

De la archivística a la gestión de datos de investigación: Diversificando el acceso a la fuentes de información

Por: Hilda Teresa Ayala-González, hilda.ayala@upr.edu
Universidad de Puerto Rico, Recinto de Mayagüez

Resumen

Con el fin de diversificar el acceso a fuentes de información y atender apropiadamente las nuevas necesidades de los usuarios, en la última década, instituciones de educación superior han comenzado a explorar nuevos servicios y por ende se ven obligadas a ampliar el rango de competencias que posee su personal en áreas como: manejo y creación de recursos digitales, digitalización, preservación y curaduría digital, sistemas de información geográfica y la gestión de datos de investigación, entre otros. Las asociaciones profesionales tales como: Association of College & Research Libraries, Association of Research Libraries y Society of American Archivist, abordan el nuevo rol del profesional de la información tomando ventaja del conocimiento previo y cómo este, puede servir de puente para desarrollar una amplia gama de servicios institucionales. Una de las bibliotecas que aspira a ser modelo para la implementación de servicios innovadores en Puerto Rico es la Biblioteca de la Universidad de Puerto Rico, Recinto de Mayagüez a través de su Centro para la Investigación e Innovación de Posgrado (GRIC). En esta presentación se hablará de la experiencia de una nueva profesional en el campo de la archivística y cómo ha logrado adaptar el conocimiento adquirido a las áreas de interés emergentes, en especial la gestión de datos de investigación y la preservación digital. Se compartirán, además estrategias exitosas para el fortalecimiento de estas destrezas y las oportunidades de adiestramiento en el campo con el fin de motivar a otros a inmiscuirse y formarse en estas modalidades de servicios de información.

Palabras claves

Archivística, Gestión de datos de investigación, Nuevos profesionales de la información

Este obra está bajo una [licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/)

To meet fully our obligations as keepers of the nation's official memory, the gaps in our archival system must be closed. As I have said, much of the work of scientists today is sponsored by government, an increasingly large share by the federal government, and unless we have programs to care for the documentation resulting from this activity the record will not be properly kept.

Wayne C. Grover, 1962

Introducción

Con el fin de diversificar el acceso a fuentes de información y atender apropiadamente las nuevas necesidades de los usuarios, en la última década, instituciones de educación superior han comenzado a explorar nuevos servicios y por ende se ven obligadas a ampliar el rango de competencias que posee su personal en áreas como: manejo y creación de recursos digitales, digitalización, preservación y curaduría digital, sistemas de información geográfica y la gestión de datos de investigación, entre otros. Una de las bibliotecas que aspira a ser modelo para la implementación de servicios innovadores en Puerto Rico es la Biblioteca de la Universidad de Puerto Rico, Recinto de Mayagüez a través de su Centro para la Investigación e Innovación de Posgrado (GRIC). El GRIC, auspiciado por la Iniciativa Transformadora para la Educación y la Investigación Graduada (TIGER), abrió sus puertas en septiembre 2016 con el fin de diseñar e implementar una nueva serie de servicios para la comunidad de profesores y estudiantes graduados.

GRIC: Servicios para la comunidad

El GRIC, además de ser un espacio fundamentado en facilitar el trabajo colaborativo, en un ambiente cómodo, adaptable a las necesidades de los usuarios y con tecnologías de primera, ha desarrollado sus servicios a partir de las etapas del ciclo de la investigación. Este modelo aspira proveerle a los usuarios una guía de las áreas de especialidad que el equipo de trabajo del GRIC está listo para apoyar durante su investigación. Algunas de estas áreas son tradicionalmente reconocidas e implementadas como parte de los servicios que ofrecen las

bibliotecas, estando relacionadas a las competencias de información como: la búsqueda de información, la revisión de literatura, los manuales de estilo, la integridad académica y los derechos de autor. El GRIC, además de reconocer la importancia de estas destrezas, ha logrado ir más allá, ofreciendo apoyo en otras áreas como la comunicación académica, los perfiles e identidad virtual y el apoyo en la disseminación de sus productos mediante el repositorio digital institucional. Finalmente, este modelo expandió su alcance para incluir servicios especializados en la gestión de datos de investigación.

