



Fundamentación

Entre las dificultades con que los investigadores se encuentran para llevar adelante sus proyectos, la disponibilidad en línea de documentación, bases de datos (de investigación y bibliográficas), bibliografías y mapas no son las menos frecuentes.

Si bien muchos de esos contenidos están disponibles en la web, debemos considerar que su disposición en diferentes repositorios, (muchos de los cuales no están orientados al uso de los investigadores) implica enormes pérdidas de tiempo, en muchos casos ocasionadas por etiquetados faltantes o deficientes sumado a la falta de referencias y vínculos que relacionen los elementos almacenados en cada uno de esos sitios.

Iniciativas vinculadas a la promoción del Acceso Abierto de fuentes y datos de investigación en tanto información de gobierno se encuentran enmarcadas por la Ley 26.899/13 de Creación de Repositorios Digitales Institucionales de Acceso Abierto Propios o Compartidos, en el marco del cual el Ministerio de Ciencia y Tecnología desarrolló el Repositorio Institucional Conicet Digital. Asimismo, la Ley 27.275/16 de Derecho de Acceso a la Información Pública, que presume que toda la información producida por los organismos del Estado, salvo excepciones debidamente especificadas, es pública y debe disponerse en tal sentido.

En articulación con la redacción de la Ley 26.899/13, se creó el Sistema Nacional de Repositorios Digitales (SNRD) entre el Ministerio de Ciencia y Tecnología y el Consejo Interinstitucional de Ciencia y Tecnología, que estableció programas de financiamiento para el desarrollo de grandes sistemas nacionales de datos por temáticas, el más reciente de los cuales fue el Sistema Nacional de Datos Históricos¹.

A su vez, las universidades nacionales han conducido sus propios programas, como en el caso de la Universidad Nacional de Rosario, que complementó su Repositorio Hipermedial con un Repositorio de Datos Académicos.

¹ <https://www.argentina.gob.ar/ciencia/sistemasnacionales/documentacion-historica>

En algunos documentos, como Indart, Ríos Hilario y Rodríguez Bravo (2021) también el Decreto 117/2016 es recogido como un antecedente relevante para la evolución de la transparencia de los datos de investigación. Sin embargo, esta disposición conmina a los organismos públicos a declarar ante el recientemente creado Ministerio de Modernización la nómina de empleados, con sus salarios y recursos, dentro de un plazo determinado. El uso que este fugaz Ministerio ejerció sobre esta información no redundó en ninguna política de acceso a los datos públicos ni mucho menos de investigación.

En cumplimiento de estos marcos legales, las iniciativas institucionales que se realizaron tuvieron un éxito dispar o se encuentran actualmente en desarrollo. Por otro lado, como señala Borgman (2012) cuando afirma que el problema de la publicación de datos es que a pesar de la presión de los organismos de financiamiento, hay muchas causas que frenan el desarrollo adecuado de bases de datos. Por un lado, la falta de conocimiento técnico, recursos e incentivos de parte de los investigadores. Por el otro, la propia característica de determinados datos, que pueden existir en formatos intransferibles, o estar restringidos por razones éticas o epistemológicas. Por último, en muchos casos, no queda claro cuales son “los datos” asociados a un determinado proyecto o más aún, qué es específicamente un dato. Estas experiencias muestran la importancia de institucionalizar procesos que trasciendan las conducciones individuales y promuevan proyectos de mediano plazo con objetivos y saberes técnico profesionales consolidados.

En el marco internacional, la OECD (2007) define a los datos como

registros factuales (numéricos, textuales, imagen o sonido) usados como fuentes primarias para investigaciones científicas, comúnmente aceptadas por la comunidad científica como necesarias para validar hallazgos de investigación. Un set de datos constituye una representación sistemática y parcial del objeto investigado (p.13).²

Si bien es indudable la proliferación de reflexiones e iniciativas institucionales alrededor del tema de los datos científicos, uno de los problemas que se perciben tiene que ver con su estructura. Los Sistemas Nacionales de Repositorios Digitales (SNRD), por ejemplo, se sostienen a través de plataformas para el sostenimiento de Repositorios Institucionales como DSpace, o de bibliotecas virtuales, como KOHA. Estructuralmente, son plataformas que pueden sostener archivos informáticos de

