos roles pese a una serie de barreras institucionales. Las realidades por áreas geográficas también son bastante diferentes. En muchos países anglosajones los servicios de datos desde las bibliotecas están completamente integrados con los laboratorios de ciencia de datos.*

**Gestión de datos de investigación y alfabetización en datos como procesos que llegaron para quedarse**

*En todo este contexto, hay dos gaps importantes que superar: the skills gap y the management gap, que giran en torno a que los bibliotecarios aprenden nuevas habilidades, pero necesitan continuar realizando las anteriores (Burton et al., 2018). Y ello versa sobre la base de que ciertamente si las competencias en datos no son parte de la estrategia institucional, es difícil encontrar tiempo y recursos para actualizar el conjunto de competencias mientras cumple con sus obligaciones laborales. Además, las propias barreras institucionales de falta de recursos humanos, financiación, estrategia, visión, etc. Sin embargo, nos alienta saber y pensar que la influencia conjunta de varios factores como la omnipresencia y relevancia de los datos, el avance de las políticas en ciencia abierta y evaluación, la propia evolución de las funciones de la biblioteca en el contexto universitario y de investigación, incremento de acciones formativas en datos para los bibliotecarios a todos los niveles, fortalecimiento del rol de los servicios de apoyo a la investigación, recodificación de los propios modelos mentales de los directivos de las bibliotecas en este sentido y de los propios bibliotecarios, está llevando a un avance en este sentido y a que cada vez más los bibliotecarios asuman estos roles. De hecho, aunque falta terreno podría variar bastante de un área del conocimiento a otra, ya hay evidencias de ese acompañamiento informacional y en materia de datos por parte del bibliotecario al investigador. No estamos ante una quimera. Es una realidad imaginar a un bibliotecario más que sobrecargado, siendo co-autor junto a los investigadores (Borrego et al., 2018; Rethlefsen et al., 2015).*

*Hay muchos programas ya dedicados a la formación en datos para bibliotecarios. Algunas iniciativas exitosas en este sentido incluyen Data Carpentry, Library Carpentry, MANTRA Research Data Management Training del Servicio de datos de la Universidad de Edimburgo, Essentials 4 Data Support, o el Servicio de datos de Uk Data Service) con una gama de recursos y videos de formación. En el ámbito del apoyo a la investigación,*

**Gestión de datos de investigación y alfabetización en datos como procesos que llegaron para quedarse**

*el programa del Título Experto en Apoyo a la Investigación de la Universidad Pablo de Olavide, incluye también módulos enfocados a la preparación de los bibliotecarios y documentalistas en materia de datos.*

*En resumen, el ecosistema de los datos comprende mediciones, tanto numéricas como no-numéricas, diseño, visualización y graficado; ofrece herramientas, procesos normalizados, y metodologías; beneficia acciones a largo plazo, demanda orden, organización, gestión cuidadosa y, desde luego, nuevos aprendizajes, lo que estamos denominando alfabetización en datos, que permita aprender a leer y escribir datos pero también cómo acceder a ellos, interpretarlos, discernirlos, analizarlos y utilizarlos conscientemente. En una de sus publicaciones A. Merrill (2011) propuso al gremio una reflexión que pareciera hoy más vigente que nunca: “la alfabetización en datos necesita una sede ¿por qué no la biblioteca?” (Merrill, 2011). Quienes están dedicando un tiempo de su agenda a leer estas líneas seguramente tienen ya la respuesta.*

## **Referencias bibliográficas**

Borrego, Á., Ardanuy, J., & Urbano, C. (2018). Librarians as Research Partners: Their Contribution to the Scholarly Endeavour Beyond Library and Information Science. *The Journal of Academic Librarianship*, 44(5), 663-670.  
<https://doi.org/10.1016/j.acalib.2018.07.012>



**Gestión de datos de investigación y alfabetización en datos como procesos que llegaron para quedarse**

Burton, M., Lyon, L., Erdmann, C., & Tijerina, B. (2018). *Shifting to Data Savvy: The Future of Data Science In Libraries* [Monograph]. University of Pittsburgh. <https://d-scholarship.pitt.edu/33891/>

Dal-Ré, R. (2020). Análisis de los artículos españoles de biomedicina retractados entre 1970 y 2018. *Medicina Clínica*, 154(4), 125-130. <https://doi.org/10.1016/j.medcli.2019.04.018>

Measuring digital development. Facts and Figures 2022. (s. f.). ITU. Recuperado 15 de febrero de 2023, de <https://www.itu.int:443/en/ITU-D/Statistics/Pages/facts/default.aspx>

Merrill, A. (2011). Library+. *Public Services Quarterly*, 7(3-4), 144-148. <https://doi.org/10.1080/15228959.2011.623598>

Rebecca Bryant, B. L. (2021, noviembre 15). *The Realities of Research Data Management Part Two: Scoping the University RDM Service Bundle*. OCLC. <https://www.oclc.org/research/publications/2017/oclcresearch-rdm-part-two-scoping-service-bundle.html>

Research Data Services in Academic Libraries: Where are We Today? (2021, junio 28). *Choice* 360. <https://www.choice360.org/research/research-data-services-in-academic-libraries-where-are-we-today/>

Rethlefsen, M. L., Farrell, A. M., Osterhaus Trzasko, L. C., & Brigham, T. J. (2015). Librarian co-authors correlated with higher quality reported search strategies in general internal medicine systematic reviews. *Journal of Clinical Epidemiology*, 68(6), 617-626. <https://doi.org/10.1016/j.jclinepi.2014.11.025>

## Editorial

### **Gestión de datos de investigación y alfabetización en datos como procesos que llegaron para quedarse**

Servicios de datos de investigación. Especial IFLA journal. (2017, abril 18).  
*bibliotecaceu*. <https://bibliotecaceu.wordpress.com/2017/04/18/servicios-de-datos-de-investigacion-especial-ifla-journal/>

Los autores declaran que no existe conflicto de intereses.