La Gestión de Datos de Investigación (GDI):

Es el proceso activo de manejo de los datos generados en una investigación. La GDI se realiza de forma continua y cubre todas las decisiones relacionadas con la gestión de los datos a lo largo de su ciclo vital, comenzando en la etapa de planificación de la investigación y abarcando su ejecución, la disseminación de sus resultados y la preservación de los conjuntos de datos de forma que estos sean precisos, completos, auténticos y fiables, y se mantengan accesibles y reutilizables a lo largo del tiempo (CEPAL, 2019, “Módulo 1”).

Si observamos detenidamente esta definición, específicamente desde el contexto digital, como archivistas, debemos vernos identificados y comprender el llamado que, como profesionales, tenemos ante estas fuentes de información y las -no tan- nuevas necesidades de nuestros usuarios. Más aún, desde el GRIC, hemos aprovechado esta coyuntura para promover las competencias no tan solo en la gestión de sus datos de investigación, sino también en el manejo de sus documentos electrónicos.

Desde la archivística, apoyando la GDI en la academia

En el contexto de la Universidad de Puerto Rico, Recinto de Mayagüez, anualmente se administran millones de dólares provenientes de fondos externos, en su gran mayoría de instituciones federales. En el 2013 se emitió un memorando desde la Oficina de Políticas para la Ciencia y Tecnologías de la Oficina Ejecutiva del Presidente de los Estados Unidos con el fin de mejorar la inversión monetaria que se realiza cuando se subvencionan investigaciones (OSTP, 2013). El documento expone que los proyectos que obtengan fondos federales deberán fomentar el acceso abierto (o con las menores restricciones posibles) a los datos que se compilan y hacerlos accesibles a través del depósito de estos en un repositorio. A la par, diversas casas publicadoras y revistas académicas de renombre (como por ejemplo: Nature, Plos, Science, etc.) continúan integrando entre sus políticas para la aceptación de artículos, el requisito de la entrega o disponibilidad de los datos para promover la transparencia y reproducibilidad de las investigaciones (Open Access Directory, 2018).

Enfocándonos en el contenido sugerido de un plan de manejo de datos, observamos que varias de las secciones, tanto en el lenguaje utilizado como en los requisitos señalados, van a requerir el apoyo de profesionales de la información. Por ejemplo, áreas como la organización y nomenclatura de las carpetas y documentos, la creación de metadatos y documentación del proyecto, estrategias de almacenamiento y seguridad, protección de datos, derechos de autor y limitaciones para el acceso y finalmente las tareas de preservación a largo plazo y licencias de uso para su reutilización (CEPAL, 2019, “Módulo 2”) están mucho más familiarizadas con lo que hacemos desde el archivo que con lo que hace un investigador en su día a día.

Estas dos tendencias, tanto el requisito de elaborar los planes de manejo de datos, como la disponibilidad de estos en repositorios, apuestan a lo mismo, la preservación a largo plazo y el

promover el acceso y la reutilización en el futuro. Ambas, tareas fundamentales a la gestión de los archivos históricos y ligadas intrínsecamente a las destrezas y conocimiento que posee un archivista.

Partiendo de esta premisa, es que se diseña la serie de GDI en el GRIC, con el fin de proveer a los estudiantes graduados, profesores e investigadores las destrezas necesarias para asegurar que sus datos puedan cumplir su ciclo de vida y mantenerse accesibles en la posteridad.

GDI en el GRIC

La serie de GDI del GRIC, cuenta con 5 talleres que aspiran ofrecerse semestralmente para atender las necesidades de la comunidad de investigadores en el Recinto. Estos son presentaciones de 1.5 horas con contenido en inglés y dictado en español, para beneficio de la diversidad de personas que componen nuestra población. Varios de los talleres cuentan con un elemento práctico y discuten herramientas de utilidad para su proceso investigativo. Además, se crearon 2 talleres adicionales específicamente para profesores. Uno sobre la creación de planes de manejo de datos, diseñado y ofrecido por la Prof. Jaquelina Álvarez y otro sobre las mejores prácticas, que busca contextualizar los requisitos de las agencias y casas publicadoras. También se creó una guía informativa con contenido suplementario, realizada por la Prof. Alessandra Otero. Los usuarios pueden acceder desde allí las presentaciones en cualquier momento.

La serie en sí parte de la premisa de que los plan de manejo de datos, a pesar de que se promueven solo como un requisito de las agencias federales que otorgan fondos, son una excelente herramienta de planificación y debería ser utilizado de todos los que elaboran un proyecto de investigación.