² Factual records (numerical scores, textual records, images and sounds) used as primary sources for scientific research, and that are commonly accepted in the scientific community as necessary to validate research findings. A research data set constitutes a systematic, partial representation of the subject being investigated.

cualquier tipo con descriptores a configurar, pero en rigor, el *dataset* se publica como documento, y no constituye una base de datos propiamente dicha. Entendemos que la base de datos es la estructura por la cual se sistematizan los datos científicos y su construcción es específica a la naturaleza de los datos en cuestión y a la necesidad de optimizar su usabilidad e interoperabilidad. La confusión respecto de la naturaleza y definición de los datos emerge incluso en publicaciones e investigaciones rigurosas. Un ejemplo interesante es el de de Indart, Ríos Hilario y Rodríguez Bravo (2021), que excluyeron de su relevamiento, acertadamente, al Archivo Digital Comunitario de la UNCPBA por contener fotografías, pero a su vez, con un criterio diferente, excluyeron también a Humadoc (UNMDP) por contener fichas de entrevistas sin las entrevistas propiamente dichas. Con este último criterio, cualquier base de datos de documentación histórica utilizada en una investigación debería incluir sus originales, pero con el primer criterio, quedarían excluidas si lo hicieran. Por último, Indart, Ríos Hilario y Rodríguez Bravo también excluyeron a la plataforma abierta de datos espaciales de la Argentina, *Poblaciones*, con el argumento de que solo producían visualizaciones georreferenciadas de datos. Para nosotros, los datos siempre se “visualizan” de algún modo, y la georreferenciación es en sí misma un dato, o una interfaz particular de una base de datos. Por lo tanto, la construcción de bases de datos es específica de la estructura del set de datos, e idealmente debe ser elaborada *ad hoc*, en función de las propias características de cada base de datos en particular, pero estructurados con criterios generales que permitan la futura integración e interoperabilidad.

Laboratorio de Procesamiento de datos y Objetos Digitales (OPRODi)

En vistas a estas problemáticas, proponemos la conformación de un Laboratorio (LAPRODi) a cargo de un equipo de trabajo dedicado a la elaboración y concreción de Planes de Gestión de Datos de Investigación y de los recursos digitales vinculados con ellos para la comunidad académica, con miras a la elaboración de bases de datos y a la publicación de fuentes de relevancia para la investigación científica que no se encuentren en fondos o archivos digitalizados.

Proyectos específicos del LAPRODi

Constitución de Bases de datos de investigación en base a datos proporcionados por investigadores del LAHISCO y del programa MyAS

- Capacitación a investigadores para la elaboración de bases de datos
- Sistematización y aplicación de criterios de interoperabilidad
- Publicación en red de bases de datos elaboradas por nuestros investigadores

Creación de una Biblioteca de documentos para la historia regional

- Publicación de fuentes digitales, previo proceso de reconocimiento de caracteres y revisión, o en formato facsímil, con su propio ISBN.
- Elaboración de un plan de adquisiciones con un perfil regional con el CEHISO

Cartografías digitales

- Mapas antiguos y modernos.
- Mapas temáticos elaborados por los investigadores
- Utilización de sistema QGIS para el desarrollo de mapas y cartografías especializadas

Autoevaluación y metodología de trabajo

La tarea de elaboración de bases de datos estructuradas y edición de documentos de historia regional tienen como función básica la facilitación de la tarea de los historiadores. Seguimos a Malavassi Aguilar (2012) cuando afirmamos que permiten:

- “1. Unificación de la información, en ciertos casos procedente de variados focos documentales, en torno a un eje temático específico, ya que una base responde al planteamiento de una problemática muy específica, donde las diversas entradas permiten medir las variables de la investigación a través de sus respectivos indicadores
2. Automatización y agilización en la consulta de los datos
3. Maleabilidad al posibilitar la introducción de cambios en la estructura de los datos y la actualización de los mismos
4. Depuración de los datos con la consecuente eliminación de las duplicaciones”

La operacionalización de esos cuatro puntos puede dar a luz a un formulario de autoevaluación que permita autodiagnosticar criterios básicos de calidad técnica y científica en la elaboración de los datos. En esa misma línea, los lineamientos establecidos por la OECD son una referencia ineludible a la hora de pensar

principios básicos para el acceso de datos de investigación con fondos públicos. Estos son: apertura, flexibilidad, transparencia, conformidad legal, protección de la propiedad intelectual, responsabilidad formal, profesionalismo, interoperabilidad, calidad, seguridad, eficiencia, sostenibilidad (OECD, 2007).