Es por esto que la primera presentación, *10 elementos a considerar antes de comenzar un proyecto de investigación*, se fundamenta en lo recomendado por el Digital Curation Center

(2013) y su “Checklist for a Data Management Plan”. Como apoyo, a los participantes se les comparte una hoja adaptada con información suplementaria donde pueden encontrar más información sobre cada uno de los requisitos. La segunda presentación, *Organización, estandarización y control de documentos* se enfoca en pura archivística y manejo de documentos. Cubre temas como nomenclaturas, creación de series y sub-series (pero no se usa esta terminología), manejo de originales y duplicados, metadatos, entre otras cosas. El taller es práctico y se discuten herramientas como Batch Renaming, Easy Duplicate Finder y Git.

La tercera presentación, *Estrategias de almacenamiento*, discute la importancia de la retención de documentos versus los riesgos de pérdida de contenido desde: la diversidad de localidades (soportes) donde almacenamos, la volatilidad de los formatos, la obsolescencia de tecnologías y programados, entre otros. También, provee estrategias para monitorear, crear copias de resguardo y asegurar que los datos se mantengan accesibles a largo plazo. Como elemento práctico, se explora el uso de la herramienta Fixity de AVPreserve para corroborar si un documento sufrió alguna alteración, a través de la comparación de la suma de verificación. La cuarta presentación, en colaboración con la Prof. Jaquelina Álvarez, *Búsqueda y reutilización de datos*, provee a los participantes una idea de la diversidad de motores de búsqueda y repositorios donde pueden encontrar datos para integrar a sus investigaciones. Además, se discute cómo evaluar los mismos y distinguir datos de buena calidad versus aquellos de dudosa proveniencia. También se exploran aspectos sobre la ética en la reutilización de contenido y se provee un resumen de cómo interpretar las licencias Creative Commons y citar correctamente los datos incorporados. El taller incluye herramientas para facilitar el procesamiento de datos como Unscan-it, goodtables y Openrefine y se hace un ejercicio práctico para buscar datos a través del directorio Re3Data y DataSearch de Elsevier.

La quinta y última presentación de la serie de GDI del GRIC, *Difusión de datos de investigación*, presenta el panorama actual para ampliar la accesibilidad y reproducibilidad de datos de investigación. Se discuten los agentes involucrados, con un enfoque en las agencias federales y casas publicadoras y se presentan las diversas alternativas para compartir los mismos. Este aspira a que nuestra comunidad pueda cumplir responsablemente con la difusión de su contenido, aprendiendo a preparar los mismos para su depósito.

Estos talleres, aún pueden abarcar otras áreas, como la compilación, análisis y presentación de datos. Pero, con el fin de preservar la documentación científica en nuestro País, mi rol, es proveerle las destrezas a estudiantes graduados, profesores e investigadores para que sus datos sobrevivan los años de la investigación y que puedan pasar a mejor vida una vez culmine el proyecto.

Sobre esto, podemos citar y hacer eco al llamado de Wayne C. Grover, tercer archivista de los Estados Unidos cuando en 1962 habló sobre el rol de los archivos para preservar la documentación científica. En su escrito Grover indica que: a menos que contemos con un sistema para cuidar los documentos que resultan de estas actividades, los mismos no serían mantenidos de la forma adecuada. Y añade que: como custodios de la memoria [y yo añadiría, evidencia] oficial de una nación, es nuestra obligación cerrar todas las lagunas que pueden existir en el sistema (1962, p.56).

En el GRIC, nosotros escogimos comenzar con la primera línea de defensa contra la pérdida de información digital, es decir, con los creadores. Ya que son ellos quienes tiene la responsabilidad inicial en la preservación de sus objetos digitales (Task Force on Archiving of Digital Information, 1996, p.22, 40). Esta necesidad nos inspira a cumplir con nuestro deber de archivista como educador en asuntos relacionados a la gestión de la información digital (Tibbo,

2003) para asegurar el acceso a estas fuentes en la posteridad. Además, auguramos que los talleres, no tan solo apoyen su rol como investigador en el siglo XXI, sino que le faciliten ejercer cabalmente su papel como ciudadanos en la era digital.

GDI para profesionales de la información

Esto, también quiere decir que nosotros, como profesionales de la información, tenemos la responsabilidad de educarnos en estos asuntos, no tan solo para apoyar a nuestros usuarios, sino también para ejercer nuestras tareas del día a día. Más aún, implica que las escuelas desde donde se forman los archivistas y bibliotecarios deben atemperar sus currículos y proveer a la nueva cepa de profesionales, por lo menos la base para afrontar los retos de la era digital.