Para asegurar el cumplimiento de los lineamientos de la OECD, consideramos que las bases de datos deben ser elaboradas con una arquitectura abierta y reconocida, basada en la web. Por eso, proponemos una estructura basada en lenguaje php y MySQL, que puede operar tanto en la página del ISHIR como en servidores externos. Para la constitución de la biblioteca regional, los formatos de salida deberán ser pdf maquetados en formato libro y epub, estableciendo series o colecciones coherentes y bien estructuradas. A la descripción de ambos los unifica en un esquema de metadatos bien conocido por el equipo de trabajo, ya utilizado para la descripción de colecciones digitalizadas: Dublin Core, que permite además homogeneizar criterios de contenidos y en un futuro, quizá, posibilitarían elaborar bases de datos documentales sobre la propia biblioteca de documentos que estamos iniciando.

Resultados esperados

Uno de los objetivos principales es consolidar un método de trabajo que sirva de soporte y ayude a potenciar a las investigaciones locales, además de constituirse en un servicio para aquellos académicos de cualquier parte del mundo que estén interesados en introducirse en problemáticas de nuestra región. El mismo constituirá además una herramienta de vinculación tecnológica, que permitirá la realización de STAN, convenios de colaboración institucional y fortalecimiento de proyectos de investigación.

La producción y elaboración de bases de datos y colecciones digitales es una tarea de alcance técnico profesional, que requiere la participación de personal específico, capacitado, con conocimiento actualizado en la materia y con alta integración a las actividades científicas que se desarrollan en el ISHIR. La consolidación de un equipo de trabajo es un resultado que se obtiene *in situ*, sobre proyectos concretos en los que se releven necesidades y fortalezas propias, y requieren un soporte institucional estable y consolidado.

La creación de un fondo bibliográfico de documentos para la historia regional permitiría además recircular obras de difícil acceso, agotados o sin reediciones recientes, potenciar el impacto y alcance de fuentes de investigación de alcance limitado pero de una utilidad extendida, como una tarea que permita complementar con la creación de bases de datos, de modo de constituir al ISHIR como un punto de referencia en la producción de conocimiento de base regional.

Bibliografía

Borgman, C.L. (2012), The conundrum of sharing research data. J Am Soc Inf Sci Tec, 63: 1059-1078. <https://doi.org/10.1002/asi.22634>

Indart, C., Ríos Hilario, A. B., & Rodríguez Bravo, B. (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>

Organisation for Economic Cooperation and Development (OECD) (2007). *OECD Principles and Guidelines for Access to Research Data from Public Funding*. Paris: OECD Publishing.

<<https://www.oecd.org/sti/inno/38500813.pdf>> [Consulta: 05 septiembre 2023].

Vasallo, J. & García, N. (2013). Derecho y Patrimonio documental en Argentina: La protección jurídica a través de la acción de amparo y la denuncia penal. *Documentalistas*, 32 (6) (pp. 20-26).

Toso, R. (2010). Los archivos, el Estado y las políticas de la Memoria. Algunas reflexiones críticas. *Anuario digital*, 23(2) (pp. 151-166).

Weibel, S. Kunze, J. Lagoze, C. & Wolf, M. (1998). Dublin Core Metadata for Resource Discovery. *The Internet Society*: <http://www.ietf.org/rfc/rfc2413.txt>

Malavassi Aguilar, Ana Paulina. (2012). Las bases de datos como herramienta para la investigación histórica.

Diálogos Revista Electrónica de Historia, 13(1), 193-197. <https://doi.org/10.15517/dre.v13i1.6370>

ANEXO

Colección de bases de datos estructuradas que sirven como modelo

<https://artempire.cica.es/search>

<http://studium-parisiense.univ-paris1.fr/home>

<https://www.slavevoyages.org/american/database#>

<http://www.chacoarchive.org/cra/>

<https://www.agapepatristics.net/public/frontend>

<https://brasilhis.usal.es/es>

Para lograr esos objetivos, el proyecto deberá sostener dos criterios fundamentales:

- un recorte y definición respecto de la referencia geográfica de los materiales publicados
- una sólida estructura interna que relacione los artículos para facilitar y promover lecturas cruzadas a partir de una correcta edición de los vínculos al interior del sitio.