Como recién graduada de la Maestría en Archivística de la Universidad de British Columbia en Vancouver Canadá (en el 2016) tuve la oportunidad de tomar varios cursos que hicieron fácil la transferencia de conocimiento, entre estos: tecnologías de la información, desarrollo de bases de datos, preservación digital, diplomática y forense digital. Pero, hago la salvedad que gran parte del conocimiento adquirido ha sido gracias a diversas actividades de desarrollo profesional.

Actualmente existen un sinnúmero de recursos, en su mayoría gratuitos, que pueden ayudar a cualquier persona interesada en aprender sobre esta temática. En mi caso, estos fueron en su mayoría en el idioma inglés y comencé con el curso en línea *Research Data Management and Sharing*, que se ofrece a través de la plataforma Coursera, por las instructoras, Helen Tibbo y Sarah Jones y *MANTRA, Research Data Management Training* diseñado por la Universidad de Edinburgh. También participé de dos ediciones del Research Data Access & Preservation Summit, lo cuál se convirtió hace poco más de un año en una asociación profesional, una excelente comunidad de apoyo para los novatos y expertos. Finalmente, formé parte de uno de

los cohortes del *Preserving Digital Objects With Restricted Resources Institute* en el verano de 2018. Esto, sin contar los diversas lecturas, talleres y conferencias en línea que continúan aportando a desarrollarme en este campo y me otorgan mayor confianza de compartir con otros lo aprendido. Unas descubiertas muy recientemente, y en español, son: la Biblioguía “[Gestión de datos de investigación](#)”, con contenido de MANTRA, diseñado por la Comisión Económica para América Latina y el Caribe (CEPAL) y la guía por James Baker y traducida por Víctor Gayol “[Preservar tus datos de investigación](#)” publicado en *The Programming Historian*. Ambos con excelentes recomendaciones para iniciarse en la materia.

Conclusión

Una de las cosas que más aprecio de ser una nueva profesional de la información son las ganas constantes de adquirir conocimiento y emprender nuevos retos. A los archivistas que aún están dudosos de la relevancia de nuestra profesión en estos tiempo, les invito a asumir una nueva mirada y ejercer el rol protagónico que amerita este siglo. Hoy más que nunca nuestro conocimiento es necesario, pero no solo para unos pocos, si no para todos.

Referencias

Comisión Económica para América Latina y el Caribe (CEPAL). (2019). Gestión de datos de investigación [Biblioguía]. Disponible en: <https://biblioguias.cepal.org/gestion-de-datos-de-investigacion>

Digital Curation Center. (2013). Checklist for a Data Management Plan. v.4.0. Disponible en http://www.dcc.ac.uk/sites/default/files/documents/resource/DMP/DMP_Checklist_2013.pdf

Grover, W.C. (1962). The role of the archivist in the preservation of scientific records. *ISIS, The Conference on Science Manuscripts*, 53(1), 55-62. Disponible en

<https://www.jstor.org/stable/227829>

Office of Science and Technology Policy (OSTP) (2013, February 22). Increasing access to the results of federally funded scientific research [memorandum]. Disponible en

<https://www2.icsu-wds.org/files/ostp-public-access-memo-2013.pdf>

Open Access Directory. (2018). Journal open-data policies. Disponible en

http://oad.simmons.edu/oadwiki/Journal_open-data_policies

Task Force on Archiving of Digital Information. (1996, May). Preserving Digital Information: Report of the Task Force on Archiving of Digital Information. Disponible en

<https://www.clir.org/wp-content/uploads/sites/6/pub63watersgarrett.pdf>

Tibbo, H. R. (2003). On the nature and importance of archiving in the digital age. *Advances in Computers*, 53 1–67. [https://doi.org/10.1016/s0065-2458\(03\)57001-2](https://doi.org/10.1016/s0065-2458(03)57001-2)

Este trabajo ha sido realizado gracias al apoyo del Centro para la Investigación e Innovación de Posgrado de la Biblioteca General de la Universidad de Puerto Rico en Mayagüez y a la Iniciativa Transformadora para la Educación y la Investigación Graduada del Programa de Título V, Parte B para la Promoción de Oportunidades Post-bachillerato para Hispanos del Departamento de Educación Federal (#P031M140035).